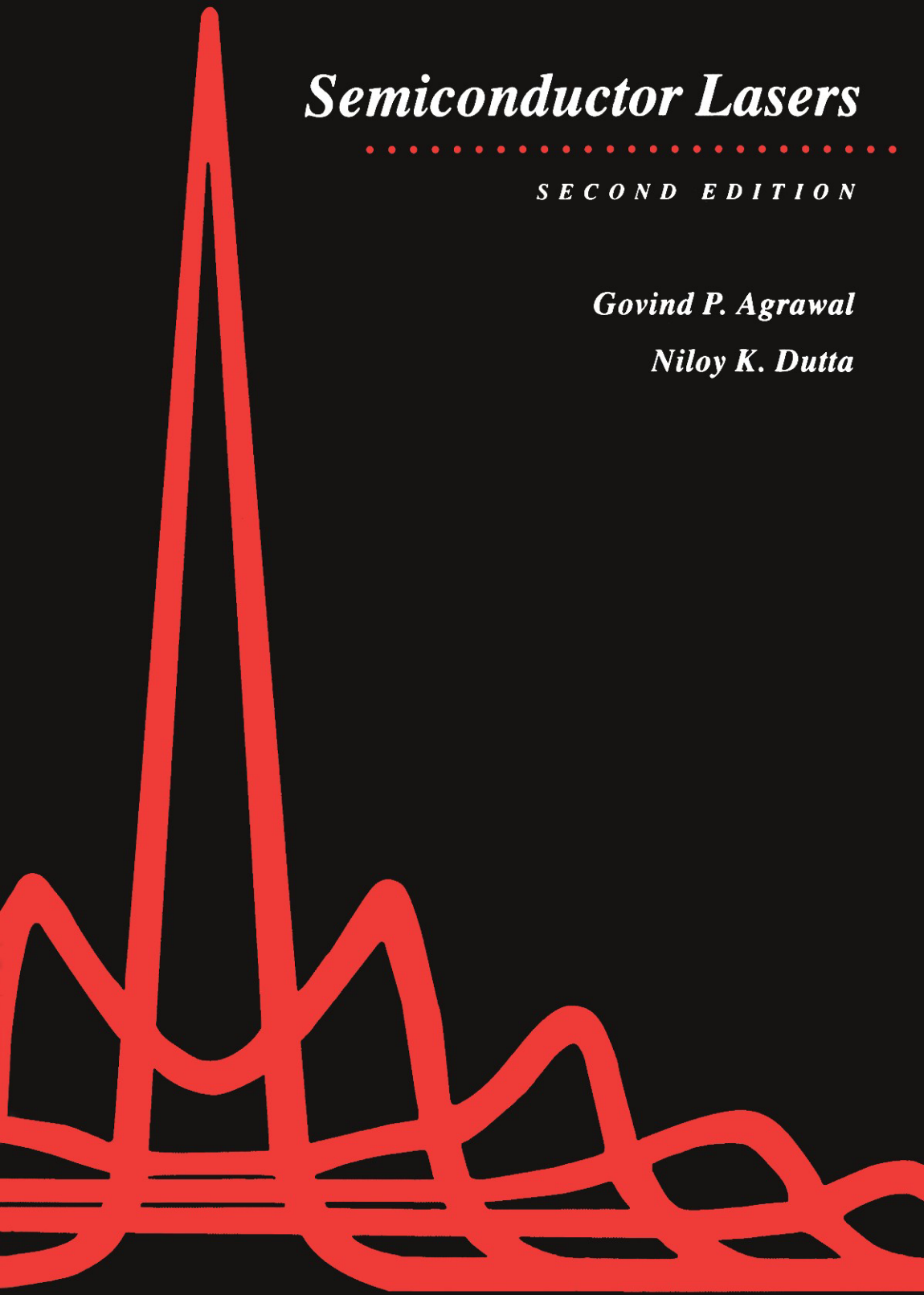
Semiconductor Lasers

SECOND EDITION

Govind P. Agrawal

Niloy K. Dutta



SEMICONDUCTOR LASERS

SEMICONDUCTOR LASERS

Second Edition

Govind P. Agrawal

The Institute of Optics, University of Rochester
Rochester, New York

and

Niloy K. Dutta

AT&T Bell Laboratories
Murray Hill, New Jersey

KLUWER ACADEMIC PUBLISHERS
BOSTON/DORDRECHT/LONDON

Distributors for North, Central and South America:
Kluwer Academic Publishers
101 Philip Drive
Assinippi Park
Norwell, Massachusetts 02061 USA
Telephone (781) 871-6600
Fax (781) 871-6528
E-Mail <kluwer@wkap.com>

Distributors for all other countries:
Kluwer Academic Publishers Group
Distribution Centre
Post Office Box 322
3300 AH Dordrecht, THE NETHERLANDS
Telephone 31 78 6392 392
Fax 31 78 6546 474
E-Mail services@wkap.nl>


Electronic Services <http://www.wkap.nl>

Library of Congress Cataloging-in-Publication

Agrawal, G.P. (Govind P.), 1951--
Semiconductor lasers / Govind P. Agrawal and Niloy K. Dutta.--2nd ed.
p. cm.
First ed. published 1986 under title: Long-wavelength semiconductor lasers.
Includes bibliographical references and index.
ISBN-13: 978-1-4612-7579-4 e-ISBN-13: 978-1-4613-0481-4
DOI: 10.1007/978-1-4613-0481-4
1. Semiconductor lasers. I. Dutta, N.K. (Niloy K.), 1953--
II. Agrawal, G.P. (Govind P.), 1951-- Long-wavelength semiconductor lasers. III. Title.
TA1700.A37 1993 92-46782
621.36'61--dc20 CIP

Third printing 2001 by Kluwer Academic Publishers
Softcover reprint of the hardcover 2nd edition 2001

Printed on acid-free paper.

This printing is a digital duplication of the original edition.

CONTENTS

PREFACE TO THE FIRST EDITION

Since its invention in 1962, the semiconductor laser has come a long way. Advances in material purity and epitaxial growth techniques have led to a variety of semiconductor lasers covering a wide wavelength range of 0.3–100 μm. The development during the 1970s of GaAs semiconductor lasers, emitting in the near-infrared region of 0.8–0.9 μm, resulted in their use for the first generation of optical fiber communication systems. However, to take advantage of low losses in silica fibers occurring around 1.3 and 1.55 μm, the emphasis soon shifted toward long-wavelength semiconductor lasers. The material system of choice in this wavelength range has been the quaternary alloy InGaAsP. During the last five years or so, the intense development effort devoted to InGaAsP lasers has resulted in a technology mature enough that lightwave transmission systems using InGaAsP lasers are currently being deployed throughout the world.

This book is intended to provide a comprehensive account of long-wavelength semiconductor lasers. Particular attention is paid to InGaAsP lasers, although we also consider semiconductor lasers operating at longer wavelengths. The objective is to provide an up-to-date understanding of semiconductor lasers while incorporating recent research results that are not yet available in the book form. Although InGaAsP lasers are often used as an example, the basic concepts discussed in this text apply to all semiconductor lasers, irrespective of their wavelengths.

The book is aimed at researchers already engaged in or wishing to enter the field of semiconductor lasers. It should serve as a useful reference for engineers who are interested in optical fiber communications and want to know about the semiconductor-laser sources employed therein. The book can also be useful for a graduate-level course on semiconductor lasers as part of a program in optical communications. We have attempted to make the book self-contained and to provide sufficient details of the mathematical derivations. Furthermore, each chapter refers to a large number of published papers that can be consulted for further study.

The book is organized as follows. The first three chapters introduce the basic concepts and provide the mathematical derivations useful for

understanding the operation of semiconductor lasers. Chapters 4 and 5 describe epitaxial techniques and various device structures employed to fabricate semiconductor lasers. The operating characteristics of these lasers are considered in Chapter 6, including static, dynamic, spectral, noise, and modulation aspects. The next two chapters are devoted to single-frequency semiconductor lasers employing the distributed-feedback and coupled-cavity schemes, while Chapter 9 considers quantum-well semiconductor lasers. The degradation mechanisms and reliability issues of semiconductor lasers are discussed in Chapter 10. Finally, Chapter 11 considers lead-salt semiconductor lasers emitting at relatively longer wavelengths in the far-infrared region 3–34 μm.

We wish to thank the members of the semiconductor laser development department and other colleagues at AT&T Bell Laboratories for numerous discussions and for providing a stimulating working environment. We are thankful to D. P. Wilt and C. H. Henry for their comments on several chapters. The support of the AT&T Bell Laboratories management for this project is gratefully acknowledged. We particularly thank R. W. Dixon, J. E. Geusic, and P. J. Anthony for their encouragement.

G. P. Agrawal
N. K. Dutta

PREFACE TO THE SECOND EDITION

The field of semiconductor lasers has advanced considerably since the publication of the first edition in 1986. Among the recent advances, to name a few topics, are surface-emitting semiconductor lasers, high-power laser arrays, visible semiconductor lasers, and strained layer quantum-well lasers. The second edition is intended to bring this book up to date so that it remains a source of comprehensive coverage on semiconductor lasers. Whereas the first edition focused mainly on long-wavelength semiconductor lasers (mostly InGaAsP lasers), the scope of this edition has been widened to include all kinds of semiconductor lasers, as reflected by the change in the title of the book. Since the first edition has occasionally been used as a textbook in some graduate-level courses, we have added selective problems at the end of each chapter to help teachers and students. It is our hope that the second edition can serve as a textbook for graduate courses dealing with semiconductor lasers. Selective chapters can also be useful for other courses related to lasers, optoelectronics, and optical communications.

The list of topics that could have been included in the second edition was quite large. Size limitations forced us to make a selection. It was felt that surface-emitting semiconductor lasers, semiconductor laser amplifiers, and optoelectronic integration needed enough coverage that a new chapter was justified for each of them. The main change consists of adding three new chapters (Chapters 10, 11, and 12) and several new sections to the existing chapters. Major changes are made in Chapters 6, 7, and 9 while other chapters are updated to bring the discussion up to date. Specifically, a section on mode-locked semiconductor lasers is added to Chapter 6. The advances in the field of distributed feedback semiconductor lasers are covered by adding a section on tunable semiconductor lasers and three sections on phase-shifted, quantum-well, and gain-coupled distributed feedback lasers. Chapter 9 has a new section on strained-layer quantum-well lasers, a topic that has attracted considerable attention in recent years. Visible semiconductor lasers are included in Chapter 13. We feel that these additions have improved the text enough that it should serve the need of the scientific community during the 1990s. We would welcome suggestions and comments from the readers.

G. P. Agrawal
N. K. Dutta

Chapter 1
INTRODUCTION

1.1 HISTORICAL PERSPECTIVE

The advent of the laser dates back to 1958, the year in which the seminal paper of Schawlow and Townes[1] appeared. It was followed by the successful operation of a solid-state ruby laser[2] in May 1960 and of an He-Ne gas laser[3] in December 1960. The feasibility of stimulated emission in semiconductor lasers was considered during this period,[4–6] and in 1962 several groups[7–10] reported the lasing action in semiconductors. The device consisted of a forward-biased GaAs *p-n* junction.[7–9] Electron-hole recombination in the depletion region of the *p-n* junction provided the optical gain, and the polished facets perpendicular to the junction plane provided the optical feedback (by forming a resonant cavity)—the two necessary ingredients for any laser. Soon *p-n* junctions of other direct-band-gap semiconductor materials such as InAs, InP, GaAsP, GaInAs, and InPAs were used to obtain semiconductor lasers at different wavelengths. Practical utility of these earlier devices was, however, limited since a large value of the threshold current density ($J_{th} \gtrsim 50$ kA/cm^2) inhibited their continuous operation at room temperature.

As early as 1963 it was suggested[11,12] that semiconductor lasers might be improved if a layer of one semiconductor material were sandwiched between two cladding layers of another semiconductor that has a relatively wider band gap. Such a device consisting of two dissimilar semiconductors is commonly referred to as a *heterostructure laser*, in contrast to the single-semiconductor devices, which are labeled as *homostructure lasers*. Both of these structures are shown schematically in Fig. 1.1, which also indicates their typical physical dimensions. Heterostructure lasers are further classified as *single-heterostructure* or *double-heterostructure* devices depending on whether the active region, where lasing occurs, is surrounded on one or both sides by a cladding layer of higher band gap. The use of a heterostructure, however, requires a careful matching of the lattice constants of the two semiconductors. It was only in 1969 that the successful room-temperature operation of a

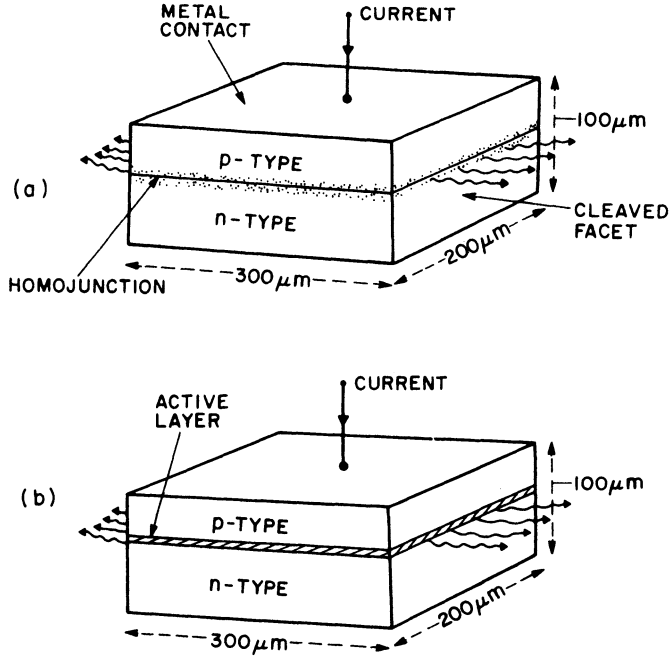


Fig. 1.1 Schematic illustration of (a) homostructure and (b) double-heterostructure semiconductor lasers with their typical physical dimensions. The dotted area represents the depletion region in the vicinity of the homojunction. The hatched area shows the thin (~ 0.2 μm) active layer of a semiconductor material whose band gap is slightly lower than that of the surrounding cladding layers.

heterostructure laser was demonstrated[13–15] using the liquid-phase epitaxial technique[16] for the growth of GaAs and $Al_xGa_{1-x}As$ layers. However, these lasers operated in the pulsed mode. Further work led in 1970 to heterostructure lasers operating continuously at room temperature.[17,18] Notation such as (Ga,Al)As or AlGaAs/GaAs is often used to emphasize the heterostructure nature of these GaAs lasers. However, since homostructure lasers are no longer used, we shall simplify the notation in this book, whenever no confusion is likely to arise, by denoting a heterostructure laser only by the composition of its active layer.

Already in 1969 double-heterostructure GaAs lasers with a room-temperature value of $J_{th} \cong 5$ kA/cm^2 were reported.[15] This value was reduced[17] to about 1.6 kA/cm^2 in 1970, and by 1975 AlGaAs layers with $J_{th} \cong 0.5$ kA/cm^2 were demonstrated using thin (~ 0.1 μm thick) active layers.[19] This was an improvement by more than two orders of magnitude over the simple homostructure lasers first made in 1962. It converted the semiconductor

laser from a laboratory curiosity to a practical, compact, coherent light source useful for numerous applications.

The physical reason for the reduction in the threshold current density with the use of a heterostructure device is twofold.[11,12] The cladding layers surrounding the active layer have a higher band gap and at the same time a lower refractive index compared with those of the active layer (see Fig. 1.2). The band-gap difference helps to confine electrons and holes to the active layer, where they recombine to produce the optical gain. At the same time the refractive-index difference confines the optical mode close to the active layer, which acts as a dielectric waveguide. The optical-mode confinement significantly reduces the internal loss[13] that would otherwise occur in the absence of index guiding due to the spreading of the optical mode in the lossy regions.

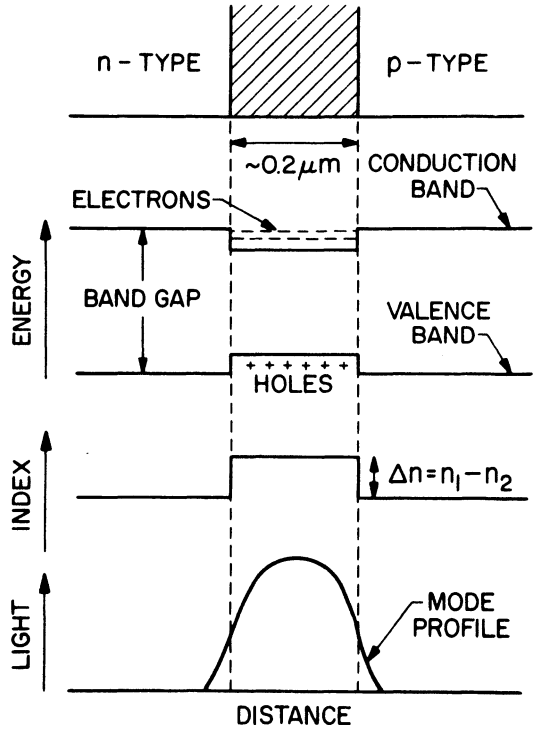


Fig. 1.2 Schematic illustration of the simultaneous confinement of the charge carriers and the optical mode to the active region occurring in a double-heterostructure semiconductor laser. The active layer has a lower band gap and a higher refractive index than those of the cladding layers. (After Ref. 24)

A double-heterostructure semiconductor laser such as shown in Fig. 1.1 is sometimes called a *broad-area laser* since it does not incorporate any mechanism for the lateral (parallel to the junction plane) confinement of the injected current or the optical mode. As early as 1967, stripe-geometry homostructure lasers were proposed[20] to limit the lateral spread of the injected carriers inside the active layer. In these lasers the current is injected over a narrow ($\sim 10\ \mu m$) central region using a stripe contact. The stripe geometry was adopted for heterostructure lasers[21] in 1971. Such lasers are also referred to as gain-guided since it is the lateral variation of the optical gain that confines the optical mode to the stripe vicinity.[22] By contrast, heterostructure lasers where the optical mode confinement occurs mainly through lateral variations of the refractive index are termed index-guided. A large number of index-guided structures have been proposed and demonstrated. The literature is too vast to cite them individually, and a discussion of their relative merits can be found in several books[23–25] that also describe the research and development effort for GaAs lasers up to 1980.

GaAs lasers are of continued interest. Using the growth techniques of vapor-phase epitaxy[26] and molecular-beam epitaxy,[27] multiquantum-well laser structures have been developed.[28] In these devices the active region is not a single GaAs layer but rather consists of several ultrathin ($\sim 0.01\ \mu m$) layers composed alternatively of GaAs and AlGaAs materials. Considerable effort has been directed toward developing high-power GaAs lasers and phased-array semiconductor lasers.[29,30] In one approach, multiple stripes are used to generate distinct regions of optical gain in the junction plane and the near field consists of several spots. However, since the stripes are not widely separated, the optical field in the gain region of each emitter overlaps with that of the neighboring emitter. Such a coupling leads to a phase-locked array of emitters providing well-collimated high power output. Whereas the output power from a conventional GaAs laser is usually below 50 mW, more than 100 W of power has been obtained[30] from monolithic laser arrays.

So far we have followed the development of GaAs lasers operating usually in the wavelength range of 0.8–0.9 μm. Long-wavelength semiconductor lasers in the range of 1.1–1.6 μm are of considerable interest for optical fiber communications.[31] Although several material systems were considered, the combination InGaAsP-InP turned out to be the most suitable in view of its nearly perfect lattice match.[32] The active layer is composed of the $In_{1-x}Ga_xAs_yP_{1-y}$ quaternary alloy. By varying the mole fractions x and y, almost any wavelength in the 1.1–1.6-μm range can be selected. The cladding layers in this heterostructure laser (see Fig. 1.1) consist of either InP or InGaAsP itself with different mole fractions x and y.

Room-temperature operation of a 1.1-μm InGaAsP laser in the pulsed

mode[33] was reported in 1975. The adoption of stripe geometry led to continuous operation[34] of such lasers in 1976. In 1977 the wavelength was extended[35] to 1.3 μm. Since low-loss dispersion-free fibers at 1.3 μm were already available,[36] considerable attention was focused on developing a practical InGaAsP laser at this wavelength. Motivated by the realization[37] of an ultra-low-loss (~0.2 dB/km) fiber at the 1.55-μm wavelength, several groups[38–43] in 1979 reported on InGaAsP lasers operating in the vicinity of 1.55 μm. After that, the development effort for InGaAsP lasers operating in the wavelength range of 1.3–1.6 μm proceeded at an enormous pace.[44–46] The primary motivation was due to their application in optical fiber communications, and by 1984 the use of InGaAsP lasers in long-haul optical communication systems had reached the commercial stage.[47]

1.2 SEMICONDUCTOR MATERIALS

The previous section has followed the development of heterostructure lasers based on two semiconductor materials, AlGaAs and InGaAsP. However, in view of their potential application in such diverse fields as optical fiber communication,[31] optical data recording,[48] high-speed printing,[49] and molecular spectroscopy, the list of semiconductor materials that have exhibited lasing action has continued to grow. Figure 1.3 shows the range of emission wavelengths for various semiconductor lasers. Taken together, these materials cover the optical spectrum from near ultraviolet to far infrared. Recently CdZnSe semiconductor lasers operating near 0.46 μm have been developed.[50]

The most important criterion in selecting the semiconductor material for a specific heterostructure laser is related to the quality of the heterojunction interface between the two semiconductors of different band gaps. To reduce the formation of lattice defects, the lattice constants of the two materials should typically match to better than 0.1%. Figure 1.4 shows the interrelationship between the band gap E_g and the lattice constant a for several ternary and quaternary compounds. Solid dots represent the binary compounds and solid lines correspond to the ternary compounds. The clear region bounded by the polygon (whose edges represent ternary compounds) denote the possible values of E_g and a for the quaternary solid solution of $In_{1-x}Ga_xAs_yP_{1-y}$ obtained by varying the mole fractions x and y. The dotted line shows the range of band-gap values that can be achieved by varying the compositions x and y to obtain a quaternary material that is lattice-matched to the binary InP. Figure 1.5 shows the constant band-gap contours (solid curves) in the x–y compositional plane. The dashed lines correspond to the fixed values of the lattice constant. For given values of E_g and a, the intersection of solid and dashed curves provides the compositional

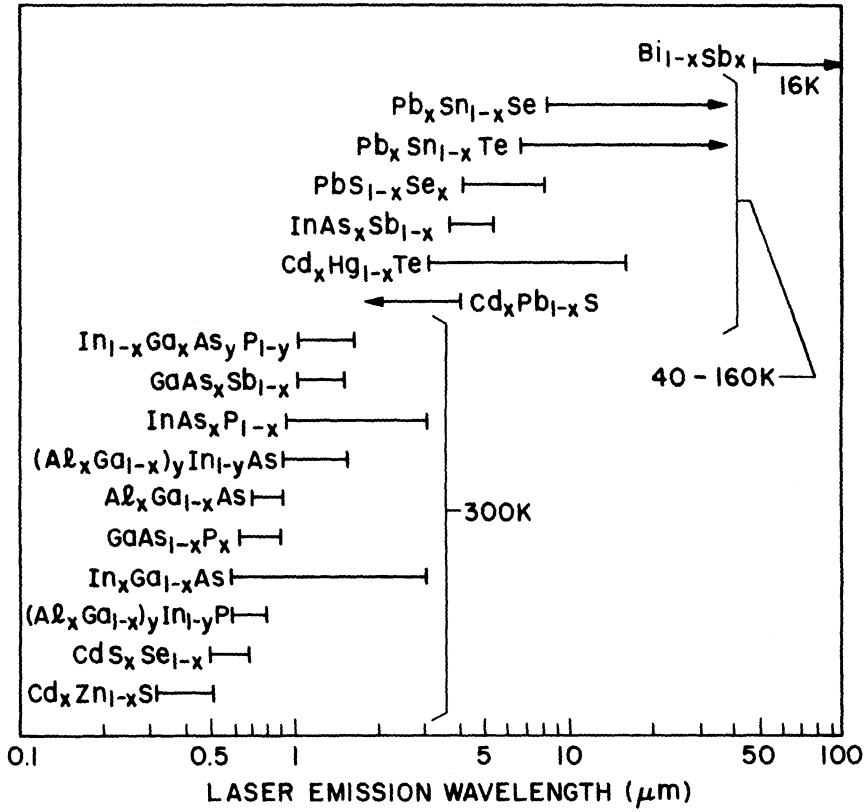


Fig. 1.3 Wavelength range of semiconductor lasers covered by different material systems. Semiconductor lasers emitting at $\lambda > 3$ μm usually require low-temperature operation.

values x and y used to obtain the active-layer quaternary material. Since the photon energy $E = hv$ is approximately equal to the band-gap energy, the lasing wavelength λ is obtained using $E_g = hc/\lambda$, where h is the Planck constant and c is the speed of light in vacuum. If E_g is expressed in electron volts, the lasing wavelength λ in micrometers is given by

$$\lambda \cong \frac{1.24}{E_g}. \tag{1.2.1}$$

For $In_{1-x}Ga_xAs_yP_{1-y}$ lasers a wavelength range of 1.1–1.65 μm can be covered by choosing x and y according to Fig. 1.5 such that the active layer is lattice-matched to InP ($a = 0.587$ nm). Semiconductor lasers emitting

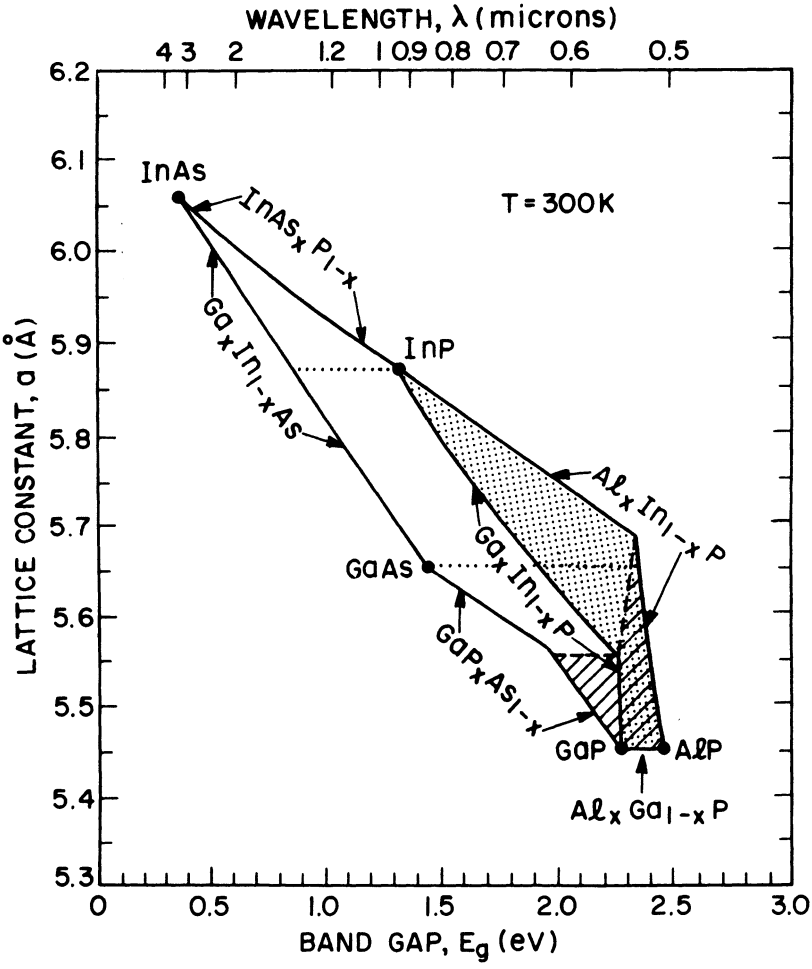


Fig. 1.4 Band gap and lattice constant for $In_{1-x}Ga_xAs_{1-y}P_y$ (clear region) and $(Al_xGa_{1-x})_y In_{1-y}P$ (shaded region) obtained by varying compositions x and y. Dashed lines separate indirect-band-gap regions (shown hatched). Dotted lines show the wavelength range (top scale) for a semiconductor laser whose quaternary active layer is lattice-matched to the binary compound. (After Ref. 24)

at 1.3-μm and 1.55-μm wavelengths are of particular interest because of their application in optical fiber communications. From Fig. 1.5 the active-layer composition at 1.3 μm ($E_g \cong 0.95$ eV) corresponds to $x = 0.28$ and $y = 0.6$. The cladding-layer composition can also be chosen from Fig. 1.5; the only requirement is that its band gap should be somewhat larger than that of the active layer. Diagrams similar to Figs. 1.4 and 1.5 have been

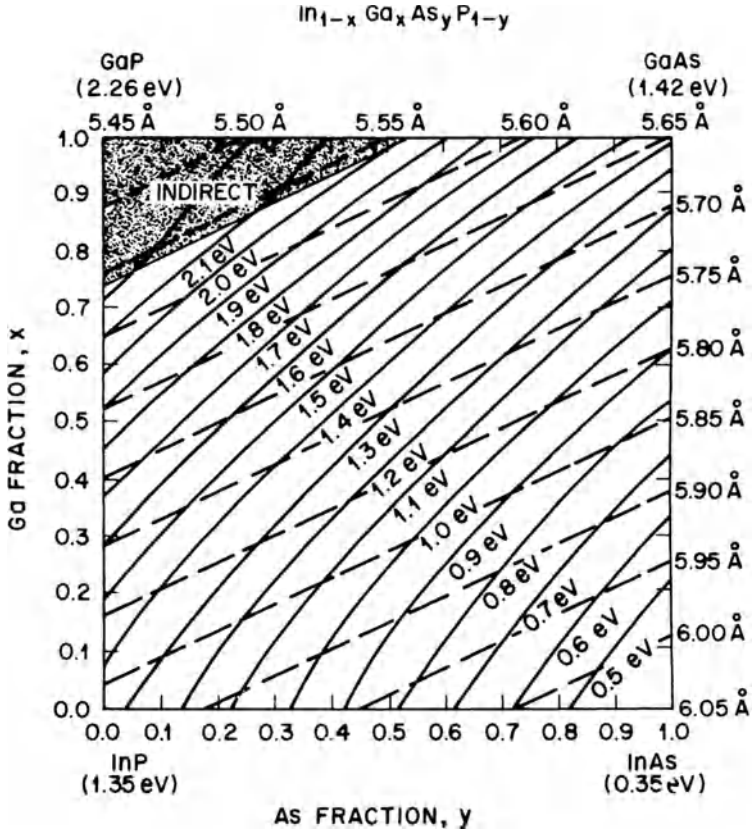


Fig. 1.5 Contours of constant band gap (solid lines) and constant lattice spacing (dashed lines) in the x–y compositional plane for $In_{1-x}Ga_xAs_yP_{1-y}$. The composition values x and y can be chosen to obtain a particular band gap (or laser wavelength) for a given lattice constant. Band gap is indirect in the shaded area. (After Ref. 32)

constructed for other heterostructure materials[24] in order to provide guidance for the lattice-matched growth of the active and cladding layers. Recent advances in the epitaxial growth techniques permit a lattice mismatch of up to a few percent without degrading the interface quality significantly. Such semiconductor lasers are referred to as *strained-layer lasers* and have attracted considerable attention because of their superior performance.[51]

To cover the longer wavelength region ($\lambda > 1.6$ μm), other material systems shown in Fig. 1.3 have been successfully exploited. Lattice-matched active layers of the quaternary $In_xGa_{1-x}As_ySb_{1-y}$ material have been grown[52] on GaSb substrate and can cover the wavelength range of 1.7–4.4 μm. Another

important class of materials is lead salts. Lead salts have been used to make semiconductor lasers emitting in the far-infrared region ranging from 3–34 μm,[53–57] although low-temperature operation is required. Two material systems of particular interest are PbTe-$Pb_{1-x}Sn_xTe$ and PbS-$PbS_{1-x}Se_x$. The longest wavelength of about 100 μm is obtained using the semiconductor material $Bi_{1-x}Sb_x$. By contrast, short-wavelength semiconductor lasers require the use of so-called II–VI semiconductors such as ZnS or ZnSe, often doped with Cd.[50]

1.3 OPERATING PRINCIPLES

The purpose of this section is to provide a qualitative understanding of the physics behind the semiconductor laser. Most of the concepts introduced here are discussed in detail in subsequent chapters. Two things are required to operate a laser: (i) a gain medium that can amplify the electromagnetic radiation propagating inside it and provide the spontaneous-emission noise input and (ii) a feedback mechanism that can confine the electromagnetic field through the well-defined optical modes. As the name itself implies, the *gain medium* for a semiconductor laser consists of a semiconductor material (see Sec. 1.2). The optical feedback is obtained using the cleaved facets that form a Fabry-Perot (FP) cavity, and the mode confinement is achieved through dielectric waveguiding. To provide the optical gain, a semiconductor laser needs to be externally pumped, and both electrical and optical pumping techniques have been used for this purpose. A simple, practical, and most commonly used method employs current injection through the use of a forward-biased *p-n* junction. Such semiconductor lasers are sometimes referred to as injection lasers or laser diodes.

1.3.1 *p-n* Junction

At the heart of a semiconductor injection laser is the *p-n* junction.[58] A *p-n* junction is formed by bringing a *p*-type and an *n*-type semiconductor into contact with each other. When they first come into contact, their quasi-Fermi levels do not match since the two are not in equilibrium. An equilibrium is, however, quickly established through diffusion of electrons from the *n* side to the *p* side, while the reverse occurs for holes. These diffusing electrons and holes recombine in the junction region. Eventually a steady state is reached in such a way that further diffusion of electrons and holes is opposed by the built-in electric field across the *p-n* junction arising from the negatively charged acceptors on the *p* side and the positively charged donors on the *n* side. The Fermi level is then continuous across the *p-n* junction, as shown

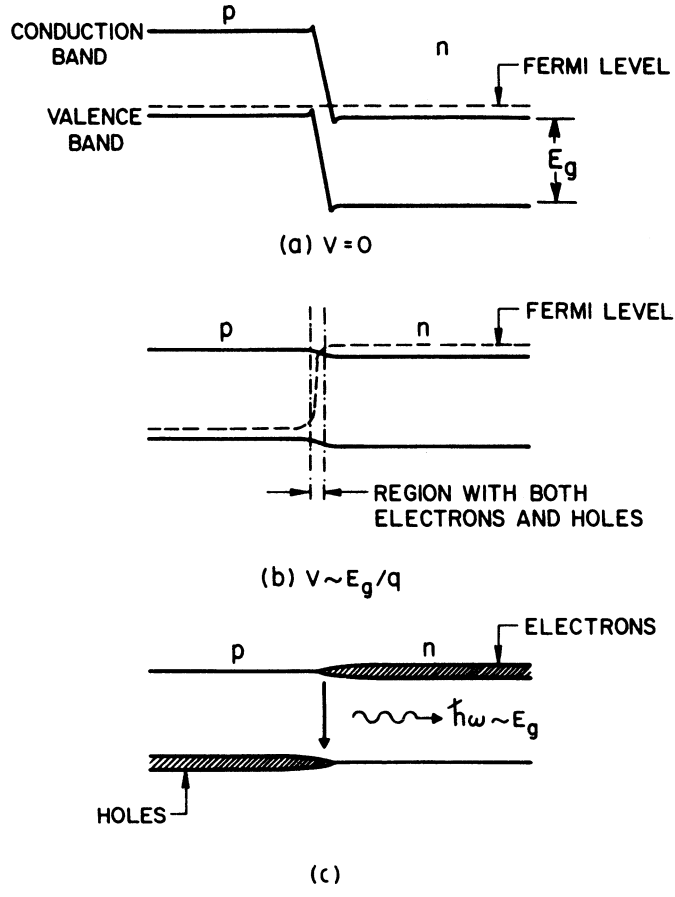


Fig. 1.6 Energy-band diagram of a *p-n* junction at (a) zero bias and (b) forward bias. (c) Schematic representation of the electron and hole densities under forward bias. Radiative recombination of electrons and holes in the narrow overlapping region generates light.

in Fig. 1.6 where the energy-band diagram of the *p-n homojunction* (junction between two similar semiconductors) is shown.

When a *p-n* junction is forward-biased by applying an external voltage, the built-in electric field is reduced, making possible a further diffusion of electrons and holes across the junction. As Fig. 1.6b shows, in a narrow depletion region both electrons and holes are present simultaneously and can recombine either radiatively or nonradiatively. Photons of energy $h\nu \cong E_g$ are emitted during radiative recombination. However, these photons can also be absorbed through a reverse process that generates electron-hole

pairs. When the external voltage exceeds a critical value, a condition known as *population inversion* is achieved, in which the rate of photon emission exceeds that of absorption. The *p-n* junction is then able to amplify the electromagnetic radiation, whose wavelength satisfies Eq. (1.2.1), and is said to exhibit *optical gain.* However, for a homojunction the thickness of the region where gain is sufficiently high is very small (~ 0.01 μm) since there is no mechanism to confine the charge carriers.

The carrier-confinement problem is solved through the use of a *p-n heterojunction.* Figure 1.7 shows the energy-band diagram for a double-heterostructure laser wherein the thin *p*-type active region has a lower band gap compared to that of the two *p*-type and *n*-type cladding layers. Electrons and holes can move freely to the active region under forward bias. However, once there, they cannot cross over to the other side because of the potential barrier resulting from the band-gap difference. This allows for a substantial build-up of the electron and hole populations inside the active region, where they can recombine to produce optical gain. The width of the gain region is determined by the active-layer thickness, typically 0.1–0.3 μm. As mentioned earlier, it was the adoption of the heterostructure scheme that resulted in significantly lower threshold current densities (compared with a homojunction) and led to the room-temperature operation of semiconductor lasers.

1.3.2 Dielectric Waveguide

The successful operation of a laser requires that the generated optical field should remain confined in the vicinity of the gain region. In double-heterostructure semiconductor lasers the optical confinement occurs by virtue of a fortunate coincidence. The active layer with a smaller band gap also has a higher refractive index compared with that of the surrounding cladding layers (see Fig. 1.2). Because of the index difference, the active layer in effect acts as a dielectric waveguide. The physical mechanism behind the confinement is total internal reflection, as illustrated in Fig. 1.8. When a ray traveling at an angle θ (measured from the interface normal) hits the interface, it is reflected back if the angle θ exceeds the critical angle given by

$$\theta = \sin^{-1} \frac{\mu_1}{\mu_2}, \tag{1.3.1}$$

where μ_1 and μ_2 are the refractive indices of the cladding and active layers respectively. Thus, rays traveling nearly parallel to the interface are trapped and constitute the waveguide mode. A more detailed discussion of waveguide modes requires the use of Maxwell's equations and is given in Chapter 2.

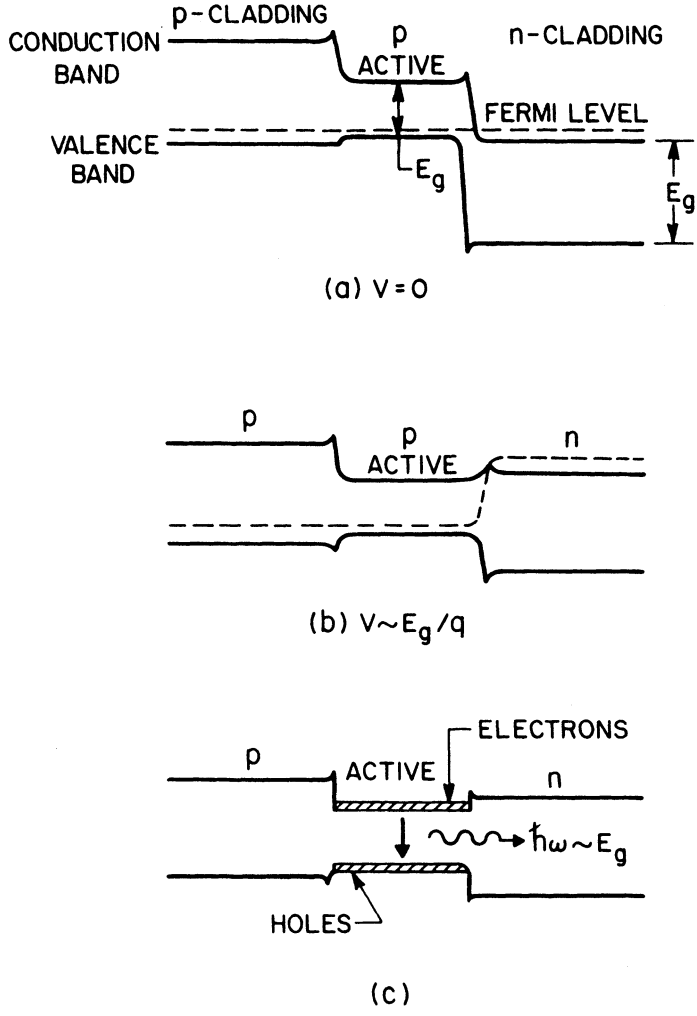


Fig. 1.7 Energy-band diagram of a double-heterostructure semiconductor laser at (a) zero bias and (b) forward bias. (c) The band-gap discontinuities at the two heterojunctions help to confine electrons and holes inside the active region, where they recombine to produce light.

1.3.3 Recombination Mechanisms

When the current flowing through a semiconductor laser is increased, charge carriers (electrons and holes) are injected into the thin active region, where they recombine through radiative or nonradiative mechanisms. As one may expect, nonradiative recombinations are not helpful for laser operation, and

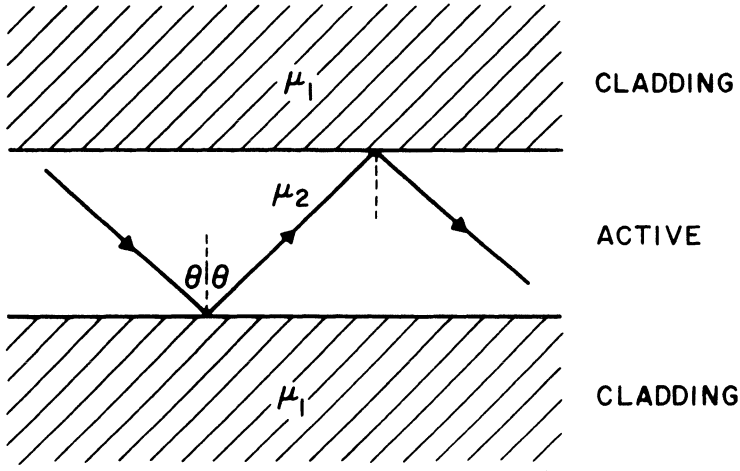


Fig. 1.8 Dielectric waveguiding in a heterostructure semiconductor laser. The relatively higher refractive index ($\mu_2 > \mu_1$) of the active layer allows total internal reflection to occur at the two interfaces for angles such that $\sin\theta > \mu_1/\mu_2$.

attempts are made to minimize their occurrence by controlling point defects and dislocations. However, a nonradiative recombination mechanism, known as the *Auger process*, is intrinsic and becomes particularly important for long-wavelength semiconductor lasers operating at room temperature and above. It is described in detail in Chapter 3. Physically speaking, during the Auger process the energy released by the electron-hole recombination is taken by a third charge carrier and is eventually lost to lattice phonons.

During a radiative recombination, the energy E_g released by the electron-hole pair appears in the form of a photon whose frequency ν or wavelength λ satisfies the energy conservation relation $E_g = h\nu = hc/\lambda$. This can happen through two optical processes known as spontaneous emission and stimulated emission. These are shown schematically in Fig. 1.9. In the case of *spontaneous emission*, photons are emitted in random directions with no phase relationship among them. *Stimulated emission*, by contrast, is initiated by an already existing photon. The remarkable feature is that the emitted photon matches the original photon not only in its wavelength but also in direction of propagation. It is this relationship between the incident and emitted photons that renders the light emitted by a laser coherent.

1.3.4 Laser Threshold

Although stimulated emission can occur as soon as current is applied to the semiconductor laser, the laser does not emit coherent light until the

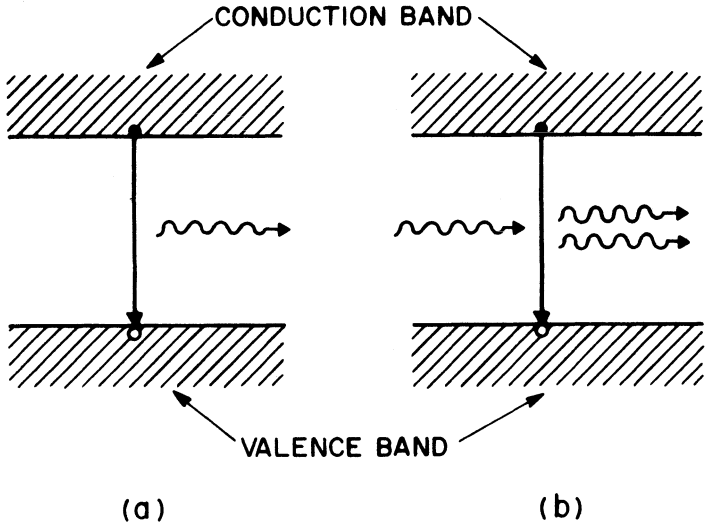


Fig. 1.9 Schematic illustration of (a) spontaneous-emission and (b) stimulated-emission processes wherein an electron-hole pair recombines to generate a photon. In the case of stimulated emission the two outgoing photons match in their frequency and direction of propagation.

current exceeds a critical value, known as the *threshold current* (I_{th}). This is so because stimulated emission has to compete against the absorption processes during which an electron-hole pair is generated at the expense of an absorbed photon. Since the electron population in the valence band generally far exceeds that of the conduction band, absorption dominates. At a certain value of the external current, a sufficient number of electrons are present in the conduction band to make the semiconductor optically transparent. With a further increase in current, the active region of the semiconductor laser exhibits optical gain and can amplify the electromagnetic radiation passing through it. Spontaneously emitted photons serve as the noise input for the amplification process.

However, optical gain alone is not enough to operate a laser. The other necessary ingredient is *optical feedback*. In semiconductor lasers it is provided by the cleaved facets that form an *FP cavity*. The role of the FP cavity is twofold. First, it provides a direction selectivity for the process of stimulated emission, since only photons traveling along its axis are reflected back and forth. Second, it provides a wavelength selectivity since the feedback is strongest for wavelengths corresponding to the longitudinal modes of the FP cavity.

Because of the optical feedback, the number of photons traveling perpendicular to the facets increases when the current is large enough to satisfy

the condition of net stimulated emission. However, some photons are lost through the partially transmitting facets and some get scattered or absorbed inside the cavity. If the loss exceeds the gain, stimulated emission cannot sustain a steady supply of photons. This is precisely what happens below threshold, when the laser output consists of mainly spontaneously emitted photons. At threshold, gain equals loss and stimulated emission begins to dominate. Over a narrow current range in the vicinity of the threshold current, the output power jumps by several orders of magnitude and the spectral width of the emitted radiation narrows considerably because of the coherent nature of stimulated emission.

In the above-threshold regime, laser output increases almost linearly with the current. Almost all electrons and holes injected into the active region now recombine through stimulated emission, and the internal quantum efficiency approaches 100%. The performance of a semiconductor laser is governed by a large number of emission characteristics related to the static, dynamic, and spectral behavior of the light output. These are discussed in Chapter 6.

1.4 OPTICAL FIBER COMMUNICATIONS

As mentioned before, an important application of semiconductor lasers is in the field of optical fiber communications. Compared to other kinds of lasers, a semiconductor laser offers two distinct advantages, namely its compact size and the possibility of direct modulation through variations in applied electric current. This section briefly discusses the important features of an optical communication system,[31] since the desirable characteristics of an InGaAsP semiconductor laser are often dictated by the performance requirements of the whole system.

In an optical communication system, information is transmitted by light propagation inside an optical fiber in the form of a coded sequence of optical pulses. The signal weakens during propagation because of fiber loss. After some distance it becomes necessary to regenerate the signal through the use of a *repeater* (essentially a detector-amplifier-transmitter combination). System cost considerations require that the repeater spacing, L, be as large as possible. Another consideration is related to the transmission capacity as determined by the bit rate, B (number of bits transmitted per second). The objective of many lightwave transmission systems is to maximize the bit rate–distance product, BL. The bit rate is limited inherently by chromatic dispersion in the fiber, which is responsible for the broadening of optical pulses during their propagation inside a single-mode fiber. The choice of the operating wavelength is therefore related to the loss and dispersion characteristics of the fiber. Figures 1.10 and 1.11 show the variation of the

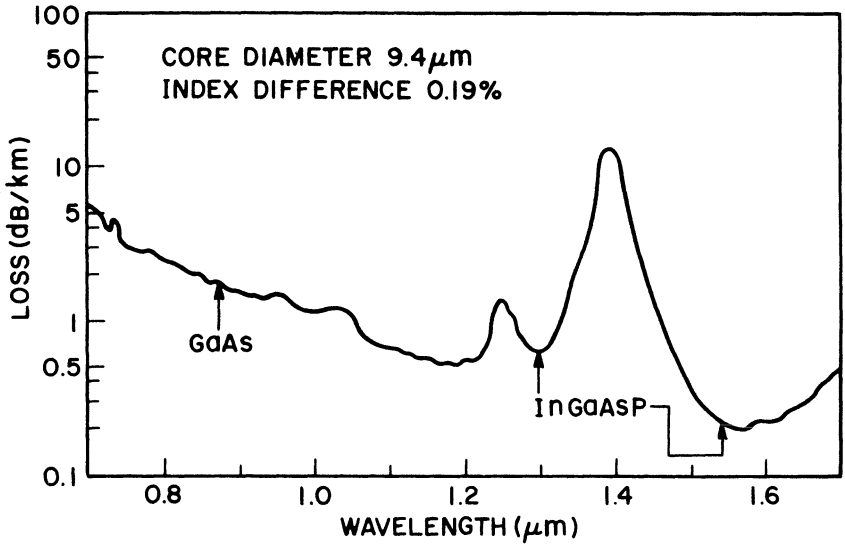


Fig. 1.10 Measured loss in a single-mode silica fiber as a function of wavelength. Arrows indicate wavelength regions used for optical fiber communications. Minimum loss occurs around the 1.55-μm wavelengths. (After Ref. 37)

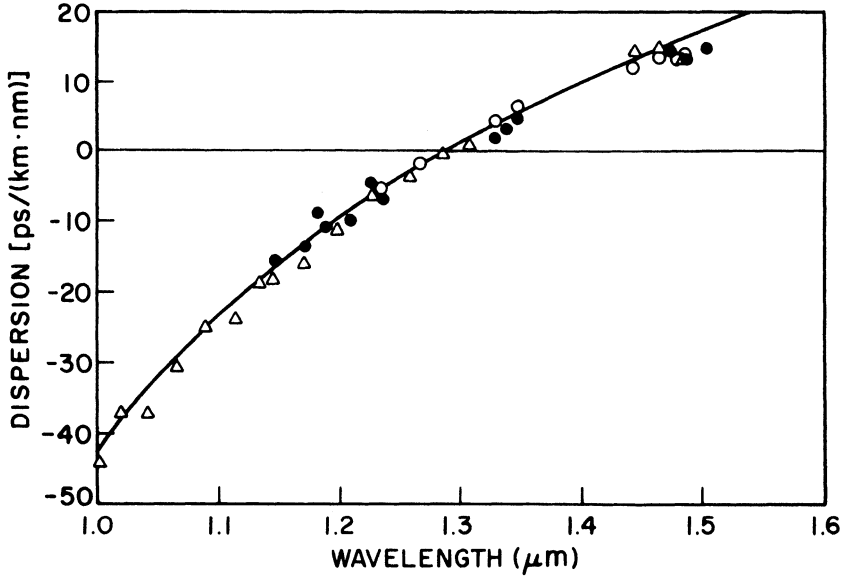


Fig. 1.11 Measured material dispersion for typical single-mode silica fibers. Data points for three fibers are shown. (After Ref. 59 © 1980 IEEE)

optical loss[37] (in dB/km) and the *dispersion coefficient*[59] [in ps/(km·nm)] as functions of the wavelength for typical single-mode silica fibers.

The first generation of lightwave transmission systems utilized GaAs lasers operating at a wavelength of about 0.85 μm. Relatively high values of the loss and dispersion coefficients restricted the repeater spacing to ~10 km and the bit rate to ~100 Mb/s. The second generation of systems made use of the wavelength region around 1.3 μm, where fiber dispersion is negligible (see Fig. 1.11). The use of InGaAsP lasers coupled with the relatively low fiber loss allowed a repeater spacing of about 20 km. However, the bit rate had to be below ~100 Mb/s because of modal dispersion in multimode fibers. This problem was overcome with the use of single-mode fibers; the absence of chromatic dispersion near 1.3 μm then allowed much higher bit rates (up to 2 Gb/s). However, the repeater spacing (~50 km) was limited by the fiber loss at this wavelength.

As is clear from Fig. 1.10, minimum fiber loss occurs around 1.55 μm. The third generation of optical communication systems is therefore based on 1.55-μm InGaAsP lasers. At this wavelength the repeater spacing can easily exceed 100 km for moderate bit rates. At high bit rates ($B \gtrsim 1$ Gb/s) the repeater spacing is limited not by the fiber loss but rather by the extent of fiber dispersion. Two distinct routes are being followed to overcome this problem. In one approach the zero-dispersion wavelength, which is about 1.3 μm for conventional silica fibers (see Fig. 1.11), is shifted towards the desirable 1.55-μm region by modifying the fiber characteristics.[60–62] In the other approach the effect of fiber dispersion is minimized by reducing the spectral width of the 1.55-μm InGaAsP laser source.[63]

To understand the role of the spectral width, $\Delta\lambda_s$, note that, during its propagation inside a fiber of length L, an optical pulse broadens and its width increases by an amount given by

$$\Delta\tau = D_f(\Delta\lambda_s)L \tag{1.4.1}$$

where D_f is the fiber dispersion coefficient having a value of ~17 ps/(km·nm) at the 1.55-μm wavelength (see Fig. 1.11). In order to arrive within its assigned time slot, the pulse should not be wider than B^{-1}, i.e.,

$$B\Delta\tau = BLD_f(\Delta\lambda_s) < 1. \tag{1.4.2}$$

Unless designed specially, an InGaAsP laser simultaneously oscillates in several longitudinal modes and the spectral width $\Delta\lambda_s \cong 5$ nm. Using the appropriate values for D_f and $\Delta\lambda_s$, the bit rate–distance product BL is limited to less than ~10 GHz·km, too small for the present lightwave transmission

systems. A repeater spacing of 100 km limits the bit rate to less than or equal to 100 Mb/s.

It is with the desire of reducing the source spectral width, $\Delta\lambda_s$, that the third generation of lightwave systems forced the development of the *single-frequency semiconductor laser*,[63] a device in which almost all of the output power appears in a single longitudinal mode. Two schemes commonly used for this purpose make use of distributed feedback and coupled-cavity mechanisms (discussed in Chapters 7 and 8). With proper design, a value of $\Delta\lambda_s$ less than ~ 0.1 nm can be achieved, increasing BL to a value of about 600 GHz·km. A BL value of 412 GHz·km was realized in 1985 in a transmission experiment[64] employing a 1.55-μm distributed feedback laser. Essentially error-free transmission (less than one error per billion bits) was achieved over a 103-km-long fiber with a bit rate of 4 Gb/s. The bit rate of third-generation lightwave systems has been extended to beyond 10 Gb/s by using distributed feedback lasers in combination with dispersion-shifted fibers.[65,66]

The fourth-generation of optical communication systems makes use of homodyne or heterodyne detection techniques. Because of the phase-sensitive nature of such systems, they are referred to as coherent communication systems.[67,68] For the same reason, they require tunable semiconductor lasers with a narrow linewidth (usually < 10 MHz or 10^{-4} nm). Special multisection lasers have been developed to meet the demands of fourth-generation lightwave systems.[69–71] The fifth generation of lightwave systems employs erbium-doped fiber amplifiers[72,73] for amplifying the transmitted optical signal periodically to compensate for fiber loss. The transmission distance can exceed a few thousand kilometers for such systems as long as fiber dispersion does not limit it.[74–76] A novel approach makes use of optical solitons that use fiber nonlinearity for compensating fiber dispersion.[77,78] Such solitons can travel along the fiber for thousands of kilometers without experiencing temporal broadening as long as fiber loss is compensated through periodic amplification.[79,80] The development of such fifth-generation lightwave systems requires new kinds of semiconductor lasers. For instance, semiconductor lasers operating at wavelengths of 0.98 and 1.48 μm have been developed for pumping erbium-doped fiber amplifiers.[81,82] The use of optical solitons requires semiconductor lasers capable of producing ultrashort optical pulses (pulse width < 50 ps) at high repetition rates. Laboratory experiments have been performed at a bit rate as high as 32 Gb/s by using mode-locked semiconductor lasers.[83] Transmission distances in excess of 10,000 km at a bit rate of up to 10 Gb/s have been demonstrated by using a recirculating-loop configuration.[84–86] Advances in the semiconductor-laser technology play an important role in realizing such high-performance lightwave systems.

1.5 OVERVIEW

The book is intended to provide a comprehensive account of semiconductor lasers. As Fig. 1.3 shows, a large number of semiconductor materials can be used to cover the wavelength range of 0.4 to 100 μm. An enormous amount of the research and development effort is, however, directed towards AlGaAs and InGaAsP lasers because of their commercial use in optical recording and fiber communication systems. For this reason, particular attention is paid to AlGaAs and InGaAsP lasers, which cover the near-infrared range of 0.8–1.6 μm. Chapter 13 considers longer-wavelength semiconductor lasers. It should be stressed however, that the theoretical aspects of semiconductor lasers discussed in this text apply to all semiconductor lasers. Following is a brief outline of the various chapters.

Chapter 2 introduces the basic concepts involved in operating a semiconductor laser. Starting from Maxwell's equations, the gain and loss considerations are used to obtain the threshold condition. The required material gain and its dependence on the injected current density are discussed within the framework of a simple phenomenological model. The text then introduces the longitudinal, transverse, and lateral modes associated with the optical field. In particular, the dependence of the lateral modes on gain and index guiding is discussed. Finally, emission characteristics used to characterize the laser output are presented after considering the static, spatial, spectral, and dynamic aspects.

Chapter 3 describes recombination mechanisms, both radiative and nonradiative. For the case of radiative recombination, the expressions are obtained for the spontaneous-emission rate and the absorption coefficient using a simple but fairly realistic model of the band structure. Varying the absorption coefficient with the injected current density shows how the optical gain depends on various parameters such as carrier density and doping levels. For the case of nonradiative recombination, particular attention is paid to the Auger process, which becomes increasingly important for long-wavelength semiconductor lasers. The last section is devoted to studying the temperature dependence of the threshold behavior of semiconductor lasers.

Chapter 4 reviews the epitaxial growth techniques used to make heterostructure semiconductor lasers. The liquid-phase, the vapor-phase, and the molecular epitaxial growth techniques are introduced with emphasis on their use for the fabrication of InGaAsP lasers. The chapter also provides a brief discussion of the material parameters and their measurement techniques.

Chapter 5 introduces various laser structures that have been used for semiconductor lasers. Broad-area semiconductor lasers are considered together with the dependence of their threshold current density on the active-layer thickness. Next, gain-guided structures and their performance are described.

The attention is then paid to various index-guided structures and to a comparison of their light-current characteristics.

The emission characteristics of semiconductor lasers are described in Chapter 6 by considering their static, dynamic, spectral, noise, and modulation performance. These properties are described theoretically using a set of rate equations. The emphasis has been on studying the performance of InGaAsP lasers when used as a source in optical communication systems. The discussion pays particular attention to the dynamic line broadening that results from the frequency chirping associated with direct modulation.

Chapters 7 and 8 deal with single-frequency semiconductor lasers, which emit predominantly in a single longitudinal mode. Distributed-feedback (DFB) semiconductor lasers are considered in Chapter 7. After a brief discussion of DFB structures, the theory of wave propagation in a periodic medium is used to obtain the longitudinal modes and corresponding threshold gains for a DFB semiconductor laser. The performance of state-of-the-art DFB lasers is discussed with particular emphasis on their application in recent transmission experiments. Chapter 8 considers coupled-cavity semiconductor lasers. In such lasers, single-longitudinal-mode operation is achieved using an interference effect that requires the laser mode to be simultaneously resonant with respect to both cavities. Various schemes and their relative performance are discussed with particular attention paid to the cleaved-coupled cavity semiconductor laser.

Chapter 9 describes quantum-well semiconductor lasers. In these lasers the active layer is made so thin that the quantum-mechanical effects associated with the confinement of charge carriers to the thin active region become important, leading to quantization of the conduction- and valence-band energies in one direction. This modifies the density of the states, which significantly affects radiative and nonradiative recombination rates. Changes in the emission characteristics of a semiconductor laser arising from the quantum-well effect are discussed both for single and multiple quantum wells.

Chapter 10 is devoted to the discussion of surface-emitting semiconductor lasers. Particular emphasis is placed on the design of vertical-cavity surface-emitting semiconductor lasers. Such lasers have attracted considerable interest because of their potential applications in the optical interconnect technology and other technologies based on the use of optoelectronic integrated circuits. The laser cavity is extremely short (~ 1 μm) in such lasers and is often referred to as the microcavity. It is formed by depositing alternating layers of low- and high-index dielectric materials that act as a Bragg reflector of high reflectivity ($>99\%$). Surface-emitting semiconductor lasers based on both the GaAs and InP technologies are discussed.

Chapter 11 considers the design and performance of semiconductor laser amplifiers. It begins by introducing the general concepts common to all

amplifiers such as the gain spectrum, amplifier bandwidth, gain saturation, and amplifier noise. It then describes the techniques used to realize ultrasmall facet reflectivities needed to make high-quality amplifiers. The chapter then describes the operating characteristics of semiconductor laser amplifiers such as the gain, bandwidth, noise, gain saturation, and the amplifier performance when multiple channels and short optical pulses are amplified.

Chapter 12 describes the progress made in the fabrication of photonic and optoelectronic integrated circuits. It begins with the description of photonic integrated circuits such as multiple lasers in the form of an array fabricated on the same chip and integration of the laser with detectors and modulators. Optoelectronic integrated circuits are described next by considering integrated transmitters, receivers, and regenerators. Optoelectronic integrated circuits capable of performing logic functions are also discussed briefly.

Chapter 13 considers semiconductor lasers that emit in the infrared and visible regions. Particular attention is paid to lead-salt lasers, which, through proper choice of material composition, can be made to cover a wide wavelength range from 3–34 μm. Starting from a brief description of their band structure and other physical properties, we consider radiative and nonradiative recombinations in lead-salt compounds. Auger recombination becomes particularly important in these long-wavelength semiconductor lasers; low-temperature operation is necessary to reduce the rate of Auger recombination. The fabrication and operating characteristics of lead-salt lasers are discussed with emphasis on their wavelength tuning range.

Chapter 14 discusses issues of device reliability related to the degradation of semiconductor lasers with age. Various kinds of defects and the techniques used for their identification are described. The chapter discusses the concept of accelerated aging, its use for reliability assurance, and experimental results.

PROBLEMS

1.1 Calculate the optical frequency (in units of THz) for semiconductor lasers operating at 0.8, 1.55, and 10 μm. What should be the band gap of the active-layer material in each case assuming that the photon energy nearly equals the band-gap energy? Use Fig. 1.3 to determine the semiconductor materials which can be used to make such lasers.

1.2 Explain what is meant by direct or indirect band gap in semiconductors. Give an example of each type of semiconductor.

1.3 Explain the difference beween a homojunction and a heterojunction. How does the use of heterojunctions lead to confinement of charge carriers and photons to the active region of semiconductor lasers?

1.4 The quaternary material $In_{1-x}Ga_xAs_yP_{1-y}$ is lattice-matched to InP whenever $x = 0.45y$. The band gap is found to vary with y as

$$E_g(y) = 1.35 - 0.72y + 0.12y^2.$$

Find the composition of the active layer for semiconductor lasers designed to operate at 1.3- and 1.55-μm wavelengths. What are the shortest and the longest wavelengths at which such lasers can operate?

1.5 What are the main differences between a semiconductor laser and a gas laser?

1.6 Describe the five generations of optical communication systems. What kind of semiconductor lasers are needed for each generation?

REFERENCES

1. Schawlow, A. L., and C. H. Townes. *Phys. Rev.* **112,** 1940 (1958).
2. Maiman, T. H. *Nature* **187,** 493 (1960).
3. Javan, A., W. R. Bennett, Jr., and D. R. Herriot. *Phys. Rev. Lett.* **6,** 106 (1961).
4. Basov, N. G., O. N. Krokhin, and Yu. M. Popov. *Sov. Phys. JETP* **13,** 1320 (1961).
5. Bernard, M. G. A., and G. Duraffourg. *Phys. Status Solidi* **1,** 699 (1961).
6. Dumke, W. P. *Phys. Rev.* **127,** 1559 (1962).
7. Hall, R. N., G. E. Fenner, J. D. Kingsley, T. J. Soltys, and R. O. Carlson. *Phys. Rev. Lett.* **9,** 366 (1962).
8. Nathan, M. I., W. P. Dumke, G. Burns, F. H. Dill, Jr., and G. Lasher. *Appl. Phys. Lett.* **1,** 62 (1962).
9. Quist, T. M., R. H. Rediker, R. J. Keyes, W. E. Krag, B. Lax, A. L. McWhorter, and H. J. Zeiger. *Appl. Phys. Lett.* **1,** 91 (1962).
10. Holonyak, N., Jr., and S. F. Bevacqua. *Appl. Phys. Lett.* **1,** 82 (1962).
11. Kroemer, H. *Proc. IEEE* **51,** 1782 (1963).
12. Alferov, Zh. I., and R. F. Kazarinov. Authors certificate 181737 (U.S.S.R.), 1963.
13. Kressel, H., and H. Nelson. *RCA Rev.* **30,** 106 (1969).
14. Hayashi, I., M. B. Panish, and P. W. Foy. *IEEE J. Quantum Electron.* **QE-5,** 211 (1969).
15. Alferov, Zh. I., V. M. Andreev, E. L. Portnoi, and M. K. Trukan. *Sov. Phys. Semicond.* **3,** 1107 (1970). [Translated from *Fiz. Tekh. Poluprovodn.* **3,** 1328 (1969)].
16. Rupprecht, H., J. M. Woodall, and G. D. Pettit. *Appl. Phys. Lett.* **11,** 81 (1967).
17. Hayashi, I., M. B. Panish, P. W. Foy, and S. Sumski. *Appl. Phys. Lett.* **17,** 109 (1970).
18. Alferov, Zh. I., V. M. Andreev, D. Z., Garbuzov, Yu. V. Zhilyaev, E. P. Morozov, E. L. Portnoi, and V. G. Trofim. *Sov. Phys. Semicond.* **4,** 1573 (1971). [Translated from *Fiz. Tekh. Poluprovodn.* **4,** 1826 (1970)].
19. Ettenberg, M. *Appl. Phys. Lett.* **27,** 652 (1975).
20. Dyment, J. C. *Appl. Phys. Lett.* **10,** 84 (1967).
21. Ripper, J. E., J. C. Dyment, L. A. D'Asaro, and T. L. Paoli. *Appl. Phys. Lett.* **18,** 155 (1971).
22. Cook, D. D., and F. R. Nash. *J. Appl. Phys.* **46,** 1660 (1975).
23. Kressel, H., and J. K. Butler. *Semiconductor Lasers and Heterojunction LEDs.* New York: Academic Press, 1977.
24. Casey, H. C., Jr., and M. B. Panish. *Heterostructure Lasers,* Parts A and B. New York: Academic Press, 1978.

25. Thompson, G. H. B. *Physics of Semiconductor Laser Devices.* Chichester: John Wiley & Sons, 1980.
26. Dupuis, R. D. *J. Crys. Growth* **55,** 213 (1981).
27. Cho, A. Y. *J. Vac. Sci. Tech.* **16,** 275 (1979).
28. Dingle, R., W. Wiegmann, and C. H. Henry. *Phys. Rev. Lett.* **33,** 827 (1974); Tsang, W. T., C. Weisbuch, R. C. Miller, and R. Dingle. *Appl. Phys. Lett.* **35,** 673 (1979).
29. Scifres, D. R., W. Streifer, and R. D. Burnham. *Appl. Phys. Lett.* **33,** 1015 (1978).
30. Sakamoto, M., J. G. Endriz, and D. R. Scifres. *Electron. Lett.* **28,** 178 (1992); **28,** 197 (1992).
31. Suematsu, Y. *Proc. IEEE* **71,** 692 (1983); Li, T. *IEEE J. Sel. Areas Commun.* **SAC-1,** 356 (1983).
32. Moon, R. L., G. A. Antypas, and L. W. James. *J. Electron. Mater.* **3,** 635 (1974).
33. Bogatov, A. P., L. M. Dolginov, P. G. Eliseev, M. G. Mil'vidskii, B. N. Sverdlov, and E. G. Shevchenko. *Sov. Phys. Semicond.* **9,** 1282 (1975).
34. Hsieh, J. J., J. A. Rossi, and J. P. Donnelly. *Appl. Phys. Lett.* **28,** 709 (1976).
35. Oe, K., S. Ando, and K. Sugiyama. *Jpn. J. Appl. Phys.* **16,** 1273 (1977).
36. Horiguchi, M., and H. Osani. *Electron. Lett.* **12,** 310 (1976).
37. Miya, T., Y. Terunuma, T. Hosaka, and T. Miyoshita. *Electron. Lett.* **15,** 106 (1979).
38. Kobayashi, N., and Y. Horikoshi. *Jpn. J. Appl. Phys.* **18,** 1005 (1979).
39. Akiba, S., K. Sakai, Y. Matsushima, and T. Yamamoto. *Electron. Lett.* 606 (1979).
40. Henshall, G. D., and P. D. Greene. *Electron. Lett.* **15,** 621 (1979).
41. Kawaguchi, H., T. Takahei, Y. Toyoshima, H. Nagai, and G. Iwane. *Electron. Lett.* **15,** 669 (1979).
42. Kaminow, I. P., R. E. Nahory, M. A. Pollack, L. W. Stulz, and J. C. Dewinter. *Electron. Lett.* **15,** 763 (1979).
43. Arai, S., M. Asada, Y. Suematsu, and Y. Itaya. *Jpn. J. Appl. Phys.* **18,** 2333 (1979).
44. Botez, D., and G. Herskowitz. *Proc. IEEE* **68,** 689 (1980).
45. Suematsu, Y., K. Iga, and K. Kishino, in *GaInAsP Alloy Semiconductors*, ed. T. P. Pearsall. New York: John Wiley & Sons, 1982.
46. Nelson, R. J., and N. K. Dutta, in *Semiconductor and Semimetals*, Vol. 22, Part C, ed. W. T. Tsang. New York: Academic Press, 1985.
47. Runge, P. K., and P. R. Trischitta. *J. Lightware Technol.* **LT-2,** 744 (1984).
48. Bartolini, R. A., A. E. Bell, and F. W. Spong. *IEEE J. Quantum Electron.* **QE-17,** 69 (1981).
49. Asbeck, P. M., D. A. Cammack, J. J. Daniele, D. Lou, J. P. J. Heemskerk, W. J. Klenters, and W. H. Ophey. *Appl. Phys. Lett.* **34,** 835 (1979).
50. Hasse, M. A., J. Qiu, J. M. DePuydt, and H. Cheng. *Appl. Phys. Lett.* **59,** 1272 (1991).
51. Thijs, P. J. A., L. F. Tiemeijer, P. I. Kuindersma, J. J. M. Binsma, and T. van Dongen. *IEEE J. Quantum Electron.* **27,** 1426 (1991).
52. Kobayashi, N., Y. Horikoshi, and D. Uemura. *Jpn. J. Appl. Phys.* Part 2, **19,** L30 (1980).
53. Groves, S. H., K. W. Nill, and A. J. Strauss. *Appl. Phys. Lett.* **25,** 331 (1974).
54. Tomasetta, L. R., and C. G. Fonstand. *Appl. Phys. Lett.* **25,** 440 (1974).
55. Walpole, J. N., A. R. Calawa, T. C. Harman, and S. H. Groves. *Appl. Phys. Lett.* **28,** 552 (1976).
56. McLane, G. F., and K. J. Sleger, *J. Electron. Mater.* **4,** 465 (1975).
57. Preier, H., M. Bleicher, W. Riedel, and H. Maier. *Appl. Phys. Lett.* **28,** 669 (1976).
58. Sze, S. M. *Physics of Semiconductor Devices*, Second Ed. New York: John Wiley & Sons, 1981.
59. Sugimura, A., K. Daikoku, N. Imoto, and T. Miya. *IEEE J. Quantum Electron.* **QE-16,** 215 (1980).
60. Cohen, L. G., C. Lin, and W. G. French. *Electron. Lett.* **15,** 334 (1979).
61. Tsuchiya, H., and N. Imoto. *Electron. Lett.* **15,** 476 (1979).

62. Okamoto, K., T. Edahiro, A. Kawana, and T. Miya. *Electron. Lett.* **15,** 729 (1979).
63. Agrawal, G. P. Vol. 26, Chap. 3, in *Progress in Optics*, ed. by E. Wolf. Amsterdam: North-Holland, 1988.
64. Gnauck, A. H., B. L. Kasper, R. A., Linke, R. W. Dawson, T. L. Koch, T. J. Bridges, E. G. Burkhardt, R. T. Yen, D. P. Wilt, J. C. Campbell, K. C. Nelson, and L. G. Cohen. *J. Lightwave Technol.* **LT-3**, 1032 (1985).
65. Gimlett, J. L., M. J. Iqbal, J. Young, L. Curtis, R. Spicer, and N. K. Cheung. *Electron. Lett.* **25,** 596 (1989).
66. Fujita, S., M. Mitamura, T. Torikai, N. Henmi, H. Yamada, T. Suzaki, I. Takano, and M. Shikada. *Electron. Lett.* **25,** 702 (1989).
67. Okoshi, T., and K. Kikuchi. *Coherent Optical Fiber Communications.* Boston: Kluwer Academic, 1988.
68. Agrawal, G. P. Chap. 6, in *Fiber-Optic Communication Systems.* New York: Wiley Interscience, 1992.
69. Kobayashi, K., and I. Mito. *J. Lightwave Technol.* **6,** 1623 (1988).
70. Kotaki, Y., M. Matsuda, H. Ishikawa, and H. Imai. *Electron. Lett.* **24,** 503 (1988); Y. Kotaki and H. Ishikawa. *IEE Proc. J.* **138,** 171 (1991).
71. Koch, T. L., U. Koren, and B. J. Miller. *Appl. Phys. Lett.* **53,** 1036 (1988); T. L. Koch and U. Koren. *J. Lightwave Technol.* **8,** 274 (1990).
72. Special issue on optical amplifiers. *J. Lightwave Technol.* **9,** 145–296 (1991).
73. Desurvire, E. *Scientific American* **266** (1), 114 (1992).
74. Saito, S., T. Imai, and T. Ito. *J. Lightwave Technol.* **9,** 161 (1991).
75. Ryu, S., S. Yamamoto, H. Taga, N. Edagawa, Y. Yoshida, and H. Wakabayashi. *J. Lightwave Technol.* **9,** 251 (1991).
76. Nakagawa, K., S. Nishi, K. Aida, and E. Yoneda. *J. Lightwave Technol.* **9,** 198 (1991).
77. Hasegawa, A. *Solitons in Optical Fibers.* Berlin: Springer-Verlag, 1989.
78. Agrawal, G. P., Chap. 2, in *Contemporary Nonlinear Optics*, ed. by G. P. Agrawal and R. W. Boyd. Boston: Academic Press, 1992.
79. Mollenauer, L. F., J. P. Gordon, and S. G. Evangelides. *Laser Focus World* **27** (11), 159 (1991).
80. Nakazawa, M. Proc. Europ. Conf. Opt. Commun., Paris, Sept. 1991, pp. 150–64.
81. Okayasu, M. M., Fukuda, T. Takeshita, and S. Uehara. *IEEE Photon. Technol. Lett.* **2,** 689 (1990).
82. Asano, H., S. Takano, M. Kawaradani, M. Kitamura, and I. Mito. *IEEE Photon. Technol. Lett.* **3,** 415 (1991).
83. Andrekson, P. A., N. A. Olsson, M. Haner, J. R. Simpson, T. Tanbun-Ek, R. A. Logan, D. Coblentz, H. M. Presby, and K. W. Wecht. *IEEE Photon. Technol. Lett.* **4,** 76 (1992).
84. Mollenauer, L. F., M. J. Neubelt, M. Haner, E. Lichtman, S. G. Evangelides, and B. M. Nyman. *Electron. Lett.* **27,** 2055 (1991).
85. Bergano, N. S., J. Aspell, C. R. Davidson, P. R. Trischitta, B. M. Nyman, and F. W. Kerfort. *Electron. Lett.* **27,** 1889 (1991).
86. Mollenauer, L. F., E. Lichtman, G. T. Harvey, M. J. Neubelt, and B. M. Nyman. *Electron. Lett.* **28,** 792 (1992).

Chapter 2
BASIC CONCEPTS

2.1 INTRODUCTION

This chapter is intended to provide a review of the basic concepts in laser theory[1–3] as they apply particularly to semiconductor lasers.[4–6] In general terms, a laser is an externally pumped self-sustained oscillator and consists of a gain medium that is placed inside an optical cavity to provide the necessary feedback. Various kinds of lasers differ only in their choice of a suitable gain medium and the pumping mechanism. In semiconductor lasers a semiconductor material is electrically pumped using a forward-biased *p-n* diode structure, and charge carriers injected into a thin active layer provide the optical gain. No external cavity is required since cleaved facets of the semiconductor gain medium itself can provide sufficient optical feedback. The injected current density J governs the strength of external pumping. The laser threshold is reached when J reaches a critical value J_{th} at which the gain is sufficient to overcome the cavity losses. Any further increase in J leads to light emission by stimulated emission. The general objective is to study the static, dynamic, and spectral characteristics of the emitted light as a function of the pump parameter J.

In spite of the generic similarity among lasers, a semiconductor laser differs significantly in the details of its operation[4–6] from other kinds such as a gas laser, where the gain medium is modeled as a collection of independent two-level atoms or molecules.[2] In both cases the electrons participate in the same three types of basic interactions, namely spontaneous emission, absorption, and stimulated emission. However, because of the delocalized nature of electronic states in the energy bands of a semiconductor laser, the electrons interact with each other through various intraband collision processes and should be treated collectively. Further, the charge carriers, both electrons and holes, can diffuse spatially. The inclusion of carrier diffusion leads to a spatially inhomogeneous gain profile that affects the spatial variation of the optical mode through stimulated emission.

The nature of optical modes supported by the semiconductor laser is also different from those encountered in other kinds of lasers. In most lasers the modes are of the *transverse-electromagnetic* (TEM) type, which are supported by the external cavity and are well approximated by the free-space Hermite-Gaussian modes.[7] In a semiconductor laser the optical modes are *transverse-electric* (TE) or *transverse-magnetic* (TM),[8] both resulting from dielectric waveguiding in a heterostructure device. Further, the cold-cavity modes are often influenced by external pumping that changes both the gain and the index of refraction. The significant variation in refractive index with external pumping (the injected current density J) is unique to semiconductor lasers and affects their operating characteristics in important ways. Among other things, it makes the cavity longitudinal modes shift their frequencies with the pumping.

This chapter is organized as follows. In Sec. 2.2 we introduce Maxwell's equations and the optical parameters required to characterize the laser medium, such as the dielectric constant, the refractive index, and the absorption coefficient. Section 2.3 considers the plane-wave solutions of the wave equation and obtains the threshold condition for a semiconductor laser. This introduces the concept of longitudinal modes associated with a Fabry-Perot cavity. The dependence of the gain on the pumping parameter, the current density, is considered in Sec. 2.4. Current injection in a forward-biased p-n junction provides a high density of charge carriers (electrons and holes) confined to the active region by the heterostructure. These charge carriers recombine to produce gain. Various recombination mechanisms are discussed briefly in Sec. 2.4. Section 2.5 is devoted to the transverse and lateral modes supported by the dielectric waveguide. Finally, Sec. 2.6 considers the emission properties of semiconductor lasers and introduces important concepts such as the threshold current, the differential quantum efficiency, near and far fields, the power spectrum, the line width, and the frequency chirp.

2.2 MAXWELL'S EQUATIONS

Since the mathematical description of all optical phenomena is based on Maxwell's equations, it is appropriate to start our discussion of semiconductor lasers by considering these equations in some detail. In the MKS system of units, the field equations take the following form:[9]

$$\nabla \times \mathscr{E} = -\partial \mathscr{B}/\partial t \tag{2.2.1}$$

$$\nabla \times \mathscr{H} = \mathscr{J} + \partial \mathscr{D}/\partial t \tag{2.2.2}$$

$$\nabla \cdot \mathscr{D} = \rho_f \tag{2.2.3}$$

$$\nabla \cdot \mathscr{B} = 0 \tag{2.2.4}$$

where $\mathscr{E}$ and $\mathscr{H}$ are the electric and magnetic field vectors, respectively, and $\mathscr{D}$ and $\mathscr{B}$ are the corresponding electric and magnetic flux densities. The current density vector $\mathscr{J}$ and the charge density ρ_f represent the sources for the electromagnetic field.

The flux densities $\mathscr{D}$ and $\mathscr{B}$ arise in response to the electric and magnetic fields $\mathscr{E}$ and $\mathscr{H}$ propagating inside the medium. In general their relationship depends on details of the matter-radiation interaction. For a nonmagnetic dielectric medium the relationship can be expressed in terms of the constitutive relations given by

$$\mathscr{D} = \epsilon_0 \mathscr{E} + \mathscr{P} \tag{2.2.5}$$

$$\mathscr{B} = \mu_0 \mathscr{H} \tag{2.2.6}$$

$$\mathscr{J} = \sigma \mathscr{E} \tag{2.2.7}$$

where ϵ_0 is the vacuum permittivity, μ_0 is the vacuum permeability, and σ is the conductivity of the medium. The induced electric polarization $\mathscr{P}$ is calculated quantum mechanically; for a semiconductor material its evaluation requires the knowledge of the Bloch wave functions and the density of states for the conduction and valence bands.

Maxwell's equations can be used to obtain the wave equation that describes the propagation of an optical field inside the medium. We take the curl of Eq. (2.2.1) to obtain

$$\nabla \times \nabla \times \mathscr{E} = -\mu_0 \frac{\partial}{\partial t} (\nabla \times \mathscr{H}) \tag{2.2.8}$$

where Eq. (2.2.6) has been used. With the help of Eqs. (2.2.2), (2.2.5), and (2.2.7), we can eliminate $\mathscr{H}$, $\mathscr{J}$, and $\mathscr{D}$ in favor of $\mathscr{E}$ and $\mathscr{P}$ to obtain

$$\nabla \times \nabla \times \mathscr{E} = -\mu_0 \sigma \frac{\partial \mathscr{E}}{\partial t} - \mu_0 \epsilon_0 \frac{\partial^2 \mathscr{E}}{\partial t^2} - \mu_0 \frac{\partial^2 \mathscr{P}}{\partial t^2}. \tag{2.2.9}$$

The left-hand side of Eq. (2.2.9) can be simplified by using the vector identity

$$\nabla \times \nabla \times \mathscr{E} = \nabla(\nabla \cdot \mathscr{E}) - \nabla^2 \mathscr{E}. \tag{2.2.10}$$

In the absence of free charges, $\rho_f = 0$, and from Eqs. (2.2.3) and (2.2.5) we obtain

$$\nabla \cdot \mathscr{D} = \epsilon_0 \nabla \cdot \mathscr{E} + \nabla \cdot \mathscr{P} = 0. \tag{2.2.11}$$

The term $\nabla \cdot \mathscr{P}$ is negligible in most cases of practical interest and, consequently, to a good degree of approximation, $\nabla \cdot \mathscr{E} = 0$ in Eq. (2.2.10). Equation (2.2.9) then becomes

$$\nabla^2 \mathscr{E} - \frac{\sigma}{\epsilon_0 c^2} \frac{\partial \mathscr{E}}{\partial t} - \frac{1}{c^2} \frac{\partial^2 \mathscr{E}}{\partial t^2} = \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \mathscr{P}}{\partial t^2} \tag{2.2.12}$$

where we have used the familiar relation

$$\mu_0 \epsilon_0 = \frac{1}{c^2} \tag{2.2.13}$$

and c is the speed of light in vacuum. The wave equation (2.2.12) is valid for arbitrary time-varying fields. Of particular interest are the optical fields with harmonic time variations since any field can be decomposed into its sinusoidal Fourier components. Using the complex notation, we write

$$\mathscr{E}(x, y, z, t) = \text{Re}[\mathbf{E}(x, y, z) \exp(-i\omega t)] \tag{2.2.14}$$

$$\mathscr{P}(x, y, z, t) = \text{Re}[\mathbf{P}(x, y, z) \exp(-i\omega t)] \tag{2.2.15}$$

where $\omega = 2\pi\nu$ and is the angular frequency and $\nu = c/\lambda$ and is the oscillation frequency of the optical field at the vacuum wavelength λ. The notation Re stands for the real part of the bracketed expression. Note that $\mathbf{E}$ and $\mathbf{P}$ are generally complex since they contain the phase information. Using Eqs. (2.2.14) and (2.2.15) in Eq. (2.2.12), we obtain

$$\nabla^2 \mathbf{E} + k_0^2[1 + i\sigma/(\epsilon_0 \omega)]\mathbf{E} = -(k_0^2/\epsilon_0)\mathbf{P} \tag{2.2.16}$$

where $k_0 = \omega/c = 2\pi/\lambda$ is the vacuum wave number.

Under steady-state conditions the response of the medium to the electric field is governed by the susceptibility χ defined by

$$\mathbf{P} = \epsilon_0 \chi(\omega)\mathbf{E} \tag{2.2.17}$$

where the frequency dependence of χ has been explicitly shown to emphasize the dispersive nature of the medium response. In general, χ is a second-rank tensor. For an isotropic medium, χ is a scalar. It is useful to decompose χ into two parts

$$\chi = \chi_0 + \chi_p \tag{2.2.18}$$

where χ_0 is the medium susceptibility in the absence of external pumping and χ_p is the additional contribution to the susceptibility related to the strength of pumping. In the case of semiconductor lasers, current injection is the source of pumping and χ_p depends on the concentration of charge carriers (electrons and holes) in the active layer. For the moment we leave χ_p unspecified except for noting that both χ_0 and χ_p are generally complex and frequency-dependent.

If we use Eq. (2.2.17) to eliminate the electric polarization $\mathbf{P}$ in Eq. (2.2.16), we obtain the time-independent wave equation

$$\nabla^2\mathbf{E} + \epsilon k_0^2\mathbf{E} = 0 \tag{2.2.19}$$

where we have introduced the complex dielectric constant

$$\begin{aligned}\epsilon &= \epsilon' + i\epsilon'' \\ &= \epsilon_b + i\,\mathrm{Im}(\chi_0) + \chi_p + i\sigma/(\epsilon_0\omega)\end{aligned} \tag{2.2.20}$$

and $\epsilon_b = 1 + \mathrm{Re}(\chi_0)$ is the background dielectric constant of the unpumped material and is real, as defined. The notations Re and Im stand for the real and imaginary parts, respectively.

The wave equation (2.2.19) can be used to obtain the spatial mode structure of the optical field. However, considerable insight can be gained by considering plane-wave solutions of Eq. (2.2.19) even though these are not the spatial modes of the semiconductor laser. In place of using the complex dielectric constant given by Eq. (2.2.20), the propagation characteristics of a plane wave in a medium are conveniently described in terms of two optical constants, the index of refraction μ and the absorption coefficient α. Consider a plane wave propagating in the positive z direction such that

$$\mathbf{E} = \hat{\mathbf{x}}E_0 \exp(i\tilde{\beta}z) \tag{2.2.21}$$

where $\hat{\mathbf{x}}$ is the polarization unit vector and E_0 is the constant amplitude. The complex propagation constant $\tilde{\beta}$ is determined by substituting Eq. (2.2.21) in Eq. (2.2.19) and is given by

$$\tilde{\beta} = k_0\sqrt{\epsilon} = k_0\tilde{\mu} \tag{2.2.22}$$

where $\tilde{\mu}$ is the complex index of refraction. It can be written as

$$\tilde{\mu} = \mu + i(\alpha/2k_0) \tag{2.2.23}$$

where μ is the refractive index of the medium, α is the power-absorption coefficient, and usually $\alpha \ll \mu k_0$. Using $\epsilon = \tilde{\mu}^2$ and equating the real and imaginary parts, we find

$$\mu = (\epsilon')^{1/2} = [\epsilon_b + \mathrm{Re}(\chi_p)]^{1/2} \tag{2.2.24}$$

$$\alpha = \frac{k_0 \epsilon''}{\mu} = \frac{k_0}{\mu}[\mathrm{Im}(\chi_0 + \chi_p) + \sigma/(\epsilon_0 \omega)] \tag{2.2.25}$$

where Eq. (2.2.20) has been used. Equations (2.2.24) and (2.2.25) show explicitly how the refractive index and the net absorption coefficient are affected by external pumping of the semiconductor material.

2.3 THRESHOLD CONDITION AND LONGITUDINAL MODES

The plane-wave solution, Eq. (2.2.21), of the wave equation obtained in the previous section can be used to obtain an estimate of the laser frequency and the optical gain required for the onset of oscillations. It should be kept in mind that the lasing modes are never plane waves; Sec. 2.5 considers the spatial variations of the lasing modes. Nonetheless, the threshold condition derived here is reasonably accurate and is helpful in understanding the essential physics of the lasing process.

Consider a semiconductor laser of length L shown schematically in Fig. 2.1. The thin central region provides the optical gain. The cleaved facets form a Fabry-Perot (FP) cavity and provide the feedback for the onset of laser oscillations. If the z axis is taken along the cavity length, the optical field in the plane-wave approximation is given by Eq. (2.2.21), i.e.,

$$\mathbf{E} = \hat{\mathbf{x}} E_0 \exp(i\tilde{\beta} z) \tag{2.3.1}$$

where from Eqs. (2.2.22) and (2.2.23) the complex propagation constant takes the form

$$\tilde{\beta} = \mu k_0 + i\alpha/2. \tag{2.3.2}$$

Equation (2.2.24) for the refractive index shows that μ changes with external pumping of the semiconductor laser, as also observed experimentally.[10–16] The physical reason for this change is related to the high density of charge carriers inside the active region. Several mechanisms contribute to $\mathrm{Re}(\chi_p)$. The most dominant contribution is due to the band-gap shrinkage occurring in the presence of charge carriers. Generally, however,

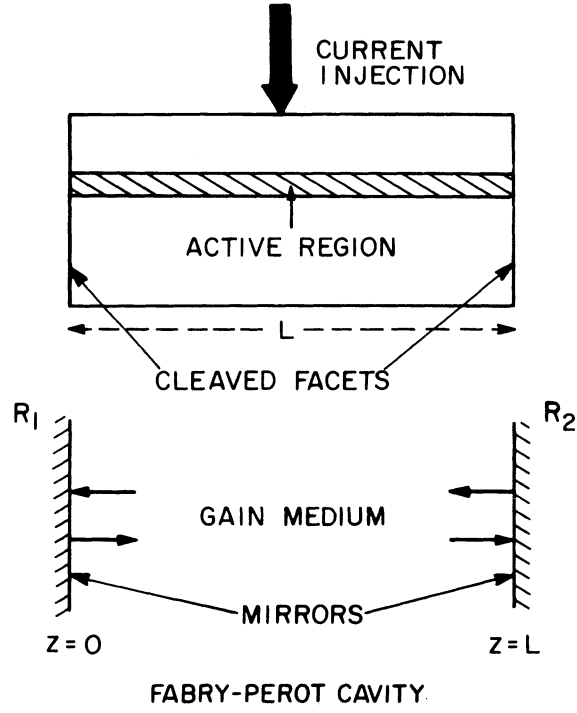


Fig. 2.1 Schematic illustration of a semiconductor laser and its associated Fabry-Perot (FP) cavity. The cleaved facets act as partially reflecting mirrors.

$|\mathrm{Re}(\chi_p)| \ll \epsilon_b$, and Eq. (2.2.24) can be approximated by

$$\mu \cong \mu_b + \Delta\mu_p \tag{2.3.3a}$$

where $\mu_b = \sqrt{\epsilon_b}$ is the background refractive index of the unpumped material and $\Delta\mu_p$, the amount by which it changes in the presence of charge carriers, is given by

$$\Delta\mu_p \cong \mathrm{Re}(\chi_p)/2\mu_b. \tag{2.3.3b}$$

Usually $\mathrm{Re}(\chi_p)$ is negative, making $\Delta\mu_p$ also negative. Even though the reduction in the refractive index is often less than 1%, it significantly affects the static, dynamic, and spectral characteristics of a semiconductor laser. This is in contrast to other kinds of lasers, such as a gas laser, where $\Delta\mu_p \cong 0$.

Next consider the absorption coefficient given by Eq. (2.2.25). It has three contributions arising from different sources. The term $\mathrm{Im}(\chi_0)$ accounts for the material absorption, while $\mathrm{Im}(\chi_p)$ is responsible for its reduction with the external pumping. It is often convenient to describe their combined effect as the net gain g, defined as

$$g = -\frac{k_0}{\mu_b}\,\mathrm{Im}(\chi_0 + \chi_p). \tag{2.3.4}$$

The last term in Eq. (2.2.25) accounts for other internal losses that generally occur in a semiconductor laser. Several mechanisms such as free-carrier absorption and scattering at the heterostructure interfaces may contribute to the internal loss.[5] Since the individual contribution of these and other internal losses is often difficult to estimate, they are collectively accounted for through $\alpha_{\mathrm{int}} = k_0\sigma/(\epsilon_0\omega\mu)$. The net absorption coefficient then becomes

$$\alpha = -\Gamma g + \alpha_{\mathrm{int}}. \tag{2.3.5}$$

The constant factor Γ is introduced here phenomenologically and its use will be justified in Sec. 2.5. Physically, it accounts for the reduction in gain that occurs because of the spreading of the optical mode beyond the active region. It is known as the *confinement factor* or the filling factor and represents the fraction of the mode energy contained in the active region.

To obtain the threshold condition, we require that the optical field Eq. (2.3.1) should reproduce itself after each round trip under steady-state or continuous-wave (CW) operating conditions. If R_1 and R_2 are the facet reflectivities at two ends, the net change in the amplitude after one round trip is set to unity at the laser threshold. This leads to the condition

$$(R_1R_2)^{1/2}\exp(2i\tilde{\beta}L) = 1. \tag{2.3.6}$$

Equating the real and imaginary parts of Eq. (2.3.6) and substituting $\tilde{\beta}$ from Eq. (2.3.2), we obtain

$$(R_1R_2)^{1/2}\exp(-\alpha L) = 1 \tag{2.3.7}$$

$$\sin(2\mu k_0 L) = 0. \tag{2.3.8}$$

The first condition (2.3.7) gives the threshold gain. Using Eq. (2.3.5), it can be written in the form

$$\Gamma g = \alpha_m + \alpha_{\mathrm{int}} \tag{2.3.9}$$

where

$$\alpha_m = \frac{1}{2L} \ln\left(\frac{1}{R_1 R_2}\right) \tag{2.3.10}$$

is the mirror loss and accounts for the radiation escaping from the FP cavity because of finite facet reflectivities. Equation (2.3.9) expresses the fact that the gain due to external pumping must balance the total losses. We remark that this is only approximately true since the effect of spontaneous emission has been ignored in this simple analysis. The contribution of spontaneous emission, invariably present in any laser, is to slightly lower the gain required to reach threshold from that given by Eq. (2.3.9). The effects of spontaneous emission are discussed in Chapter 6.

The condition (2.3.8) can be used to obtain the lasing frequency. However, because of the periodic nature of trigonometric functions, Eq. (2.3.8) has multiple solutions

$$2\mu k_0 L = 2m\pi \tag{2.3.11}$$

where m is an integer. Using $k_0 = 2\pi\nu/c$, the lasing frequency ν is given by

$$\nu = \nu_m = mc/(2\mu L) \tag{2.3.12}$$

where ν_m, often referred to as the cavity-resonance frequency, is the frequency of the mth longitudinal mode of an FP cavity of the optical length μL. Equation (2.3.12) shows that the laser tends to oscillate at a frequency that coincides with that of a longitudinal mode supported by the FP cavity. Which one and how many of them reach threshold depends on details of the gain spectrum such as the gain bandwidth and the gain-broadening mechanism[2] (whether homogeneous or inhomogeneous). In the case of homogeneous broadening only one longitudinal mode, whose frequency nearly coincides with the gain-peak frequency, reaches threshold, and the laser maintains the single-longitudinal-mode operation even in the above-threshold regime (see Fig. 2.2).

The longitudinal-mode spacing can be obtained from Eq. (2.3.12). However, it is important to remember that in a semiconductor laser the refractive index μ varies with the frequency ν. Using the relation

$$\Delta(\mu\nu) = \mu(\Delta\nu) + \nu(\Delta\mu) \tag{2.3.13}$$

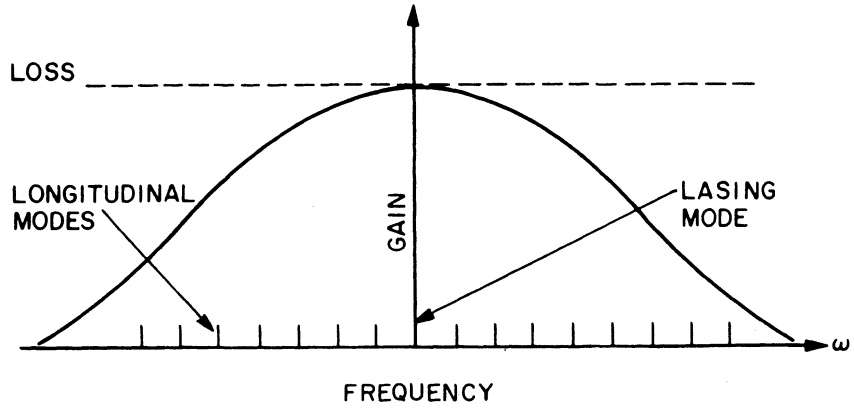


Fig. 2.2 Schematic illustration of the gain profile and longitudinal modes of a semiconductor laser. For the lasing mode in the vicinity of the gain peak, the threshold is reached when gain equals loss.

where Δ denotes a small change, the intermode spacing is given by

$$\Delta\nu = c/(2\mu_g L) \tag{2.3.14}$$

where

$$\mu_g = \mu + \nu(\partial\mu/\partial\nu) \tag{2.3.15}$$

is the group index of the dispersive semiconductor material. A feature unique to semiconductor lasers is that the longitudinal-mode frequencies and their separation vary with the external pumping because of the refractive-index variations as indicated by Eq. (2.3.3.a).

For the purpose of a rough estimate, consider an InGaAsP semiconductor laser in which cavity length $L = 250$ μm. For this material in the wavelength range of 1.3–1.6 μm, $\mu \cong 3.5$ and $\mu_g \cong 4$. The longitudinal mode spacing from Eq. (2.3.14) is estimated to be $\Delta\nu \cong 150$ GHz or $\Delta\lambda = \lambda^2\Delta\nu/c \cong 1$ nm. If we use the cleaved-facet reflectivity ($R_1 = R_2 \cong 0.32$) the facet loss α_m from Eq. (2.3.10) is $\sim 45\,\text{cm}^{-1}$. Using typical values for Γ and α_{int} of 0.5 and $\sim 30\,\text{cm}^{-1}$ respectively in Eq. (2.3.9), a material gain g of $150\,\text{cm}^{-1}$ is required to achieve threshold.

From the viewpoint of device operation, the quantity of practical interest is not the threshold gain itself but the threshold current density J_{th} required to achieve the gain. To obtain a relation between the gain g and the injected current density J, it becomes essential to consider the response of the semiconductor material to the optical field. This is considered in the next section.

2.4 GAIN AND STIMULATED EMISSION

In the semiclassical laser theory[1–3] the medium response is governed by the polarization $\mathscr{P}$ induced by the optical field $\mathscr{E}$ and leads to the susceptibility χ, as defined in Eq. (2.2.17). In terms of the density-matrix operator ρ, the induced polarization is given by

$$\mathscr{P} = \mathrm{Tr}(\rho \mathbf{p}) = \sum_{c,v} (\rho_{cv}\mathbf{p}_{vc} + \rho_{vc}\mathbf{p}_{cv}) \tag{2.4.1}$$

where $\mathbf{p}$ is the dipole-moment operator and the sum is over all the energy states per unit volume in the conduction and valence bands. The dynamic evolution of the density-matrix operator is governed by[2]

$$\frac{d\rho}{dt} = \frac{1}{i\hbar}[H_0 - \mathbf{p}\cdot\mathbf{E}, \rho] - \tfrac{1}{2}[\gamma\rho + \rho\gamma] + \Lambda \tag{2.4.2}$$

where H_0 is the unperturbed Hamiltonian of the semiconductor, γ is the decay operator, and Λ takes into account the carrier generation in the active region because of external pumping. The decay operator γ in general should include all decay mechanisms[17–19] through which electrons in a given energy state can decay out of that state. These can be divided into two categories corresponding to the intraband and interband decay mechanisms. *Intraband processes* in a semiconductor laser constitute electron-electron scattering and electron-phonon scattering[18–21] and occur at a fast time scale of $\sim$0.1 ps. By contrast, *interband processes* occur at a time scale of a few nanoseconds; they consist of radiative recombination leading to spontaneous and stimulated emissions as well as nonradiative recombination.[4–6] In long-wavelength semiconductor lasers an important source of nonradiative recombination is due to the Auger process,[22–24] as discussed in detail in Chapter 3.

A first-principle approach based on Eqs. (2.4.1) and (2.4.2) has proved useful for gas and solid-state lasers.[1,2] Such an approach has also been used for semiconductor lasers by considering both the many-body effects and the intraband scattering processes.[25,26] However, its usefulness for semiconductor lasers is limited because a realistic analysis is extremely complex. The reasons are severalfold. The inclusion of all the decay processes governed by γ in Eq. (2.4.2) is usually difficult since intraband scattering processes are not well understood. Further, the description of the unperturbed system through H_0 in Eq. (2.4.2) requires knowledge of the band structure and density of states in the conduction and the valence bands. A further complication arises in a semiconductor laser because the active region is often heavily doped and the band-tail effects become important.[4–6]

An alternative approach[27–30] has been used to calculate the optical gain. It is based on a generalization of Einstein's A and B coefficients,[31] which relate the rate of spontaneous and stimulated emissions to the net gain or absorption coefficient.[4–6] This method is discussed in detail in Chapter 3. A shortcoming of this approach is that it provides only the small-signal gain, and the gain-saturation effects cannot be treated. It is, however, useful for studying the dependence of the gain on various parameters such as the doping levels and the carrier density.

In view of the above-mentioned difficulties, a phenomenological approach is generally used for describing a semiconductor laser and has proven to be extremely successful. It is based on the observation that the numerically calculated gain[18,30] at the lasing frequency (corresponding to the value at which the gain spectrum peaks for a given current density J) varies almost linearly with the injected carrier density n for all values of J. The gain g can therefore be approximated by

$$g(n) = a(n - n_0) \tag{2.4.3}$$

where a is the gain coefficient and n_0 is the carrier density required to achieve transparency (corresponding to the onset of population inversion). Both of the parameters can be estimated from the numerical calculations or determined experimentally.[14,32,33] Note that the product an_0 is just the absorption coefficient of the unpumped material. To complete the phenomenological description, the refractive index is also assumed to vary linearly with the carrier density; i.e., $\Delta\mu_p$ in Eq. (2.3.3) is given by

$$\Delta\mu_p = bn \tag{2.4.4}$$

where $b = \partial\mu/\partial n$ and is often determined experimentally.[10–16] The parameters a, b, and n_0 are the three parameters of the phenomenological model. The assumption that both g and $\Delta\mu_p$ vary linearly with n appears to be quite drastic at first sight. However, the carrier density n changes little above threshold, and the linear variation is a reasonable approximation for small changes in n.

On comparing Eqs. (2.4.3) and (2.4.4) with Eqs. (2.3.3) and (2.3.4), we can see that the present model assumes a linear variation of the complex susceptibility χ_p with the carrier density n, i.e.,

$$\chi_p = \mu_b(2b - ia/k_0)n. \tag{2.4.5}$$

A parameter that is found to be quite useful is the ratio of the real to the

imaginary parts of χ_p and is given by

$$\beta_c = \frac{\mathrm{Re}(\chi_p)}{\mathrm{Im}(\chi_p)} = -\frac{2k_0 b}{a} = -2k_0\left(\frac{\partial\mu/\partial n}{\partial g/\partial n}\right). \tag{2.4.6}$$

Since $\beta_c \propto b$, it is often used in place of b in the phenomenological description of a semiconductor laser. Note, however, that since b is negative, β_c is a positive dimensionless number. Depending on the context, different names and symbols have been attributed to β_c. It is commonly referred to as the antiguiding parameter[34–36] or the line-width enhancement factor.[37–39]

The phenomenological description is complete once the carrier density n is related to the pump parameter, the current density J. This is accomplished through a rate equation that incorporates all the mechanisms by which the carriers are generated or lost inside the active region. In general, the continuity equations for both electrons and holes should be considered. The two are interrelated because of charge neutrality, and it suffices to consider one rate equation for electrons. In its general form, the carrier-density rate equation is

$$\frac{\partial n}{\partial t} = D(\nabla^2 n) + \frac{J}{qd} - R(n). \tag{2.4.7}$$

The first term accounts for carrier diffusion, and D is the diffusion coefficient. The second term governs the rate at which the carriers, electrons or holes, are injected into the active layer because of the external pumping. The electron and hole populations are assumed to be the same to maintain charge neutrality. In the second term, q is the magnitude of the electron charge and d is the active-layer thickness. Finally, the last term $R(n)$ takes into account the carrier loss owing to various recombination processes, both radiative and nonradiative. A rigorous derivation of Eq. (2.4.7) has to be based on the density-matrix approach. Yamada has carried out this procedure using Eq. (2.4.2) with some simplifying assumptions and has shown that carrier diffusion is a consequence of intraband scattering.[17,18] Carrier diffusion in general plays an important role in semiconductor lasers and complicates significantly their analysis. Depending on the device geometry, diffusion effects sometimes are of minor nature. This is the case, for example, for strongly index guided semiconductor lasers where the active-region dimensions (in the plane perpendicular to the cavity axis) are often small compared to the diffusion length. Since the carrier density in that case does not vary significantly over the active-region dimensions, it can be assumed to be approximately constant and the diffusion term in Eq. (2.4.7) can be neglected. In the steady state,

$\partial n/\partial t = 0$ and we obtain

$$J = qdR(n). \tag{2.4.8}$$

The situation is entirely different for gain-guided semiconductor lasers where the lateral variation (parallel to the heterojunction) of the carrier density arising from carrier diffusion makes the gain in Eq. (2.4.3) spatially inhomogeneous and helps to confine the optical mode.[34] The inclusion of the diffusion term is then a necessity.[35] This case is considered in Sec. 2.5.

The threshold condition (2.3.9) together with Eqs. (2.4.3) and (2.4.8) can be used to model the light-current characteristics of semiconductor lasers. To complete the description, we need a suitable form of the carrier-recombination rate $R(n)$ appearing in Eq. (2.4.8). The charge carriers recombine through several radiative and nonradiative mechanisms described in detail in Chapter 3. Radiative recombination may lead to either spontaneous or stimulated emission. A suitable form for $R(n)$ is[40,41]

$$R(n) = A_{nr}n + Bn^2 + Cn^3 + R_{st}N_{ph} \tag{2.4.9}$$

where it is assumed that the doping level of the active layer is well below the injected carrier density. The quadratic term Bn^2 is due to spontaneous radiative recombination wherein an electron in the conduction band recombines with a hole in the valence band and a photon is spontaneously emitted. The cubic term Cn^3 is due to Auger recombination, and its inclusion is particularly important for long-wavelength semiconductor lasers ($\lambda > 1$ μm). The nonradiative Auger process is described more fully in Chapter 3. The last term $R_{st}N_{ph}$ is due to stimulated recombination that leads to coherent emission of light. It is directly proportional to the intracavity photon density N_{ph} and to the net rate of stimulated emission[4–6]

$$R_{st} = (c/\mu_g)g(n) \tag{2.4.10}$$

where $g(n)$ is the optical gain given by Eq. (2.4.3).

The present phenomenological model assumes that the coefficients A_{nr}, B and C in Eq. (2.4.9) are constants independent of the external pumping. However, the bimolecular radiative coefficient B is known[42,43] to depend on the carrier density and is often approximated by[41]

$$B \cong B_0 - B_1 n. \tag{2.4.11}$$

This effect can be incorporated in Eq. (2.4.9) by using effective values of B and C. Similarly the effect of active-layer doping can be included by suitably

modifying the coefficients A_{nr}, B, and C.[41] The numerical values of these coefficients and the method of calculation are described in Chapter 3.

2.5 WAVEGUIDE MODES

The discussion in the previous sections is based on the plane-wave solutions of the wave equation (2.2.19). However, the light emitted by a laser has finite transverse dimensions, since it should be confined in the vicinity of the thin active region, which provides gain for stimulated emission. In semiconductor lasers the output is in the form of a narrow beam with an elliptic cross section. Depending on the laser structure, the field distribution across the beam can take certain well-defined forms, often referred to as the *laser modes.* Mathematically, a laser mode is the specific solution of the wave equation (2.2.19) that satisfies all the boundary conditions imposed by the laser structure. In the general multimode case the optical field is denoted by E_{pqm}, where the subscript m denotes the longitudinal or axial modes encountered in Sec. 2.3. The subscripts p and q take integer values; they stand for the *transverse* and *lateral* modes specifying the field distribution in the direction perpendicular and parallel to the junction plane, respectively. An understanding of the number of allowed modes and the resulting field distributions is essential for their control, since it is often desirable to design semiconductor lasers that emit light predominantly in a single mode. Further, important laser characteristics such as the near- and far-field widths depend on details of laser modes. For semiconductor lasers the far-field distribution plays an important role as it controls the amount of power coupled into a fiber.

In heterostructure semiconductor lasers the field confinement in the transverse direction, perpendicular to the junction plane, occurs through dielectric waveguiding.[4–6] This mechanism is often referred to as *index guiding* since the refractive-index discontinuity between the active and cladding layers is responsible for the mode confinement through the total internal reflection occurring at the interface.[8] However, field confinement in the lateral direction, which is parallel to the junction plane, is not always due to index guiding. Semiconductor lasers can be classified as gain-guided or index-guided depending on whether it is the lateral variation of the optical gain or the refractive index that confines the mode. Index-guided lasers can further be subclassified as weakly or strongly index guided depending on the magnitude of the lateral index step. Figure 2.3 shows schematically the three kinds of devices. These and other related structures are discussed in Chapter 5. Historically, gain-guided devices based on the stripe geometry were developed first in view of their ease of fabrication.[44–45] However, such devices have a number of undesirable characteristics that become worse as the laser wavelength

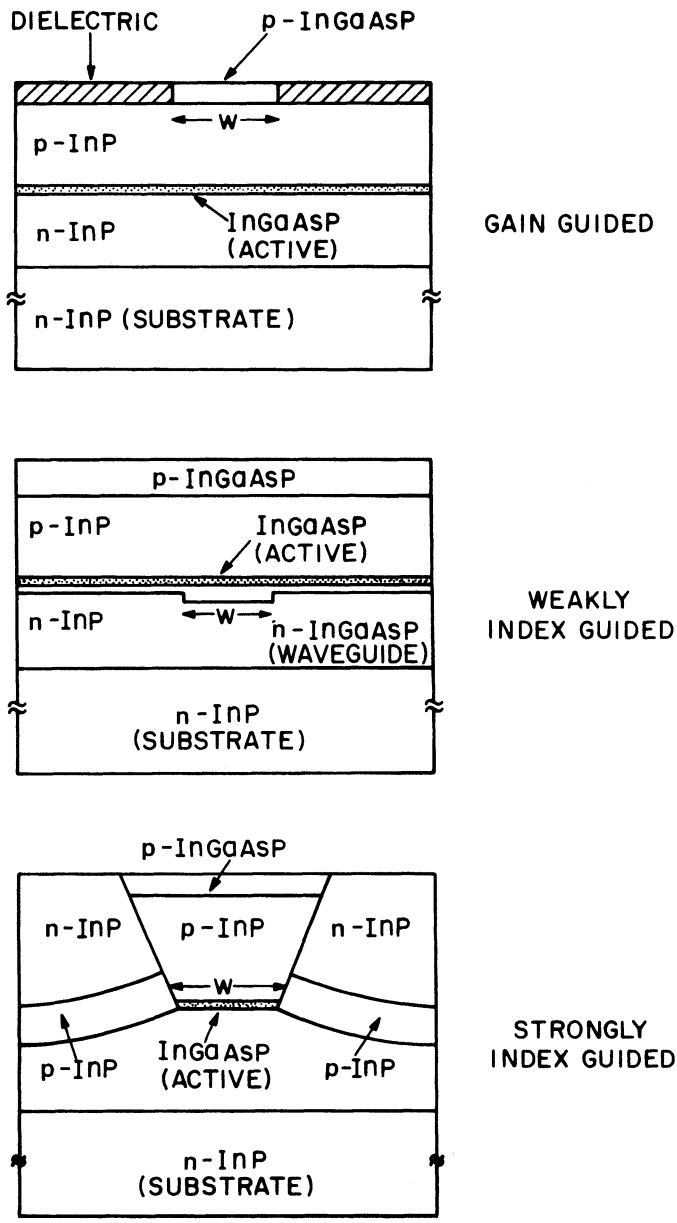


Fig. 2.3 Three kinds of InAsP laser structures shown schematically. The active-layer material (shown shaded) has a smaller band gap compared with that of the surrounding cladding layers.

increases. Index guiding is therefore almost invariably used in most practical semiconductor lasers.

2.5.1 Effective Index Approximation

The mathematical description of the laser modes is based on the time-independent wave equation [see Eq. (2.2.19)]

$$\nabla^2 \mathbf{E} + \epsilon(x, y) k_0^2 \mathbf{E} = 0 \tag{2.5.1}$$

where the x axis is parallel and the y axis is perpendicular to the heterojunction in Fig. 2.3. The dielectric constant ϵ may vary with x and y but is assumed to be independent of z, the direction of field propagation. For some semiconductor lasers (e.g., distributed feedback lasers), ϵ varies with z. As discussed in Chapter 7, such variations are small enough that they can be ignored in the discussion of waveguide modes. The spatially varying dielectric constant is generally of the form,

$$\epsilon(x, y) = \epsilon_j(x) \tag{2.5.2}$$

where the subscript j numbers various layers in a heterostructure laser. To account for the absorption, the dielectric constant ϵ_j is complex in each layer. Further, as shown by Eq. (2.2.20), within the active layer it also varies with external pumping.

To obtain an exact solution of Eq. (2.5.1) is a difficult task. It is essential to make certain simplifying assumptions whose nature and validity vary from device to device. For a strongly index guided laser such as the buried heterostructure laser shown schematically in Fig. 2.3, the effect of gain or loss on the passive-cavity modes can often be ignored. The resulting rectangular waveguide problem can then be solved[46,47] to obtain the transverse and lateral modes supported by the device. This approach is, however, not suitable for gain-guided lasers where the lateral modes arise precisely because of the active-region gain.

An alternative approach is based on the effective index approximation.[48–50] Instead of solving the two-dimensional wave equation, the problem is split into two one-dimensional parts whose solutions are relatively easy to obtain. Such an approach is partially successful for both gain-guided and index-guided lasers.[51–53] At the same time, it is helpful for a physical understanding of the guiding mechanism. The physical motivation behind the effective index approximation is that often the dielectric constant $\epsilon(x, y)$ varies slowly in the lateral x direction compared to its variation in the transverse y direction. To a good approximation, the slab-waveguide problem in the y direction

can be solved for each x and the resulting solution can then be used to account for the lateral variation. The electric field in Eq. (2.5.1) is thus approximated by

$$\mathbf{E} \cong \hat{\mathbf{e}}\phi(y; x)\psi(x) \exp(i\beta z) \tag{2.5.3}$$

where β is the propagation constant of the mode and $\hat{\mathbf{e}}$ is the unit vector in the direction along which the mode is polarized. On substituting Eq. (2.5.3) in Eq. (2.5.1) we obtain

$$\frac{1}{\psi}\frac{\partial^2\psi}{\partial x^2} + \frac{1}{\phi}\frac{\partial^2\phi}{\partial y^2} + [\epsilon(x, y)k_0^2 - \beta^2] = 0. \tag{2.5.4}$$

In the effective index approximation, the transverse field distribution $\phi(y; x)$ is obtained first by solving

$$\frac{\partial^2\phi}{\partial y^2} + [\epsilon(x, y)k_0^2 - \beta_{\text{eff}}^2(x)]\phi = 0 \tag{2.5.5}$$

where $\beta_{\text{eff}}(x)$ is the effective propagation constant for a fixed value of x. The lateral field distribution $\psi(x)$ is then obtained by solving

$$\frac{\partial^2\psi}{\partial x^2} + [\beta_{\text{eff}}^2(x) - \beta^2]\psi = 0. \tag{2.5.6}$$

For a given laser structure Eqs. (2.5.5) and (2.5.6) can be used to obtain the transverse and lateral modes, respectively. Since $\epsilon(x, y)$ is generally complex, $\beta_{\text{eff}}(x)$ is also complex. The effective index of refraction is defined as

$$\mu_{\text{eff}}(x) = \beta_{\text{eff}}(x)/k_0 \tag{2.5.7}$$

and is itself complex.

Equation (2.5.5) is a one-dimensional eigenvalue equation and can be solved using the methods developed for dielectric slab waveguides.[54–57] Although it is in general possible to include the gain or loss occurring in each layer, the resulting analysis is cumbersome. A simpler approach is to treat the effect of gain or loss as a small perturbation to the eigenvalue problem. This is justified for heterostructure semiconductor lasers since the mode confinement in the y direction occurs mainly because of the index step at the heterostructure interfaces. The dielectric constant $\epsilon(x, y)$ is of the form

$$\epsilon(x, y) = \mu_b^2(y) + \Delta\epsilon(x, y) \tag{2.5.8}$$

where μ_b is the background (real) refractive index, constant for each layer. The small perturbation $|\Delta\epsilon| \ll \mu_b^2$ includes the loss and the contribution of external pumping. If we use the first-order perturbation theory,[58] the eigenvalue given by μ_{eff} becomes

$$\mu_{\text{eff}}(x) = \mu_e(x) + \Delta\mu_e(x) \tag{2.5.9}$$

where $\mu_e(x)$ is obtained by solving the unperturbed eigenvalue equation

$$\frac{d^2\phi}{dy^2} + k_0^2[\mu_b^2(y) - \mu_e^2(x)]\phi = 0 \tag{2.5.10}$$

and the perturbation $\Delta\mu_e$ is obtained using

$$2\mu_e\,\Delta\mu_e = \frac{\int_{-\infty}^{\infty} \Delta\epsilon(x, y)\phi^2(y; z)\,dy}{\int_{-\infty}^{\infty} \phi^2(y; x)\,dy}. \tag{2.5.11}$$

Since $\Delta\epsilon(x, y)$ is constant within each layer, Eq. (2.5.11) can be simplified to become

$$\Delta\mu_e(x) = \frac{1}{2\mu_e}\sum_j \Gamma_j(x)\Delta\epsilon_j(x) \tag{2.5.12}$$

where the sum is over the number of layers, $\Delta\epsilon_j$ is the dielectric perturbation of jth layer, and

$$\Gamma_j(x) = \frac{\int_{d_j} \phi^2(y; x)\,dy}{\int_{-\infty}^{\infty} \phi^2(y; x)\,dy} \tag{2.5.13}$$

is the fraction of the mode intensity contained in that layer. For the active layer, Γ_j is referred to as the confinement or filling factor since it indicates the extent to which the mode is confined to the active region. Both Γ_j and the effective index μ_e vary with x if the active layer is not laterally uniform in thickness. This is the case, for example, for a channeled-substrate device[52,53] discussed in Chapter 5.

2.5.2 Transverse Modes

The transverse modes are obtained by solving Eq. (2.5.10) and depend on the thicknesses and refractive indices of the various layers used to fabricate a semiconductor laser. The number of layers to be considered depends on the specific laser structure, and it is often necessary to consider four or five layers for a reasonably accurate description of the transverse modes.[4,5] However, the basic concepts involved in dielectric waveguiding may be understood using a symmetric three-layer slab waveguide shown schematically in Fig. 2.4. The active layer of thickness d is surrounded on both sides by cladding layers. If the cladding layers are sufficiently thick such that the mode is largely confined within the three layers, the remaining layers can be ignored.

The slab-waveguide problem has been extensively studied.[54–56] Starting from Maxwell's equations (2.1)–(2.4), a slab waveguide is found to support two sets of modes, the TE and TM modes, which are distinguished on the basis of their polarization. For TE modes the electric field E is polarized along the heterojunction plane, i.e., the polarization vector in Eq. (2.5.3) is along the x axis. For TM modes, it is the magnetic field H that is polarized along the x axis. In both cases the same equation (2.5.10) can be used to obtain the field distributions. The boundary condition that the tangential components of the electric and magnetic fields be continuous at the dielectric interface requires that for TE modes ϕ and $d\phi/dy$ should match at $|y| = d/2$ (see Fig. 2.4). The other boundary condition for the guided modes is that the field $\phi(y)$ should vanish as y tends to infinity. It should be pointed

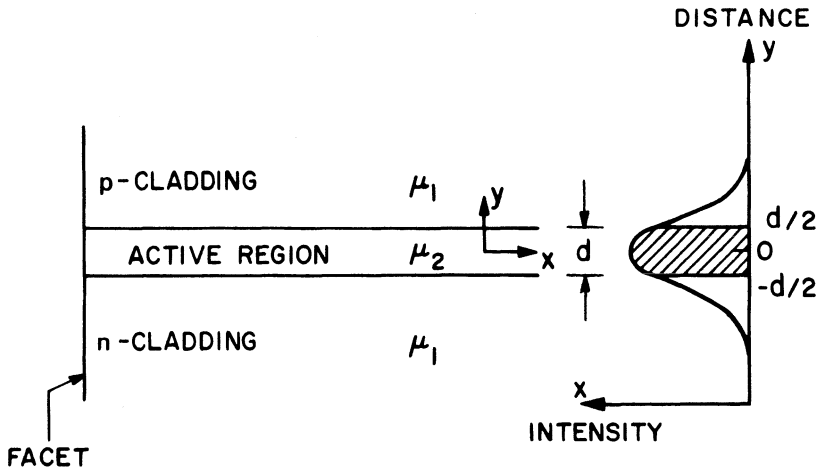


Fig. 2.4 Three-layer slab-waveguide model of a semiconductor laser with refractive indices such that $\mu_2 > \mu_1$. The intensity distribution of the fundamental waveguide mode is also shown. The hatched region represents the fraction of the mode within the active region.

out that a slab waveguide also supports unguided modes, the so-called radiation modes, for which the latter boundary condition does not apply.[8] For a semiconductor laser, these radiation modes are not of significant interest and would not be considered. However, their inclusion is crucial if an arbitrary optical field is expanded in terms of a complete set of waveguide modes.

The mode analysis is considerably simplified[8] if the even and odd solutions of the eigenvalue equation (2.5.10) are considered separately. Let us first consider the even TE modes. A general solution of Eq. (2.5.10) is of the form

$$\phi(y) = \begin{cases} A_e \cos(\kappa y) & \text{for } |y| \le d/2 \\ B_e \exp[-\gamma(|y| - d/2)] & \text{for } |y| \ge d/2 \end{cases} \tag{2.5.14}$$

where

$$\kappa = k_0(\mu_2^2 - \mu_e^2)^{1/2} \tag{2.5.15}$$

$$\gamma = k_0(\mu_e^2 - \mu_1^2)^{1/2} \tag{2.5.16}$$

and μ_2 and μ_1 are the material refractive indices for the active and cladding layers, respectively, with $\mu_2 > \mu_1$. The continuity of ϕ and $d\phi/dy$ at $|y| = d/2$ requires that

$$B_e = A_e \cos(\kappa d/2) \tag{2.5.17}$$

$$\gamma B_e = \kappa A_e \sin(\kappa d/2). \tag{2.5.18}$$

If we divide Eq. (2.5.18) by Eq. (2.5.17), we obtain the eigenvalue equation

$$\gamma = \kappa \tan(\kappa d/2) \tag{2.5.19}$$

whose solutions yield the effective mode index μ_e. In general, multiple solutions are possible corresponding to different even TE modes.

A similar analysis can be carried out for the odd TE modes with the only difference that the $\cos(\kappa y)$ in Eq. (2.5.14) is replaced by $\sin(\kappa y)$. The application of the boundary conditions now yields the eigenvalue equation

$$\gamma = -\kappa \cot(\kappa d/2) \tag{2.5.20}$$

whose solutions yield μ_e for the odd TE modes. For all guided modes, the inequality $\mu_2 > \mu_e > \mu_1$ is satisfied.

The TM modes are obtained using the same procedure. The only difference lies in the application of the boundary conditions. The continuity of the tangential component of the electric field, E_z, requires that $\mu_j^{-2}(d\phi/dy)$ be

continuous across the heterostructure interfaces at $|y| = d/2$. Here μ_j is either μ_1 or μ_2 depending on the side from which the interface is approached. The origin of this difference between the TE and TM modes can be traced back to Maxwell's equations (2.2.1) and (2.2.2). Similar to TE-mode analysis, the even and odd TM modes are considered separately. The resulting eigenvalue equations are

$$\mu_2^2\gamma = \mu_1^2\kappa \tan(\kappa d/2) \tag{2.5.21}$$

$$\mu_2^2\gamma = -\mu_1^2\kappa \cot(\kappa d/2) \tag{2.5.22}$$

for the even and odd TM modes respectively. In heterostructure semiconductor lasers, the TE modes are generally favored over the TM modes since the facet reflectivity is higher for TE modes, as discussed in Sec. 2.6. In the following discussion we therefore consider only TE modes.

The TE-mode eigenvalues are obtained using Eqs. (2.5.19) and (2.5.20) together with the relation

$$\kappa^2 + \gamma^2 = k_0^2(\mu_2^2 - \mu_1^2) \tag{2.5.23}$$

obtained by squaring and adding Eqs. (2.5.15) and (2.5.16). Equation (2.5.23) describes a circle in the κ–γ plane, and its intersection with the curves obtained using Eqs. (2.5.19) and (2.5.20) yields κ_p and γ_p values for the pth TE mode. Multiple solutions occur because of the periodic nature of trigonometric functions. The number of allowed waveguide modes can be determined by noting that a solution is no longer bounded if $\gamma \leq 0$ since it leads to an exponential growth of the field distribution ϕ in the cladding layers [see Eq. (2.5.14)]. The cut-off condition is thus determined by $\gamma = 0$ and occurs when

$$\kappa d = p\pi \tag{2.5.24}$$

where p is an integer whose even and odd values correspond to even and odd TE modes governed respectively by Eqs. (2.5.19) and (2.5.20). If we substitute κ from Eq. (2.5.23) after using $\gamma = 0$, we obtain the simple relation

$$D = p\pi \tag{2.5.25}$$

where

$$D = k_0(\mu_2^2 - \mu_1^2)^{1/2}d \tag{2.5.26}$$

is the normalized waveguide thickness. This parameter plays a central role in determining the mode characteristics of a three-layer slab waveguide. In particular, if $D < \pi$, the waveguide can support only the lowest-order

($p = 0$), fundamental TE mode. A heterostructure semiconductor laser can therefore be made to emit light in a single transverse mode if the active layer thickness d is chosen to satisfy the condition $D < \pi$ or from Eq. (2.5.26) after using $k_0 = 2\pi/\lambda$,

$$d < \frac{\lambda}{2}(\mu_2^2 - \mu_1^2)^{-1/2}. \tag{2.5.27}$$

Since typically $d \lesssim 0.2$ μm, the single-transverse-mode condition is almost always satisfied in practical devices. For an InGaAsP laser with InP cladding layers, it has been noted[59] that to a good approximation

$$\lambda(\mu_2^2 - \mu_1^2)^{-1/2} \cong 0.95 \text{ μm} \tag{2.5.28}$$

for the wavelength range of 1.1–1.65 μm. Condition (2.5.27) then simply becomes $d < 0.48$ μm. Similarly, using Eqs. (2.5.26) and (2.5.28), one finds that

$$D \cong 6.6d \tag{2.5.29}$$

where d is the active-layer thickness in micrometers. This relation is independent of λ, μ_1, and μ_2 in the entire wavelength region of 1.1–1.65 μm for InGaAsP lasers with InP cladding layers.

As mentioned before, a quantity that plays an important role for heterostructure semiconductor lasers is the transverse confinement factor Γ_T because it represents the fraction of the mode energy within the active layer that is available for interaction with the injected charge carriers. Using $\phi(y)$ from Eq. (2.5.14) in

$$\Gamma_T = \frac{\int_{-d/2}^{d/2} \phi^2(y)\, dy}{\int_{-\infty}^{\infty} \phi^2(y)\, dy} \tag{2.5.30}$$

and carrying out the integrations, we obtain

$$\Gamma_T = \frac{1 + 2\gamma d/D^2}{1 + 2/\gamma d}. \tag{2.5.31}$$

The evaluation of Γ_T requires the knowledge of κ and γ, and in general it is necessary to solve the eigenvalue equation (2.5.19) numerically. For the fundamental transverse mode, however, a remarkably simple expression

$$\Gamma_T \cong D^2/(2 + D^2) \tag{2.5.32}$$

was found to be accurate to within 1.5%,[59] and can be used to obtain Γ_T for any slab waveguide with the help of Eq. (2.5.26). For an InGaAsP laser with a 0.15-μm thick active layer, $D \cong 1$ from Eq. (2.5.29) and $\Gamma_T \cong \frac{1}{3}$. In Fig. 2.5 we have shown the variation of Γ_T with the active-layer thickness for a 1.3-μm InGaAsP laser. The transverse confinement factor for the fundamental TM mode is also shown for comparison. It is found to be lower, indicating that the TE mode is confined more than the TM mode. Finally, the effective index μ_e of the fundamental TE mode can be approximated by[59]

$$\mu_e^2 \cong \mu_1^2 + \Gamma_T(\mu_2^2 - \mu_1^2). \tag{2.5.33}$$

Both μ_e and Γ_T are required for our discussion of the lateral modes.

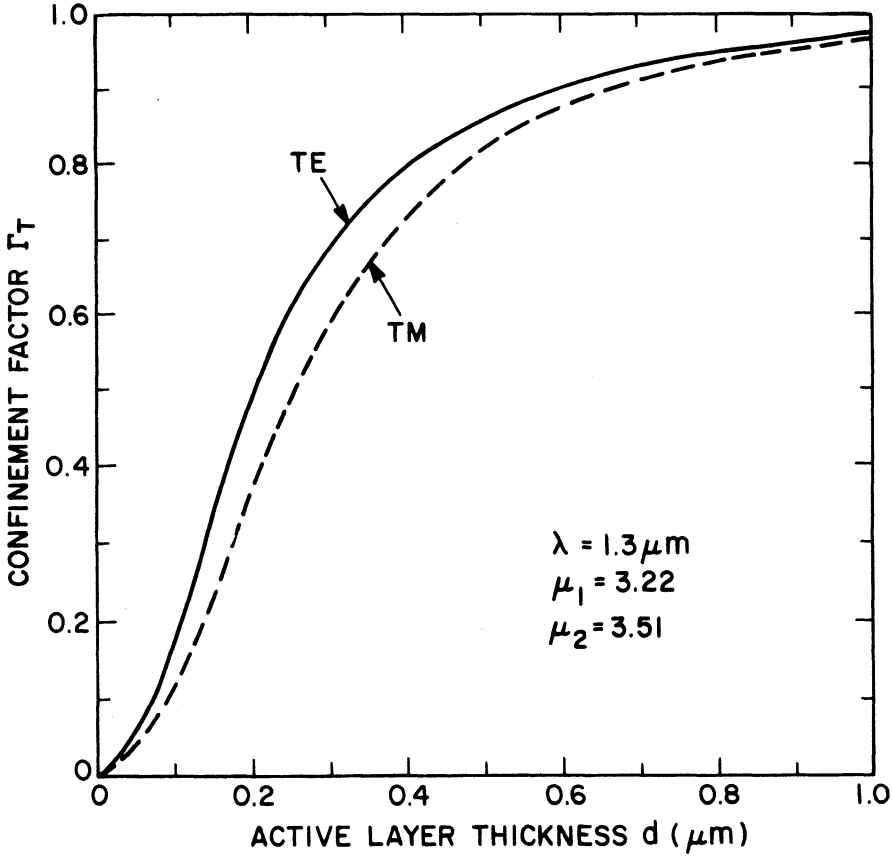


Fig. 2.5 Transverse confinement factor of the fundamental TE and TM modes as a function of the active-layer thickness for a 1.3-μm InGaAsP laser with InP cladding layers.

2.5.3 Lateral Modes

The lateral modes are obtained by solving Eq. (2.5.6) which after using Eqs. (2.5.7) and (2.5.9) becomes

$$\frac{\partial^2 \psi}{\partial x^2} + \{k_0^2[\mu_e(x) + \Delta\mu_e(x)]^2 - \beta^2\}\psi = 0. \tag{2.5.34}$$

The lateral-mode behavior in semiconductor lasers is different depending on whether gain guiding or index guiding is used to confine the lateral modes. In a gain-guided device, $\mu_e(x)$ is a constant given by Eq. (2.5.33). By contrast, in an index-guided device, structural lateral variations are used to make μ_e larger in a central region of width w. For the latter case the slab-waveguide problem discussed in Sec. 2.5.2 is solved separately in the two regions, and

$$\mu_e(x) = \begin{cases} \mu_e^{\text{in}} & \text{if } |x| \le w/2 \\ \mu_e^{\text{out}} & \text{otherwise} \end{cases} \tag{2.5.35}$$

where μ_e^{in} and μ_e^{out} are the effective indices corresponding to the two regions. Their magnitude depends on structural details, and the lateral index step

$$\Delta\mu_L = \mu_e^{\text{in}} - \mu_e^{\text{out}} \tag{2.5.36}$$

determines the extent of index guiding. Whether the lateral mode is index-guided or gain-guided depends on the relative magnitudes of $\Delta\mu_L$ and $\Delta\mu_e(x)$, and in general both should be considered.

Strongly Index Guided Semiconductor Lasers. Index-guided semiconductor lasers can often be classified as strongly index guided or weakly index guided depending on the structural modifications that give rise to $\Delta\mu_L$. A buried heterostructure device (see Fig. 2.3) falls in the category of strongly index guided devices since the index step $\Delta\mu_L \gg |\Delta\mu_e(x)|$. In that case the effect of gain can be treated as a small perturbation to the index-guided lateral mode and one can follow a perturbation procedure similar to that outlined in Eqs. (2.5.9)–(2.5.13). The mode-propagation constant is given by

$$\beta = k_0\bar{\mu} + i\bar{\alpha}/2 \tag{2.5.37}$$

where $\bar{\mu}$ and $\bar{\alpha}$ are the refractive index and the absorption coefficient of the mode supported by the rectangular waveguide of width w and thickness d. The lateral modes are obtained by solving the three-layer slab-waveguide

problem

$$\frac{\partial^2\psi}{\partial x^2} + k_0^2[\mu_e^2(x) - \bar{\mu}^2]\psi = 0. \tag{2.5.38}$$

The mode-absorption coefficient $\bar{\alpha}$ is obtained using the first-order perturbation theory[58] and is given by

$$\bar{\alpha} = \frac{k_0}{\bar{\mu}} \operatorname{Im}\left(\frac{\int 2\mu_e(x)\,\Delta\mu_e(x)\psi^2(x)\,dx}{\int \psi^2(x)\,dx}\right) \tag{2.5.39}$$

where $\Delta\mu_e(x)$ is given by Eq. (2.5.12). For the symmetric three-layer waveguide considered in Sec. 2.5.2, $\Gamma_2 = \Gamma_T$, and $\Gamma_1 = \Gamma_3 = (1 - \Gamma_T)/2$. If we use Eqs. (2.2.20)–(2.2.25) to obtain $\Delta\epsilon_j$ for the active and cladding layers, $\Delta\mu_e(x)$ becomes

$$\Delta\mu_e = \frac{1}{2\mu_e}[\Gamma_T\mu_2(2\Delta\mu_p - ig/k_0) + i(1 - \Gamma_T)\mu_1\alpha_c/k_0] \tag{2.5.40}$$

where α_c is the cladding-layer absorption coefficient, $\Delta\mu_p$ is the carrier-induced index change given by Eq. (2.3.3), and g is the active-layer gain given by Eq. (2.3.4). In general, both $\Delta\mu_p$ and g vary with x due to carrier diffusion. For buried heterostructure lasers (see Fig. 2.3), only the central region of width w is externally pumped. Furthermore, if w is small compared to the diffusion length, the carrier density is approximately uniform across the active width. In that case, g and $\Delta\mu_p$ can be assumed to be independent of x for $|x| \leq w/2$. Outside the active region, $g = \alpha_c$. Substituting Eq. (2.5.40) in Eq. (2.5.39) and carrying out the integrations, we obtain the simple expression

$$\bar{\alpha} \cong -\Gamma_L\Gamma_T g + (1 - \Gamma_L\Gamma_T)\alpha_c \tag{2.5.41}$$

where we used the approximation $\bar{\mu} \cong \mu_1 \cong \mu_2$ and

$$\Gamma_L = \frac{\int_{-w/2}^{w/2} |\psi(x)|^2\,dx}{\int_{-\infty}^{\infty} |\psi(x)|^2\,dx} \tag{2.5.42}$$

is the lateral confinement factor.

The lateral modes are obtained by solving the waveguide problem governed by Eqs. (2.5.35) and (2.5.38). Note that the injected carriers reduce μ_e^{in} by a small amount approximately given by $\Gamma_T\,\Delta\mu_p$ [see Eq. (2.5.40)]. For strongly

index guided lasers, this is a minor effect and can be neglected. It is clear from the form of Eq. (2.5.35) that a three-layer waveguide problem is to be solved. The properties of the lateral modes for a strongly index guided buried heterostructure laser can therefore be described using the results of Sec. 2.5.2. In analogy to Eqs. (2.5.25) and (2.5.26), one can define the normalized waveguide width W as

$$W = k_0 w[(\mu_e^{\text{in}})^2 - (\mu_e^{\text{out}})^2]^{1/2} \tag{2.5.43}$$

and the cut-off condition becomes

$$W = q\pi \tag{2.5.44}$$

where q is an integer whose even and odd values correspond to even and odd lateral modes, respectively. In particular, only the lowest-order ($q = 0$) lateral mode is supported by the waveguide if $W \leq \pi$ or equivalently, the active-layer width

$$w \leq \lambda/(8\mu_e \, \Delta\mu_L)^{1/2} \tag{2.5.45}$$

where $\Delta\mu_L$ is the lateral index step given by Eq. (2.5.35) and μ_e is the average effective index. In analogy to Eqs. (2.5.32) and (2.5.33), the lateral confinement factor Γ_L and the mode refractive index $\bar{\mu}$ for the fundamental lateral mode are given by

$$\Gamma_L \cong W^2/(2 + W^2) \tag{2.5.46}$$

and

$$\bar{\mu}^2 \cong (\mu_2^{\text{out}})^2 + \Gamma_L[(\mu_e^{\text{in}})^2 - (\mu_e^{\text{out}})^2]. \tag{2.5.47}$$

The description of waveguiding in strongly index-guided lasers is now complete. If the active-layer dimensions are chosen to satisfy Eqs. (2.5.27) and (2.5.45), only a single transverse and lateral mode is supported by the passive waveguide. Its complex propagation constant is given by Eq. (2.5.37), where the mode index $\bar{\mu}$ is obtained using Eq. (2.5.47). The mode loss is governed by $\bar{\alpha}$ given by Eq. (2.5.41), which can be rewritten as

$$\bar{\alpha} = -\Gamma g + \alpha_{\text{int}} \tag{2.5.48}$$

where $\Gamma = \Gamma_L \Gamma_T$ and is the mode confinement factor and α_{int} is the internal loss due to absorption in the cladding layers. Equation (2.5.48) should be compared with Eq. (2.3.5), its counterpart obtained in the plane-wave

approximation. Whereas the confinement factor Γ in Eq. (2.3.5) was introduced phenomenologically, in the waveguide description,

$$\Gamma = \Gamma_L \Gamma_T \tag{2.5.49}$$

where Γ_T and Γ_L are the transverse and lateral confinement factors defined by Eqs. (2.5.30) and (2.5.42), respectively. Clearly Γ represents the fraction of the mode energy contained within the active region. For typically used values $w \cong 2\ \mu\text{m}$, $\Gamma_L \cong 1$ and Γ_T can be used for Γ. The internal loss α_{int} in Eq. (2.5.48) should be generalized to include all sources of loss. Its suitable form is

$$\alpha_{\text{int}} = \Gamma\alpha_a + (1 - \Gamma)\alpha_c + \alpha_{\text{scat}} \tag{2.5.50}$$

where α_a is the active-layer loss (mainly due to free-carrier absorption) and α_{scat} is the scattering loss at heterostructure interfaces.

Gain-guided Semiconductor Lasers. For gain-guided devices the effective index μ_e in Eq. (2.5.34) is constant along the lateral direction x, and the mode confinement occurs through $\Delta\mu_e(x)$ given by Eq. (2.5.12). In contrast to index-guided lasers where the index changes discontinuously [see Eq. (2.5.35)], $\Delta\mu_e(x)$ varies continuously. For the three-layer waveguide model, $\Delta\mu_e(x)$ is obtained using Eq. (2.5.40). Equation (2.5.34) then becomes

$$\frac{\partial^2\psi}{\partial x^2} + k_0^2[\mu_e^2 + \Gamma\mu_2(2\Delta\mu_p - ig/k_0) + i\mu_1\alpha_{\text{int}}/k_0]\psi = \beta^2\psi \tag{2.5.51}$$

where α_{int} is given by Eq. (2.5.50) and $|\Delta\mu_e| \ll \mu_e$ is assumed. Both g and $\Delta\mu_p$ vary with the carrier density n as given by Eqs. (2.4.3) and (2.4.4). The most distinctive feature of gain-guided lasers is that carrier diffusion plays an important role and n, obtained by solving Eq. (2.4.7), is laterally nonuniform. It is precisely the variation in the optical gain $g(x)$ that leads to the confinement of the lateral modes. For a known carrier-density profile $n(x)$, the eigenvalue equation (2.5.51) can be solved to obtain the complex propagation constant β and the mode profile $\psi(x)$ corresponding to various lateral modes.

The problem of lateral-mode determination for a gain-guided semiconductor laser is, however, exceedingly complex, and in general a numerical approach is necessary. This is so because stimulated emission couples the carrier-diffusion equation (2.4.7) and the wave equation (2.5.51), and the two should be solved self-consistently for each value of the device current. A further complication is that the injected current density $J(x)$ is itself laterally

nonuniform because of current spreading in the cladding layer below the contact stripe (see Fig. 2.3). Several numerical models[60–72] have been developed for gain-guided lasers and are reasonably successful in predicting the operating characteristics of a gain-guided semiconductor laser. Most of these models have been applied to AlGaAs lasers. They can be generalized for InGaAsP and other long-wavelength semiconductor lasers with some modifications such as the inclusion of Auger recombination.[72,73] A combination of increased Auger recombination and increased index antiguiding [governed by the parameter β_c defined by Eq. (2.4.6)] makes gain-guided semiconductor lasers relatively unattractive at wavelengths greater than 1 μm.

To obtain some physical insight in the process of gain guiding, it is useful to consider a specific carrier-density profile. One excellent example is provided by the parabolic profile

$$n(x) = n(0) - n_2 x^2 \tag{2.5.52}$$

which is often a good approximation near the stripe center at $x = 0$. If we substitute Eq. (2.5.52) in Eqs. (2.4.3) and (2.4.4) and then use g and $\Delta\mu_p$ in Eq. (2.5.51), we obtain

$$\frac{\partial^2\psi}{\partial x^2} + k_0^2[\epsilon(0) - \bar{a}^2 x^2]\psi = \beta^2\psi \tag{2.5.53}$$

where

$$\epsilon(0) = \epsilon_{0r} + i\epsilon_{0i} \tag{2.5.54}$$

is the complex dielectric constant at the stripe center and

$$\bar{a} = a_r + ia_i \tag{2.5.55}$$

governs its quadratic decrease with x. The parameters ϵ_{0r}, ϵ_{0i}, a_r, and a_i are determined by the material constants and phenomenological parameters introduced in Sec. 2.4.

The eigenvalue equation (2.5.53) can be readily solved to obtain the eigenfunctions $\psi(x)$ and the eigenvalues β corresponding to various lateral modes labeled by the subscript q. The lateral modes are the well-known Hermite-Gaussian functions given by[74]

$$\psi_q(x) = H_q[(k_0\bar{a})^{1/2}x] \exp(-\tfrac{1}{2}k_0\bar{a}x^2) \tag{2.5.56}$$

where $H_q(z)$ is the qth-order Hermite polynomial for its complex argument z. The corresponding eigenvalues are

$$\beta_q^2 \cong k_0^2\epsilon(0) - (2q + 1)\bar{a}k_0. \tag{2.5.57}$$

The real and imaginary parts of β_q can be used to obtain the mode index $\bar{\mu}$ and the absorption coefficient $\bar{\alpha}$ using Eq. (2.3.39).

The Hermite-Gaussian modes in a homostructure semiconductor laser were observed[44] as early as 1967. However, their origin was unclear at that time and a quadratic variation of the refractive index was used to explain the experimental results.[45] It was only in 1973 that the mechanism of gain guiding[57,74] was identified as being responsible for the lateral confinement of the optical mode in stripe-geometry heterostructure semiconductor lasers. A characteristic feature of gain guiding is that the phase front is curved,[34] in contrast to index guiding that leads to a planar phase front [see Eq. (2.5.14)]. This can be seen in Eq. (2.5.56) by noting that $\bar{a}$ is complex. For the lowest-order lateral mode ($q = 0$), the phase varies quadratically. Its curvature can be accounted by defining a radius of curvature R_c that is given by[34]

$$R_c = \bar{\mu}/a_i \tag{2.5.58}$$

where $\bar{\mu}$ is the mode index of the fundamental lateral mode. Physically, the origin of a curved wavefront in gain-guided devices can be traced back to the carrier-induced index change $\Delta\mu_p$ given by Eq. (2.4.4). The antiguiding parameter β_c defined by Eq. (2.4.6) therefore plays an important role in gain-guided devices, and it is easy to verify that $a_i = 0$ if $\beta_c = 0$.

The number of lateral modes excited in a gain-guided device depends on the stripe width w, which is a rough equivalent of the active-layer width w used in the discussion of index-guided devices. Although the quadratic model used above does not predict a mode cut-off condition such as given by Eq. (2.5.45), experimentally it is observed that only the fundamental lateral mode is excited when $w \lesssim 10\ \mu\text{m}$. For narrow-stripe lasers, the parabolic model is generally not suitable and a numerical solution of Eq. (2.5.51) is often necessary.[60–73] Some analytical insight can be gained using an alternative model that assumes that the complex dielectric constant varies as $\text{sech}^2(x)$.[75] A detailed discussion of this model can be found in Ref. 6.

Weakly Index Guided Semiconductor Lasers. Gain-guided semiconductor lasers have an inherent drawback. The lateral-gain profile responsible for the mode confinement in general changes with external pumping, and the lateral-mode control is difficult to achieve in the above-threshold regime. Strongly index guided lasers such as the buried heterostructure do not suffer from this problem but are relatively difficult to fabricate. During the

development of GaAs lasers it was realized that some index guiding can be induced in gain-guided lasers if the cladding-layer thickness is designed to be laterally nonuniform (see Fig. 2.3). A number of structures, including the rib waveguide, ridge waveguide, and channeled-substrate planar waveguide devices, have been proposed.[76–78] We shall collectively call them weakly index guided devices. From the viewpoint of the lateral-mode control, the important point is that the effective index μ_e in Eq. (2.5.34) is slightly larger over a narrow central region of width w. It should be stressed that w is not the active-layer width (see Fig. 2.3). Further, carrier diffusion inside the active layer is responsible for the continuous lateral variation of the gain and the refractive index.

The mathematical description of lateral modes is based on Eq. (2.5.34). The lateral variation of μ_e is represented by Eq. (2.5.35). The index step $\Delta\mu_L$ given by Eq. (2.5.35) governs the extent of index guiding. However, in contrast to strongly index guided lasers where $\Delta\mu_L \gg |\Delta\mu_e|$, now the two terms are comparable in magnitude and should be considered simultaneously. The resulting lateral-mode behavior is complex since gain guiding, carrier-induced antiguiding, and built-in index guiding all participate in the formation of lateral modes. A numerical solution of Eq. (2.5.34) is generally necessary[72,79] to analyze a weakly index guided device. Depending on the magnitude of $\Delta\mu_L$, such a device exhibits lateral-mode features reminiscent of pure gain guiding or pure index guiding. As $\Delta\mu_L$ progressively increases, a transition from gain guiding to index guiding occurs.[72] Clearly the device is of practical use only when it is designed to operate in the index-guided regime. Numerical calculations for a 1.3-μm InGaAsP laser show[72,79] that an index step $\Delta\mu_L$ in the range of 0.005–0.01 is enough to achieve a stable lateral mode.

2.6 EMISSION CHARACTERISTICS

Previous sections have discussed such basic concepts as gain-loss considerations and characteristics of the longitudinal, transverse, and lateral modes supported by the semiconductor laser cavity. We now consider the emission characteristics used to characterize the performance of a semiconductor laser. They can be classified into several groups: (i) light-current, (ii) spatial-mode, (iii) spectral, and (iv) transient or dynamic characteristics. In the following discussion we consider each group separately.

2.6.1 Light-Current Characteristics

The light emitted by one facet of a semiconductor laser is measured as a function of the device current I. The resulting curve is often referred to as

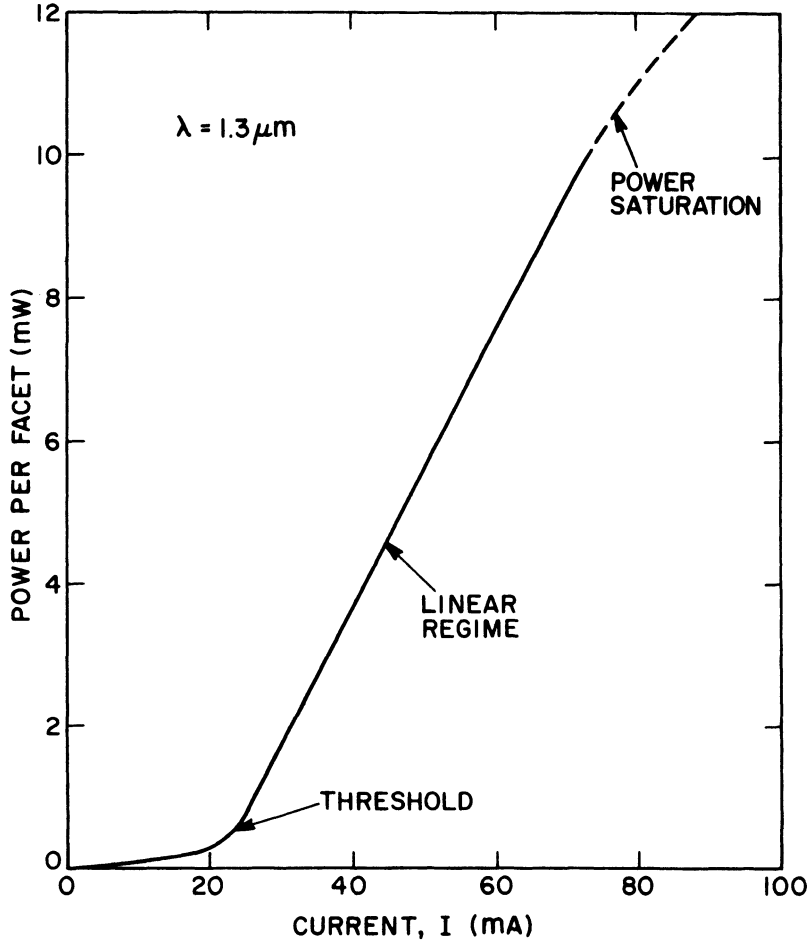


Fig. 2.6 Light-current (L-I) curve shown schematically for a strongly index guided InGaAsP laser. The dashed line represents the onset of power saturation corresponding to a sublinear increase of power with the current.

the light-current (L-I) curve and is strongly temperature dependent. Fig. 2.6 shows the behavior schematically for a strongly index guided laser at a given temperature. The current I is related to the current density J that is injected into the active layer by

$$I = I_a + I_L = wLJ + I_L \tag{2.6.1}$$

where I_a is the current passing through the active region, L is the cavity

length and w is the active-region width. The leakage current I_L accounts for the fact that, depending on a specific structure, a part of the total current may not pass through the active layer. Equation (2.6.1) is applicable for a broad-area laser or a strongly index guided laser for which the injected current density is nearly uniform over the active region width. If the current density is laterally nonuniform, such as in a gain-guided device, J in Eq. (2.6.1) should be interpreted as the average injected current density over the width w.

The form of the L-I curve shown in Fig. 2.6 is typical for any laser. The turning point, at which the light output abruptly starts increasing, corresponds to the laser threshold. The threshold current I_{th} (or equivalently the threshold current density J_{th}) is an important device parameter, and its minimization is often sought. When $I < I_{th}$, the light output mainly consists of spontaneous emission and its magnitude is governed by the term Bn^2 in Eq. (2.4.9). When $I > I_{th}$, stimulated emission, which is governed by the last term in Eq. (2.4.9), starts to dominate. In the vicinity of the threshold, the power increases by several orders of magnitude. This phase-transition-like behavior is discussed in Chapter 6 after including the contribution of spontaneous emission to the lasing mode.

An expression for the threshold current density of a strongly index guided semiconductor laser can be obtained using the phenomenological model of Sec. 2.4. As J increases, the gain also increases owing to an increase in the carrier density n [see Eqs. (2.4.3) and (2.4.8)]. The threshold is reached when n achieves the threshold value n_{th} for which the gain equals loss. Using Eqs. (2.3.9) and (2.4.3), the threshold carrier density is given by

$$n_{th} = n_0 + (\alpha_m + \alpha_{int})/(a\Gamma) \tag{2.6.2}$$

where n_0 is the transparency value at which population inversion occurs. The threshold current density J_{th} is obtained using Eqs. (2.4.8) and (2.4.9). Near threshold, stimulated recombination can be neglected and we obtain

$$J_{th} = qdn_{th}/\tau_e(n_{th}) \tag{2.6.3}$$

where

$$\tau_e(n) = (A_{nr} + Bn + Cn^2)^{-1} \tag{2.6.4}$$

is the carrier-recombination time that in general depends on n. Once threshold is reached, the carrier density n remains pinned to its threshold value n_{th} and a further increase in J leads to light emission through the process of stimulated emission. Equations (2.4.8) and (2.4.9) can be used to obtain the

intracavity photon density given by

$$N_{\text{ph}} = \eta_i(\tau_p/qd)(J - J_{\text{th}}) \tag{2.6.5}$$

where the *internal quantum efficiency* η_i is introduced phenomenologically, the photon lifetime τ_p is defined as

$$\tau_p^{-1} = v_g(\alpha_m + \alpha_{\text{int}}) \tag{2.6.6}$$

and $v_g = c/\mu_g$ is the group velocity. Equation (2.6.5) shows that once the threshold is reached, the intracavity photon density increases linearly with a further increase in J. The carrier-recombination time τ_e and the photon lifetime τ_p play an important role in determining the transient response of a semiconductor laser, as discussed in Chapter 6.

To obtain the L-I curve, Eq. (2.6.5) should be expressed in terms of the output power P_{out} and the device current I. Since photons escape out of the laser cavity at a rate of $v_g\alpha_m$, the power emitted by each facet is related to the intracavity photon density by the relation

$$P_{\text{out}} = \tfrac{1}{2}hvv_g\alpha_m VN_{\text{ph}} \tag{2.6.7}$$

where $V = Lwd$ is the active volume and the factor of $\frac{1}{2}$ is due to the assumption of equal fact reflectivities so that the same power is emitted by the two facets. The power emitted from each facet becomes different when the reflectivities are changed by coating the laser facets. Using Eqs. (2.6.1) and (2.6.7) in Eq. (2.6.5), we obtain

$$P_{\text{out}} = \frac{hv}{2q}\frac{\eta_i\alpha_m}{\alpha_m + \alpha_{\text{int}}}(I - I_{\text{th}} - \Delta I_L) \tag{2.6.8}$$

where Eq. (2.6.6) was used to eliminate τ_p. The quantity ΔI_L represents a possible increase in the leakage current with the current I. The threshold current I_{th} is obtained using Eqs. (2.6.1) and (2.6.3) and is given by

$$I_{\text{th}} = qVn_{\text{th}}/\tau_e + I_L. \tag{2.6.9}$$

Equation (2.6.8) expresses the laser output as a function of the current I, the external pumping parameter. It predicts that the output power should vary linearly with I. The slope dP_{out}/dI is a measure of the device efficiency, and a typical value for an InGaAsP laser emitting near 1.3 μm is 0.25 mW/mA per facet. In practice the slope efficiency does not remain constant, and the output power saturates for large values of I (see Fig. 2.6). Equation (2.6.8)

predicts that three factors may contribute to power saturation. (i) The leakage current may increase with I ($\Delta I_L \neq 0$) so that a smaller fraction of the device current leads to carrier injection into the active layer. (ii) The value of I_{th} is current dependent and increases with I. A possible mechanism is junction heating that may reduce the carrier-recombination time τ_e as the laser power increases. The decrease in τ_e can be due to Auger recombination,[24–26] which increases significantly with temperature. (iii) The internal loss α_{int} increases with I so that a smaller fraction of generated photons is useful as the output power. The relative importance of the three possible mechanisms is discussed in Chapter 3.

Before the power-saturation mechanisms set in, the slope dP_{out}/dI is reasonably constant and may be used to obtain the differential quantum efficiency η_d of the laser device. The *differential (external) quantum efficiency* is defined as

$$\eta_d = \eta_i \frac{\text{photon escape rate}}{\text{photon generation rate}} \tag{2.6.10}$$

where η_i is the internal quantum efficiency indicating what fraction of injected carriers is converted into photons. To maintain the steady-state, the photons are generated at a rate of τ_p^{-1} and escape out of the laser cavity at a rate of $v_g \alpha_m$. Using Eqs. (2.6.6) and (2.6.10), we obtain

$$\eta_d = \eta_i (v_g \alpha_m \tau_p) = \eta_i \frac{\alpha_m}{\alpha_m + \alpha_{int}}. \tag{2.6.11}$$

If we use Eq. (2.6.8), we find that η_d is directly proportional to the slope of the L-I curve, i.e.,

$$\eta_d = \frac{2q}{h\nu} \frac{dP_{out}}{dI}. \tag{2.6.12}$$

The internal quantum efficiency represents the fraction of injected carriers that recombine radiatively and generate photons. Using Eq. (2.4.9), it is given by

$$\eta_i = \frac{Bn^2 + R_{st} N_{ph}}{A_{nr} n + Bn^2 + Cn^3 + R_{st} N_{ph}} \tag{2.6.13}$$

and is the ratio of the radiative to the total recombination rates. In the above-threshold regime the stimulated-recombination term $R_{st} N_{ph}$ dominates

and $\eta_i \cong 1$. Using α_m from Eq. (2.3.10) in Eq. (2.6.11), η_d becomes

$$\eta_d = \eta_i \left[1 + \frac{\alpha_{\text{int}} L}{\ln(1/R_m)} \right]^{-1} \tag{2.6.14}$$

where $R_m = R_1 = R_2$ and is the mirror reflectivity assumed to be equal for both facets. Equation (2.6.14) was first obtained by Biard et al.[80] and predicts that η_d increases with a reduction in the cavity length L. The inverse dependence of η_d on L has been verified[81] and can be used to estimate η_i. For a typical strongly index guided 1.3-μm semiconductor laser with length L of 250 μm, the slope efficiency $dP_{\text{out}}/dI \cong 0.25$ mW/mA; from Eq. (2.6.12), $\eta_d \cong 0.5$ if we assume that $\eta_i \cong 1$. From Eq. (2.6.14) this value corresponds to an internal loss α_{int} of about 40 cm^{-1}.

The emitted power and the slope efficiency become different for the two facets when their reflectivities are made different by coating the laser facets. For instance, it is common to reduce the reflectivity of one facet through an antireflection coating to increase the output power from the facet. The other facet is often coated with a high-reflection coating in order to maintain the laser threshold. The calculation of output powers emitted from each facet is slightly more involved since the photon density or the intracavity power is far from being uniform along the cavity length when the facet reflectivities are made different. The average intracavity power is obtained by using

$$P_{\text{in}} = (\alpha_m + \alpha_{\text{int}}) \int_0^L |A \exp(i\tilde{\beta} z + B \exp(-i\tilde{\beta} z)|^2 \, dz \tag{2.6.15}$$

where A and B are the amplitudes of the forward and backward propagating waves normalized such that $|A|^2$ and $|B|^2$ represent the corresponding optical power at $z = 0$. The complex propagation constant $\tilde{\beta}$ is obtained from Eqs. (2.3.2) and (2.3.4) and is given by

$$\tilde{\beta} = \mu k_0 - \frac{i}{2}(\Gamma g - \alpha_{\text{int}}). \tag{2.6.16}$$

The boundary conditions at the laser facets relate A and B such that

$$A = \sqrt{R_1}\, B, \qquad B = \sqrt{R_2}\, A \exp(2i\tilde{\beta} L) \tag{2.6.17}$$

where R_1 and R_2 are the reflectivities of the facets located at $z = 0$ and $z = L$ respectively. The threshold condition (2.3.6) is readily obtained from Eq.

(2.6.17). In particular, the gain g in Eq. (2.6.16) is given by Eq. (2.3.9) or by $\Gamma g = \alpha_m + \alpha_{\text{int}}$, where the mirror loss α_m is defined as in Eq. (2.3.10).

The integration in Eq. (2.6.15) is easily performed. The interference term can be neglected in most cases of practical interest because of its rapidly oscillating nature. By using Eqs. (2.6.15) and (2.6.16) together with $\alpha_m = \Gamma g - \alpha_{\text{int}}$, the result is

$$P_{\text{in}} = \frac{\alpha_m + \alpha_{\text{int}}}{\alpha_m} \frac{(\sqrt{R_1} + \sqrt{R_2})(1 - \sqrt{R_1 R_2})}{R_1 \sqrt{R_2}} |A|^2. \tag{2.6.18}$$

The average intracavity power P_{in} is related to the photon density N_{ph} as

$$P_{\text{in}} = h\nu V N_{\text{ph}} / \tau_p \tag{2.6.19}$$

where $h\nu V N_{\text{ph}}$ represents the total energy. Equations (2.6.18) and (2.6.19) can be used to obtain $|A|^2$ in terms of N_{ph}. The powers emitted from the facets located at $z = 0$ and $z = L$ are related to $|B|^2$ and $|A|^2$ as

$$P_1 = (1 - R_1)|B|^2, \qquad P_2 = (1 - R_2)|A \exp(i\tilde{\beta} L)|^2. \tag{2.6.20}$$

By using Eqs. (2.6.18)–(2.6.20) the output powers are related to N_{ph} as

$$P_1 = \frac{(1 - R_1)\sqrt{R_2}}{(\sqrt{R_1} + \sqrt{R_2})(1 - \sqrt{R_1 R_2})} h\nu v_g \alpha_m V N_{\text{ph}} \tag{2.6.21a}$$

$$P_2 = \frac{(1 - R_2)\sqrt{R_1}}{(\sqrt{R_1} + \sqrt{R_2})(1 - \sqrt{R_1 R_2})} h\nu v_g \alpha_m V N_{\text{ph}}. \tag{2.6.21b}$$

These equations should be compared with Eq. (2.6.7) to which they reduce when the facet reflectivities are equal ($R_1 = R_2$) since $P_1 = P_2 = P_{\text{out}}$ in that case. Note that $P_1 + P_2 = 2P_{\text{out}}$ for all values of R_1 and R_2, as it should since a change in facet reflectivities does not affect the total power but only its partition between the two facets. The reflectivity-dependent factor in Eqs. (2.6.21) shows how the total power is partitioned; it reduces to 1/2 when $R_1 = R_2$.

The threshold current I_{th} of a gain-guided laser is considerably higher than that of a strongly index guided device. Further, I_{th} depends strongly on the stripe width w over which the current is injected (see Fig. 2.3). The variation of I_{th} with w is shown in Fig. 2.7 for 1.3-μm and 1.55-μm gain-guided InGaAsP lasers.[73] The curves were obtained using a numerical

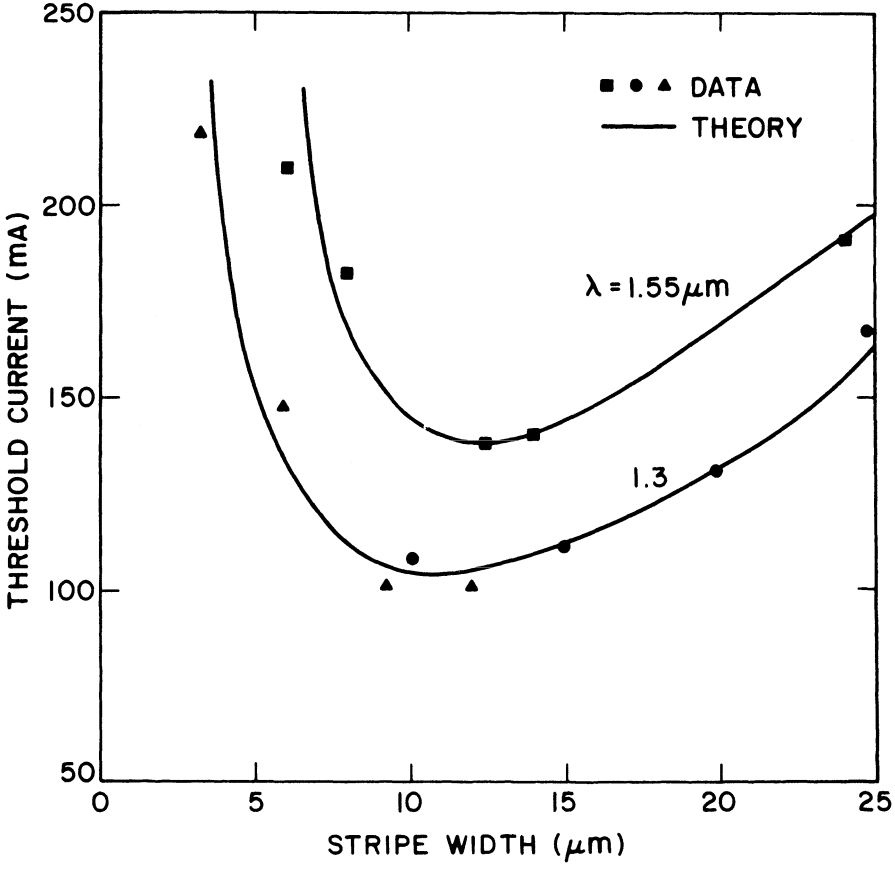


Fig. 2.7 Variation of the threshold current with the stripe width for 1.3-μm and 1.55-μm gain-guided InGaAsP lasers. (After Ref. 73)

model[70] that includes carrier diffusion, current spreading, and various radiative and nonradiative recombination mechanisms. The inclusion of Auger recombination was essential to obtain a good fit between the theory and the experimental data.[73] The minimum value of the threshold current at 1.3 μm is more than 100 mA and occurs for $w \cong 10$ μm. The rapid increase in I_{th} for narrower stripes is due to the combined effect of index antiguiding and Auger recombination. The differential quantum efficiency η_d is also lower compared to an index-guided device because of additional diffraction losses.

As discussed in Sec. 2.5, a weakly index guided device (see Fig. 2.3) overcomes these shortcomings through a central region of a relatively higher index. Compared to a strongly index guided device such as a buried

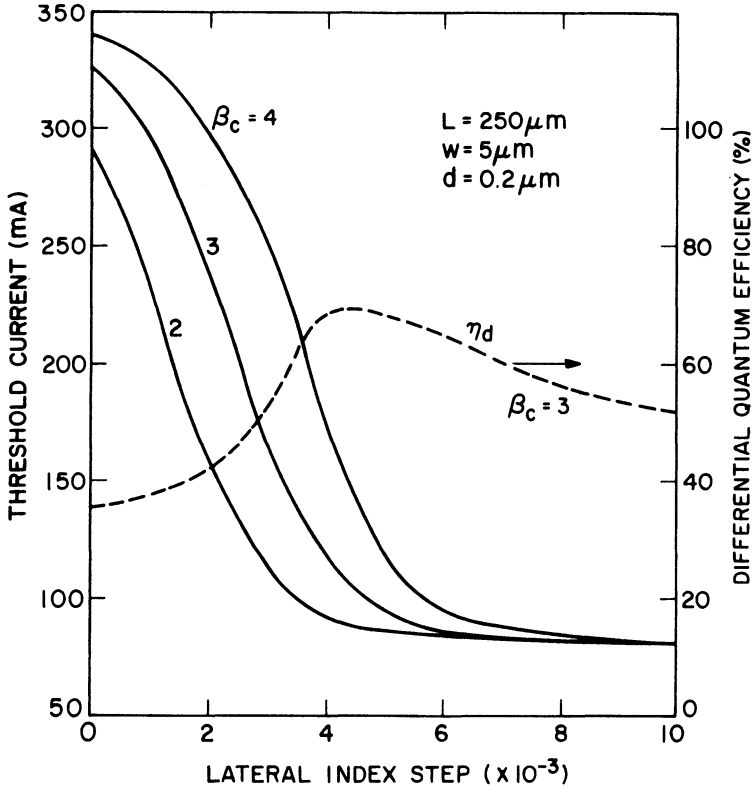


Fig. 2.8 Calculated variation of the threshold current with the lateral index step $\Delta\mu_L$ for a 1.3-μm weakly index guided laser for several values of the antiguiding parameter β_c. Dashed line shows the corresponding variation of the differential quantum efficiency for the specific case where $\beta_c = 3$. (After Ref. 72 © 1984 IEEE)

heterostructure laser, the lateral index step $\Delta\mu_L$ is more than an order of magnitude smaller. Figure 2.8 shows the calculated variation of the threshold current I_{th} and the differential quantum efficiency η_d with the index step $\Delta\mu_L$ using parameters appropriate for a 1.3-μm InGaAsP laser.[72] In the absence of index guiding ($\Delta\mu_L = 0$), the device is gain-guided with $I_{th} \cong 300$ mA and $\eta_d \cong 0.35$. As $\Delta\mu_L$ increases, I_{th} rapidly decreases and η_d increases. For $\Delta\mu_L \geq 5 \times 10^{-3}$, the transition from gain guiding to index guiding is complete. The slight decrease in η_d with a further increase in $\Delta\mu_L$ is due to a mode profile that becomes narrower than the gain profile. This profile mismatch reduces the internal quantum efficiency since a part of the gain is not used for stimulated emission.[72] The important point to note is that an index step as small as $\Delta\mu_L = 5 \times 10^{-3}$ is enough for index guiding. The

limiting threshold current I_{th} is higher for a weakly index guided device than that of a strongly index guided device because of a larger mode volume, a weaker mode confinement, and increased carrier diffusion.

The threshold current I_{th} and the differential quantum efficiency η_d are two important quantities invariably used to characterize the performance of a semiconductor laser. As Eqs. (2.6.2) and (2.6.11) show the mirror loss α_m associated with the cleaved-facet reflectivity R_m plays a significant role in their determination. Further, the selection of TE or TM modes in a heterostructure is also based on the mode reflectivity R_m, which is generally different for the two polarizations. A considerable amount of theoretical work[82–88] has been done to calculate the dependence of R_m on the waveguide parameters such as the active-layer thickness d and the index difference $\mu_2 - \mu_1$. In general, a numerical approach[83] is necessary. The main feature is that the effect of dielectric waveguiding is to increase R_m for the TE modes and to decrease R_m for the TM modes relative to their plane-wave value (~ 0.32), which is the same for both sets of modes. Together with the confinement factor, the reflectivity difference helps to select the TE mode of operation in heterostructure semiconductor lasers.

2.6.2 Spatial-Mode Characteristics

A semiconductor laser emits light in the form of a narrow spot of elliptical cross section. The spatial-intensity distribution of the emitted light near the laser facet is known as the *near field*. During its propagation the spot size grows in size due to beam divergence. The dimensions of the elliptical spot and its divergence angles, both parallel and perpendicular to the junction plane, are important beam parameters associated with the laser mode. The angular intensity distribution far from the laser facet is known as the *far field*.

In Sec. 2.5 we solved the wave equation and associated the spatial distribution of the optical field with the transverse mode $\phi(y)$ and the lateral mode $\psi(x)$. In general, several transverse and lateral modes may be excited, and the resulting near field is formed by a superposition of them. However, the active-layer dimensions are often chosen such that only the lowest-order transverse and lateral modes are supported by the waveguide. The near field just outside the laser facet can then be written as

$$E = E_0 \psi(x) \phi(y) \exp(ik_0 z) \tag{2.6.22}$$

where E_0 is a constant related to the output power ($P_{out} \propto |E_0|^2$). $\psi(x)$ and $\phi(y)$ are the lateral and transverse mode profiles discussed in Sec. 2.5. Experimentally, $|\psi(x)|^2$ and $|\phi(y)|^2$ can be measured by scanning a photodetector across the beam dimensions. Their full widths at half maximum

(FWHM), $w_{\parallel}$ and $w_{\perp}$ respectively, are taken to be a measure of the dimensions of the elliptical spot.

The far field is obtained by taking a two-dimensional Fourier transform of the near field given by Eq. (2.6.22). It can be expressed as a product of the two one-dimensional Fourier transforms corresponding to the lateral and transverse directions x and y. The lateral far-field intensity distribution is given by

$$|\tilde{\psi}_{FF}(\theta)|^2 = \cos^2\theta \left| \int_{-\infty}^{\infty} \psi(x) \exp(ik_0 x \sin\theta)\, dx \right|^2 \tag{2.6.23}$$

with a similar expression for the transverse far field. Again the quantity of practical interest is the width (FWHM) of the far-field distribution. The angles $\theta_{\parallel}$ and $\theta_{\perp}$, corresponding to the lateral and transverse directions respectively, are used as a measure of the angular spread of the emitted light.

Near Field. The near field $\phi(y)$ perpendicular to the junction plane depends on the thickness and composition of the various layers used to make a heterostructure semiconductor laser. For a three-layer waveguide model, $\phi(y)$ is given by Eq. (2.5.14) and is shown schematically in Fig. 2.4. The width $w_{\perp}$ depends on the active-layer thickness d and increases with a reduction in d because of a loss in the strength of mode confinement. An analytic expression for $w_{\perp}$ is difficult to obtain. However, by fitting $\phi(y)$ with a Gaussian distribution, the width can be approximated by[89]

$$w_{\perp} \cong d(2 \ln 2)^{1/2}(0.321 + 2.1D^{-3/2} + 4D^{-6}) \tag{2.6.24}$$

where the normalized thickness D is given by Eq. (2.5.26). The expression is reasonably accurate for $1.8 < D < 6$.

The near field $\psi(x)$ parallel to the junction plane depends on the lateral guiding mechanism and displays qualitatively different features for gain-guided and index-guided lasers. For a strongly index guided laser such as a buried-heterostructure device, the near-field behavior is similar to that of $\phi(y)$, and Eq. (2.6.24) can be applied after replacing D by W, where W is the normalized width given by Eq. (2.5.43). Figure 2.9 shows typical near-field intensity profiles $|\psi(x)|^2$ obtained for a buried-heterostructure laser with an active layer 0.15 μm thick and 1.8 μm wide.[90] Curves for several drive currents are shown, and $I_{th} = 52$ mA for this device. The width (FWHM) $w_{\parallel}$ is about 2 μm, indicating that the near field is largely confined within the active layer. By contrast, the lateral near field for gain-guided lasers extends considerably beyond the strip width. For weakly index guided lasers, the lateral confinement is remarkably improved for an index step $\Delta\mu_L$ as

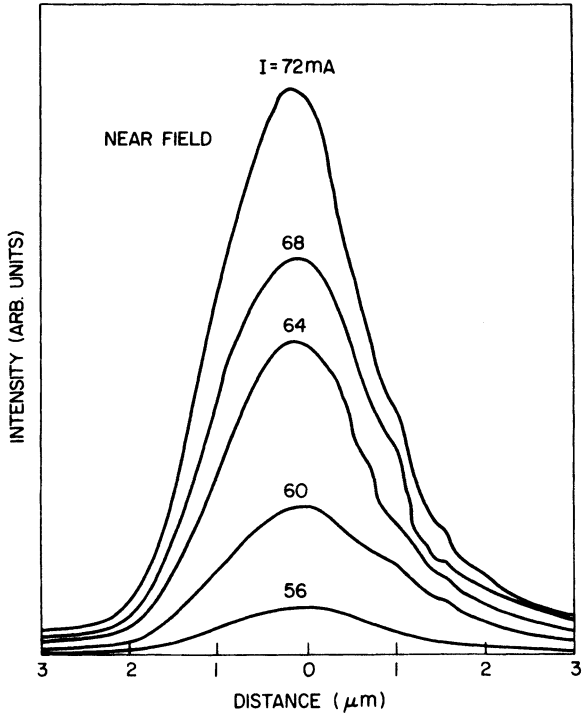


Fig. 2.9 Near-field scans along the junction plane of a 1.3-μm buried-heterostructure laser at various current levels. (After Ref. 90 © 1981 IEEE)

small as 5×10^{-3}, since the device then operates in the index-guided regime (see Fig. 2.8).

Far Field. The far-field patterns in the directions parallel and perpendicular to the junction plane indicate the angular spread of the laser mode and are important in determining the coupling efficiency between the semiconductor laser and a fiber. Figure 2.10 shows the far-field scans[90] in the two directions for the same BH device of Fig. 2.9. The angular widths $\theta_{\parallel}$ and $\theta_{\perp}$ at 64 mA are 33° and 57°, respectively, and represent typical values for strongly index guided semiconductor lasers. Mathematically the far-field pattern is obtained by taking the Fourier transform of the near field as indicated by Eq. (2.6.23).

As one may expect, the lateral far field (parallel to the junction plane) depends strongly on the lateral guiding mechanism and exhibits qualitatively different behavior for gain-guided and for index-guided lasers. In particular, for narrow-stripe gain-guided lasers the far field has two widely separated

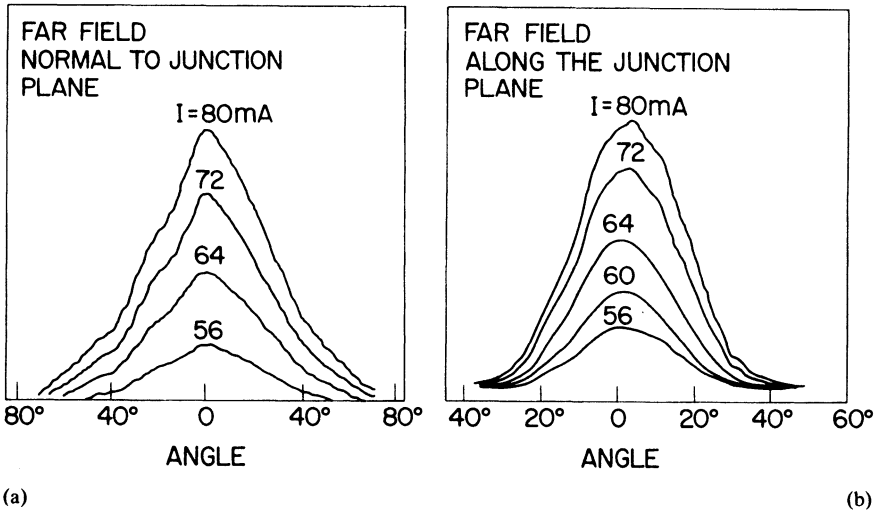


Fig. 2.10 Far-field scans (a) perpendicular and (b) parallel to the junction plane of a 1.3-μm buried-heterostructure laser at various current levels. (After Ref. 90 © 1981 IEEE)

peaks with a minimum at the center. Such twin-lobe far fields are known[63,65] to arise from the carrier-induced index reduction, which is governed by the antiguiding parameter β_c [see Eqs. (2.4.4) and (2.4.6)], that leads to a curved wavefront for the lateral mode. The introduction of weak index guiding has a dramatic effect on the far field, as shown in Fig. 2.11 for a 1.3-μm InGaAsP laser.[72] Only one of the two peaks for each curve is shown. As $\Delta\mu_L$ increases, the two peaks move closer and eventually merge to yield a single-lobe far-field pattern. The transition is almost complete for $\Delta\mu_L = 4 \times 10^{-3}$, at which point the device is then index-guided (see also Fig. 2.8).

2.6.3 Spectral Characteristics

The power spectrum of a semiconductor laser is an important device characteristic since in many applications the spectral control of the laser output is required. Below threshold the output takes the form of spontaneous emission with a large spectral width of ~ 30 nm. As the threshold is approached, the spectrum narrows considerably, and several peaks whose frequencies coincide with the longitudinal mode frequencies (see Sec. 2.3) appear. In the above-threshold regime, the longitudinal mode closest to the gain peak increases in power, while the power in remaining side peaks saturates. This behavior is shown in Fig. 2.12 for a 1.3-μm buried-heterostructure InGaAsP laser.

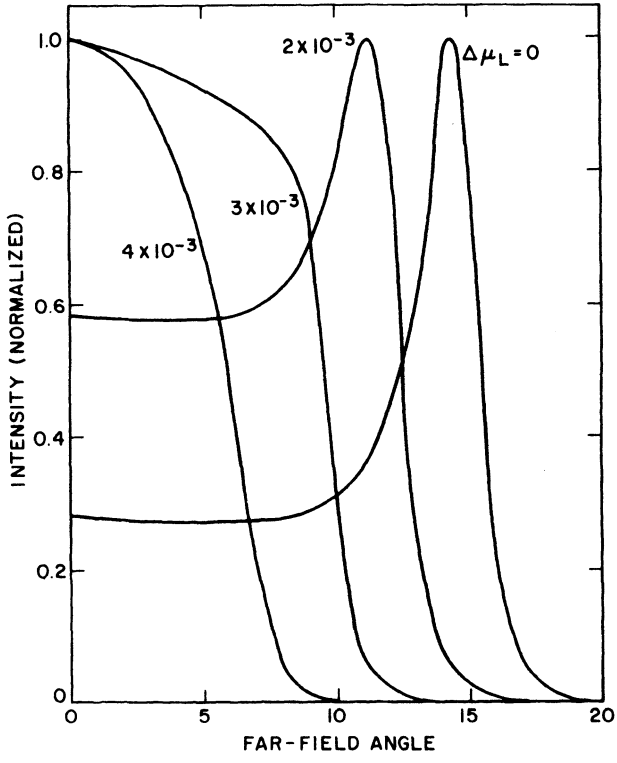


Fig. 2.11 Calculated far fields of 1.3-μm weakly index guided InGaAsP lasers for different values of the lateral index step $\Delta\mu_L$. Due to its symmetric nature, only half of the profile is shown. The device is purely gain-guided when $\Delta\mu_L = 0$. (After Ref. 72 © 1984 IEEE)

A somewhat surprising feature of semiconductor lasers is that even though the gain profile is largely homogeneously broadened, many longitudinal modes oscillate simultaneously and a significant amount of laser power is carried out by side modes even in the above-threshold regime. Several mechanisms such as spatial-hole burning, spectral-hole burning, and a high rate of spontaneous emission into the lasing mode may contribute to the observed spectral features. Chapter 6 gives a theoretical discussion based on the multimode rate equations.

For the application of a semiconductor laser in high-bit-rate, long-haul, optical fiber communication systems, the longitudinal side modes are undesirable because they cause pulse spreading in the presence of fiber dispersion, thereby degrading system performance. An important parameter used to describe the spectral purity of a semiconductor laser is the *mode suppression ratio* (MSR). It is defined as the ratio of the main-mode power

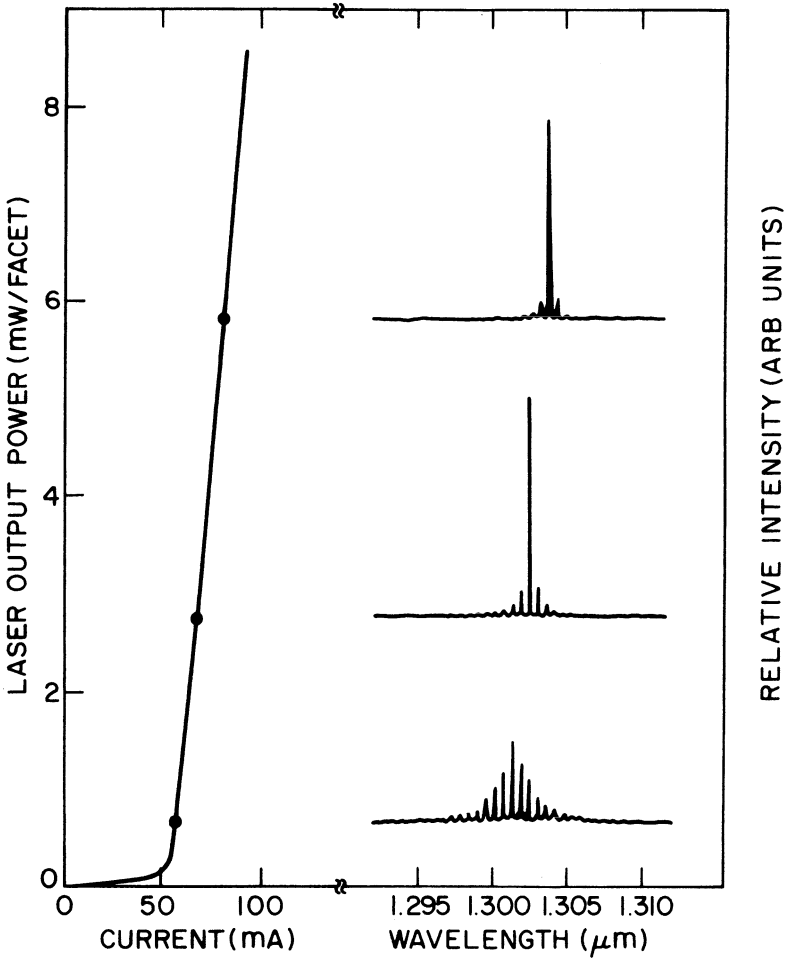


Fig. 2.12 Optical power spectra at three different power output levels as indicated in the L-I curve of a 1.3 μm buried-heterostructure laser. (After Ref. 90 © 1981 IEEE)

to the power carried by the most intense side mode. At the 1.55-μm wavelength, MSR values exceeding 30 dB are generally desirable. However, mode discrimination in semiconductor lasers is usually poor, and such values of MSR are difficult to realize. In recent years distributed-feedback and coupled-cavity mechanisms have been used to provide additional side-mode discrimination and to increase the MSR. Such lasers are collectively referred to as single-longitudinal-mode lasers[91] since their power spectrum consists of a dominant single longitudinal mode. Chapters 7 and 8 examine

distributed-feedback and coupled-cavity semiconductor lasers in view of their increasing importance in optical communication systems.

Another quantity of practical interest is the spectral width of a single longitudinal mode. Even when a semiconductor laser operates predominantly in a single longitudinal mode, quantum fluctuations associated with the process of spontaneous emission lead to a broadening of the laser line. Spectral line width is typically in the range of 10–100 MHz and decreases inversely with the laser power. The line width is calculated using single-mode rate equations generalized to take into account various noise mechanisms. Chapter 6 gives a detailed discussion of the noise phenomena and the line width.

2.6.4 Dynamic Characteristics

An understanding of the transient or dynamic response of a semiconductor laser is crucial since in some applications, particularly in optical fiber communications, the device current is modulated periodically and the laser output takes the pulsed form. When the laser is turned on by increasing the device current from 0 to its final above-threshold value, a few nanoseconds (a time interval known as the turn-on delay) elapse before the laser power starts increasing rapidly. Further, because of the nonlinear nature of interaction between photons and charge carriers, the power does not increase monotonically. The laser output oscillates periodically for several nanoseconds before attaining its steady-state CW value. The oscillations are referred to as relaxation oscillations, and their frequency is governed by the nonlinear dynamics of photon-carrier interaction. The relaxation-oscillation frequency is typically in the gigahertz range and increases with an increase in the device current or the output power.

When the device current is modulated at frequencies approaching a few gigahertz, transient effects play an increasingly important role. The power distribution among various longitudinal modes changes with time. Under high-frequency direct modulation, the side modes are considerably enhanced compared to those in the CW case. Further, the line width of an individual longitudinal mode increases considerably. The line broadening is related to the carrier-induced index change [governed by the parameter β_c defined in Eq. (2.4.6)] that leads to frequency chirping. Chapter 6 considers the dynamic and modulation characteristics of semiconductor lasers in detail.

PROBLEMS

2.1 Prove that the spacing between the longitudinal modes of a semiconductor laser of length L is given by $\Delta\nu = c/(2\mu_g L)$ where μ_g is the group index.

2.2 The refractive index of the active layer of a semiconductor laser varies with the carrier density. What changes do you expect to occur as a result of such variations in the refractive index?

2.3 Explain the concept of the linewidth enhancement factor. Why is its introduction necessary for semiconductor lasers?

2.4 Derive the eigenvalue equation for the TM modes of a planar waveguide of thickness d and refractive index μ_2 sandwiched between two cladding layers of refractive index μ_1. Assume infinitely thick cladding layers.

2.5 The active region of a 1.55-μm InGaAsP laser is 0.1-μm thick, 1-μm wide, and 250-μm long. Calculate the confinement factor and the modal index of the fundamental waveguide mode supported by the laser. Assume $\mu_2 = 3.5$ and $\mu_1 = 3.2$.

2.6 Calculate the active-region gain required for the laser of problem 2.5 to reach threshold. Assume an internal loss of $30\ \text{cm}^{-1}$ for the laser cavity.

2.7 Calculate the maximum allowed thickness of the active layer for a 1.3-μm semiconductor laser designed to support a single transverse model. How does this value change if the laser operates at 1.55 μm. Assume $\mu_2 = 3.5$ and $\mu_1 = 3.2$.

2.8 Calculate the maximum allowed width of the active region for a 1.55-μm semiconductor laser with 0.1-μm-thick active layer if the laser is designed to operate in a single lateral mode. Assume $\mu_2 = 3.5$ and $\mu_1 = 3.2$.

2.9 Estimate the spot size (FWHM) in the lateral and transverse directions for a 1.3-μm semiconductor laser whose active region is 0.2-μm thick and 1-μm wide. Assume $\mu_2 = 3.5$ and $\mu_1 = 3.2$.

2.10 Calculate the angular width (FWHM) of the far field in the lateral and transverse directions for the laser of problem 2.9.

REFERENCES

1. Siegman, A. E. *Lasers.* Mill Valley, CA: University Science Books, 1986.
2. Sargent III, M., M. O. Scully, and W. E. Lamb, Jr. *Laser Physics.* Reading, Mass.: Addison-Wesley Publishing Co., 1974.
3. Milonni, P. W., and J. H. Eberly. *Lasers.* New York: John Wiley & Sons, 1988.
4. Kressel, H., and J. K. Butler. *Semiconductor Lasers and Heterojunction LEDs.* New York: Academic Press, 1977.
5. Casey, H. C., Jr., and M. B. Panish. *Heterostructure Lasers.* New York: Academic Press, 1978.
6. Thompson, G. H. B. *Physics of Semiconductor Laser Devices.* Chichester: John Wiley & Sons, 1980.
7. Kogelnik, H., and T. Li. *Proc. IEEE* **54,** 1312 (1966).
8. Marcuse, D. *Light Transmission Optics.* New York: Van Nostrand Reinhold Co., 1982.

9. Magid, L. M. *Electromagnetic Fields, Energy, and Waves*, Chap. 9. New York: John Wiley & Sons, 1972.
10. Henry, C. H., R. A. Logan, and K. A. Bertness, *J. Appl. Phys.* **52,** 4457 (1981).
11. Olsson, A., and C. L. Tang. *Appl. Phys. Lett.* **39,** 24 (1981).
12. Manning, J. S., and R. Olshansky. *Electron. Lett.* **17,** 506 (1981).
13. Stubkjaer, K., M. Asada, S. Arai, and Y. Suematsu. *J. Appl. Phys.* **20,** 1499 (1981).
14. Turley, S. E. H. *Electron. Lett.* **18,** 590 (1982).
15. Dutta, N. K., N. A. Olsson, and W. T. Tsang. *Appl. Phys. Lett.* **45,** 836 (1984).
16. Manning, J., R. Olshansky, and C. B. Su. *IEEE J. Quantum Electron.* **QE-19,** 1525 (1983).
17. Yamada, M., and Y. Suematsu. *IEEE J. Quantum Electron.* **QE-15,** 743 (1979); *J. Appl. Phys.* **52,** 2653 (1981).
18. Yamada, M. *IEEE J. Quantum Electron.* **QE-19,** 1365 (1983).
19. Kazarinov, R. F., C. H. Henry, and R. A. Logan. *J. Appl. Phys.* **53,** 4631 (1982).
20. Yamada, M., H. Ishiguro, and H. Nagato. *Jpn. J. Appl. Phys.* **19,** 135 (1980).
21. Yamanishi, M., I. Suemune, K. Nonomura, and N. Mikoshiba. *Jpn. J. Appl. Phys.* **21,** 1240 (1982).
22. Dutta, N. K., and R. J. Nelson. *Appl. Phys. Lett.* **38,** 407 (1980); *J. Appl. Phys.* **53,** 74 (1982).
23. Sugimura, A. *IEEE J. Quantum Electron.* **QE-17,** 627 (1981); **QE-19,** 930 (1983).
24. Haug, A. *Appl. Phys. Lett.* **42,** 512 (1983); *J. Phys. C: Solid State Phys.* **16,** 4159 (1983).
25. Haug, H., and S. W. Koch. *Phys. Rev. A* **39,** 1887 (1989).
26. Chow, W. W., S. W. Koch, and M. Sargent III. *IEEE J. Quantum Electron.* **QE-26,** 1052 (1990); *Appl. Phys. Lett.* **58,** 328 (1991).
27. Lasher, G., and F. Stern. *Phys. Rev.* **133,** A553 (1964).
28. Stern, F. *Phys. Rev.* **148,** 186 (1966).
29. Hwang, C. J. *Phys. Rev. B* **2,** 4126 (1970).
30. Stern, F. *Phys. Rev. B* **3,** 2636 (1971).
31. Einstein, A. *Z. Phys.* **18,** 121 (1917).
32. Henry, C. H., R. A. Logan, and F. R. Merritt. *J. Appl. Phys.* **51,** 3042 (1980).
33. Asada, M., and Y. Suematsu. *IEEE J. Quantum Electron.* **QE-19,** 917 (1983).
34. Cook, D. D., and F. R. Nash. *J. Appl. Phys.* **46,** 1660 (1975).
35. Kirkby, P. A., A. R. Goodwin, G. H. B. Thompson, and P. R. Selway. *IEEE J. Quantum Electron.* **QE-13,** 705 (1977).
36. Streifer, W., R. D. Burnham, and D. R. Scifres. *IEEE J. Quantum Electron.* **QE-15,** 136 (1979).
37. Henry, C. H. *IEEE J. Quantum Electron.* **QE-18,** 259 (1982).
38. Harder, C., K. Vahala, and A. Yariv. *Appl. Phys. Lett.* **42,** 328 (1983).
39. Osinski, M., and J. Buus. *IEEE J. Quantum Electron.* **QE-23,** 9 (1987).
40. Asada, M., and Y. Suematsu. *Appl. Phys. Lett.* **41,** 353 (1982).
41. Olshansky, R., C. B. Su, J. Manning, and W. Powazinik. *IEEE J. Quantum Electron.* **QE-20,** 838 (1984).
42. Stern, F. *J. Appl. Phys.* **47,** 5382 (1976).
43. Su, C. B., and R. Olshanky. *Appl. Phys. Lett.* **40,** 833 (1982).
44. Dyment, J. C. *Appl. Phys. Lett.* **10,** 84 (1967).
45. Zachos, T. H., and J. E. Ripper. *IEEE J. Quantum Electron.* **QE-5,** 29 (1969).
46. Marcatili, E. J. *Bell Syst. Tech. J.* **48,** 2133 (1969).
47. Goell, J. E. *Bell Syst. Tech. J.* **48,** 2133 (1969).
48. Knox, R. M., and P. P. Toulios, in *Proc. MRI Symposium on Submillimeter Waves*, ed. J. Fox. Brooklyn, N.Y.: Polytechnic Press, 1970.
49. Streifer, W., and E. Kapon. *Appl. Opt.* **18,** 3724 (1979).
50. Buus, J. *IEEE J. Quantum Electron.* **QE-18,** 1083 (1982).

51. Paoli, T. L. *IEEE J. Quantum Electron.* **QE-13,** 662 (1976).
52. Kirkby, P. A., and G. H. B. Thompson. *J. Appl. Phys.* **47,** 4578 (1976).
53. Streifer, W., R. D. Burnham, and D. R. Scifres. *Appl. Phys. Lett.* **37,** 121 (1980).
54. Marcuse, D. *Theory of Dielectric Optical Waveguides.* New York: Academic Press, 1991.
55. Unger, H.-G. *Planar Optical Waveguides and Fibers.* Oxford: Clarendon Press, 1977.
56. Adams, M. J. *An Introduction to Optical Waveguides.* New York: John Wiley & Sons, 1981.
57. Scholosser, W. O. *Bell Syst. Tech. J.* **52,** 887 (1973).
58. Morse, P. M., and H. Feshbach. *Methods of Theoretical Physics,* Chap. 9. New York: McGraw-Hill Book Company, 1953.
59. Botez, D. *IEEE J. Quantum Electron.* **QE-17,** 178 (1981).
60. Lang, R. *IEEE J. Quantum Electron.* **QE-15,** 718 (1979).
61. Seki, K., T. Kamiya, and H. Yanai. *Trans. Inst. Electron. Commun. Eng. Jpn.* **E62,** 73 (1979).
62. Buus, J. *IEEE J. Quantum Electron.* **QE-15,** 734 (1979).
63. Asbeck, P. M., D. A. Cammack, J. J. Daniele, and V. Klebanoff. *IEEE J. Quantum Electron.* **QE-15,** 727 (1979).
64. Shore, K. A., T. E. Rozzi, and G. H. in't Veld. *IEE Proc. I* **127,** 221 (1980).
65. Streifer, W., R. D. Burnham, and D. R. Scifres. *IEEE J. Quantum Electron.* **QE-17,** 736 (1981); *IEEE J. Quantum Electron.* **QE-18,** 856 (1982).
66. Wilt, D. P., and A. Yariv. *IEEE J. Quantum Electron.* **QE-17,** 1941 (1981).
67. Butler, J. K., and D. Botez. *Appl. Phys. Lett.* **41,** 1118 (1982).
68. Buus, J. *IEEE J. Quantum Electron.* **QE-19,** 953 (1983).
69. Agrawal, G. P., W. B. Joyce, R. W. Dixon, and M. Lax. *Appl. Phys. Lett.* **43,** 11 (1983).
70. Agrawal, G. P. *J. Appl. Phys.* **56,** 3100 (1984).
71. Meissner, P., E. Patzak, and D. Yevick. *IEEE J. Quantum Electron.* **QE-20,** 899 (1984).
72. Agrawal, G. P. *J. Lightwave Technol.* **LT-2,** 537 (1984).
73. Agrawal, G. P., and N. K. Dutta. *Electron. Lett.* **19,** 974 (1983).
74. Nash, F. R. *J. Appl. Phys.* **44,** 4696 (1973).
75. Unger, K. *Annalen der Physik* **19,** 64 (1967).
76. Lee, T. P., C. A. Burrus, B. I. Miller, and R. A. Logan. *IEEE J. Quantum Electron.* **QE-11,** 432 (1975).
77. Kawaguchi, H., and T. Kawakami. *IEEE J. Quantum Electron.* **QE-13,** 556 (1977).
78. Aiki, K., M. Nakamura, T. Kuroda, and J. Umeda. *Appl. Phys. Lett.* **48,** 649 (1977).
79. Ueno, M., T. Lang, S. Matsumoto, H. Kawano, T. Furuse, and I. Sakuma. *IEE Proc. I* **129,** 218 (1982).
80. Biard, J. R., W. N. Carr, and B. S. Reed. *Trans. Metall. Soc. AIME* **230,** 286 (1964).
81. Henshall, G. D. *Appl. Phys. Lett.* **31,** 205 (1977).
82. Reinhart, F. K., I. Hayashi, and M. B. Panish. *J. Appl. Phys.* **42,** 4466 (1971).
83. Ikegami, T. *IEEE J. Quantum Electron.* **QE-8,** 470 (1972).
84. Lewin, L. *IEEE Trans. Microwave Theory Tech.* **MTT-23,** 576 (1975).
85. Davis, R. W., and J. N. Walpole. *IEEE J. Quantum Electron.* **QE-12,** 291 (1976).
86. Buus, J. *IEEE J. Quantum Electron.* **QE-17,** 2256 (1981).
87. Kardontchiek, J. E. *IEEE J. Quantum Electron.* **QE-18,** 1279 (1982).
88. Hardy, A. *J. Opt. Soc. Am. A* **1,** 550 (1984).
89. Botez, D., and M. Ettenberg. *IEEE J. Quantum Electron.* **QE-14,** 827 (1978).
90. Nelson, R. J., R. B. Wilson, P. D. Wright, P. A. Barnes, and N. K. Dutta. *IEEE J. Quantum Electron.* **QE-17,** 202 (1981).
91. Agrawal, G. P., Vol. 26, Chap. 3, in *Progress in Optics,* ed. by E. Wolf. Amsterdam: North-Holland Pub. Co., 1988.

Chapter 3
RECOMBINATION MECHANISMS IN SEMICONDUCTORS

3.1 INTRODUCTION

This chapter describes the electron-hole recombination mechanisms in a direct-band-gap semiconductor. Recombination mechanisms can in general be classified into two groups, radiative and nonradiative. *Radiative recombination* occurs when an electron in the conduction band recombines with a hole in the valence band and the excess energy is emitted in the form of a photon. Radiative recombination is thus the radiative transition of an electron in the conduction band to an empty state (hole) in the valence band. The optical processes associated with radiative transitions are (i) spontaneous emission, (ii) absorption or gain, and (iii) stimulated emission. Stimulated emission, in which the emitted photon has exactly the same energy and momentum as the incident photon, forms the basis for laser action. The concept of stimulated emission dates back to the work of Einstein[1] in 1917.

In thermal equilibrium, a direct-band-gap semiconductor (e.g., GaAs, InP, or GaSb) has a few electrons in the conduction band and a few holes (empty electron states) in the valence band. When a photon of energy greater than the band gap passes through such a semiconductor, the photon has a high probability of being absorbed, giving its energy to an electron in the valence band, thereby raising the electron to the conduction band. In principle, such a photon could stimulate the emission of an identical photon with the transition of an electron from the conduction to the valence band. The emitted photon derives its energy from the energy lost by the electron. In thermal equilibrium the number of electrons in the conduction band is very small ($\sim 10^6$ cm^{-3} for GaAs), so the probability of stimulated emission is negligible compared to the probability for absorption. However, external excitation, can sufficiently increase the number of electrons in the conduction band such that the probability of stimulated emission eventually becomes higher than the probability of absorption. This situation corresponds to population

inversion in a laser medium and is necessary for optical gain. The external excitation which generates a high density of electron-hole pairs in a semiconductor is usually provided by current injection. It can also be achieved by optical pumping (absorption of radiation higher in energy than the band gap). In Sec. 3.2 we discuss the absorption and emission rates. These are then used to study the dependence of optical gain on the injected current and the carrier density.

Nonradiative recombination of an electron-hole pair, as the name implies, is characterized by the *absence* of an emitted photon in the recombination process. This of course makes the experimental study, and hence identification, of such processes very difficult. In indirect-band-gap semiconductors such as Ge or Si, the probability of nonradiative recombination dominates that of radiative recombination by several orders of magnitude. The measurable quantities associated with nonradiative recombinations are the internal quantum efficiency and the carrier lifetime. The variation of these quantities with parameters like temperature, pressure and carrier concentration is, by and large, the only way to identify a particular nonradiative recombination process. One of the effects of nonradiative recombination on the performance of injection lasers is to increase the threshold current.

The nonradiative recombination processes that affect the performance of long-wavelength semiconductor lasers are Auger recombination, surface recombination, and recombination at defects. The Auger recombination mechanism involves four particle states (three electrons and one hole) and is believed to be important at high temperatures and for low band-gap semiconductors. Section 3.3 gives a detailed discussion of the various Auger recombination mechanisms. Section 3.4 presents the experimental results on the measurement of radiative and nonradiative recombination coefficients. Section 3.5 considers the threshold current density and the contribution of carrier leakage to it. Finally, in Sec. 3.6 we discuss the temperature dependence of the threshold current.

3.2 RADIATIVE RECOMBINATION

The processes associated with the radiative recombination of electron-hole pairs in semiconductors are spontaneous emission, optical absorption or gain, and stimulated emission. The rates of these processes are related to each other by the Einstein relations. These relations were first derived for electronic transitions between two discrete energy levels (i.e., in a gaseous medium). Similar relations also hold for a semiconductor where the recombining electrons and holes can occupy a continuous band of energy eigenstates.[2]

Figure 3.1 shows a simplified energy-versus-wave-vector diagram for a direct-band-gap semiconductor. An accurate description of the band structure

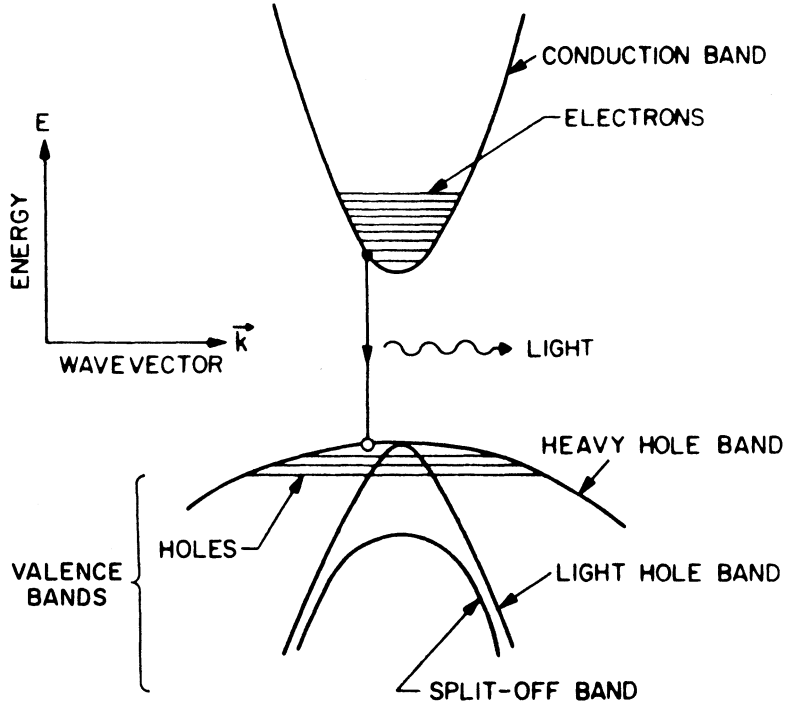


Fig. 3.1 Energy versus wave-vector diagram of a direct-band-gap semiconductor showing schematically the conduction and valence bands. Three valence bands are required to model band-to-band transitions realistically. Horizontal lines show the filled energy states. Radiative recombination of electrons and holes generates photons.

requires sophisticated numerical techniques. A commonly used approximation of the exact band structure in a direct-gap semiconductor is the parabolic band model. In this model the energy-versus-wave vector (E versus $\mathbf{k}$) relation is assumed to be parabolic, that is,

$$E_c = \frac{\hbar^2 k^2}{2m_c} \qquad \text{for electrons} \tag{3.2.1a}$$

$$E_v = \frac{\hbar^2 k^2}{2m_v} \qquad \text{for holes} \tag{3.2.1b}$$

where m_c and m_v are the effective masses of electrons and holes, respectively, and k is the magnitude of the wave vector $\mathbf{k}$.

In a direct-gap semiconductor, the minimum in the conduction band curve and the maximum in the valence band curve occur at the same value of

the wave vector $\mathbf{k}$ ($\mathbf{k} = 0$ in Fig. 3.1). Since a photon carries negligible momentum compared with the carrier momentum $\hbar\mathbf{k}$, radiative transitions occur between free electrons and free holes of essentially identical wave vectors.

At a given temperature T, the available number of electrons and holes are distributed over a range of energies. The occupation probability f_c of an electron with energy E_c is given by the Fermi-Dirac statistics, and is

$$f_c(E_c) = \frac{1}{\exp[(E_c - E_{\mathrm{fc}})/k_B T] + 1} \tag{3.2.2}$$

where E_{fc} is the quasi-Fermi level for the conduction band. Similarly for the valence-band holes, the occupation probability of a hole with energy E_v is

$$f_v(E_v) = \frac{1}{\exp[(E_v - E_{\mathrm{fv}})/k_B T] + 1} \tag{3.2.3}$$

where E_{fv} is the quasi-Fermi level for the valence band. The notations used here assume that E_c and E_{fc} are measured from the conduction-band edge and are positive into the conduction band. On the other hand, E_v and E_{fv} are measured from the valence-band edge and are positive into the valence band. Note that f_v represents the occupation probability of the hole and not of the electron.

Consider the transition shown in Fig. 3.1 in the presence of an incident photon whose energy $E = h\nu = E_c + E_v + E_g$, where E_g is the band gap. The photon can be absorbed creating an electron of energy E_c and a hole of energy E_v. The absorption rate is given by

$$R_a = B(1 - f_c)(1 - f_v)\rho(E) \tag{3.2.4}$$

where B is the transition probability, $\rho(E)$ is the density of photons of energy E, and the factors $1 - f_c$ and $1 - f_v$ represent the probabilities that the electron and hole states of energy E_c and E_v are not occupied. The stimulated-emission rate of photons is given by

$$R_e = B f_c f_v \rho(E) \tag{3.2.5}$$

where the Fermi factors f_c and f_v are the occupation probabilities of the electron and hole states of energy E_c and E_v respectively. The stimulated-emission process is accompanied by the recombination of an electron-hole

pair. The condition for net stimulated emission or optical gain is

$$R_e > R_a. \tag{3.2.6}$$

Using Eqs. (3.2.4) and (3.2.5), this condition becomes

$$f_c + f_v > 1. \tag{3.2.7}$$

With the use of f_c and f_v given by Eqs. (3.2.2) and (3.2.3), this becomes

$$E_{fc} + E_{fv} > E_c + E_v. \tag{3.2.8}$$

By adding E_g to both sides and noting that $E_c + E_v + E_g = h\nu$, it follows that the separation of the quasi-Fermi levels must exceed the photon energy in order for the stimulated-emission rate to exceed the absorption rate. This is the necessary condition for net stimulated emission or optical gain.[3]

3.2.1 Absorption and Emission Rates for Discrete Levels

The spontaneous-emission rate and the absorption rate for photons in a semiconductor can be calculated using time-dependent perturbation theory and summing over the available electron and hole states. We first consider the case of two discrete energy levels of energy E_i and E_f. The electron makes a transition from an initial state of energy E_i to a final state of energy E_f, and in the process a photon of energy $E = E_i - E_f$ is emitted. The transition probability or the emission rate W of such a process is given by Fermi's golden rule

$$W = \frac{2\pi}{\hbar} |H'_{if}|^2 \rho(E_f) \delta(E - E_i + E_f) \tag{3.2.9}$$

where H'_{if} is the matrix element $\langle i|H'|f\rangle$ of the time-independent part of the perturbation Hamiltonian H_I between the initial state $|i\rangle$ and final state $|f\rangle$, and $\rho(E_f)$ is the density of the final states. In obtaining Eq. (3.2.9), the perturbation Hamiltonian is assumed to be of the form

$$H_I = 2H' \sin(\omega t). \tag{3.2.10}$$

The time-dependent sinusoidal part gives rise to the delta function of Eq. (3.2.9) arising from the energy conservation during the emission process.

For photons interacting with free electrons of mass m_0, the perturbation

Hamiltonian is given by[5]

$$H_I = -\frac{q}{m_0}\mathbf{A}\cdot\mathbf{p} \tag{3.2.11}$$

where $\mathbf{A}$ is the vector potential of the electromagnetic field and $\mathbf{p}$ is the electron-momentum vector. The vector potential is related to the electric field $\mathscr{E}$ by the equation

$$\mathscr{E} = -\frac{\partial \mathbf{A}}{\partial t}. \tag{3.2.12}$$

For a traveling plane wave with the angular frequency ω, the electric field is written as

$$\mathscr{E} = E_0\hat{\boldsymbol{\epsilon}}\cos(\omega t - \mathbf{k}\cdot\mathbf{r}) \tag{3.2.13}$$

where $\hat{\boldsymbol{\epsilon}}$ is a unit vector determining the polarization of the field, E_0 is its amplitude, and $\mathbf{k}$ denotes the propagation vector such that $\mathbf{k}\times\hat{\boldsymbol{\epsilon}} = 0$. The magnitude of the electric field E_0 associated with one photon can be obtained as follows. The magnitude of the Poynting vector that represents the energy flux is given by

$$S = |\langle \mathscr{E}\times\mathscr{H}\rangle| \tag{3.2.14}$$

where angular brackets denote time averaging and $\mathscr{H}$ is the magnetic field, which can be obtained using Maxwell's equations (see Sec. 2.2). Using them, we obtain

$$S = \tfrac{1}{2}E_0^2\bar{\mu}\epsilon_0 c \tag{3.2.15}$$

where $k = \bar{\mu}\omega/c$ and $\bar{\mu}$ is the refractive index of the medium. The factor 1/2 arises from time averaging over one optical cycle. The energy flux per photon is also given by the product of the photon energy $\hbar\omega$ and the group velocity $c/\bar{\mu}$. Hence it follows that

$$S = \hbar\omega c/\bar{\mu}. \tag{3.2.16}$$

From Eqs. (3.2.12), (3.2.15), and (3.2.16), the vector potential $\mathbf{A}$ is given by

$$\mathbf{A} = -\hat{\boldsymbol{\epsilon}}\left(\frac{2\hbar}{\epsilon_0\bar{\mu}^2\omega}\right)^{1/2}\sin(\omega t - \mathbf{k}\cdot\mathbf{r}) \tag{3.2.17}$$

and using Eq. (3.2.11) the perturbation Hamiltonian becomes

$$H_I = \frac{q}{m_0}\left(\frac{2\hbar}{\epsilon_0 \bar{\mu}^2 \omega}\right)^{1/2} (\hat{\boldsymbol{\epsilon}} \cdot \mathbf{p}) \sin(\omega t - \mathbf{k} \cdot \mathbf{r}). \tag{3.2.18}$$

The square of the transition matrix element H'_{if} is now given by

$$|H'_{if}|^2 = \frac{q^2}{m_0^2} \frac{2\hbar}{\epsilon_0 \bar{\mu}^2 \omega} \frac{1}{4} |\langle i|\hat{\boldsymbol{\epsilon}} \cdot \mathbf{p}|f\rangle|^2 \tag{3.2.19}$$

where we have made the usual assumption that the electron wave functions spread over a linear dimension shorter than the wavelength ($r \ll \lambda$) so that $\mathbf{k} \cdot \mathbf{r} \ll 1$.

For stimulated emission or absorption, $\rho(E_f) = 1$, and the absorption or emission rate at a photon energy $E = \hbar\omega$ for transitions between two discrete levels obtained from Eqs. (3.2.9) and (3.2.19) is given by

$$W = \frac{\pi q^2}{m_0^2 \epsilon_0 \bar{\mu}^2 \omega} |M_{if}|^2 \delta(E_i - E_f - \hbar\omega) \tag{3.2.20}$$

where $M_{if} = \langle i|\hat{\boldsymbol{\epsilon}} \cdot \mathbf{p}|f\rangle$ is the momentum matrix element between the initial and final electron states. In the case of absorption, the number of photons absorbed per second is W. Since a photon travels a distance $c/\bar{\mu}$ in 1 s, the number of photons absorbed per unit distance or the absorption coefficient $\alpha = \bar{\mu}W/c$. Using Eq. (3.2.20), the absorption coefficient for photons of energy $E = \hbar\omega$ becomes

$$\alpha(E) = \frac{q^2 h}{2\epsilon_0 m_0^2 c \bar{\mu} E} |M_{if}|^2 \delta(E_i - E_f - E). \tag{3.2.21}$$

For spontaneous emission, the quantity $\rho(E_f)$ in Eq. (3.2.9) equals the number of states for photons of energy E per unit volume per unit energy. It is given by

$$\rho(E_f) = (2) \frac{4\pi k^2}{(2\pi)^3} \frac{dk}{dE} \tag{3.2.22}$$

where the factor 2 arises from two possible polarizations of the electromagnetic

field. Since $E = \hbar\omega$ and $k = \bar{\mu}\omega/c$, Eq. (3.3.22) reduces to

$$\rho(E_f) = \frac{\bar{\mu}^3\omega^2}{\pi^2 c^3 \hbar}. \tag{3.2.23}$$

Thus using Eqs. (3.2.9), (3.2.19), and (3.2.23), the spontaneous emission rate per unit volume at the photon energy E is given by

$$r_{sp}(E) = \frac{4\pi q^2 \bar{\mu} E}{m_0^2 \epsilon_0 c^3 h^2} |M_{if}|^2 \, \delta(E_i - E_f - E). \tag{3.2.24}$$

3.2.2 Absorption and Emission Rates in Semiconductors

We now calculate the spontaneous-emission rate and the absorption coefficient for direct-gap semiconductors. These quantities are obtained by integrating Eqs. (3.2.21) and (3.2.24) over the occupied electron and hole states in a semiconductor.[6] A realistic band model for a III–V direct-gap semiconductor is the four-band model[7] shown in Fig. 3.1. In this model the valence band is represented by three subbands: the heavy-hole band, the light-hole band, and the spin-split-off-hole band. The light- and heavy-hole bands are degenerate at $\mathbf{k} = 0$. Usually the split-off energy Δ is large compared with the thermal energy $k_B T$; hence the split-off band is full of electrons, that is, no holes are present. The quasi-Fermi levels for the conduction and valence bands can be obtained from the known effective masses and the density of states in each band. For the conduction band, E_{fc} is obtained using

$$n = \int \frac{\rho_c(E)\, dE}{1 + \exp[(E - E_{fc})/k_B T]} \tag{3.2.25}$$

where $\rho_c(E)$ is the conduction-band density of states and n is the number of electrons in the conduction band. The density of states for electrons of energy E can be obtained using Eq. (3.2.22) with the dispersion relation (3.2.1a) and becomes

$$\rho_c(E) = (2)\frac{4\pi k^2}{(2\pi)^3}\frac{dk}{dE} = 4\pi\left(\frac{2m_c}{h^2}\right)^{3/2} E^{1/2} \tag{3.2.26}$$

where $E = \hbar^2 k^2/2m_c$, and m_c is the conduction-band effective mass. The factor 2 in Eq. (3.2.26) arises from the two electronic spin states. Equation (3.2.25) can be rewritten as

$$n = N_c \frac{2}{\pi^{1/2}} \int \frac{\epsilon^{1/2}\, d\epsilon}{1 + \exp(\epsilon - \epsilon_{fc})} \tag{3.2.27}$$

where $N_c = 2(2\pi m_c k_B T/h^2)^{3/2}$ and $\epsilon_{fc} = E_{fc}/k_B T$. Equation (3.2.27) can be used to calculate ϵ_{fc} for a given electron concentration. A useful approximation for ϵ_{fc} is given by Joyce and Dixon.[8] They show that ϵ_{fc} can be represented by a convergent series of the form

$$\epsilon_{fc} = \ln\left(\frac{n}{N_c}\right) + \sum_{i=1}^{\infty} A_i \left(\frac{n}{N_c}\right)^i \tag{3.2.28}$$

where the first few coefficients are

$$\begin{aligned} A_1 &= 3.53553 \times 10^{-1} \\ A_2 &= -4.95009 \times 10^{-3} \\ A_3 &= 1.48386 \times 10^{-4} \\ A_4 &= 4.42563 \times 10^{-6}. \end{aligned}$$

Another approximation, although not as elegant but showing faster convergence, is the Nilson approximation.[9]

For a nondegenerate electron gas ($n \ll N_c$), all terms except the first in Eq. (3.2.28) can be neglected and the Fermi energy is given by

$$E_{fc} = k_B T \ln\left(\frac{n}{N_c}\right). \tag{3.2.29}$$

Using Eq. (3.2.29) in Eq. (3.2.2) and neglecting 1 in the denominator, the occupation probability becomes

$$f_c(E) \cong \frac{n}{N_c} \exp\left(\frac{-E}{k_B T}\right). \tag{3.2.30}$$

This is often referred to as the *Boltzmann approximation.* The use of Eq. (3.2.30) simplifies considerably the calculation of recombination rates in semiconductors. The nondegenerate approximation $n \ll N_c$, however, does not always hold for semiconductor lasers and should be used with caution.

The quasi-Fermi level for the holes in the valence band can be similarly calculated. The hole density is given by

$$p = \sum_{i=l,h} \int \frac{\rho_{vi}(E)\, dE}{1 + \exp[(E - E_{fv})/k_B T]} \tag{3.2.31}$$

where p is the density of holes and the summation is over the light- and heavy-hole bands. Using an expression similar to Eq. (3.2.26) for the density

of states, we obtain

$$p = N_v \frac{2}{\pi^{1/2}} \int \frac{\epsilon^{1/2}\, d\epsilon}{1 + \exp(\epsilon - \epsilon_{fv})} \tag{3.2.32}$$

where

$$N_v = 2(2\pi k_B T/h^2)^{3/2}(m_{lh}^{3/2} + m_{hh}^{3/2})$$

and m_{lh} and m_{hh} are the effective masses of light and heavy holes, respectively. For the nondegenerate case where $p \ll N_v$, Eq. (3.2.32) can be simplified to obtain the following expressions for the quasi-Fermi energy and the hole-occupation probability:

$$E_{fv} = k_B T \ln\left(\frac{p}{N_v}\right) \tag{3.2.33a}$$

$$f_v(E) = \frac{p}{N_v} \exp\left(\frac{-E}{k_B T}\right). \tag{3.2.33b}$$

So far we have considered the idealized case of parabolic bands with the density of states given by Eq. (3.2.26). Under high injected carrier densities or for high doping levels, the density of states is modified. In the case of heavy doping, this occurs because randomly distributed impurities create an additional continuum of states near the band edge. These are called *band-tail states*[10,11] and are schematically shown in Fig. 3.2.

Several models of band-tail states exist in the literature. Principal among them are the Kane model[10] and the Halperin-Lax model.[11] Figure 3.3 shows a comparison of the calculated density of states for the two models. The curves were calculated by Hwang[12] for p-type GaAs with a hole concentration of 3×10^{18} cm^{-3}. The material was heavily compensated with acceptor and donor concentrations N_A and N_D of 6×10^{18} cm^{-3} and 3×10^{18} cm^{-3}, respectively. Figure 3.3 shows that for the Kane model the band tails extend much more into the band gap than for the Halperin-Lax model. The absorption measurements on heavily compensated samples indicate that the Halperin-Lax model is more realistic. However, this model is not valid for states very close to the band edge. To remedy the situation, Stern[13] approximated the density of states by fitting a density-of-state equation of the Kane form (which is Gaussian) to the Halperin-Lax model of band-tail states. The fit is done at *one* energy value in the band tail, which lies within the validity of the Halperin-Lax model. The Kane form of band tail merges with the parabolic band model above the band edge as seen in Fig. 3.3.

The importance of band-tail states in the operation of semiconductor lasers arises from two considerations. First, band-tail states can contribute

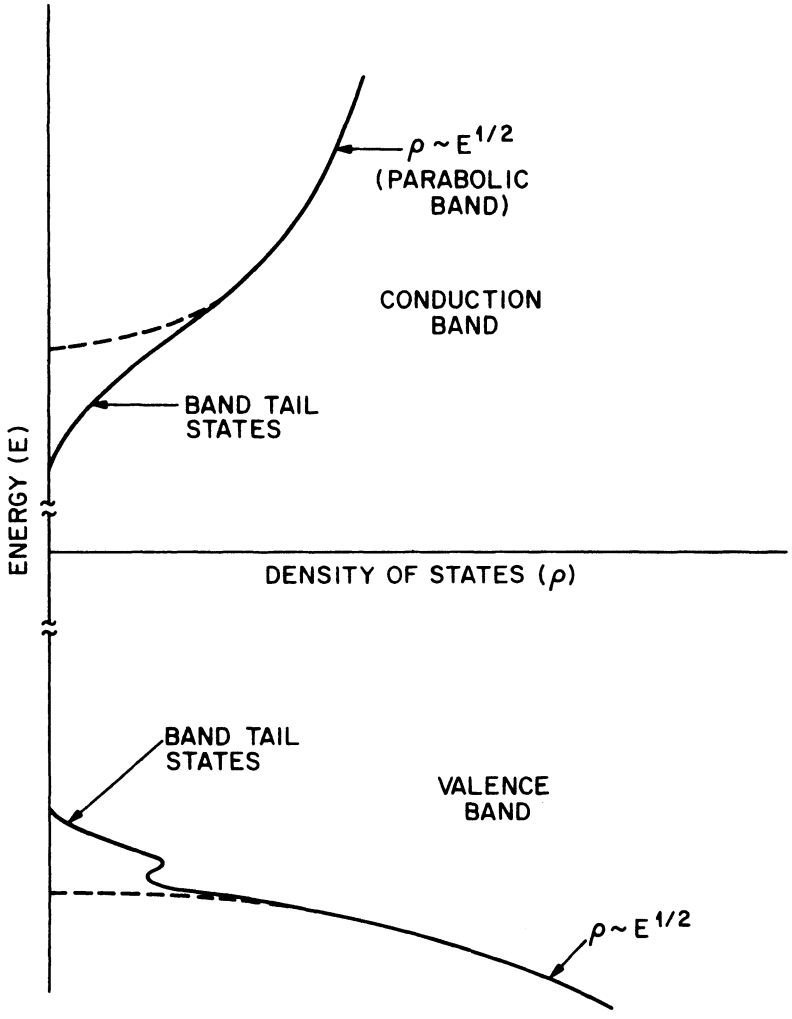


Fig. 3.2 Energy for the conduction and valence bands plotted against density of states. Dashed curves are for the parabolic-band model, while solid curves show schematically the effect of band-tail states.

significantly to the total spontaneous-emission rate and hence to the injection current needed to reach threshold. Second, band-tail transitions can have significant optical gain so that the laser can lase on these transitions, especially for heavily doped semiconductors at low temperatures.

The matrix element for transitions between band-tail states differs from that involving free electron and hole states in one important aspect. The

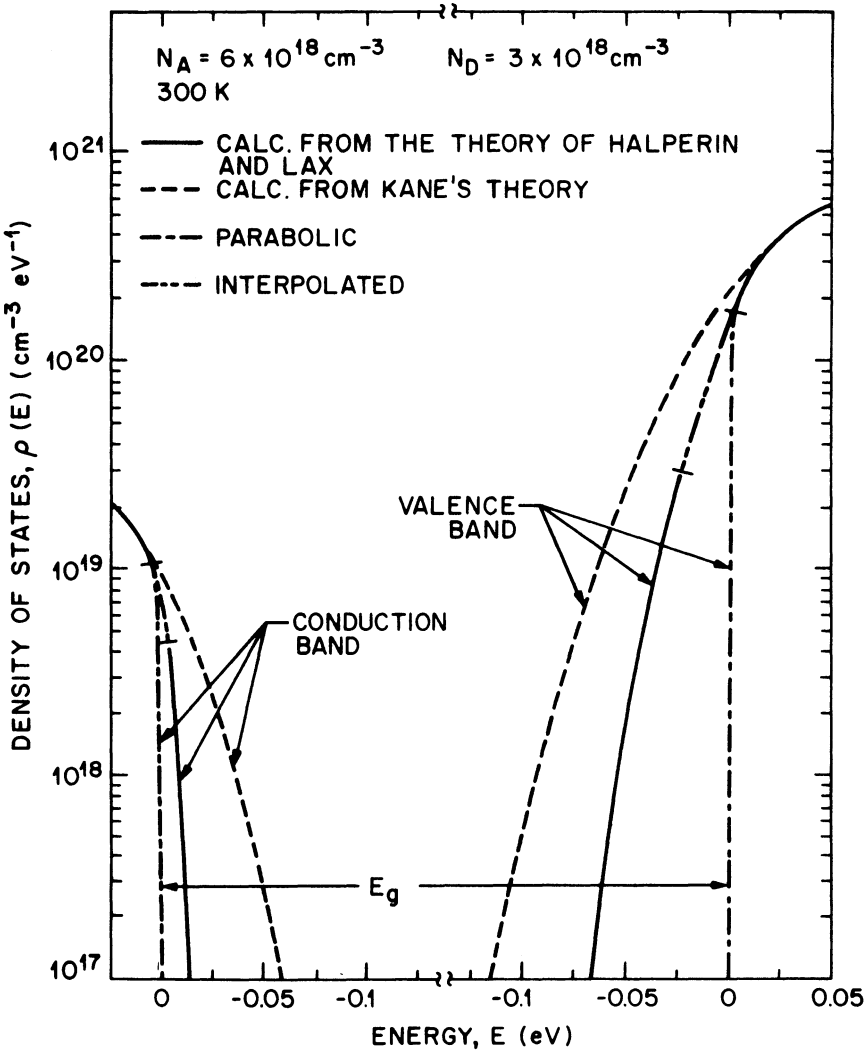


Fig. 3.3 Calculated density of states versus energy in the conduction and valence bands for p-type GaAs with a net carrier concentration of $3 \times 10^{18}\,cm^{-3}$. The acceptor concentration $N_A = 6 \times 10^{18}\,cm^{-3}$, and the donor concentration $N_D = 3 \times 10^{18}\,cm^{-3}$. (After Ref. 12)

former does not satisfy momentum conservation; i.e., since the electron and holes states are not states of definite momentum, the **k**-selection rule does not apply. By contrast, for transitions involving parabolic band states, the initial and final particle states obey the **k**-selection rule. Thus the exact matrix element for optical transitions in heavily compensated semiconductors must

not invoke the **k**-selection rule for transitions involving band-tail states and, at the same time, should extrapolate to that obeying the **k**-selection rule for above-band-edge transitions involving parabolic band states. The model for such a matrix element has been considered by Stern.[14–16] Before discussing his model in the next section, we consider the simpler case where the **k**-selection rule holds and obtain the absorption coefficient and the spontaneous emission rate.

The k-Selection Rule. When the **k**-selection rule is obeyed, $|M_{if}|^2 = 0$ unless $\mathbf{k}_c = \mathbf{k}_v$. If we consider a volume V of the semiconductor, the matrix element $|M_{if}|^2$ is given by

$$|M_{if}|^2 = |M_b|^2 \frac{(2\pi)^3}{V} \delta(\mathbf{k}_c - \mathbf{k}_v). \tag{3.2.34}$$

The δ function accounts for the momentum conservation between the conduction-band and valence-band states. The quantity $|M_b|$ is an average matrix element for the Bloch states. Using the Kane model,[7] $|M_b|^2$ in bulk semiconductors is given by[7,17]

$$|M_b|^2 = \frac{m_0^2 E_g(E_g + \Delta)}{12m_c(E_g + 2\Delta/3)} = \xi m_0 E_g \tag{3.2.35}$$

where m_0 is the free-electron mass, E_g is the band gap, and Δ is the spin-orbit splitting. For GaAs, using $E_g = 1.424$ eV, $\Delta = 0.33$ eV, $m_c = 0.067\, m_0$, we get $\xi = 1.3$.

We are now in a position to calculate the spontaneous-emission rate and the absorption coefficient for a bulk semiconductor. Equations (3.2.24) and (3.2.34) can be used to obtain the total spontaneous-emission rate per unit volume. Summing over all states in the band, we obtain

$$r_{sp}(E) = \frac{4\pi\bar{\mu}q^2E}{m_0^2\epsilon_0 h^2c^3} |M_b|^2 \frac{(2\pi)^3}{V} (2)\left(\frac{V}{(2\pi)^3}\right)^2 \frac{1}{V}$$
$$\times \sum \int \cdots \int f_c(E_c) f_v(E_v)\, d^3\mathbf{k}_c\, d^3\mathbf{k}_v\, \delta(\mathbf{k}_c - \mathbf{k}_v)\delta(E_i - E_f - E) \tag{3.2.36}$$

where f_c and f_v are the Fermi factors for electrons and holes. The factor 2 arises from the two spin states. In Eq. (3.2.36), $\sum$ stands for the sum over the three valence bands (see Fig. 3.1). For definiteness, we first consider transitions involving electrons and heavy holes. The integrals in Eq. (3.2.36)

can be evaluated with the following result:

$$r_{sp}(E) = \frac{2\bar{\mu}q^2E|M_b|^2}{\pi m_0^2\epsilon_0 h^2c^3}\left(\frac{2m_r}{\hbar^2}\right)^{3/2}(E - E_g)^{1/2}f_c(E_c)f_v(E_v) \qquad (3.2.37)$$

where

$$E_c = \frac{m_r}{m_c}(E - E_g)$$

$$E_v = \frac{m_r}{m_{hh}}(E - E_g)$$

$$m_r = \frac{m_c m_{hh}}{m_c + m_{hh}}$$

and m_{hh} is the effective mass of the heavy hole. Equation (3.2.37) gives the spontaneous-emission rate at the photon energy E. To obtain the total spontaneous-emission rate, a final integration should be carried out over all possible energies. Thus the total spontaneous-emission rate per unit volume due to electron-heavy-hole transitions is given by

$$R = \int_{E_g}^{\infty} r_{sp}(E)\,dE = A|M_b|^2 I \qquad (3.2.38)$$

where

$$I = \int_{E_g}^{\infty}(E - E_g)^{1/2}f_c(E_c)f_v(E_v)\,dE \qquad (3.2.39)$$

and A represents the remaining constants in Eq. (3.2.37). A similar equation holds for the electron–light-hole transitions if we replace m_{hh} by the effective light-hole mass m_{ll}.

The absorption coefficient $\alpha(E)$ can be obtained in a similar way using Eqs. (3.2.21) and (3.2.34) and integrating over the available states in the conduction and valence bands. The resulting expression is

$$\begin{aligned}\alpha(E) &= \frac{q^2h}{2\epsilon_0 m_0^2 c\bar{\mu}E}|M_b|^2\frac{(2\pi)^3}{V}(2)\left(\frac{V}{(2\pi)^3}\right)^2\left(\frac{1}{V}\right)\\ &\quad\times\int\cdots\int(1 - f_c - f_v)\,d^3\mathbf{k}_c d^3\mathbf{k}_v\,\delta(\mathbf{k}_c - \mathbf{k}_v)\,\delta(E_i - E_f - E)\\ &= \frac{q^2h|M_b|^2}{4\pi^2\epsilon_0 m_0^2 c\bar{\mu}E}\left(\frac{2m_r}{\hbar^2}\right)^{3/2}(E - E_g)^{1/2}[1 - f_c(E_c) - f_v(E_v)]. \qquad (3.2.40)\end{aligned}$$

For the case of a constant matrix element, the assumption is that $|M_{if}|^2$ is independent of the energy and wave vector of the initial and final states. In such a case, the total spontaneous emission rate and absorption coefficient are obtained by summing over the initial and final states using Eqs. (3.2.21) and (3.2.24). The calculation is similar to that presented above; the only difference is that in the latter case the delta function representing momentum conservation is absent.

3.2.3 Absorption Coefficient and Optical Gain

We now describe the calculation of the spontaneous-emission rate and the absorption coefficient using the Gaussian Halperin-Lax band-tail model and Stern's matrix element. The results in this section are presented for the quaternary InGaAsP alloy emitting near 1.3 μm or 1.55 μm. The gain calculations for the InGaAsP material have been given in Ref. 18–21.

The optical absorption (or gain) coefficient for transitions between the valence and conduction bands is obtained by integrating Eq. (3.2.21) over all states, and is given by

$$\alpha(E) = \frac{q^2 h}{2\epsilon_0 m_0^2 c \bar{\mu} E} \int_{-\infty}^{\infty} \rho_c(E')\rho_v(E'')|M_{if}|^2[1 - f_v(E'') - f_c(E')]\, dE' \quad (3.2.41)$$

where $E'' = E' - E$ and ρ_c and ρ_v are, for the conduction and valence bands respectively, the densities of states per unit volume per unit energy. The integral is evaluated numerically and summed for both light-hole and heavy-hole bands. The matrix element M_{if} can be expressed as a product of two terms, $M_{if} = M_b M_{env}$, where M_b is the previously defined contribution [Eq. (3.2.35)] from the band-edge Bloch functions and M_{env} is the matrix element of the envelope wave functions representing the effect of band-tail states. The envelope wave function is a plane wave above the band edge and takes the form of the ground state of a hydrogen atom for impurity states in the band tail. The calculation described here uses the envelope matrix element of Stern.[14,15] The square of the Bloch-function matrix element $|M_b|^2$ is given by Eq. (3.2.35) for the Kane model. Correction to the value of $|M_b|^2$ can arise from the contribution of other conduction bands.[22,23] These corrections would simply scale the gain values calculated using Eq. (3.2.41).

Figure 3.4 shows the calculated absorption curves for p-type InGaAsP ($\lambda = 1.3$ μm) with different doping levels. The parameters used in the calculation are listed in Table 3.1. The heavy-hole mass, the light-hole mass, the dielectric constant, and the spin-orbit splitting are obtained from a linear extrapolation of the measured binary values. The conduction-band effective

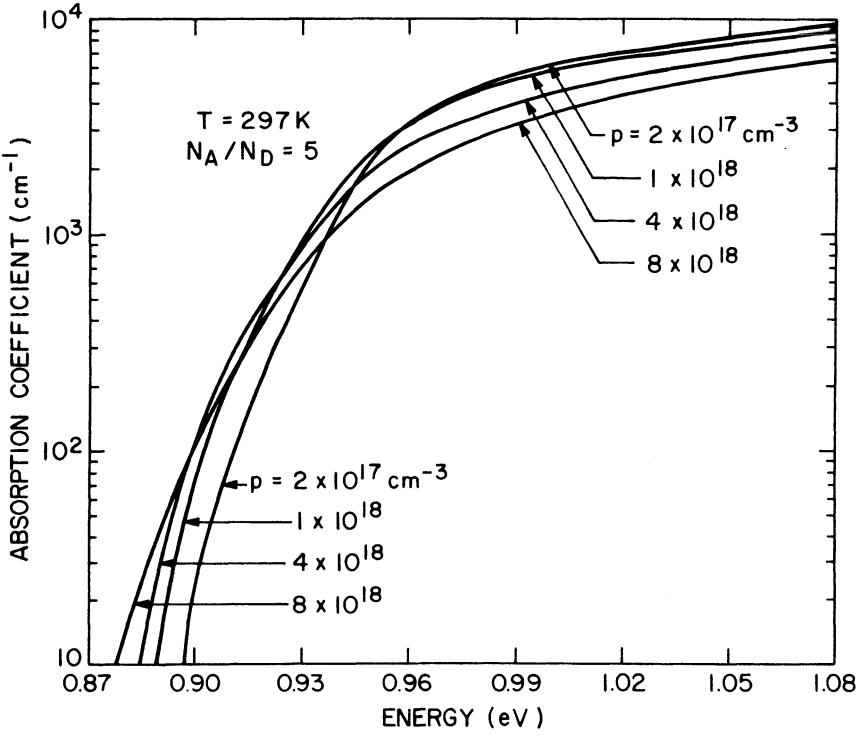


Fig. 3.4 Calculated absorption curves for p-type InGaAsP ($\lambda = 1.3$ μm) for different majority carrier concentrations. The majority carrier density $p = N_A - N_D$. (After Ref. 18)

Table 3.1 Band-Structure Parameters of $In_{1-x}Ga_xAs_yP_{1-y}$ Lattice-Matched to InP [$x = 0.4526y/(1 - 0.031y)$]

PARAMETER	DEPENDENCE ON THE MOLE FRACTIONS x AND y
Energy gap at zero doping	E_g (in eV) $= 1.35 - 0.72y + 0.12y^2$
Heavy-hole mass	$m_{hh}/m_0 = (1 - y)[0.79x + 0.45(1 - x)] + y[0.45x + 0.4(1 - x)]$
Light-hole mass	$m_{lh}/m_0 = (1 - y)[0.14x + 0.12(1 - x)] + y[0.082x + 0.026(1 - x)]$
Dielectric constant	$\epsilon = (1 - y)[8.4x + 9.6(1 - x)] + y[13.1x + 12.2(1 - x)]$
Spin-orbit splitting	Δ (in eV) $= 0.11 + 0.31y - 0.09y^2$
Conduction-band mass	$m_c/m_0 = 0.080 - 0.039y$

mass is obtained from the data in Ref. 24. Since we are dealing with a composition that is lattice-matched to InP, the energy gap E_{g0} (for zero doping) is obtained from the data in Ref. 25. The reported experimental results are in good agreement with the extrapolated light-hole mass.[26] The carrier-concentration-dependent band-gap reduction is assumed to be the same as for GaAs,[16]

$$\Delta E_g \text{ (in eV)} = (-1.6 \times 10^{-8})(p^{1/3} + n^{1/3}) \tag{3.2.42}$$

where n and p are in units of cm^{-3}. The effect of ΔE_g is only to shift the position of the absorption edge or the gain peak, and the magnitude of the maximum gain at a given current density remains unchanged.

The results shown in Fig. 3.4 assume that acceptors are fully ionized, i.e. $N_A - N_D = p$ where N_A and N_D are the acceptor and the donor concentrations, respectively and p is the majority carrier concentration. The ratio $N_A/N_D = 5$ was chosen on the basis of the experimental studies on Zn-doped GaAs. The calculation shows that with an increase in the doping level, the effect of band-tail states becomes more pronounced and the absorption edge is less steep. For the same reason, the absorption on the high-energy side is reduced with an increase in the doping level.

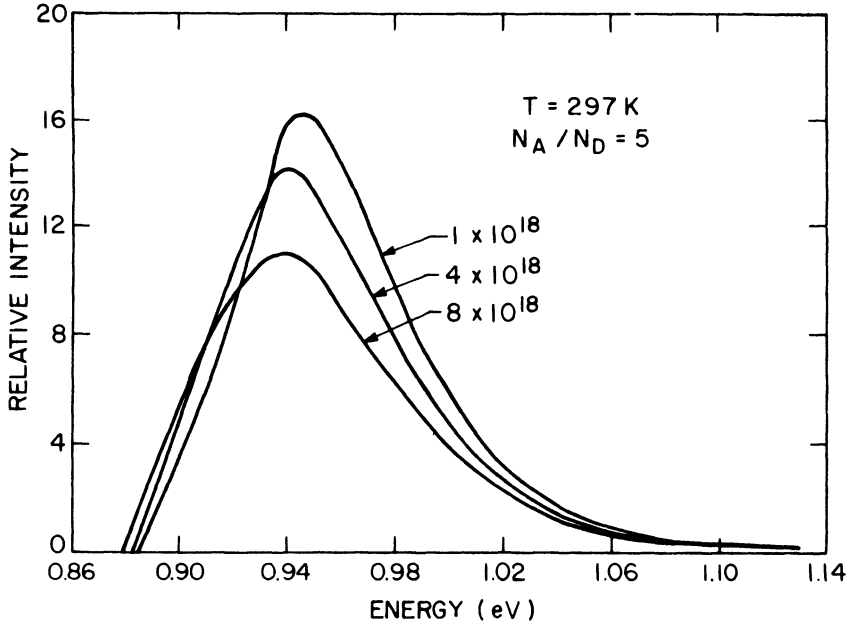


Fig. 3.5 Spontaneous emission spectra for p-type InGaAsP ($\lambda = 1.3$ μm). The parameters used are the same as those in Fig. 3.4. (After Ref. 18)

The total spontaneous-emission rate per unit volume can be calculated in a similar manner. Integrating Eq. (3.2.24) over all states, it is given by

$$R(E) = \frac{4\pi q^2 \bar{\mu} E}{m_0^2 \epsilon_0 c^3 h^2} \int_{-\infty}^{\infty} \rho_c(E')\rho_v(E'')f_c(E')f_v(E'')|M_{if}|^2 \, dE' \quad (3.2.43)$$

where $E'' = E' - E$. The integral is evaluated numerically. Equation (3.2.43) yields the spontaneous-emission spectrum if R is calculated as a function of the photon energy E.

Figure 3.5 shows the calculated spontaneous-emission spectra for different hole concentrations. Note that with increasing carrier concentration the peak of the spontaneous emission shifts to lower energies while its height (maximum emission intensity) decreases. Also note that the width of the emission spectrum increases with increasing carrier concentration as a result of the band-tail states.

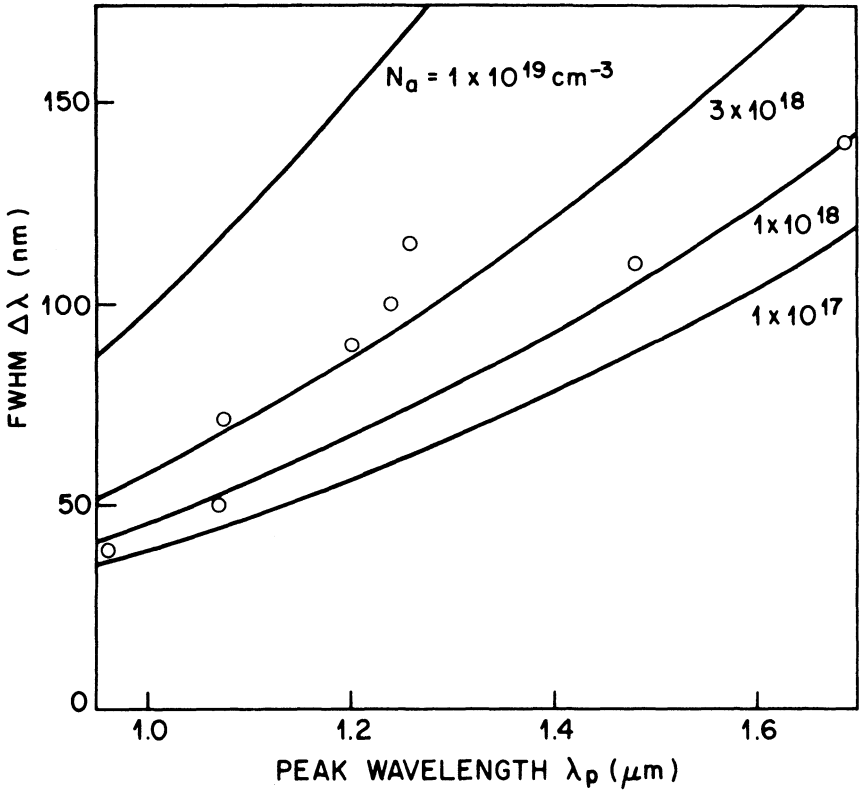


Fig. 3.6 Effect of acceptor concentration on the width of the spontaneous-emission spectrum for p-type InGaAsP (λ = 1.3 μm). Circles are experimental data points. (After Ref. 20)

Figure 3.6 shows the effect of acceptor concentration on the width (FWHM) of the spontaneous emission spectrum. The measured data of Ref. 20 are also shown. The absorption edge shifts to lower energies due to the formation of band-tail states with increasing impurity concentration. The emission and absorption spectra of *n*-type semiconductors have essentially the same features as those of *p*-type semiconductors.

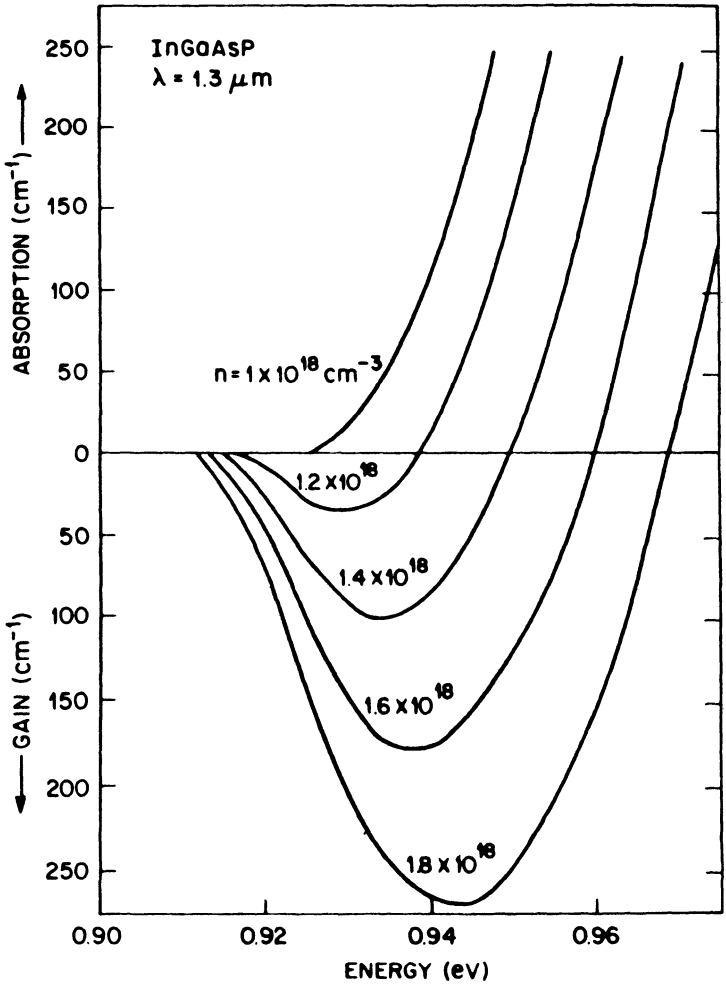


Fig. 3.7 Calculated gain spectra at different injected carrier densities for undoped InGaAsP ($E_g = 0.96$ eV, $\lambda = 1.3$ μm). The absorption coefficient is plotted as a function of the photon energy. (After Ref. 18)

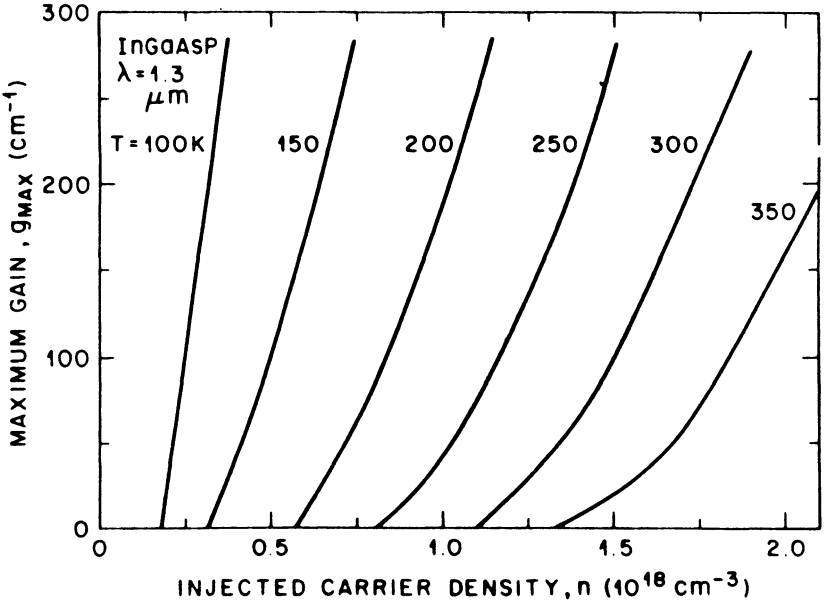


Fig. 3.8 Calculated maximum gain (g_{max}) as a function of injected carrier density (n) in undoped InGaAsP ($\lambda = 1.3$ μm) at various temperatures. (After Ref. 27)

We now calculate the gain in InGaAsP excited by external pumping. Under sufficient optical excitation or current injection, the number of electrons and holes may be large enough to satisfy Eq. (3.2.8); the semiconductor material would then exhibit optical gain. The rate of excitation can be expressed in terms of the injected carrier density or by the nominal current density defined below by Eq. (3.2.44). Figure 3.7 shows the gain spectra at various injected carrier densities. For injected carrier densities less than 10^{18} cm^{-3}, there is no net gain at any photon energy. As the carrier density is further increased, the absorption coefficient α becomes negative over a limited energy range and gain occurs. The photon energy at which the maximum gain occurs shifts to higher energy with increasing injection while, at the same time, the gain appears at a lower energy with increasing injection. This effect is due to a shift in the band edge given by Eq. (3.2.42). Figure 3.8 shows the maximum gain (peak value in Fig. 3.7) plotted as a function of the injected carrier density at different temperatures.[27] The calculation is done for undoped material with a residual concentration of 2×10^{17} cm^{-3} acceptors and donors. Note that the gain is higher at a lower temperature for the same excitation. The temperature plays a role through the Fermi factors and affects the distribution of electrons and holes in the conduction and valence bands. At higher

temperatures the carriers are distributed over a wider energy range, so the maximum gain is less.

A convenient way to express the excitation rate is to use the nominal current density J_n, which is defined as the total spontaneous-emission rate per unit volume and equals the injected current density at unit quantum efficiency. The nominal current density is usually expressed in A/(cm^2·μm) and is given by

$$J_n = qR \tag{3.2.44}$$

where R is the total spontaneous-emission rate per unit volume given by Eq. (3.2.43). Figure 3.9 shows the calculated maximum gain as a function of nominal current density for InGaAsP at 1.3 μm. The nominal current density equals the current lost to radiative recombination in a double-heterostructure laser. Figure 3.9 shows that to achieve the same gain, higher current density is needed at higher temperatures. Figures 3.8 and 3.9 show the basis of the increase in threshold current of injection lasers as the temperature is increased. Section 3.6 discusses the temperature dependence in detail.

Results of a similar calculation for the variation of the maximum gain with the injected carrier density and the nominal current density at 1.55 μm

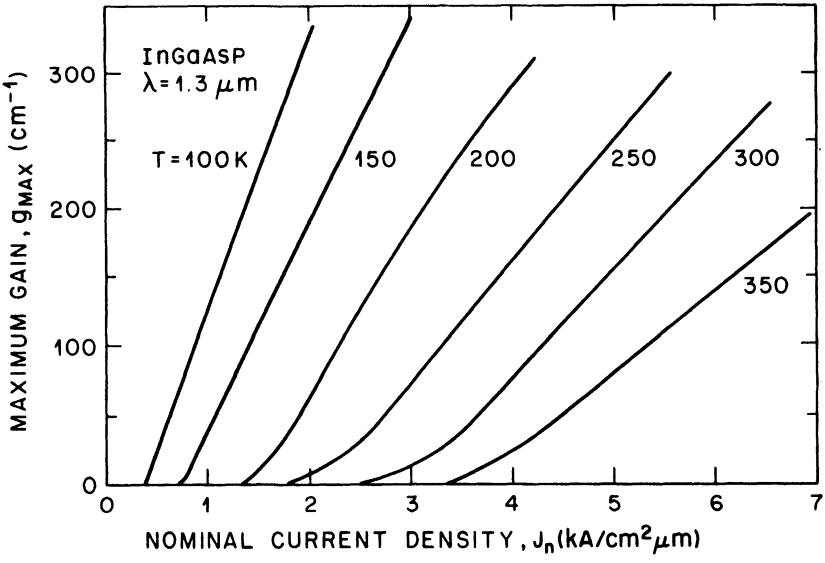


Fig. 3.9 Calculated maximum gain (g_{max}) as a function of nominal current density (J_{nom}) for undoped InGaAsP (λ = 1.3 μm) at various temperatures. (After Ref. 27)

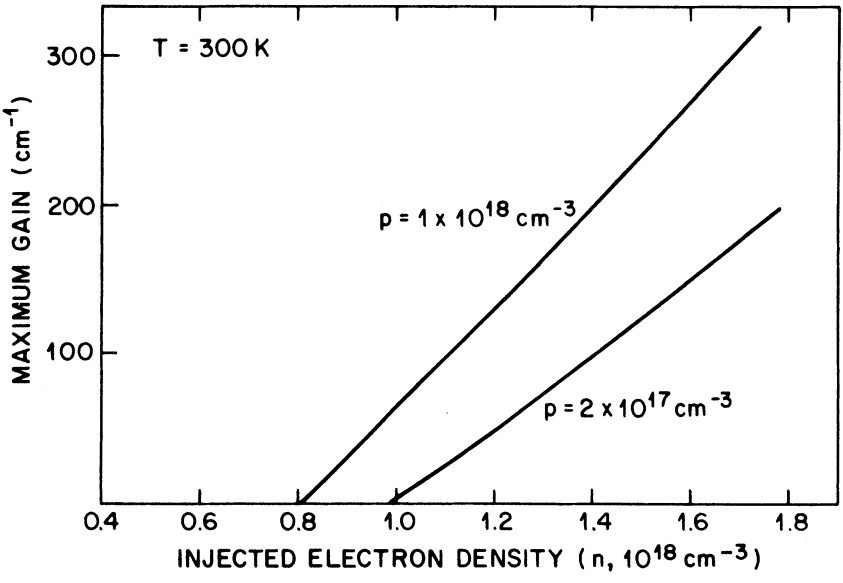


Fig. 3.10 Calculated maximum gain as a function of injected carrier density in p-type InGaAsP for two different majority carrier (hole) concentrations.

are given in Ref. 27. The results are very similar to those of Figs. 3.8 and 3.9 except that a smaller injected carrier density and a smaller nominal current density are needed at the longer wavelength to achieve the same gain. This arises from the smaller conduction-band effective electron mass at the longer wavelength, which allows the gain condition (3.2.8) to be satisfied at a lower carrier density.

We next consider the effect of doping on the optical gain. Figure 3.10 shows the calculated maximum gain as a function of the injected electron density in p-type InGaAsP ($\lambda = 1.3$ μm) with different hole concentrations. The calculation assumes $N_A/N_D = 5$. Figure 3.11 shows the result for n-type InGaAsP as a function of the injected hole density. The two figures show that the excitation rate required for transparency ($g = 0$) decreases with higher doping both for p-and n-type materials. The decrease is more rapid for n-type than for p-type materials. The difference in behavior between the n- and p-type materials is due to the small effective mass of electrons compared with that of holes. This may be seen as follows. The condition for net stimulated emission or gain is given by Eq. (3.2.8). Increasing the doping level increases both E_{fc} and E_{fv}; hence fewer additional injected carriers are needed to satisfy the condition for gain. Since the effective mass

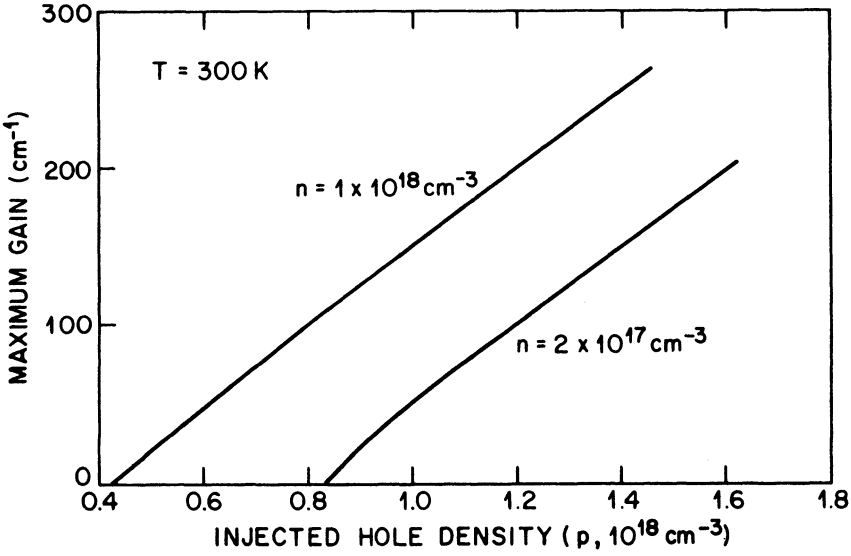


Fig. 3.11 Calculated maximum gain as a function of injected hole density in n-type InGaAsP for two different majority carrier (electron) concentrations.

of the conduction band is smaller than that of the valence band,

$$\left|\frac{\partial E_{fc}}{\partial n}\right| > \left|\frac{\partial E_{fv}}{\partial p}\right|.$$

This explains why the effect of doping is more pronounced in n-type than in p-type material.

Figures 3.8–3.11 show that the optical gain varies almost linearly with the injected carrier density and can be approximately written as

$$g = a(n - n_0) \tag{3.2.45}$$

where the slope a is called the gain coefficient and n_0 is the injected carrier density required to achieve transparency (i.e., $g = 0$ when $n = n_0$). From Eq. (3.2.8), the condition for transparency is given by

$$E_{fc} + E_{fv} = 0. \tag{3.2.46}$$

The quantity n_0 can be calculated from Eq. (3.2.27) using the approximation

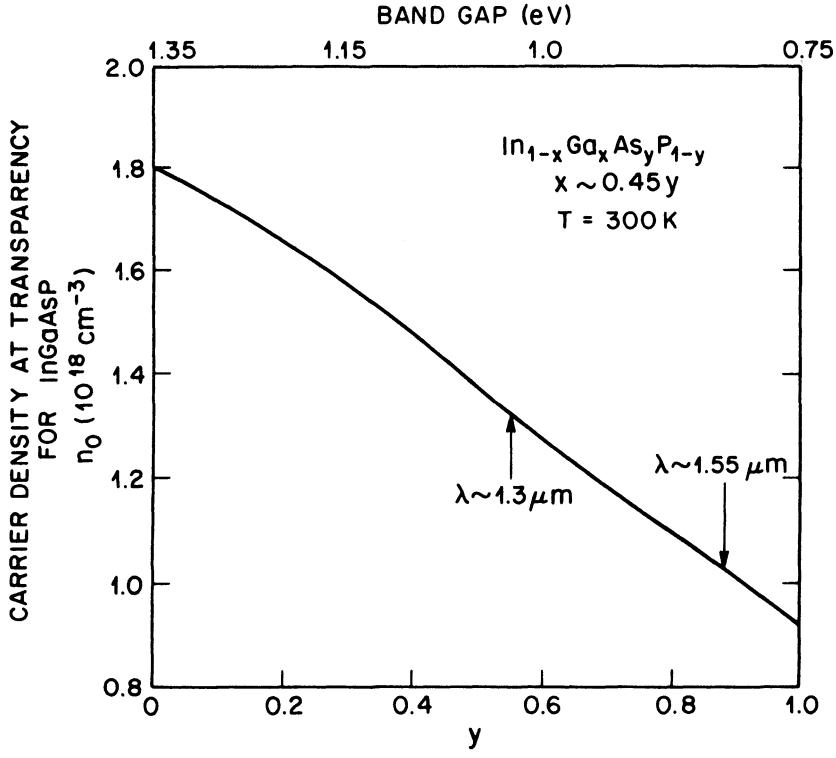


Fig. 3.12 Injected carrier density n_0 needed for transparency ($g = 0$) for InGaAsP lattice-matched to InP. Arrows indicate values for 1.3- and 1.55-μm InGaAsP lasers.

(3.2.28) for parabolic bands. Figure 3.12 shows the calculated n_0 for undoped InGaAsP lattice-matched to InP. The smaller n_0 for low band gap is due to a smaller effective mass.

The variation of the gain with the carrier density is approximately linear. As such, it is useful for a simple but fairly accurate description of semiconductor lasers as was presented in Chapter 2. Equation (3.2.45) was used in the phenomenological model introduced in Sec. 2.4. The gain parameters a and n_0 can be determined either using numerical results presented in Figs. 3.8–3.12 or deduced from the experimental data. Typical values for InGaAsP lasers are in the range of $1.2–2.5 \times 10^{-16}$ cm^2 for a and $0.9–1.5 \times 10^{18}$ cm^{-3} for n_0 depending on the laser wavelength and doping levels. It should be stressed that because of uncertainties in the band-structure parameters, the numerical results are only accurate to within a factor of 2.

The total spontaneous emission rate R given by Eq. (3.2.43) can be

approximated by

$$R = Bnp \tag{3.2.47}$$

where B is the radiative recombination coefficient and n and p are the electron density and the hole density respectively. For undoped semiconductors, $R = Bn^2$. The calculated $B = 1.2 \times 10^{-10}$ cm^3/s for InGaAsP ($\lambda = 1.3$ μm) at a carrier concentration of 1×10^{18} cm^{-3}. Calculation of the radiative recombination rate shows that B decreases with increasing carrier density. This has been confirmed by Olshansky et al.[28] using carrier-lifetime measurements.

3.3 NONRADIATIVE RECOMBINATION

As mentioned in Sec. 3.1, electrons and holes in a semiconductor can also recombine nonradiatively. Nonradiative mechanisms include recombination at defects, surface recombination, and Auger recombination, among others. For long-wavelength semiconductor lasers, however, the Auger process is generally the predominant nonradiative mechanism, and most of this section is devoted to its study. The surface- and defect-recombination processes are discussed briefly in Sec. 3.3.4.

The Auger recombination process involves four particle states (three electrons and one hole, two electrons and two holes, and so forth). In this process, the energy released during the electron-hole recombination is transferred to another electron (or hole), which gets excited to a high energy state in the band. This electron or hole then relaxes back to achieve thermal equilibrium by losing its energy to lattice vibrations or phonons. The Auger recombination rate R_a may be approximately written as

$$R_a = Cn^3 \tag{3.3.1}$$

where n is the injected carrier density in an undoped semiconductor. It is useful to define the carrier lifetime τ_A for the Auger process as

$$\tau_A = n/R_a = (Cn^2)^{-1}. \tag{3.3.2}$$

The quantity C is called the Auger coefficient. The inverse-square dependence of carrier lifetime on carrier density is used to identify the Auger effect experimentally. Departures from the simple inverse-square relation are discussed later in the section.

It is generally accepted that Auger recombination can be a major nonradiative mechanism in narrow-gap semiconductors.[29–33] The observed strong temperature dependence of threshold current in long-wavelength

semiconductor lasers has focused attention on the Auger effect. There are several different types of Auger recombination processes. The three major types are (i) band-to-band processes,[27,32–37] (ii) phonon-assisted Auger processes[38–41], and (iii) trap-assisted processes.[42–45]

3.3.1 Band-to-Band Auger Processes

The band-to-band Auger processes in a direct-gap semiconductor are shown in Fig. 3.13. The three processes are labeled CCCH, CHHS, and CHHL where C stands for the conduction band and H, L, and S stand for heavy-hole, light-hole, and split-off valence bands, respectively. The CCCH process involves three electrons and one hole and is dominant in *n*-type material.

The Auger recombination rate for this process was first derived in Ref. 29. The basis of the Auger effect is the Coulomb interaction between the two electrons 1 and 2 in the conduction band. In the CCCH mechanism, electron 1 recombines with hole 1′ (i.e., electron 1 makes a transition to empty state 1′) and the excess energy is transferred to electron 2, which is excited to state 2′. Electron 2′ loses its energy to optical phonons when it relaxes back to thermal equilibrium. The CHHS process involves one electron, two heavy holes, and a split-off-band hole. CHHL is similar to CHHS except that it involves a light hole. CHHS and CHHL are dominant in *p*-type

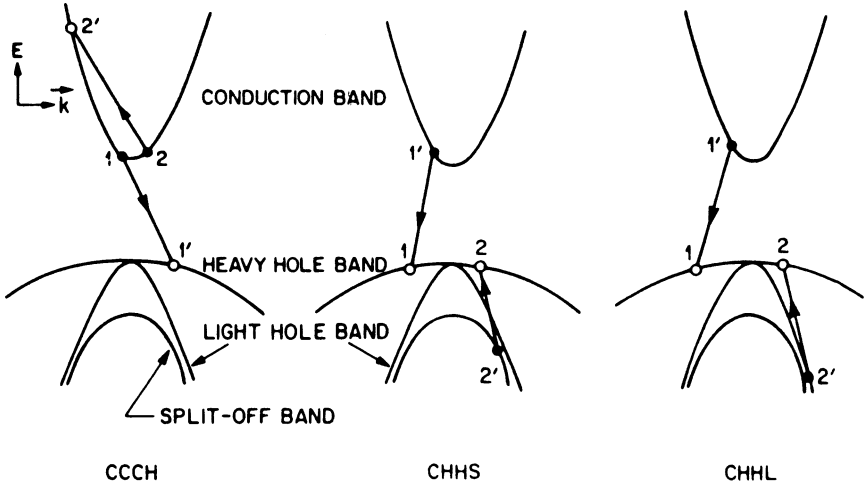


Fig. 3.13 Three different band-to-band Auger processes shown schematically. Electrons are represented by closed circles and holes by open circles.

semiconductors. Under the high-injection conditions commonly present in semiconductor lasers, all three mechanisms must be considered.

Band-to-band Auger processes are characterized by a strong temperature dependence and band-gap dependence, i.e., the Auger rate decreases rapidly either for low temperatures or for high band-gap materials. These dependencies arise from the laws of energy and momentum conservation that the four free particle states involved (1, 2, 1′, 2′ in Fig. 3.13) must satisfy. This may be seen in the following way for CCCH. The momentum and energy conservation laws give rise to a threshold energy E_T for each process. For CCCH, if we assume $E_1 \cong E_2 = 0$, only holes with energies greater than $\sim(E_T - E_g) = \delta E_g$ can participate (δ is a constant that depends on effective masses). The number of such holes varies approximately as $\exp(-\delta E_g/k_B T)$ for nondegenerate semiconductors, and thus the Auger rate varies as $\exp(-\delta E_g/k_B T)$ in the nondegenerate case. The strong temperature dependence does not appear if the individual particle states are not states of definite momentum, such as trap states, or if conservation of momentum is satisfied through phonon assistance. These processes are discussed later in this section.

We now calculate the recombination rates of the various band-to-band Auger processes. Most of the notation and methods of calculation are from Ref. 27.

CCCH Auger Process. Beattie and Landsberg[29] derived an expression for the CCCH Auger process using Bloch functions for particle states. Using Fermi's golden rule, the expression for the Auger rate is

$$R_a = \frac{2\pi}{\hbar}\left(\frac{1}{(2\pi)^3}\right)^4 \iiiint |M_{\text{if}}|^2 P(1, 1', 2, 2') \times \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_1' - \mathbf{k}_2')\delta(E_i - E_f)\, d^3\mathbf{k}_1\, d^3\mathbf{k}_2\, d^3\mathbf{k}_1'\, d^3\mathbf{k}_2' \qquad (3.3.3)$$

where R_a denotes the Auger rate per unit volume. $P(1, 1', 2, 2')$ denotes the net probability for the CCCH process. The momentum conservation (represented by the first δ function) arises from the representation of the particle states as the product of a plane wave and a Bloch function. M_{if} is the matrix element of the Coulomb interaction potential V between two electrons, i.e.,

$$V(\mathbf{r}_1 - \mathbf{r}_2) = \frac{q^2}{\epsilon|\mathbf{r}_1 - \mathbf{r}_2|}\exp[-\lambda|\mathbf{r}_1 - \mathbf{r}_2|] \qquad (3.3.4)$$

where λ is the inverse of the screening length, ϵ is the dielectric constant, and $\mathbf{r}_1$ and $\mathbf{r}_2$ are the positions of the interacting electrons.

The matrix element after summing over the singlet and triplet state of electrons 1 and 2 takes the following form:

$$M_{if} = \iint [\phi_1^*(\mathbf{r}_1)\phi_2^*(\mathbf{r}_2)\,\Delta_{21} - \phi_2^*(\mathbf{r}_1)\phi_1^*(\mathbf{r}_2)\,\Delta_{12}] \times V(\mathbf{r}_1 - \mathbf{r}_2)\phi_1'(\mathbf{r}_1)\phi_2'(\mathbf{r}_2)\,d^3\mathbf{r}_1\,d^3\mathbf{r}_2 \tag{3.3.5}$$

where $\phi_1(\mathbf{r}_1)$ and $\phi_2(\mathbf{r}_2)$ denote the initial state of electrons 1 and 2 while $\phi_1'(\mathbf{r}_1)$, $\phi_1'(\mathbf{r}_2)$ denote the corresponding final states. The quantities $\Delta_{21} = \Delta_{12} = 1$ when electrons 1 and 2 have like spins, $\Delta_{21} = 1$, $\Delta_{12} = 0$ when they have opposite spins preserved by the transition, and $\Delta_{12} = 1$, $\Delta_{21} = 0$ when they have opposite spins both changed by the transition. The wave functions $\phi_1(\mathbf{r}_1)$ can be written as

$$\phi_1(\mathbf{r}_1) = \psi_1(\mathbf{k}_1, \mathbf{r}_1)\exp(i\mathbf{k}_1\cdot\mathbf{r}_1) \tag{3.3.6}$$

where $\psi_1(\mathbf{k}_1, \mathbf{r}_1)$ denotes the localized part of the Bloch function and $\exp(i\mathbf{k}_1\cdot\mathbf{r}_1)$ is the plane-wave part. Consider one term in the matrix element and carry out the integration after using Eqs. (3.3.4) and (3.3.6). The result is

$$\iint \phi_1^*(\mathbf{r}_1)\phi_2^*(\mathbf{r}_2)V(\mathbf{r}_1 - \mathbf{r}_2)\phi_1'(\mathbf{r}_1)\phi_1'(\mathbf{r}_2)\,d^3\mathbf{r}_1\,d^3\mathbf{r}_2 = \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_1' - \mathbf{k}_2')\frac{4\pi q^2}{\epsilon(\lambda^2 + g^2)}F(1, 1')F(2, 2') \tag{3.3.7}$$

where $g = |\mathbf{k}_1 - \mathbf{k}_1'|$ and $F(1, 1')$ and $F(2, 2')$ are the overlap integrals defined using

$$F(1, 1') = \int \psi_1^*(\mathbf{r}_1)\psi_1'(\mathbf{r}_1)\,d^3\mathbf{r}_1$$

$$F(2, 2') = \int \psi_2^*(\mathbf{r}_2)\psi_2'(\mathbf{r}_2)\,d^3\mathbf{r}_2.$$

In deriving the above, we have used the relation $\mathbf{r}_{12} = \mathbf{r}_1 - \mathbf{r}_2$ and the integral

$$\int \frac{d^3\mathbf{r}_{12}}{r_{12}}\exp(-\lambda r_{12} + i\mathbf{g}\cdot\mathbf{r}_{12}) = \frac{4\pi}{\lambda^2 + g^2}. \tag{3.3.8}$$

Using a similar procedure for the other terms in Eq. (3.3.5), the matrix

element $|M_{if}|^2$ is given by[29]

$$|M_{if}|^2 = \left(\frac{4\pi q^2}{\epsilon}\right)^2 \left[\left| \frac{F(1,1')F(2,2')}{\lambda^2 + g^2} \right|^2 + \left| \frac{F(1,2')F(2,1')}{\lambda^2 + h^2} \right|^2 + \left| \frac{F(1,1')F(2,2')}{\lambda^2 + g^2} - \frac{F(1,2')F(2,1')}{\lambda^2 + h^2} \right|^2 \right] \tag{3.3.9}$$

where $g^2 = |\mathbf{k}_1 - \mathbf{k}_1'|^2 = |\mathbf{k}_2' - \mathbf{k}_2|^2$ and $h^2 = |\mathbf{k}_1 - \mathbf{k}_2'|^2 = |\mathbf{k}_1' - \mathbf{k}_2|^2$. $F(i, j)$'s denote the overlap integral of the Bloch functions. These overlap integrals have been evaluated using the Kane model for the band structure.[34,46] The result is that

$$|F(i, j)|^2 \cong 1 \tag{3.3.10a}$$

when i and j are in the same band, and

$$|F(i, j)|^2 \cong \frac{\alpha_{ij}}{E_g} |\mathbf{k}_i - \mathbf{k}_j|^2 \tag{3.3.10b}$$

when i and j are in different bands. Here α_{ij} is a constant. Calculations of the overlap integrals using a more detailed band structure have been carried out.[47,48] Since the matrix element given by Eq. (3.3.9) is fairly complicated, some approximate form is generally employed. An approximation to the matrix element $|M_{if}|^2$ obtained using Eq. (3.3.10) is ($0.5|M_{if}'|^2 \leq |M_{if}|^2 \leq 1.5|M_{if}'|^2$):

$$|M_{if}|^2 \cong |M_{if}'|^2 = 4\left(\frac{4\pi q^2}{\epsilon}\right)^2 \frac{\alpha_{CH}}{E_g} \frac{k_2'^2}{|k_2'^2 + \lambda^2|^2} = 4|\bar{M}_{if}|^2. \tag{3.3.11}$$

Here $\alpha_{CH} = \alpha_{11}'$, and we have used the approximation $k_2'^2 \gg k_1^2, k_2^2$ so that $g^2 = h^2 = k_2'^2$. Equation (3.3.11) defines the matrix element $|M_{if}|^2$, and the expression (3.3.3) for the Auger rate can be written as

$$R_a = \left(\frac{8\pi}{\hbar}\right)\left(\frac{1}{8\pi^3}\right)^4 \iiiint |\bar{M}_{if}|^2 P(1, 1', 2, 2') \times \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_1' - \mathbf{k}_2')\delta(E_i - E_f)\, d^3\mathbf{k}_1\, d^3\mathbf{k}_2\, d^3\mathbf{k}_1'\, d^3\mathbf{k}_2'. \tag{3.3.12}$$

The quantity $P(1, 1', 2, 2')$ in Eq. (3.3.12) is the difference between the probabilities of the Auger recombination process of an electron-hole pair shown in Fig. 3.13 and the inverse process of creating a pair. For the CCCH process to occur, there must be electrons in states $|1\rangle$, $|2\rangle$, a hole in state $|1'\rangle$, and no electrons in state $|2'\rangle$. The appropriate occupation probabilities

are given by the Fermi factors (3.2.2) and (3.2.3). Using them, we obtain

$$P(1, 1', 2, 2') = f_c(1)f_c(2)f_v(1')[1 - f_c(2')] - [1 - f_c(1)][1 - f_c(2)][1 - f_v(1')]f_c(2') \quad (3.3.13)$$

where $f_c(i)$ stands for $f_c(E_i)$, E_i being the energy of ith particle. The second term in Eq. (3.3.13) is almost negligible since $f_c(2') \cong 0$ in most cases of practical interest where $E_g \gg k_B T$. To a good degree of approximation, we can therefore use

$$P(1, 1', 2, 2') \cong f_c(1)f_c(2)f_v(1')[1 - f_c(2')]. \quad (3.3.14)$$

The energies of the four particles involved in the Auger process satisfy the energy conservation relation

$$E_1 + E_2 = E_2' - E_g - E_1'. \quad (3.3.15)$$

Note that the hole energies are measured positive downward from the top of the valence band.

For the nondegenerate case, the probability factor $P(1, 1', 2, 2')$ can be simplified using Eqs. (3.2.30) and (3.2.33). The result is

$$P(1, 1', 2, 2') = \frac{n^2 p}{N_c^2 N_v} \exp\left(\frac{-E_2' + E_g}{k_B T}\right) \quad (3.3.16)$$

where we have used Eq. (3.3.15) and assumed that $f_c(2') \ll 1$. The total Auger recombination rate is now given by

$$R_a = \frac{8\pi}{\hbar}\left(\frac{1}{8\pi^3}\right)^4 \xi_1 \iiiint \exp\left(\frac{-E_2'}{k_B T}\right)|\bar{M}_{if}(k_2')|^2 \times \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_1' - \mathbf{k}_2')\delta(E_i - E_f)\, d^3\mathbf{k}_1\, d^3\mathbf{k}_2\, d^3\mathbf{k}_1'\, d^3\mathbf{k}_2' \quad (3.3.17)$$

where we have assumed that $\bar{M}_{if}$ depends only on k_2' as in Eq. (3.3.11) and ξ_1 is given by

$$\xi_1 = \frac{n^2 p}{N_c^2 N_v} \exp\left(\frac{E_g}{k_B T}\right). \quad (3.3.18)$$

The integrals can be considerably simplified if we define the new variables

$$\mathbf{h} = \mathbf{k}_1 + \mathbf{k}_2 \qquad \mathbf{j} = \mathbf{k}_1 - \mathbf{k}_2.$$

Integrating over them, we obtain

$$R_a = \frac{m_{c0}\xi_1}{64\pi^7\hbar^3}\iint \delta^3\mathbf{k}_1'\, d^3\mathbf{k}_2' |j_0|\,|\bar{M}_{if}|^2 \exp\left(\frac{-E_2'}{k_B T}\right) \tag{3.3.19}$$

where

$$j_0^2 = 2\left[\mu_c k_2'^2 - \mu k_1'^2 - \frac{2m_{c0}E_g}{\hbar^2}\right] - |\mathbf{k}_1' + \mathbf{k}_2'|^2$$

$$\mu_c = \frac{m_{c0}}{m_c} \qquad \mu = \frac{m_{c0}}{m_v}.$$

Here m_{c0} and m_c are the conduction-band effective masses at the band edge and at an energy E_2', respectively. Defining the new variables

$$\mathbf{Z}_1 = \mathbf{k}_1' + \frac{\mathbf{k}_2'}{1+2\mu} \qquad \mathbf{Z}_2 = \mathbf{k}_2' \tag{3.3.20}$$

we get the following result:

$$R_a = \frac{m_{c0}\xi_1}{4\pi^5\hbar^3} I \tag{3.3.21}$$

where

$$I = \iint dZ_1\, dZ_2\, Z_1^2 Z_2^2 |J_0|\,|\bar{M}_{if}|^2 \exp\left(\frac{-E_2'}{k_B T}\right) \tag{3.3.22}$$

and

$$J_0^2 = 2\alpha_c Z_2^2 - Z_1^2(1+2\mu) - \frac{4m_{c0}E_g}{\hbar^2} \tag{3.3.23}$$

where

$$\alpha_c = \mu_c - \frac{\mu}{1+2\mu}.$$

Before simplifying the integral I further, we note the following. Since I should be real, J_0^2 must be a positive quantity. Since $E_2' = \hbar^2 Z_2^2/2m_c$, Eq. (3.3.23) for J_0^2 shows that there is a minimum value of E_2' for the Auger process to take place. A lower bound on this minimum value is obtained

using

$$2\alpha_c Z_2^2 > \frac{4m_{c0}E_g}{\hbar^2} \qquad \text{or} \qquad E_2' > \frac{\mu_c E_g}{\alpha_c}.$$

Thus the threshold energy, or the minimum energy for the electron 2′, is given by

$$E_T = \frac{\mu_c E_g}{\alpha_c} = \frac{2m_{c0} + m_v}{2m_{c0} + m_v - m_c} E_g. \tag{3.3.24}$$

Most of the contribution to the integral I arises from energies $E_2' \cong E_T$ because of the exponential term $\exp(-E_2'/k_BT)$ in Eq. (3.3.22). Hence we can approximate $|\bar{M}_{if}(k_2')|^2$ by its value $|M_a|^2$ at energy E_T, i.e.,

$$|\bar{M}_{if}(k_2')|^2 \cong |M_a|^2. \tag{3.3.25}$$

Integration over Z_1 can be readily performed and Eq. (3.3.22) becomes

$$I = \frac{\pi\alpha_c^2|M_a|^2}{(1+2\mu)^{3/2}} \int_{Z_{2t}}^{\infty} dZ_2\,(Z_2^2 - Z_{2t}^2)^2 Z_2^2 \exp\left(\frac{-E_2'}{k_BT}\right)$$

where $Z_{2t} = (2m_c/\hbar^2)E_T$. Defining $y = Z_2 - Z_{2t}$ as the new integration variable, we obtain

$$I = \frac{\pi\alpha_c^2|M_a|^2}{(1+2\mu)^{3/2}} \exp\left(\frac{-E_T}{k_BT}\right) I_0 \tag{3.3.26a}$$

where

$$I_0 = \int_0^{\infty} dy\, y^2(Z_{2t}+y)^2(2Z_{2t}+y)^2 \exp\left(-\frac{\hbar^2 y}{2m_c k_B T}(2Z_{2t}+y)\right) \tag{3.3.26b}$$

and can be analytically evaluated. Using Eqs. (3.3.21) and (3.3.26), the expression for the Auger rate in the nondegenerate approximation becomes

$$R_a = \frac{m_{c0}\alpha_c^2|M_a|^2}{4\pi^4\hbar^3(1+2\mu)^{3/2}} \frac{n^2 p I_0}{N_c^2 N_v} \exp\left(\frac{-\Delta E}{k_BT}\right) \tag{3.3.27}$$

where $\Delta E = E_T - E_g$. A similar expression was initially derived in Ref. 29 for the case $m_c = m_{c0}$. The main points to note are that the Auger rate varies at n^2p and is temperature-dependent through the exponential term. In particular, the Auger rate increases with increasing temperature in the

nondegenerate approximation. For the degenerate case, the Auger-rate calculations can be done using the same procedure as above except that the integral I in Eq. (3.3.22) is modified.[27]

CHHS Auger Process. The Auger recombination rate for CHHS, which involves two heavy holes, one electron and one split-off-band hole, can be calculated using the same procedure as for CCCH. The expression for the Auger rate is given by Eq. (3.3.12) with the matrix element[34] (obtained using the Kane model):

$$|\bar{M}_{\text{if}}|^2 \cong \left(\frac{4\pi q^2}{\epsilon}\right)^2 \frac{\alpha_{\text{CH}}}{E_g} \frac{\alpha_{\text{SH}}}{E_g} \frac{k_2'^4}{(k_2'^2 + \lambda^2)^2} \tag{3.3.28}$$

where we have made the approximation that $k_2' \gg k_1, k_2$.

In the nondegenerate approximation, the probability factor $P(1, 1', 2, 2')$ can be simplified using Eqs. (3.2.30) and (3.2.33) to obtain

$$P(1, 1', 2, 2') = \frac{np^2}{N_c N_v^2} \exp\left(\frac{-E_2' + E_g}{k_B T}\right). \tag{3.3.29}$$

The total Auger recombination rate R_a using Eqs. (3.3.12) and (3.3.29) is given by

$$R_a = \frac{8\pi}{\hbar}\left(\frac{1}{8\pi^3}\right)^4 \xi_s \iiiint \exp\left(\frac{-E_2'}{k_B T}\right)|\bar{M}_{\text{if}}(k_2')|^2 \times \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_1' - \mathbf{k}_2')\delta(E_i - E_f)\, d^3\mathbf{k}_1\, d^3\mathbf{k}_2\, d^3\mathbf{k}_1'\, d^3\mathbf{k}_2' \tag{3.3.30}$$

where ξ_s is defined by

$$\xi_s = \frac{np^2}{N_c N_v^2} \exp\left(\frac{E_g}{k_B T}\right). \tag{3.3.31}$$

We have made the approximation that the matrix element $|\bar{M}_{\text{if}}|^2$ depends on k_2' in accordance with Eq. (3.3.28).

The integrals in Eq. (3.3.30) can be simplified using the same procedure as for the CCCH process. We define the variables $\mathbf{h}$ and $\mathbf{j}$ as before, such that $\mathbf{h} = \mathbf{k}_1 + \mathbf{k}_2$ and $\mathbf{j} = \mathbf{k}_1 - \mathbf{k}_2$, and integrate over them. If we further define new variables

$$\mathbf{z}_1 = \mathbf{k}_1' + \frac{\mathbf{k}_2'}{2/\mu + 1} \qquad \mathbf{z}_2 = \mathbf{k}_2' \tag{3.3.32}$$

we obtain

$$R_a = \frac{m_{c0}\xi_s}{4\pi^5\hbar^3\mu} I \tag{3.3.33}$$

where

$$I = \iint dz_2\, dz_1\, z_1^2 z_2^2 |\bar{M}_{if}|^2 |J_0| \exp\left(\frac{-E_2'}{k_B T}\right) \tag{3.3.34}$$

and

$$J_0^2 = \frac{2}{\mu}\alpha_s z_2^2 - \frac{2}{\mu}(1 + \mu/2) z_1^2 - \frac{4m_{c0}}{\hbar^2\mu}(E_g - \Delta) \tag{3.3.35}$$

$$\alpha_s = \mu_s - 1 + \left(\frac{1}{1 + \mu/2}\right) \qquad \mu_s = \frac{m_{c0}}{m_s} \qquad \mu = \frac{m_{c0}}{m_v}. \tag{3.3.36}$$

For J_0^2 to be positive, there is a minimum value of z_2 (assuming $E_g > \Delta$) for the Auger process to take place. This minimum value is obtained using

$$\alpha_s z_2^2 > 2m_{c0}(E_g - \Delta)/\hbar^2.$$

Since $E_2' = \hbar^2 z_2^2/2m_s$, the above inequality implies that E_2' must be greater than the threshold energy E_T for CHHS given by

$$E_T = \frac{m_{c0}(E_g - \Delta)}{\alpha_s m_s} = \frac{2m_v + m_{c0}}{2m_v + m_{c0} - m_s}(E_g - \Delta). \tag{3.3.37}$$

Making the same approximations as for the CCCH process, viz. replace $|\bar{M}_{if}|^2$ by its value $|M_s|^2$ at $E_2' = E_T$, the integral I may be simplified to

$$I = \frac{\pi\alpha_s^2 |M_s|^2}{(1 + \mu/2)^{3/2}} \left(\frac{2}{\mu}\right)^{1/2} \exp\left(\frac{-E_T}{k_B T}\right) I_0 \tag{3.3.38}$$

where I_0 is given by Eq. (3.3.26b) except that m_c is replaced by m_s. Using Eqs. (3.3.31), (3.3.33), and (3.3.38), the expression for the CHHS Auger rate in the nondegenerate approximation is given by

$$R_a = \frac{m_{c0}\alpha_s^2 |M_s|^2}{4\pi^4\hbar^3\mu(1 + \mu/2)^{3/2}} \left(\frac{2}{\mu}\right)^{1/2} \frac{np^2 I_0}{N_c^2 N_v} \exp\left(-\frac{\Delta E}{k_B T}\right) \tag{3.3.39}$$

where

$$\Delta E = E_T - E_g. \tag{3.3.40}$$

The temperature dependence and the band-gap dependence of the CHHS Auger process arises from the exponential term in Eq. (3.3.39). For the case where $E_g < \Delta$, there is no threshold energy for CHHS. The expression for the CHHS Auger rate under degeneracy has also been derived.[27,37]

CHHL Auger Process. The expression of the Auger rate for CHHL is similar to that for CHHS except that the split-off-band mass m_s is replaced by the light-hole mass m_l and $E_g - \Delta$ is replaced by E_g in Eq. (3.3.37). The threshold energy for CHHL is given by

$$E_T = \frac{2m_v + m_{c0}}{2m_v + m_{c0} - m_l} E_g. \tag{3.3.41}$$

The quantities m_c, m_s, and m_l that appear in Eqs. (3.3.24), (3.3.37), and (3.3.41) are the effective masses in the conduction, split-off, and light-hole bands at the threshold energies of CCCH, CHHS, and CHHL, respectively. Since the band structure away from the band edge is not well known experimentally, simplified theoretical models must be used to calculate these effective masses. As a result, the accuracy of the calculated Auger rate is limited by the accuracy of the band-structure model used.[49,50]

Calculated Auger Lifetime. As stated previously, the Auger effect is mainly responsible for the observed high temperature dependence of the threshold current in long-wavelength semiconductor lasers, in particular, lasers fabricated using the InGaAsP material system and emitting near 1.3 μm and 1.55 μm. We now present the results of calculated Auger lifetimes in InGaAsP lattice-matched to InP under low injection.

The Auger lifetime of injected carriers is obtained using Eq. (3.3.2) after the Auger rate R_a has been calculated for a given carrier density n. For n-type semiconductors, the CCCH process is important. Figure 3.14 shows the calculated Auger lifetime for n-type InGaAsP (curve labeled CCCH), with a majority carrier concentration for $n_0 = 10^{18}\ \text{cm}^{-3}$. This calculation is for low injection ($n = 10^{15}\ \text{cm}^{-3}$). For p-type semiconductors, both CHHS and CHHL may be important. Figure 3.14 shows the calculated result for both processes at low injection ($n = 10^{15}\ \text{cm}^{-3}$) for p-type InGaAsP with a majority carrier concentration of $p_0 = 10^{18}\ \text{cm}^{-3}$. In all cases the Auger lifetime decreases (or the Auger recombination rate increases) rapidly with decreasing band gap.

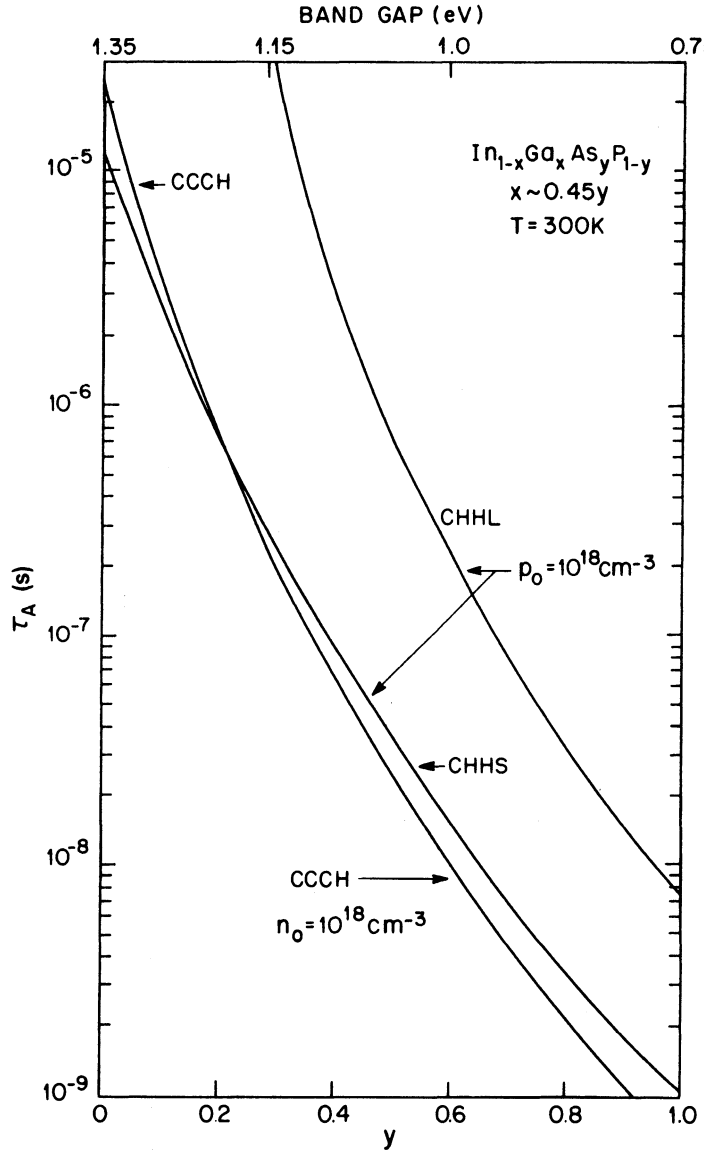


Fig. 3.14 Calculated Auger lifetime τ_A for *n*-type and *p*-type InGaAsP at low injection as a function of band gap (top scale) or mole fraction (bottom scale) with a majority carrier concentration of 10^{18} cm^{-3}. For the *n*-type material CCCH dominates; for *p*-type material, both CHHS and CHHL are considered. (After Ref. 27)

Figure 3.15 shows the temperature dependence of the Auger processes under low injection. The three curves correspond to the three Auger processes. The band gap is chosen to correspond to $\lambda = 1.3$ μm. The Auger lifetimes decrease (or the Auger rates increase) rapidly with increasing temperature. The temperature dependence and the band-gap dependence essentially arise

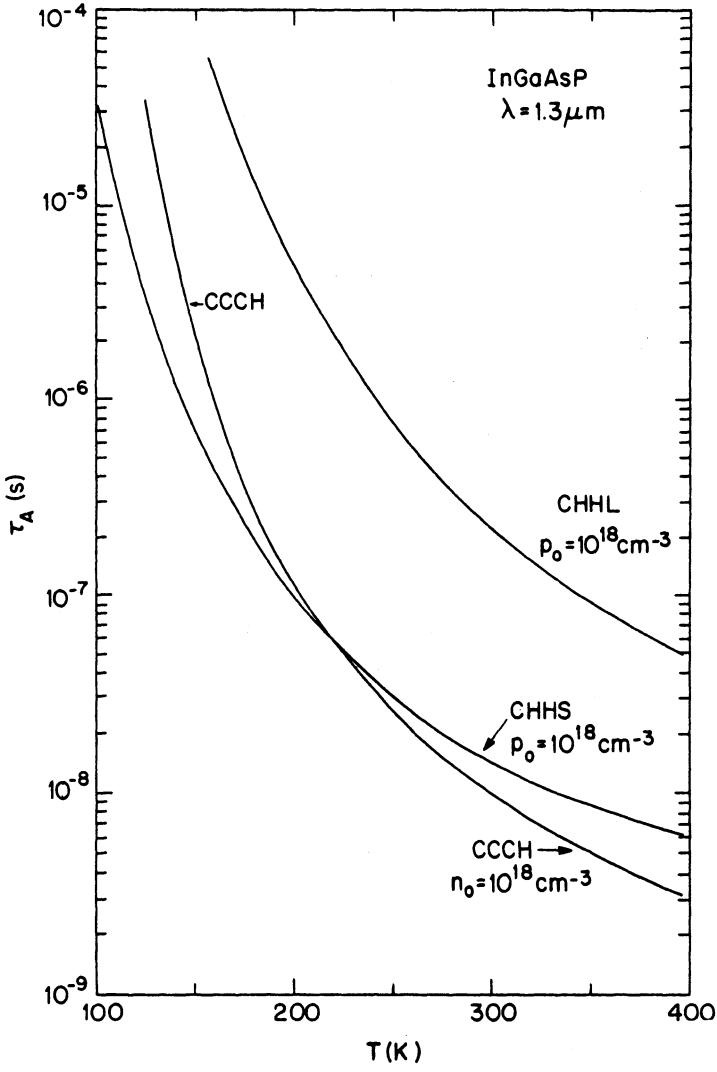


Fig. 3.15 Calculated Auger lifetime for *n*-type and *p*-type InGaAsP as a function of temperature. The majority carrier concentration is 10^{18} cm^{-3}. (After Ref. 27)

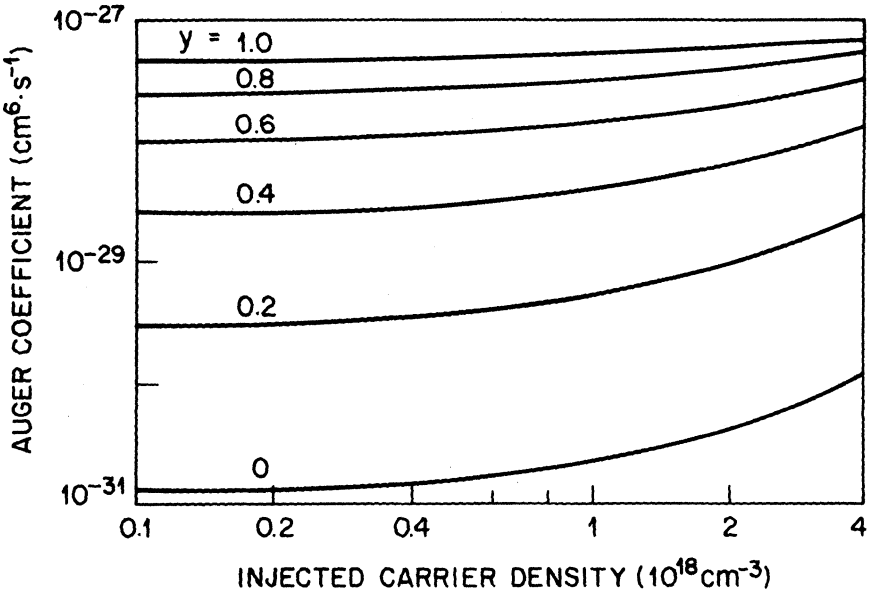


Fig. 3.16 Calculated Auger coefficient (due to CHHS) as a function of injected carrier concentration in $In_{1-x}Ga_xAs_yP_{1-y}$ ($x \cong 0.45y$) for several compositions. (After Ref. 51)

from the conservation of momentum and energy that the four particle states must satisfy. Figures 3.14 and 3.15 show qualitatively that the Auger effect is expected to play an important role in the performance of long-wavelength devices, particularly semiconductor lasers (such as InGaAsP lasers) that operate at higher carrier densities.

The screening of the Coulomb interaction and the use of the Fermi-Dirac statistics can alter the Auger rate at high injection from that expected by simple extrapolation at low injection. Figure 3.16 shows the calculated Auger coefficient[51] for the CHHS Auger processes as a function of injected carrier concentration for undoped $In_{1-x}Ga_xAs_yP_{1-y}$ lattice-matched to InP ($x \cong 0.45y$). Different curves correspond to different values of the mole fraction y. The Auger coefficient C is obtained using the relation $R_a = Cn^3$. In Sec. 2.4 we used C as a phenomenological parameter [see Eq. (2.4.9)]. Figure 3.16 can be used to estimate its numerical value. If we use a typical value for n of $2 \times 10^{18}\,cm^{-3}$ for the injected carrier density in InGaAsP lasers, we find that $C \cong 2 \times 10^{-28}\,cm^6/s$ at $\lambda = 1.3\,\mu m$ ($y \cong 0.6$) and $C \cong 5 \times 10^{-28}\,cm^6/s$ at $\lambda = 1.55\,\mu m$ ($y \cong 0.95$). In Sec. 3.4 we discuss the values of C deduced experimentally. It should be stressed that because of various approximations and uncertainties in the band-structure parameters,

the numerical calculations provide only an order-of-magnitude estimate of the Auger coefficient.

3.3.2 Phonon-Assisted Auger Processes

The temperature dependence and the band-gap dependence of band-to-band Auger processes arise from the conservation laws of energy and momentum that the four "free-particle" states involved must satisfy. These laws, as seen in Sec. 3.3.1, give rise to a minimum kinetic-energy requirement or to a threshold energy E_T for the processes. In phonon-assisted Auger processes, conservation of momentum is satisfied through phonon participation and hence the four particle states need not satisfy conservation of momentum. This eliminates the requirement of minimum kinetic energy.

Phonon-assisted Auger processes in the direct-gap semiconductors were first discussed in Eagles.[38] Figure 3.17 shows these processes schematically, which are labeled CCCHP, CHHSP, and CHHLP. CCCHP is similar to the CCCH process of Fig. 3.13, except that the Coulomb interaction produces a heavy hole in an intermediate state I, which then interacts with a phonon and

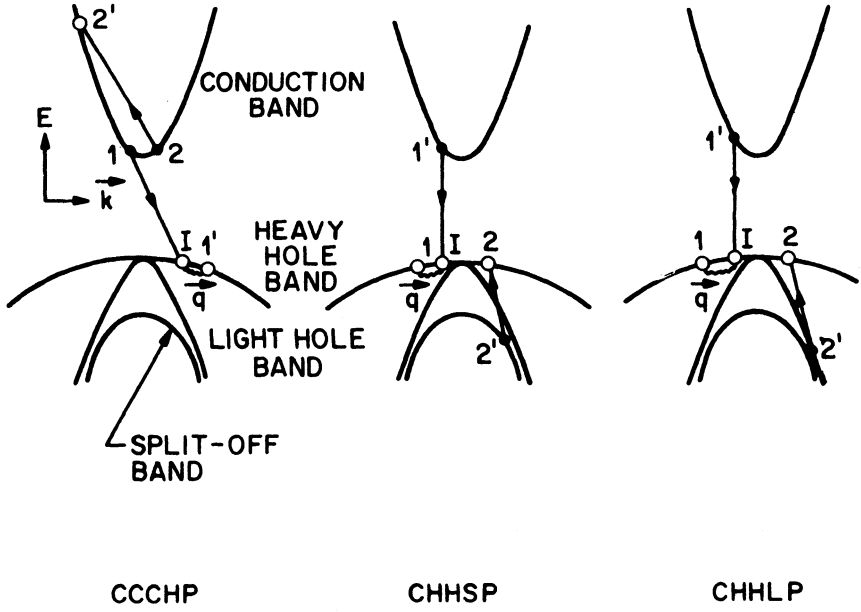


Fig. 3.17 Three different kinds of phonon-assisted Auger processes shown schematically. As in Fig. 3.14, electrons are represented by closed circles and holes by open circles. The wavy line shows the emission or absorption of a phonon of wave vector $\mathbf{q}$.

reaches a final state 1′. Thus the phonon-assisted Auger process is a second-order process. All phonon interactions in Fig. 3.17 are shown to take place in the heavy-hole band because, as will be shown later, heavy holes interact more strongly with phonons compared to electrons and light holes. The calculation of the phonon-assisted Auger processes described here is from Ref. 27.

Phonon-assisted CCCH Process. We first consider the phonon-assisted CCCH process. Using the same approximation for the matrix elements as before, the expression for the Auger rate per unit volume is given by Eq. (3.3.12) but without the δ function representing the momentum conservation, i.e.,

$$R_p = \left(\frac{8\pi}{\hbar}\right)\left(\frac{1}{8\pi^3}\right)^4 \times \iiiint d^3\mathbf{k}_1\, d^3\mathbf{k}_2\, d^3\mathbf{k}_1'\, d^3\mathbf{k}_2'\, P_p(1, 1', 2, 2')|M_{if}|^2\delta(E_i - E_f) \tag{3.3.42}$$

where

$$M_{if} = \sum_m \frac{\langle i|H_{ee}|m\rangle\langle m|H_{ep}|f\rangle}{E_i - E_m} \tag{3.3.43}$$

$$H_{ee} = V(\mathbf{r}_1 - \mathbf{r}_2) = \frac{q^2}{\epsilon|\mathbf{r}_1 - \mathbf{r}_2|}\exp(-\lambda|\mathbf{r}_1 - \mathbf{r}_2|) \tag{3.3.44}$$

$$\langle i|H_{ep}|f\rangle = M(q)[(N_\mathbf{q} + 1/2) \pm 1/2]^{1/2}. \tag{3.3.45}$$

H_{ee} is the Coulomb-interaction Hamiltonian between two electrons as in Eq. (3.3.4), and H_{ep} is the electron-phonon-interaction Hamiltonian.[52] The upper sign in the $\pm$ of Eq. (3.3.45) is for emission of a phonon, and the lower sign is for absorption of a phonon. $M(\mathbf{q})$ is the strength of interaction for the emission or absorption of a phonon of wave vector $\mathbf{q}$, and $N_\mathbf{q}$ denotes the number of phonons of wave vector $\mathbf{q}$. We now use the approximation

$$E_m - E_i \cong E_v(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_2') - E_v(\mathbf{k}_1') \cong \frac{\hbar^2 k_2'^2}{2m_v} = \frac{\mu}{\mu_c}E_g \tag{3.3.46}$$

where, as before, $\mu = m_{c0}/m_v$, $\mu_c = m_{c0}/m_c$, and m_{c0} and m_c denote the conduction-band masses at the band edge and at energy E_g above the band edge. The matrix element of H_{ee} between two two-electron states is given by $|\bar{M}_{if}|^2$ in Eq. (3.3.11). The quantity $P_p(1, 1', 2, 2')$ in Eq. (3.3.42) is

given by

$$P_p = (1 + 2N_q)P(1, 1', 2, 2')$$
$$= (1 + 2N_q)\frac{n^2 p}{N_c^2 N_v}\exp\left(\frac{-E_2' + E_g}{k_B T}\right) \tag{3.3.47}$$

where the second line is obtained for the nondegenerate case after using $P(1, 1', 2, 2')$ from Eq. (3.3.16). The factor $1 + 2N_q$ in Eq. (3.3.47) arises from the matrix elements of H_{ep}, as shown in Eq. (3.3.45), after including both emission and absorption processes. The summation over the intermediate states $|m\rangle$ in Eq. (3.3.43) reduces to a single term because of the momentum conservation:

$$\mathbf{q} = \mathbf{k}_1' + \mathbf{k}_2' - \mathbf{k}_1 - \mathbf{k}_2. \tag{3.3.48}$$

With these simplifications, we get the following expression for the phonon-assisted Auger rate in the nondegenerate approximation:

$$R_p = \left(\frac{8\pi}{\hbar}\right)\left(\frac{1}{8\pi^3}\right)^4\left(\frac{\mu_c}{\mu E_g}\right)^2 \frac{n^2 p}{N_c^2 N_v}\exp\left(\frac{E_g}{k_B T}\right) I \tag{3.3.49}$$

where

$$I = \iiiint |M(\mathbf{q})|^2(1 + 2N_q)|\bar{M}_{if}(k_2')|^2$$
$$\times \exp\left(-\frac{\hbar^2 k_2'^2}{2m_c k_B T}\right)\delta(E_i - E_f)\, d^3\mathbf{k}_1\, d^3\mathbf{k}_2\, d^3\mathbf{k}_1'\, d^3\mathbf{k}_2'. \tag{3.3.50}$$

The quantities $|M(\mathbf{q})|^2$ and N_q depend on the type of phonon-scattering process. We first consider the case of acoustic phonons. Here, the scattering is caused by acoustic waves and is expressed in terms of a deformation potential D.[52] The corresponding matrix element is

$$|M_{AC}(\mathbf{q})|^2 = \frac{D^2 \hbar\omega_p}{2c_1} \tag{3.3.51}$$

where c_1 is the averaged longitudinal elastic constant and $\hbar\omega_p$ is the phonon energy. The number of acoustic phonons N_q at a temperature T is given by

$$N_q = [\exp(\hbar\omega_p/k_B T) - 1]^{-1} \tag{3.3.52}$$

and for $k_B T \gg \hbar\omega_p$ becomes $N_{\mathbf{q}} = k_B T/\hbar\omega_p \gg 1$. Since $|M(\mathbf{q})|^2(1 + 2N_{\mathbf{q}}) \cong D^2 k_B T/c_1$ is independent of $\mathbf{q}$, the integral I may be simplified to yield[27]

$$I_{AC} = \frac{D^2 k_B T}{c_1} \frac{128\pi^5 m_{c0} \mu^2}{105\hbar^2} \left(\frac{\mu_c}{\mu}\right)^{7/2} I_1, \tag{3.3.53}$$

where

$$I_1 = \int_{k_g}^{\infty} dk_2'(k_2'^2 - k_g^2)^{7/2} k_2'^2 \exp\left(-\frac{\hbar^2 k_2'^2}{2m_c k_B T}\right) |\bar{M}_{\text{if}}(k_2')|^2$$

and $k_g^2 = 2m_c E_g/\hbar^2$. In deriving the above, we have neglected the contribution of phonon energy in $E_f - E_i$. Several other phonon-scattering mechanisms have been discussed;[39,40] these include piezoelectric scattering, nonpolar-optical-deformation-potential scattering, and polar optical phonon scattering. These scattering mechanisms are described by different electron-phonon-interaction matrix elements and lead to different values of I in Eq. (3.3.49) for the Auger rate. In the results presented below, all of these mechanisms were included.

We have discussed the case when phonon scattering takes place in the heavy-hole band. When phonon scattering occurs in the conduction band, the quantity $E_m - E_i \cong E_g$ in Eq. (3.3.46). Thus the Auger rate for conduction-band phonon scattering is smaller by a factor $\sim|\mu/\mu_c|^2$. For 1.3-μm InGaAsP, this factor is ~ 20. The calculations of CHHSP for CHHLP phonon-assisted Auger processes are discussed in detail in Ref. 27.

Figure 3.18 shows the calculated phonon-assisted Auger lifetime $\tau_A = n/R_p$ under low injection ($n = 10^{15}$ cm^{-3}) for n-type and p-type InGaAsP with a majority carrier concentration of 10^{18} cm^{-3}. The calculation was carried out using the Kane band model and the electron-phonon-interaction parameters of GaAs. Figure 3.18 shows the variation of τ_A with the band gap of $In_{1-x}Ga_xAs_yP_{1-y}$ lattice-matched to InP ($x = 0.45y$ and different values of y correspond to different band gaps). CCCH dominates for n-type material, while both CHHL and CHHS are considered for the p-type material. Figure 3.18 should be compared with Fig. 3.14. A direct comparison shows that the dependence of the Auger rate on the band gap is considerably weaker for phonon-assisted processes than for direct band-to-band processes. Further, in both cases the Auger lifetime is shorter for the n-type material when CCCH is dominant.

Figure 3.19 shows the relative contributions of phonon-assisted and band-to-band Auger processes together with the total Auger lifetime for p-type InGaAsP. Figure 3.19a shows the variation of τ_A with the band gap at a

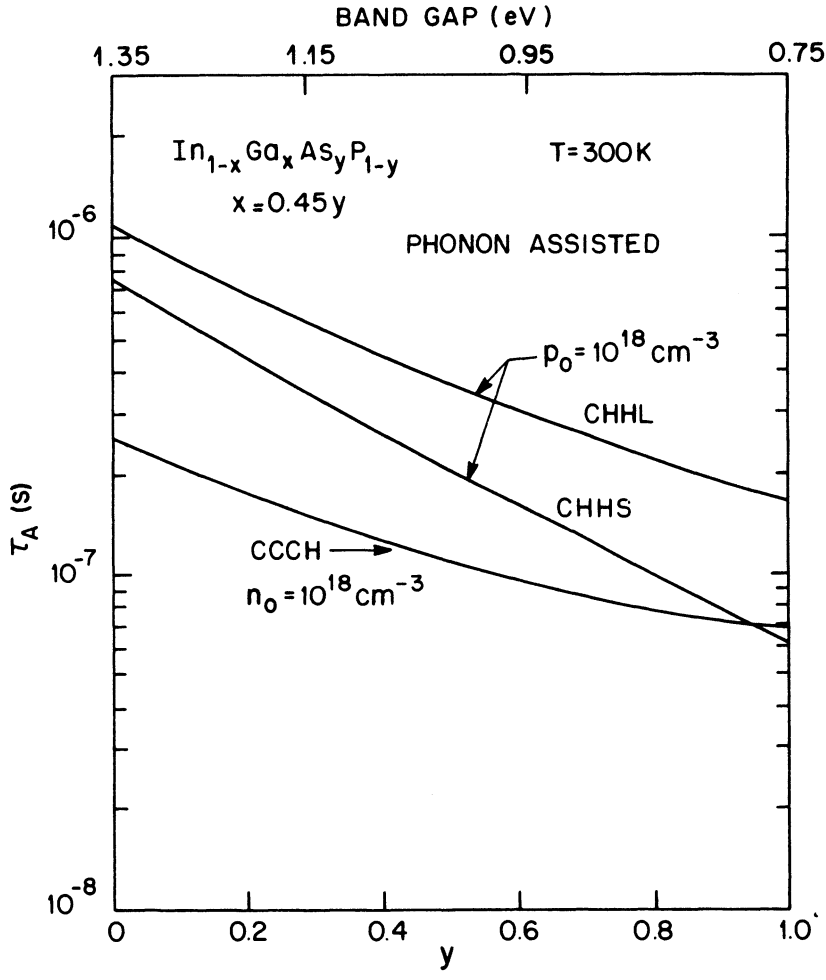


Fig. 3.18 Calculated phonon-assisted Auger lifetime for n-type InGaAsP as a function of the composition (bottom scale) or band gap (top scale). The majority carrier concentration in each case is 1×10^{18} cm^{-3}. Compare with Fig. 3.14 drawn for the direct (unassisted) band-to-band Auger process. (After Ref. 27)

constant temperature $T = 300$ K. We note that when $E_g < 1.14$ eV, the band-to-band Auger process dominates. Figure 3.19b shows the variation of τ_A with temperature at a band gap of 0.96 eV ($y = 0.6$) corresponding to the wavelength $\lambda = 1.3$ μm. The band-to-band Auger process dominates for temperatures exceeding $T = 160$ K. Although these results are strongly influenced by the choice of material parameters and band-structure

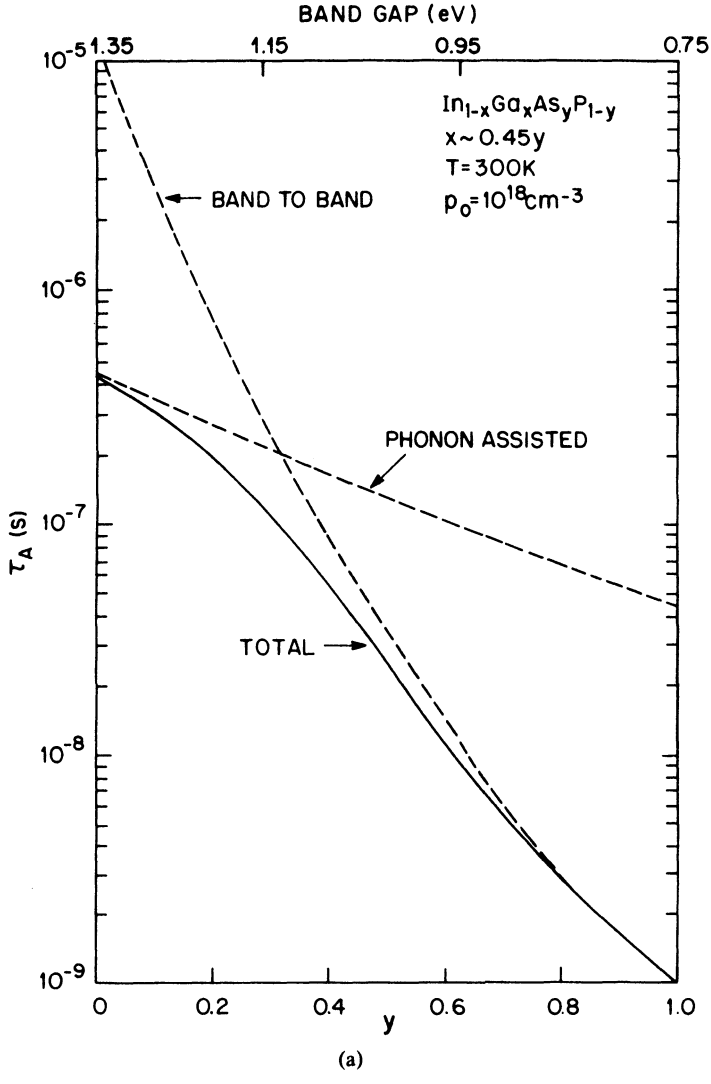
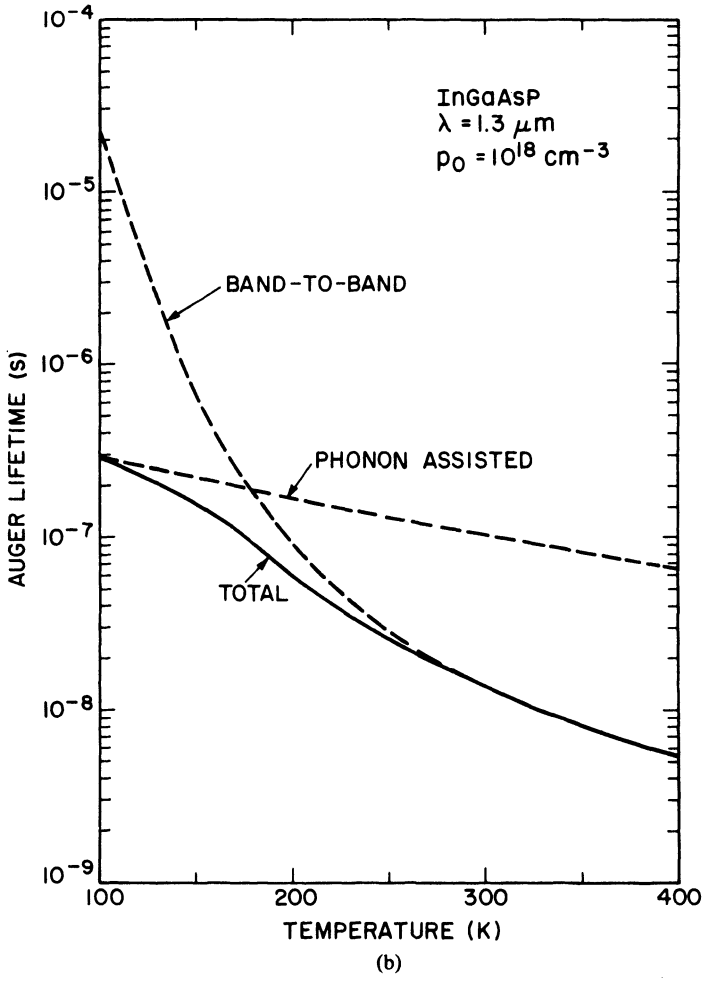


Fig. 3.19 Comparison of band-to-band and phonon-assisted Auger processes in InGaAsP. (a) Auger lifetime as a function of composition (bottom scale) or band gap (top scale). *(continued)*

parameters, the qualitative features are expected to be valid. Haug[49] has calculated the phonon-assisted and band-to-band Auger rates for 1.3-μm InGaAsP using the Chelikowsky and Cohen[50] model of the band structure. He finds that phonon-assisted CCCH and band-to-band CHHS



(b)

Fig. 3.19 (*continued*) (b) Auger lifetime as a function of temperature for $E_g = 0.96$ eV. (After Ref. 27).

are comparable in magnitude. It is important to point out that the numerical values calculated for the various Auger processes are strongly influenced (by as much as one order of magnitude) not only by the magnitude of the matrix elements but also by the band structure model used to calculate the effective masses far from the band edge. Consequently the experimental results provide the only guide for assessing the relative importance of the various Auger processes. These are discussed in Sec. 3.4.

3.3.3 Defect and Surface Recombination

Recombination at Defects. Defects in the active region of an injection laser can be formed in several different ways. In many cases they are grown-in during epitaxy. They can also occur, multiply, or propagate during a stress-aging test. Defects can propagate along a specific crystal axis in a strained lattice. The well-known dark-line defect (dark region of linear aspect) is generally believed to be responsible for the high degradation rates (short life span) of early AlGaAs lasers. Chapter 14 discusses defects and their effects on laser reliability in more detail.

Defects produce a continuum of states in a localized region. Electrons or holes that are within a diffusion length from the edge of the defect may recombine nonradiatively via the continuum of states. The rate of recombination at a defect or trap is usually written as

$$R_d = A_{\text{nr}} n \tag{3.3.54}$$

where

$$A_{\text{nr}} = \sigma v N_t \tag{3.3.55}$$

σ is the capture cross section of the trap, N_t is the density of traps, and v is the velocity of the electrons or holes. The quantity A_{nr} is the inverse of the nonradiative lifetime in the presence of a trap, introduced in Sec. 2.4 to account for nonradiative recombinations.

Surface Recombination. In an injection laser the cleaved facets are surfaces exposed to the ambient. In addition, in many index-guided laser structures, the edges of the active region can be in contact with curved surfaces that may not form a perfect lattice. In general, a surface is a strong perturbation of the crystal lattice, creating many dangling bonds that can absorb impurities from the ambient. Hence a high concentration of defects can occur on these surfaces, and the defects can act as centers of nonradiative recombination. Such localized nonradiative centers, in addition to increasing the threshold current, can cause other performance problems in lasers (e.g., sustained oscillations). Another possible source for surface recombination is the heterostructure interface between the active and cladding layers.

The recombination rate of carriers at the surface is expressed in terms of a surface recombination velocity. If A is the surface area and n_{th} is the threshold carrier density, then the surface-recombination rate is given by

$$R_s = SAn_{\text{th}} \tag{3.3.56}$$

where S is the surface recombination velocity. For an InP surface exposed to air S is about two orders of magnitude smaller than that for GaAs. For the modeling of semiconductor lasers, it is customary to lump defect and surface recombination and introduce a single parameter A_{nr}, as done in Eq. (2.4.9).

3.4 EXPERIMENTAL RESULTS

In this section we present the experimental results on the measurement of radiative and Auger coefficients (the parameters B and C) for InGaAsP material. Many of these measurements were made in an effort to understand the observed higher temperature dependence of the threshold current of InGaAsP lasers ($\lambda \cong 1.3$ μm and 1.55 μm) compared with that of AlGaAs lasers ($\lambda \cong 0.82$ μm). Measurements under conditions of both optical and electrical excitations have been reported.

We first discuss the measurements under optical excitation of InGaAsP from a mode-locked Nd:YAG ($\lambda \cong 1.06$ μm) laser.[53] Figure 3.20 shows the experimental setup. The mode-locked laser pulses (~ 120 ps long) are focused to a diameter of ~ 8 μm onto a 1.3-μm InGaAsP-InP double heterostructure. The optical radiation is absorbed in the InGaAsP layer ($\lambda \cong 1.3$ μm) and not in the InP cladding layers. From the measured transmission of the laser pulse and the spot size, the carrier density generated in the double-heterostructure well (InGaAsP layer) can be estimated. The optically generated carriers can recombine through various mechanisms including radiative and Auger recombinations. The carrier lifetime is measured by measuring the decay of the photoluminescence signal at 1.3 μm. Figure 3.21 shows the experimental data for carrier lifetime as a function of the carrier density. In the absence of recombination at defects and surfaces, the carrier lifetime τ may be written as

$$\frac{1}{\tau} = B(n)n + Cn^2 \tag{3.4.1}$$

where $B(n)$ is the radiative recombination coefficient, C is the Auger coefficient, and n is the carrier density. For undoped InGaAsP material, it is found that $B \cong 2 \times 10^{-10}$ cm^3/s for $n = 1 \times 10^{18}$ cm^{-3} and $C \cong 2.3 \times 10^{-29}$ cm^6/s when $\lambda = 1.3$ μm and $T \cong 300$ K.

Henry et al.[54] have reported similar measurements of the carrier lifetime in InGaAsP-InP double heterostructures at $\lambda = 1.3$ μm. However, they varied the carrier concentration by doping the active layer (n- or p-type). A low-power mode-locked dye laser was used to generate a small number of excess carriers, and the decay rate of these carriers was obtained by

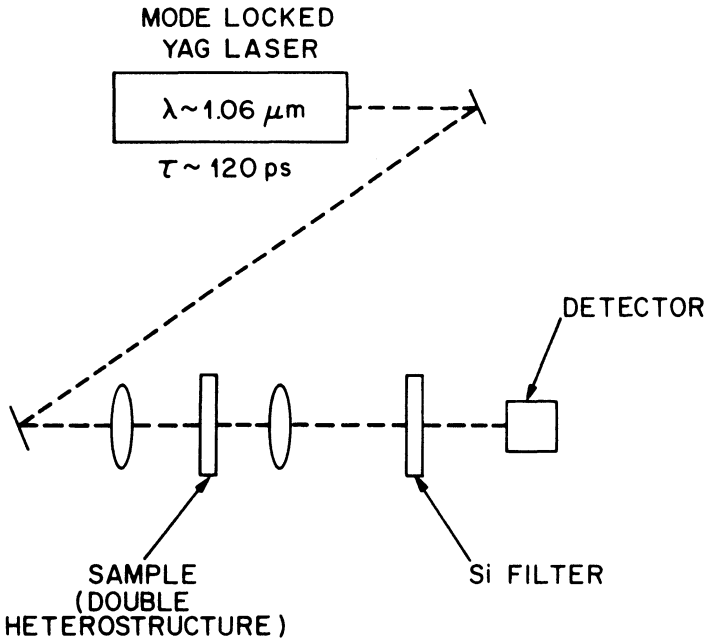


Fig. 3.20 Experimental setup for the measurement of carrier lifetime using the method of optical excitation. Mode-locked YAG pulses of 120-ps duration were used in the experiment. (After Ref. 53)

measuring the luminescence ($\lambda \cong 1.3$ μm) decay time. The lifetime τ is given by Eq. (3.4.1) as before. The estimated Auger coefficient is 1×10^{-29} cm^6/s from these measurements.

The radiative and Auger coefficients in InGaAsP have also been estimated[55] by the technique of absorption bleaching using 1.06-μm pump and probe pulses. In this experiment, a train of high-power pulses (85-ps width, 8-W power) from a mode-locked Nd: YAG laser was focused onto the InGaAsP double-heterostructure sample. This produced a high density of carriers in the InGaAsP layer. Further absorption of a probe beam, also at 1.06 μm, is reduced (or bleached) because the available states are already occupied. This is known as the Burstein effect.[56] Thus, by delaying the probe beam with respect to the pump beam, one can measure the change in absorption, from which the carrier decay time can be obtained. The estimated radiative and Auger coefficients for 1.3-μm InGaAsP are 1.2×10^{-10} cm^3/s and $1.5\text{–}2.8 \times 10^{-29}$ cm^6/s, respectively.

In the CHHS Auger process (Fig. 3.13), two heavy holes interact and

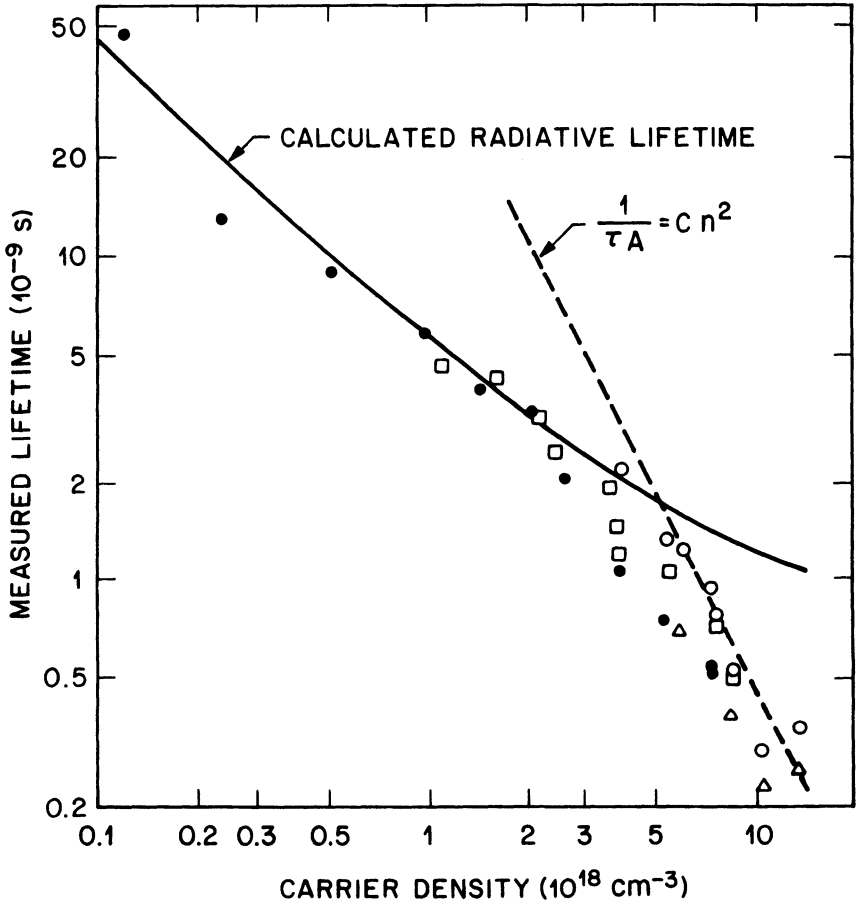


Fig. 3.21 Measured carrier lifetime as a function of injected carrier density. At low carrier densities, radiative recombination dominates and the solid line is used to estimate B. At high carrier densities, Auger recombination dominates and the dashed-line fit yields the value of C. (After Ref. 53)

create an electron in the conduction band and a hole in the split-off band. The electron and the split-off-band hole can recombine with the emission of phonons at an energy $E_g + \Delta$. The CHHS Auger coefficient has been estimated by measuring the intensity of this high-energy ($E_g + \Delta$) luminescence.[57] For 1.3-μm InGaAsP, $E_g + \Delta \cong 1.31$ eV. Figure 3.22 shows the experimental results. Toward low energies, the onset of the band-to-band transition (at ~ 0.96 eV) can be seen. The estimated CHHS Auger coefficient is $\sim 5 \times 10^{-29}$ cm^6/s at 300 K.

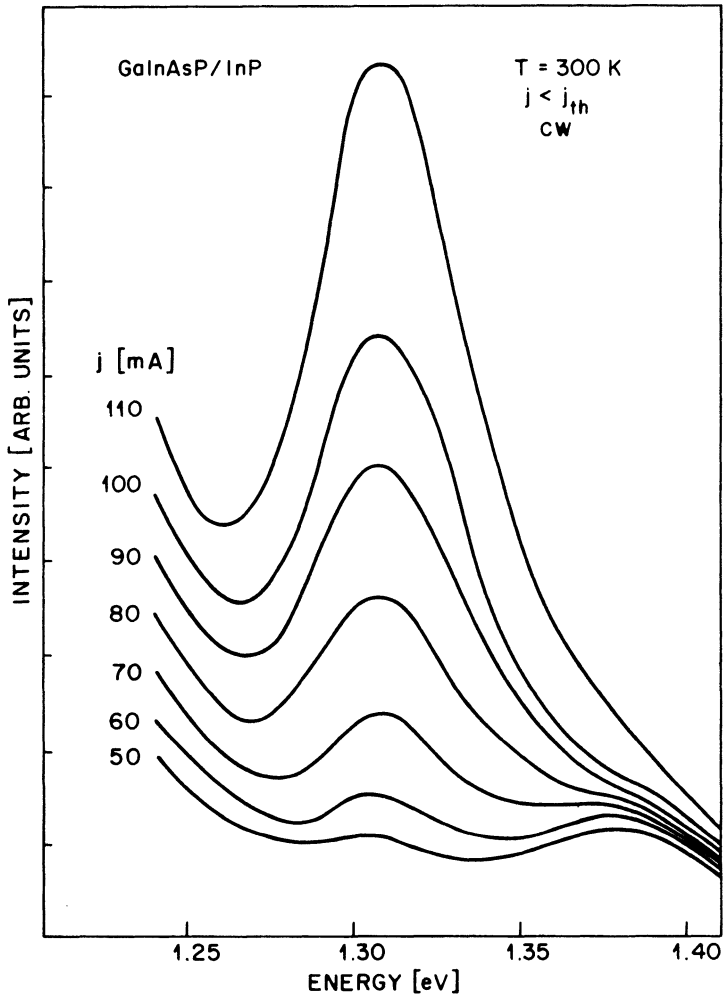


Fig. 3.22 Intensity of photoluminescence near energy $E_g + \Delta$ ($\Delta \cong 1.31$ eV) produced by the recombination of an electron and a split-off-band hole. Different curves are for different injected current densities. (After Ref. 57 © 1983 IEEE)

Measurements of the radiative and Auger coefficients have also been reported using differential carrier-lifetime measurements.[58,59] The differential carrier lifetime τ_n is defined as

$$\frac{1}{\tau_n} = \frac{\partial R}{\partial n} \tag{3.4.2}$$

where $R(n)$ is the total recombination rate [below threshold; see Eq. (2.4.9)] given by

$$R(n) = A_{nr}n + B(n)n^2 + Cn^3. \tag{3.4.3}$$

The injected current I is related to the carrier density n by the relation

$$I = qVR(n) \tag{3.4.4}$$

where V is the active volume. Since the spontaneous-emission rate is proportional to $B(n)n^2$, Eqs. (3.4.3) and (3.4.4) show that in the presence of the Auger effect, the spontaneous-emission rate should vary sublinearly with increasing I. For $C = 0$ and a constant B, the relation between τ_n and I becomes

$$\frac{1}{\tau_n^2} = A_{nr}^2 + \frac{4B}{qV} I. \tag{3.4.5}$$

Thus for $C = 0$, τ_n^{-2} varies linearly with the current I. In the presence of the third term in Eq. (3.4.3), $1/\tau_n^2$ versus I becomes superlinear with increasing I. Figure 3.23 shows an example each of the measured $1/\tau_n^2$ versus I and the measured spontaneous-emission rate versus I for 1.3-μm InGaAsP lasers. The coefficients A_{nr}, $B(n)$ and C in Eq. (3.4.3) can be obtained by

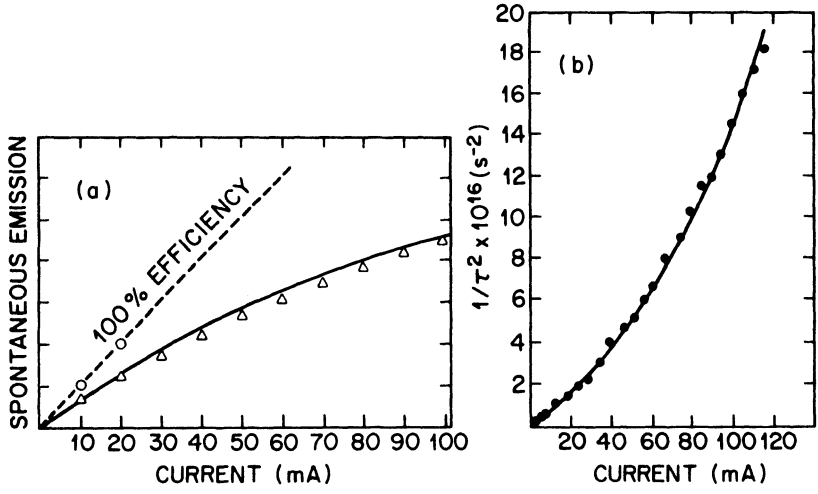


Fig. 3.23 (a) Spontaneous emission intensity plotted against injection current. (b) $1/\tau^2$ plotted against injection current. (After Ref. 28 © 1984 IEEE)

curve-fitting the data of Fig. 3.23. An extensive analysis of such data shows that $B(n)$ for 1.3-μm InGaAsP may be written as

$$B(n) = B_0 - B_1 n \tag{3.4.6}$$

with $B_0 = 0.5\text{–}0.7 \times 10^{-10}$ cm^3/s and $B_1/B_0 = 1.7\text{–}2.2 \times 10^{-19}$ cm^3 at 300 K. The Auger coefficient $C = 1\text{–}2 \times 10^{-29}$ cm^6/s for undoped InGaAsP and 9×10^{-29} cm^6/s for p-type InGaAsP. Thompson[60] arrived at an Auger coefficient of 3×10^{-29} cm^6/s from an independent analysis of the data in Ref. 58.

Uji et al.[61] analyzed the output-power-saturation characteristics of 1.3-μm InGaAsP light-emitting diodes. They showed that by measuring the light-current characteristics of diodes of different active-region thicknesses, the effects of heterojunction leakage and Auger recombination can be separated. They estimate an Auger coefficient in the range of $3\text{–}8 \times 10^{-29}$ cm^6/s and a radiative recombination coefficient in the range $0.7\text{–}1.5 \times 10^{-10}$ cm^3/s at an injected carrier density of $\sim 1 \times 10^{18}$ cm^{-3}.

In an indirect method, the Auger coefficient in InGaAsP material is estimated by fitting the measured threshold current to stripe width for stripe-geometry lasers (see Fig. 2.7).[62] The estimated values are 3×10^{-29} cm^6/s and 9×10^{-29} cm^6/s for 1.3-μm and 1.55-μm InGaAsP respectively.

The various experimental values for the radiative and Auger recombination rates in InGaAsP are listed in Tables 3.2 and 3.3 respectively. Note that different methods lead to considerable differences in the values. In spite of these large differences, it appears that the measured radiative coefficient is smaller (probably by 30%) than the calculated value and that the measured Auger coefficient is smaller (by a factor of 3 to 5) than the theoretical values obtained using the Kane model of the band structure. The Auger rate calculation using the Chelikowsky and Cohen model of the band structure agrees well with the experimental data.

Table 3.2 Measured Radiative Coefficient in InGaAsP

B, cm^3/s	REFERENCE	COMMENT
$0.7\text{–}1.5 \times 10^{-10}$	Uji, Iwamoto, and Lang[61]	$\lambda = 1.3$ μm, LEDs
1.2×10^{-10}	Wintner and Ippen[55]	$\lambda = 1.3$ μm, optically pumped
1.0×10^{-10}	Wintner and Ippen[55]	$\lambda = 1.55$ μm
0.4×10^{-10}	Wintner and Ippen[55]	$\lambda = 1.65$ μm, InGaAs
$B_0 = 0.5\text{–}0.7 \times 10^{-10}$ $B_1/B_0 = 1.7\text{–}2.2 \times 10^{-19}$	Olshansky et al.[28]	$\lambda = 1.3$ μm, $B = B_0 - B_1 n$

Table 3.3 Measured Auger Coefficients in InGaAsP

C, cm^6/s	REFERENCE	COMMENT
5×10^{-29}	Mozer et al.[57]	$\lambda = 1.3$ μm, CHHS process
$2.3 \pm 1 \times 10^{-29}$	Sermage et al.[53]	$\lambda = 1.3$ μm, optically pumped
1×10^{-29}	Henry et al.[54]	$\lambda = 1.3$ μm, doping dependence of τ
$1\text{–}2 \times 10^{-29}$	Su et al.[58]	$\lambda = 1.3$ μm
9×10^{-29}	Su et al.[59]	$\lambda = 1.3$ μm, p-type material
3×10^{-29}	Thompson[60]	Fit to data of Su et al.[58]
$3\text{–}8 \times 10^{-29}$	Uji, Iwamoto, and Lang[61]	$\lambda = 1.3$ μm, LEDs
$1.5\text{–}2.8 \times 10^{-29}$	Wintner and Ippen[55]	$\lambda = 1.3$ μm, optically pumped
7.5×10^{-29}	Wintner and Ippen[55]	$\lambda = 1.55$ μm
9.8×10^{-29}	Wintner and Ippen[55]	$\lambda = 1.65$ μm, InGaAsP

3.5 THRESHOLD CURRENT DENSITY

In this section we show how the calculations presented in Secs. 3.2 and 3.3 can be used for estimating the threshold current density J_{th} of a semiconductor laser. Because of the different facet reflectivities for the TE and TM modes,[63] the TE mode generally has a lower threshold. In Sec. 2.6.1 we obtained J_{th} using the linear-gain model based on Eq. (2.4.3). From Eqs. (2.6.3) and (2.6.4), J_{th} is given by

$$J_{th} = qd(A_{nr}n_{th} + Bn_{th}^2 + Cn_{th}^3) + J_L, \tag{3.5.1}$$

where d is the active-layer thickness and n_{th} is the injected carrier density at threshold. To account for the carrier leakage from the active region, we have introduced J_L into Eq. (3.5.1). In general, the carrier leakage is structure-dependent, as will be discussed in Chapter 5. For simplicity and generality, we consider a broad-area laser for which the leakage current density J_L is only due to carrier leakage over the *heterojunctions* (interfaces between the active and cladding layers).

The threshold carrier density n_{th} is obtained using the threshold condition (2.3.9). Using Eqs. (2.3.10) and (2.5.50) we obtain

$$\Gamma g_{th} = \frac{1}{2L}\ln\left(\frac{1}{R_1R_2}\right) + \Gamma\alpha_a + (1-\Gamma)\alpha_c + \alpha_{scat}, \tag{3.5.2}$$

where the first term represents facet loss, while the remaining terms account for internal losses. Typically, $L \cong 250$ μm, $R_1 = R_2 \cong 0.32$, and the facet

loss is about 45 cm^{-1}. If we consider a 1.3-μm InGaAsP laser with InP cladding layers, the confinement factor Γ for the TE mode is about 0.47 for a 0.2-μm-thick active layer (see Fig. 2.5). We calculate internal losses using $\alpha_{scat} \cong 0$ and $\alpha_a = \alpha_c = 30\ cm^{-1}$. With these parameters, the threshold gain $g_{th} \cong 166\ cm^{-1}$. The injected carrier density at room temperature ($T = 300$ K) required to achieve this gain can be obtained using Fig. 3.8; $n_{th} \cong 1.7 \times 10^{18}\ cm^{-3}$.

The threshold current density J_{th} can now be obtained using Eq. (3.5.1) provided the parameters A_{nr}, B, and C have been estimated from the results presented in Secs. 3.2 and 3.3. For good-quality InGaAsP lasers, the contribution of A_{nr} (trap and surface recombinations) is often negligible compared with the other terms in Eq. (3.5.1), and J_{th} can be rewritten as

$$J_{th} \cong J_r + J_{nr} + J_L \tag{3.5.3}$$

where

$$J_r = qd(Bn_{th}^2) = qdR \tag{3.5.4}$$

$$J_{nr} = qd(Cn_{th}^3) = qdR_a \tag{3.5.5}$$

are the radiative and nonradiative components of the current density, respectively. The radiative spontaneous-emission rate R and the Auger recombination rate R_a can be obtained from the results presented in Secs. 3.2 and 3.3. For the example under consideration, using $d = 0.2$ μm and $g_{th} = 166\ cm^{-1}$, we obtain $J_r \cong 1\ kA/cm^2$ from Fig. 3.9. Further, if we use $C = 2.3 \times 10^{-29}\ cm^6/s$ as the experimentally deduced value (see Sec. 2.4), $J_{nr} = 0.36\ kA/cm^2$ from Eq. (3.5.5). Thus the calculated J_{th} in our example is $1.36\ kA/cm^2$ in the absence of carrier leakage. This compares well with the measured values, which are typically in the range 1–1.5 kA/cm^2. In the next section we calculate the carrier leakage over the heterojunctions in order to estimate J_L.

3.5.1 Carrier Leakage over the Heterojunctions

Heterojunction carrier leakage is caused by diffusion and drift of electrons and holes from the edges of the active region to the cladding layers, and is schematically shown in Fig. 3.24a. The heterojunction leakage in a double heterostructure has been extensively studied.[64–71] In thermal equilibrium, at the boundary between the active and cladding layers, a certain number of electrons and holes are present. Figure 3.24b shows an energy-level diagram for the heterojunction of active and p-cladding layers; the electron and hole

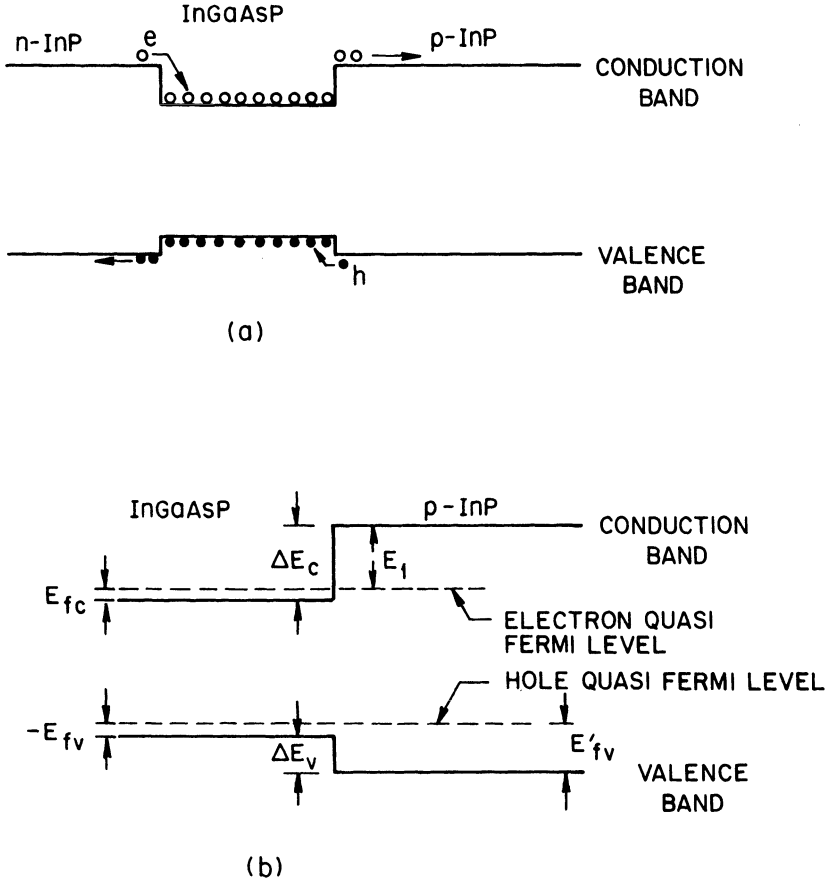


Fig. 3.24 (a) Schematic illustration of the heterojunction leakage current. (b) Energy level diagram near the boundary of *P*-InP cladding layer and InGaAsP active layer.

quasi-Fermi levels E_{fc} and E_{fv} are assumed to be continuous at the boundary.[66] The number of electrons n_b at the boundary of the p-cladding layer is given by [from Eq. (3.2.27)]

$$n_b = N_{cc} \frac{2}{\pi^{1/2}} \int_{\epsilon_c}^{\infty} \frac{\epsilon^{1/2}\, d\epsilon}{1 + \exp(\epsilon - \epsilon_{fc})} \tag{3.5.6}$$

where

$$N_{cc} = 2\left(\frac{2\pi m_c k_B T}{h^2}\right)^{3/2} \qquad \epsilon_c = \Delta E_c / k_B T$$

and m_c is the conduction-band mass of the p-cladding layer. Note that n_b is the number of electrons with energy greater than the conduction-band barrier height ΔE_c. Using the Boltzmann approximation for the non-degenerate case, Eq. (3.5.6) may be simplified to yield

$$n_b = N_{cc} \frac{2}{\pi^{1/2}} \int_{\epsilon_c}^{\infty} \epsilon^{1/2}\, d\epsilon \exp(-\epsilon + \epsilon_{fc}) = N_{cc} \exp\left(\frac{-E_1}{k_B T}\right). \tag{3.5.7}$$

The quantity $E_1 = \Delta E_c - E_{fc}$ is shown in Fig. 3.24. In deriving the above, we have used the relation $\epsilon_{fc} = E_{fc}/k_B T$. We now show that E_1 is related to the band-gap-difference ΔE_g between the active and cladding layers. From Fig. 3.24 we note that

$$E_1 = \Delta E_c - E_{fc} = \Delta E_g - \Delta E_v - E_{fc}. \tag{3.5.8}$$

Further, ΔE_v is given by

$$\Delta E_v = E'_{fv} + E_{fv} \tag{3.5.9}$$

where E'_{fv} and E_{fv} are the hole quasi-Fermi levels in the p-cladding layer and active layer respectively. From Eqs. (3.5.7), (3.5.8), and (3.5.9) it follows that

$$n_b = \frac{N_{cc} N_{vc}}{P} \exp\left(-\frac{\Delta E_g + E_{fc} + E_{fv}}{k_B T}\right) \tag{3.5.10}$$

where we have used the relation

$$P = N_{vc} \exp\left(\frac{-E'_{fv}}{k_B T}\right) \tag{3.5.11}$$

for the majority (hole) carrier density and

$$N_{vc} = 2\left(\frac{2\pi k_B T}{h^2}\right)^{3/2} (m_{hh}^{3/2} + m_{lh}^{3/2}) \tag{3.5.12}$$

is the valence-band density of states for the p-cladding layer. Note that m_c, m_{hh}, and m_{lh} are the effective masses for the p-cladding layer and can be obtained using Table 3.1.

The quantities E_{fc} and E_{fv} can be calculated from the known carrier density in the active region [using Eq. (3.2.28)]. An equation similar to (3.5.10)

can be derived for the density of holes (p_b) at the boundary between the n-cladding layer and the active layer. Equation (3.5.10) shows that n_b increases rapidly with increasing temperature and suggests that the carrier leakage can be a major carrier-loss mechanism at high temperatures, especially for low heterojunction-barrier heights.

The electron leakage current density j_n at the p-cladding layer is given by[71]

$$j_n = -qD_n \frac{dn}{dx} + qn\mu_n E \tag{3.5.13}$$

where the first term represents diffusive leakage and the second term represents drift leakage in the presence of an electric field E. D_n is the electron diffusivity, μ_n is the minority carrier mobility and $n(x)$ is the density of electrons at a distance x from the boundary between the active region and the P-cladding layer. The current j_n also satisfies the continuity equation

$$\frac{1}{q}\frac{dj_n}{dx} + \frac{n}{\tau_n} = 0 \tag{3.5.14}$$

where τ_n is the minority carrier lifetime. Equation (3.5.13) can be solved using the boundary condition $n(x = 0) = n_b$ and $n(x = h) = 0$. The second condition assumes that the minority carrier density at the contact, which is at a distance h from the boundary, is 0. The result for the electron leakage current at $x = 0$ is

$$j_n = qD_n n_b \frac{(Z - Z_1)\exp(Z_2 h) + (Z_2 - Z)\exp(Z_1 h)}{\exp(Z_2 h) - \exp(Z_1 h)} \tag{3.5.15}$$

where

$$Z = qE/k_B T \tag{3.5.16}$$

$$Z_{2,1} = \tfrac{1}{2}Z \pm \left(\frac{1}{L_n^2} + \tfrac{1}{4}Z^2\right)^{1/2} \tag{3.5.17}$$

$$L_n = (D_n \tau_n)^{1/2}. \tag{3.5.18}$$

L_n is the electron diffusion length. In the above, the relation $D_n = \mu_n k_B T/q$ has been used. In the limit $E = 0$, Eq. (3.5.15) reduces to the case of pure diffusive leakage, and for $L_n \gg h$, it reduces to $i_n = qn_b\mu_n E$, which holds if only drift leakage is present. A similar equation can be derived for the hole leakage current j_p in the n-cladding layer. However, since the

diffusion length and mobility of electrons are large compared to those of holes, the electron leakage is considerably larger than the hole leakage. The total leakage is given by the sum

$$J_L = j_n + j_p. \tag{3.5.19}$$

Figure 3.25 shows the calculated J_L for the diffusive heterojunction leakage in an InGaAsP-InP laser.[72] At higher temperatures, the leakage current is

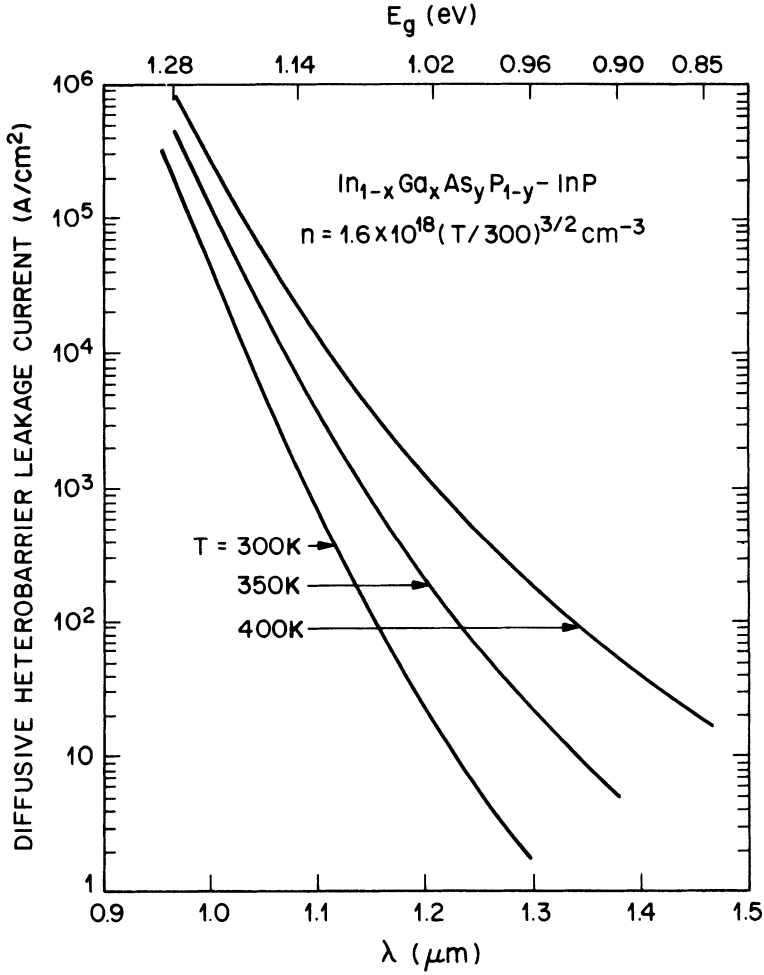


Fig. 3.25 Calculated diffusive heterojunction leakage current (J_L) in an InGaAsP-InP double heterostructure as a function of wavelength (bottom scale) or band gap (top scale). Curves for three different temperatures are shown. (After Ref. 72 © 1982 IEEE)

considerably higher. Also, the leakage current increases rapidly when the barrier height decreases. At 1.3-μm and 1.55-μm wavelengths, however, the calculated $J_L \ll 100$ A/cm^2 near room temperature, implying that the hetero-barrier carrier leakage is not significant for such InGaAsP layers. The calculation assumes that the electron and hole concentrations in the *n*- and *p*-cladding layers are 3×10^{17} cm^{-3}, respectively, and that the electron and hole diffusion lengths are 5 μm and 1 μm, respectively. The thickness h of the *p*-cladding layer is assumed to be 2 μm. The drift leakage current increases rapidly with a decrease in *p*-cladding-layer doping. For a carrier concentration of 3×10^{17} cm^{-3} or higher, which is normally the case for semiconductor lasers, the drift leakage is small compared to the diffusive leakage.

3.6 TEMPERATURE DEPENDENCE OF THRESHOLD CURRENT

In this section we address the important issue of the high-temperature performance of semiconductor lasers. The threshold current of double-heterostructure lasers is found to vary with temperature T as

$$I_{th}(T) = I_0 \exp(T/T_0) \tag{3.6.1}$$

where I_0 is a constant and T_0 is a characteristic temperature often used to express the temperature sensitivity of threshold current. For AlGaAs lasers the observed $T_0 \geq 120$ K near room temperature, while for InGaAsP lasers, T_0 values lie in the range of 50–70 K.[73–79] A lower T_0 value implies that the threshold current increases more rapidly with increasing temperature. The high temperature sensitivity of the threshold current of InGaAsP lasers limits their performance under high-temperature operation. Furthermore, under CW operation at room temperature, the maximum power emitted by these lasers is limited by a thermal runaway process: more and more current is required to offset the effect of the internal temperature increase and this in turn further increases the temperature. Because of these practical limitations, a considerable amount of experimental and theoretical work has been done to understand the higher temperature sensitivity of the threshold current of InGaAsP lasers.[80–98]

In some InGaAsP laser structures, a part of the injected current can flow around the active region, the magnitude of which may vary with temperature; this results in an anomalously low or high temperature sensitivity of the threshold current.[81] The purpose of this section is not to consider such structure-related effects but to concentrate on the fundamental aspects that can give rise to a high temperature dependence of the threshold current. For this reason, most results are presented for broad-area or stripe-geometry layers.

Figure 3.26 shows the measured threshold current of a broad-area 1.3-μm

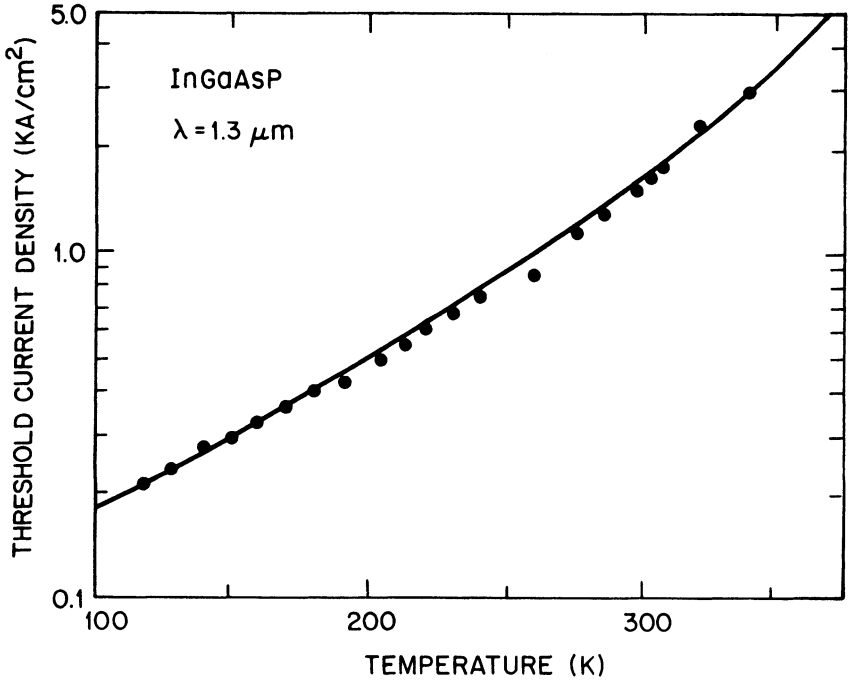


Fig. 3.26 Measured threshold current density as a function of temperature for a 1.3-μm InGaAsP-InP laser. The observed exponential variation can be used to deduce the characteristic temperature T_0. Note the semilog plot.

InGaAsP-InP laser as a function of the operating temperature. The threshold current can be represented by Eq. (3.6.1) with $T_0 \cong 110$ K for 100 K $\leq T \leq$ 240 K and $T_0 \cong 60$ K for 240 K $\leq T \leq$ 340 K. At higher temperatures, even lower T_0 values are observed. Similar T_0 values have been reported by optical pumping using a YAG laser[82] or a pulsed dye laser[83,84] as the excitation source.

Several authors have measured the threshold current as a function of temperature in the temperature range of 10–70°C, which is generally the operating temperature range of these devices when they are used in practical systems. Figure 3.27 shows the estimated T_0 obtained using Eq. (3.6.1) from these measurements for InGaAsP-InP lasers as a function of the emission wavelength. The measured T_0 values generally lie in the range of 40–70 K. Several mechanisms have been proposed to explain the observed high temperature sensitivity of the threshold current of InGaAsP lasers. These are carrier leakage over the heterojunction,[67–70,87–90] Auger recombination,[27–33,58–64,91–95] and intervalence band absorption.[96–98] The differences

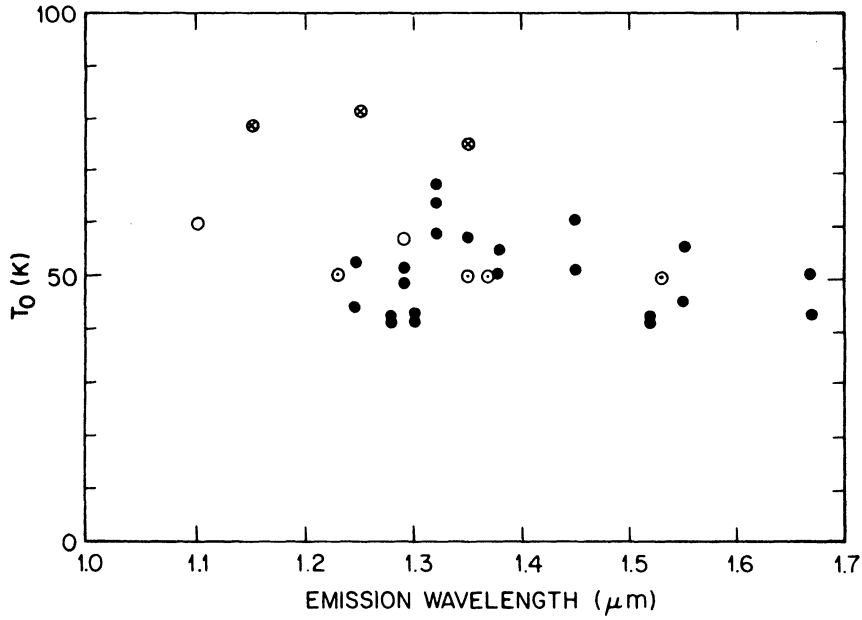


Fig. 3.27 Measured T_0 values for InGaAsP-InP lasers as a function of emission wavelength. The data points are from Ref. 76 (⊙), Ref. 78 (●), Ref. 85 (⊗), and Ref. 73 (○).

between the temperature dependences of other lasing characteristics (e.g., carrier lifetime at threshold, external differential quantum efficiency, and optical gain) of GaAs-AlGaAs and InGaAsP-InP lasers have been extensively studied. The results of these studies are consistent with the observed higher temperature dependence of the threshold current of InGaAsP-InP lasers compared with that of GaAs-AlGaAs lasers. We now describe some of these measurements.

3.6.1 Carrier-Lifetime Measurements

The carrier lifetime at threshold τ_{th} has been measured by the turn-on delay technique (see Sec. 6.4.2) for both GaAs-AlGaAs and InGaAsP-InP lasers. Figure 3.28 shows the measured τ_{th} as a function of temperature. The results obtained are analyzed in the following fashion. The threshold current density J_{th} of a broad-area laser is given by Eq. (2.6.3), or

$$J_{th} = qdn_{th}/\tau_{th} \tag{3.6.2}$$

where d is the active-layer thickness and n_{th} is the carrier density at threshold.

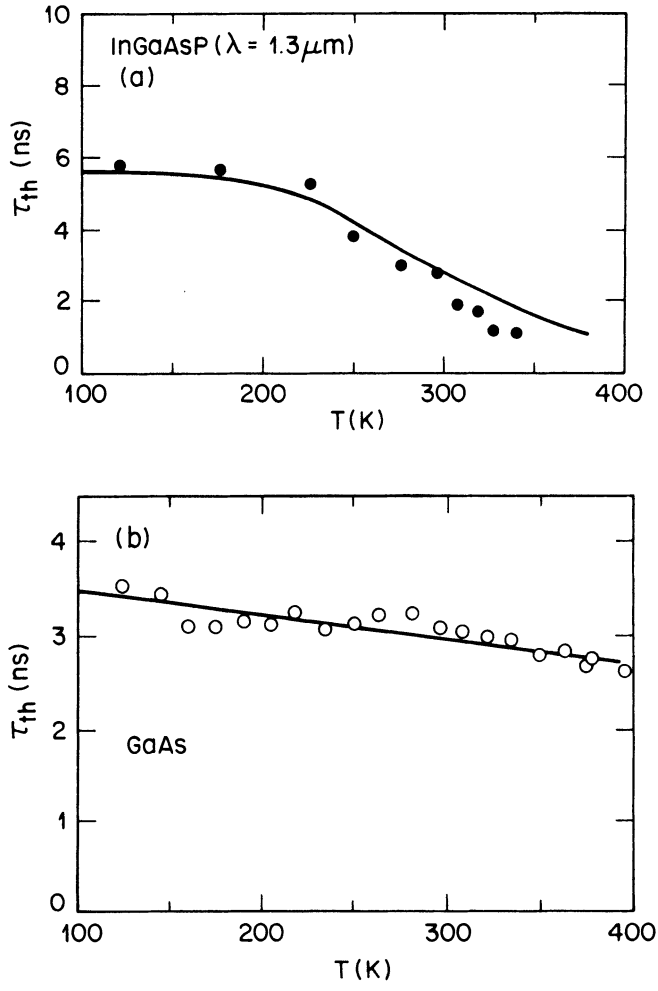


Fig. 3.28 The measured carrier lifetime at threshold as a function of temperature for (a) 1.3-μm InGaAsP-InP and (b) GaAs-AlGaAs lasers. (After Ref. 72 © 1982 IEEE)

The observed decrease of τ_{th} with increasing temperature (in the high temperature range) is more rapid for 1.3-μm InGaAsP lasers than for AlGaAs lasers; this is consistent with the observation that J_{th} increases with increasing temperature more rapidly in InGaAsP lasers than in AlGaAs lasers. Furthermore, it is possible to conclude that the threshold carrier density n_{th} does not increase as rapidly with increasing temperature as does J_{th}. This is consistent with the measurements of n_{th} using short electrical or optical pulse excitation.

3.6.2 Optical-Gain Measurements

The temperature dependence of the optical gain has been measured by both current injection and optical pumping. In the current-injection technique, the net optical gain is obtained from measuring the intensity modulation in spontaneous-emission spectra that are observed from the facet of a laser at currents below threshold.[99] The modulation is caused by the Fabry-Perot resonances (the longitudinal modes) of the cavity. The net gain G is negative below threshold and is given by

$$G = \frac{1}{L} \ln\left(\frac{r^{1/2} - 1}{r^{1/2} + 1}\right) \tag{3.6.3}$$

where $G = \Gamma g - \alpha_t$, $r = P_{max}/P_{min}$ and α_t is the total loss in the cavity including mirror losses. The quantities P_{max} and P_{min} are the maximum and minimum intensities in the spontaneous-emission spectrum. G is found to vary linearly with the current I. The measured slope of the gain-current curve dG/dI is shown in Fig. 3.29 for both InGaAsP-InP and GaAs-AlGaAs lasers. As the temperature increases, the quantity dG/dI decreases more rapidly for InGaAsP (for $T > 230$ K) than for GaAs. An explanation for these results is that a fraction of the injected current in InGaAsP lasers is lost (by nonradiative recombination or carrier leakage) at high temperatures and that this fraction does not contribute to radiative recombination or optical gain.

Measurements of the optical gain have been reported using excitation from a pulsed N_2-laser-pumped dye laser.[100] The gain is determined by measuring the dependence of the amplified intensity on the length of the excited region. Measurements of both the gain spectra and its temperature dependence have been reported. Figure 3.30a shows the measured gain spectra at 300 K, and Fig. 3.30b shows the optical pump intensity required to maintain a constant gain as a function of temperature. The measured gain spectrum is similar to that calculated in Fig. 3.7. The optical pump power needed to maintain a constant gain is found to vary with temperature T as $P = P_0 \exp(T/T_0)$. The T_0 values are noted in Fig. 3.30b. The observed lower T_0 above a certain temperature ($T = 265$ K) is consistent with the higher temperature dependence of threshold current in the same temperature range (see Fig. 3.26).

3.6.3 External Differential Quantum Efficiency

The external differential quantum efficiency for a laser with cleaved facets is given by Eq. (2.6.11), or

$$\eta_d = \eta_i \frac{\alpha_m}{\alpha_m + \alpha_{\text{int}}} \tag{3.6.4}$$

where η_i is the fraction of injected carriers than recombine to produce a stimulated photon, $\alpha_m = \ln(1/R_m)/L$ is the mirror loss, L is the length of the laser, R is the facet reflectivity, and α_{int} is the sum of internal losses in the active and cladding layers. A measurement of η_d as a function of

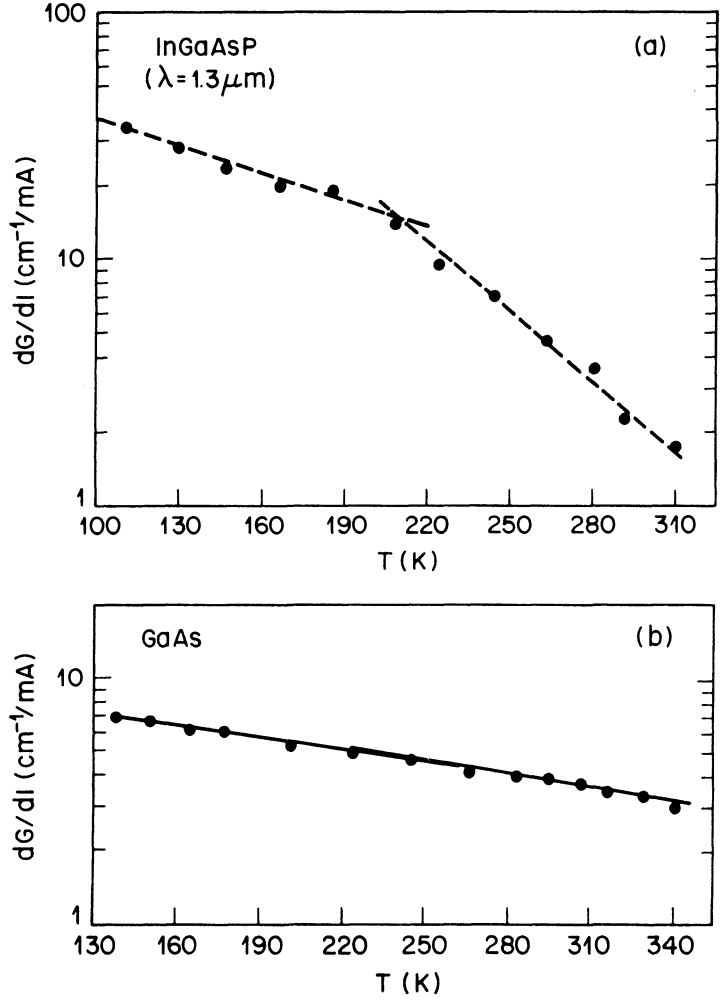


Fig. 3.29 Measured slope of the gain-versus-current curve (dG/dI) as a function of temperature for (a) 1.3-μm InGaAsP-InP and (b) GaAs-AlGaAs lasers. (After Ref. 72 © 1982 IEEE)

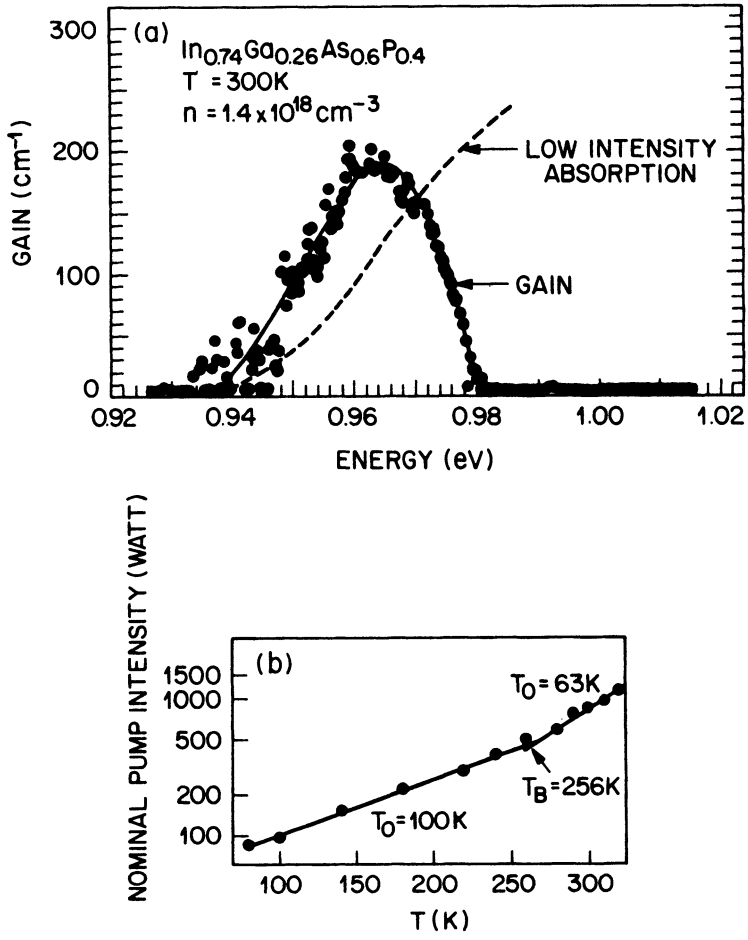


Fig. 3.30 (a) Measured gain spectra of 1.3-μm InGaAsP-InP obtained by optical pumping. Dashed line shows the absorption when the pump intensity is so small that the carrier density is below the transparency level. (After Ref. 83). (b) The optical excitation power needed to maintain the constant gain as a function of temperature. Note the change in slope at $T_B = 256$ K. (After Ref. 100)

temperature may be used to infer the variation of α_{int} with temperature. However, for an absolute determination of α_{int}, η_i must be known, and for a determination of the temperature dependence of α_{int}, the variation of η_i with temperature must be known. Incomplete knowledge of η_i may lead to misleading conclusions. For example, in some laser structures a large fraction of current can flow around the active region; as a result the observed η_d is small. The measured η_d as a function of temperature for GaAs and InGaAsP

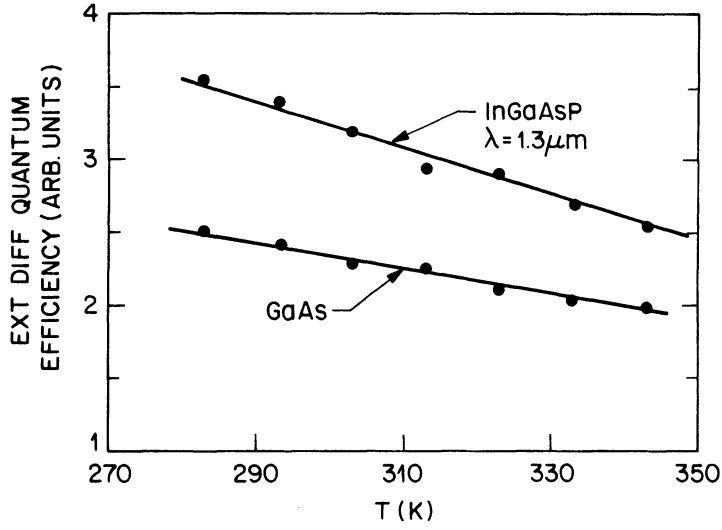


Fig. 3.31 The measured external differential quantum efficiency per facet as a function of temperature for InGaAsP ($\lambda = 1.3$ μm) and GaAs lasers.

lasers is shown in Fig. 3.31. The lasers are oxide-stripe devices that do not have significant current leakage paths. The temperature dependence of η_d is somewhat steeper for InGaAsP lasers than that for GaAs lasers.

3.6.4 Discussion

The experimental observations on the temperature dependence of the threshold current, the carrier lifetime at threshold, the optical gain, and the sublinearity of spontaneous emission (Sec. 3.4) suggest that InGaAsP lasers suffer from a nonradiative carrier loss at high temperatures. This carrier loss may be due to heterobarrier leakage that, as shown in Sec. 3.5, increases with increasing temperature.[67–70] Heterobarrier leakage can be a major carrier loss mechanism for InGaAsP-InP lasers with emission wavelengths less than 1.1 μm due to a small heterobarrier height. For lasers emitting near 1.3 μm and 1.55 μm, which is the region of interest for optical communication applications, the calculated drift-and-diffusion type of leakage is small, as shown in Fig. 3.25. Experimentally, the leakage current is about 10–30% of the total current and is mostly due to electron leakage.[88,89] Since the magnitude of the conventional drift-and-diffusion-type heterobarrier leakage depends on the doping and thickness of the *p*-cladding layer, laser structures can be designed with very small carrier leakage. The threshold current in

such lasers nonetheless exhibits a high temperature dependence, suggesting that other mechanisms are also operative.

Another mechanism for carrier loss at high temperatures is Auger recombination. The measured Auger coefficient is large enough to account for a lower T_0 value for 1.3-μm and 1.55-μm InGaAsP lasers than that observed for AlGaAs lasers. The threshold current density of a broad-area laser can be calculated using Eq. (3.5.1). The calculated value is determined principally by the magnitudes of coefficients B and C and the threshold carrier density n_{th} if we assume that $A_{nr} = 0$ and $J_L = 0$. The value of n_{th} depends on the threshold gain g_{th} which is determined by the laser structure (e.g., through the mode confinement factor) and optical absorption.

There have been several suggestions[96–98] based on indirect measurements that intervalence band absorption in the active layer due to transitions between the split-off band and the heavy-hole band is large and may be responsible for the strong temperature sensitivity of J_{th}. However, direct measurements yield small values. Henry et al.[101] obtained values of 14 cm^{-1} and 12 cm^{-1} for intervalence band absorption in p-type InP and p-type InGaAsP, respectively, at $\lambda = 1.3$ μm for a carrier concentration of 10^{18} cm^{-3}. When $\lambda = 1.6$ μm, these values are 24 cm^{-1} and 25 cm^{-1} for p-type InP and p-type InGaAsP, respectively, at a carrier concentration of 10^{18} cm^{-3}. Casey and Carter[102] reported similar values of intervalence band absorption in p-type InP. These measurements show that intervalence band absorption is significantly smaller that what is needed to explain the observed temperature dependence of the threshold current. The internal absorption losses are determined by free carrier absorption in addition to intervalence band absorption, which is determined by the carrier concentrations in the active and cladding layers. Waveguide scattering losses are also included among the internal losses. For the purpose of calculations described in this section, we assume that a temperature-independent total optical loss $\alpha_a\Gamma + (1 - \Gamma)\alpha_c = 30$ cm^{-1}.

Figure 3.32 shows the calculated radiative and Auger components of the threshold current density plotted as a function of temperature for 1.3-μm and 1.55-μm InGaAsP-InP broad-area lasers with an active-layer thickness of 0.2 μm. The radiative component of the current is calculated using the gain-versus-nominal-current-density relation of Fig. 3.9. Figure 3.32a shows that the threshold current density of 1.55-μm InGaAsP lasers should be lower than that for 1.3-μm InGaAsP lasers in the absence of Auger recombination. The Auger component of the total current density as calculated from Eq. (3.5.5) is shown in Fig. 3.32b. The total calculated threshold current density J_{th} is shown in Fig. 3.32c. If the calculated threshold current I_{th} is expressed as $\sim I_0 \exp(T/T_0)$, the calculated T_0 values of the radiative component J_r and $T_0 \cong 100$ K for 100 K $< T <$ 200 K and $T_0 \cong 200$ K for 270 K $< T <$ 350 K. The calculated temperature dependence of J_r in the low

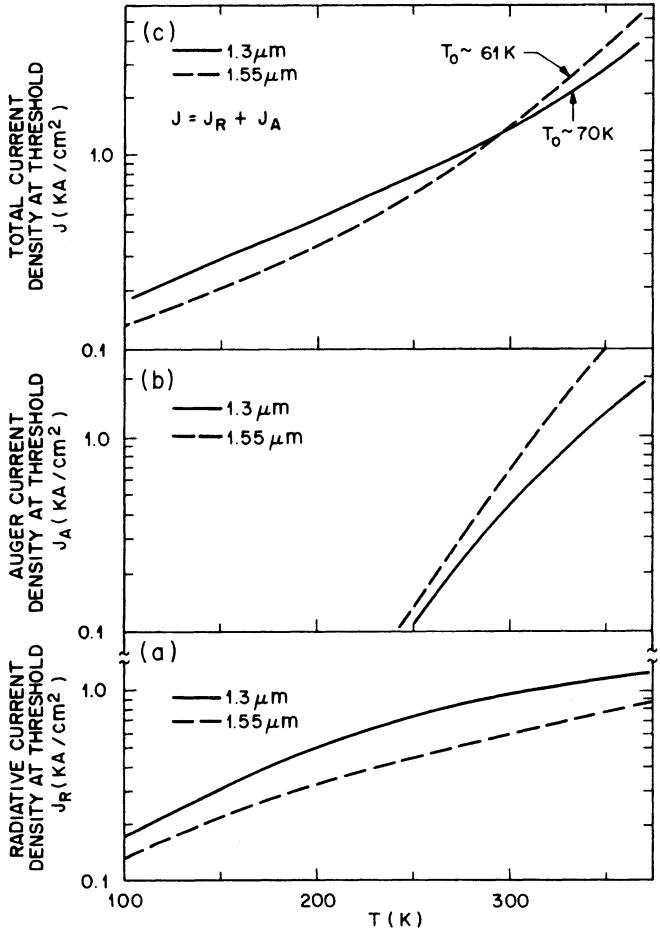


Fig. 3.32 Calculated radiative and nonradiative components of the threshold current density as a function of temperature for 1.3-μm and 1.55-μm InGaAsP lasers: (a) radiative, (b) nonradiative (Auger), and (c) total. Note the crossover for the total threshold current density.

temperature range agrees with the measured temperature dependence characterized by T_0 values in the range of 100–110 K. However, the radiative component alone cannot explain the measured temperature dependence of J_{th} in the high temperature range. If Auger recombination is included, the calculated values for T_0 in the high temperature range (270 K < T < 350 K) using Fig. 3.30c are 70 K and ~61 K for 1.3-μm and 1.55-μm InGaAsP lasers, respectively. (These values increase to 90 K and 72 K if a temperature-independent Auger coefficient is used in the calculation.) Experimental T_0

values are slightly lower than the values calculated above. Considering many uncertainties in the calculation, it is possible to conclude that the Auger processes, which become increasingly dominant with decreasing band gap, play a significant role in determining the observed high temperature sensitivity of long-wavelength InGaAsP lasers.

It has been pointed out that the carrier leakage can also take place due to Auger recombination.[103] In the CCCH Auger process, hot electrons with energies far in excess of the heterobarrier height are produced. A fraction of these electrons can "leak" out of the active region to the cladding layers, constituting a leakage current, the value of which may not depend strongly on barrier height. In some experiments[104,105] the carrier temperature is found to be considerably higher than the lattice temperature near threshold in InGaAsP lasers. However, other measurements[106,107] show that the two temperatures are nearly equal. An understanding of the interaction of the hot Auger electrons with the lattice will be helpful in estimating the importance of Auger recombination for carrier leakage in InGaAsP laser structures.[108]

PROBLEMS

3.1 Derive Eq. (3.2.26) for the density of states in the conduction band by counting available states in the $\mathbf{k}$ space for an electron enclosed in a cubic box of volume V. Show that the result is independent of the volume V.

3.2 Determine the dependence of the Fermi level E_{fc} on the electron density n by evaluating the integral in Eq. (3.2.27) under the approximation that the Fermi-level lies within the bandgap such that $E - E_{fc} \gg k_B T$.

3.3 Show that Eq. (3.2.37) for the spontaneous emission rate follows from Eq. (3.2.36).

3.4 Explain Auger nonradiative recombination. How many kinds of Auger processes are possible in semiconductor lasers? Use diagrams as necessary.

3.5 Perform the integration in Eq. (3.3.17) for the Auger recombination rate and show that the final result is given by Eq. (3.3.27).

3.6 Explain carrier lifetime and differential carrier lifetime. How can the measurement of differential carrier lifetime be used to determine the Auger rate?

3.7 Derive an expression for the threshold current density by including radiative and nonradiative recombination processes. Calculate it for a 250-μm-long semiconductor laser with 30% facet reflectivity and 0.1-μm thick active layer. Assume an internal loss of 30 cm^{-1}. Assume that the mode-confinement factor has a value of 0.2.

3.8 The threshold current of a semiconductor laser doubles when its temperature is increased by 40°C. Determine the characteristic temperature T_0.

3.9 Why is the characteristic temperature lower for an InGaAsP laser compared with that of GaAs lasers?

3.10 Derive Eq. (3.6.3) by considering multiple reflections in a Fabry-Perot cavity (see Sec. 2.3).

3.11 External differential quantum efficiency of semiconductor lasers generally decreases with an increase in the temperature. Give reasons for such a decrease with physical justification in each case.

3.12 The carrier lifetime of a semiconductor laser decreases by a factor of 2 when its temperature increases by 50°C. Calculate the increase in the Auger rate and the threshold current by assuming that the dominant contribution to the carrier recombination rate comes from the Auger process. You can assume that the transparency value of the carrier density remains unchanged with temperature.

REFERENCES

1. Einstein, A. *Phy. Z.* **18,** 121 (1917).
2. McCumber, D. E. *Phys. Rev.* **136,** A954 (1964).
3. Bernard, M. G. A., and G. Duraffourg. *Phys. Status Solidi,* **1,** 699 (1961).
4. Schiff, L. I. *Quantum Mechanics,* Third Ed, pp. 279–286. New York: McGraw-Hill Book Co., 1968.
5. Merzbacher, E. *Quantum Mechanics,* Second Ed., Chap. 18, Sec. 4. New York: John Wiley & Sons, 1970.
6. Lasher, G., and F. Stern. *Phys. Rev.* **133,** A553 (1964).
7. Kane, E. O. *J. Phys. Chem. Solids* **1,** 249 (1957).
8. Joyce, W. B., and R. W. Dixon. *Appl. Phys. Lett.* **31,** 354 (1977).
9. Nilson, N. *Appl. Phys. Lett.* **33,** 653 (1978).
10. Kane, E. O. *Phys. Rev.* **131,** 79 (1963).
11. Halperin, B. I., and M. Lax. *Phys. Rev.* **148,** 722 (1966).
12. Hwang, C. J. *Phys. Rev. B* **2,** 4126 (1970).
13. Stern, F. *Phys. Rev.* **148,** 186 (1966).
14. Stern, F. *Phys. Rev. B* **3,** 2636 (1971).
15. Stern, F. *Phys. Rev. B* **3,** 3559 (1971).
16. Stern, F. *J. Appl. Phys.* **47,** 5382 (1976).
17. Casey, H. C., Jr., and M. B. Panish. *Heterostructure Lasers,* Part A, Chap. 3. New York: Academic Press, 1978.
18. Dutta, N. K. *J. Appl. Phys.* **51,** 6095 (1980).
19. Dutta, N. K. *J. Appl. Phys.* **52,** 55 (1981).
20. Takagi, T. *Jpn. J. Appl. Phys.* **18,** 2017 (1979).
21. Osinski, M., and M. J. Adams. *IEE Proc. I* **129,** 229 (1982).
22. Hermann, C., and C. Weisbuch. *Phys. Rev. B* **15,** 823 (1977).
23. Chadi, D. J., A. H. Clark, and R. D. Burnham. *Phys. Rev. B* **13,** 4466 (1976).
24. Nicholas, R. J., J. C. Portal, C. Houlbert, P. Perrier, and T. P. Pearsall. *Appl. Phys. Lett.* **34,** 492 (1979).

25. Nahory, R. E., M. A. Pollack, W. D. Johnston, J., and R. L. Burns. *Appl. Phys. Lett.* **33,** 659 (1978).
26. Hermann, C., and T. P. Pearsall. *Appl. Phys. Lett.* **38,** 450 (1981).
27. Dutta, N. K., and R. J. Nelson. *J. Appl. Phys.* **53,** 74 (1982).
28. Olshansky, R., C. B. Su, J. Manning, and W. Powaziik. *IEEE J. Quantum Electron.* **QE-20,** 838 (1984).
29. Beattie, A. R., and P. T. Landsberg. *Proc. R. Soc. London Ser. A* **249,** 16 (1959).
30. Horikoshi, Y., and Y. Furukawa. *Jpn. J. Appl. Phys.* **18,** 809 (1979).
31. Thompson, G. H. B., and G. D. Henshall. *Electron. Lett.* **16,** 42 (1980).
32. Dutta, N. K., and R. J. Nelson. *Appl. Phys. Lett.* **38,** 407 (1981).
33. Sugimura, A. *IEEE J. Quantum Electron.* **QE-17,** 627 (1981).
34. Beattie, A. R., and G. Smith. *Phys. Status Solidi* **19,** 577 (1967).
35. Takashima, M. *J. Appl. Phys.* **43,** 4114 (1972).
36. Takashima, M. *J. Appl. Phys.* **44,** 4717 (1973).
37. Sugimura, A. *J. Appl. Phys.* **51,** 4405 (1980).
38. Eagles, D. M. *Proc. Phys. Soc. London* **78,** 204 (1961).
39. Lochmann, W. *Phys. Status Solidi A* **40,** 285 (1977).
40. Lochmann, W. *Phys. Status Solidi A* **42,** 181 (1977).
41. Huldt, L. *Phys. Status Solidi A* **33,** 607 (1976).
42. Bess, L. *Phys. Rev.* **105,** 1469 (1957).
43. Sinha, K. P., and M. DiDomenico. *Phys. Rev. B* **1,** 2623 (1970).
44. Bonch-Bruevich, V. L., and T. V. Gulyev. *Sov. Phys.–Solid State* **2,** 431 (1960).
45. Landsberg, P. T., C. Rhys-Roberts, and P. Lal. *Proc. Phys. Soc. London* **84,** 915 (1964).
46. Antoncik, E., and P. T. Landsberg. *Proc. Phys. Soc. London* **82,** 337 (1963).
47. Brand, S., and R. A. Abram. *J. Phys. C: Solid State Phys.* **17,** L201 (1984).
48. Burt, M. G., and C. Smith. *J. Phys. C: Solid State Phys.* **17,** L47 (1984).
49. Haug, A. *Appl. Phys. Lett.* **42,** 512 (1983).
50. Chelikowsky, J. R., and M. L. Cohen. *Phys. Rev. B* **14,** 556 (1976).
51. Sugimura, A. *Appl. Phys. Lett.* **39,** 21 (1981).
52. Conwell, E. *Solid State Phys. Suppl.* **9,** 105 (1967).
53. Sermage, B., H. J. Eichler, J. Heritage, R. J. Nelson, and N. K. Dutta. *Appl. Phys. Lett.* **42,** 259 (1983).
54. Henry, C. H., B. F. Levine, R. A. Logan, and C. G. Bethea. *IEEE J. Quantum Electron.* **QE-19,** 905 (1983).
55. Wintner, E., and E. P. Ippen. *Appl. Phys. Lett.* **44,** 999 (1984).
56. Burstein, E. *Phys. Rev.* **93,** 632 (1954).
57. Mozer, A., K. M. Romanek, O. Hildebrand, W. Schmid, and M. H. Pilkhun. *IEEE J. Quantum Electron.* **QE-19,** 913 (1983).
58. Su, C. B., J. Schlafer, J. Manning, and R. Olshansky. *Electron. Lett.* **18,** 595 (1982).
59. Su, C. B., J. Schlafer, J. Manning, and R. Olshansky. *Electron. Lett.* **18,** 1108 (1982).
60. Thompson, G. H. B. *Electron. Lett.* **19,** 154 (1983).
61. Uji, T., K. Iwamoto, and R. Lang. *IEEE Trans. Electron. Devices* **ED-30,** 316 (1983).
62. Agrawal, G. P., and N. K. Dutta. *Electron. Lett.* **19,** 154 (1983).
63. Ikegami, T. *IEEE J. Quantum Electron.* **QE-8,** 470 (1972).
64. Rode, D. L. *J. Appl. Phys.* **45,** 3887 (1974).
65. Goodwin, A. R., J. R. Peters, M. Pion, G. H. B. Thompson, and J. E. A. Whiteaway. *J. Appl. Phys.* **46,** 3126 (1975).
66. Casey, H. C., Jr. *J. Appl. Phys.* **49** (1978).
67. Anthony, P. J., and N. E. Schumaker. *IEEE Electron Device Lett.* **EDL-1,** 58 (1980); *J. Appl. Phys.* **51,** 5038 (1980).

68. Yano, M., H. Imai, K.-I. Hori, and M. Takusagawa. *IEEE J. Quantum Electron.* **QE-17,** 619 (1981).
69. Yano, M., H. Imai, and M. Takusagawa. *J. Appl. Phys.* **52,** 3172 (1981); *IEEE J. Quantum Electron.* **QE-17,** 1954 (1981).
70. Yano, M., Y. Nishitani, K.-I. Hori, and M. Takusagawa. *IEEE J. Quantum Electron.* **QE-19,** 1319 (1983).
71. Dutta, N. K. *J. Appl. Phys.* **52,** 70 (1981).
72. Dutta, N. K., and R. J. Nelson. *IEEE J. Quantum Electron.* **QE-18,** 871 (1982).
73. Horikoshi, Y., and Y. Furukawa. *Jpn. J. Appl. Phys.* **18,** 809 (1979).
74. Ettenberg, M., C. J. Nuese, and H. Kressel. *J. Appl. Phys.* **50,** 2949 (1979).
75. Yano, M., H. Nishi, and M. Takusagawa. *IEEE J. Quantum Electron.* **QE-15,** 571 (1979).
76. Nahory, R. E., M. A. Pollack, and J. C. DeWinter. *Electron. Lett.* **15,** 695 (1979).
77. Thompson, G. H. B., and G. D. Henshall. *Electron. Lett.* **65,** 42 (1980).
78. Arai, S., Y. Suematsu, and Y. Itaya. *IEEE J. Quantum Electron.* **QE-16,** 197 (1980).
79. Dutta, N. K., R. J. Nelson, and P. A. Barnes. *Electron. Lett.* **16,** 653 (1980).
80. Horikoshi, Y., Chap. 15 in *GaInAsP Alloy Semiconductors*, ed. T. P. Pearsall. New York: John Wiley & Sons, 1982.
81. Dutta, N. K., P. D. Wright, R. J. Nelson, R. B. Wilson, and P. R. Besomi. *IEEE J. Quantum Electron.* **QE-18,** 1414 (1982).
82. Chraplyvy, A. R., I. P. Kaminow, and A. G. Dentai. *IEEE J. Quantum Electron.* **QE-19,** 1335 (1983).
83. Göbel, E. O., G. Luz, and E. Schlosser. *IEEE J. Quantum Electron.* **QE-15,** 697 (1979).
84. Jung, H., E. O. Göbel, K. M. Romanek, and M. H. Pilkuhn. *Appl. Phys. Lett.* **39,** 468 (1981).
85. Greene, P. D., and G. D. Henshall. *IEE J. Solid State Electron Devices* **3,** 174 (1979).
86. Hsieh, J. J. *IEEE J. Quantum Electron.* **QE-15,** 694 (1979).
87. Yamakoshi, S., T. Sanada, O. Wada, I. Umebu, and T. Sakurai. *Appl. Phys. Lett.* **40,** 144 (1982).
88. Chen, T. R., S. Margalit, U. Koren, K. L. Yu, L. C. Chiu, A. Hasson, and A. Yariv. *Appl. Phys. Lett.* **42,** 1000 (1983).
89. Chen, T. R., B. Chang, L. C. Chiu, K. L. Yu, S. Margalit, and A. Yariv. *Appl. Phys. Lett.* **43,** 217 (1983).
90. Chiu, L. C., K. L. Yu, S. Margalit, T. R. Chen, U. Koren, A. Hasson, and A. Yariv. *IEEE J. Quantum Electron.* **QE-19,** 1335 (1983).
91. Thompson, G. H. B. *IEE Proc. I* **128,** 37 (1981).
92. Dutta, N. K., and R. J. Nelson. *IEEE Trans. Electron Devices* **ED-28,** 1222 (1981).
93. Sugimura, A. *IEEE J. Quantum Electron.* **QE-17,** 441 (1981).
94. Heinen, J., H. Albrecht, and C. Weyrich. *J. Appl. Phys.* **53,** 1800 (1982).
95. Haug, A. *IEEE J. Quantum Electron.* **QE-21,** 716 (1985).
96. Adams, A. R., M. Asada, Y. Suematsu, and S. Arai. *Jpn. J. Appl. Phys.* **19,** L621 (1980).
97. Asada, M., A. R. Adams, K. E. Stubkjaer, Y. Suematsu, Y. Itaya, and S. Arai. *IEEE J. Quantum Electron.* **QE-17,** 611 (1981).
98. Asada, M., and Y. Suematsu. *IEEE J. Quantum Electron.* **QE-19,** 917 (1983).
99. Hakki, B. W., and T. L. Paoli. *J. Appl. Phys.* **46,** 1299 (1975).
100. Göbel, E. O., Chap. 13 in *GaInAsP Alloy Semiconductors*, ed. T. P. Pearsall. New York: John Wiley & Sons, 1982.
101. Henry, C. H., R. A. Logan, F. R. Merritt, and J. P. Luongo. *IEEE J. Quantum Electron.* **QE-19,** 947 (1983).
102. Casey, H. C., Jr., and P. L. Carter. *Appl. Phys. Lett.* **44,** 82 (1984).
103. Chiu, L. C., and A. Yariv. *Jpn. J. Appl. Phys.* **21,** L305 (1982).

104. Shah, J., R. F. Leheny, R. E. Nahory, and H. Temkin. *Appl. Phys. Lett.* **39,** 618 (1981).
105. Etienne, B., J. Shah, R. F. Leheny, and R. E. Nahory. *Appl. Phys. Lett.* **41,** 1018 (1982).
106. Henry, C. H., R. A. Logan, H. Temkin, and F. R. Merritt. *IEEE J. Quantum Electron.* **QE-19,** 941 (1983).
107. Manning, J., R. Olshansky, C. B. Su, and W. Powazinik. *Appl. Phys. Lett.* **43,** 134 (1983).
108. Casey, H. C., Jr. *J. Appl. Phys.* **56,** 1959 (1984).

Chapter 4
EPITAXY AND MATERIAL PARAMETERS OF InGaAsP

4.1 INTRODUCTION

The lattice-matched crystalline growth of one semiconductor over another is called *epitaxy*. The development of sophisticated epitaxial techniques has been of major significance in the development of high-quality, reliable semiconductor lasers. The commonly used techniques are liquid-phase epitaxy (LPE), vapor-phase epitaxy (VPE), and molecular-beam epitaxy (MBE). In LPE a saturated solution of the constituents of the layer to be grown is cooled while in contact with the substrate. In VPE, the epitaxial layer is grown by the reaction of gaseous elements or compounds at the surface of a heated substrate. The VPE technique has also been called chemical vapor deposition (CVD) depending on the constituents of the reactants. A variant of the same technique is metal-organic chemical vapor deposition (MOVPE), in which metal alkyls are used as the compound source. In MBE the epitaxial layer is grown by reacting atomic or molecular beams of the constituent elements (of the layer to be grown) with a crystalline substrate held at high temperature in an ultrahigh vacuum.

In this chapter we describe the three epitaxial techniques and their use for the fabrication of double-heterostructure semiconductor lasers. The apparatus and the growth methods used for LPE are discussed in Sec. 4.2 with particular emphasis on the growth of InGaAsP material on InP substrates. VPE and the variant technique, MOVPE, are considered in Secs. 4.3 and 4.4 respectively, while Sec. 4.5 presents the details of MBE. Lattice-mismatch effects such as misfit dislocations, which are common to all epitaxial techniques, are described in Sec. 4.6. Section 4.7 is devoted to a discussion of the material parameters of InGaAsP. Epitaxial growth of strained layers is discussed in Sec. 4.8.

4.2 LIQUID-PHASE EPITAXY

The LPE technique was first demonstrated in 1963.[1] Since then, it has been successfully used to fabricate various types of III–V compound semiconductor devices including injection lasers, light-emitting diodes, photodetectors, solar cells, bipolar transistors, and field-effect transistors. In LPE a supersaturated solution of the material to be grown is brought into contact with the substrate for a desired period of time. If the substrate is single crystalline and the material to be grown has nearly the same lattice constant as the substrate, some of the material precipitates on the substrate while maintaining the crystalline quality. The precipitated material forms a lattice-matched epitaxial layer on the surface of the substrate.

4.2.1 LPE Apparatus

Three basic types of growth apparatus have been used for LPE. They are (i) the tipping furnace,[1] in which the substrate is brought into contact with the solution by tipping the furnace; (ii) the vertical furnace,[2] in which the substrate is dipped into the solution; and (iii) the multibin furnace,[3,4] in which the substrate can be brought into contact with different solutions kept in successive bins. The multibin furnace type of growth apparatus is extensively used for the fabrication of laser structures that require the successive growth of several epitaxial layers.

Figure 4.1 shows the tipping furnace originally used by Nelson.[1] The substrate is held at one end of a graphite boat inside a quartz tube. The solution is at the other end of the graphite boat. A thermocouple connected to the boat is used to control the temperature of the furnace. A flow of

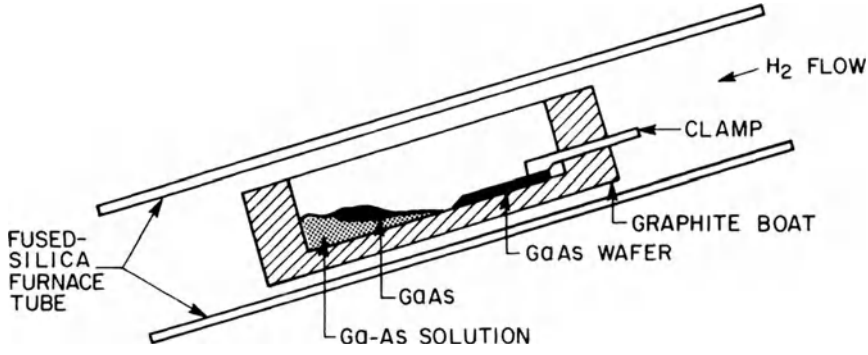


Fig. 4.1 Schematic illustration of the tipping surface used for liquid-phase epitaxy (LPE). (After Ref. 1)

hydrogen through the system prevents oxidation. The temperature of the furnace is reduced slowly; when the desired temperature is reached, the furnace is tipped so that the solution is in contact with the substrate, which allows an epitaxial layer to grow on the substrate. The furnace is then tipped back to the original position after the desired epitaxial layer thickness is obtained.

Figure 4.2 shows a vertical growth apparatus, in which the epitaxial layer is grown by dipping the substrate in a saturated solution. The solution is kept in a graphite or Al_2O_2 chamber. The substrate is held by a holder just above the solution. Growth can be started and terminated by dipping and withdrawing the substrate from the solution at a desired temperature.

Figure 4.3 shows the multibin-boat apparatus generally used for growing double-heterostructure lasers. The graphite boat has a number of reservoirs, each of which contains a saturated solution corresponding to the epitaxial layers to be grown. The substrate (seed) is placed in a graphite slider that has a groove to hold the substrate. The slider is attached to a long rod that allows an operator outside the furnace to position the substrate under different reservoirs. In this way, several epitaxial layers of different materials and desired thicknesses can be successfully grown on the substrate. Hydrogen or helium is generally used as the ambient gas during the growth process. Instead of a horizontal furnace and slider as shown in Fig. 4.3, a rotary

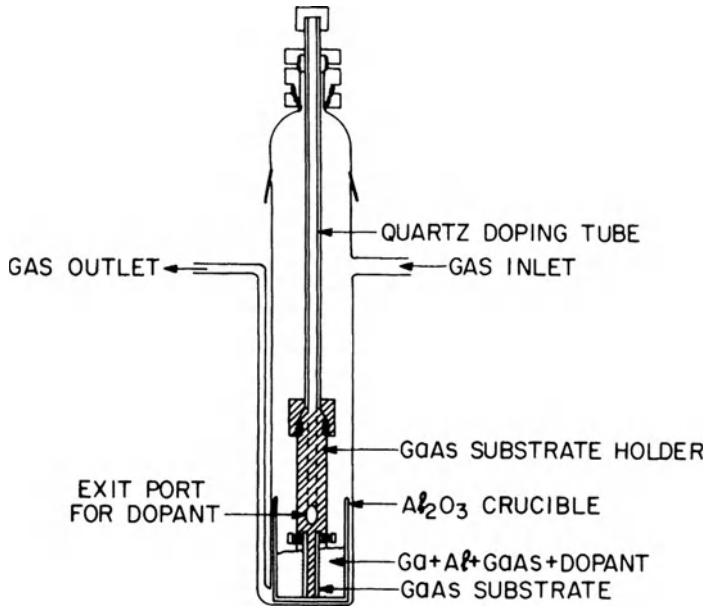


Fig. 4.2 Schematic illustration of a vertical LPE apparatus. (After Ref. 2)

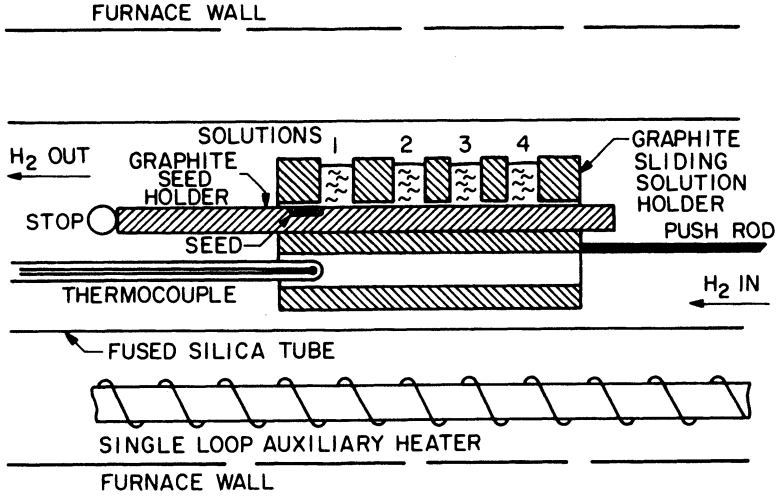


Fig. 4.3 Schematic illustration of a multibin-boat LPE apparatus used for growing several epitaxial layers. (After Ref. 4)

boat with a vertical furnace has also been used to grow several epitaxial layers successively on the substrate. In this case the graphite boat is in the form of a circular cylinder, and the substrate is moved from underneath one reservoir to another by rotating the disc-shaped slider that holds the substrate.

4.2.2 Growth Methods

Several LPE methods have been used to grow InGaAsP material.[5–8] They are step cooling,[5] equilibrium cooling,[5,8] supercooling,[7] and the two-phase method.[5,6] The kinetics of the LPE growth process have been studied extensively.[9–13]

In the step-cooling technique, the substrate and the growth solution are cooled to a temperature ΔT below the saturation temperature of the solution. The substrate is slid under the solution and a constant temperature is maintained during the growth period. The growth is terminated by sliding the substrate (with the epitaxially grown layer) out of the solution. The growth rate is determined by the diffusion rate of layer constituents from the solution to the substrate surface. The thickness d of the grown layer is related to ΔT and the growth time t by the relation[5,14]

$$d = K \Delta T t^{1/2} \tag{4.2.1}$$

where K is a constant that depends on the diffusivity of each solute and on the solute's mole fraction in the solution at the growth temperature.

In the equilibrium cooling technique, both substrate and solution are at the saturation temperature of the solution. Growth begins when the substrate is brought into contact with the solution and both are cooled at a uniform rate. The growth is terminated by sliding the substrate with the grown layer out of the solution. The thickness of the grown layer is given by[5,14]

$$d = \tfrac{2}{3}KRt^{3/2} \tag{4.2.2}$$

where t is the growth time and the cooling rate $R = dT/dt$.

The supercooling technique is a combination of step-cooling and equilibrium cooling. The substrate is brought into contact with the solution when both are at temperature ΔT below the saturation temperature of the solution. The growth solution and the substrate are further cooled during growth at a rate R. The thickness of the grown layer is given by the sum of Eqs. (4.2.1) and (4.2.2), i.e.,

$$d = K(\Delta T t^{1/2} + \tfrac{2}{3}Rt^{3/2}). \tag{4.2.3}$$

The value of ΔT used in the supercooling technique is generally smaller than that for the step-cooling technique.

In the two-phase technique, the cooling procedure is the same as in the equilibrium technique except that a piece of solid InP is added on top of the solution. The solid InP is in equilibrium with the solution during growth. The two-phase technique can in principle be used to grow very thin layers because the presence of the solid InP piece reduces the growth rate.

The experimental results on the material quality obtained by the above methods or their variations are available in the literature.[5–8,15] All of these methods produce good-quality wafers. Run-to-run variations, observed in an LPE process are influenced by wafer quality, surface preparation, and thermal decomposition prior to growth.

When an InP substrate is exposed to high temperatures ($\sim$650°C) occurring in an LPE reactor, the phosphorus evaporates from the surface, leaving behind In-rich regions. Epitaxy on this thermally decomposed surface has poor morphology and results in a low photoluminescence efficiency. Several methods have been used[16–20] to protect the substrate surface prior to growth by creating a phosphorus overpressure. These include (i) using a cover wafer in close proximity to the substrate to prevent large phosphorus loss[16]; (ii) providing excess phosphorus in the vicinity of the InP substrate using phosphorus powder[17] or phosphine gas[18]; and (iii) using an external

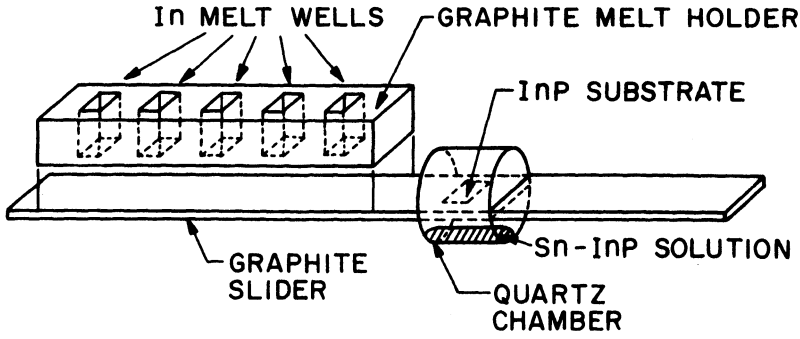


Fig. 4.4 Schematic illustration of a multibin LPE boat with a quartz chamber used to prevent thermal decomposition of the substrate prior to growth. (After Ref. 20)

chamber with phosphorus overpressure. For the last scheme, one of the common methods is the use of a graphite[19] or quartz[20] chamber with an Sn-InP solution. Figure 4.4 shows a schematic diagram of a multibin LPE boat with a quartz chamber. The substrate is kept inside this chamber prior to growth. The excess phosphorus overpressure created by the Sn-InP solution prevents any loss of phosphorus from the substrate surface.

4.2.3 LPE of InGaAsP

In order to grow lattice-matched layers of InGaAsP on InP, it is necessary to determine the phase diagram of In-Ga-As-P. The liquidus and solidus isotherms of this material system have been extensively studied.[21–31] Nakajima et al. determined the liquidus isotherm using the seed-dissolution technique.[24] An undersaturated solution with known amounts of In, Ga, and As obtained from In, InAs and GaAs was saturated with P at 650°C using an InP seed. The solution and the seed were kept in contact for $\sim$1 h. The amount of P in the equilibrated solution was calculated from the loss of weight of the seed after removal from the solution. Figure 4.5 shows the liquidus isotherms at 650°C for several atomic fractions of various elements in the liquid.[28]

In order to grow high-quality epitaxial layers, one needs the correct solution composition for lattice matching. The composition for lattice-matched growth is determined by growing $In_{1-x}Ga_xAs_yP_{1-y}$ layers on InP from In-Ga-As-P solutions whose compositions are determined by liquidus isotherms. The lattice constant of the grown layer is determined by the x-ray double-crystal diffraction technique. Precise lattice matching is verified using the substrate as a standard.

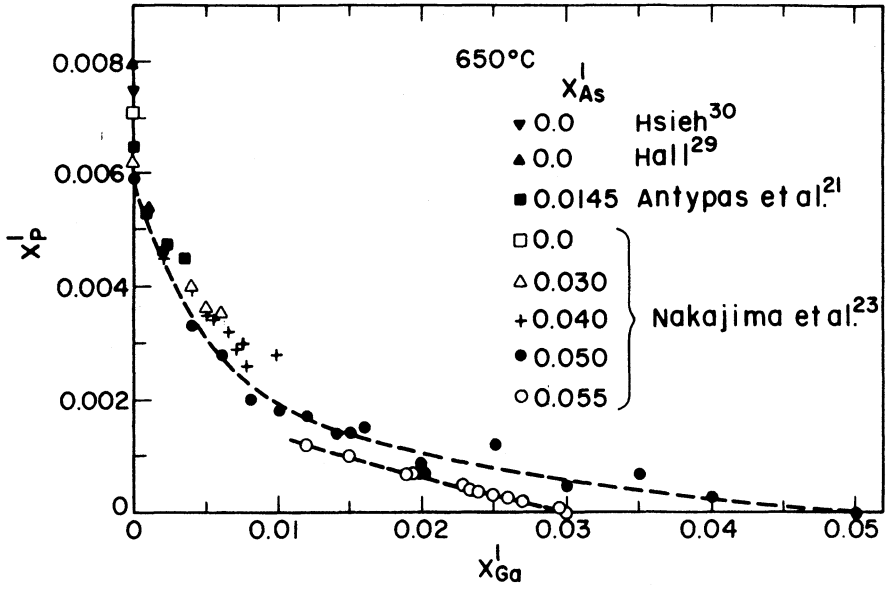


Fig. 4.5 Liquidus isotherms of the In-Ga-As-P alloy system. Atomic fractions of Ga and P are shown for several atomic fractions of As. Dashed curves are drawn for $x^l_{As} = 0.05$ and $x^l_{As} = 0.056$. (After Ref. 28)

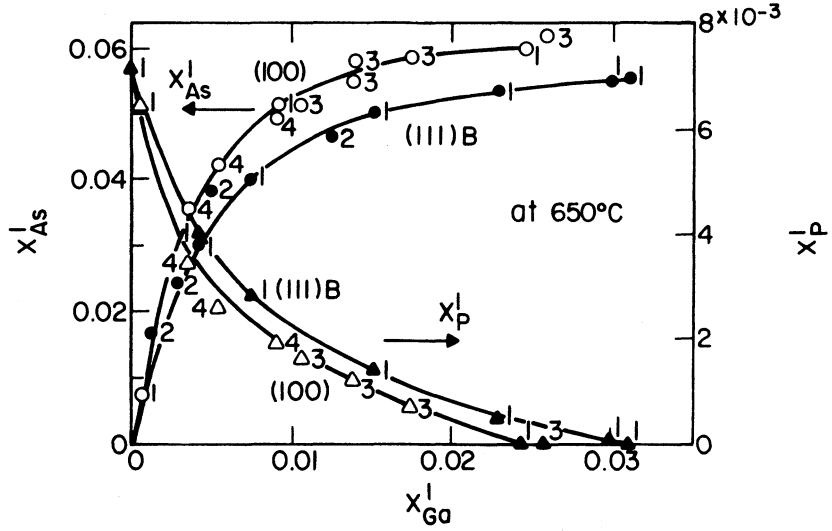


Fig. 4.6 The atomic fractions x^l_{As}, x^l_P, and x^l_{Ga} in the melt for the growth of lattice-matched $In_{1-x}Ga_xAs_yP_{1-y}$ on (111) B and (100) InP substrates at 650°C. Data points are from (1) Refs. 23 and 24, (2) Ref. 25, (3) Ref. 31, and (4) Ref. 27. (After Ref. 28)

The melt compositions needed for lattice-matched growth depend strongly on the orientation of the substrate. Nakajima et al.[24] and Sankaran et al.[25] have determined the melt composition for the lattice-matched growth on (111) B-oriented (P-rich planes) InP, while Nagai and Noguchi[31] and Feng et al.[27] have determined the melt compositions for (100)-oriented InP. Their results as compiled by Nakajima[28] are shown in Fig. 4.6. Pollack et al.[26] have reported similar measurements of the lattice-matched melt compositions for growth at 620°C on (100) InP.

The composition of the quaternary solid $In_{1-x}Ga_xAs_yP_{1-y}$ is determined using electron-microprobe analysis and wavelength-dispersive x-ray detection which uses a comparison of the intensities of the Ga–K_α, As–K_α and P–K_α lines from the grown sample and from GaP and InAs standards. Figure 4.7 shows the measured composition of the lattice-matched quaternary solid. The solid line is the calculated result on the basis of Vegard's law, according to which the lattice constant $a(x, y)$ of $In_{1-x}Ga_xAs_yP_{1-y}$ is given by

$$a(x, y) = xy\,a(\text{GaAs}) + x(1-y)\,a(\text{GaP}) + (1-x)y\,a(\text{InAs}) + (1-x)(1-y)\,a(\text{InP}). \quad (4.2.4)$$

Table 4.1 lists the room-temperature lattice constants of GaP, GaAs, InP, and InAs. Using $a(x, y) = a(\text{InP})$, Eq. (4.2.4) yields[32]

$$x = \frac{0.4526y}{1 - 0.031y} \quad (4.2.5)$$

Table 4.1 Band Gap and Lattice Constants of Binaries GaP, GaAs, InP, and InAs

COMPOUND	LATTICE, CONSTANT (Å)	REF.	BAND GAP E_g (eV) 0 K	REF.	300 K	REF.
GaP	5.45117	a	2.338	d	2.261	d
GaAs	5.65325	b	1.519	e	1.424	h
InP	5.86875	c	1.421	f	1.351	f
InAs	6.0584	c	0.420	g	0.360	i

[a] Casey, H. C., Jr., and M. B. Panish. *Heterostructure Lasers*, Part B. New York: Academic Press, 1978, p. 8.
[b] Driscoll, C. M. H., A. F. W. Willoughby, J. B. Mullin, and B. W. Straughan. *Gallium Arsenide and Related Compounds*, Inst. of Physics, London, 1975, p. 275.
[c] Giesecke, G., and H. Pfister. *Acta Crystallogr.* **11**, 369 (1958).
[d] Lorenz, M. R., G. D. Pettit, and R. C. Taylor. *Phys. Rev.* **171**, 876 (1968). GaP has indirect band gap.
[e] Sell, D. D., S. E. Stokowski, R. Dingle, and J. V. DiLorenzo. *Phys. Rev. B* **7**, 4568 (1973).
[f] Turner, W. J., and W. E. Reese. *Radiative Recombination in Semiconductors*. Paris: Dunod, 1965, p. 59.
[g] Pidgeon, C. R., D. L. Mitchell, and R. N. Brown. *Phys. Rev.* **154**, 737 (1967).
[h] Sell, D. D., and H. C. Casey, Jr. *J. Appl. Phys.* **45**, 800 (1974).
[i] Zwerdling, S., W. H. Kleiner, and J. P. Theriault. *J. Appl. Phys.* **32**, 2118 (1961).

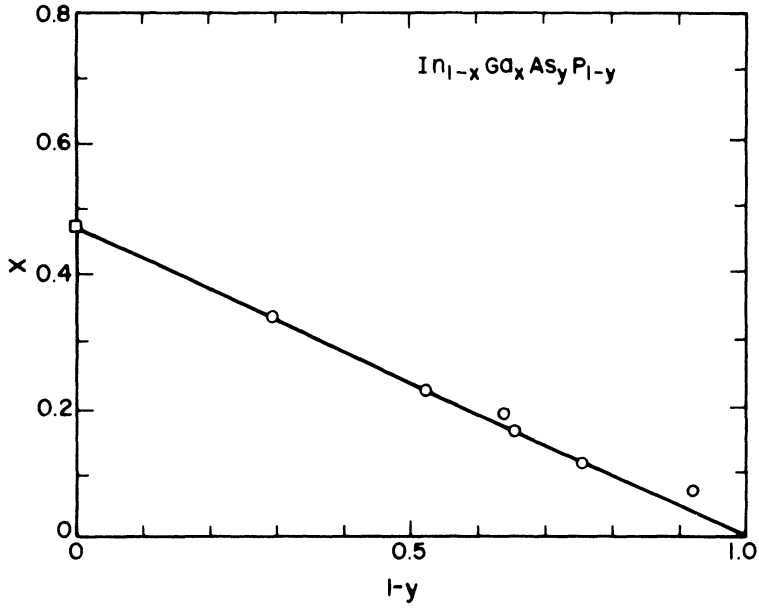


Fig. 4.7 Relationship between mole fractions x and y for the quaternary solid $In_{1-x}Ga_xAs_yP_{1-y}$ that is lattice-matched to InP. Data points are shown by open circles, while the straight line is the linear fit based on Vegard's law. (After Ref. 24)

as the the relation between the mole fractions x and y for the lattice-matched layer. Figure 4.7 shows that InGaAsP quaternary alloys obey Vegard's law.

Figure 4.8 shows the scanning-electron-microscope (SEM) cross section of a typical 1.3-μm InGaAsP double-heterostructure wafer grown by the

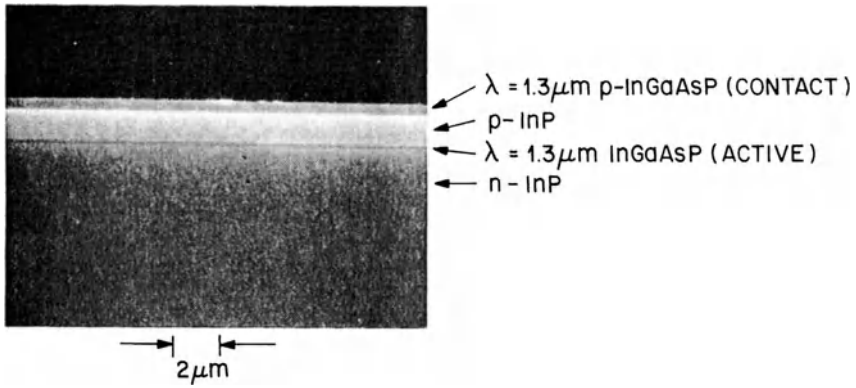


Fig. 4.8 Scanning-electron-microscope (SEM) cross section of an InGaAsP double-heterostructure laser grown by LPE.

equilibrium cooling technique on (100)-oriented *n*-type InP. The four layers grown are *n*-type InP (buffer layer), undoped InGaAsP ($\lambda = 1.3$ μm, active layer), *p*-type InP (cladding layer), and *p*-type InGaAsP ($\lambda = 1.3$ μm, contact layer).

4.3 VAPOR-PHASE EPITAXY

In vapor-phase epitaxy (VPE), the source chemicals from which the epitaxial layers are grown are gaseous. The technique has been widely used for the growth of several III–V compound semiconductors. VPE is often classified as one of two different methods, the chloride and the hydride techniques. In the *chloride method*, $AsCl_3$ or PCl_3 is passed over elemental Ga or In to form metal chlorides. These metal chlorides then react with AsH_3 or PH_3 near the InP substrate to form epitaxial layers of InGaAsP on InP. The metal chlorides can also be formed by using source pieces of GaAs or InP instead of elemental Ga or In. In the *hydride method*, metal chlorides are formed by passing HCl gas over hot In or Ga metal. VPE of InGaAsP material for laser devices has often used the hydride method. The chloride method is more useful for fabrication of the field-effect transistors because it produces layers with lower background doping levels.

Several workers have reported the growth of InGaAsP alloys by VPE.[33–41] A review by Olsen[33] describes in detail the growth systems they used. Tietjen and Amick[34] first reported the growth of GaAsP on GaAs using AsH_3 and PH_3. Figure 4.9 shows a schematic of the growth reactor for the growth of InGaAsP. The deposition is started by passing HCl gas over hot (850–900°C) In or Ga metal, which forms metal chlorides. Arsine and phosphine are mixed with the metal chlorides in the mixing zone. The InP

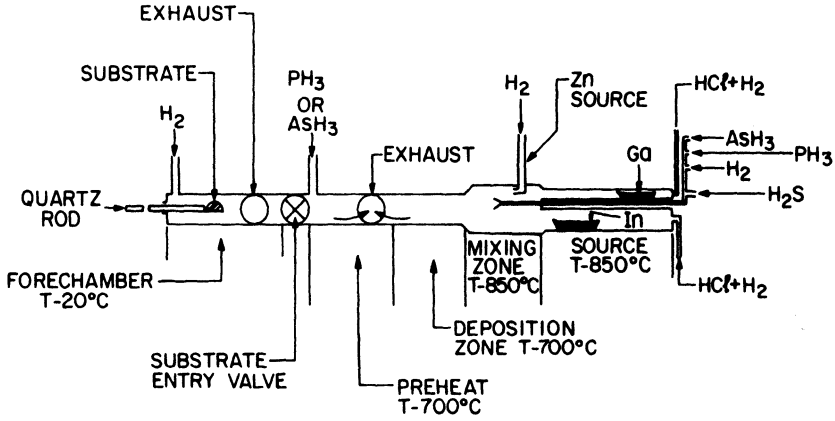


Fig. 4.9 Schematic illustration of a vapor-phase epitaxy (VPE) reactor. (After Ref. 33)

substrate is held in a chamber at room temperature that is flushed with H_2. Prior to growth, the substrate is heated to a temperature of $\sim 700°C$ in an atmosphere of AsH_3 and PH_3 to prevent surface decomposition. The substrate is then inserted into the growth chamber for a desired period of time. Here the growth takes place. The growth rate generally lies in the range of 0.1–1 μm/min. All reactant flows are controlled by precision mass-flow controllers. P-type doping (with Zn) can be accomplished by flowing H_2 over hot Zn metal, which carries Zn vapor to the mixing chamber. Similarly, *n*-type doping can be accomplished using 100 ppm H_2S along with AsH_3 and PH_3 in the mixing chamber.

The composition of the layers to be grown in a VPE reactor is changed by altering the flow rates of the reactants. It takes a certain amount of time to establish gaseous equilibrium in the growth chamber after the flow rates are changed. Because of this time delay, it is difficult to grow successive epitaxial layers with different compositions and abrupt interfaces in a one-chamber reactor of the form shown in Fig. 4.9.

For the growth of high-quality layers of different compositions, multi-chamber growth reactors have been used. Figure 4.10 shows a schematic of a dual-chamber growth reactor. In each chamber the gas-flow ratio, which is specific to a certain composition, is first established. Then the substrate is introduced into one chamber for the first growth to take place. For the second growth, the substrate is transferred to the other chamber. The transfer time is ~ 2 s. Heterojunction-interface widths of about 5–6 nm have been obtained using this technique.

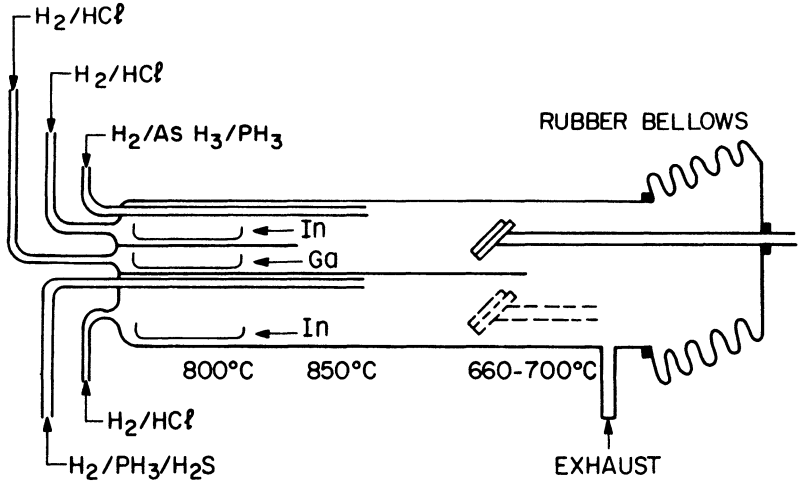


Fig. 4.10 Schematic illustration of a VPE dual-chamber growth reactor. (After Ref. 41)

4.4 METAL-ORGANIC VAPOR-PHASE EPITAXY

Metal-organic vapor-phase epitaxy (MOVPE), also known as metal-organic chemical vapor deposition (MOVPE), is a variant of the VPE technique that uses metal alkyls as sources from which the epitaxial layers form.[42,43] MOVPE has been extensively studied by Dupuis, Dapkus, and coworkers[44–46] using the AlGaAs material system. The low-pressure MOVPE technique, wherein the reaction takes place at a gas pressure of ~0.1 atm, has been used for the growth of InGaAsP by Hirtz, Razeghi, and coworkers.[47–51] Several other researchers have also reported on the growth of III–V compounds by MOVPE.[52–54]

Figure 4.11 shows a schematic of a low-pressure MOVPE system. Group III alkyls [$Ga(C_2H_5)_3$ and $In(C_2H_5)_3$] and group V hydrides [AsH_3 and PH_3] are introduced into a quartz reaction chamber that contains a substrate placed on a radio-frequency (RF) heated (~500°C) carbon susceptor. The gas flow near the substrate is laminar, with velocities in the range of 1–15 cm/s

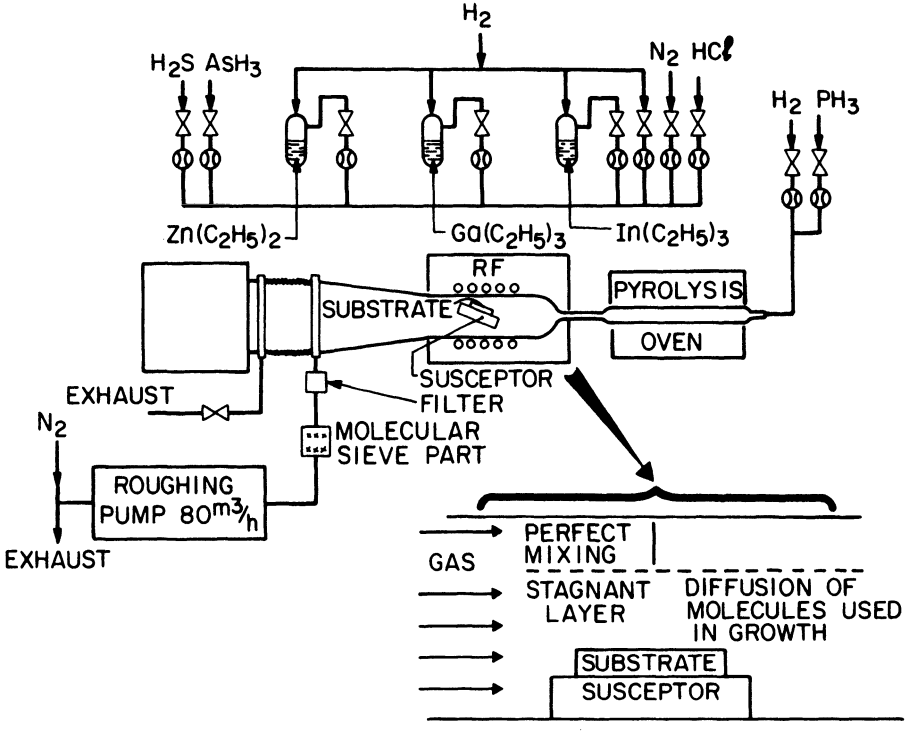


Fig. 4.11 Schematic illustration of a low-pressure MOVPE system. (After Ref. 47)

for a working pressure between 0.1 and 0.5 atm. It is thought[47] that a stagnant boundary layer is formed near the hot susceptor and gas molecules diffuse to the hot surface of the substrate. At the hot surface the metal alkyls and the hydrides decompose, producing elemental In, Ga, P, and As. The elemental species deposit on the substrate, forming an epitaxial layer. The gas-flow rates are controlled by mass-flow controllers. $Zn(C_2H_5)_3$ and H_2S are used as sources for *p*-type and *n*-type doping respectively.

In the low-pressure MOVPE technique, the velocities of the reactants are higher than those for hydride VPE. This allows quick changes in composition at heterojunction interfaces since new gas compositions are rapidly established. The growth rates are typically in the range of 2–4 μm/h. Dupuis et al.[55] in 1985 were able to grow high-quality InGaAsP double-heterostructures using an atmospheric pressure MOVPE system.

4.5 MOLECULAR-BEAM EPITAXY

In the molecular-beam epitaxial (MBE) technique, epitaxial layers are grown by impinging atomic or molecular beams on a heated substrate kept in an ultrahigh vacuum. The constituents of the beam "stick" to the substrate, resulting in a lattice-matched layer. The beam intensities can be separately controlled to take into consideration the difference in sticking coefficients of the various constituents of the epitaxial layers.

The widespread use of MBE for the growth of different III–V semiconductors resulted from the original work of Arthur and Cho.[56–61] The lattice-matched growth of AlGaAs on GaAs substrates by MBE was first reported in 1971.[58] Since then, extensive work on heterostructure lasers, microwave devices, quantum-well lasers, and superlattice structures has been reported using AlGaAs materials prepared by MBE.

Figure 4.12 shows a schematic of an MBE system for the growth of AlGaAs heterostructures. The substrate is heated by a Mo heating block inside a vacuum chamber at a pressure of 10^{-7}–10^{-10} torr. The electron-diffraction gun, the Auger analyzer, and the mass spectrometer can be used to study the layer characteristics during growth. The sources Ga, GaAs, and so forth are kept in independently heated effusion ovens enclosed in liquid-nitrogen-cooled shrouds. The dopants also have separate effusion ovens. The effusion ovens are heated to a temperature high enough to produce an adequate beam flux on the surface. For the growth of GaAs on AlGaAs, the Ga flux is in the range of 10^{12}–10^{14} atoms/(cm^2·sec). The flux needed depends on the material's *sticking coefficient* (the fraction of atoms in the beam sticking to the substrate). For many group III elements (e.g., Al or Ga) the sticking coefficient is nearly unity. However, the coefficient differs significantly for group V elements. This makes it necessary to determine

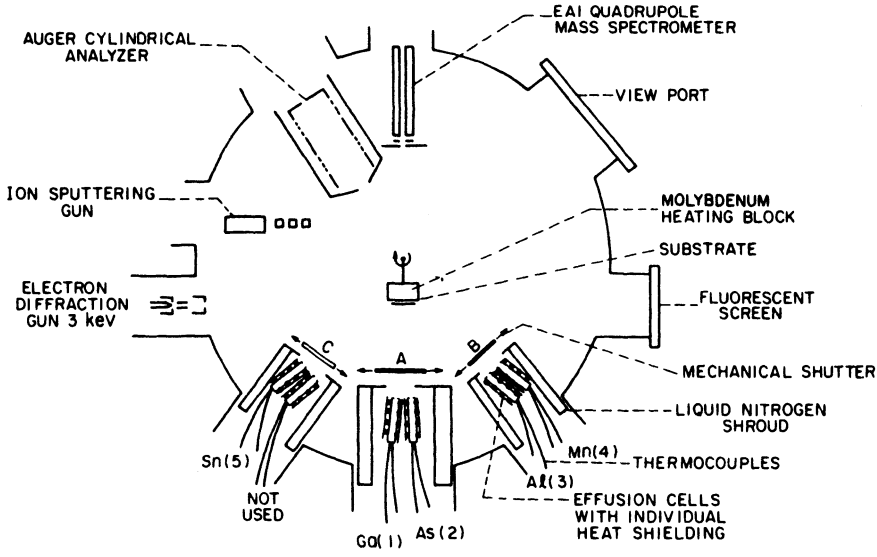


Fig. 4.12 Schematic illustration of a molecular-beam-epitaxy (MBE) system for epitaxial growth of AlGaAs. (After Ref. 59)

empirically the sticking coefficient for all the components as a function of substrate temperature and beam intensity.

Tsang et al.[62] reported on InGaAsP ($\lambda = 1.3$ μm) double-heterostructure lasers fabricated by MBE. Separate In and Ga beams were used, and the substrate temperature was approximately 580–600°C. To maintain the correct As-to-P ratio, a single beam of As and P_2 was used. It was obtained by passing As_4 and P_4 (evaporated from elemental arsenic and red phosphorus) through a common high-temperature zone.

Panish and coworkers[63,64] have demonstrated the growth of high-quality InGaAsP layers using gas sources for As_2 and P_2 beams. The use of gas sources for the growth of GaAs by MBE has also been reported.[65] The As_2 and P_2 molecules are generated by decomposing AsH_3 and PH_3 in a heated chamber and then allowing them to leak out to the effusion section connected to the MBE growth chamber. Figure 4.13 shows a schematic drawing of two types of gas sources. The major difference between the two types is that one uses separate decomposition tubes and the other uses the same tube. The decomposition tubes are operated at 900–1200°C and are filled with gases at pressures in the range of 0.3–2.0 atm. The tubes have a leaky seal at one end. At the operating temperatures and pressures stated above, AsH_3 and PH_3 decompose to produce As_4, and P_4, and H_2. The

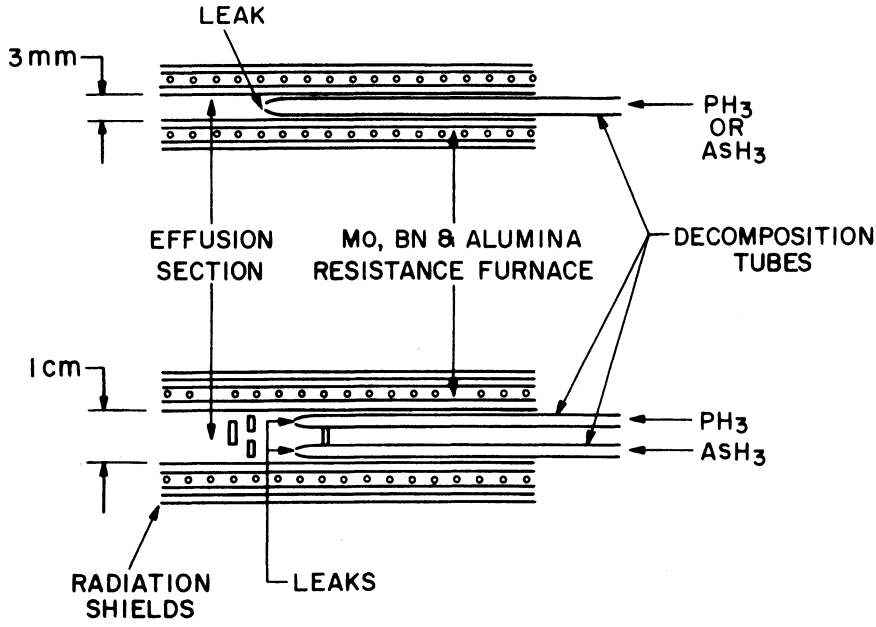


Fig. 4.13 Schematic illustration of two-types of high-pressure gas sources used for MBE growth of InGaAsP. (After Ref. 63)

molecules leak into a low-pressure (few millitorr), heated region where As_4 and P_4 decompose further into molecules of As_2 and P_2.

Panish[64] has used a low-pressure gas source where AsH_3 and PH_3 decompose at a pressure of less than 0.1 torr. Calawa[65] used similar low-pressure sources for MBE growth of high quality GaAs. Well-controlled molecular beams of As_2 and P_2 can be obtained[63] using gas sources of the type shown in Fig. 4.13. The beam intensity can be easily controlled by varying the pressure of AsH_3 and PH_3 in the decomposition chamber. The gas-source MBE technique has been used for the fabrication of high-quality InGaAsP double-heterostructure lasers.[66]

Another growth technique, known as chemical-beam epitaxy (CBE), has been demonstrated by Tsang[67] for the growth of InP and GaAs. In this technique, all sources are gaseous and are derived from group III and group IV alkyls. Figure 4.14 shows a schematic diagram of a CBE system. The growth chamber is similar to that of a conventional MBE system and is kept at high vacuum ($< 5 \times 10^{-4}$ torr). In and Ga are obtained by pyrolysis of either trimethylindium (TMIn) or triethylindium (TEIn), and trimethylgallium (TMGa) or triethylgallium (TEGa) at the heated substrate surface. The As_2 and P_2 molecules are obtained by thermal decomposition of

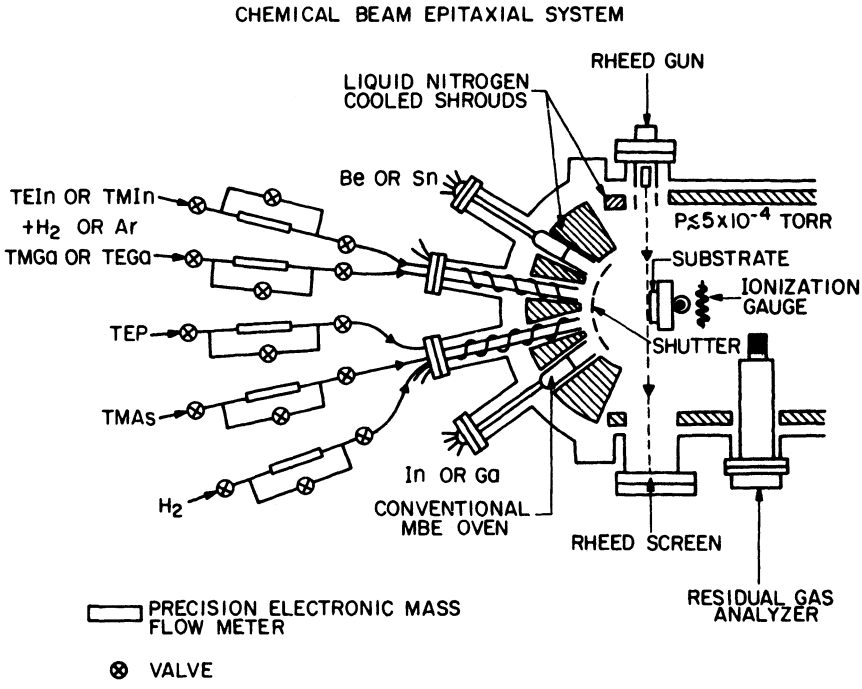


Fig. 4.14 Schematic illustration of a chemical-beam epitaxy (CBE) system obtained by modifying a conventional MBE apparatus (see Fig. 4.12). (After Ref. 67)

trimethylarsine (TMAs) and triethylphosphine (TEP) passing through a Ta- or Mo-buffered, heated alumina tube at a temperature of 950–1200°C. The CBE technique differs from MOVPE in the following way: in CBE the metal alkyls impinge directly on the heated surface, leaving the metal at a high substrate temperature, whereas in low-pressure MOVPE the metal alkyls along with H_2 form a stagnant boundary layer (as shown in Fig. 4.11) over the heated surface. Thus growth rate of the epitaxial layer for MOVPE is diffusion-limited, whereas for CBE it is limited by the flow rate of metal alkyls.

4.6 LATTICE-MISMATCH EFFECTS

A defect-free epitaxial growth of one crystal lattice over another takes place if the lattice constants of the two materials are nearly identical. In the presence of a small lattice mismatch (less than 0.1%), growth occurs with an approximate match of the lattice sites in the interface region of two lattices. This approximate match is possible if there is an elastic strain at the interface,

that is, each atom is slightly displaced from its original position at the boundary layer. Although a small amount of strain can be tolerated for thin layers and can even be beneficial in the case of quantum-well lasers (see Sec. 4.8), in general the strain energy stored in the crystal is reduced by the formation of a misfit dislocation (a missing row of atoms in one of the lattices). This is schematically illustrated in Fig. 4.15. If a is the lattice constant of the substrate and $a' = a - \Delta a$ is that of the grown layer, the separation between each row of missing atoms is approximately given by

$$L \cong a^2/\Delta a. \tag{4.6.1}$$

At the interface between the two crystal lattices, the rows of missing atoms exist along two perpendicular directions. The separation between the rows along a principal crystallographic axis (e.g., 100) is approximately given by Eq. (4.6.1). This type of imperfection at the interface is called a *dislocation*; and since it arises from lattice mismatch (or misfit), it is called *misfit dislocation.*

Near a misfit dislocation, the lattice is imperfect, containing many dangling bonds that cause nonradiative recombination of electrons and holes. Thus misfit-dislocation-free layers are needed to fabricate high-quality electro-optical devices. The generation of misfit dislocation depends on both the lattice mismatch and the thickness of the epitaxial layer grown. Oe et al.[68] found that misfit dislocations did not form in InGaAsP-InP double-heterostructure layers (0.4 μm thick) grown on (100) InP if the lattice mismatch $\Delta a/a$ was in the range of -5×10^{-3} to 5×10^{-3}.

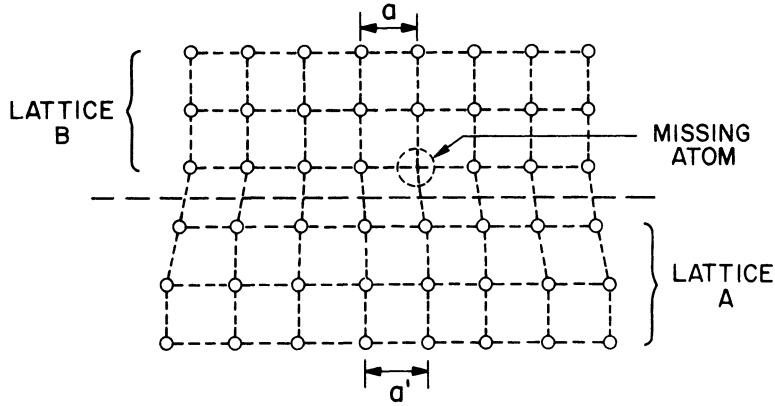


Fig. 4.15 Schematic illustration of a misfit dislocation formed during the epitaxial growth in the (100) plane of a cubic lattice. The two semiconductors have slightly different lattice constants.

Nakajima et al.[69] have studied the occurrence of misfit dislocations as a function of lattice mismatch for InGaAs layers of different thicknesses grown on (100) InP at 650°C. The measured data are shown in Fig. 4.16. The solid lines represent the boundary where no misfit dislocations were observed. For the growth of thick, dislocation-free InGaAs layers, the tolerable room-temperature lattice mismatch is found to lie between -6.5×10^{-4} and

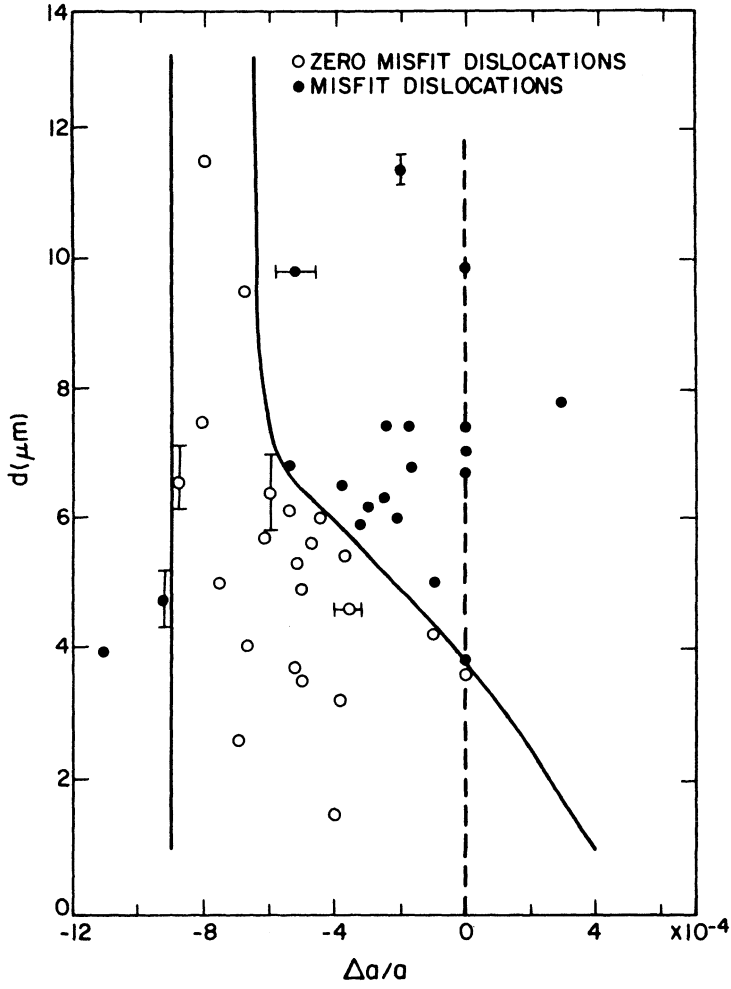


Fig. 4.16 Occurrence of misfit dislocations as a function of lattice mismatch for different layer thicknesses of InGaAs grown by LPE on (100) InP. No misfit dislocations were observed in the region bounded by solid lines. (After Ref. 69)

-9×10^{-4}. The reason for this negative lattice mismatch is due to different thermal-expansion coefficients of InGaAs and InP that introduce a negative room-temperature lattice mismatch for perfectly matched layers at the growth temperature of 650°C. Since misfit dislocations are formed near the growth temperature, lattice matching at the growth temperature is important to produce dislocation-free layers.

4.7 MATERIAL PARAMETERS

This section describes the numerical values of various material parameters of InGaAsP quaternary alloy grown lattice-matched on InP. A knowledge of the band-structure parameters, such as the band gap and the effective masses of the conduction and valence bands, is necessary to calculate the radiative and nonradiative Auger recombination rates discussed in Chapter 3. The low-field minority carrier mobilities are also useful for calculating the diffusion coefficient that plays an important role in device performance (see Sec. 2.4). Tables 4.1 to 4.3 list the band gap, the lattice constant, the effective masses, and the dielectric constant of the four binaries InP, GaAs, InAs, GaP that constitute the InGaAsP alloy. A review of the material parameters of $In_{1-x}Ga_xAs_yP_{1-y}$ has also been published.[70]

Table 4.2 Conduction-Band Effective Mass, Dielectric Constant, and Refractive Index of GaP, GaAs, InP, and InAs

	CONDUCTION-BAND EFFECTIVE MASS		dc DIELECTRIC CONSTANT		REFRACTIVE INDEX AT E_g	
COMPOUND	m_c/m_0	REF.	ϵ	REF.	μ	REF.
GaP	0.82	a	8.4	d	3.452	e
GaAs	0.067	b	13.1	d	3.655	f
InP	0.077	c	9.6	d	3.450	g
InAs	0.023	c	12.2	d	3.52	h

[a] Onton, A. *Phys. Rev.* **186,** 786 (1969).
[b] Vrehen, Q. H. F. *J. Phys. Chem. Solids* **29,** 129 (1968).
[c] Palik, E. D., and R. F. Wallis. *Phys. Rev.* **123,** 131 (1961).
[d] Ashcroft, N. W., and N. D. Mermin. *Solid State Physics.* New York: Holt, Rinehart & Winston, 1976, Chap. 27.
[e] Nelson, D. F., and E. H. Turner. *J. Appl. Phys.* **39,** 3337 (1968).
[f] Sell, D. D., H. C. Casey, Jr., and K. W. Wecht. *J. Appl. Phys.* **45,** 2650 (1974).
[g] Pettit, G. D., and W. J. Turner. *J. Appl. Phys.* **36,** 2081 (1965).
[h] Lorimor, O. G., and W. G. Spitzer. *J. Appl. Phys.* **36,** 1841 (1965).

Table 4.3 Valence-Band Effective Masses of GaP, GaAs, InP, and InAs

	HEAVY-HOLE MASS		LIGHT-HOLE MASS		SPLIT-OFF-HOLE MASS	
COMPOUND	m_{hh}/m_0	REF.	m_{lh}/m_0	REF.	m_{sh}/m_0	REF.
GaP	0.79	a	0.14	a	0.24	a
GaAs	0.45	b	0.082	b	0.15	a
InP	0.45	b	0.12	b	0.17	a
InAs	0.4	b	0.026	b	0.089	a

[a] Lawaetz, P. *Phys. Rev. B* **4**, 3460 (1971).
[b] Alavi, K., R. L. Aggarwal, and S. H. Groves. *J. Magn. Magn. Mater.* **11**, 136 (1979); R. L. Aggarwal, *Semiconductors and Semimetals*, Vol. 9, ed. R. K. Willardson and A. C. Beer. New York: Academic Press, 1972, p. 151.

4.7.1 Band-Structure Parameters

The band structure of InGaAsP is similar to that of InP or GaAs with a direct gap at the Γ-point in **k**-space. A simplified model of the band structure is the four-band model of Kane[71] shown in Fig. 3.1. Under photoexcitation, electrons and holes are created in the conduction and valence bands respectively; when they recombine, light at the band-gap energy E_g is emitted. This technique is called *photoluminescence*, and has been used to measure the direct band gap of $In_{1-x}Ga_xAs_yP_{1-y}$ lattice-matched to InP.[23,32] Figure 4.17 shows the measured data of Ref. 32. From the data, the room-temperature band gap of $In_{1-x}Ga_xAs_yP_{1-y}$ lattice-matched to InP is represented by the following expression:

$$E_g \text{ (in eV)} = 1.35 - 0.72y + 0.12y^2 \qquad \text{at 300 K.} \tag{4.7.1}$$

On the other hand, Nakajima et al.[28] have suggested the following linear interpolations of the band gap from their measured data:

$$E_g \text{ (in eV)} = \begin{cases} 0.74 + 0.61(1-y) & \text{at 300 K} \\ 0.80 + 0.61(1-y) & \text{at 77 K.} \end{cases} \tag{4.7.2}$$

The measurement of the spin-orbit splitting (Δ) at the zone center has been carried out[72] using electroreflectance. Figure 4.18 shows the measured data. The electroreflectance technique has also been used to measure the direct band gap, and the results agree well with those obtained using the photoluminescence technique.

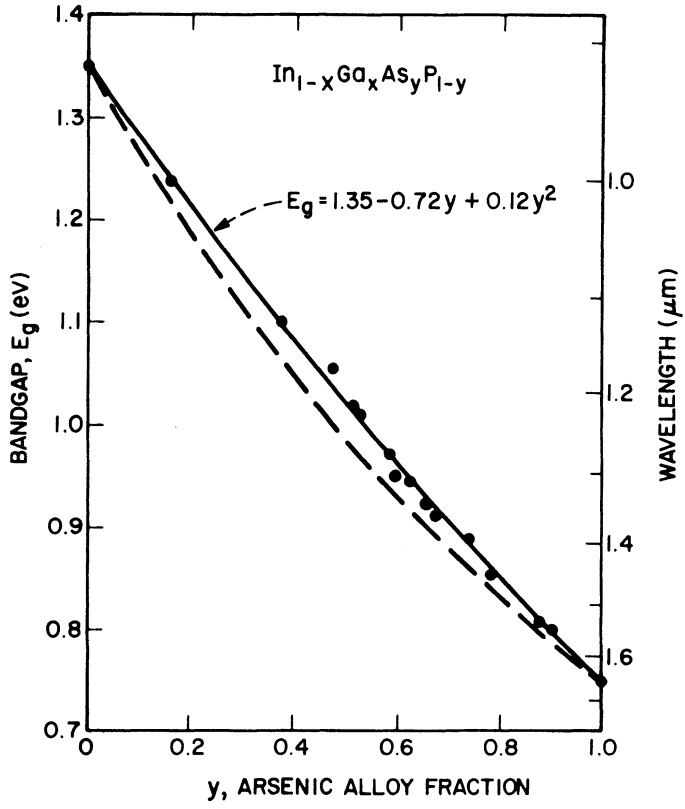


Fig. 4.17 Band gap of $In_{1-x}Ga_xAs_yP_{1-y}$ lattice-matched to InP obtained from photoluminescence measurements as a function of As fraction y. Solid line is a fit to the data points (circles), while dashed line shows interpolation using the binary values. (After Ref. 32)

The temperature dependence of the band gap ($\lambda \cong 1.3$ μm) of InGaAsP lattice-matched to InP has been obtained from measurements of the wavelength of a double-heterostructure injection laser as a function of the ambient temperature.[73] The result is shown in Fig. 4.19. As temperature increases, the band gap decreases at a rate of 0.325 meV/K.

The conduction-band-edge effective mass of $In_{1-x}Ga_xAs_yP_{1-y}$ lattice-matched to InP has been measured using techniques based on cyclotron resonance,[74,75] magneto-absorption,[76,77] the Shubnikov–de Haas effect,[78,79] and the magneto-photon effect.[74] Figure 4.20 shows the results of various measurements as compiled by Pearsall.[80] The results of these independent measurements agree very well. The following analytic expression has been

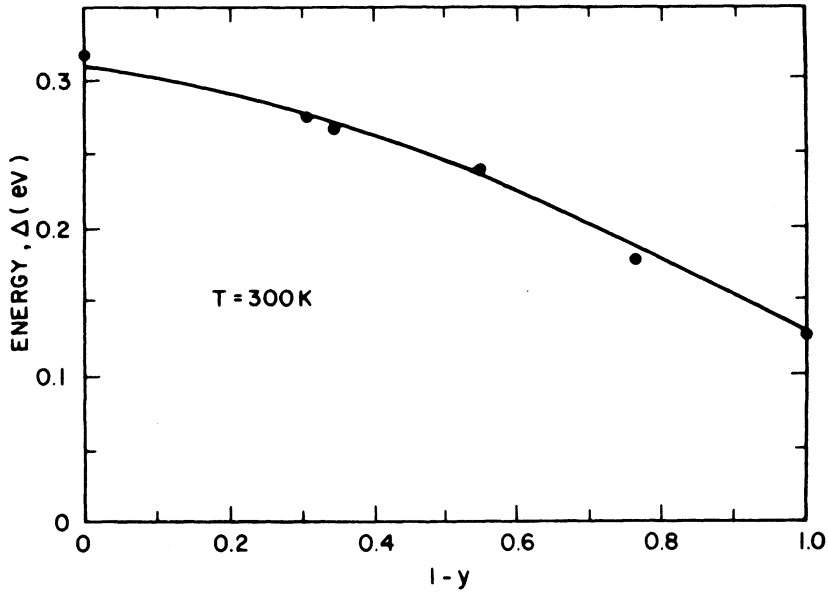


Fig. 4.18 Spin-orbit splitting (Δ) as a function of As fraction y for $In_{1-x}Ga_xAs_yP_{1-y}$ lattice-matched to InP. (After Ref. 72)

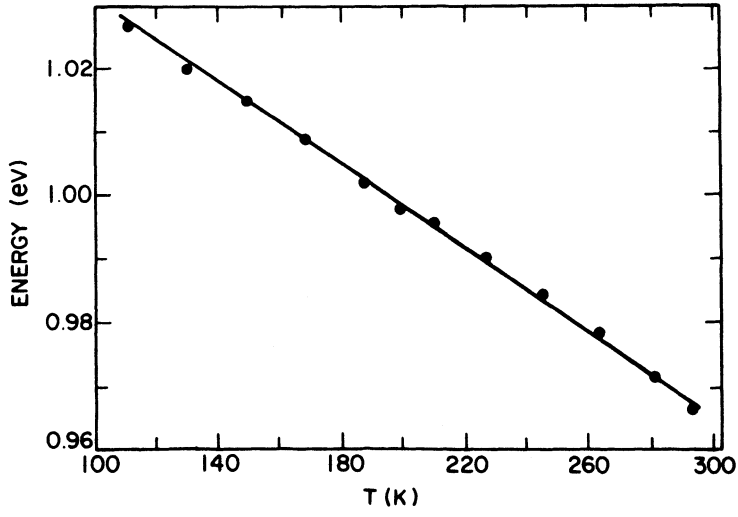


Fig. 4.19 Variation of band gap (obtained through measurements of the lasing wavelength) as a function of ambient temperature for 1.3-μm InGaAsP-InP double-heterostructure lasers. (After Ref. 73 © 1982 IEEE)

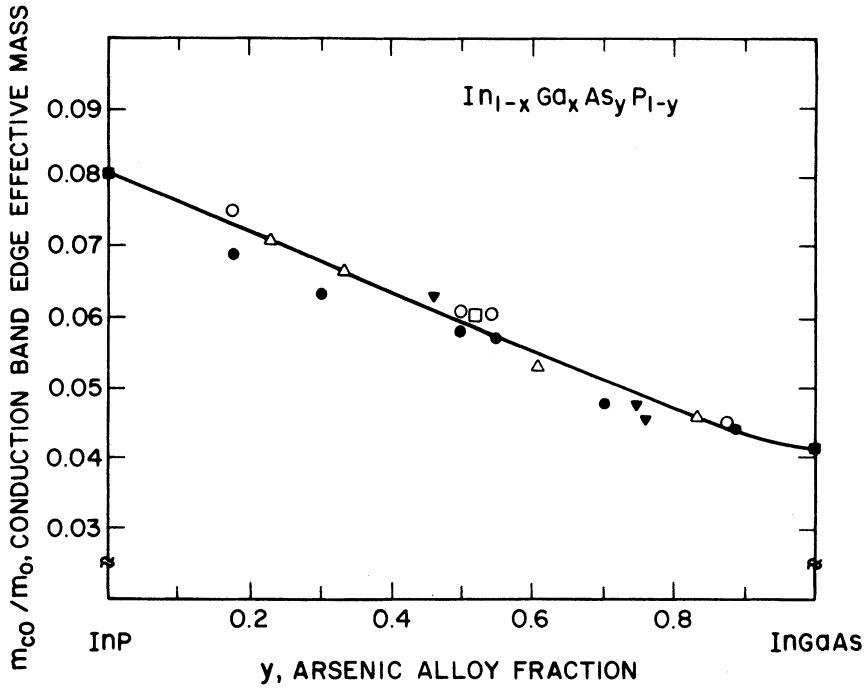


Fig. 4.20 Conduction-band-edge effective mass as a function of As fraction y for $In_{1-x}Ga_xAs_yP_{1-y}$ lattice-matched to InP. (After Ref. 80)

proposed[74] for the conduction-band-edge effective mass m_{c0} based on cyclotron measurements:

$$m_{c0}/m_0 = 0.080 - 0.039y \tag{4.7.3}$$

where m_0 is the free-electron mass and y is the As mole fraction. The valence-band effective masses of the InGaAsP alloy are not so well known because of difficulties in interpreting the experimental data. The heavy-hole mass and the split-off-band mass are generally estimated from an interpolation of the measured binary values. The light-hole effective mass of $In_{1-x}Ga_xAs_yP_{1-y}$ has been obtained using the optical pumping technique.[81] The measured data shown in Fig. 4.21 agree well with the value obtained from an interpolation of the measured binary values. The valence-band effective masses obtained from an interpolation of the binary values were used to calculate the radiative and nonradiative Auger recombination rates in Chapter 3.

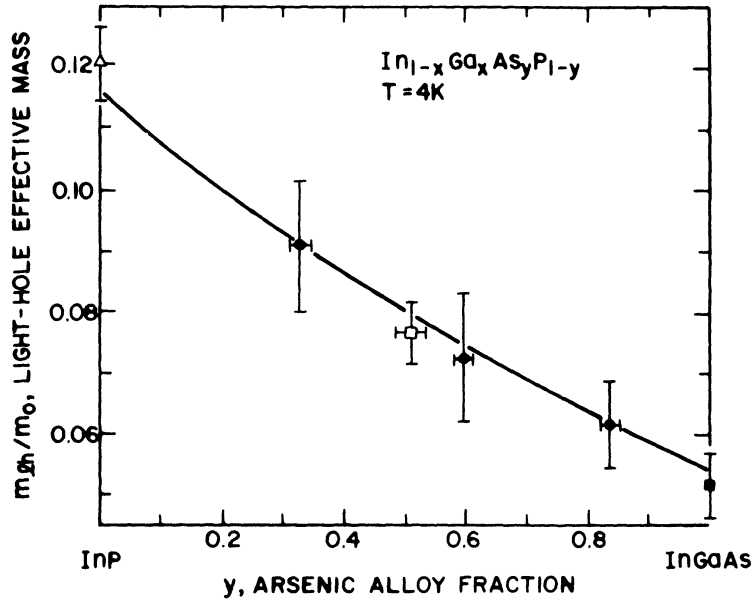


Fig. 4.21 Light-hole effective mass as a function of As fraction y for $In_{1-x}Ga_xAs_yP_{1-y}$ lattice-matched to InP. (After Ref. 81)

4.7.2 Mobility

The low-field carrier mobility is an important parameter in the performance of semiconductor injection lasers. The diffusion coefficient D is related to the mobility μ by the relation $D = \mu k_B T/q$. The diffusion of carriers can influence spatial hole-burning, multitransverse mode operation, and waveguiding in many semiconductor laser structures. The low-field carrier mobility in InGaAsP alloys has been extensively discussed.[82,83] The phenomena related to hot-electron effects and high-field transport have also been discussed in several review articles.[84,85] In this section we limit ourselves to a discussion of the measured low-field carrier mobility.

The room-temperature electron mobility of $In_{1-x}Ga_xAs_yP_{1-y}$ has been extensively studied.[86–89] Figure 4.22 shows the measured electron mobility as a function of the alloy composition in samples with a free-carrier concentration in the range of 1×10^{16}–4×10^{16} cm^{-3}. The measurements were done on single layers of InGaAsP grown by LPE on (100)-oriented semi-insulating InP. The mobility increases with decreasing temperature. The influence of various scattering mechanisms (such as ionized impurity scattering, optical

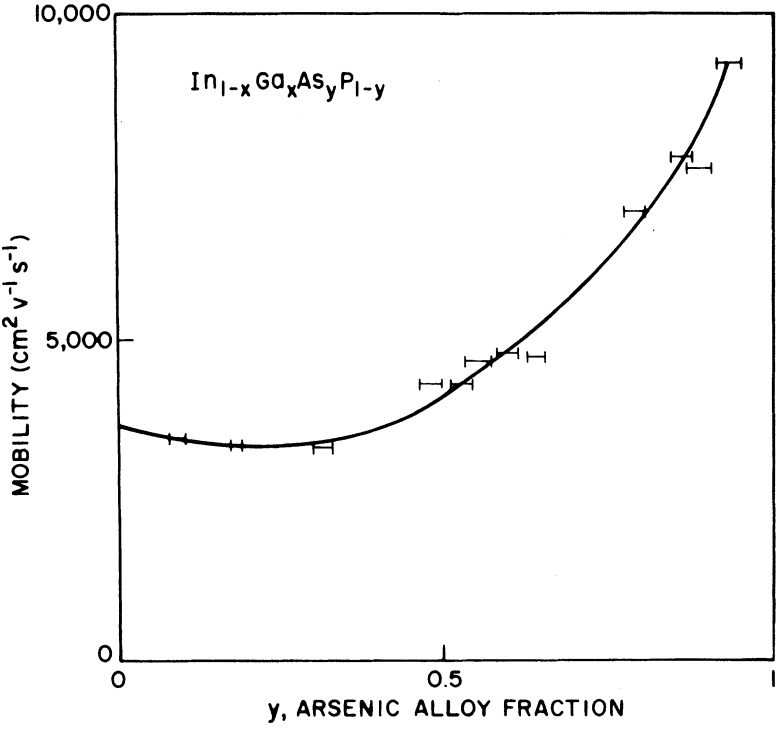


Fig. 4.22 Measured electron mobility as a function of $In_{1-x}Ga_xAs_yP_{1-y}$ alloy composition lattice-matched to InP. (After Ref. 86)

phonon scattering, and alloy scattering) on the temperature dependence of the mobility has also been studied.[87]

Hayes et al.[83,88] have measured the hole mobility as a function of the mole fractions x and y in $In_{1-x}Ga_xAs_yP_{1-y}$. Their experimental data are shown in Fig. 4.23. The samples used in the experiment were grown by LPE on (100)-oriented semi-insulating InP substrates. The shape of the mobility curve as a function of the alloy composition is principally governed by alloy scattering, which is large near the middle range of y ($y \cong 0.5$).

4.7.3 Refractive Index

As was seen in Chapter 2, a knowledge of the refractive index is important for understanding the waveguiding properties of semiconductor lasers. The InGaAsP double-heterostructures used for the fabrication of laser diodes generally have p-type InP and n-type InP as cladding layers. The wavelength

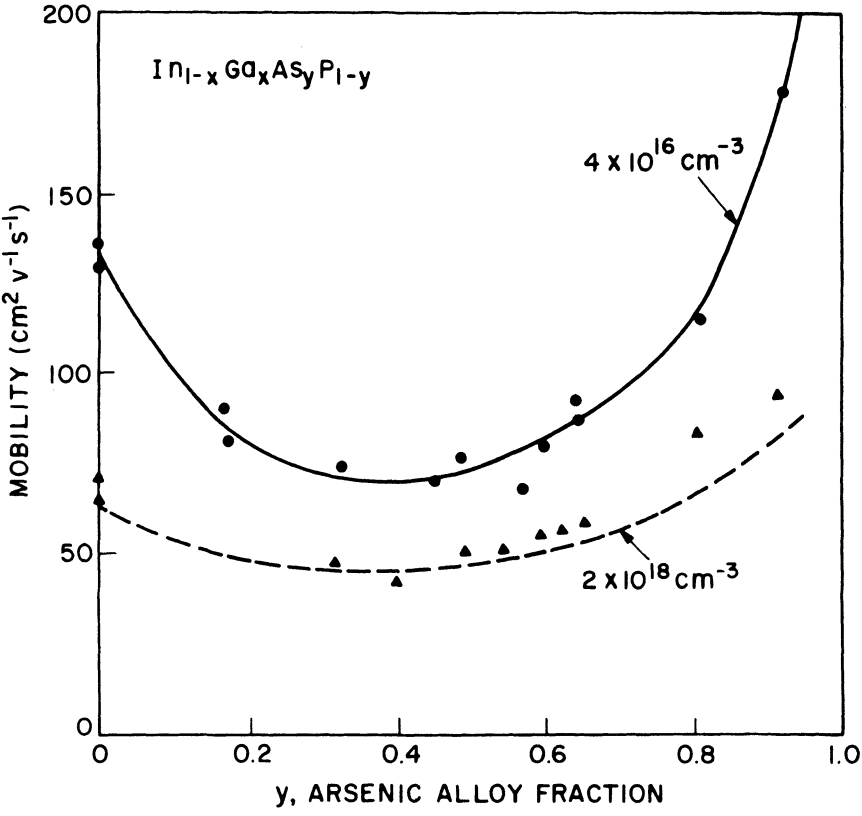


Fig. 4.23 Measured hole mobility as a function of alloy fraction y for $In_{1-x}Ga_xAs_yP_{1-y}$ lattice-matched to InP with two different doping levels. (After Ref. 83)

dependence of the refractive index of InP has been studied by Petit and Turner.[90] For InGaAsP, several interpolation schemes using the measured binary values have been proposed and compared.[91] Nahory and Pollack[92] have suggested the following analytic form for the refractive index (at a wavelength corresponding to the band gap) of $In_{1-x}Ga_xAs_yP_{1-y}$ lattice-matched to InP (for which $x \cong 0.45y$):

$$\mu(y) = 3.4 + 0.256y - 0.095y^2. \tag{4.7.4}$$

Values of the refractive indices have also been estimated using indirect methods such as far-field measurements and mode spacing of injection lasers.[93–95] All of these values agree quite well.

Chandra et al.[96] reported direct measurement of the refractive index of $In_{1-x}Ga_xAs_yP_{1-y}$ lattice-matched to InP. About 1-μm thick layers of InGaAsP were grown on InP by vapor-phase epitaxy. Transmission and reflection measurements were done on each sample as a function of the incident wavelength. The spectral line width (~ 5 Å) of the light source was large enough to average out forward-backward wave interference effects occurring in thick substrates. The refractive index was obtained from the observed modulation in reflection or transmission caused by interference of the light reflecting between the two surfaces of the epitaxial layer. Figure 4.24

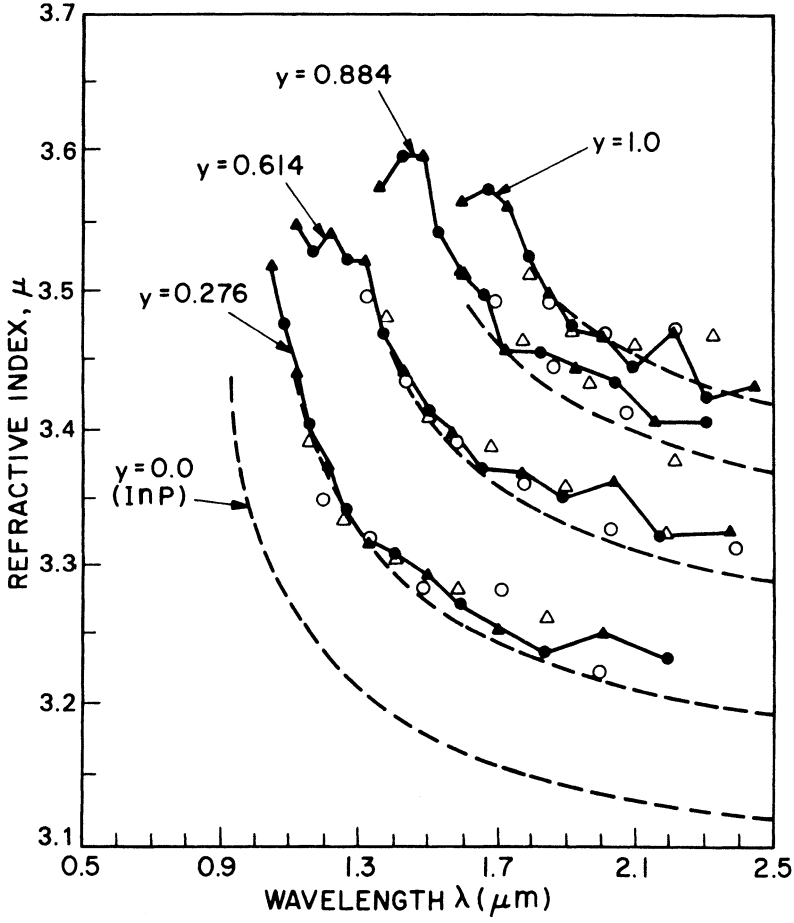


Fig. 4.24 Refractive index as a function of wavelength for four different As fractions (y) of $In_{1-x}Ga_xAs_yP_{1-y}$. Different symbols refer to measurements on different samples. Dashed curves are obtained using the interpolation scheme of Ref. 97. (After Ref. 96)

shows the measured refractive index as a function of wavelength for four lattice-matched compositions. The measured values agree well with the calculated values obtained using the interpolation method of Afromowitz.[97]

4.8 STRAINED-LAYER EPITAXY

Although the growth of lattice-matched layers is very important for the fabrication of reliable semiconductor lasers, it is possible to make high-quality semiconductor lasers using materials with a small degree of lattice mismatch among them.[98–100] This lattice mismatch introduces strain on the epitaxial layers, altering the semiconductor band structure. Typical values of tolerable strain ($\Delta a/a$, where a is the lattice constant of the substrate and Δa is the difference in lattice constants between the substrate and the epitaxial layer) are less than 1.5%. The larger the amount of strain, the thinner is the layer that can be grown free of dislocations on the substrate. Beyond a certain thickness, known as the critical thickness,[101] very large numbers of dislocations are generated, and the luminescence properties of the material become quite poor.

The two material systems that have been studied extensively for the fabrication of strained-layer lasers are (i) InGaAsP active region grown over InP and (ii) InGaAs active region grown over GaAs. The former emits in the wavelength range of 1.2–1.6 μm whereas the latter covers the wavelength range of 0.85–1.05 μm. Both types of strained-layer lasers have some advantages discussed in Sec. 9.9 in the context of quantum-well devices.

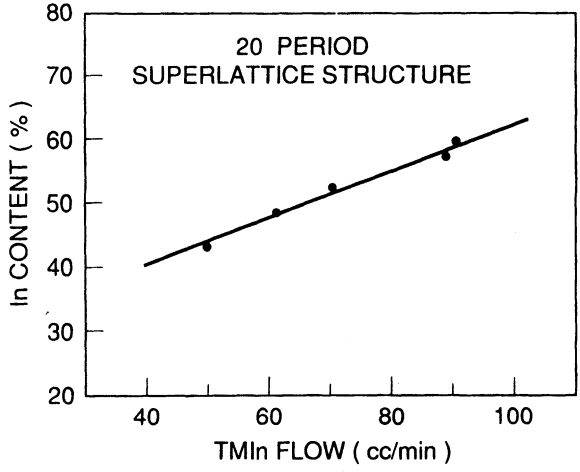


Fig. 4.25 Indium concentration in $In_{1-x}Ga_xAs$ epitaxial layer grown over InP as a function of TMIn flow. (After Ref. 102 © 1990 IEEE)

The InGaAsP semiconductor alloy can also be grown lattice-matched on InP as discussed in the previous sections. Non-lattice-matched layers have been grown mainly by using MOVPE and CBE growth techniques. Layer composition can be changed by varying the flow rate of one of the species. Changes in the indium concentration in the strained $In_{1-x}Ga_xAs$ solid as a function of trimethylindium (TMIn) flow are shown in Fig. 4.25.[102] For the layer lattice-matched to InP, the In content is 53%. The data were obtained from superlattice samples grown over the (100)-oriented InP substrate by atmospheric pressure MOVPE, containing 20–50 periods of strained quantum wells separated by 20-nm thick InP barriers.[102] Strained layers retain their structural integrity as judged by sharp and intense satellite peaks in x-ray diffraction measurements. The linear relationship between TMIn flux and solid-state In composition in $In_{1-x}Ga_xAs$ suggests TMIn is

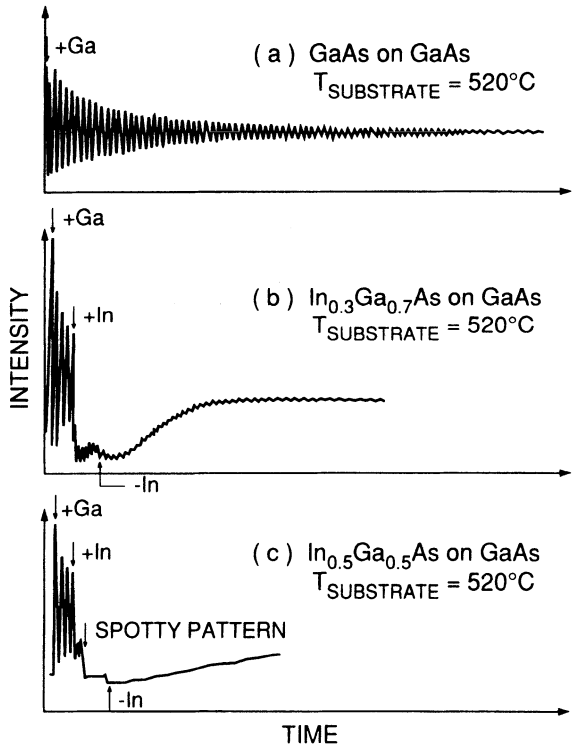


Fig. 4.26 RHEED-oscillation data for growth of (a) GaAs on GaAs, (b) $In_{0.3}Ga_{0.7}As$ on GaAs, and (c) $In_{0.5}Ga_{0.5}As$ on GaAs. "+" and "−" represent the addition or removal of those species, e.g., +Ga means Ga beam is turned on at the time indicated. (After Ref. 103)

completely decomposed at the growth temperature (625°C) and is proportionally incorporated in the epitaxial $In_{1-x}Ga_xAs$ layer.

Strained $In_{1-x}Ga_xAs$ layers have been grown on GaAs substrates using both MBE and MOVPE growth techniques. When the growth takes place by MBE, it is possible to monitor the growth front in situ using RHEED (reflection high-energy electron diffraction) oscillations. Figure 4.26 shows RHEED oscillations for the three cases (a) GaAs on GaAs, (b) $In_{0.3}Ga_{0.7}As$ on GaAs, and (c) $In_{0.5}Ga_{0.5}As$ on GaAs.[103] In the first case, which corresponds to the lattice-matched growth, RHEED oscillations show a smooth layer-by-layer growth. However, in the strained case (when the In source is turned on) the growth mode changes rapidly to one that suggests island formation, i.e., RHEED oscillations disappear. As the strain is increased (higher In mole fraction), the transition to the island-growth mode becomes more abrupt, as seen in Fig. 4.26c.

PROBLEMS

4.1 What is meant by epitaxial growth? Describe the three main techniques used for such a growth.

4.2 Explain how the multiple epitaxial layer can be grown over a substrate by using multibin-boat LPE apparatus. Use diagrams as necessary.

4.3 Use Eq. (4.2.5) to determine the composition of the active region of InGaAsP lasers operating at 1.3 and 1.55 μm. What is the longest wavelength at which such lasers can be operated? The band gap of an $In_{1-x}Ga_xAs_yP_{1-y}$ layer at room temperature is given by $E_g(y) = 1.35 - 0.72y + 0.12y^2$.

4.4 Explain the MBE growth technique by drawing the growth apparatus schematically. How does chemical-beam epitaxy differ from MBE?

4.5 Consider the InGaAsP semiconductor lasers operating at 1.3 and 1.55 μm. Both lasers are identical in structure except for the composition of the 0.1-μm thick active layer. Calculate the confinement factor for each laser by calculating refractive indices from Eq. (4.7.4). Assume InP cladding layers to be thick enough that the theory of Sec. 2.5.2 can be used.

REFERENCES

1. Nelson, H. *RCA Rev.* **24,** 603 (1963).
2. Woodall, J. M., H. Rupprecht, and W. Reuter. *J. Electrochem. Soc.* **116,** 899 (1969).
3. H. Nelson, U.S. Patent 3,565,702 (1971).
4. Hayashi, I., M. B. Panish, P. W. Foy, and S. Sumski. *Appl. Phys. Lett.* **17,** 109 (1970).
5. Feng, M., L. W. Cook, M. M. Tashima, and G. E. Stillman. *J. Electron. Mater.* **9,** 241 (1980).

6. Ladany, I., and F. Z. Hawrylo. *J. Cryst. Growth* **54,** 69 (1981).
7. Hsieh, J. J. *Appl. Phys. Lett.* **28,** 283 (1976).
8. Nelson, R. J. *Appl. Phys. Lett.* **35,** 654 (1979).
9. Hsieh, J. J. *J. Cryst. Growth* **27,** 49 (1974).
10. Rode, D. L. *J. Cryst. Growth* **20,** 13 (1973).
11. Tiller, W. A. *J. Cryst. Growth* **2,** 69 (1968).
12. Small, M. B., and I. Crossley. *J. Cryst. Growth* **27,** 35 (1974).
13. Moon, R. L. *J. Cryst. Growth* **27,** 62 (1974).
14. Casey, H. C., Jr. and M. B. Panish. *Heterostructure Lasers*, Part B, Sec. 6.5. New York: Academic Press, 1978.
15. Besomi, P., J. Degami, N. K. Dutta, W. R. Wagner, and R. J. Nelson. *J. Appl. Phys.* **56,** 2879 (1984).
16. Doi, A., N. Chinone, A. Aiki, and R. Ito. *Appl. Phys. Lett.* **34,** 393 (1979).
17. DiGuiseppe, M. A., H. Temkin, and W. A. Bonner. *J. Cryst. Growth* **58,** 279 (1982).
18. Clawson, A. R., W. Y. Lum, and G. W. McWilliams. *J. Cryst. Growth* **46,** 300 (1979).
19. Antypas, G. A. *Appl. Phys. Lett.* **37,** 64 (1980).
20. Besomi, P., R. B. Wilson, W. R. Wagner, and R. J. Nelson. *J. Appl. Phys.* **54,** 535 (1983).
21. Antypas, G. A., and R. L. Moon. *J. Electrochem. Soc.* **120,** 1574 (1973).
22. Antypas, G. A., and J. Edgecumbe. *J. Cryst. Growth* **34,** 132 (1976).
23. Nakajima, K., T. Tanahashi, K. Akita, and T. Yamaoka. *J. Appl. Phys.* **50,** 4975 (1979).
24. Nakajima, K., T. Kusunoki, K. Akita, and T. Kotani. *J. Electrochem. Soc.* **125,** 123 (1978).
25. Sankaran, R., G. A. Antypas, R. L. Moon, J. S. Escher, and L. W. James. *J. Vac. Sci. Technol.* **13,** 932 (1976).
26. Pollack, M. A., R. E. Nahory, J. C. DeWinter, and A. A. Ballman. *Appl. Phys. Lett.* **33,** 314 (1978).
27. Feng, M., T. H. Windhorn, M. M. Tashima, and G. E. Stillman. *Appl. Phys. Lett.* **32,** 758 (1978).
28. For a review see Nakajima, K. Chap 2 in *GaInAsP Alloy Semiconductors*, ed. T. P. Pearsall. New York: John Wiley & Sons, 1982.
29. Hall, R. N. *J. Electrochem. Soc.* **110,** 385 (1963).
30. Hsieh, J. J. pp. 74–80 in *North American Conference on GaAs and Related Compounds, Sept. 1976, St. Louis*, ed. L. F. Eastman. Conf. Series 33b. London: Institute of Physics, 1977.
31. Nagai, H. and Y. Noguchi. *Appl. Phys. Lett.* **32,** 234 (1978).
32. Nahory, R. E., M. A. Pollack, W. D. Johnston, Jr., and R. L. Barns. *Appl. Phys. Lett.* **33,** 659 (1978).
33. Olsen, G. H., Chap. 1 in *GaInAsP Alloy Semiconductors*, ed. T. P. Pearsall. New York: John Wiley & Sons, 1982.
34. Tietjen, J. J., and J. A. Amick. *J. Electrochem. Soc.* **113,** 724 (1966).
35. Olsen, G. H., and T. J. Zamerowski. *IEEE J. Quantum Electron.* **QE-17,** 128 (1981).
36. Enda, H. *Jpn. J. Appl. Phys.* Part 1, **18,** 2167 (1979).
37. Enstrom, R. E., C. J. Nuese, J. R. Appert, and N. J. DiGuiseppe. *Electrochem. Soc. Ext. Abstr.* **78-1,** 504 (1978).
38. Hyder, S. B., R. R. Saxena, and C. B. Cooper. *Appl. Phys. Lett.* **34,** 584 (1979).
39. Seki, H. A., Konkitu, and M. Matsumura. *J. Cryst. Growth* **54,** 615 (1981).
40. Johnston, W. D., Jr., and K. E. Strege. *38th Annual IEEE Device Research Conf. Abstr.*, Cornell University, Ithaca, N.Y., June 1980, Vol. IVB–3.
41. Mizutani, T., M. Yoshida, A. Usai, H. Watanabe, T. Yuasa, and I. Hayashi. *Jpn. J. Appl. Phys.* Part 2, **19,** L113 (1980).
42. Manasevit, H. M., and W. I. Simpson. *J. Electrochem. Soc.* **116,** 1725 (1969).
43. Seki, Y., K. Tanno, K. Iida, and E. Ichiki. *J. Electrochem. Soc.* **122,** 1108 (1975).

44. Dupuis, R. D., P. D. Dapkus, R. D. Yingling, and L. A. Moudy. *Appl. Phys. Lett.* **31,** 201 (1978).
45. Dupuis, R. D. *J. Cryst. Growth* **55,** 213 (1981).
46. Dapkus, P. D. *Annu. Rev. Mater. Sci.* **12,** 243 (1982).
47. Hirtz, J. P., M. Razeghi, M. Bonnet, and J. P. Duchemin, Chap. 3 in *GaInAsP Alloy Semiconductors*, ed. T. P. Pearsall. New York: John Wiley & Sons, 1982; Hersee, S. D., and J. P. Duchemin. *Annu. Rev. Mater. Sci.* **12,** 65 (1982).
48. Hirtz, J. P., J. P. Larivain, J. P. Duchemin, T. P. Pearsall, and M. Bonnet. *Electron. Lett.* **16,** 415 (1980).
49. Hirtz, J. P., J. P. Larivain, D. Leguen, M. Razeghi, and J. P. Duchemin, pp. 29–35 in *Gallium Arsenide and Related Compounds 1980*, Conf. series 56, Institute of Physics, Bristol, 1981.
50. Hirtz, J. P., M. Razeghi, J. P. Larivain, S. Hersee, and J. P. Duchemin. *Electron. Lett.* **17,** 113 (1981).
51. Razeghi, M., S. Hersee, J. P. Hirtz, R. Blondeau, B. de Cremoux, and J. P. Duchemin. *Electron. Lett.* **19,** 336 (1983).
52. Fukui, T. and Y. Horikoshi. *Jpn. J. Appl. Phys.* Part 1, **19,** 395 (1980).
53. Cooper, C. B., R. R. Saxena, and M. J. Ludowise. *Electron. Lett.* **16,** 892 (1980).
54. Yoshino, J., T. Iawnot, and H. Kukimoto. *J. Cryst. Growth* **55,** 74 (1981).
55. Dupuis, R. D., H. Temkin, L. C. Hopkins. *Electron. Lett.* **21,** 60 (1985).
56. Arthur, J. R. *J. Appl. Phys.* **39,** 4032 (1968).
57. Arthur, J. R., and J. J. LePore. *J. Vac. Sci. Technol.* **6,** 545 (1969).
58. Cho, A. Y. *J. Vac. Sci. Technol.* **8,** S31 (1971).
59. Cho, A. Y. *J. Appl. Phys.* **41,** 2780 (1970).
60. Cho, A. Y. *J. Appl. Phys.* **42,** 2074 (1971).
61. Cho, A. Y., and J. R. Arthur. Vol. 10, p. 157, in *Progress in Solid State Chemistry*, ed. J. O. McCladin and G. Somorjai. New York: Pergamon Press, 1975.
62. Tsang, W. T., F. K. Reinhart, and J. A. Ditzenberger. *Appl. Phys. Lett.* **41,** 1094 (1982).
63. Panish, M. B., and S. Sumski. *J. Appl. Phys.* **55,** 3571 (1984).
64. Panish, M. B. *J. Electrochem. Soc.* **127,** 2729 (1980).
65. Calawa, A. R. *Appl. Phys. Lett.* **38,** 701 (1981).
66. Temkin, H., M. B. Panish, R. A. Logan, and J. P. Van der Ziel. *Appl. Phys. Lett.* **45,** 330 (1984).
67. Tsang, W. T. *Appl. Phys. Lett.* **45,** 1234 (1984).
68. Oe, K., Y. Shinoda, and K. Sugiyama. *Appl. Phys. Lett.* **33,** 962 (1978).
69. Nakajima, K., S. Komiya, K. Akita, T. Yamaoka, and O. Ryuzan. *J. Electrochem. Soc.* **127,** 1568 (1980).
70. Adachi, S. *J. Appl. Phys.* **53,** 8775 (1982).
71. Kane, E. O. *J. Phys. Chem. Solids* **1,** 249 (1957).
72. Perea, E. H., E. E. Mendez, and C. G. Fonstad. *Appl. Phys. Lett.* **36,** 978 (1980).
73. Dutta, N. K., and R. J. Nelson. *IEEE J. Quantum Electron.* **QE-18,** 44 (1982).
74. Nicholas, R. J., S. J. Sessions, and J. C. Portal. *Appl. Phys. Lett.* **37,** 178 (1980).
75. Nicholas, R. J., J. C. Portal, C. Houlbert, P. Perrier, and T. P. Pearsall. *Appl. Phys. Lett.* **34,** 492 (1979).
76. Rochon, P., and E. Fortin. *Phys. Rev. B* **12,** 5803 (1975).
77. Alavi, K., R. L. Aggarwal, and S. H. Groves. *J. Magn. Magn. Mater.* **11,** 136 (1979).
78. Perea, E. H., E. Mendez, and C. G. Fonstad. *J. Electron. Mater.* **9,** 459 (1980).
79. Restorff, J. B., B. Houston, R. S. Allgaier, M. A. Littlejohn, and S. B. Phatak. *J. Appl. Phys.* **51,** 2277 (1980).

80. Pearsall, T. P., Chap. 12 in *GaInAsP Alloy Semiconductors*, ed. T. P. Pearsall. New York: John Wiley & Sons, 1982.
81. Hermann, C., and T. P. Pearsall. *Appl. Phys. Lett.* **38,** 450 (1980).
82. Takeda, Y., Chap. 9 in *GaInAsP Alloy Semiconductors*, ed. T. P. Pearsall. New York: John Wiley & Sons, 1982.
83. Hayes, J. R., A. R. Adams, and P. D. Greene, Chap. 8 in *GaInAsP Alloy Semiconductors*, ed. T. P. Pearsall. New York: John Wiley & Sons, 1982.
84. Littlejohn, M. A., T. H. Glisson, and J. R. Hauser, Chap. 10 in *GaInAsP Alloy Semiconductors*, ed. T. P. Pearsall. New York: John Wiley & Sons, 1982.
85. Leheny, R. F., Chap. 11 in *GaInAsP Alloy Semiconductors*, ed. T. P. Pearsall. New York: John Wiley & Sons, 1982.
86. Greene, P. D., S. A. Wheeler, A. R. Adams, A. N. El-Sabbahy, and C. N. Ahmad. *Appl. Phys. Lett.* **35,** 78 (1979).
87. Leheny, R. F., A. A. Ballman, J. C. DeWinter, R. E. Nahory, and M. A. Pollack. *J. Electron. Mater.* **9,** 561 (1980).
88. Hayes, J. R., A. R. Adams, and P. D. Greene. *Electron. Lett.* **16,** 282 (1980).
89. Bhattacharya, P. A., J. W. Ku, S. J. T. Owen, G. H. Olsen, and S. H. Chiao. *IEEE J. Quantum Electron.* **QE-17,** 150 (1981).
90. Petit, G. D., and W. J. Turner. *J. Appl. Phys.* **36,** 2081 (1965); Seraphin, B. O., and H. E. Bennett, pp. 499–543, in *Semiconductors and Semimetals*, Vol. 3, ed. R. K. Willardson and A. C. Beer. New York: Academic Press, 1967.
91. Buus, J., and M. J. Adams. *IEEE Proc. I* **3,** 189 (1979).
92. Nahory, R. E., and M. A. Pollack. *Electron. Lett.* **14,** 727 (1978).
93. Utaka, K., Y. Suematsu, K. Kobayashi, and H. Kawanishi. *Jpn. J. Appl. Phys.* Part 2, **19,** L137 (1980).
94. Olsen, G. H., C. J. Nuese, and M. Ettenberg. *Appl. Phys. Lett.* **34,** 262 (1979).
95. Yano, M., H. Nishi, and M. Takusagawa. *IEEE J. Quantum Electron.* **QE-15,** 571 (1979).
96. Chandra, P., L. A. Coldren, and K. E. Strege. *Electron. Lett.* **17,** 6 (1981).
97. Afromowitz, M. A. *Solid State Commun.* **15,** 59 (1974).
98. Osbourne, G. C. *J. Vac. Sci. Technol.* B **1**, 379 (1983).
99. Massalink, W. T., T. Henderson, J. Klem, R. Fischer, P. Pearah, H. Morkoc, M. Hafich, P. D. Wang, and G. Y. Robinson. *Appl. Phys. Lett.* **45,** 1309 (1984).
100. Fischer, S. E., D. Fekete, G. B. Feak, and J. A. Ballantyne. *Appl. Phys. Lett.* **50,** 714 (1984).
101. Matthews, J. W., and A. E. Blakeslee. *J. Cryst. Growth* **27,** 118 (1974).
102. Tanbun-Ek, T., R. A. Logan, H. Temkin, S. N. G. Chu, N. A. Olsson, A. M. Sergent, and K. W. Wecht. *IEEE J. Quantum Electron.* **QE-26,** 1323 (1990).
103. Berger, P. R., K. Chang, P. Bhattacharya, J. Singh, and K. K. Bajaj. *Appl. Phys. Lett.* **53,** 684 (1988).

Chapter 5
LASER STRUCTURES AND THEIR PERFORMANCE

5.1 INTRODUCTION

Semiconductor lasers operating in the wavelength range of 1.1–1.65 μm can be fabricated using the InGaAsP quaternary material which has been grown lattice-matched on an InP substrate.[1] Room-temperature continuous operation of InGaAsP-InP double-heterostructure lasers was first reported[2] in 1976. Since then, a large number of laser structures have been developed guided by the performance requirements of specific applications. This chapter discusses different InGaAsP laser structures with particular emphasis on their performance in terms of the light-current characteristics, the threshold current, and the threshold current's temperature dependence. Other performance characteristics that make a laser structure suitable as a source for high-speed digital lightwave transmission systems are discussed in detail in Chapter 6.

In a double-heterostructure laser, the optical mode is confined perpendicular to the junction plane because the cladding layers have a lower index of refraction compared with that of the active region. For stable fundamental-mode operation with a low threshold current, additional confinement of the optical mode along the junction plane is required. In the absence of this lateral-mode confinement (which may be induced by a region of well-defined optical gain or a refractive index step), the laser behaves as a broad-area laser with a threshold current density of ~ 1 kA/cm^2; i.e., a threshold current of ~ 1 A is required for a laser of cross section 250 μm × 380 μm. Such a high threshold current would limit the usefulness of semiconductor injection lasers. The laser structures described in this chapter are essentially different ways of realizing the lateral-mode confinement; the eventual aim is to obtain a laser structure with a low threshold current that can operate in the fundamental lateral mode up to high powers. State-of-the-art InGaAsP lasers emitting near 1.3 μm generally have threshold currents in the range of 10–15 mA at 20°C.

As discussed in Sec. 2.5, the laser structures can in general be classified into two groups, gain-guided and index-guided. In the *gain-guided* structure,

the width of the optical mode along the junction plane is mainly determined by the width of the optical gain region, which in turn is determined by the width of the current-pumped region (typically in the range of 5–10 μm). In *index-guided* lasers, a narrow central region of relatively higher refractive index in the junction plane confines the lasing mode to that region. The index-guided lasers can be divided into two general subgroups, weakly index guided and strongly index-guided lasers. In weakly index-guided lasers, the active region is continuous and the effective index discontinuity is provided by a cladding layer of varying thickness. Depending on the specific design, these structures are known under various names, such as rib waveguide, ridge waveguide, plano-convex waveguide, and channeled-substrate planar waveguide. By contrast, the strongly index-guided lasers employ a buried heterostructure. In these lasers, the active region is bounded by low-index, epitaxially grown layers both along and normal to the junction plane. The lateral index difference is ~ 0.01 for weakly index-guided lasers and ~ 0.2 for strongly index-guided, buried-heterostructure lasers.

Gain-guided lasers are easier to fabricate than index-guided lasers. However, gain-guided lasers exhibit higher threshold currents than do index-guided lasers. Typical threshold currents of gain-guided lasers are in the range of 100–150 mA compared with 10–15 mA for buried-heterostructure lasers. The fabrication of strongly index-guided lasers generally requires either a single epitaxial growth over nonplanar surfaces or two epitaxial growths and, in addition, careful attention to processing. In spite of these difficulties in fabrication, their superior performance characteristics—low threshold current, stable fundamental-mode operation at high powers, and good high-speed modulation characteristics—make them a prime candidate for high-performance applications. In particular, their use as a source in lightwave transmission systems is highly attractive. In fact, the development of various types of strongly index-guided laser structures over the last few years principally has been carried out using InGaAsP. This is the material of choice for long-haul lightwave transmission system applications. Weakly index-guided lasers are easier to fabricate than strongly index-guided lasers but generally have a higher threshold current (~ 50 mA). These lasers can be a useful alternative to strongly index-guided lasers in less demanding applications.

5.2 BROAD-AREA LASERS

The laser structures described in this chapter are generic, in principle, to any material system, although we focus our attention on the use of InGaAsP. The InGaAsP alloy can be grown lattice-matched on InP and can emit light in the wavelength of 1.1–1.65 μm. A useful figure of merit, indicative of

the "quality" of a semiconductor laser material, is the broad-area threshold current density. A broad-area laser does not employ any scheme for current confinement. The processing of these devices is relatively simple. Most of the early laser work using InGaAsP was carried out using broad-area structures grown by liquid-phase epitaxy (LPE) on an *n*-type InP substrate.[2–8] The as-grown wafer, typically 250–400 μm thick, is thinned down to a thickness of 75–100 μm, and metallic contacts are deposited on the substrate side (typically AuSn or AuGe alloy) and the epitaxially grown side (typically CrAu or AuZn alloy). The thinning of the wafer is necessary to facilitate cleaving along a crystallographic plane. The cleaved planes [typically (110)] form the mirrors of the laser cavity. The other sides are saw-cut. Figure 5.1 shows a schematic of a broad-area laser chip, together with typical device dimensions.

The threshold current density J_{th} depends on the confinement of the optical mode and hence varies with the thickness d of the active region [see Eqs. (2.6.2) and (2.6.3)]. Figure 5.2 shows experimental data[4] for J_{th} as a function of d for 1.3-μm InGaAsP-InP double-heterostructure lasers. We note that J_{th} is minimum for an optimum value of d. The rapid increase for d when $d < 0.1$ μm is due to the loss in mode confinement for thin active layers. The measured threshold current density varies linearly with d for thicknesses greater than 0.2 μm. Thus J_{th}/d is a useful universal parameter for the characterization of lasers emitting at different wavelengths or lasers fabricated

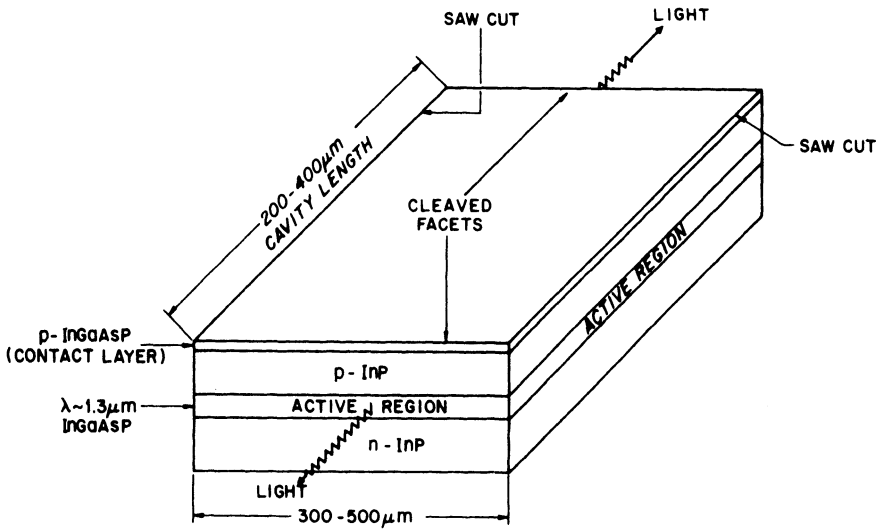


Fig. 5.1 Schematic illustration of a typical broad-area laser chip for a 1.3-μm InGaAsP double-heterostructure laser.

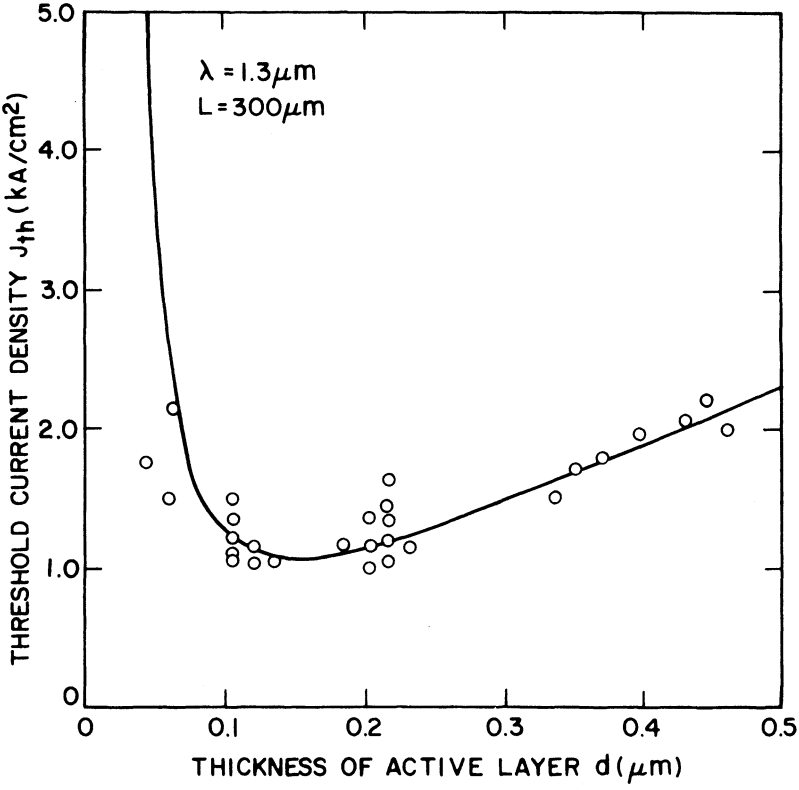


Fig. 5.2 Measured threshold current density at room temperature as a function of the active-layer thickness for 1.3-μm InGaAsP double-heterostructure lasers. (After Ref. 4)

using different material systems. The quantity J_{th}/d equals the nominal current density defined by Eq. (3.2.44) in the absence of nonradiative recombination mechanisms. Figure 5.3 shows the experimental data for J_{th}/d for InGaAsP lasers emitting at different wavelengths.[5] For lasers emitting at wavelengths greater than 1.5 μm it is often necessary to grow an additional layer (called the antimeltback layer) on top of the active layer to prevent melting back of the active region during the LPE growth of the *p*-InP cladding layer. The band gap of the antimeltback layer is intermediate to that of surrounding layers. Figure 5.3 shows the data for lasers with and without antimeltback layers. In all cases the average value of J_{th}/d is ~ 5 kA/(cm^2 · μm) and is found to be independent of wavelength in the range of 1.2–1.65 μm. This value is comparable to the values reported for AlGaAs lasers.[9]

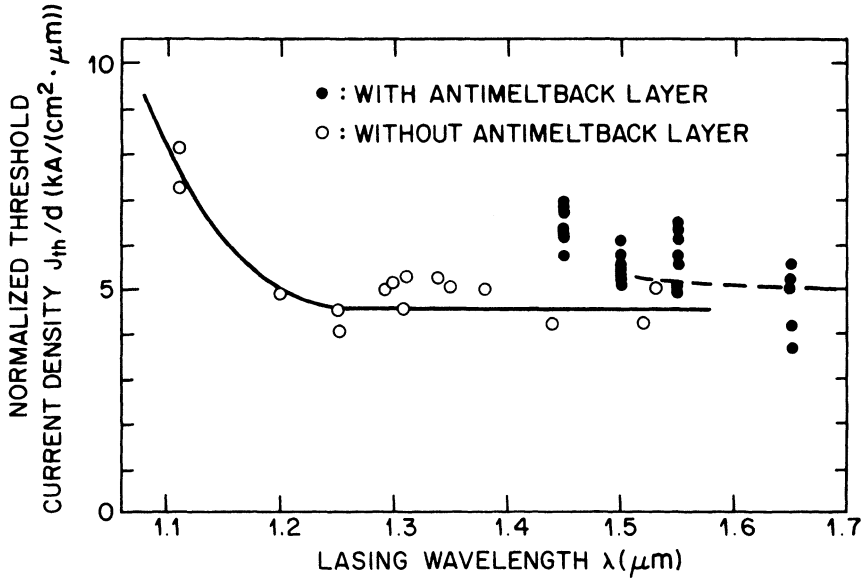


Fig. 5.3 Experimental data for J_{th}/d for broad-area InGaAsP lasers emitting at different wavelengths. (After Ref. 5 © 1980 IEEE)

Various factors that can affect the threshold current density of a broad-area laser have been theoretically discussed in Sec. 2.6. One of the factors is the carrier concentration in the active region. It is generally accepted that increased p doping of the active region can increase the threshold current. The effect is partly due to an increase in the nonradiative Auger recombination rate and partly due to an increase in optical absorption. The former manifests itself in a decrease in the carrier lifetime [see Eq. (2.6.3)], and the latter causes a decrease in the external differential quantum efficiency [see Eq. (2.6.11)].

In the fabrication of InGaAsP-InP lasers, the active layer is not intentionally doped. However, during LPE of the p-InP cladding layer that is usually doped using Zn, a diffusion of Zn into the active layer takes place and makes it p type. Experimental data[4] for the observed threshold current density as a function of the acceptor concentration of the active layer are shown in Fig. 5.4 for three different thicknesses of the active layer. The donor concentration of the n-InP layer in these experiments was 7.5×10^{18} cm^{-3}. Figure 5.4 shows that thicker active layers are less affected than thinner layers in the range of 10^{17}–10^{18} cm^{-3}. The increase of threshold current for cladding-layer doping levels less than 10^{17} cm^{-3} is probably due to a displacement of the p-n junction into the p cladding layer during growth.

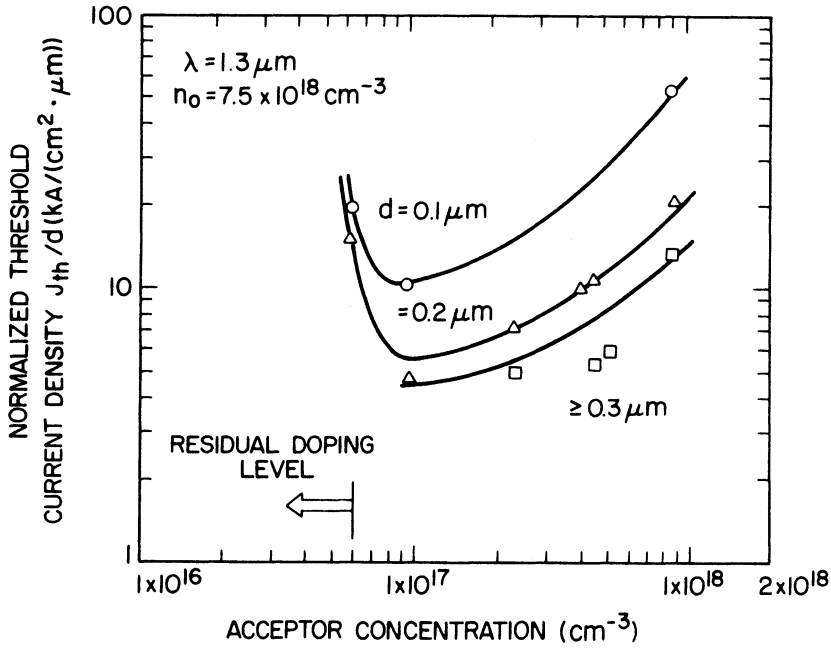


Fig. 5.4 Measured threshold current density of 1.3-μm InGaAsP-InP broad-area lasers as a function of the acceptor concentration in the *p*-InP cladding layer. The donor concentration is kept constant. (After Ref. 4)

The threshold current density of a broad-area laser also depends on the cavity length L. This dependence can be seen from Eqs. (2.6.2) and (2.3.10). For long cavities, the distributed mirror loss α_m is small compared to the internal loss α_{int}. Itaya et al.[10] have reported a threshold current density of 770 A/cm^2 for a cavity length of 1500 μm in an InGaAsP laser emitting at 1.3 μm. More recently, values as low as $\sim$500 A/cm^2 have been observed with design optimization for 1.3-μm InGaAsP lasers of about 400 μm in length.[8]

5.3 GAIN-GUIDED LASERS

Most injection lasers intended for commercial applications have a built-in feature that restricts current injection to a small region along the junction plane. This restriction serves several purposes: (i) it allows continuous-wave (CW) operation with reasonably low threshold currents (10–100 mA) compared with an unacceptably high value ($\sim$1 A) for broad-area lasers; (ii) it can allow fundamental-mode operation along the junction plane, which is

necessary for applications where the light is coupled into an optical fiber (iii) the requirements for heat sinking are considerably less severe than those for a broad-area laser; and (v) the low operating current allows operation at high temperatures and increases the operating life of the device.

The simplest current-restricting structure is an oxide-stripe device, originally fabricated by Dyment[11] and used to study the transverse-mode structure along the junction plane in AlGaAs injection lasers. A variation of the above structure, the proton-stripe AlGaAs laser,[12] has become the workhorse for many commercial applications in systems employing AlGaAs lasers.

Figure 5.5 shows three gain-guided laser structures fabricated using the InGaAsP material system. These are (i) the oxide-stripe laser, where an SiO_2 layer on the p contact confines the injected current flow to a small region through an opening in the dielectric[6]; (ii) the proton-stripe[13] or deuteron-stripe[14] laser, where the implanted protons or deuterons create a region of high resistivity that restricts the current flow to an opening in the implanted region; and (iii) the junction-stripe laser,[15] where Zn diffusion converts a small region of the top n-type layer into p type, thus providing a current

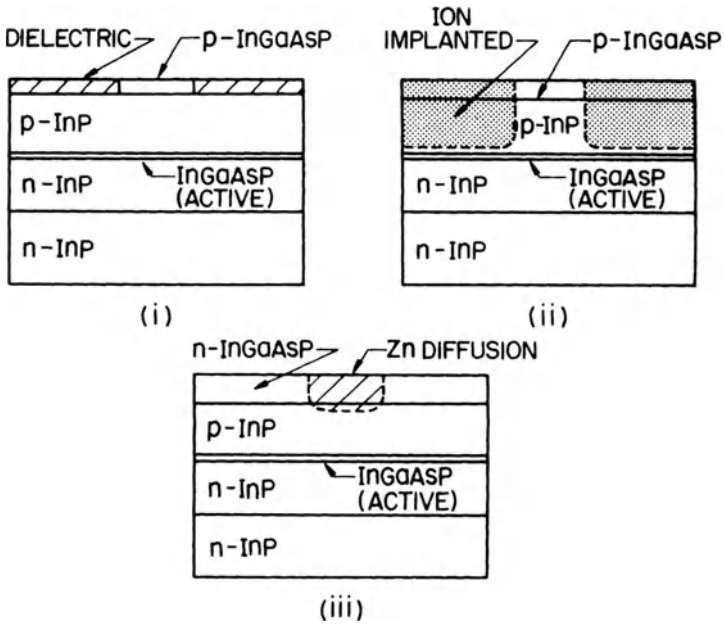


Fig. 5.5 Schematic cross section of different types of gain-guided laser structures: (i) oxide-stripe, (ii) proton- or deuteron-stripe, and (iii) junction-stripe.

path while the reverse-biased junction over the remaining region provides current confinement. Variations of the above structures can also be found in the literature.

In all these structures, the active region is planar and continuous. The stimulated-emission characteristics of such a laser are determined by the carrier distribution (which provides optical gain) along the junction plane. Since the optical mode distribution along the junction is determined by the optical gain, these lasers are called gain-guided lasers. The physics behind the gain-guiding mechanism has been discussed in Sec. 2.5.3.

Figure 5.6a shows the typical light-current (L-I) characteristics of a gain-guided deuteron-stripe InGaAsP laser ($\lambda = 1.3\ \mu m$). These devices were fabricated using an 8-μm-thick wire as a bombardment mask. The 250-μm-long lasers had threshold currents in the range 100–150 mA. Similar threshold current values have been reported for other types of gain-guided lasers.

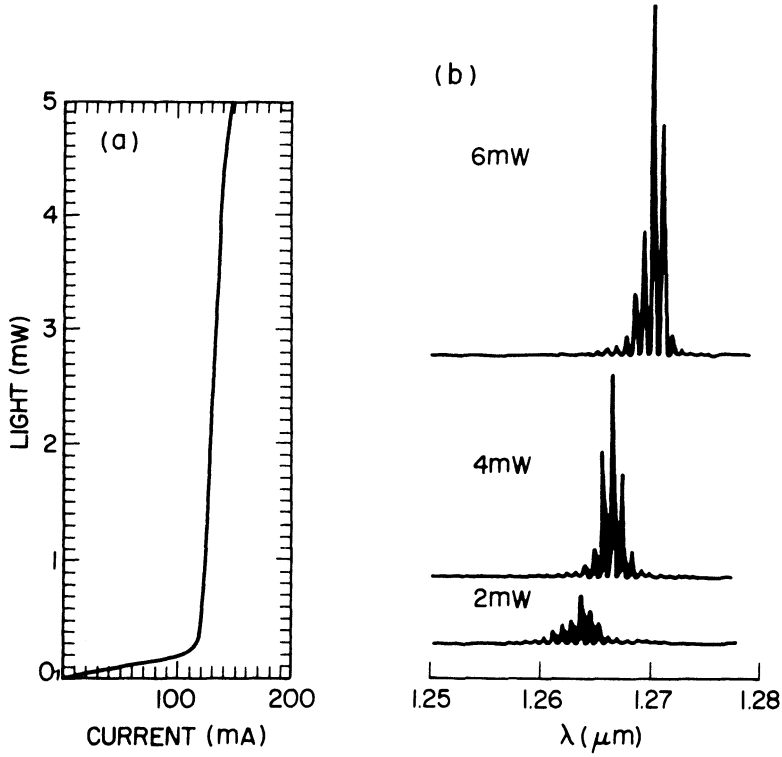


Fig. 5.6 (a) Measured CW light-current characteristic of a deuteron-stripe gain-guided InGaAsP laser. (b) Optical spectra at several power levels. (After Ref. 14 © 1984 IEEE)

Typical spectral emission characteristics of these lasers are shown in Fig. 5.6b. The laser emits at several wavelengths, each of which corresponds to a different longitudinal mode of the laser cavity, as discussed in Sec. 2.3. Note that the longitudinal modes shift toward longer wavelengths as the drive current increases. This is a consequence of the temperature-induced shift of the gain peak as a result of device heating at high currents. We now discuss some specific characteristics (generally undesirable) of gain-guided lasers.

Kinks in Light-Current Characteristics. As the current through the laser is increased, the L-I characteristics of a gain-guided laser often exhibit a "kink," or nonlinearity. The nonlinearity can be associated with a movement of the optical mode along the junction plane, a transition to higher-order modes, or a transition from the TE to the TM mode. Such a nonlinear response can severely alter the amount of light coupled into an optical fiber and thus limits the usefulness of the laser in an optical communication system. Figure 5.7 shows an example of the kink for an oxide-stripe laser.

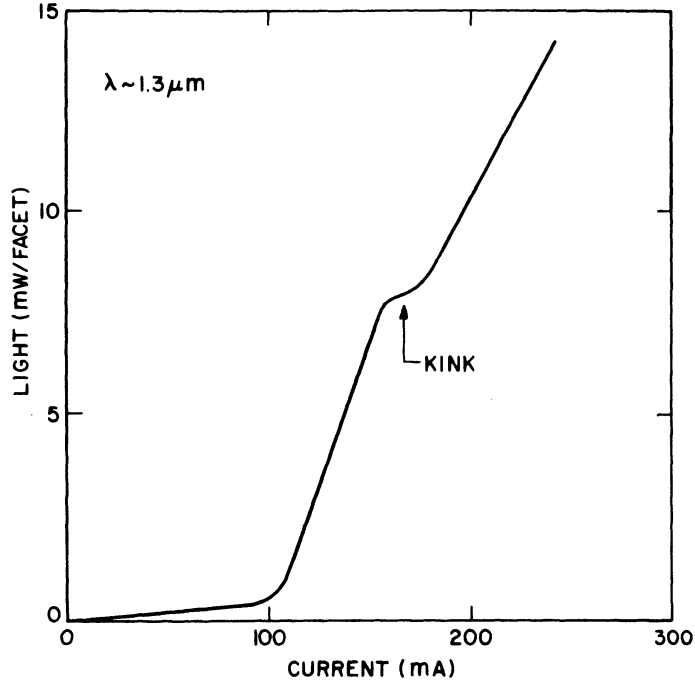


Fig. 5.7 Light-current characteristic of a gain-guided InGaAsP laser ($\lambda = 1.3$ μm) showing the occurrence of a kink at an output power of 8 mW/facet.

The lasers with narrow stripes exhibit kinks at higher operating powers than do lasers with wide stripes, as observed for AlGaAs proton-stripe lasers.[16]

Pulsations. Many gain-guided injection lasers exhibit sustained oscillations or pulsations; i.e., the emitted light pulsates at a certain frequency (typically in the range of 200–500 MHz). Figure 5.8 shows an example of a self-pulsing laser. The bottom trace shows the long duration current pulse, and the top trace shows the emitted light pulses at a repetition rate of ~200 MHz.[17]

The self-pulsing phenomenon has been extensively studied for AlGaAs proton-stripe gain-guided lasers.[18–20] Similar characteristics have been observed for InGaAsP lasers.[21] In general, the pulsation frequency increases with increasing drive current, approximately obeying the simple relation

$$\Omega_R \propto \left(\frac{I}{I_{\text{th}}} - 1\right)^{1/2} \tag{5.3.1}$$

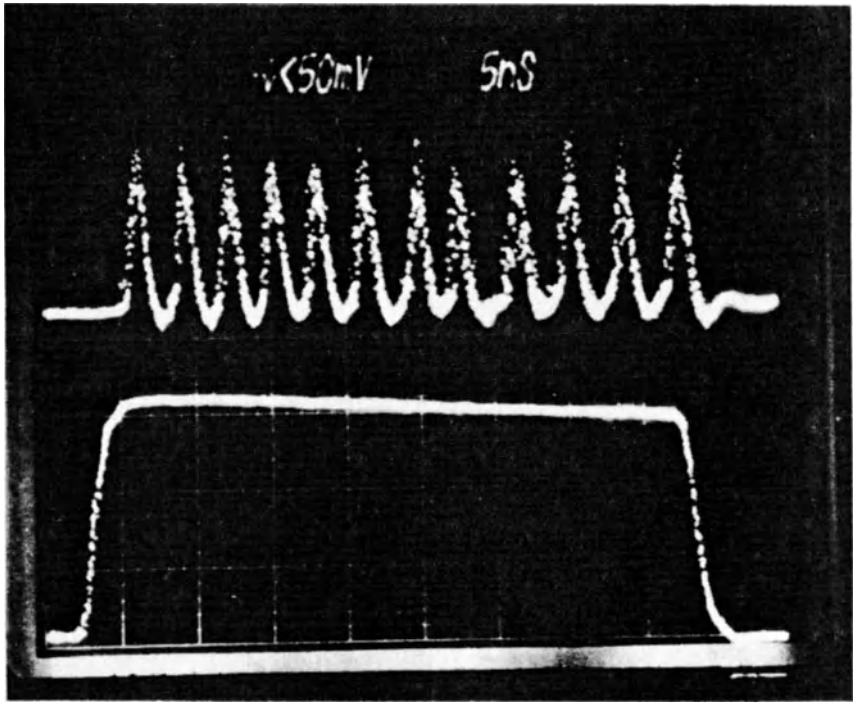


Fig. 5.8 Self-pulsing in a 1.3-μm InGaAsP oxide-stripe laser. The top trace shows the light output, and the bottom shows the current pulse. The horizontal scale is 5 ns per division. (After Ref. 17)

where Ω_R is the pulsation frequency, I_{th} is the threshold current, and I is the injection current. It will be seen in Chapter 6 that Ω_R is the relaxation-oscillation frequency of an intrinsic laser resonance. Self-pulsing semiconductor lasers have found commercial applications in optical-disk systems (e.g., compact-disk players) because of their relative insensitivity to external optical feedback.

Several models exist in the literature that explain the observation of pulsation in injection lasers. A saturable-absorption model proposed by Joyce and Dixon[22] satisfactorily accounts for pulsations. This model relies on a fundamental nonlinear relation between the optical gain and the injected carrier density in semiconductor lasers.[23] This relation is shown schematically in Fig. 5.9. The pulsation phenomenon can be understood in the following qualitative way. Consider an injection laser that has a nonuniform carrier density along the cavity length. Such nonuniformity is generally localized and can arise from defects in the active layer, nonradiative recombination at mirror facets, or local variation in current injection caused by processing defects. The carrier density in these regions is smaller than the average value across the cavity length. Let n_1 denote the carrier density at the defects and n_0 be the average carrier density. When a current ΔI is injected into the laser, the carrier density n_1 changes by Δn_1 and the average carrier density n_0 changes by Δn_0 (see Fig. 5.9). The corresponding changes in the gain or loss are $(dg/dn_0)\Delta n_0$ and $(dg/dn_1)\Delta n_1$. If l is the length of the defective region and L is the length of the laser, a net optical gain results if

$$l\left(\frac{dg}{dn_1}\right)\Delta n_1 > L\left(\frac{dg}{dn_0}\right)\Delta n_0. \tag{5.3.2}$$

The condition can be satisfied for some values of l and L since $dg/dn_1 > dg/dn_0$ as seen in Fig. 5.9. If Eq. (5.3.2) is satisfied, an increase in the injected current results in a net increase in the above-threshold gain, which causes the stimulated emission to rise suddenly. This sudden increase in stimulated emission causes a depletion of carriers to a below-threshold value, and the stimulated emission stops. Stimulated emission begins when the carrier density is replenished to an above-threshold value by the injected current and the process repeats itself. Thus Eq. (5.3.2) may be viewed as the condition for pulsation in an injection laser. Central to this model is the nonlinear-gain-versus-carrier-density relation, which makes regions with smaller carrier density act as saturable absorbers. A discussion of the pulsation instability based on the rate equations is given in Sec. 6.4.3.

The regions near the cleaved facets of an injection laser are generally regions where carrier density is depleted by surface recombination. The recombination rate at a surface is usually expressed in terms of surface

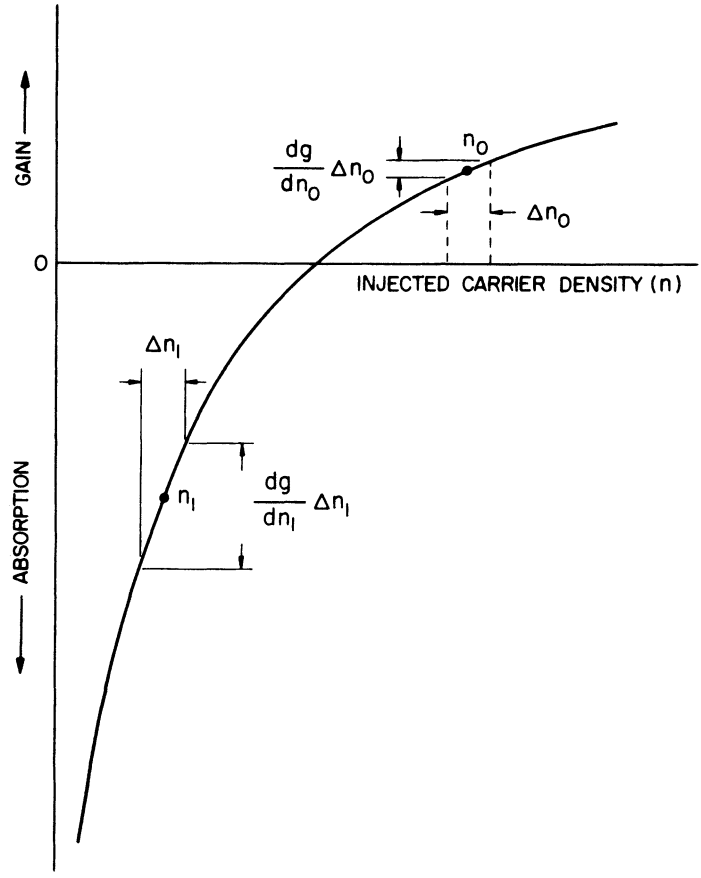


Fig. 5.9 Optical gain and absorption versus injected carrier density in a semiconductor laser show schematically. A small change in carrier density affects absorption more strongly than gain.

recombination velocity S (see Sec. 3.3.3), which is the product of the defect density N_d, capture cross section σ, and carrier velocity v, i.e.,

$$S = N_d \sigma v. \tag{5.3.3}$$

Nash et al.[24] have shown that the surface recombination velocity of a cleaved GaAs surface can be reduced by coating the surface with Al_2O_3, which also reduces the rate at which pulsations occur in AlGaAs lasers. The occurrence of pulsation is generally more frequent in AlGaAs lasers than in long-wavelength InGaAsP lasers. This is generally attributed to the fact that

a cleaved InP surface has a smaller surface recombination velocity ($\sim 10^5$ cm/s) than does a GaAs surface exposed to air ($\sim 10^7$ cm/s).[25]

Van der Ziel[26] has studied the self-focusing (contraction of the beam diameter near the pulse center) associated with pulsations in AlGaAs proton-stripe gain-guided lasers. Figure 5.10 shows the similar self-focusing behavior for 1.3-μm InGaAsP lasers. The top trace shows the light pulse and the bottom trace shows the measured near-field width during the evolution of the optical pulse. As a consequence of self-focusing, the width decreases as the light intensity increases. Lang[27] has shown that index guiding along the junction can reduce the self-focusing effect by stabilizing the optical mode and can also reduce the probability of pulsations. His analysis agrees with the performance of strongly index-guided InGaAsP lasers, which generally do not exhibit pulsations.[21]

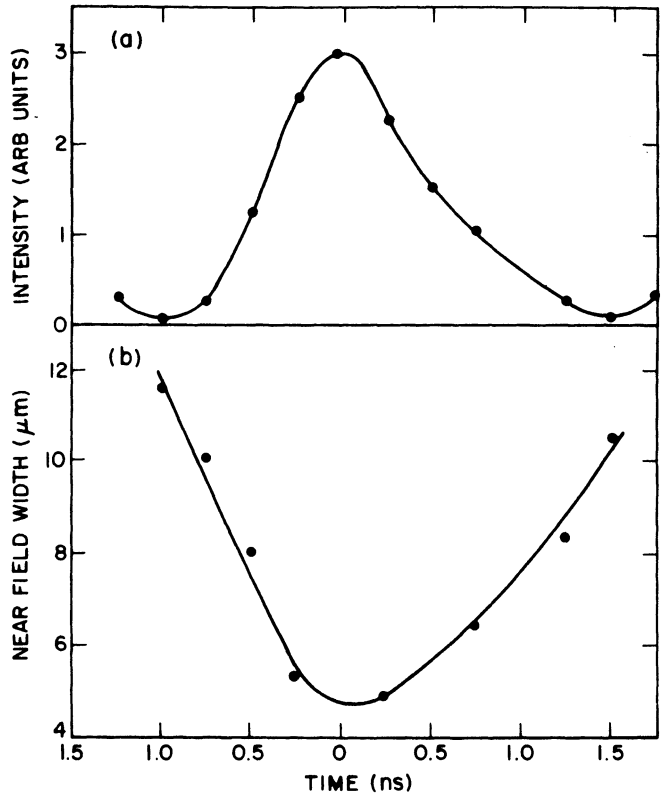


Fig. 5.10 Self-focusing observed in a self-pulsing 1.3-μm InGaAsP laser. (a) The light pulse. (b) The corresponding measured near-field width during the evolution of the pulse. Note that the width decreases as the optical intensity increases.

Electrical Derivative Characteristics. Electrical characterization, especially the measurement of current-voltage derivatives, is a useful way of understanding various current paths in an injection laser. The electrical characteristics of injection lasers were first developed for the study of stripe-geometry gain-guided lasers.[28–30] Below threshold, the current-voltage characteristic of a broad-area junction laser is similar to that of a diode with a series resistance R. The series resistance accounts for the contact resistance and that of the various layers. The current-voltage characteristics of a diode are given by the Shockley equation

$$I = I_s[\exp(\beta V_d) - 1] \tag{5.3.4}$$

where $\beta = q/\eta k_B T$, V_d is the voltage across the diode, I_s is the saturation parameter, k_B is the Boltzmann constant, T is the absolute temperature, q is the electron charge, and η is called the ideality factor. Usually $\exp(\beta V_d) \gg 1$, so that the second term in Eq. (5.3.4) can be neglected. Above threshold, the voltage across the laser diode saturates because the carrier density saturates. The measured voltage V across the laser diode then is

$$V = V_d + IR. \tag{5.3.5}$$

From Eqs. (5.3.4) and (5.3.5), it follows that

$$I\frac{dV}{dI} = \begin{cases} \dfrac{\eta k_B T}{q} + RI & I < I_{\text{th}} \\ RI & I > I_{\text{th}}. \end{cases} \tag{5.3.6}$$

The quantity $I\,dV/dI$ drops by $\eta k_B T/q$ at threshold for an injection laser. Measurement of this quantity can provide information about the junction characteristics of the laser diode. Figure 5.11 shows the measured optical and electrical characteristics of a stripe-geometry gain-guided laser. A kink at threshold in the $I\,dV/dI$ curve is in agreement with Eq. (5.3.6). Equation (5.3.6) is modified in the presence of carrier leakage over the heterojunction (see Sec. 3.5.1). The modified equation shows that observation of sublinearity in the above-threshold variation of $I\,dV/dI$ with I is an indication of heterobarrier leakage.[31]

Wright et al.[32] have examined the electrical characteristics in the presence of a shunt path across the laser diode. Such shunt paths can arise in strongly index-guided lasers. For simplicity, we consider a resistive shunt path. Figure 5.12 shows the $I\,dV/dI$-I characteristic of a laser with a resistive shunt. The slow turn-on of the I-V curve is due to the current flowing through

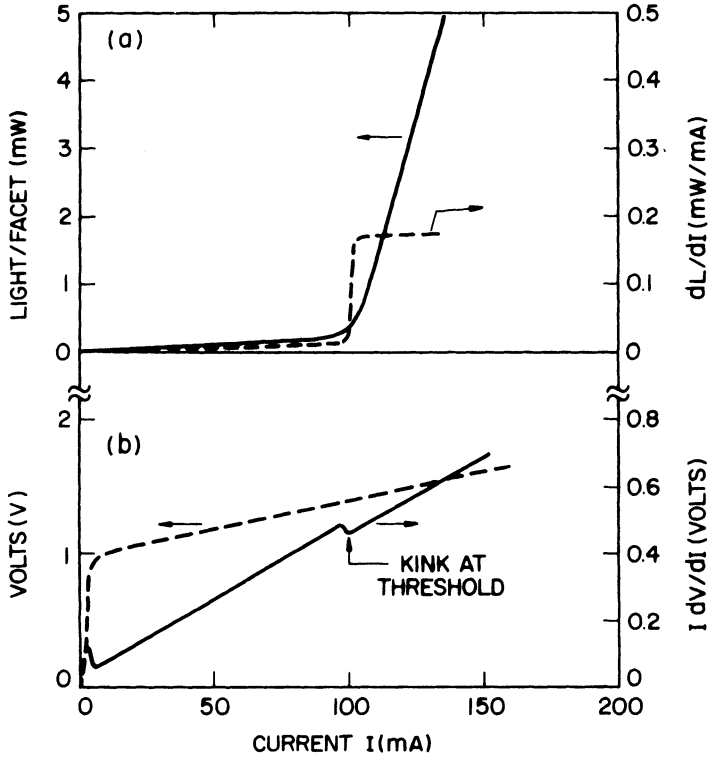


Fig. 5.11 Measured (a) optical and (b) electrical characteristics of a stripe-geometry gain-guided InGaAsP laser operating at 1.3 μm. In (a), solid and dashed lines show L-I and *dL/dI-I* curves respectively. In (b), solid and dashed lines show *V-I* and *I dV/dI-I* curves respectively.

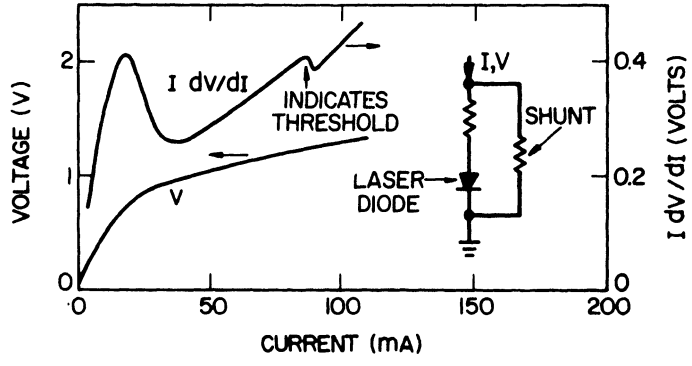


Fig. 5.12 Electrical derivative characteristics of a laser with a linear shunt. Peak in the *I dV/dI* (top) curve at about 20 mA is a signature of the shunt path. (After Ref. 32)

the resistive shunt, and its slope at low currents equals the shunt resistance. The corresponding $I\,dV/dI$-versus-I curve shows a "bump" before threshold is reached. Observation of such a bump in the measured $I\,dV/dI$-versus-I characteristic of a laser diode indicates the presence of a shunt path.

Optical characteristics also manifest themselves through the measured electrical derivative characteristics. For example, a kink in the L-I characteristic is generally associated with a kink in the $I\,dV/dI$-I characteristic. This can be understood as follows. An L-I kink is usually associated with a change in the optical mode (mode movement, mode transition, or appearance of additional higher-order modes) along the junction plane. This also changes the average carrier density in the active region, which in turn changes the Fermi energies. Since the voltage across the diode equals the separation of the quasi-Fermi levels in the conduction and valence bands, a change in voltage occurs with a change in the optical mode along the junction plane. The voltage change appears as a kink in the $I\,dV/dI$-versus-I characteristic.

The sustained oscillations of the emitted light or pulsations are often associated with sharp kinks in the electrical derivative characteristics. An example of such behavior is shown in Fig. 5.13. The pulsation-induced sharp spike in the dL/dI curve manifests as a sharp dip in $I\,dV/dI$. Note that the $I\,dV/dI$ curve near threshold is lower than what would be expected from the series resistance alone. Thus the device apparently exhibits negative resistance. This type of negative-resistance behavior was first reported by Anthony et al.[33] During pulsation, the average carrier density oscillates rapidly, which

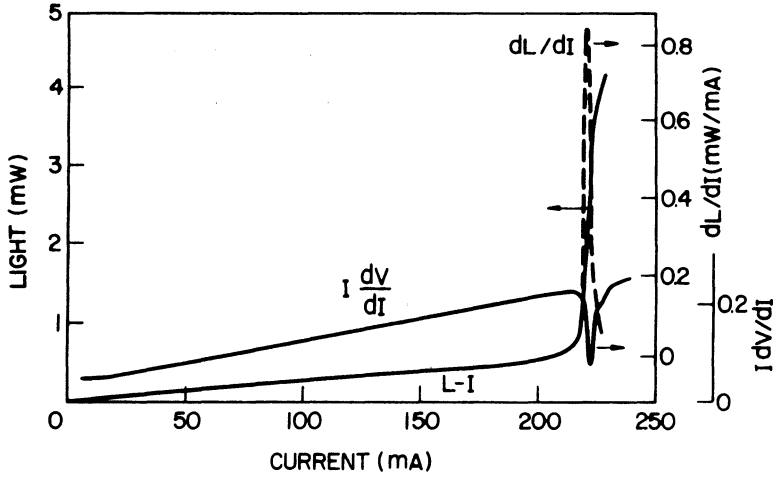


Fig. 5.13 An example of the observed negative-resistance-like behavior in the electrical-derivative characteristic of a self-pulsing laser. (After Ref. 17)

leads to an oscillation in voltage across the diode. This rapidly oscillating voltage appears as an apparent negative resistance in the measured electrical derivative characteristics.

5.4 WEAKLY INDEX-GUIDED LASERS

The lateral mode in gain-guided lasers is determined by the distribution of optical gain along the junction plane. The optical gain is determined by the carrier distribution that is influenced by both current spreading (due to the sheet resistance of the p cladding layer) and carrier diffusion in the active region. As discussed in Chapter 2, gain-guided InGaAsP lasers have undesirable characteristics, such as a high threshold current and a low differential quantum efficiency, as a consequence of the carrier-induced index reduction leading to index antiguiding. The effective index depression caused by carriers at the waveguide center is $\sim 5 \times 10^{-3}$. In weakly index-guided laser structures, the laser structure is modified so as to introduce an effective index step of $\sim 10^{-2}$, which is larger than the carrier-induced reduction. Hence the optical mode along the junction plane in a weakly index-guided laser is essentially determined by the device structure. The transition from gain guiding to index guiding as the lateral effective index step is varied was discussed in Sec. 2.6.

The design of a weakly index-guided laser requires that the thickness of at least one layer be laterally nonuniform. A number of structures, known under various names such as rib waveguide, ridge waveguide, plano-convex waveguide, channeled-substrate planar waveguide, have been used[34–44] for this purpose, and some of them are shown schematically in Fig. 5.14. For ease of reference, we group them into two categories, collectively called the ridge waveguide type and the rib waveguide type. In the *ridge waveguide* type of structure, the use of a dielectric around the ridge [see schemes (i) and (ii) in Fig. 5.14] inhibits current spreading in the p-InP layer. By contrast, in the *rib waveguide* type of structure [see schemes (iii) and (iv) in Fig. 5.14], current spreading in the p-InP layer can influence the threshold current. Both schemes are capable, however, of providing an effective index step of $\sim 10^{-2}$ for the lateral mode under the ridge or rib region, and the mode is essentially index-guided. Two types of ridge waveguide lasers are shown: (i) the structure used by Kaminow et al.[38] for InGaAsP lasers and (ii) the heteroepitaxial ridge-overgrown laser.[43] The difference between the two types of rib waveguide lasers shown in Fig. 5.14 is that in one case the lateral index step is provided by a waveguide layer of varying thickness, while in the other case the lateral index step arises from an active layer of varying thickness. Note that the rib can also be made on the p side.[34] To distinguish the two cases, the structure (iv) of Fig. 5.14 is sometimes referred to as the

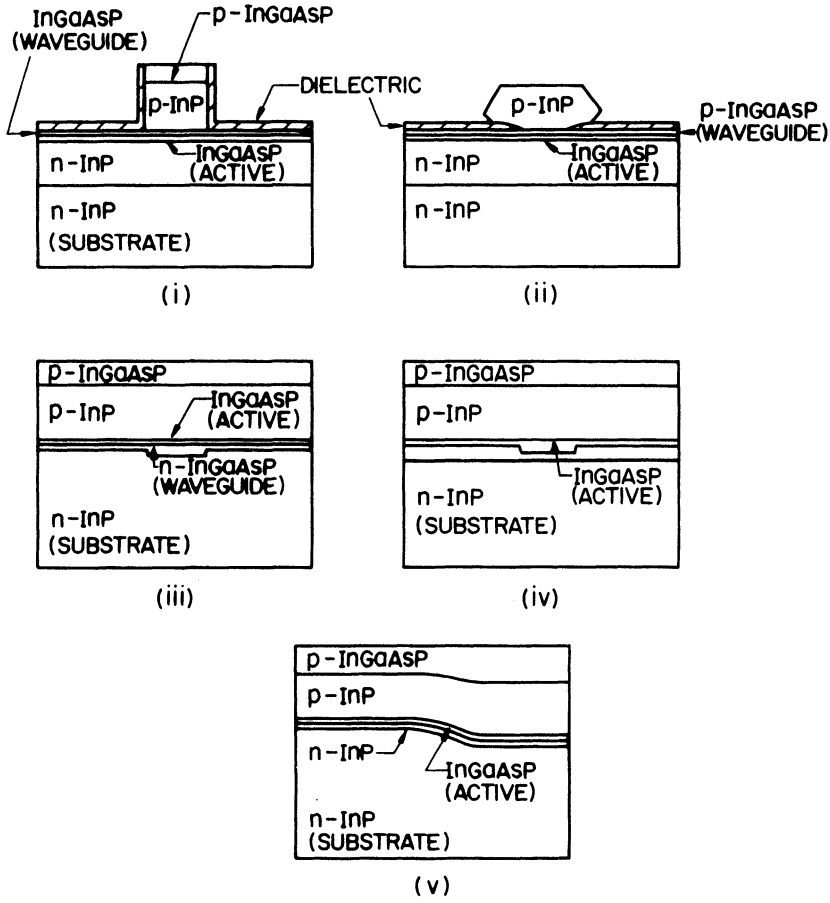
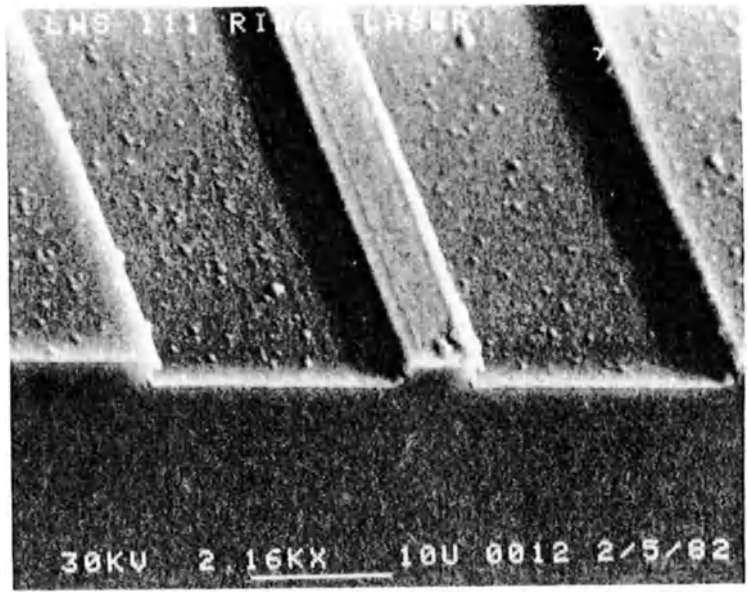


Fig. 5.14 Schematic cross section of several types of weakly index-guided laser structures: (i) ridge waveguide laser; (ii) heteroepitaxial ridge-overgrown laser; (iii) channeled-substrate planar waveguide laser; (iv) simple rib waveguide laser; and (v) terraced-substrate laser.

inverted rib waveguide.[41] Similarly, structure (iii) is also called plano-convex waveguide.[39]

Ridge Waveguide Laser. The ridge waveguide (RWG) laser has been extensively studied.[42–44] Figure 5.15 shows scanning electron microscope (SEM) photomicrographs of the laser cross section. Device fabrication involves the following steps: The planar epitaxial layers are successively grown as shown in Fig. 5.14 by either LPE, VPE, or MOVPE on a (100)-oriented *n*-InP substrate. The wafer is then etched to form a ridge



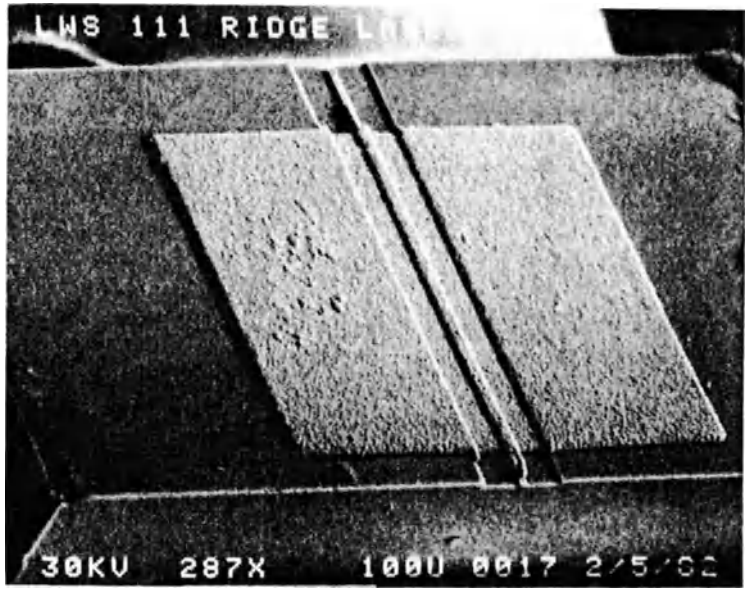
(a)

Fig. 5.15(a) SEM cross section of a ridge waveguide laser showing the ridge formed by chemical etching. The nominal width of the ridge is about 5 μm.

with a width of ~5 μm. The 1.1-μm InGaAsP waveguide layer acts as a stop-etch layer. The wafer is then processed to produce lasers using standard dielectric deposition, photolithographic, and metallization techniques. Figure 5.15(a) shows the SEM cross section of an RWG laser showing the ridge formed by etching. Figure 5.15(b) shows the same laser with a gold pad used for *p*-contact.

The planar active layer of a ridge waveguide laser extends beyond the ridge, and the threshold current is affected by carrier diffusion in the active region. Figure 5.16 shows the typical light-current characteristics of a 1.3-μm InGaAsP ridge waveguide laser.[42] Also shown are the far-field patterns along and normal to the junction plane. Typical threshold currents of these devices are in the range of 40–60 mA.

Lateral index guiding in a ridge waveguide laser is schematically shown in Fig. 5.17a. A fraction of the lasing mode overlaps with the dielectric, which has a considerably lower refractive index ($\mu \cong 1.83$) than that of the InP cladding layer ($\mu = 3.22$). This overlap introduces an effective lateral index step as discussed in detail in Sec. 2.5. The magnitude of the lateral



(b)

Fig. 5.15(b) SEM cross section of the same RWG laser showing the gold pad used for *p*-contact.

index step is determined by the thicknesses of the waveguide layer. For a very thick waveguide layer, the lateral index step may become so small that it is comparable to the carrier-induced index change. The performance of a ridge waveguide laser may then become very similar to that of a gain-guided laser.

Rib Waveguide Laser. Rib waveguide lasers have attracted considerable attention.[34–36,39,41] The threshold current of these lasers is influenced by current spreading in the *p* cladding layer, by carrier diffusion in the active region, and by the lateral index step. The measured threshold currents are typically in the range of 60–80 mA. The lateral index guiding in a rib waveguide laser is schematically shown in Fig. 5.17b. A fraction of the lasing mode overlaps with the lower-index waveguide layer on either side of the active region. This overlap introduces a lateral index step. The effective index step depends on the width of the active region, the thickness of the waveguide region, and the thickness variation of the waveguide layer. An excellent control of layer thicknesses is necessary for fabrication of these

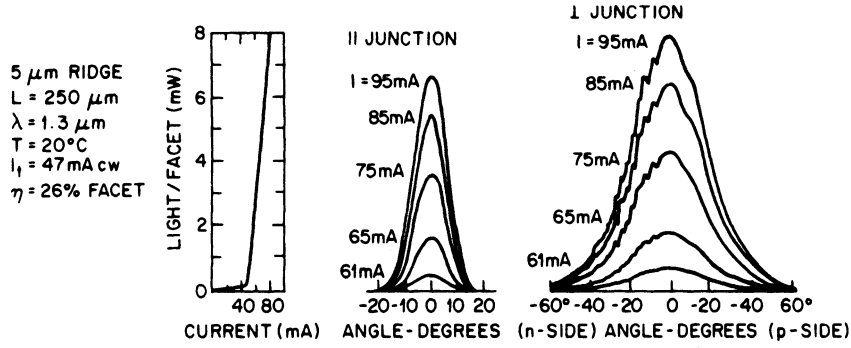


Fig. 5.16 Typical light-current (L-I) characteristics of a 1.3-μm InGaAsP ridge waveguide laser. Far fields along and normal to the junction plane are also shown. (After Ref. 42 © 1983 IEEE)

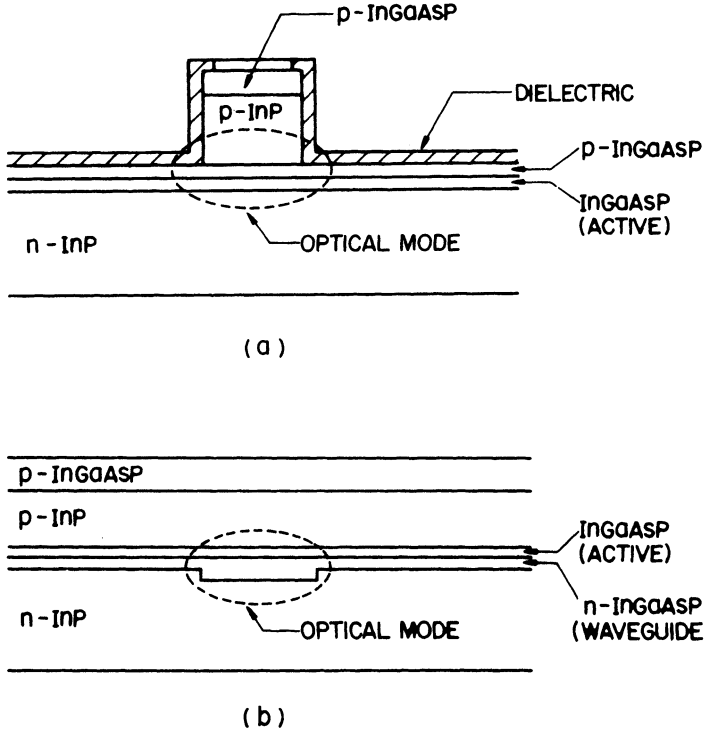


Fig. 5.17 Schematic representation of the optical mode in (a) ridge waveguide and (b) rib waveguide lasers. The dashed elliptical curve shows the spatial extent of the optical mode.

devices. Large values of active and waveguide layer thicknesses can reduce the lateral index step sufficiently so that the rib waveguide laser behaves as a gain-guided laser.

5.5 STRONGLY INDEX-GUIDED LASERS

The lateral mode control in injection lasers can be achieved using index guiding along the junction plane. The mode control is necessary for improving the light-current linearity and the modulation response of injection lasers. In strongly index-guided lasers, the active region is buried in higher band-gap layers (e.g., InP) on all sides. For this reason, these lasers are called buried-heterostructure lasers. The lateral index step along the junction plane is ~ 0.2 in these laser structures and is about two orders of magnitude larger than the carrier-induced effects. As a result, the lasing characteristics of buried-heterostructure lasers are primarily determined by the rectangular waveguide that confines the mode inside the buried active region. Lateral and transverse modes of buried-heterostructure lasers were discussed in Sec. 2.5.

The buried heterostructure was first proposed in 1974 for GaAs lasers.[45] Since then, several types of InGaAsP buried-heterostructure lasers have been developed.[46–54] These lasers can be divided into two categories: (i) planar-active-layer structures and (ii) nonplanar-active-layer structures. The planar-active-layer structure is more suitable for single-frequency distributed-feedback-type lasers because the waveguide modes of a planar active layer are more strongly coupled to a grating than are the waveguide modes of a nonplanar active region. The fabrication and performance characteristics of distributed-feedback buried-heterostructure lasers are described in Chapter 7. Generally, an optimized buried-heterostructure laser is characterized by a low threshold current (10–15 mA), can be operated to high output powers (~ 10 mW) in the fundamental transverse mode, and can be modulated at high frequencies (~ 2 GHz). Buried-heterostructure lasers are more difficult to fabricate than weakly index-guided or gain-guided lasers. However, these strongly index-guided structures are most suitable for demanding applications. In particular, they are useful as a source in high-bit-rate long-haul lightwave transmission systems.

Figure 5.18 shows a schematic cross section of several different types of buried-heterostructure lasers with a planar active layer. They are (i) the buried heterostructure[46] that we shall call the etched-mesa buried heterostructure (EMBH) in order to distinguish it from other buried-heterostructure lasers, (ii) the double-channel planar buried heterostructure[47] (DCPBH), (iii) the planar buried heterostructure,[48] and (iv) the strip buried heterostructure.[49] Variations of these structures also exist in the literature.

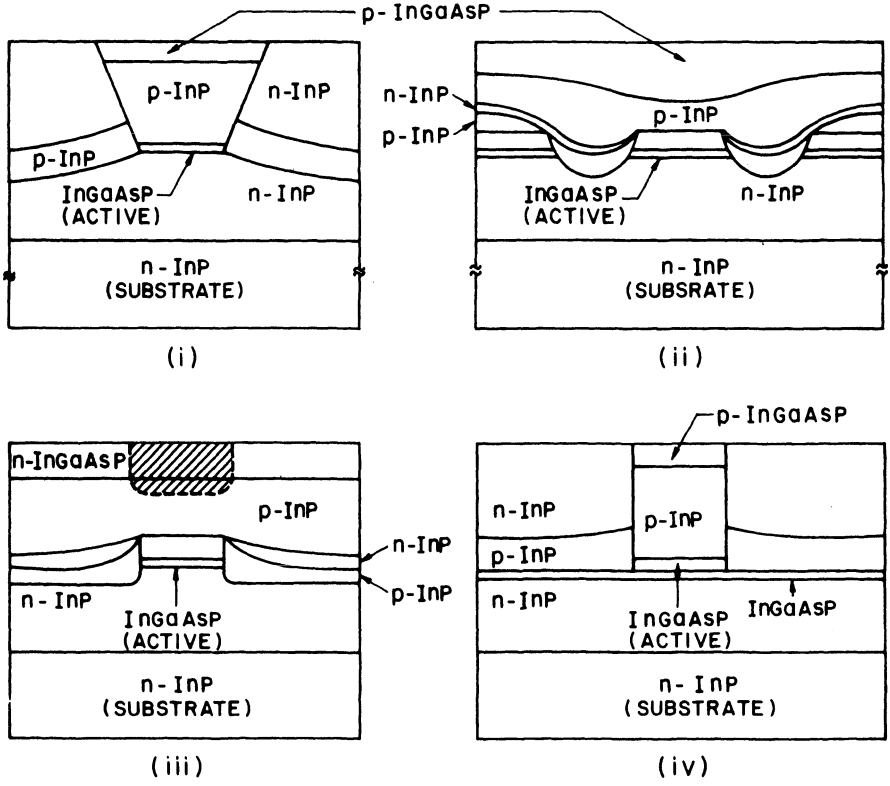


Fig. 5.18 Schematic cross section of different types of buried-heterostructure lasers with a planar active layer: (i) etched-mesa buried heterostructure; (ii) double-channel planar buried heterostructure; (iii) planar buried heterostructure, and (iv) strip buried heterostructure.

Figure 5.19 shows the schematic cross section of several nonplanar-active-layer laser structures. They are (i) the V-grooved substrate buried-heterostructure laser,[50] also known as the channeled-substrate buried-heterostructure (CSBH) laser[51] or the buried crescent[52] laser; (ii) the mesa-substrate buried-heterostructure (MSBH) laser,[53] and (iii) the buried-crescent buried-heterostructure (BCBH) laser.[54] Variations of these laser structures utilizing different current-confining schemes have also been proposed. We now describe the fabrication and performance characteristics of some of these laser structures.

Etched-Mesa Buried-Heterostructure (EMBH) Laser. Figure 5.20 shows an SEM photomicrograph of an EMBH laser. This device structure was first reported by Hirao et al.[46] The fabrication of this structure involves two

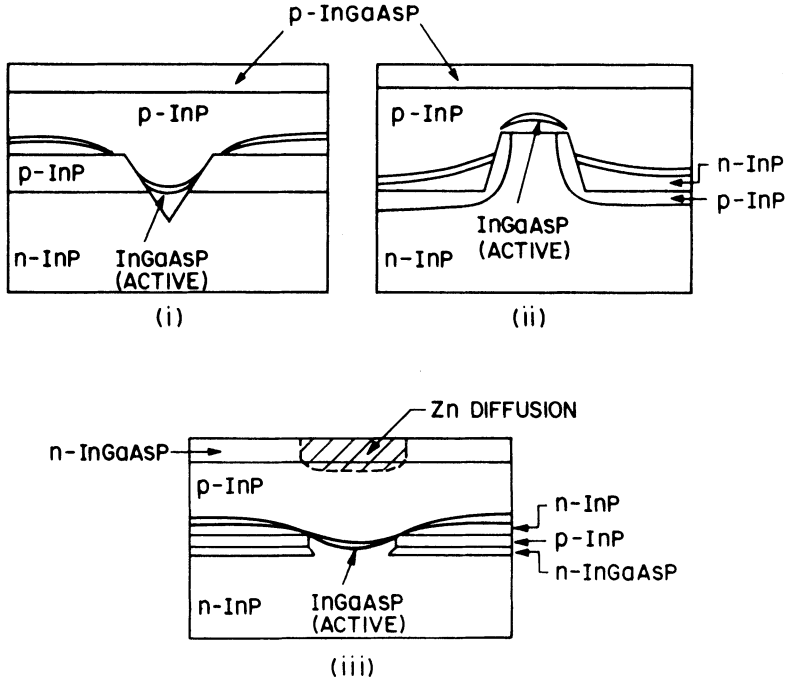


Fig. 5.19 Schematic cross section of different buried-heterostructure lasers with a nonplanar active layer: (i) V-grooved substrate or channeled-substrate buried heterostructure; (ii) mesa-substrate buried heterostructure; and (iii) buried crescent.

epitaxial growths. In the first, four layers are grown (by LPE, VPE, or MOVPE) on a (100)-oriented *n*-InP substrate. The four epitaxial layers are: (i) *n*-InP buffer layer ($\sim 2 \times 10^{18}$ cm^{-3} doping level, 3–4 μm thick); (ii) undoped InGaAsP active layer ($\lambda = 1.3$ μm, 0.2 μm thick); (iii) *p*-InP cladding layer ($\sim 10^{18}$ cm^{-3} doping level, 2 μm thick); and *p*-InGaAsP contact layer ($\lambda \cong 1.1$ μm, $\sim 10^{19}$ cm^{-3} doping level, 0.5 μm thick). Mesas 2–3 μm wide are then etched on the as-grown wafer using a solution of 1% Br_2 and CH_3OH with plasma-deposited SiO_2 acting as the etching mask. The etching mask is then removed and the *p*-InP and *n*-InP blocking layers are grown using a second LPE growth. The wafer is then cleaved to produce laser diodes after *p* and *n* contact metallizations.

Figure 5.21 shows the L-I characteristics of an InGaAsP ($\lambda = 1.3$ μm) EMBH laser at different temperatures. These lasers typically have room-temperature threshold currents in the range of 10–15 mA and can be operated in the fundamental transverse mode to powers greater than 10 mW/facet. The lateral mode control in strongly index-guided lasers such

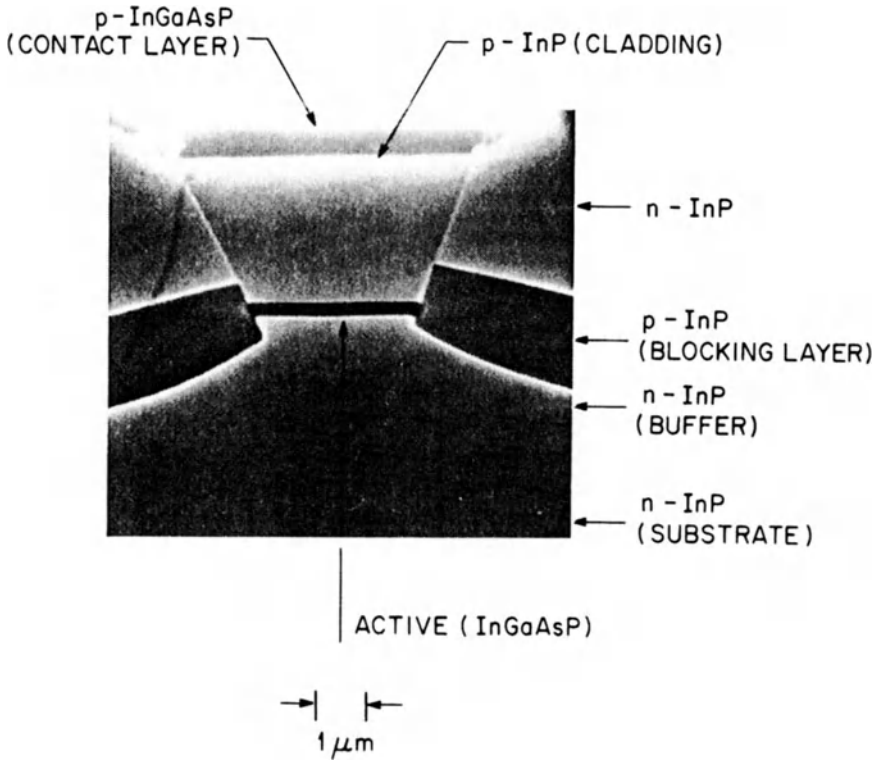


Fig. 5.20 SEM cross section of an EMBH laser.

as the EMBH laser can be achieved by controlling the width of the active region. As discussed in Sec. 2.5, lasers with a wide active region allow higher-order lateral modes. Hence, with increasing current, such lasers may switch from one lateral mode to another or a higher-order lateral mode can appear in addition to the fundamental mode. Such mode transitions are associated with a kink in the light-current characteristics similar to that shown in Fig. 5.7 for a gain-guided laser. Figure 5.22 shows the L-I characteristic at room temperature of an EMBH laser that exhibits a kink. The inset in Fig. 5.22 shows the measured far-field intensity along the junction plane at currents above and below the kink. The appearance of side lobes in the far field provides clear evidence that a higher-order lateral mode is lasing together with the fundamental mode.

The output power at which a kink appears is called the kink power. In Fig. 5.22 the kink power is about 1.5 mW. We expect the kink power to increase as the width of the active region decreases because the higher-order

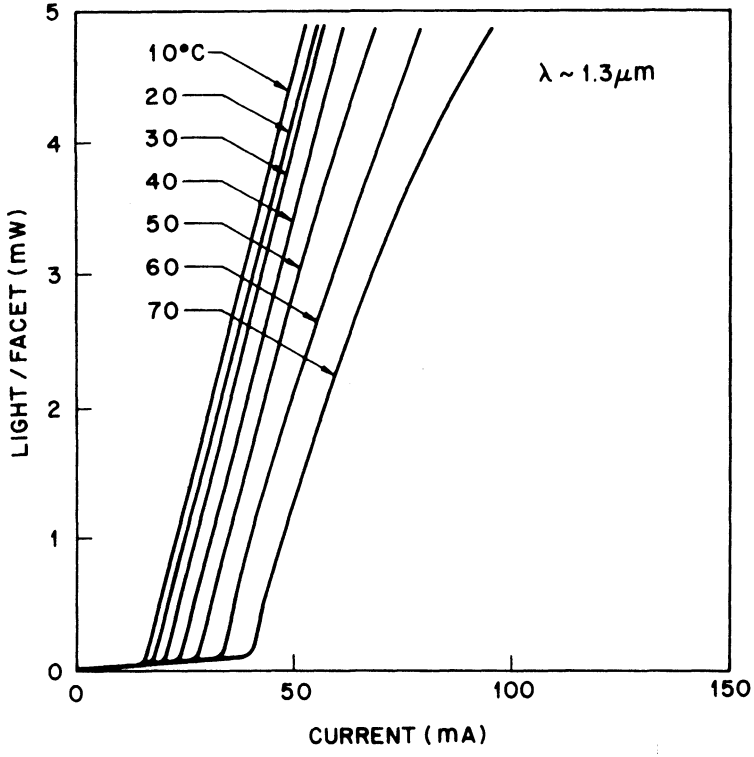


Fig. 5.21 L-I characteristics of an EMBH laser at different temperatures.

modes are then below or close to the cut-off level. Figure 5.23 shows the measured kink power as a function of the active-region width of several EMBH lasers. The active-region thickness of these lasers was 0.2 μm. The dashed curve in Fig. 5.23 is drawn as a visual aid indicating that for kink powers greater than 10 mW, the active-region width must be less than 1.6 μm for a 0.2-μm-thick active layer.[55]

In an EMBH laser structure, the *p-n-p-n* and other current-blocking junctions (see Fig. 5.20) confine the current to the active region.[56–58] This current confinement is generally not complete, and a fraction of the injection current flows (or leaks) around the active region in a buried-heterostructure laser. This leakage current should be small in an optimized device structure. An excessive leakage current can significantly affect the device threshold, the temperature performance, the modulation characteristics, and the output power characteristics. The dependence of leakage current on various device parameters and its effect on the performance characteristics are described in detail in Sec. 5.6.

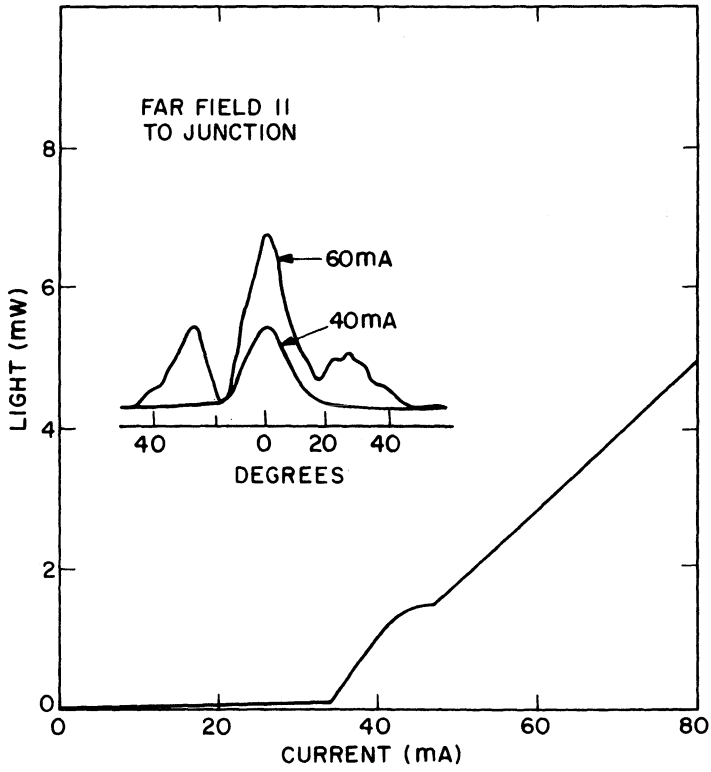


Fig. 5.22 L-I characteristics of an EMBH laser showing a kink. The measured far fields along the junction plane at currents above and below the kink are also shown. (After Ref. 55 © 1984 IEEE)

Double-Channel Planar Buried-Heterostructure (DCPBH) Lasers. Figure 5.24 shows an SEM photomicrograph of a DCPBH laser. The active region in this device consists of a 1.3-μm InGaAsP layer. The device fabrication involves two epitaxial growths. In the first, four layers are grown by LPE in a manner similar to that for the EMBH laser. The top InGaAsP layer is removed using an etching mixture of H_2SO_4, H_2O_2, and H_2O in proportions of 10:1:1. Two channels, about 3 μm deep and 5 μm wide, are then etched on the wafer using plasma-deposited SiO_2 as the etching mask. The mask is then removed and four epitaxial layers—*p*-InP, *n*-InP, *p*-InP, and *p*-InGaAsP ($\lambda = 1.3$ μm)—are grown on the wafer during a second LPE growth. The last quaternary layer serves as a contact layer. The wafer is then processed to produce laser chips in a procedure similar to that for gain-guided lasers

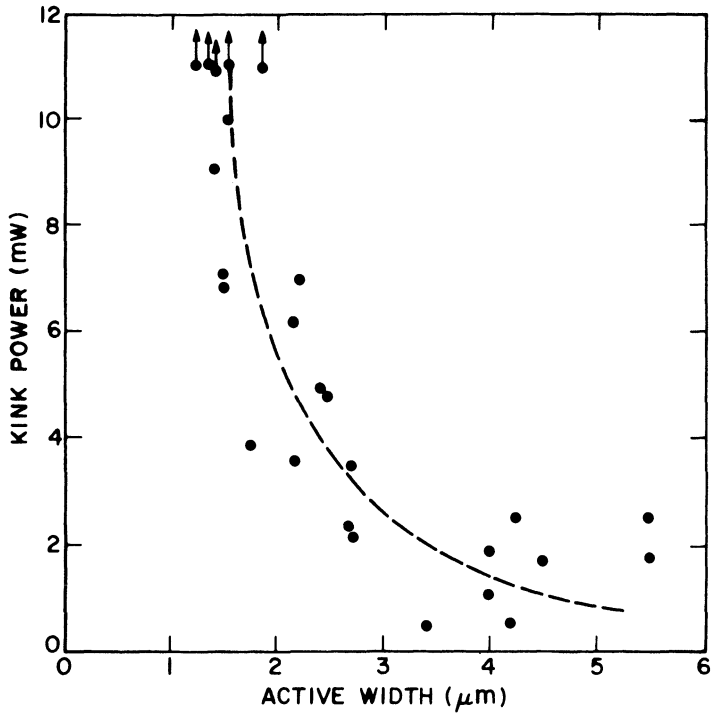


Fig. 5.23 Measured kink power (light output at which a kink appears) as a function of the active-region width of EMBH lasers. Arrows indicate that the kink power was more than 11 mW for these lasers. Dashed line is drawn as a visual aid. (After Ref. 55 © 1984 IEEE)

except that the contact stripe is aligned over the active region using photolithographic techniques.

Figure 5.25 shows the light-current characteristics of a 1.3-μm InGaAsP DCPBH laser at different temperatures. The device can operate CW up to 130°C. DCPBH lasers typically exhibit room-temperature threshold currents in the range of 10–15 mA. The lateral mode control is achieved by reducing the dimension of the active region. Experimental data for the kink power as a function of the active region width are similar to those shown in Fig. 5.23.

InGaAsP lasers emitting near 1.55 μm are also of considerable interest as a source in optical fiber communication systems. The fabrication procedure for these lasers is similar to that described above except that the active-layer composition is chosen so that the band gap corresponds to the 1.55-μm wavelength. Figure 5.26 shows the light-current characteristics at different temperatures of a DCPBH laser emitting near 1.55 μm.[59] The laser operates

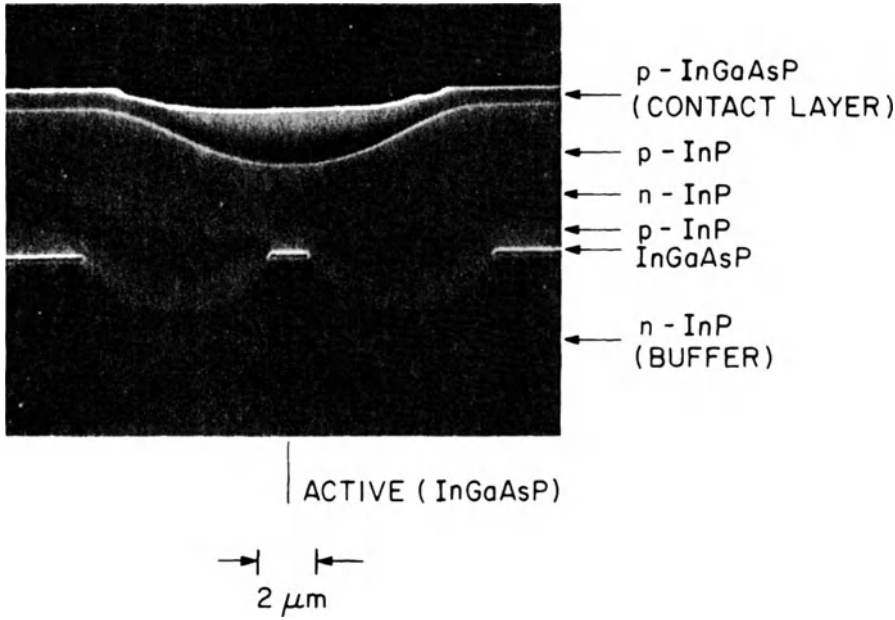


Fig. 5.24 SEM cross section of a DCPBH laser.

CW up to 110°C. A comparison of Figs. 5.25 and 5.26 shows that 1.55-μm InGaAsP lasers have a higher threshold current and a lower maximum operating temperature than do 1.3-μm lasers. As discussed in Chapter 3, increased Auger recombination at longer wavelengths is partly responsible for this behavior. The lasers emitting near 1.55 μm also exhibit a lower differential quantum efficiency than do lasers emitting near 1.3 μm. This may be due to increased intervalence band absorption at longer wavelengths.[60]

V-Grooved or Channeled-Substrate Buried-Heterostructure (CSBH) Laser. Figure 5.27 shows an SEM photomicrograph of a CSBH laser. This device structure differs from the EMBH and the DCPBH in that the active region is nonplanar. The *p*-InP current-blocking layer (bottom *p*-InP layer in Fig. 5.27) can be grown by LPE or VPE on a (100)-oriented *n*-InP substrate. Sealed-tube Cd-diffusion using a Cd_3P_2 source at 600–650°C can also be used to type-convert a thin layer (~ 2 μm) on the surface of *n*-InP substrate.[61] This latter method allows the fabrication of the device using a single LPE step. V-shaped grooves are then etched on the wafer parallel to (110) using a 3:1 mixture of HCl and H_3PO_4 and an SiO_2 etching mask. Four epitaxial

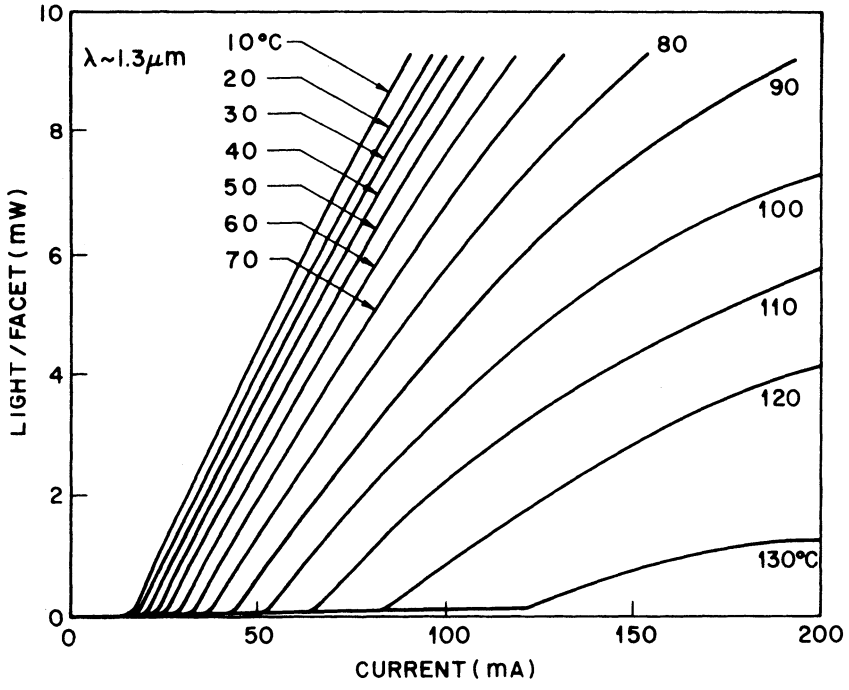


Fig. 5.25 L-I characteristics of a 1.3-μm InGaAsP DCPBH laser at different temperatures.

layers are then grown by LPE. These are (i) an *n*-InP buffer layer, (ii) an undoped InGaAsP active layer ($\lambda = 1.3$ μm), (iii) a *p*-InP cladding layer, and (iv) a *p*-InGaAsP contact layer. The wafer is then processed to produce CSBH lasers with a procedure very similar to that used for oxide-stripe gain-guided lasers.

Figure 5.28 shows the CW light-current characteristics of a 1.3-μm CSBH laser at different temperatures. The threshold current of these lasers is typically in the range of 10–15 mA at room temperature, which is similar to that for EMBH and DCPBH lasers. Lasers emitting near 1.55 μm have also been fabricated using the CSBH structure with performance characteristics similar to those obtained for DCPBH lasers.

Lateral mode control, needed for the kink-free operation up to high powers, can be achieved by controlling the dimensions of the active region. Experimental data for the kink power as a function of the area of the active region is shown in Fig. 5.29 for CSBH lasers. The active area A is defined as $\frac{2}{3}wd$, where w and d are the width and the thicknesses of the crescent-shaped active region. Figure 5.29 shows that fundamental transverse mode

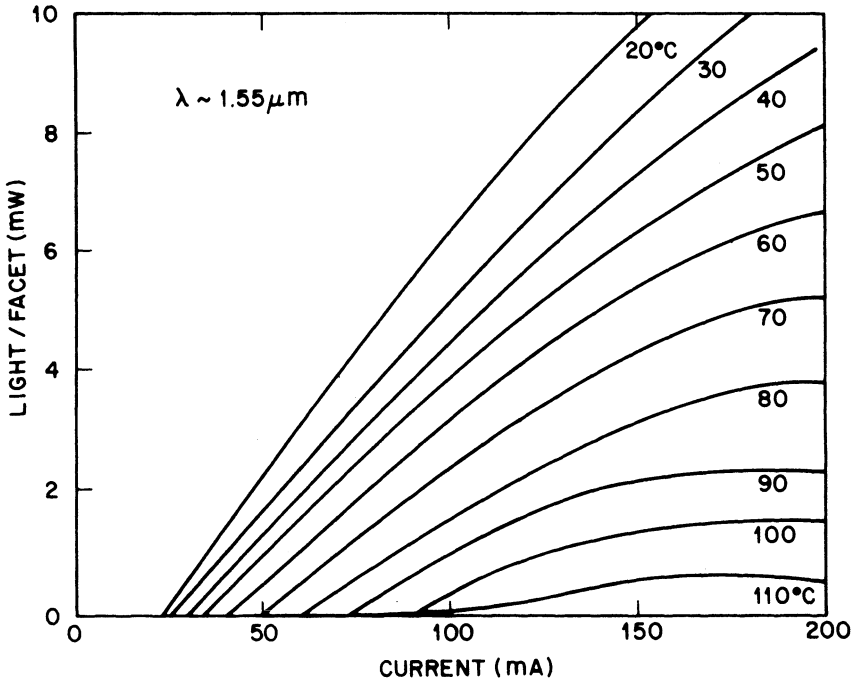


Fig. 5.26 L-I characteristics of a 1.55-μm InGaAsP DCPBH laser at different temperatures. (After Ref. 59)

operation can be achieved for lasers with an active area less than 0.3 μm^2. This result is similar to that shown in Fig. 5.23 for EMBH lasers.

Mesa-Substrate Buried-Heterostructure (MSBH) Laser. Figure 5.30 shows an SEM cross section of an MSBH laser. This device structure can be fabricated using one LPE growth. The device fabrication involves the following steps:[62] A mesa 7–8 μm wide and 5 μm deep is etched on a (100)-oriented *n*-InP substrate. Mesa etching is done using a solution of HCl and H_3PO_4 and a plasma-deposited SiO_2 stripe mask oriented parallel to the (011) direction. After etching, the SiO_2 stripe is removed and a type-converted region (*p*-InP) is created using sealed-ampule Cd diffusion. Then four epitaxial layers are grown: (i) an *n*-InP buffer layer, (ii) an undoped InGaAsP ($\lambda = 1.3$ μm) active layer, (iii) *p*-InP cladding layer, and (iv) a *p*-InGaAsP contact layer. The wafer is then processed into laser chips using standard fabrication procedures.

MSBH lasers typically exhibit room-temperature threshold currents in the range of 15–30 mA. The MSBH laser was initially fabricated without the

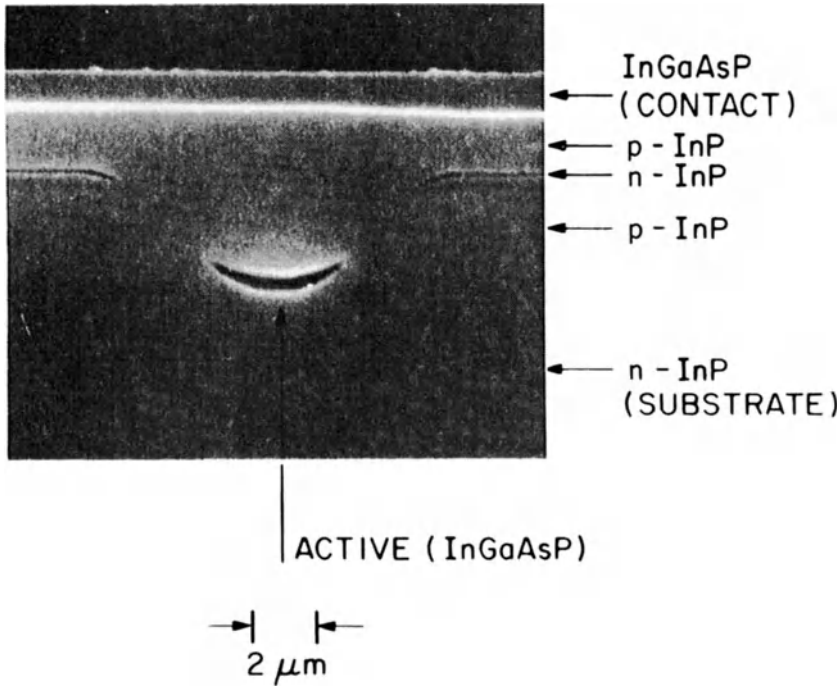


Fig. 5.27 SEM cross section of a CSBH laser.

Cd-diffused layer.[53] In the absence of a *p*-type current-blocking layer, these devices have a somewhat higher threshold current compared with that of the EMBH, DCPBH, and CSBH lasers due to a higher leakage current; i.e., the current flowing around the active region is higher in the absence of a current-blocking layer. The lateral mode control in this laser structure is also achieved by controlling the dimensions of the active region in a manner similar to CSBH lasers.

Other Device Structures. One of the technological issues associated with the fabrication of high-performance index-guided lasers is the control of leakage current. The device structures described above utilize junctions between different semiconductors for confining the current to the active region. Many laser structures, some of which are variations of these structures, use semi-insulating layers[63,64] (e.g., Be- or Fe-doped InP) or highly resistive layers (undoped InP) for current confinement. Figure 5.31 shows three such structures. They are (a) a CSBH laser with an Fe-implanted high-resistivity layer for current confinement[63]; (b) a vapor-phase

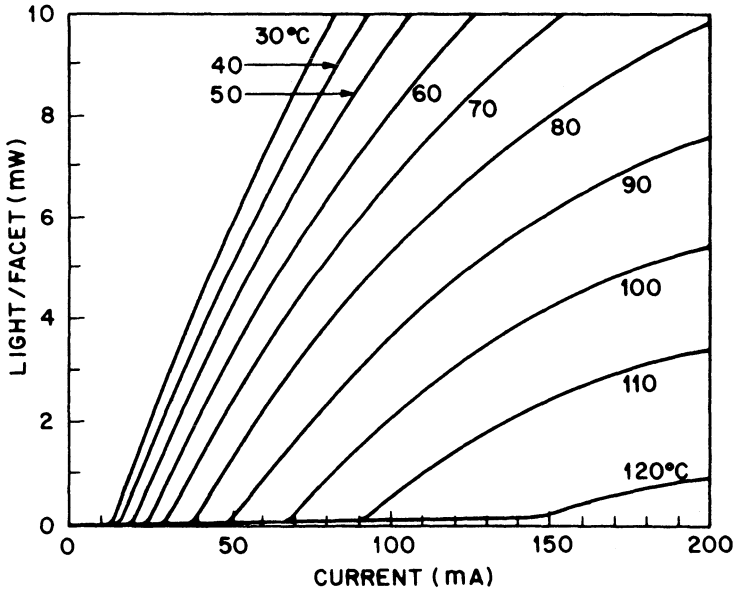


Fig. 5.28 L-I characteristics of a 1.3-μm InGaAsP CSBH laser at different temperatures.

mass-transported buried-heterostructure laser employing an undoped InP layer for current confinement[65–67]; and (c) a mushroom-stripe buried-heterostructure laser with SiO_2 layers for current confinement.[68,69] The main advantage of these structures appears to be the reduction in electrical parasitics because of the absence of reverse-biased junctions. As a result, the device is capable of high-frequency modulation.[67,69] The modulation performance of semiconductor lasers is discussed in Sec. 6.6.

5.6 LEAKAGE CURRENT

In buried-heterostructure lasers of the types shown in Figs. 5.18 and 5.19, the confinement of the injected current to the active region is achieved through reverse-biased homojunctions or forward-biased heterojunctions. The difference between the total current I injected into the laser and the current I_a passing through the active region is the leakage current I_L, i.e.,

$$I_L = I - I_a. \tag{5.6.1}$$

The magnitude of I_L and its variation with I depend on the thickness and the carrier concentration of the current-confining layers. As discussed in

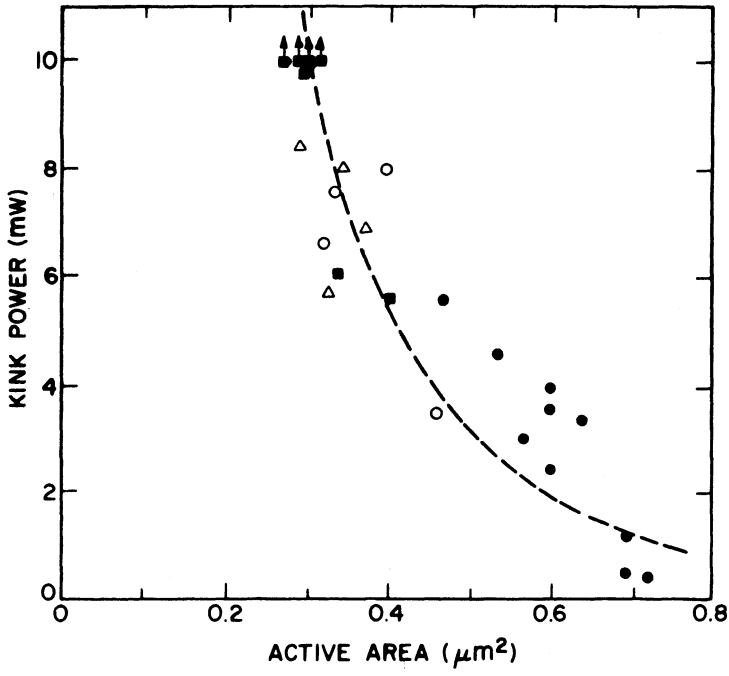


Fig. 5.29 Measured kink power as a function of active area in CSBH lasers. Data from four different wafers are shown. Arrows represent kink powers greater than 10 mW. (After Ref. 51)

Sec. 2.6.1, a large leakage may lead to a high threshold current and also to sublinearity or power saturation in the light-current characteristics. Figure 5.32 shows L-I curves for CSBH lasers with high and low leakage at two different temperatures. The dashed curves represent the slope efficiency (dL/dI versus I), which decreases with an increase in the total current for the high-leakage device. Note also that the performance of the high-leakage device degrades severely at high temperatures.

Before going into details of leakage paths in various laser structures, we first consider physically how a forward-biased junction can confine the current at low injection and become "leaky" at high injection. For the purpose of illustration, we consider an EMBH type of laser structure. Figure 5.33a shows two current paths for I_L and I_a and the corresponding equivalent electrical circuit model is shown in Fig. 5.33b. The laser and the InP homojunction are represented by forward-biased diodes in the equivalent circuit. R_s is a resistor in the laser diode path and R_2 a resistor in the path of the InP homojunction diode. Figure 5.34 shows schematically the

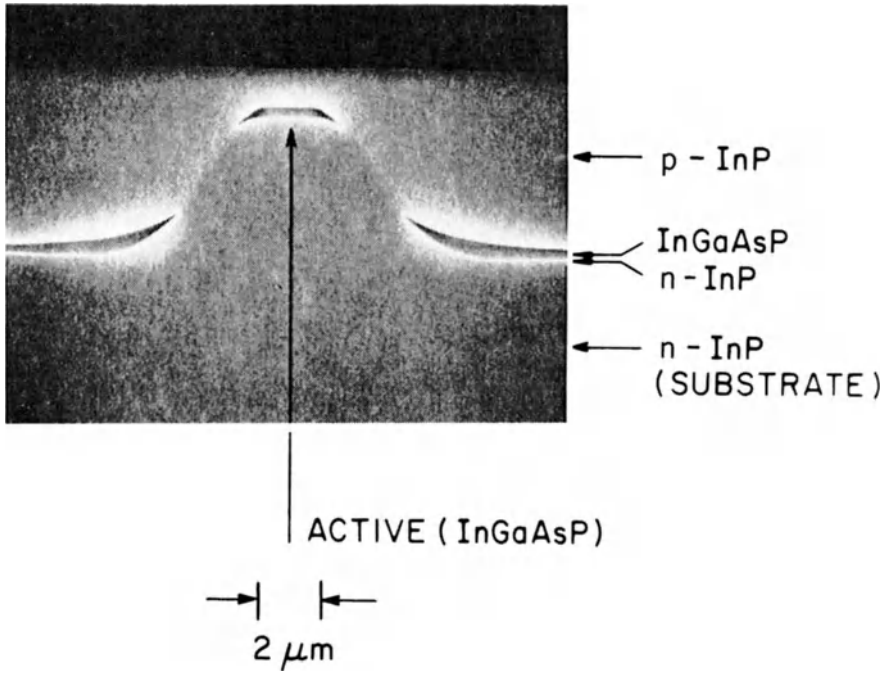


Fig. 5.30 SEM cross section of an MSBH laser.

current-voltage (I-V) characteristics of the laser diode with the series resistance R_s. The dashed line is the I-V characteristic of the homojunction-leakage diode. Figure 5.34 shows that as the external voltage is increased, most of the current initially passes through the active region, until the voltage reaches the turn-on voltage of the leakage diode. With a further increase in voltage, the fraction of leakage current increases, and at high current, the majority of the injected current flows through the leakage junction.

We now discuss the leakage current paths in an EMBH laser. The laser structure and its equivalent-electrical-circuit model are shown in Figs. 5.33a and 5.33c. For the purpose of discussion, the leakage current paths are identified by a major circuit element in each branch, e.g., the diode (p-n forward-biased InP homojunction), the transistor (n-p-n InP), and the thyristor (p-n-p-n InP). Various layers participating in the three leakage current paths are described below (refer to Fig. 5.20 for the terminology).

For the diode leakage path, the forward-biased p-InP blocking layer and the n-InP buffer layer form the diode. The effective saturation parameter for this diode is determined by current spreading in the p-InP layer. Current

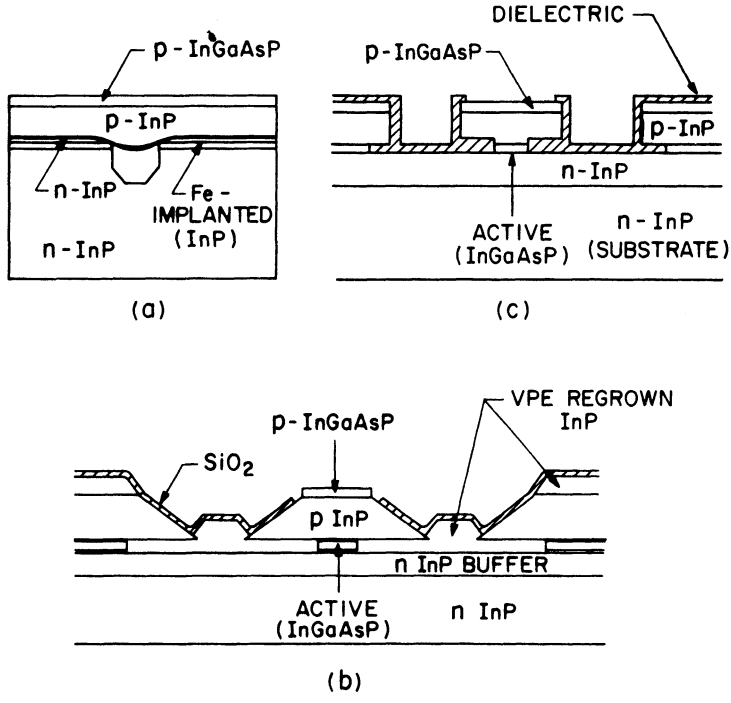


Fig. 5.31 Schematic illustration of three InGaAsP laser structures: (a) CSBH laser with an Fe-implanted layer; (b) mass-transported buried-heterostructure laser; and (c) mushroom-stripe laser.

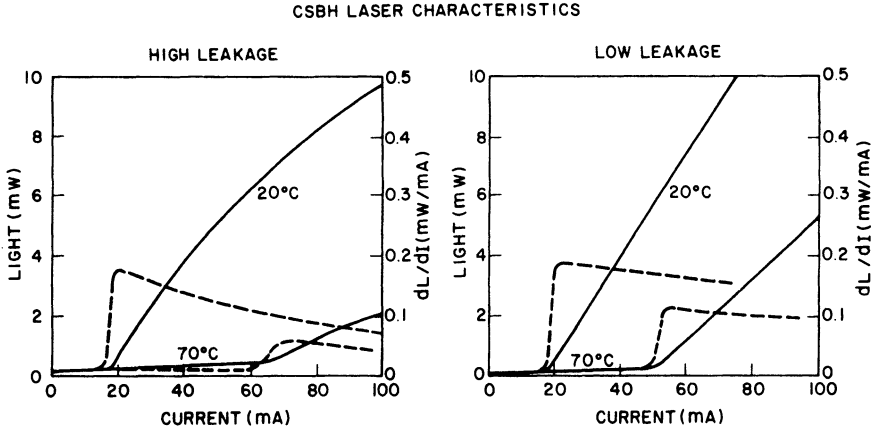


Fig. 5.32 L-I characteristics at two different temperatures for two CSBH lasers with high and low leakage currents. The dashed curves show the slope efficiency.

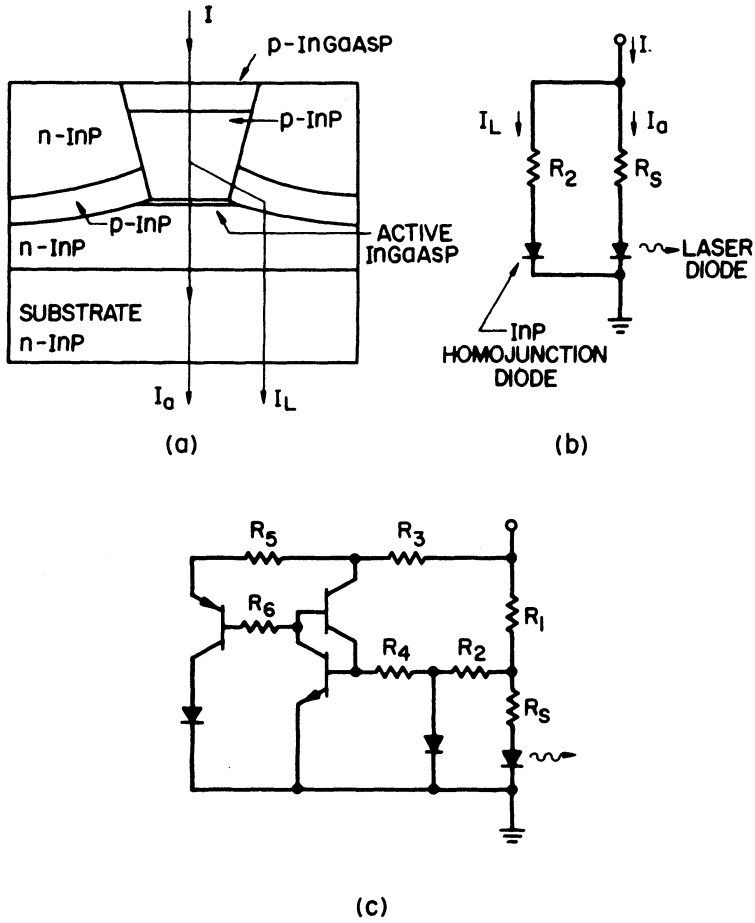


Fig. 5.33 (a) Schematic cross section of an EMBH laser. (b) Simplified electrical equivalent circuit of a laser diode in parallel with an InP homojunction diode. (c) Electrical equivalent circuit of an EMBH laser.

spreading is less for thinner *p*-InP lasers; hence the diode leakage is expected to be smaller for a thin *p*-InP layer.

The *n*-*p*-*n* transistor in the transistor leakage path consists of the *n*-type buffer layer (emitter), the *p*-InP blocking layer (base), and the *n*-InP top regrown layer (collector). Increasing the width of the *p*-InP blocking layer increases the base width of the transistor, which reduces the current gain of the transistor and hence the leakage through the transistor path.

The various layers corresponding to the *p*-*n*-*p*-*n* thyristor path are the *p*-InP cladding layer, the *n*-InP top regrown layer, the *p*-InP blocking layer,

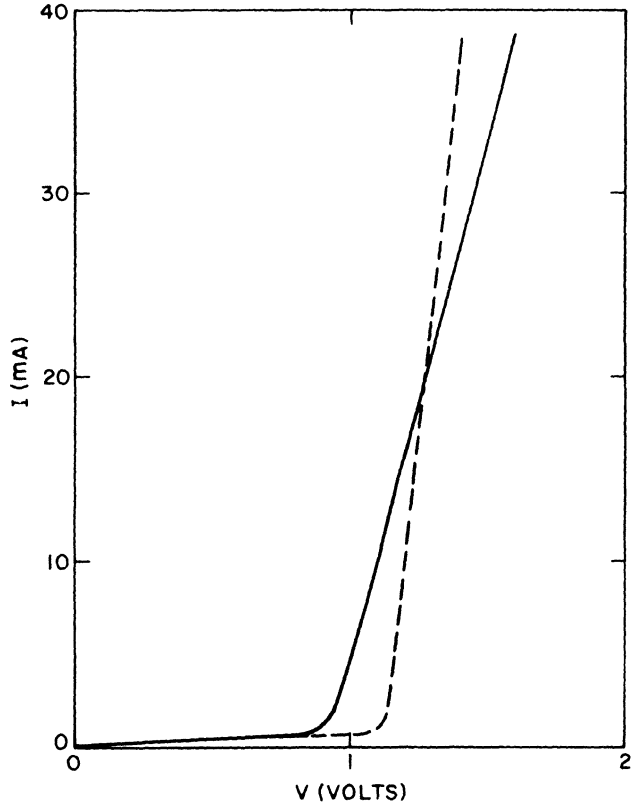


Fig. 5.34 *I-V* characteristics of the laser diode path (solid curve) and the homojunction diode path (dashed curve). The two paths are shown in Fig. 5.33b, as I_a and I_L, respectively.

and the *n*-InP buffer layer. At low currents the leakage through this path is negligible because the thyristor is off. However, at high currents when the thyristor is turned on, the leakage current through this path may be significant.

The magnitude of the leakage current through each of the leakage paths shown in Fig. 5.33c can be calculated by analyzing the equivalent electrical circuit.[58] The results of the calculation depend strongly on the saturation parameters of the various circuit elements and on the ideality factor of the diode. These quantities are determined by the quality of the interface between layers, in addition to layer thicknesses and the doping levels. The interface quality is determined by the details of device fabrication such as types of etching agents used, growth temperatures, and conditions prior to regrowth that affect the thermal decomposition of exposed surfaces. These factors

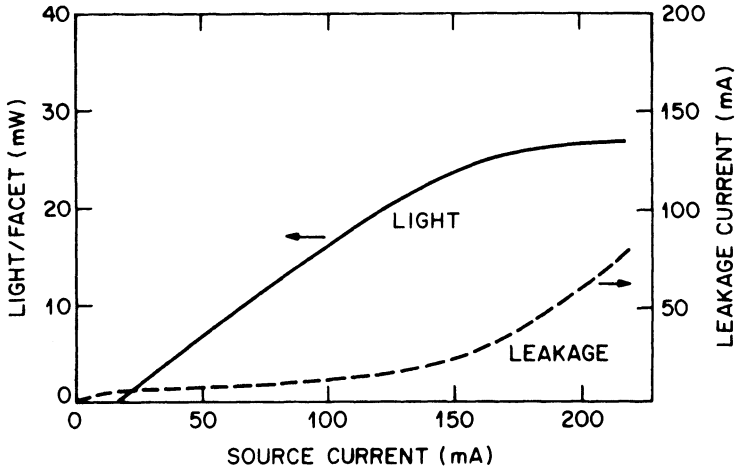


Fig. 5.35 Calculated L-I characteristic of an EMBH laser after including the current leakage paths modeled by the circuit shown in Fig. 5.33c. The dashed curve shows the leakage current.

complicate the comparison of the experimental results with the calculated light-current characteristic obtained with the electrical-equivalent-circuit model. Figure 5.35 shows representative results. The dashed curve is the leakage current, principally through the thyristor path, and the solid curve is the light-current characteristic. The calculated light output per facet is obtained using an external differential quantum efficiency of 0.2 mW/mA per facet. Due to the increase in the leakage current, the output power saturates at high currents. Under CW operation there is an additional heating effect that can "roll over" the light-current characteristics; i.e., the light output can decrease with increasing current.

5.7 LASER ARRAYS

InGaAsP semiconductor laser structures described so far have been developed for low-power (5–10 mW/facet) applications, such as providing a source in a lightwave communcation system. With a proper design, these lasers can generally emit continuously up to powers in the range 30–60 mW/facet near room temperature. This limitation in the output power mainly arises from the leakage current, which increases with an increase in the applied current. Semiconductor laser diodes for high-power operation have been extensively studied using the AlGaAs material system.

High-power lasers developed so far can be classified as one of two types, the single-emitter laser[70–76] and the laser array.[77–90] A laser array has closely

spaced multiple active regions, each of which emits light. The radiations from neighboring emitters in a laser array overlap, establishing a definite phase relation between them. As a result the array emits in a narrow, coherent beam.

Single-emitter laser structures are similar to those described in the previous section. The leakage current limits the highest operating power from these devices. Nonetheless, 140 mW of CW power has been obtained[74] from a 1.3-μm DCPBH InGaAsP laser with design optimization and facet coatings. The leakage current can be reduced in a laser structure with a *p*-InP substrate.[72,73] The use of unstable-resonator geometry[75,76] has also provided output powers of ~400 mW in a single lateral mode.

Laser-array structures using the AlGaAs material systems have been extensively studied.[77–84] More recently, InGaAsP laser arrays have attracted attention.[85,86] Laser-array structures can be classified into two groups: gain-guided arrays and weakly index-guided arrays. The difference between the two types is that the mode profile along the junction plane is determined by gain guiding for the former and by weak index guiding for the latter. Figure 5.36 shows three laser-array structures. The gain-guided array is a dielectric-stripe device. The openings in the dielectric (SiO_2) and their

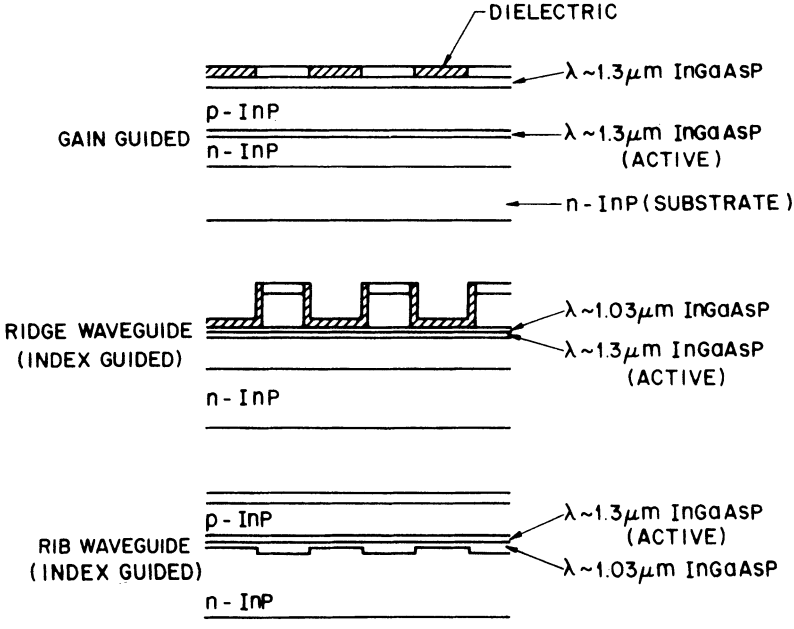


Fig. 5.36 InGaAsP laser-array structures shown schematically.

center-to-center spacings are chosen so that there is adequate overlap between oscillating modes of neighboring emitters. The weakly index-guided laser array can employ a ridge waveguide or a rib waveguide structure. The dimensions of and spacing between emitters are again chosen in such a way as to provide adequate overlap of the oscillating modes. This overlap establishes the coupling necessary for phase-locked operation. Figure 5.37 shows the L-I curves for a 1.3-μm gain-guided InGaAsP laser array with ten emitters. InGaAsP laser arrays emitting at 1.3 μm have been operated to pulsed output powers of 800 mW/facet.

An often-encountered problem in the performance of a laser array is that the far field along the junction plane consists of two lobes, a few degrees apart, whereas a single-lobe far field is desirable for practical applications. The nature of the far field depends critically on the phase relationships between neighboring emitters, and single-lobe and twin-lobe far fields are generally related to in-phase and out-of-phase operations, respectively.[77] The

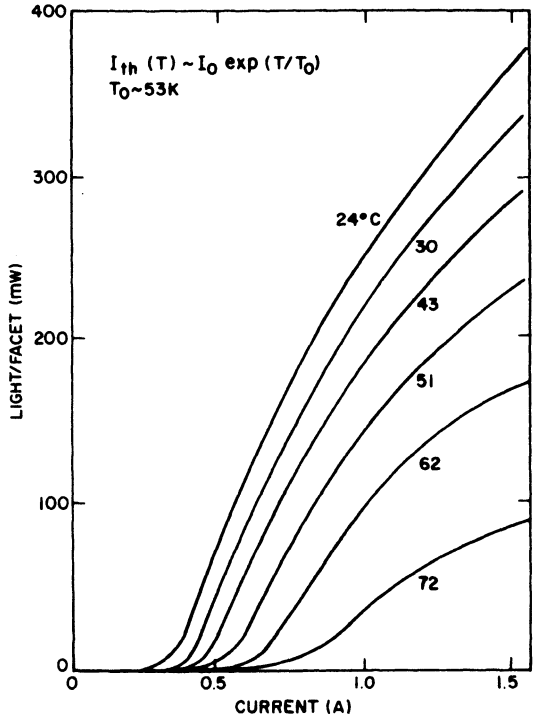


Fig. 5.37 L-I characteristics at several temperatures for a gain-guided 1.3-μm InGaAsP laser array.

relative phase between adjacent emitters, however, is generally determined by the lateral variation of the optical gain and the refractive index. With proper design, single-lobe far-field operation has been observed in some semiconductor laser arrays. Figure 5.38 shows such far fields obtained under the in-phase operation of a 1.3-μm InGaAsP gain-guided laser array. The subsidiary maxima in the far field are due to the presence of other transverse modes. Such higher-order modes have slightly different wavelengths and can

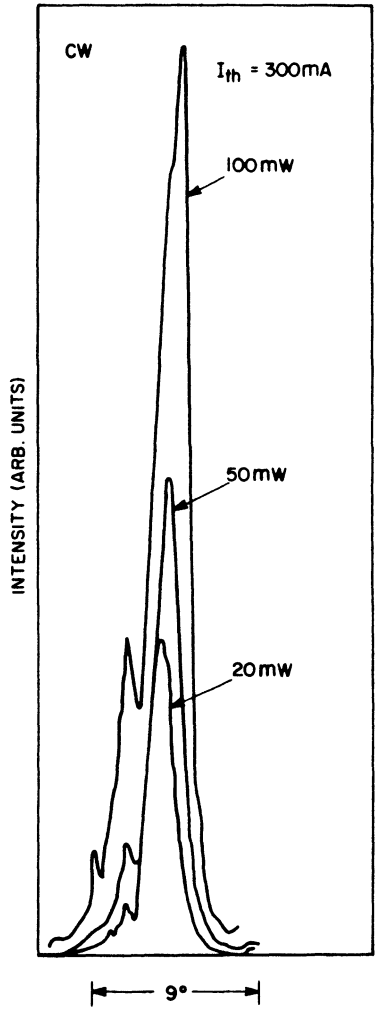


Fig. 5.38 Far fields along the junction plane of a gain-guided 1.3-μm InGaAsP laser with ten emitters showing in-phase operation.

be identified[83] in the optical spectrum. Two schemes have been proposed that lead to single-lobe far fields. These are (i) a chirped laser array, i.e., a laser array with variable emitter dimensions or a variable spacing between emitters[78,81]; and (ii) a laser structure with higher optical gain in between the emitters.[87]

The normal modes of a phase-locked laser array have been studied using coupled-mode analysis.[88,89] These normal modes, generally referred to as the supermodes or array modes, oscillate at different wavelengths with different near-field and far-field patterns. Coupled-mode analysis is very useful for a qualitative analysis of device behavior. In a more recent approach[90] the modes of the composite-array cavity are determined through a self-consistent solution of the paraxial wave equation that incorporates lateral variations of both the optical gain and the refractive index. Even though the concept of coupling between emitters is not invoked, the near field has a multipeak profile with each peak confined to a single-emitter region. This approach is useful for understanding the role of the lateral guiding mechanism and its influence on in-phase and out-of-phase coupling between emitters.

An important aspect of laser-array research is the determination of suitable structures capable of producing high output power in a diffraction-limited beam. Schemes that have been demonstrated for this purpose include a diffractive-type spatial filter,[91] Talbot cavities,[92–94] antiguided arrays[95,96] and Y-junction arrays.[97] Of all these schemes, antiguided laser arrays have produced very high power levels in a nearly diffraction-limited beam. Linear arrays of antiguided lasers have strong interelement coupling due to the presence of leaky waves. The structure of an antiguided laser array[95,96] is shown in Fig. 5.39. The fabrication process consists of two MOVPE growth

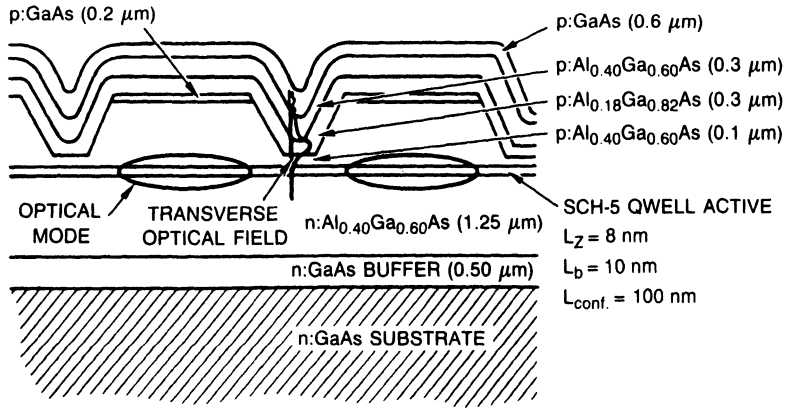


Fig. 5.39 Schematic of an antiguided phase-coupled laser array. (After Ref. 95)

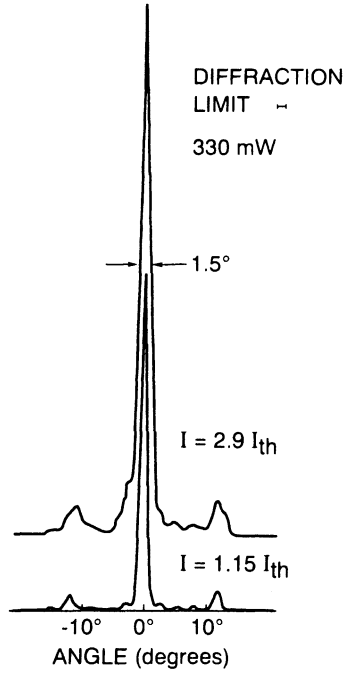


Fig. 5.40 Far-field radiation patterns at two current levels from a 20-element array with optimized facet coatings. (After Ref. 96)

steps. The first step involves the growth of the active region and the *n*- and *p*-type cladding layers. The second step involves the growth of the *p*-type $Al_{0.18}Ga_{0.82}As$ layer and the layers that follow as shown in Fig. 5.39. The lateral antiguiding property results from placing the $Al_{0.18}Ga_{0.82}As$ guide layer in close proximity (≤ 0.2 μm) to the active layer.

The lateral far-field pattern of a 20-element array is shown in Fig. 5.40. Close to threshold the output beam is diffraction-limited and at 3 times threshold the beam width is 1.5° (1.8 × diffraction limit) with 330 mW output power.

5.8 SURFACE-NORMAL EMITTING LASERS

A conventional semiconductor laser has cleaved facets that form the optical cavity (see Fig. 5.1). The mirror facets are perpendicular to the surface of the wafer and light is emitted parallel to the surface of the wafer. Many applications require semiconductor lasers emitting light normal to the wafer

surface. Two categories of laser exist where light is emitted normal to the surface of the wafers. In the first category are lasers in which the optical cavity is normal to the surface of the wafer. These lasers are also known as vertical-cavity surface-emitting lasers and are described in Chapter 10. The

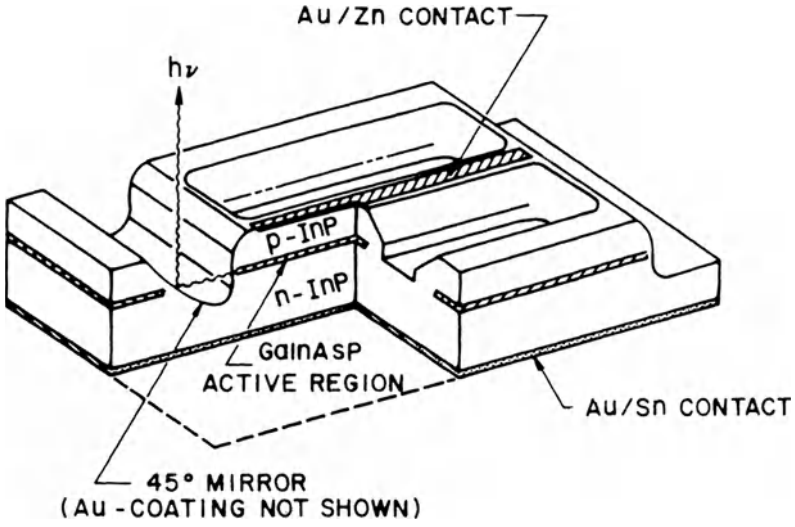


Fig. 5.41 Cross section of a buried heterostructure laser in which surface-emitting light output is obtained using an etched mirror. (After Ref. 98)

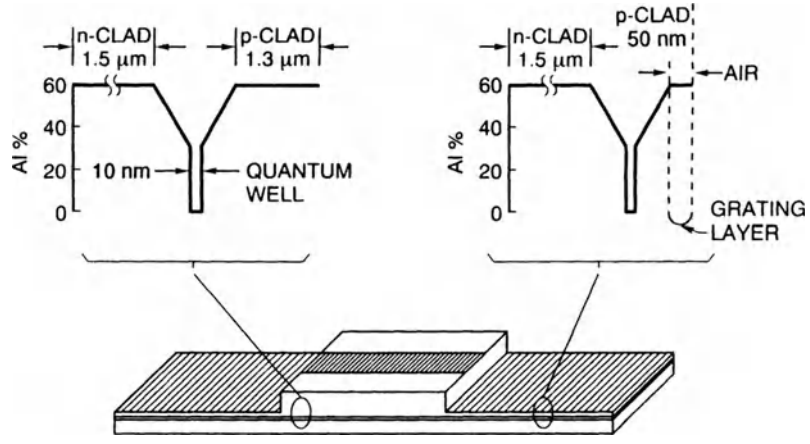


Fig. 5.42 Schematic of a grating-coupled surface-emitting laser. Lasers utilize a second-order grating and the light is emitted normal to the wafer surface. (After Ref. 99)

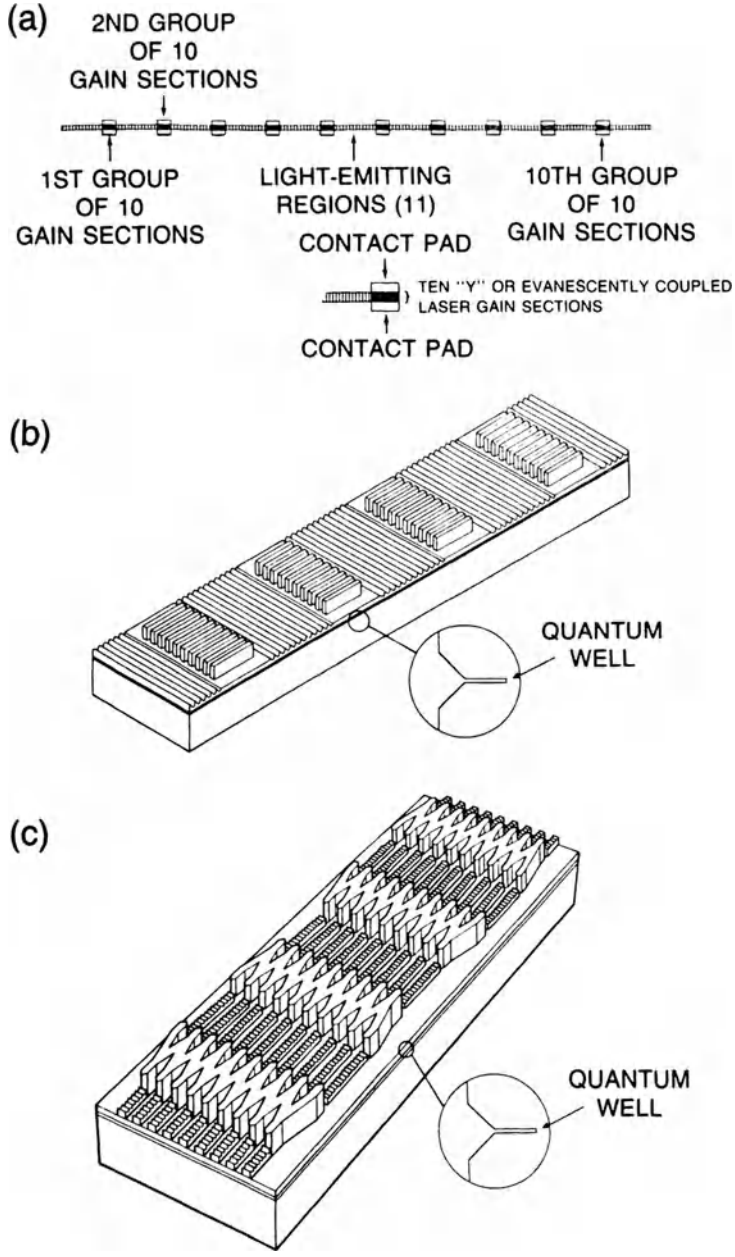


Fig. 5.43 (a) Overall view of a 10 × 10 array. (b) Detailed sketch showing four gain sections with evanescent coupling. (c) Sketch showing four gain sections with Y-coupling. (After Ref. 100 © 1989 IEEE)

second category consists of lasers where the active region has the conventional waveguide form, but the light is deflected normal to the surface using a mirror or a diffraction grating fabricated adjacent to the laser waveguide. Such lasers are described in this section.

Liau and Walpole[98] have reported a buried-heterostructure laser whose light output is deflected perpendicular to the wafer surface, so that the laser, in effect, is a surface-emitting laser. The laser structure is shown in Fig. 5.41. The mirrors of the laser are formed by the technique of vapor-phase transport. The beam is deflected normal to the surface using a monolithically integrated 45° parabolic mirror. Such lasers have room-temperature threshold currents in the range of 12–18 mA and exhibit performance characteristics comparable to those of conventional cleaved-facet lasers.

A grating can be used to deflect the laser output perpendicular to the wafer surface, such lasers are called grating-coupled surface-emitting lasers.[99–102] The schematic of the laser is shown in Fig. 5.42. A diffraction grating is etched on the *p*-type cladding layer, outside the waveguide region, to couple the light out vertically. Bragg gratings also serve as end mirrors for the laser. Threshold current densities as low as 400 A/cm^2 have been demonstrated using these devices. Diffraction gratings operate in the second-order diffraction mode. Lasers of this type have been fabricated using both the GaAs-AlGaAs ($\lambda \sim 0.85$ μm) and the InGaAs-GaAs ($\lambda \sim 0.9$–1.05 μm) material systems.[101]

The grating-coupled surface-emitting structure has been used to fabricate high-power laser arrays. A schematic of a two-dimensional 10 × 10 array is shown in Fig. 5.43a. The active region in this device is continuous throughout the wafer. In Fig. 5.43b the arrays are coupled by the evanescent light, and in Fig. 5.43c a Y-branch waveguide is used for coupling. Two-dimensional arrays of this type are useful for applications requiring high power.

PROBLEMS

5.1 Use the theory of Secs. 2.5 and 2.6 to explain the rapid increase in the threshold current density of broad-area semiconductor lasers when the active layer becomes thinner than an optimum value (see Fig. 5.2).

5.2 Describe the three commonly used techniques for making gain-guided InGaAsP lasers. Use diagrams whenever necessary.

5.3 Use Eq. (5.3.4) to show that the quantity $I\,dV/dI$ exhibits a kink at the laser threshold.

5.4 Explain the difference between weakly and strongly index-guided semiconductor lasers. Give an example of each kind of laser.

5.5 Draw a cross section of a double-channel planar buried heterostructure (DCPBH) laser. Indicate the processing steps involved in its fabrication.

5.6 What is the maximum width of the active layer of strongly index guided 1.55-μm semiconductor lasers to guarantee emission in a single lateral mode? Use the theory of Sec. 2.5 with $\mu_2 = 3.5$, and $\mu_1 = 3.2$, and $d = 0.1$ μm.

REFERENCES

1. Antypas, G. A., R. L. Moon, L. W. James, J. Edgecumbe, and R. L. Bell. III–IV Quaternary alloys, p. 48 in *Gallium Arsenide and Related Compounds*, 1972 Conf. Ser. 17. London: Institute of Physics, 1973.
2. Hsieh, J. J., J. A. Rossi, and J. P. Donnelly. *Appl. Phys. Lett.* **28,** 709 (1976).
3. Arai, S., Y. Itaya, Y. Suematsu, K. Kishino, and S. Katayama. *Jpn. J. Appl. Phys.* **17,** 2067 (1978).
4. Itaya, Y., Y. Suematsu, S. Katayama, K. Kishino, and S. Arai. *Jpn. J. Appl. Phys.* **18,** 1795 (1979).
5. Arai, S., Y. Suematsu, Y. Itaya, *IEEE J. Quantum Electron.* **QE-16,** 197 (1980).
6. Oe, K., and K. Sugiyama. *Jpn. J. Appl. Phys.* **15,** 2003 (1976).
7. Nahory, R. E., and M. A. Pollack. *Electron. Lett.* **14,** 727 (1978).
8. Dolginov, L. M., A. E. Drakin, P. G. Eliseev, B. N. Sverdlov, and E. G. Shevchenko. *IEEE J. Quantum Electron.* **QE-21,** 646 (1985).
9. Kressel, H., and M. Ettenberg. *J. Appl. Phys.* **47,** 3533 (1976).
10. Itaya, Y., S. Katayama, and Y. Suematsu. *Electron. Lett.* **15,** 123 (1979).
11. Dyment, J. C. *Appl. Phys. Lett.* **10,** 84 (1967).
12. Dyment, J. C., L. A. D'Asaro, J. C. North, B. I. Miller, and J. E. Ripper. *Proc. IEEE* **60,** 726 (1972).
13. Henshall, G. D., G. H. B. Thompson, J. E. A. Whiteway, P. R. Selway, and M. Broomfield. *IEE Proc. I* **3,** 1 (1979).
14. Schwartz, B., M. W. Focht, N. K. Dutta, R. J. Nelson, and P. Besomi. *IEEE Trans. Electron Devices* **ED-31,** 841 (1984).
15. Oe, K., S. Ando, and K. Sugiyama. *J. Appl. Phys.* **51,** 43 (1980).
16. Dixon, R. W., F. R. Nash, R. L. Hartman, and R. T. Hepplewhite. *Appl. Phys. Lett.* **29,** 372 (1976).
17. Nelson, R. J., and N. K. Dutta. *Appl. Phys. Lett.* **37,** 769 (1980).
18. Paoli, T. L. *IEEE J. Quantum Electron.* **QE-13,** 351 (1977).
19. Van der Ziel, J. P., J. L. Merz, and T. L. Paoli. *J. Appl. Phys.* **50,** 4620 (1979).
20. Paoli, T. L. *Appl. Phys. Lett.* **34,** 652 (1979).
21. Dutta, N. K., and R. J. Nelson. *IEEE J. Quantum Electron.* **QE-17,** 804 (1981).
22. Dixon, R. W., and W. B. Joyce. *IEEE J. Quantum Electron.* **QE-15,** 470 (1979).
23. Dutta, N. K. *J. Appl. Phys.* **52,** 55 (1981).
24. Nash, F. R., R. L. Hartman, T. L. Paoli, and R. W. Dixon. *Appl. Phys. Lett.* **35,** 905 (1979).
25. Casey, H. C., Jr., and E. Buehler. *Appl. Phys. Lett.* **30,** 247 (1977).
26. Van der Ziel, J. P. *IEEE J. Quantum Electron.* **QE-17,** 60 (1981).
27. Lang, R. *Jpn. J. Appl. Phys.* **19,** L93 (1980).
28. Eliseev, P. G., A. Krasilnkov, M. A. Manko, and V. P. Strakhov, pp. 150–159 in *Physics of p-n Junctions and Semiconductor Devices*, ed. S. Y. Ryvkin and Yu. V. Shmartsev. New York: Plenum Press, 1971.
29. Dixon, R. W. *Bell Syst. Tech. J.* **55,** 973 (1976).
30. Paoli, T. L., and P. A. Barnes. *Appl. Phys. Lett.* **28,** 714 (1976).
31. Anthony, P. J., and N. E. Schumaker. *IEEE Electron Device Lett.* **EDL-1,** 58 (1980).

32. Wright, P. D., W. B. Joyce, and D. C. Craft. *J. Appl. Phys.* **53,** 1364 (1982).
33. Anthony, P. J., T. L. Paoli, and R. L. Hartman. *IEEE J. Quantum Electron.* **QE-16,** 735 (1980).
34. Lee, T. P., C. A. Burrus, B. I. Miller, and R. A. Logan. *IEEE J. Quantum Electron.* **QE-11,** 432 (1975).
35. Kawaguchi, H., and T. Kawakami. *IEEE J. Quantum Electron.* **QE-13,** 556 (1977).
36. Aiki, K., M. Nakamura, T. Kuroda, and J. Umeda. *Appl. Phys. Lett.* **48,** 649 (1977).
37. Amann, M.-C. *Electron. Lett.* **15,** 441 (1979).
38. Kaminow, I. P., R. E. Nahory, M. A. Pollack, L. W. Stulz, and J. C. DeWinter. *Electron. Lett.* **15,** 763 (1979).
39. Ueno, M., I. Sakuma, T. Furuse, Y. Matsumoto, H. Kawano, Y. Ide, and S. Matsumoto. *IEEE J. Quantum Electron.* **QE-17,** 1930 (1981).
40. Kaminow, I. P., R. E. Nahory, L. W. Stulz, and J. C. DeWinter. *Electron. Lett.* **17,** 318 (1981).
41. Turley, S. E. H., G. D. Henshall, P. D. Greene, V. P. Knight, D. M. Moule, and S. A. Wheeler. *Electron. Lett.* **17,** 868 (1981).
42. Kaminow, I. P., L. W. Stulz, J. S. Ko, A. G. Dentai, R. E. Nahory, J. C. DeWinter, and R. L. Hartman. *IEEE J. Quantum Electron.* **QE-19,** 1312 (1983).
43. Tsang, W. T., and R. A. Logan. *Appl. Phys. Lett.* **45,** 1025 (1984).
44. Temkin, H., R. A. Logan, M. B. Panish, and J. P. Van der Ziel. *Appl. Phys. Lett.* **45,** 330 (1984).
45. Tsukada, T. *J. Appl. Phys.* **45,** 4899 (1974).
46. Hirao, M., S. Tsuji, K. Mizuishi, A. Doi, and M. Nakamura. *J. Opt. Commun.* **1,** 10 (1980).
47. Mito, I., M. Kitamura, K. Kobayashi, S. Murata, M. Seki, Y. Odagiri, H. Nishimoto, M. Yamaguchi, and K. Kobayashi. *J. Lightwave Technol.* **LT-1,** 195 (1983).
48. Mito, I., M. Kitamura, K. Kaede, Y. Odagiri, M. Seki, M. Sugimoto, and K. Kobayashi. *Electron. Lett.* **18,** 2 (1982).
49. Nelson, R. J., P. D. Wright, P. A. Barnes, R. L. Brown, T. Cella, and R. G. Sobers. *Appl. Phys. Lett.* **36,** 358 (1980).
50. Ishikawa, H., H. Imai, T. Tanahashi, Y. Nishitani, M. Takusagawa, and K. Takahei. *Electron. Lett.* **17,** 465 (1981).
51. Dutta, N. K., D. P. Wilt, P. Besomi, W. C. Dautremont-Smith, P. D. Wright, and R. J. Nelson. *Appl. Phys. Lett.* **44,** 483 (1984).
52. Logan, R. A., J. P. Van der Ziel, and H. Tamkin. *Proc. SPIE Int. Soc. Opt. Eng.* **380,** 181 (1981).
53. Oron, M., N. Tamari, H. Shtrikman, and C. A. Burrus. *Appl. Phys. Lett.* **41,** 609 (1982).
54. Murotani, T., E. Oomura, H. Higuchi, H. Namizaki, and W. Susaki. *Electron. Lett.* **16,** 566 (1980).
55. Dutta, N. K., R. J. Nelson, P. D. Wright, and D. C. Craft. *J. Lightwave Technol.* **LT-2,** 160 (1984).
56. Nakano, Y., K. Takahei, Y. Noguchi, Y. Suzuki, and H. Nagai. *Electron. Lett.* **18,** 501 (1982).
57. Namizaki, H., R. Hirano, H. Higuchi, E. Oomura, Y. Sakakibara, and W. Susaki. *Electron. Lett.* **18,** 703 (1982).
58. Dutta, N. K., D. P. Wilt, and R. J. Nelson. *J. Lightwave Technol.* **LT-2,** 201 (1984).
59. Besomi, P., R. B. Wilson, R. L. Brown, N. K. Dutta, P. D. Wright, and R. J. Nelson. *Electron. Lett.* **20,** 417 (1984).
60. Adams, A. R., M. Asada, Y. Suematsu, and S. Arai. *Jpn. J. Appl. Phys.* **19,** L621 (1980).
61. Ishikawa, H., H. Imai, I. Umebu, K. Hori, and M. Takusagawa. *J. Appl. Phys.* **53,** 2851 (1982).

62. Yi, M. B., L. T. Lu, E. Kapon, Z. Rav-Noy, S. Margalit, and A. Yariv. *Appl. Phys. Lett.* **46,** 328 (1985).
63. Wilt, D. P., B. Schwartz, B. Tell, E. Beebee, and R. J. Nelson. *Appl. Phys. Lett.* **44,** 290 (1984).
64. Uchiyama, S., K. Moriki, K. Iga, and S. Furukawa. *Jpn. J. Appl. Phys.* Part 2 **21,** L639 (1982).
65. Liau, Z. L., and J. N. Walpole. *Appl. Phys. Lett.* **40,** 568 (1982); Liau, Z. L., J. N. Walpole, and D. Z. Tsang. *IEEE J. Quantum Electron.* **QE-20,** 855 (1984).
66. Chen, T. R., L. C. Chiu, K. L. Yu, U. Koren, A. Hasson, S. Margalit, and A. Yariv. *Appl. Phys. Lett.* **41,** 1115 (1982).
67. Koch, T. L., L. A. Coldren, T. J. Bridges, E. G. Burkhardt, P. J. Corvini, B. I. Miller, and D. P. Wilt. *Electron. Lett.* **20,** 856 (1984).
68. Burkhard, H., and E. Kuphal. *Jpn. J. Appl. Phys.* Part 2 **22,** L721 (1983); *IEEE J. Quantum Electron.* **QE-21,** 650 (1985).
69. Bowers, J. E., B. R. Hemenway, A. H. Gnauck, T. J. Bridges, E. G. Burkhardt, D. P. Wilt, and S. Maynard. *Appl. Phys. Lett.* **47,** 78 (1985).
70. Botez, D. *IEEE Spectrum* **22** (6), 43 (1985).
71. Botez, D., J. C. Connolly, M. Ettenberg, and D. B. Gilbert. *Electron. Lett.* **19,** 882 (1983).
72. Nakano, Y., K. Takahei, Y. Noguchi, H. Nagai, and K. Nawata. *Electron. Lett.* **17,** 782 (1981); Noguchi, Y., Y. Suzuki, T. Matsuoka, and H. Nagai. *Electron. Lett.* **20, 771** (1984).
73. Asano, T., and T. Okumura. *IEEE J. Quantum Electron.* **QE-21,** 619 (1985).
74. Yamaguchi, M., H. Nishimoto, M. Kitamura, S. Yamazaki, I. Mito, and K. Kobayashi. Conf. Lasers and Electro-optics, Baltimore, May 1985, p. 180.
75. Salzman, J., T. Venkatesan, R. Land, M. Mittlestein, and A. Yariv. *Appl. Phys. Lett.* **46,** 218 (1985).
76. Craig, R. R., L. W. Casperson, O. M. Stafsudd, J. J. J. Yang, G. A. Evans, and R. A. Davidheiser. *Electron. Lett.* **21,** 63 (1985).
77. Scifres, D. R., R. D. Burnham, and W. Streifer. *Appl. Phys. Lett.* **34,** 259 (1979); Scifres, D. R., W. Streifer, and R. D. Burnham. *IEEE J. Quantum Electron.* **QE-15,** 917 (1979).
78. Ackley, D. E. *Appl. Phys. Lett.* **42,** 152 (1983); *Electron. Lett.* **20,** 695 (1984).
79. Scifres, D. R., C. Lindstrom, R. D. Burnham, W. Streifer, and T. L. Paoli. *Electron. Lett.* **19,** 169 (1983).
80. Botez, D., and J. C. Connolly. *Appl. Phys. Lett.* **43,** 1096 (1983).
81. Kapon, E., C. P. Lindsay, J. S. Smith, S. Margalit, and A. Yariv. *Appl. Phys. Lett.* **45,** 1257 (1984).
82. Mukai, S., C. Lindsay, J. Katz, E. Kapon, Z. Rav-Noy, S. Margalit, and A. Yariv. *Appl. Phys. Lett.* **45,** 834 (1984).
83. Paoli, T. L., W. Streifer, and R. D. Burnham. *Appl. Phys. Lett.* **45,** 217 (1984).
84. Temkin, H., R. D. Dupuis, R. A. Logan, and J. P. Van der Ziel. *Appl. Phys. Lett.* **44,** 473 (1984).
85. Dutta, N. K., S. G. Napholtz, R. B. Wilson, R. L. Brown, and T. Cella. *Appl. Phys. Lett.* **45,** 941 (1984); *Appl. Phys. Lett.* **46,** 803 (1985).
86. Kapon, E., L. T. Lu, Z. Rav-Noy, M. Yi, S. Margalit, and A. Yariv. *Appl. Phys. Lett.* **46,** 136 (1985).
87. Figueroa, L., C. Morrison, H. D. Law, and F. Goodwin. *J. Appl. Phys.* **56,** 3357 (1984).
88. Butler, J. K., D. E. Ackley, and D. Botez. *Appl. Phys. Lett.* **44,** 293 (1984).
89. Kapon, E., J. Katz, and A. Yariv. *Opt. Lett.* **10,** 125 (1984).
90. Agrawal, G. P. *J. Appl. Phys.* **58,** 2922 (1985).
91. Mawst, L. J., D. Botez, T. J. Roth, G. Peterson, and J. J. Yang. *Electron. Lett.* **24,** 958 (1988).
92. Golubentsev, A. A., V. V. Likhanskii, and A. P. Napartovich. *Sov. Phys. JEPT* **66,** 676 (1987).

93. Roychoudhuri, C., E. Siebert, F. D'Amato, M. Macomber, and R. Noll. Proc. LEOS Annu. Mtg., Santa Clara, Nov. 1988.
94. Wilcox, J. Z., W. W. Simmons, D. Botez, M. Jansen, L. J. Mawst, G. Peterson, T. J. Wilcox, and J. J. Yang. *Appl. Phys. Lett.* **54,** 1848 (1989).
95. Mawst, L. J., D. Botez, T. J. Roth, and G. Peterson. *Appl. Phys. Lett.* **55,** 10 (1989).
96. Mawst, L. J., D. Botez, M. Jansen, T. J. Roth, and G. Peterson. *Appl. Phys. Lett.* **55,** 2060 (1989).
97. Welch, D. F., P. S. Cross, D. R. Scifres, and W. Streifer. *Electron. Lett.* **23,** 270 (1987).
98. Liau, Z. L. and J. N. Walpole. *Appl. Phys. Lett.* **46,** 115 (1985).
99. Carlson, N. W., G. A. Evans, D. P. Bour, and S. K. Liew. *Appl. Phys. Lett.* **56,** 16 (1990).
100. Evans, G. A., N. W. Carlson, J. M. Hammer, M. Lurie, J. K. Butler, S. L. Palfrey, R. Amantea, L. A. Carr, F. Z. Hawrylo, E. A. James, C. J. Kaiser, J. B. Kirk, and W. F. Reichert. *IEEE J. Quantum Electron.* **QE-25,** 1525 (1989).
101. Evans, G. A., D. P. Bour, N. W. Carlson, J. M. Hammer, M. Lurie, J. K. Butler, S. L. Palfrey, R. Amantea, L. A. Carr, F. Z. Hawrylo, E. A. James, J. B. Kirk, S. K. Liew, and W. F. Reichert. *Appl. Phys. Lett.* **55,** 2721 (1989).
102. Evans, G. A., D. P. Bour, N. W. Carlson, R. Amantea, J. M. Hammer, H. Lee, M. Lurie, R. C. Lai, P. F. Pelka, R. E. Farkas, J. B. Kirk, S. K. Liew, W. F. Reichert, C. A. Wang, H. K. Choi, J. N. Walpole, J. K. Butler, W. F. Ferguson, Jr., R. K. DeFreez, and M. Feliski. *IEEE J. Quantum Electron.* **QE-27,** 1594 (1991).

Chapter 6
RATE EQUATIONS AND OPERATING CHARACTERISTICS

6.1 INTRODUCTION

The previous chapter has discussed the design of various semiconductor laser structures. The usefulness of a specific structure depends on its performance characteristics and how well they match the requirements for a particular application. InGaAsP lasers operating in the wavelength range of 1.3–1.6 μm have been developed primarily to serve as a light source in fiber communication systems; Sec. 1.4 briefly discussed their performance requirements. Some of the desired features are a low threshold current, a narrow spectral width, and a fast response. The purpose of this chapter is to discuss the static, spectral, and dynamic characteristics of a semiconductor laser and their dependence on various device parameters. A unified approach is provided by the rate equations that govern the interplay between photons and charge carriers. The rate equations for lasers were first formulated[1] in 1960 and have been used extensively[2–15] for semiconductor lasers after their proper generalization.

This chapter is organized as follows. Section 6.2 obtains the rate equations starting from the wave equations. Section 6.3 considers their steady-state solutions both for single-mode and multimode cases. We obtain the light-current characteristics with the inclusion of spontaneous emission and show that the sharpness of the laser threshold depends on the fraction of spontaneous emission that couples into the lasing mode. Furthermore, the same fraction also governs the longitudinal-mode spectrum. Section 6.4 studies the transient response of a semiconductor laser. A small-signal analysis of the rate equations reveals an intrinsic resonance that manifests as relaxation oscillations during which the laser shifts its stored energy back and forth between photon and carrier populations. Section 6.5 considers the effect of noise on laser output. In particular, phase noise is responsible for the line width associated with a single longitudinal mode. Section 6.6 is

devoted to the modulation response of semiconductor lasers and to concepts such as modulation bandwidth and frequency chirp. Finally, in Sec. 6.7 we consider the effect of external reflection feedback on the performance of a semiconductor laser.

6.2 RATE EQUATIONS

Since the electromagnetic field inside the laser cavity satisfies Maxwell's equations, the starting point to obtain the field rate equation is the wave equation (2.2.12). In general, the dynamics of the semiconductor material should be taken into account to obtain the induced polarization $\mathscr{P}$. However, the material response governed by the intraband scattering processes is relatively fast ($\sim$0.1 ps) compared to other time scales of interest such as the photon lifetime and the carrier recombination time. Considerable simplification occurs if we assume that the material response is instantaneous and that Eq. (2.2.17) holds even for time-dependent fields. The wave equation then becomes

$$\nabla^2 \mathscr{E} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}(\epsilon \mathscr{E}) = 0 \tag{6.2.1}$$

where the loss term related to the medium conductivity has been included in the imaginary part of the dielectric constant ϵ [see Eq. (2.2.20)]. It should be kept in mind that the use of Eq. (6.2.1) is justified only for a fast-responding semiconductor. If for a semiconductor material the intraband scattering time becomes comparable to the photon lifetime ($\sim$1 ps), Eq. (2.2.12) should be used with the induced polarization given by Eq. (2.4.1).

The optical field $\mathscr{E}$ in general may consist of a large number of lateral, transverse, and longitudinal modes oscillating at different frequencies. Furthermore, each mode forms a standing-wave pattern in the axial direction arising from a superposition of the forward and backward running waves. For simplicity, we assume that the laser structure has been designed to support a single lateral and transverse mode. This is often the case in practical semiconductor lasers. The optical field $\mathscr{E}$ can be written as

$$\mathscr{E}(x, y, z, t) = \tfrac{1}{2}\hat{\mathbf{x}}\psi(x)\phi(y)\sum_j \sin(k_j z)E_j(t)e^{-i\omega_j t} + \text{c.c.} \tag{6.2.2}$$

where $\psi(x)$ and $\phi(y)$ are the lateral and transverse field profiles discussed in Chapter 2 and c.c. stands for complex conjugate. The sinusoidal variation of the optical field in the z direction assumes facets with high reflectivity.

Even though this assumption is somewhat questionable for a semiconductor laser (facet reflectivity $\cong 32\%$), its use is essential in order to avoid solving a complicated boundary-value problem and does not introduce significant quantitative errors if the laser is above threshold.[12] The wave number k_j is related to the cavity-resonance frequency $\Omega_j = 2\pi\nu_j$ given by Eq. (2.3.12), i.e.,

$$k_j = \frac{\mu\Omega_j}{c} = \frac{m_j\pi}{L} \tag{6.2.3}$$

where L is the cavity length and m_j is an integer. The laser-mode frequency ω_j is yet undetermined except that ω_j nearly coincides with Ω_j.

For simplicity of notation, let us first consider the case of a single longitudinal mode and drop the subscript j in Eq. (6.2.2), i.e.,

$$\mathscr{E}(x, y, z, t) = \tfrac{1}{2}\hat{\mathbf{x}}\psi(x)\phi(y)\sin(kz)E(t)e^{-i\omega t} + \text{c.c.} \tag{6.2.4}$$

We substitute $\mathscr{E}$ from Eq. (6.2.4) in Eq. (6.2.1), assume that $E(t)$ varies slowly, multiply by $\psi(x)$ and $\phi(y)$, and integrate over the whole range of x and y. After some simplifications, we obtain

$$\frac{2i\omega}{c^2}\left(\langle\epsilon\rangle + \frac{\omega}{2}\frac{\partial\langle\epsilon\rangle}{\partial\omega}\right)\frac{dE}{dt} + \left(\frac{\omega^2}{c^2}\langle\epsilon\rangle - k^2\right)E = 0 \tag{6.2.5}$$

where the spatially averaged dielectric constant

$$\langle\epsilon\rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\epsilon(x, y)\psi^2(x)\phi^2(y)\,dx\,dy \tag{6.2.6}$$

and $\psi(x)$ and $\phi(y)$ are assumed to be properly normalized. The second term on the left-hand side of Eq. (6.2.5) takes into account the dispersive nature of the semiconductor material. The average in Eq. (6.2.6) can be carried out using the analysis of Sec. 2.5. We find that $\langle\epsilon\rangle$ approximately equals the effective dielectric constant of the waveguide mode and is given by

$$\langle\epsilon\rangle \cong \bar{\mu}^2 + 2\Gamma\bar{\mu}\Delta\mu_p + i\bar{\mu}\bar{\alpha}/k_0 \tag{6.2.7}$$

where $k_0 = \omega/c$, Γ is the confinement factor, and $\Delta\mu_p$ is the carrier-induced index change. Furthermore, $\bar{\mu}$ is the mode index and $\bar{\alpha}$ is the mode-absorption

coefficient given by [see Eqs. (2.5.47) and (2.5.48)]

$$\bar{\alpha} = -\Gamma g + \alpha_{\text{int}} + \alpha_m. \tag{6.2.8}$$

Equation (6.2.8) has been generalized to include the facet loss α_m [given by Eq. (2.3.10)]. A rigorous approach would require that E should be allowed to vary with z and then to solve the axial-propagation problem with appropriate boundary conditions at the laser facets.[12] However, considerable simplification occurs if we replace the localized facet loss by an equivalent distributed loss and add it phenomenologically to α_{int}. We substitute (6.2.7) into (6.2.5) and use $k = \bar{\mu}\Omega/c$ from Eq. (6.2.3). Using $(\omega^2 - \Omega^2) \cong 2\omega(\omega - \Omega)$ and $\langle \epsilon \rangle \cong \bar{\mu}^2$ on the left-hand side of Eq. (6.2.5), we obtain

$$\frac{dE}{dt} = \frac{i\bar{\mu}}{\mu_g}(\omega - \Omega)E + \frac{i\omega}{\mu_g}(\Gamma \Delta\mu_p + i\bar{\alpha}/2k_0)E \tag{6.2.9}$$

where μ_g is the group index corresponding to the mode index $\bar{\mu}$. It is useful to separate Eq. (6.2.9) into its real and imaginary parts using

$$E = A \exp(-i\phi). \tag{6.2.10}$$

We then obtain the amplitude and phase rate equations

$$\dot{A} = \tfrac{1}{2}v_g[\Gamma g - (\alpha_{\text{int}} + \alpha_m)]A \tag{6.2.11}$$

$$\dot{\phi} = -\frac{\bar{\mu}}{\mu_g}(\omega - \Omega) - \frac{\omega}{\mu_g}\Gamma \Delta\mu_p \tag{6.2.12}$$

where a dot represents time derivative, $v_g = c/\mu_g$ and is the group velocity, and we have used Eq. (6.2.8) to eliminate $\bar{\alpha}$. Equation (6.2.11) suggests that the rate of amplitude growth equals gain minus loss and may have been written directly using heuristic arguments. The advantage of the present approach where it is obtained using Maxwell's equations is that the various approximations made during its derivation have been clearly identified. Furthermore, the phase equation (6.2.12) follows self-consistently and shows that the carrier-induced index change $\Delta\mu_p$ affects the lasing-mode frequency ω. It will be seen later that this term plays an important role in the discussion of line width and frequency chirping.

It is more convenient to write the amplitude equation (6.2.11) in terms of the photon number P defined using

$$P = \frac{\epsilon_0 \bar{\mu}\mu_g}{2\hbar\omega} \int |\mathscr{E}|^2 \, dV \tag{6.2.13}$$

where $\hbar\omega$ is the photon energy. Since $P \propto A^2$, we obtain

$$\dot{P} = (G - \gamma)P + R_{sp} \tag{6.2.14}$$

where

$$G = \Gamma v_g g \tag{6.2.15}$$

is the net rate of stimulated emission and

$$\gamma = v_g(\alpha_m + \alpha_{int}) = \tau_p^{-1} \tag{6.2.16}$$

is the photon decay rate that can be used to define the photon lifetime τ_p inside the laser cavity.

The last term R_{sp} in Eq. (6.2.14) has been added to take into account the rate at which spontaneously emitted photons are added to the intracavity photon population. The contribution of spontaneous emission to the lasting mode is indispensable[9,10] and should be included for a correct treatment of laser dynamics. A rigorous approach would require quantization of the laser field. Within the semiclassical framework, the contribution of spontaneous emission can be included by adding a noise current to Maxwell's equations. It should be stressed that the photon number P is used only as a convenient dimensionless measure of the optical intensity $|E|^2$. Its use is not intended to imply the particulate nature of photons. This distinction is important for the discussion of the intensity noise.

The phase Eq. (6.2.12) can be related to the gain if we use the parameter β_c defined in Eq. (2.4.6) to replace $\Delta\mu_p$ by

$$\Delta\mu_p = -(\beta_c/2k_0)\Delta g. \tag{6.2.17}$$

Using Eqs. (6.2.15) and (6.2.17) and replacing $\Delta G = \Gamma v_g \Delta g$ by $G - \gamma$, Eq. (6.2.12) becomes

$$\dot{\phi} = -(\omega - \omega_{th}) + \tfrac{1}{2}\beta_c(G - \gamma). \tag{6.2.18}$$

This equation shows that when the gain changes from its threshold value, the phase changes as well. Physically, this is so because a gain change is always accompanied by an index change that shifts the longitudinal-mode frequencies. In Eq. (6.2.18) the longitudinal-mode frequency ω_{th} corresponds to the threshold value of the mode index. By expanding $\Omega(\omega)$ in the vicinity of ω_{th}, one can show that $(\bar{\mu}/\mu_g)[\omega - \Omega(\omega)] = \omega - \omega_{th}$.

The gain G in Eq. (6.2.14) is known in terms of the carrier density n inside the active layer. The rate equation for n was considered in Sec. 2.4

and shows that in general the effect of carrier diffusion should be included. Its inclusion, however, complicates the analysis considerably.[9,10] In this chapter we shall neglect carrier diffusion. The analysis is therefore mainly applicable to an index-guided laser. We define the number of carriers inside the active layer as

$$N = \int n \, dV = nV \tag{6.2.19}$$

where we have assumed that n is approximately constant. $V = Lwd$ and is the active volume for a laser of length L, width w, and thickness d. Using Eqs. (2.4.7) and (2.4.9), the carrier rate equation is

$$\dot{N} = I/q - \gamma_e N - GP \tag{6.2.20}$$

where $I = wLJ$, J is the current density flowing through the active region (the contribution of leakage current is excluded), and

$$\gamma_e = (A_{nr} + Bn + Cn^2) = \tau_e^{-1} \tag{6.2.21}$$

is the carrier-recombination rate that can be used to define the spontaneous carrier lifetime τ_e. Both radiative and nonradiative recombination processes discussed in Chapter 3 contribute to γ_e. The recombination rate A_{nr} is due to mechanisms such as trap or surface recombination. B is the radiative-recombination coefficient, while C is related to the Auger recombination processes. The last term in Eq. (6.2.20) is due to stimulated recombination and leads to a nonlinear coupling between photons and charge carriers.

To complete the rate-equation description, we need an expression for R_{sp} appearing in Eq. (6.2.14). If we assume that a fraction β_{sp} of spontaneously emitted photons goes into the lasing mode, R_{sp} is given by

$$R_{sp} = \beta_{sp} \eta_{sp} \gamma_e N \tag{6.2.22}$$

where $\eta_{sp} = Bn/\gamma_e$ and is the (internal) *spontaneous quantum efficiency* showing the fraction of carriers that emit photons through spontaneous recombination. The parameter β_{sp} is referred to as the *spontaneous-emission factor* and has been discussed[16–26] extensively in the literature. Petermann[17] has pointed out that β_{sp} depends on the lateral guiding mechanism and is enhanced for gain-guided lasers due to wavefront curvature. For a mode lasing at the wavelength $\lambda = 2\pi c/\omega$, his expression for β_{sp} is

$$\beta_{sp} = \frac{K\Gamma\lambda^4}{4\pi^2 \mu \bar{\mu} \mu_g V \Delta\lambda_{sp}} \tag{6.2.23}$$

where μ is the bulk-material index and $\Delta\lambda_{sp}$ is the width of the spontaneous-emission spectrum. The enhancement factor $K = 1$ for index-guided lasers but is larger for gain-guided lasers. However, there is some disagreement[17–25] over its numerical value. In practice, β_{sp} is treated as a fitting parameter. Its typical numerical value for index-guided InGaAsP lasers is in the range of 10^{-4}–10^{-5}.

An alternative expression for R_{sp} can be obtained by using the Einstein relation that relates the gain spectrum to the spontaneous-emission spectrum.[11,12] In this approach,

$$R_{sp}(\omega) = n_{sp}(\omega)G(\omega) \tag{6.2.24}$$

where

$$n_{sp}(\omega) = \left[1 - \exp\left(\frac{\hbar\omega - E_f}{k_B T}\right)\right]^{-1}$$

is the *population-inversion factor* and E_f is the energy separation between the quasi-Fermi levels. For 1.3-μm InGaAsP lasers, the estimated value[14] of n_{sp} from spectral measurements at room temperature is ~ 1.7.

Equations (6.2.14), (6.2.18), and (6.2.20) are the single-mode rate equations and will be used in this chapter to discuss the operating characteristics of a semiconductor laser. For a discussion of multimode phenomena, these equations should be generalized to include the number of possible longitudinal modes for which the gain G is positive. This number depends on the width of the gain spectrum and the frequency spacing between different longitudinal modes. The photon rate equation for each mode can be obtained using Eqs. (6.2.1) and (6.2.2) and following the preceding analysis. If P_m represents the photon population for the mth longitudinal mode oscillating at the frequency ω_m, the multimode rate equations are

$$\dot{P}_m = (G_m - \gamma_m)P_m + R_{sp}(\omega_m) \tag{6.2.25}$$

$$\dot{N} = I/q - \gamma_e N - \sum_m G_m P_m \tag{6.2.26}$$

where $G_m = G(\omega_m)$ is the mode gain and γ_m is the mode loss. In conventional Fabry-Perot semiconductor lasers, γ_m is frequency independent and $\gamma_m = \gamma = \tau_p^{-1}$ for all modes. However, frequency-dependent losses can be introduced to enhance mode selectivity (see Chapters 7 and 8). In Eq. (6.2.26) the summation is over all modes participating in the process of stimulated emission.

To simplify the notation, the rate equations have been obtained for the dimensionless variables P_m and N, which represent the photon and electron

Table 6.1 Typical parameter values for a 1.3 μm buried-heterostructure laser

PARAMETER	SYMBOL	VALUE
Cavity length	L	250 μm
Active-region width	w	2 μm
Active-layer thickness	d	0.2 μm
Confinement factor	Γ	0.3
Effective mode index	$\bar{\mu}$	3.4
Group refractive index	μ_g	4
Line-width enhancement factor	β_c	5
Facet loss	α_m	45 cm^{-1}
Internal loss	α_{int}	40 cm^{-1}
Gain constant	a	2.5×10^{-16} cm^2
Carrier density at transparency	n_0	1×10^{18} cm^{-3}
Nonradiative recombination rate	A_{nr}	1×10^{8} s^{-1}
Radiative recombination coefficient	B	1×10^{-10} cm^3/s
Auger recombination coefficient	C	3×10^{-29} cm^6/s
Threshold carrier population	N_{th}	2.14×10^{8}
Threshold current	I_{th}	15.8 mA
Carrier lifetime at threshold	τ_e	2.2 ns
Photon lifetime	τ_p	1.6 ps

populations inside the laser cavity. The quantity of practical interest is the output power emitted from each facet. As seen in Eq. (2.6.7), it is related linearly to the photon population and is given by

$$P_m^{out} = \tfrac{1}{2}\hbar\omega v_g \alpha_m P_m. \tag{6.2.27}$$

The derivation of Eq. (6.2.27) is intuitively obvious if we note that $v_g\alpha_m$ is the rate at which photons of energy $\hbar\omega$ escape through the two facets. For a 250-μm-long 1.3-μm InGaAsP laser, $P_m \cong 3.88 \times 10^4$ at a power level of 1 mW. For ease of reference, Table 6.1 gives typical parameter values for a 1.3-μm buried heterostructure laser. In the case of unequal facet reflectivities the output powers are obtained by using Eqs. (2.6.20).

6.3 STEADY-STATE CHARACTERISTICS

The rate equations of the previous section can be used to obtain the steady-state response of a semiconductor laser at a fixed value of the current I.

The steady-state solution is obtained by setting all time derivatives to zero and is applicable for continuous-wave (CW) operation after transients have died out. The steady-state or CW solution is also applicable under pulsed operation provided the duration of the current pulse is much longer than the laser response time, which is a few nanoseconds. Two steady-state features are of general interest in characterizing laser performance, namely the light-current (L-I) curve and the longitudinal-mode spectrum. The L-I curve shows how the output power varies with the device current. The longitudinal-mode spectrum indicates how the power is distributed among various modes at a given current. In the rest of Sec. 6.3, we consider the two separately.

6.3.1 Light-Current Curve

Although in general one should consider the multimode rate equations, a simple, physical description of the laser threshold and the resulting L-I curve can be based on the single-mode rate equations. From Eq. (6.2.14) the photon number P is given by

$$P = R_{sp}/(\gamma - G). \tag{6.3.1}$$

This equation stresses that a laser can be viewed as a regenerative noise amplifier.[27] Spontaneously emitted photons provide the noise input that is amplified in the presence of gain provided by the injected charge carriers. The threshold condition ($G = \gamma$) obtained in Chapter 2 ignored the contribution of spontaneous emission. Equation (6.3.1) shows that the required gain G is slightly below the cavity-loss level and approaches its asymptotic value γ as the power output increases.

To obtain the L-I curve,[8,9] we substitute Eq. (6.3.1) into Eq. (6.2.20). This leads to the relation

$$\gamma_e(N)N + R_{sp}(N)\frac{G}{\gamma - G} = \frac{I}{q} \tag{6.3.2}$$

that can be used to obtain N for a given I if the functional dependence of $G(N)$ is known. Note that G corresponds to the peak value of the gain since the longitudinal mode closest to the gain peak reaches threshold first. The simple model discussed in Sec. 2.4 used a linear dependence of G on N. Using Eqs. (2.4.3) and (6.2.15), we obtain

$$G(N) = \Gamma v_g a(N/V - n_0). \tag{6.3.3}$$

The use of Eqs. (6.2.21), (6.2.22), and (6.3.3) in Eq. (6.3.2) leads to a fourth-degree polynomial in N that can be used to obtain N as a function of the

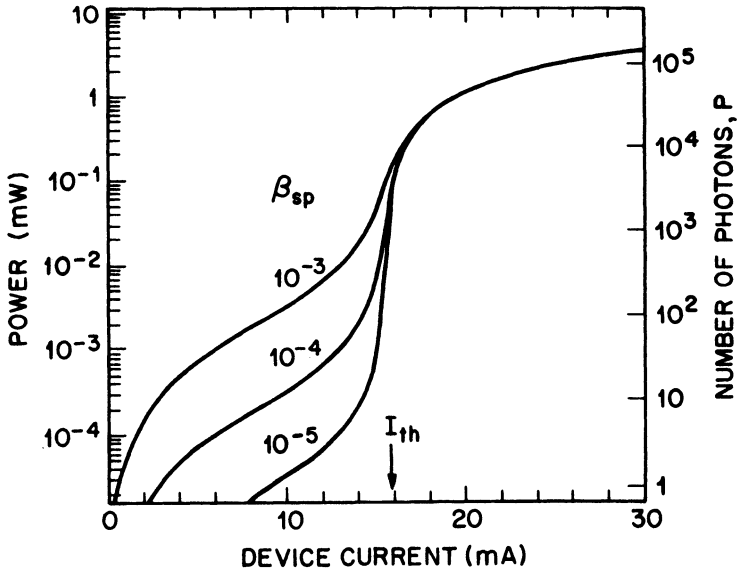


Fig. 6.1 Calculated L-I curves for a 1.3-μm InGaAsP laser showing the dramatic increase in laser power near threshold. The parameter β_{sp} governs the fraction of spontaneous emission that contributes to the lasing model. Other parameters used are given in Table 6.1.

device current I. The photon number P is then obtained using Eq. (6.3.1). The output power is related to P linearly as given by Eq. (6.2.27).

Figure 6.1 shows the variation of the output power with the device current I for a 1.3-μm InGaAsP laser using the parameter values shown in Table 6.1. Different curves correspond to different values of the spontaneous-emission factor β_{sp}. We note that P starts to increase rapidly as the laser threshold is reached.[9,10] The sudden change in the slope of the L-I curve in the vicinity of the threshold indicates that stimulated emission has taken over the spontaneous emission and is accompanied by other characteristic features such as narrowing of the spectral width. In the threshold vicinity, P changes by orders of magnitude in a narrow current range. However, the sharpness of the laser threshold depends on β_{sp}. The variation of the carrier population N with I in Fig. 6.2 shows that N is nearly constant in the threshold region and does not increase significantly with a further increase in I. As is evident from Eq. (6.3.1), a small change in N, or equivalently in the gain G, can nonetheless produce a large change in P when G is close to γ. The limiting value of N is referred to as the threshold carrier population N_{th} and corresponds to the condition where $G = \gamma$.

From a practical viewpoint the quantity of interest is the threshold current I_{th}. In the presence of spontaneous emission, however, the threshold is not

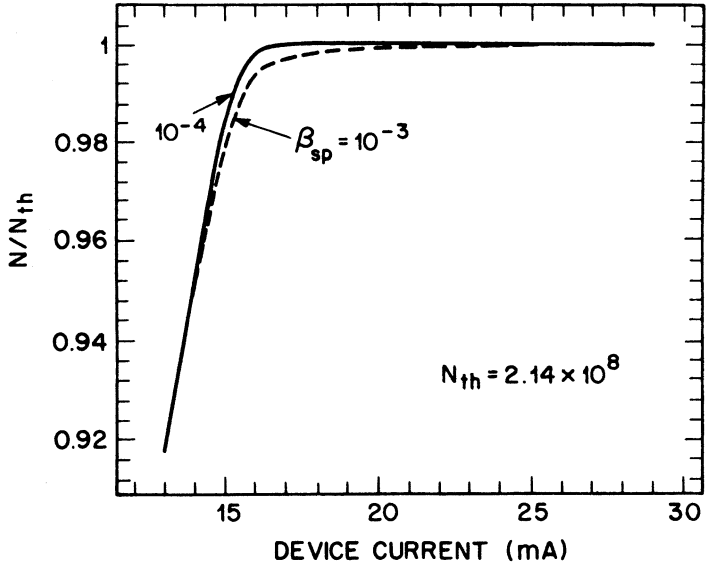


Fig. 6.2 Variation of the carrier population N with device current I corresponding to the L-I curves shown in Fig. 6.1. The curves for $\beta_{sp} = 10^{-5}$ and $\beta_{sp} = 10^{-4}$ are indistinguishable on this scale, so the curve for $\beta_{sp} = 10^{-5}$ is not given.

sharply defined. Figure 6.1 shows that the transition from the nonlasing to the lasing state becomes softer with an increase in the amount of spontaneous emission. It is customary to define I_{th} in the limiting case of $\beta_{sp} = 0$. This case was discussed in Sec. 2.6. From Eq. (6.3.2) with $R_{sp} = 0$, we obtain

$$I_{th} = q\gamma_e(N_{th})N_{th} \tag{6.3.4}$$

where N_{th} is obtained using Eq. (6.3.3) after setting $G(N_{th})$ equal to γ. For parameter values given in Table 6.1, $I_{th} \cong 16$ mA. It should be stressed that Eq. (6.3.4) expresses the current that passes through the active region. The actual device current is generally higher because of current leakage outside the active region, which invariably occurs under realistic conditions (see Sec. 5.6).

The L-I curves shown in Fig. 6.1 model reasonably well an index-guided device. In the presence of gain guiding, carrier diffusion becomes important and the L-I curve under CW operation is generally obtained numerically. The rate-equation model neglects the axial variation of mode intensities. The axial effects can be included using an FP-type approach.[28] Note also that the L-I curve is strongly temperature-dependent. In obtaining Fig. 6.1, the

parameter values corresponding to room temperature ($T = 300$ K) were used. The temperature dependence of L-I curves was discussed in Sec. 3.6.

6.3.2 Longitudinal-Mode Spectrum

The power spectrum of a semiconductor laser obtained using a spectrometer shows the presence of several longitudinal modes whose relative powers vary with the current I. The multimode nature of semiconductor lasers has been of concern from the early days.[2,5] It is believed that spontaneous emission plays an important role in determining the spectral characteristics of a semiconductor laser.[29–34] We use the multimode rate equations to obtain such spectral features as the number of longitudinal modes and their relative intensities. In its simple form, the analysis is applicable to a strongly index guided device.

To solve the multimode rate equations (6.2.25) and (6.2.26), we need to know the gain spectrum G_m where $G_m = G(\omega_m)$. The gain spectrum is either obtained experimentally[14] or calculated numerically (see Fig. 3.7); its functional form is not generally available. A simple approximation is that the gain decreases quadratically from its peak value and assumes that

$$G(\omega) = G_0\left[1 - \left(\frac{\omega - \omega_0}{\Delta\omega_g}\right)^2\right] \tag{6.3.5}$$

where ω_0 is the frequency at which the gain takes its maximum value G_0, and $\Delta\omega_g$ is the frequency spread over which the gain is nonzero on either side of the gain peak (see Fig. 6.3). We assume that the central (most intense) mode is located at the gain-peak frequency ω_0. The other longitudinal modes have frequencies

$$\omega_m = \omega_0 + m\Delta\omega_L \tag{6.3.6}$$

where $\Delta\omega_L$ is the longitudinal-mode spacing [see Eq. (2.3.14)] and the integer m varies from $-M$ to $+M$. The value of M is determined by the largest integer contained in the ratio

$$M \cong \frac{\Delta\omega_g}{\Delta\omega_L} \tag{6.3.7}$$

where $2M + 1$ represents the number of longitudinal modes for which the gain is positive. Using Eqs. (6.3.5)–(6.3.7), the mode gain is approximated by[13]

$$G_m = G_0[1 - (m/M)^2]. \tag{6.3.8}$$

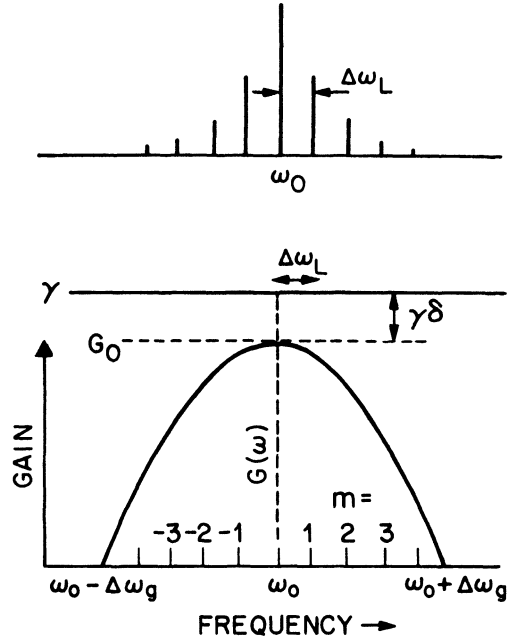


Fig. 6.3 Schematic illustration of the longitudinal-mode spectrum and its dependence on the gain spectrum. The separation between the peak gain G_0 and the cavity loss level γ is exaggerated for clarity.

The photon number P_m can be obtained using Eq. (6.2.25) and is given by

$$P_m = \frac{R_{sp}(\omega_m)}{\gamma - G_m} \tag{6.3.9}$$

where we have assumed that all modes have the same loss $\gamma = \tau_p^{-1}$. One can now follow a procedure similar to that for the single-mode case. The substitution of Eqs. (6.3.8) and (6.3.9) into Eq. (6.2.26) leads to an implicit relation that can be solved to obtain the steady-state carrier number N. This in turn determines the mode gain G_m and the steady-state photon population P_m for each value of the current I.

Figure 6.4 shows the results of such a numerical calculation using a total of 19 longitudinal modes ($M = 9$) for a 1.3-μm InGaAsP laser.[34] The output power in the central or main mode ($m = 0$) and in side modes (the four adjacent modes on both sides of the central mode) is shown as a function of the device current. Below or near threshold, the power increases in all of the longitudinal modes. However, the side-mode power saturates in the

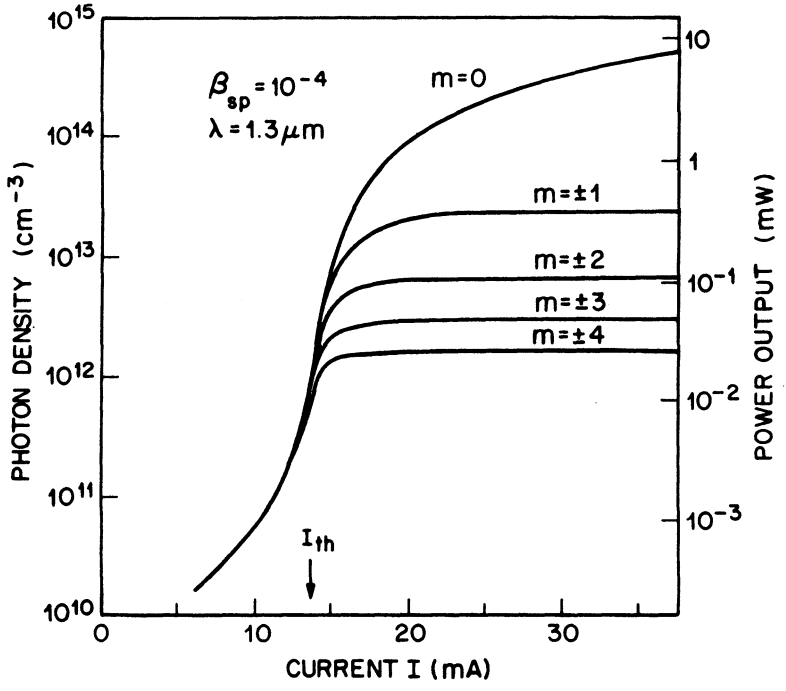


Fig. 6.4 Variation of the photon density and output power with current for several longitudinal modes of a multimode semiconductor laser. Note the power saturation for the side modes. (After Ref. 34 © 1982 IEEE)

above-threshold regime, whereas the main-mode power keeps increasing in a manner similar to the single-mode case discussed in Sec. 6.3.1. The power level at which a given side mode saturates depends on the spontaneous-emission factor β_{sp} and increases with an increase in β_{sp}. The power degeneracy for positive and negative values of m implies that the longitudinal-mode spectrum is symmetric with respect to the main mode, which is a consequence of our assumption that the gain spectrum (6.3.5) is symmetric. In practice, both the gain spectrum and the longitudinal-mode spectrum show some asymmetry.

To get some physical insight into the multimode behavior of semiconductor lasers, we note from our previous discussion for the single-mode case that the peak gain G_0 asymptotically approaches γ and that $\gamma - G_0$ is inversely proportional to the main-mode power. If we substitute $m = 0$ and use

$$G_0 = \gamma(1 - \delta) \tag{6.3.10}$$

in Eq. (6.3.9), we find that

$$\delta = \frac{R_{sp}(\omega_0)}{\gamma P_0}. \tag{6.3.11}$$

Using Eqs. (6.3.8) and (6.3.10) in (6.3.9), the photon number in the mth side mode is given by

$$P_m \cong \frac{R_{sp}(\omega_m)}{\gamma}\left(\frac{1}{\delta + (m/M)^2}\right). \tag{6.3.12}$$

This equation shows that the power distribution among various side modes is approximately Lorentzian and that the side-mode power is half the main-mode power when $m = \delta^{1/2}M$. The width (FWHM) of the spectral envelope is thus given by

$$\Delta\omega_s = 2m(\Delta\omega_L) = \delta^{1/2}(2\Delta\omega_g) \tag{6.3.13}$$

where we have used Eq. (6.3.7). Since δ decreases with an increase in the main-mode power, the spectral width $\Delta\omega_s$ decreases continuously in the above-threshold regime.

A measure of the spectral purity of a semiconductor laser is the mode suppression ratio (MSR), defined as the ratio of the main-mode power to the power of the most intense side mode,

$$\text{MSR} = \frac{P_0}{P_1} = 1 + \frac{1}{\delta M^2} \tag{6.3.14}$$

where we used Eq. (6.3.12) and assumed that $R_{sp}(\omega_0) \cong R_{sp}(\omega_1)$. Clearly the MSR increases with an increase in the main-mode power P_0, which reduces δ. For a given δ, the MSR can be improved by decreasing M. This suggests that short-cavity semiconductor lasers[34] have an improved MSR, since M can be made smaller by increasing the longitudinal-mode spacing. Using Eqs. (6.3.7) and (6.3.11) in Eq. (6.3.14), the explicit dependence of the MSR on the laser parameters is given by

$$\text{MSR} = 1 + \frac{P_0}{\tau_p R_{sp}}\left(\frac{\Delta\omega_L}{\Delta\omega_g}\right)^2 \tag{6.3.15}$$

where $\tau_p = \gamma^{-1}$ is the photon lifetime. This equation shows that the MSR increases linearly with the main-mode power. The term *single-mode operation* implies a large value of MSR. However, the exact value of the MSR above which the laser qualifies as being single mode is a matter of definition. A value of 20 for the MSR is often used for this purpose.[34]

The total output power emitted by a multimode laser can be obtained by summing over all the modes. If we use Eq. (6.2.27) to relate the power emitted per facet to the intracavity photon number, we obtain

$$P^{\text{out}} \cong \tfrac{1}{2}\hbar\omega_0 v_g \alpha_m \sum_m P_m. \tag{6.3.16}$$

We substitute P_m from Eq. (6.3.12). The summation can be performed if we replace the finite sum by an infinite sum. This does not introduce significant errors since P_m is almost zero for large m. The total intracavity photon population is given by[29]

$$P_T = \sum_m P_m = \frac{R_{\text{sp}}(\omega_0)}{\gamma\delta^{1/2}} \pi M \coth(\pi M \delta^{1/2}) \tag{6.3.17}$$

where we have assumed that $R_{\text{sp}}(\omega_m)$ can be replaced by $R_{\text{sp}}(\omega_0)$. This assumption is justified since the spontaneous-emission spectrum is much wider than the longitudinal-mode spectrum.

The previous analysis shows that the multimode characteristics of a semiconductor laser can be described in terms of two dimensionless parameters M and δ. The number $2M + 1$ corresponds to the total number of longitudinal modes that fit within the gain spectrum and experience gain. The parameter δ, defined by Eq. (6.3.10), is a measure of how closely the peak gain approaches the total cavity loss (see Fig. 6.3) and decreases with an increase in the laser power.

The numerical value of δ depends on the spontaneous-emission rate R_{sp} and is sensitive to the numerical value of the spontaneous-emission factor β_{sp} [see Eqs. (6.2.22) and (6.2.23)]. For index-guided lasers, $\beta_{\text{sp}} \lesssim 10^{-4}$, and at a power level of a few milliwatts, $\delta \cong 10^{-4}$. Using a typical value ($M = 10$) for a 250-μm-long laser, Eq. (6.3.14) predicts that an MSR of about 100 can be achieved under CW operation. Figure 6.5 shows the calculated longitudinal-mode spectra at various power levels for a 1.3-μm InGaAsP laser.[34] At a power level of 5 mW, most of the power is carried out by the main mode.

Gain-guided lasers, by contrast, exhibit longitudinal mode spectra that are strikingly different from those shown in Fig. 6.5. By way of an example, Fig. 6.6 shows the experimentally measured mode spectra of an AlGaAs gain-guided laser.[35] Many modes are present and MSR $\cong 2$ even at a power level of 11.7 mW. From Eq. (6.3.14) we note that this would be the case if $\delta \gtrsim 10^{-2}$. Such values of δ are consistent with $\beta_{\text{sp}} \cong 10^{-3}$. It has been suggested[35] that an order-of-magnitude increase in the spontaneous-emission factor is due to the enhancement factor[17,18] K in Eq. (6.2.23). The effect of

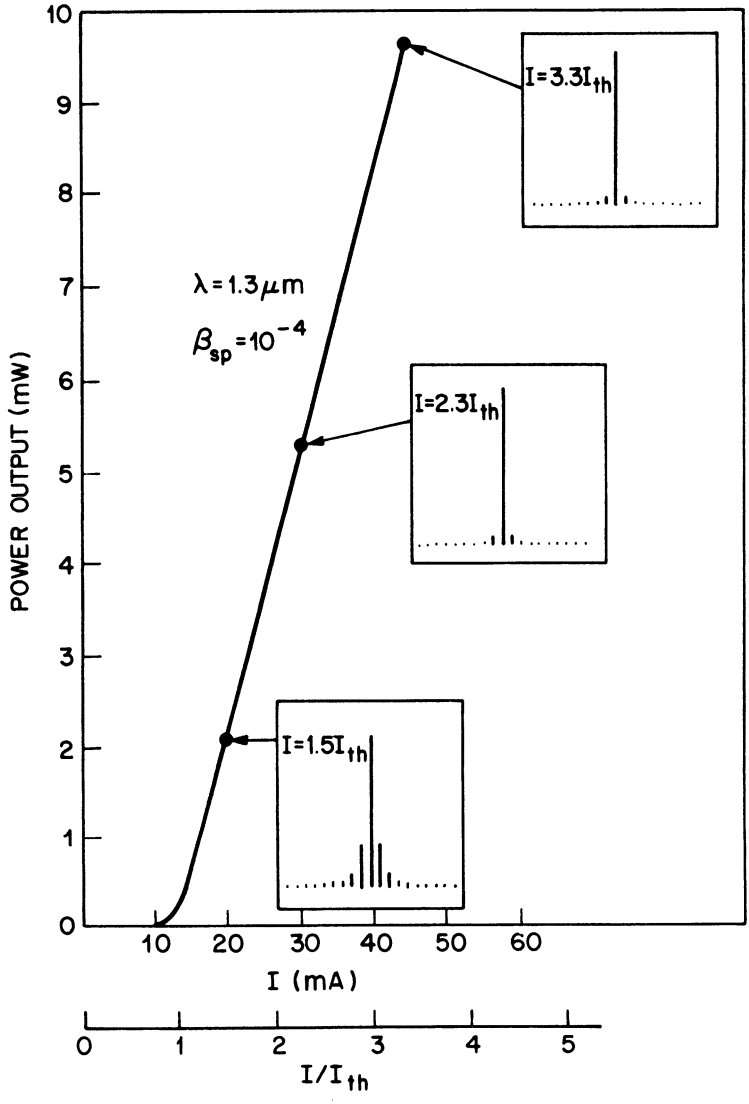


Fig. 6.5 Calculated L-I curve for the total output power of an index-guided laser. Insets show the longitudinal-mode spectra at three drive current levels. (After Ref. 34 © 1982 IEEE)

the lateral waveguiding mechanism on the longitudinal-mode spectrum is not very well understood.[36]

In the previous analysis the multimode nature of semiconductor lasers is attributed to spontaneous emission. The semiconductor laser behaves as a

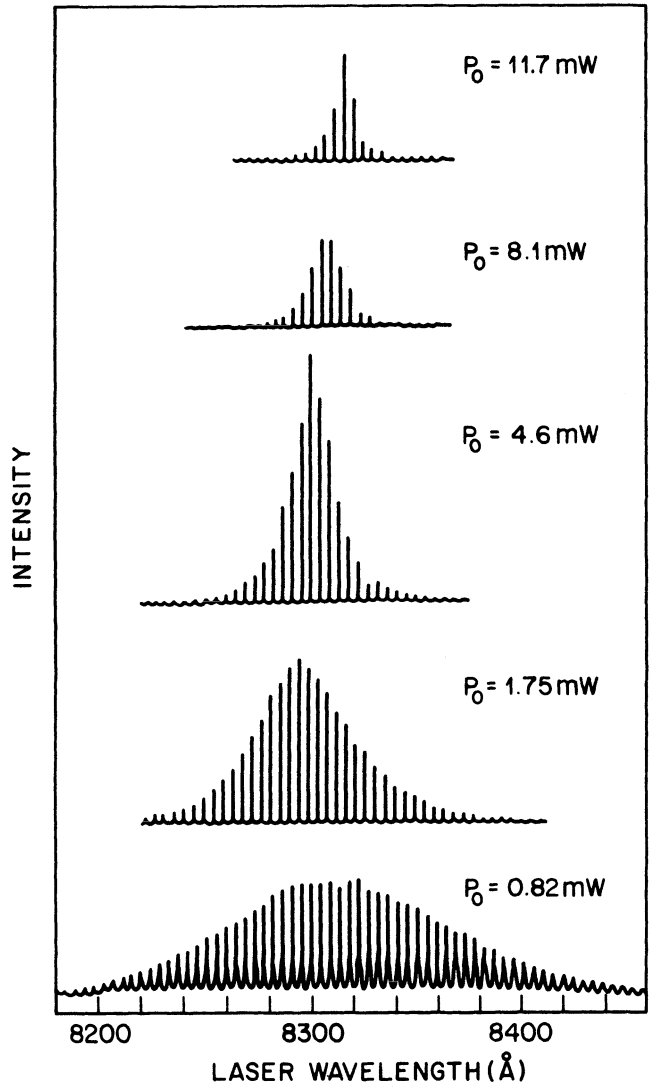


Fig. 6.6 Longitudinal-mode spectra of a gain-guided laser observed at several power levels. Experimental spectra for an index-guided laser are shown in Fig. 2.12. (After Ref. 35)

regenerative noise amplifier.[27] It amplifies all modes for which the roundtrip gain is positive. Once the threshold is reached, the gain is approximately clamped and the power in the side modes saturates.[34] In this model, the MSR increases continuously with an increase in the laser power, as is evident

from Eq. (6.3.15). In practice, however, with an increase in the laser power several other phenomena, such as spatial and spectral hole burning, start to influence the longitudinal-mode behavior.

Spatial-hole burning is a result of the standing-wave nature of the optical mode and is known[2] to lead to multimode oscillation. However, its relative importance depends on carrier diffusion, which is likely to wash out axial inhomogeneities in the carrier density.[37–39] The effectiveness of the diffusion process decreases with an increase in the laser power since the stimulated recombination time becomes much shorter relative to the carrier-diffusion time. As carriers are depleted faster at the crest of the standing wave, the main-mode gain decreases. As a result the side-mode power starts to increase when the main-mode power exceeds a critical value.[34]

Spectral-hole burning is related to the broadening mechanism of the gain profile. In semiconductor lasers the gain profile is nearly homogeneously broadened[40] because the intraband carrier scattering time is extremely short (~ 100 fs). At high laser powers the stimulated recombination time becomes small enough (~ 1 ps) that the peak gain decreases because of partial spectral-hole burning. Consequently other longitudinal modes start to grow, and the range of single-mode operation is inherently limited. The power-dependent or the nonlinear part of the gain has been used to explain the experimental results wherein the power distribution among the longitudinal modes is asymmetric and the dominant mode shifts toward longer wavelengths with an increase in device current.[41] Power-dependent changes in the mode gain are referred to as the *nonlinear gain* and have been studied extensively[42–47] by using a density-matrix approach commonly used in laser theory. We shall see later in this chapter that the nonlinear gain especially affects the dynamic response of semiconductor lasers.

A comprehensive analysis of the multimode characteristics of a semiconductor laser is extremely involved. Together with such mechanisms as spatial- and spectral-hole burning, whose main effect is to destabilize the main mode, there are other nonlinear processes that tend to stabilize the main mode. One such nonlinear process results from the beating of lasing and nonlasing modes, which modulates the carrier population at the beat frequency corresponding to the longitudinal-mode spacing (typically ~ 100 GHz).[47] Since the mode gain and the refractive index both depend on the carrier population, they are also modulated at the beat frequency, an effect referred to as the creation of dynamic gain and index gratings. Bragg diffraction from these gratings couples the longitudinal modes and affects their steady-state power distribution through a well-known nonlinear phenomenon called *four-wave mixing*.[43–47] This mechanism has been used to explain extraordinarily large MSR values observed in some index-guided

AlGaAs lasers.[43] It can also explain shifting of the main mode toward longer wavelengths with an increase in the current.

The multimode rate equations (6.2.25) and (6.2.26) are able to model reasonably well the observed features in the longitudinal-mode spectra of semiconductor lasers. However, their use neglects axial variations of the mode intensities. Recently attempts have been made to include the axial effects.[46,47] In general, the rate of spontaneous emission R_{sp} increases with a decrease in the facet reflectivity R_m. For a semiconductor laser whose $R_m = 0.32$, the increase is about 10%.[47]

6.4 TRANSIENT RESPONSE

The previous section considered device behavior under steady-state conditions. When a semiconductor laser is turned on by changing the current I, a relatively long time (~ 10 ns) elapses before the steady state is reached. In the transient regime, the power distribution among various longitudinal modes varies periodically as the laser goes through relaxation oscillations. In particular, a semiconductor laser whose CW mode spectrum is predominantly single-mode exhibits poor side-mode suppression under dynamic conditions. An understanding of the transient response is especially important for optical communication systems where the current is modulated at gigahertz frequencies and the laser never attains a steady state. The purpose of this section is to use the rate equations to describe the dynamic characteristics of a semiconductor laser.

6.4.1 Dynamic Longitudinal-Mode Spectrum

The longitudinal-mode spectrum under transient conditions is obtained by solving the multimode rate equations (6.2.25) and (6.2.26).[48–51] The laser current is abruptly increased, say when time $t = 0$, from its initial value I_0 to the final value I, which is greater than I_{th}. A numerical approach is generally required because of the nonlinear nature of the rate equations. However, approximate analytic solutions can be obtained[51] under specific conditions and provide physical insight into the dynamic process.

Figure 6.7 shows the temporal evolution of N and P_m calculated numerically using parameters corresponding to those of a 1.3-μm InGaAsP laser.[49] The laser current was increased from 0 to $1.5I_{th}$ at $t = 0$. Several features are noteworthy. We see that the photon population in all modes remains zero for a time period known as the turn-on delay time t_d, after which it increases rapidly. The turn-on delay is characteristic of any laser and indicates that stimulated emission does not occur until the carrier concentration has reached its threshold value N_{th}. The delay time $t_d \sim \tau_e$, where

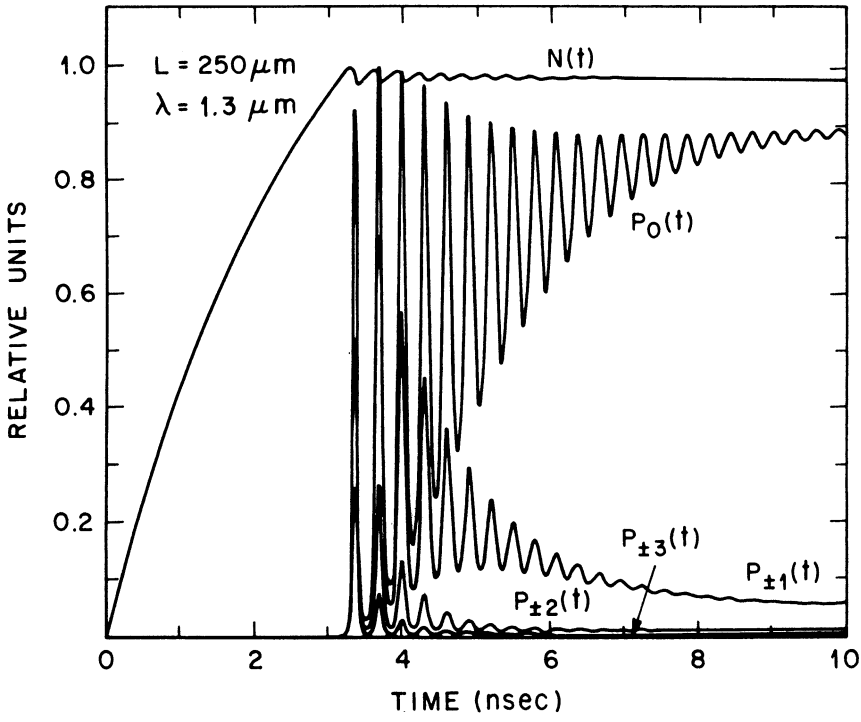


Fig. 6.7 Time evolution of the carrier and photon populations exhibiting relaxation oscillations. (After Ref. 49 © 1983 IEEE)

τ_e is the carrier-recombination time defined in Eq. (6.2.21); typically τ_e = 2–3 ns.

The most important feature of the transient response shown in Fig. 6.7 is that the electron and photon populations oscillate before attaining their steady-state values. These oscillations, referred to as *relaxation oscillations*, are manifestation of an intrinsic resonance in which energy stored in the system oscillates back and forth between the electron and photon populations. The natural frequency of relaxation oscillations is in the gigahertz range and plays an important role in determining device response. Relaxation oscillations are considered further in Sec. 6.4.3.

It is evident from Fig. 6.7 that the transient-mode spectrum is significantly different from the one obtained under CW operation. Figure 6.8 shows the mode spectra under transient and steady-state conditions obtained using Fig. 6.7. The transient spectrum corresponds to the time at which the photon populations first reach their maximum value (the peak of the first relaxation

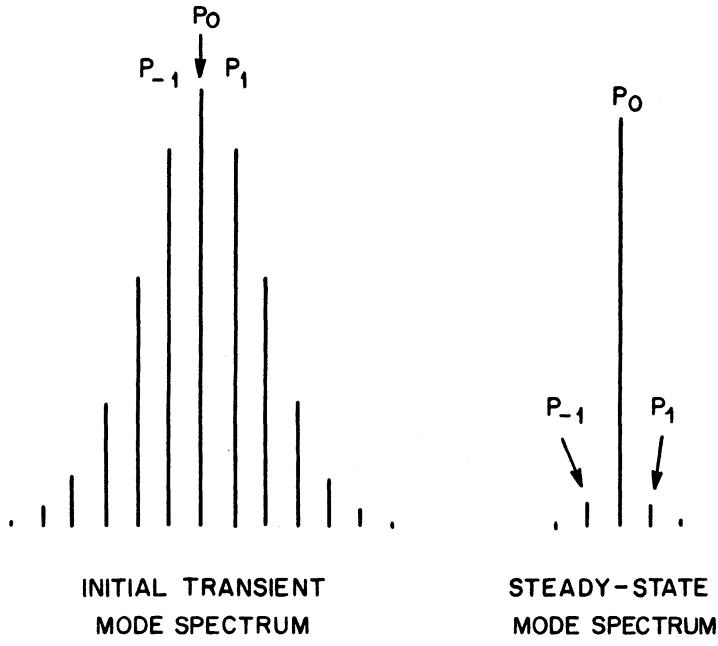


Fig. 6.8 Comparison of longitudinal-mode spectra under transient and steady-state conditions. The transient spectrum corresponds to the first relaxation-oscillations peak in Fig. 6.7. (After Ref. 49 © 1983 IEEE)

oscillation). The transient spectrum is essentially multimode, while the CW spectrum exhibits an MSR greater than 10.

Relaxation oscillations in Fig. 6.7 take several nanoseconds to become sufficiently damped for mode intensities to reach their steady-state values. However, experiments indicate that relaxation oscillations in most semiconductor lasers damp much faster (typically in <1 ns). The reason behind this disagreement can be traced to the assumption that the mode gain G_m in the rate equation (6.2.25) is independent of the mode intensities. The nonlinear gain discussed in Sec. 6.3.2 leads to a reduction in the magnitude of G_m, which depends on the mode intensities.[42–47] As discussed in Sec. 6.4.3, an important effect of the nonlinear gain is to reduce the damping time of relaxation oscillations.

6.4.2 Turn-On Delay

When the laser is turned on by increasing the device current from its initial value I_0 to the above-threshold value I greater than I_{th}, stimulated

recombination is delayed by t_d, the time during which the carrier population rises to its threshold value (see Fig. 6.7). The delay time is determined by the carrier dynamics alone. Its experimental determination is relatively easy and can be used to extract information about the carrier lifetime τ_e.[52–57] If we neglect the stimulated-recombination term in Eq. (6.2.26), the carrier density $n = N/V$ and satisfies the rate equation

$$\frac{dn}{dt} = \frac{I}{qV} - \gamma_e(n)n \tag{6.4.1}$$

where $\gamma_e = \tau_e^{-1}$ and is given by Eq. (6.2.21). Integrating from $t = 0$ to $t = t_d$, we obtain

$$t_d = qV \int_{n_0}^{n_{th}} [I - qV\gamma_e(n)n]^{-1}\, dn \tag{6.4.2}$$

where n_0 and n_{th} are obtained in terms of the current by solving the steady-state relations

$$I_0 = qV\gamma_e(n_0)n_0 \tag{6.4.3a}$$

$$I_{th} = qV\gamma_e(n_{th})n_{th}. \tag{6.4.3b}$$

It is evident that the delay time depends on the functional form of the recombination rate $\gamma_e(n)$. In the presence of Auger recombination, important for long-wavelength semiconductor lasers (see Sec. 3.3), γ_e is a quadratic polynomial in n. In this general case a closed-form expression for t_d is difficult to obtain. In a simple approximation only the constant term $\gamma_e = A_{nr}$ is used, and the result for t_d is[52]

$$t_d = \frac{1}{A_{nr}} \ln\left(\frac{I - I_0}{I - I_{th}}\right). \tag{6.4.4}$$

If only the linear term ($\gamma_e = Bn$) is used after assuming that radiative recombination dominates, Eq. (6.4.2) yields[54]

$$t_d = \left(\frac{qV}{BI}\right)^{1/2}\left[\tanh^{-1}\left(\frac{I_{th}}{I}\right)^{1/2} - \tanh^{-1}\left(\frac{I_0}{I}\right)^{1/2}\right]. \tag{6.4.5}$$

Dixon and Joyce[55] have obtained an analytic expression for t_d when $\gamma_e = A_{nr} + Bn$. As noted by them, it is useful to define a new integration variable

$$I_R = qV\gamma_e(n)n \tag{6.4.6}$$

with the physical significance of the recombination current. Equation (6.4.2) then becomes

$$t_d = \int_{I_0}^{I_{th}} \frac{\tau_e'(I_R)}{I - I_R} dI_R \tag{6.4.7}$$

where τ_e' is the differential recombination time defined as

$$dI_R/dn = qV/\tau_e'. \tag{6.4.8}$$

Using Eq. (6.2.21) in (6.4.6), one can readily see that

$$\tau_e'(n) = (A_{nr} + 2Bn + 3Cn^2)^{-1}. \tag{6.4.9}$$

Equation (6.4.7) can be used to draw several qualitative conclusions that hold for the general case that includes Auger recombination as well. If the laser is prebiased close to threshold, τ_e' is approximately constant in the entire range of integration and can be replaced by its threshold value. Similarly, $I_R \cong I_{th}$, and Eq. (6.4.7) yields the simple relation[55]

$$t_d \cong \tau_e'(n_{th})\left(\frac{I_{th} - I_0}{I - I_{th}}\right). \tag{6.4.10}$$

This equation can be used to estimate the differential recombination time τ_e' from the measurement of t_d. The important point to notice is that depending on the experimental conditions one can obtain either τ_e or τ_e' from the delay-time measurements. The value of $\tau_e(n_{th})$ is obtained when $I_0 = 0$ and $I \gg I_{th}$, so that from Eqs. (6.4.2) and (6.4.3),

$$t_d \cong \tau_e(n_{th})\frac{I_{th}}{I}. \tag{6.4.11}$$

To study the effect of Auger recombination on the delay time in the general case, we have evaluated the integral in Eq. (6.4.2) numerically using the parameter values shown in Table 6.1. The results for the specific case of no prebias ($I_0 = 0$) are shown in Fig. 6.9. For the sake of comparison, the threshold value of the carrier-recombination time has been kept constant for all three curves and is about 2.5 ns. This was achieved by setting $\tau_{nr} = (A_{nr} + Cn_{th}^2)^{-1} = 5$ ns and choosing A_{nr} accordingly for each value of the Auger coefficient C shown in Fig. 6.9. The main effect of Auger recombination is to reduce the delay time compared to its value expected when nonradiative

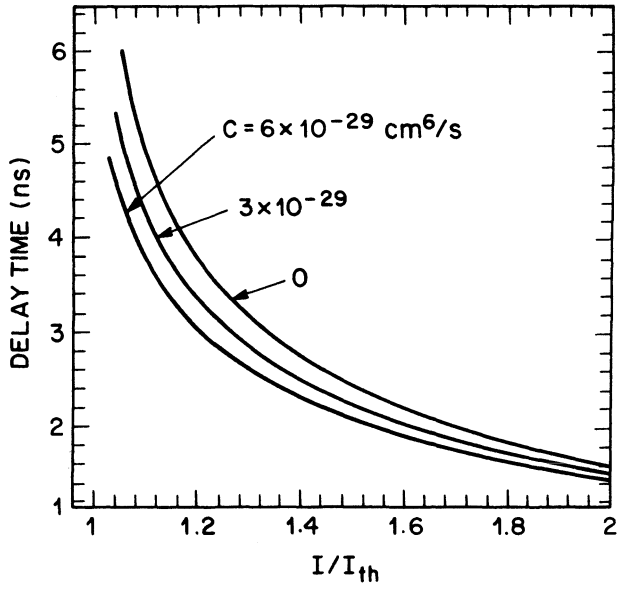


Fig. 6.9 Effect of Auger recombination on the delay time. The total carrier lifetime at threshold is the same for all curves. The parameter values are given in Table 6.1.

recombination is independent of the carrier density. The extent of reduction decreases with an increase in I, and the three curves in Fig. 6.9 merge for $I \gg I_{th}$ in accordance with Eq. (6.4.11).

6.4.3 Relaxation Oscillations

As seen in Fig. 6.7, the output of a semiconductor laser exhibits damped periodic oscillations before settling down to its steady-state value. Such relaxation oscillations are due to an intrinsic resonance in the nonlinear laser system. An expression for their frequency and decay rate can be easily obtained using the small-signal analysis of the single-mode rate equations.[3] From Eqs. (6.2.14) and (6.2.20) these are

$$\dot{P} = (G - \gamma)P + R_{sp} \tag{6.4.12}$$

$$\dot{N} = I/q - \gamma_e N - GP \tag{6.4.13}$$

where the stimulated-emission rate G is a function of N. Even though a linear dependence of G or N is often valid, for the moment we leave it unspecified. Furthermore, we allow G to vary with P to allow for the nonlinear-gain

effects.[42–47] The dependence of G on the power can occur due to several mechanisms such as the spectral-hole burning, carrier heating, and two-photon absorption.

For a given value of the device current I greater than I_{th}, the steady-state solution of each rate equation is readily obtained, as discussed in Sec. 6.3.1. In the small-signal analysis the steady-state values P and N are perturbed by a small amount δP and δN. The rate Eqs. (6.4.12) and (6.4.13) are linearized by neglecting the quadratic and higher powers of δP and δN. We then obtain

$$\delta\dot{P} = -\Gamma_p \delta P + (G_N P + \partial R_{sp}/\partial N)\delta N \tag{6.4.14}$$

$$\delta N = \Gamma_N \delta N - (G + G_P P)\delta P \tag{6.4.15}$$

where

$$\Gamma_P = R_{sp}/P - G_P P \tag{6.4.16}$$

$$\Gamma_N = \gamma_e + N(\partial\gamma_e/\partial N) + G_N P \tag{6.4.17}$$

are the decay rates of fluctuations in the photon and carrier populations, respectively. In obtaining Eqs. (6.4.14) and (6.4.15), the gain $G(N, P)$ was expanded in a truncated Taylor series

$$G(N, P) \cong G + G_N \delta N + G_P \delta P \tag{6.4.18}$$

where $G_N = \partial G/\partial N$ and $G_P = \partial G/\partial P$. The gain derivative G_P is usually negative because of gain suppression occurring at high powers. Even though the gain reduction is usually less than 1%, its inclusion is important since it contributes significantly to the damping of photon fluctuations as seen by Eq. (6.4.16). Typical values of the parameters appearing in the small-signal analysis are listed in Table 6.2 for a buried heterostructure laser with the device parameters given in Table 6.1.

The linear set of Eqs. (6.4.14) and (6.4.15) can be readily solved. If we assume an exponential time dependence

$$\left.\begin{aligned} \delta P(t) &= \delta P_0 \exp(-ht) \\ \delta N(t) &= \delta N_0 \exp(-ht) \end{aligned}\right\} \tag{6.4.19}$$

where δP_0 and δN_0 are the initial values of the perturbation, the decay constant h is complex, indicating an oscillatory approach to equilibrium, and is given by

$$h = \Gamma_R \pm i\Omega_R \tag{6.4.20}$$

Table 6.2 Typical values of the parameters appearing in the small-signal analysis of a buried-heterostructure laser operating at 2 mW of power

PARAMETER	SYMBOL	VALUE
Photon population	P	7.76×10^4
Carrier population	N	2.14×10^8
Steady-state gain	G	$6.41 \times 10^{11}\ s^{-1}$
Spontaneous-emission rate	R_{sp}	$1.28 \times 10^{12}\ s^{-1}$
Gain derivative ($\partial G/\partial N$)	G_N	$5.62 \times 10^3\ s^{-1}$
Gain derivative ($\partial G/\partial P$)	G_P	$-3.4 \times 10^4\ s^{-1}$
Small-signal carrier decay rate	Γ_N	$1.27 \times 10^9\ s^{-1}$
Small-signal photon decay rate	Γ_P	$2.56 \times 10^9\ s^{-1}$
Relaxation-oscillation frequency	$\Omega_R/2\pi$	2.65 GHz
Relaxation-oscillation decay rate	Γ_R	$1.92 \times 10^9\ s^{-1}$

where

$$\Gamma_R = \tfrac{1}{2}(\Gamma_N + \Gamma_P) \tag{6.4.21}$$

is the decay rate of relaxation oscillations and

$$\Omega_R = \left[(G + G_P P)\left(G_N P + \frac{\partial R_{sp}}{\partial N}\right) - \frac{(\Gamma_N - \Gamma_P)^2}{4}\right]^{1/2} \tag{6.4.22}$$

is the angular frequency of relaxation oscillations. The expression for Ω_R can be considerably simplified by noting that the combination of GG_NP in Eq. (6.4.22) dominates by several orders of magnitude and to a good degree of approximation

$$\Omega_R \cong (GG_NP)^{1/2}. \tag{6.4.23}$$

This is a remarkably simple expression. For practical purposes it is more useful to express it in terms of the device current. From Eq. (6.4.13) the steady-state value of P is given by

$$P = (I - I_{th})/(qG) \tag{6.4.24}$$

where $I_{th} = q\gamma_e N$. Using this equation an alternative expression for Ω_R is

$$\Omega_R = \left[\frac{G_N(I - I_{th})}{q}\right]^{1/2}. \tag{6.4.25}$$

It is evident that the most important parameter governing the relaxation-oscillation frequency is the gain derivative G_N, which equals $\partial G/\partial N$. If we use Eq. (6.3.3) and assume the linear dependence of gain on N, then

$$G_N = \Gamma v_g a/V \tag{6.4.26}$$

where a is the gain coefficient introduced in Eq. (2.4.3). Clearly Ω_R increases with a decrease in the active volume V.

The relaxation-oscillation frequency is often expressed in terms of the carrier and photon lifetimes. If we use Eqs. (6.3.3), (6.4.25), and (6.2.26), together with the equation $I_{\text{th}} = q\gamma_e N$, then Ω_R can also be written as

$$\Omega_R = \left[\frac{1 + \Gamma v_g a n_0 \tau_p}{\tau_e \tau_p}\left(\frac{I}{I_{\text{th}}} - 1\right)\right]^{1/2} \tag{6.4.27}$$

where n_0 is the transparency value of the carrier density, $\tau_e = \gamma_e^{-1}$, and $\tau_p = \gamma^{-1}$. The term $\Gamma v_g a n_0 \tau_p$ is sometimes neglected in comparison to 1. However, typically this term is ~ 1 and cannot be neglected. For a laser with parameters given in Table 6.1, its magnitude is 0.9.

The preceding small-signal analysis shows that as long as $\Gamma_R > 0$, fluctuations from the steady state exhibit damped relaxation oscillations. However, if Γ_R becomes negative, fluctuations grow exponentially and the steady state is no longer stable. From Eq. (6.4.16) this can happen only if $G_P = \partial G/\partial P > 0$, i.e., if the gain increases with power. The laser then exhibits self-pulsing that has been observed in gain-guided AlGaAs lasers (see Sec. 5.3).[58] Several possible mechanisms, such as saturable absorption and lateral-hole burning have been proposed; the common feature of these mechanisms is that they all lead to the instability condition $G_P > 0$.[58–63] Self-pulsing is rarely observed in index-guided InGaAsP lasers. It is believed that in these devices phenomena such as spectral-hole burning actually reduce the gain with an increase in the output power so that $G_P < 0$.[42–47] From Eqs. (6.4.16) and (6.4.21) we find that Γ_R then increases with power, and the relaxation oscillations are rapidly damped.[64,65]

6.5 NOISE CHARACTERISTICS

In the preceding deterministic description, laser power and frequency were assumed to remain constant in time once the steady state had been reached. In reality, however, laser output exhibits intensity as well as phase fluctuations. At the most fundamental level, the origin of these fluctuations lies in the quantum nature of the lasing process itself. A proper description therefore

requires a quantum-mechanical formulation of the rate equations.[66–69] In general, intensity noise reaches its peak in the vicinity of the laser threshold and then decreases rapidly with an increase in the drive current.[70–73] The intensity-noise spectrum shows a peak near the relaxation-oscillation frequency as a consequence of the laser's intrinsic resonance. Phase fluctuations produce spectral broadening of each longitudinal mode and are responsible for the observed line width. The noise characteristics of semiconductor lasers have been studied extensively both theoretically and experimentally.[66–99]

6.5.1 Langevin Formulation

Within the semiclassical treatment, fluctuations arising from the spontaneous-emission process and the carrier-generation-recombination process are incorporated by adding a Langevin noise source[91] to the single-mode rate equations (6.2.14), (6.2.20), and (6.2.18); these then become

$$\dot{P} = (G - \gamma)P + R_{sp} + F_P(t) \tag{6.5.1}$$

$$\dot{N} = I/q - \gamma_e N - GP + F_N(t) \tag{6.5.2}$$

$$\dot{\phi} = -(\omega_0 - \omega_{th}) + \tfrac{1}{2}\beta_c(G - \gamma) + F_\phi(t). \tag{6.5.3}$$

Physically F_P and F_ϕ arise from spontaneous emission, while F_N has its origin in the discrete nature of the carrier generation and recombination processes (shot noise). In the presence of Langevin noise sources, P, N, and ϕ become random and their dynamics are governed by the stochastic rate equations (6.5.1)–(6.5.3). The problem simplifies considerably if we make the Markovian assumption; i.e., we assume that the correlation time of the noise sources is much shorter than the relaxation times γ^{-1} and γ_e^{-1} (system has no memory). Under the Markovian assumption, the Langevin forces satisfy the general relations[91]

$$\langle F_i(t)\rangle = 0 \tag{6.5.4}$$

$$\langle F_i(t)F_j(t')\rangle = 2D_{ij}\delta(t - t') \tag{6.5.5}$$

where angle brackets denote ensemble average and D_{ij} is the diffusion coefficient associated with the corresponding noise source. In a rigorous approach,[91] D_{ij}'s are obtained by evaluating the second moments of the dynamic variables with the help of Eqs. (6.5.1)–(6.5.3). They can also be obtained using physical arguments such as those presented in Ref. 77. Their

explicit expressions are[77,91]

$$D_{PP} = R_{sp}P \qquad D_{\phi\phi} = R_{sp}/4P \qquad D_{P\phi} = 0 \tag{6.5.6}$$

$$D_{NN} = R_{sp}P + \gamma_e N \qquad D_{PN} = -R_{sp}P \qquad D_{N\phi} = 0. \tag{6.5.7}$$

Here N and P represent the steady-state *average* values of the carrier and photon populations, respectively, and are the same as obtained previously in Sec. 6.3.

To obtain the noise characteristics, the steady-state values P, N, and ϕ are perturbed by small amounts δP, δN, and $\delta\phi$, respectively. In the small-signal analysis the stochastic rate equations (6.5.1)–(6.5.3) are linearized, and we obtain

$$\delta\dot{P} = -\Gamma_P \delta P + (G_N P + \partial R_{sp}/\partial N)\delta N + F_P(t) \tag{6.5.8}$$

$$\delta\dot{N} = -\Gamma_N \delta N - (G + G_P P)\delta P + F_N(t) \tag{6.5.9}$$

$$\delta\dot{\phi} = \tfrac{1}{2}\beta_c G_N \delta N + F_\phi(t) \tag{6.5.10}$$

where the small-signal decay rates Γ_P and Γ_N are defined by Eqs. (6.4.16) and (6.4.17). Except for the presence of the noise terms, Eqs. (6.5.8) and (6.5.9) are identical to those obtained in Sec. 6.4.3 for the analysis of relaxation oscillations. The phase equation (6.5.10) is also considered here to incorporate phase fluctuations. The power dependence of gain arising from nonlinear effects[42–48] such as spectral-hole burning is included through G_P. However, its contribution to the phase is neglected since the index change associated with such nonlinear phenomena is expected to be much smaller than the carrier-induced index change.[44]

The linear set of Eqs. (6.5.8)–(6.5.10) can be solved in the frequency domain using Fourier analysis. If the Fourier transform of an arbitrary function $f(t)$ is defined as

$$\tilde{f}(\omega) = \int_{-\infty}^{\infty} f(t)\exp(-i\omega t)\,dt \tag{6.5.11}$$

the solutions for the Fourier components are

$$\delta\tilde{P}(\omega) = \frac{(\Gamma_N + i\omega)\tilde{F}_P + (G_N P + \partial R_{sp}/\partial N)\tilde{F}_N}{(\Omega_R + \omega - i\Gamma_R)(\Omega_R - \omega + i\Gamma_R)} \tag{6.5.12}$$

$$\delta\tilde{N}(\omega) = \frac{(\Gamma_P + i\omega)\tilde{F}_N - (G + G_P P)\tilde{F}_P}{(\Omega_R + \omega - i\Gamma_R)(\Omega_R - \omega + i\Gamma_R)} \tag{6.5.13}$$

and

$$\delta\tilde{\phi}(\omega) = \frac{1}{i\omega}(\tilde{F}_\phi + \tfrac{1}{2}\beta_c G_N \delta\tilde{N}) \tag{6.5.14}$$

where the decay rate Γ_R and the frequency Ω_R of relaxation oscillations are given by Eqs. (6.4.21) and (6.4.22), respectively. Equations (6.5.12)–(6.5.14) clearly show that noise is enhanced in the vicinity of $\omega = \Omega_R$ because of the intrinsic resonance exhibited by the carrier and photon populations. These equations can be used to study how processes such as spontaneous emission and carrier recombination affect the light intensity, the optical phase, and the carrier population. In the next two sections we consider separately the intensity noise and the phase noise.

6.5.2 Intensity Noise

The power emitted by a semiconductor laser fluctuates around its steady-state value. The spectrum of these fluctuations can be measured by detecting the laser output using a wide-bandwidth photodiode and a spectrum analyzer. Such a measurement yields the noise spectrum associated with the total power. The noise spectrum of a single mode can be obtained at first selecting a single longitudinal mode using a monochromator. The intensity noise at a given frequency ω is characterized by the *relative intensity noise* (RIN) defined as

$$\text{RIN} = S_P(\omega)/P^2 \tag{6.5.15}$$

where the spectral density

$$S_P(\omega) = \int_{-\infty}^{\infty} \langle \delta P(t+\tau)\delta P(t)\rangle \exp(-i\omega\tau)\, d\tau \tag{6.5.16}$$

is related to $\delta\tilde{P}(\omega)$ simply by

$$S_P(\omega) = \operatorname*{Lim}_{T\to\infty} \frac{1}{T}|\delta\tilde{P}(\omega)|^2. \tag{6.5.17}$$

The ensemble average in Eq. (6.5.16) was replaced by a time average over the interval T assuming an ergodic stochastic process. We substitute $\delta\tilde{P}(\omega)$ from Eq. (6.5.12) and make use of the relation that

$$\operatorname*{Lim}_{T\to\infty} \frac{1}{T}[\tilde{F}_i^*(\omega)\tilde{F}_j(\omega)] = 2D_{ij} \tag{6.5.18}$$

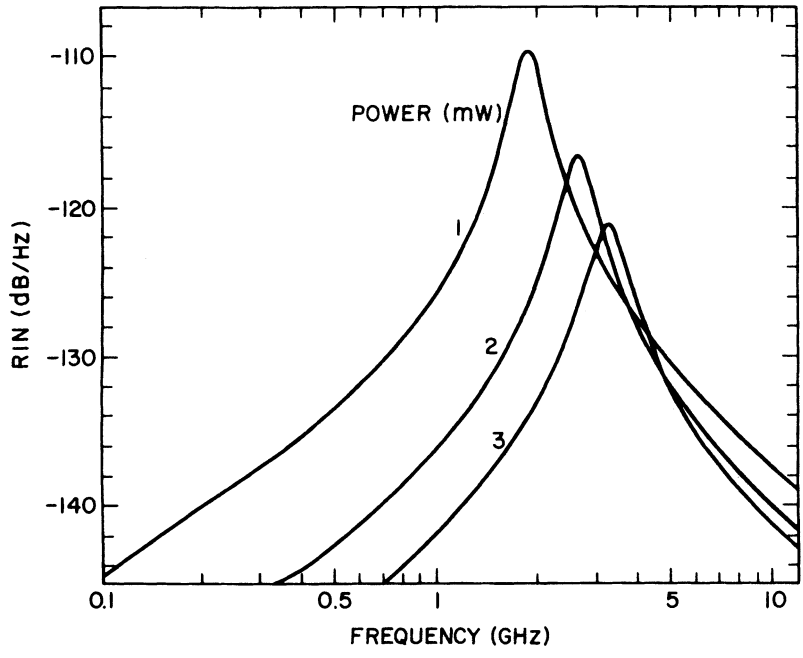


Fig. 6.10 Intensity-noise spectra at several power levels showing the relaxation-oscillation peak.

which follows from Eq. (6.5.5) and implies that the Langevin noise sources have a white (frequency-independent) spectrum. Using D_{PP}, D_{PN}, and D_{NN} from Eqs. (6.5.6) and (6.5.7), we obtain

$$\text{RIN} = \frac{2R_{sp}[(\Gamma_N^2 + \omega^2) + G_N^2 P^2(1 + \gamma_e N/R_{sp}P) - 2\Gamma_N G_N P]}{P[(\Omega_R - \omega)^2 + \Gamma_R^2][(\Omega_R + \omega)^2 + \Gamma_R^2]}. \quad (6.5.19)$$

Figure 6.10 shows the calculated intensity-noise spectra at several laser powers for a 1.3-μm InGaAsP laser using the parameter values shown in Table 6.1. At a given laser power, the noise is relatively low when $\omega \ll \Omega_R$ and is enhanced significantly in the vicinity of $\omega = \Omega_R$. At a given frequency, RIN decreases with an increase in the laser power P. To see the functional dependence of RIN on P, consider the low-frequency limit when $\omega \ll \Omega_R$. Since $\Gamma_R \ll \Omega_R$, the denominator in Eq. (6.5.19) can be replaced by $P\Omega_R^4$. Further, at low power levels the terms proportional to P and P^2 in the numerator make a relatively small contribution. If we neglect these terms and use Eq. (6.4.23) to express Ω_R in terms of P, we find that RIN decreases with the laser power as P^{-3}. The inverse cubic dependence of RIN on P has

been observed for AlGaAs as well as InGaAsP lasers.[81,86] At higher bias levels, the neglected terms in Eq. (6.5.19) become important and RIN varies more slowly approaching the P^{-1} dependence. This behavior has also been observed experimentally.[83]

Another measure of the intensity noise is proviced by the autocorrelation function $C_{pp}(\tau)$ defined as

$$C_{pp}(\tau) = \langle \delta P(t)\delta P(t+\tau)\rangle / P^2. \tag{6.5.20}$$

From Eqs. (6.5.15) and (6.5.16) $C_{pp}(\tau)$ is related to RIN(ω) by the Fourier-transform relation

$$C_{pp}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \text{RIN}(\omega)\exp(i\omega\tau)\,d\omega. \tag{6.5.21}$$

The integral in Eq. (6.5.21) can be performed by using the method of contour integration. The result is[92]

$$C_{pp}(\tau) = \frac{R_{\text{sp}}\exp(-\Gamma_R\tau)}{2\Gamma_R P}\,\text{Re}\left(\frac{\Gamma_e^2 + (\Omega_R + i\Gamma_R)^2}{\Omega_R(\Omega_R + i\Gamma_R)}\exp(i\Gamma_R\tau)\right) \tag{6.5.22}$$

where Re stands for the real part and

$$\Gamma_e^2 = \Gamma_N^2 + G_N^2P^2(1 + \gamma_e N/R_sP) - 2\Gamma_N G_N P. \tag{6.5.23}$$

The intensity autocorrelation function oscillates with τ (as a consequence of relaxation oscillations) and vanishes when τ becomes comparable to the damping time Γ_R^{-1} of relaxation oscillations.[92] The quantity of practical interest is $C_{pp}(0)$ since it is related to the variance of intensity fluctuations. Indeed, one can define the signal-to-noise ratio (SNR) of the laser light as

$$\text{SNR} = \frac{1}{\sqrt{C_{pp}(0)}} = \left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\text{RIN}(\omega)\,d\omega\right)^{-1/2}. \tag{6.5.24}$$

By substituting $C_{pp}(0)$ from Eq. (6.5.22) the SNR is given by

$$\text{SNR} = \left(\frac{2\Gamma_R P}{R_{\text{sp}}}\right)^{1/2}\left(1 + \frac{\Gamma_e^2}{\Omega_R^2 + \Gamma_R^2}\right)^{-1/2}. \tag{6.5.25}$$

The last factor can be approximated by 1 since $\Gamma_e^2 \ll \Omega_R^2 + \Gamma_R^2$ in most cases

of practical interest. The SNR is then given by a remarkably simple expression

$$\text{SNR} = \left(\frac{2\Gamma_R P}{R_{sp}}\right)^{1/2}. \tag{6.5.26}$$

It shows that the SNR depends on the rate of spontaneous emission and degrades as $R_{sp}^{-1/2}$ with an increase in it. It also shows that the damping rate Γ_R of relaxation oscillations plays a crucial role in determining the SNR. The power dependence of SNR is not immediately obvious from Eq. (6.5.26) since Γ_R itself depends on power. By using $2\Gamma_R = \Gamma_N + \Gamma_P$ from Eq. (6.4.21) and Γ_N and Γ_P from Eqs. (6.4.16) and (6.4.17), the SNR becomes

$$\text{SNR} = \left(1 + \frac{P}{R_{sp}}[\gamma_e + N(\partial\gamma_e/\partial N) + (G_N - G_P)P]\right)^{1/2}. \tag{6.5.27}$$

Near or below threshold, SNR reduces to 1, as it should, since the intensity fluctuations then correspond to those of thermal noise with an exponential distribution for the intensity. In the above-threshold regime, SNR begins to increase, first as $\sqrt{P}$ and then linearly with P. Far above threshold, the SNR varies with P as

$$\text{SNR} = \left(\frac{G_N - G_P}{R_{sp}}\right)^{1/2} P. \tag{6.5.28}$$

Since G_P is negative because of the gain suppression occurring at high powers (see Table 6.2), the nonlinear gain plays an important role in determining the SNR of the output power. In fact, $|G_P|$ is typically much larger than G_N, and the SNR is directly proportional to $\sqrt{|G_P|}$. According to Eq. (6.5.28), the SNR would keep on improving with an increase in the laser power. However, this feature is an artifact of the assumption made in Eq. (6.4.18) by truncating the Taylor series. A more accurate treatment shows that the nonlinear gain eventually limits the SNR at a level of about 30 dB at high powers.[92] At a power level of a few milliwatts the SNR of most semiconductor lasers is about 20 dB. Thus the root-mean-square value of intensity fluctuations is close to 1% under typical operating conditions.

For a multimode laser, the qualitative behavior shown in Fig. 6.10 is observed experimentally if the total power in all modes is detected. However, low-frequency noise is found to be significantly higher for individual longitudinal modes. Figure 6.11 shows this behavior for a laser with a mode-supression ratio of more than 20. At frequencies below 500 MHz the RIN for the dominant mode is higher by more than 30 dB compared to the total

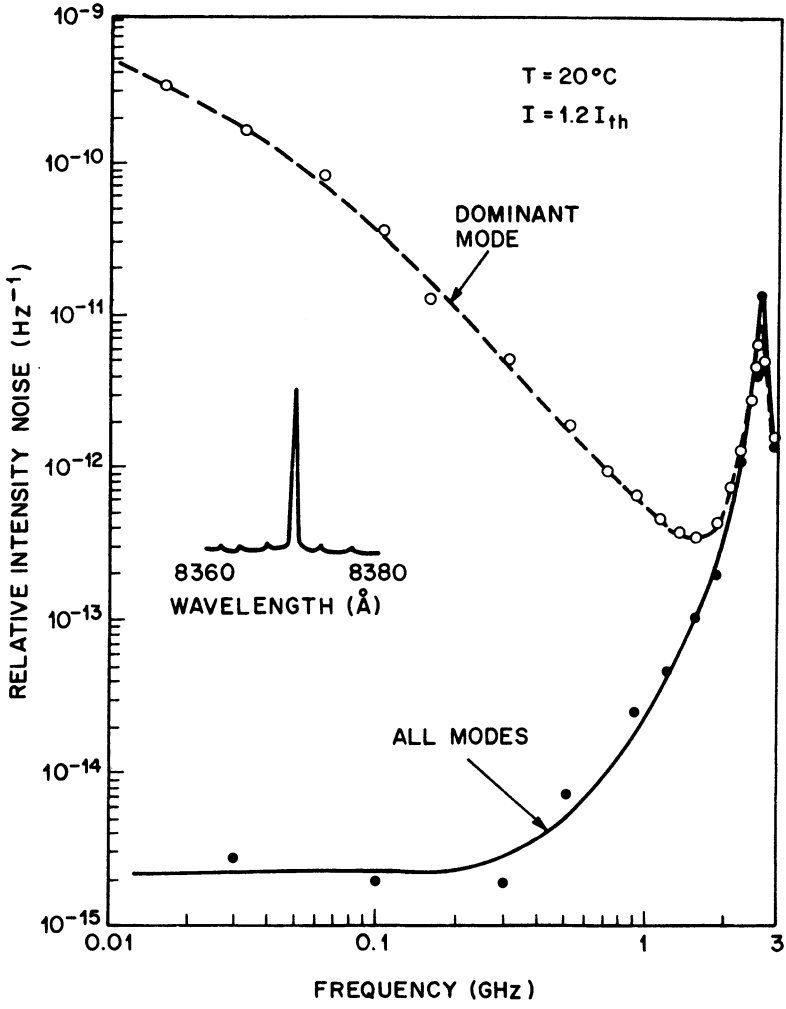


Fig. 6.11 Partition noise in a channeled-substrate planar waveguide laser. Experimentally observed intensity-noise spectra are shown when the power in all modes (solid curve) or in the dominant mode (dashed curve) is detected. Inset shows the average mode powers. (After Jäckel, Ref. 71)

intensity noise. This increase is referred to as the *mode-partition noise*[71] and has attracted considerable attention.[93–97] It arises when the main and side modes fluctuate in such a way as to leave the total intensity relatively constant.

An alternative way to characterize intensity noise is through the probability

distribution function of the random variable P. Experimentally it is obtained[94] by repetitively sampling the light intensity (using a fast detector) for a sufficiently short duration (less than ~ 100 ps) so that the signal represents instantaneous intensity and fluctuations do not average out. As one may expect, the observed distribution depends on the sampling-time window and rapidly narrows with an increase in the sampling time. The evidence for mode-partition noise was also found[94] in the recorded distribution functions for the main mode and the side modes.

Mode-partition noise can be modeled by adding the Langevin-noise terms to the multimode rate equations (6.2.25) and (6.2.26) so that

$$P_m = (G_m - \gamma_m)P_m + R_{sp} + F_m(t) \tag{6.5.29}$$

$$N = I/q - \gamma_e N - \sum G_m P_m + F_N(t) \tag{6.5.30}$$

where the sum is over the total number of modes. The Langevin forces satisfy Eqs. (6.5.4) and (6.5.5) with the diffusion coefficients

$$D_{mm} = R_{sp}P_m, \qquad D_{mn} = 0, \qquad D_{mN} = -R_{sp}P_m, \qquad D_{NN} = R_{sp}\sum P_m + \gamma_e N. \tag{6.5.31}$$

The spontaneous emission rate is assumed to be the same for all modes.

The multimode rate equations generally need to be solved numerically. However, they can be solved analytically[97] by following the linearization procedure that assumes fluctuations from the steady state to be small. The analytic solution is particularly simple when only two modes are considered. This case is of practical importance since many semiconductor lasers are designed to oscillate predominantly in a single longitudinal mode (see Chapters 7 and 8) and can be modeled by considering a single side mode accompanied by the dominant main mode. Fluctuations in the two modes are given by

$$\delta\tilde{P}_1(\omega) = [(\Gamma_2 + i\omega)\tilde{F}_1 - \theta_{12}P_1\tilde{F}_2 + G_N P_1(\Gamma_2 + i\omega - \theta_{12}P_2)\delta\tilde{N}]/D(\omega) \tag{6.5.32}$$

$$\delta\tilde{P}_2(\omega) = [(\Gamma_1 + i\omega)\tilde{F}_2 - \theta_{21}P_2\tilde{F}_1 + G_N P_2(\Gamma_1 + i\omega - \theta_{21}P_1)\delta\tilde{N}]/D(\omega) \tag{6.5.33}$$

where

$$\delta\tilde{N}(\omega) = (\tilde{F}_N - \bar{G}_1\delta\tilde{P}_1 - \bar{G}_2\delta\tilde{P}_2)/(\Gamma_N + i\omega) \tag{6.5.34}$$

$$D(\omega) = (\Gamma_1 + i\omega)(\Gamma_2 + i\omega) - \theta_{12}\theta_{21}P_1P_2 \tag{6.3.35}$$

$$\Gamma_1 = R_{sp}/P_1 + \beta_{11}P_1, \qquad \Gamma_2 = R_{sp}/P_2 + \beta_{22}P_2. \tag{6.5.36}$$

The effects of nonlinear gain, included in the single-mode case through a single parameter G_p, are now included through four parameters introduced by expanding G_1 and G_2 in Eq. (6.5.29) as

$$G_1 = \bar{G}_1 + G_N \delta N - \beta_{11} \delta P_1 - \theta_{12} \delta P_2 \tag{6.5.37}$$

$$G_2 = \bar{G}_2 + G_N \delta N - \beta_{22} \delta P_2 - \theta_{21} \delta P_1. \tag{6.5.38}$$

The parameters β_{11} and β_{22} govern self-saturation responsible for a self-induced decrease of the modal gain. The parameters θ_{12} and θ_{21} govern cross-saturation reflecting the fact that the gain of a particular mode is also affected by the intensity of other modes.

The RIN for the main mode, side mode, and the total intensity can be obtained by following the same procedure outlined before for the single-mode case. Figure 6.12 shows the RIN spectra for the main mode for several values of S/P, where P and S are average photon populations for the main and side modes. In all cases the RIN for the total intensity $P + S$ nearly coincides with the curve shown for $S/P = 10^{-5}$ in Fig. 6.12. For a 50-dB

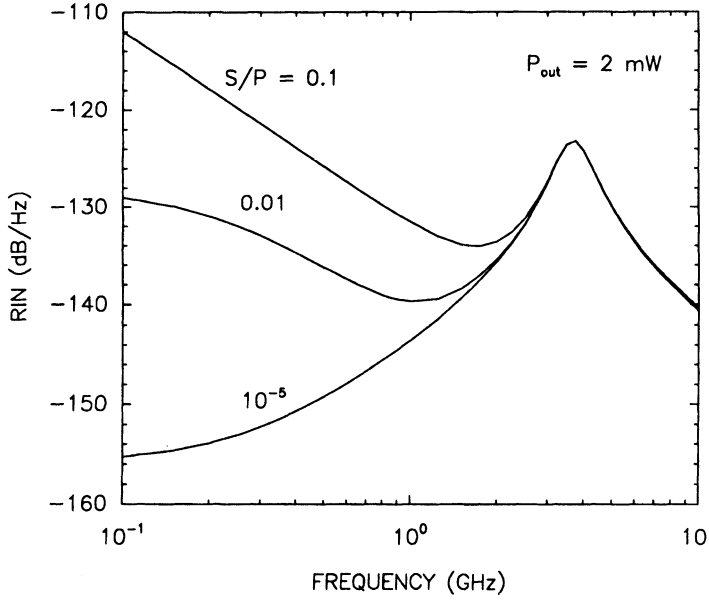


Fig. 6.12 RIN spectra of the main-mode intensity calculated for several values of the MSR for a two-mode laser operating at 2 mW. The RIN spectrum of the total intensity remains the same in the three cases and nearly coincides with the curve corresponding to $S/P = 10^{-5}$. Low-frequency RIN enhancement seen for larger values of S/P is due to mode-partition noise.

MSR the side mode is too weak to affect the main-mode RIN. However, the main-mode RIN is considerably enhanced at frequencies below 1 GHz when the MSR is 10 and 20 dB. The enhancement can exceed 40 dB at low frequencies (< 100 MHz), in qualitative agreement with the experimental data shown in Fig. 6.11. The physical origin of such an enhancement is not immediately obvious. Considerable insight is obtained by solving the multimode rate equations (6.5.29) and (6.5.30) numerically. Numerical simulations show that even though fluctuations in $P(t)$ are relatively small most of the time, occasionally the side-mode intensity becomes comparable to that of the main mode. Figure 6.13 shows such an event. These events last for a relatively small time (a few nanoseconds). However, the main-mode intensity drops considerably during such events, resulting in large fluctuations in $P(t)$ from its average value. Since such events occur on a relatively long time scale (> 10 ns), it is only the low-frequency RIN that is enhanced. The origin of such drop-off events is related to the gain saturation and to the fact that the same electron population provides gain for both the modes. The side mode exhibits larger fluctuations because of its lower average intensity. An unusually large random increase in the side-mode intensity

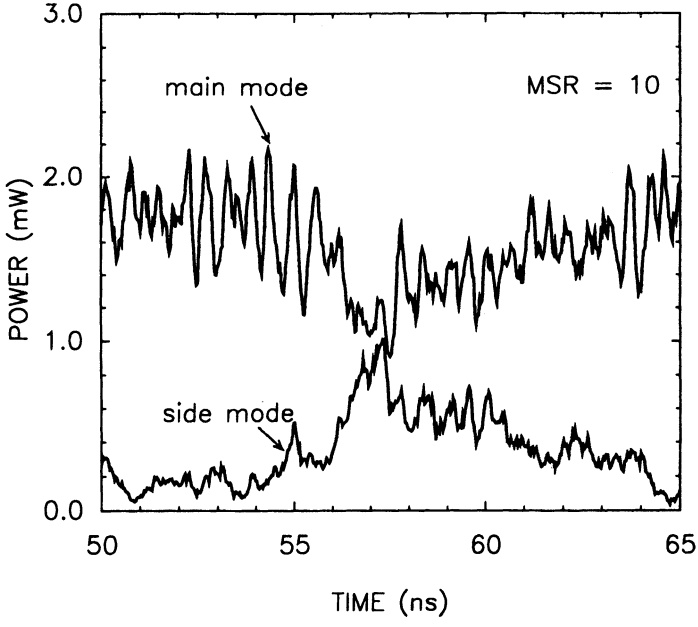


Fig. 6.13 Numerically simulated power fluctuations of main and side modes for a laser with a total average power of 2 mW. Occasionally main-mode power drops below the side-mode even though it is 10 times larger on average.

induces a simultaneous decrease in the main-mode intensity, lasting a few nanoseconds because of relaxation oscillations.

6.5.3 Phase Noise and Line Width

Quantum fluctuations associated with the lasing process affect both the intensity and the phase of the optical field. A change in phase leads to a frequency shift $\delta\omega_L = \delta\dot{\phi}$. The frequency or phase noise is of interest in evaluating the performance of coherent optical communication systems. The line width of a single longitudinal mode under CW operation is a manifestation of phase fluctuations occurring inside the laser. The solution (6.5.14) can be used to obtain the frequency-noise spectrum as well as the line width.

Two mechanisms contribute to phase fluctuations. In Eq. (6.5.14), the first term ($\tilde{F}_\phi$) is due to spontaneous emission. Each spontaneously emitted photon changes the optical phase by a random amount. The second term ($\frac{1}{2}\beta_c G_N \delta\tilde{N}$) shows that fluctuations in the carrier populations also lead to a phase change. Physically, this is so because a change in N affects not only the optical gain but also the refractive index (or the optical path length) and consequently the phase. The parameter β_c provides a proportionality between the gain and index changes [see Eq. (6.2.17)]. It is sometimes referred to as the line-width enhancement factor since it was found[77] that the second term in Eq. (6.5.14) increases the line width by a factor of $1 + \beta_c^2$.

The spectral density of the frequency noise is defined similar to the intensity-noise case [see Eqs. (6.5.16) and (6.5.17)] and is

$$S_{\dot{\phi}}(\omega) = \langle |\omega \delta\tilde{\phi}(\omega)|^2 \rangle \tag{6.5.39}$$

where $\delta\tilde{\phi}(\omega)$ is given by Eq. (6.5.14). The general expression for $S_{\dot{\phi}}(\omega)$ is fairly complicated. It can however be considerably simplified by noting that the main contribution to $\delta\tilde{N}$ comes from the term proportional to $\tilde{F}_P$ in Eq. (6.5.13). Using Eqs. (6.5.6) and (6.5.18), an approximate expression for the frequency-noise spectrum is

$$S_{\dot{\phi}}(\omega) \cong \frac{R_{sp}}{2P}\left(1 + \frac{\beta_c^2 \Omega_R^4}{[(\Omega_R^2 - \omega^2)^2 + (2\omega\Gamma_R)^2]}\right) \tag{6.5.40}$$

where we have used Eq. (6.4.23).

Equation (6.5.40) shows that the frequency-noise spectrum is relatively flat when $\omega \ll \Omega_R$ and peaks at the relaxation-oscillation frequency Ω_R. The qualitative behavior is similar to the intensity-noise spectrum shown in

Fig. 6.10. In both cases, even though the Langevin noise is white (frequency-independent), noise in the vicinity of Ω_R is selectively enhanced by the intrinsic resonance manifested by a semiconductor laser. This enhancement has been observed experimentally.[78]

The line width of a single longitudinal mode can be obtained by recording the optical power spectrum using a high-resolution FP interferometer.[74–76] Mathematically, the power spectrum is obtained by taking the Fourier transform of the field autocorrelation and is given by [see Eq. (6.5.16)]

$$S_E(\omega) = \int_{-\infty}^{\infty} \langle E^*(t+\tau)E(t)\rangle \exp(-i\omega\tau)\, d\tau \tag{6.5.41}$$

where $E(t)$ is the optical field such that

$$E(t) = (P + \delta P)^{1/2} \exp[-i(\omega_0 t + \phi + \delta\phi)]. \tag{6.5.42}$$

P and ϕ are the steady-state values, and ω_0 is the frequency of the single longitudinal mode under consideration.

The power spectrum $S_E(\omega)$ can be calculated[77–79] using the small-signal solution given by Eqs. (6.5.12)–(6.5.14). It consists of a central peak at the mode frequency ω_0 followed by much weaker satellite peaks at $\omega_0 \pm n\Omega_R$ (n being an integer). These side peaks arise from relaxation oscillations and have been experimentally observed.[74,78,79] Furthermore, a small symmetry in the amplitudes of the side peaks was observed[79] and is interpreted to be due to a correlation between intensity and phase fluctuations.

To simplify the analysis, we neglect intensity fluctuations and set δP equal to 0 in Eq. (6.5.42). The only effect of this simplification is that side-peak asymmetry is excluded from the theoretical treatment. Using Eqs. (6.5.41) and (6.5.42), we then obtain

$$S_E(\omega) = P \int_{-\infty}^{\infty} \langle \exp(i\Delta_\tau\phi)\rangle \exp[-i(\omega - \omega_0)\tau]\, d\tau \tag{6.5.43}$$

where

$$\Delta_\tau\phi = \delta\phi(t+\tau) - \delta\phi(t) \tag{6.5.44}$$

is the difference in phase fluctuations at times t and $t + \tau$. Assuming that $\Delta_\tau\phi$ has a Gaussian distribution,[91] the average in Eq. (6.5.43) is easily performed to obtain

$$\langle \exp(i\Delta_\tau\phi)\rangle = \exp[-\tfrac{1}{2}\langle(\Delta_\tau\phi)^2\rangle]. \tag{6.5.45}$$

To evaluate the variance $\langle(\Delta_\tau\phi)^2\rangle$ associated with the stationary random process $\Delta_\tau\phi$, we note by taking the Fourier transform of Eq. (6.5.44) that

$$\Delta_\tau\phi = \frac{1}{2\pi}\int_{-\infty}^{\infty} \delta\tilde{\phi}(\omega)[\exp(i\omega\tau) - 1]\exp(i\omega t)\,d\omega. \tag{6.5.46}$$

Since cross-correlation between the frequency components $\delta\tilde{\phi}(\omega)$ vanishes, it can be easily shown that

$$\langle(\Delta_\tau\phi)^2\rangle = \frac{1}{\pi}\int_{-\infty}^{\infty} \langle|\delta\tilde{\phi}(\omega)|^2\rangle(1 - \cos\omega\tau)\,d\omega. \tag{6.5.47}$$

If we use $\delta\tilde{\phi}(\omega)$ from Eq. (6.5.14) and evaluate the average using Eq. (6.5.17), the integral can be done by the method of contour integration. The result is[77]

$$\langle(\Delta_\tau\phi)^2\rangle = \frac{R_{sp}}{2P}\left((1 + \beta_c^2 b)\tau + \frac{\beta_c^2 b}{2\Gamma_R\cos\delta}\right.$$
$$\left.\times\,[\cos(3\delta) - \exp(-\Gamma_R\tau)\cos(\Omega_R\tau - 3\delta)]\right) \tag{6.5.48}$$

where

$$b = \Omega_R/(\Omega_R^2 + \Gamma_R^2)^{1/2} \qquad \text{and} \qquad \delta = \tan^{-1}(\Gamma_R/\Omega_R). \tag{6.5.49}$$

The optical spectrum is obtained by substituting Eqs. (6.5.45) and (6.5.48) in Eq. (6.5.43) and evaluating the Fourier transform numerically. Figure 6.14 shows an example of the calculated spectrum for a semiconductor laser by using parameter values from Table 6.2 together with $\beta_c = 5$. Only half of the spectrum is shown since the spectrum is symmetric around $\omega = \omega_0$. It consists of a dominant central peak located at ω_0 and multiple satellite peaks located at $\omega = \omega_0 \pm m\Omega_R$, where m is an integer. The amplitude of the satellite peaks decreases with the frequency spacing. In practice, often only the first pair of satellite peaks located at $\omega = \omega_0 \pm \Omega_R$ is visible experimentally. Even the amplitude of this pair of satellite peaks is much smaller compared with the central peak (typically 1% or less at a power level of 1 mW). As the laser power increases, the central peak narrows and the amplitude of satellite peaks decreases even more. These spectral features have been observed experimentally.[79] The physical origin of the satellite peaks is related to relaxation oscillations which are responsible for the term proportional to b in Eq. (6.5.48). As a result, the location and the amplitude of satellite peaks is strongly affected by the frequency and the damping rate

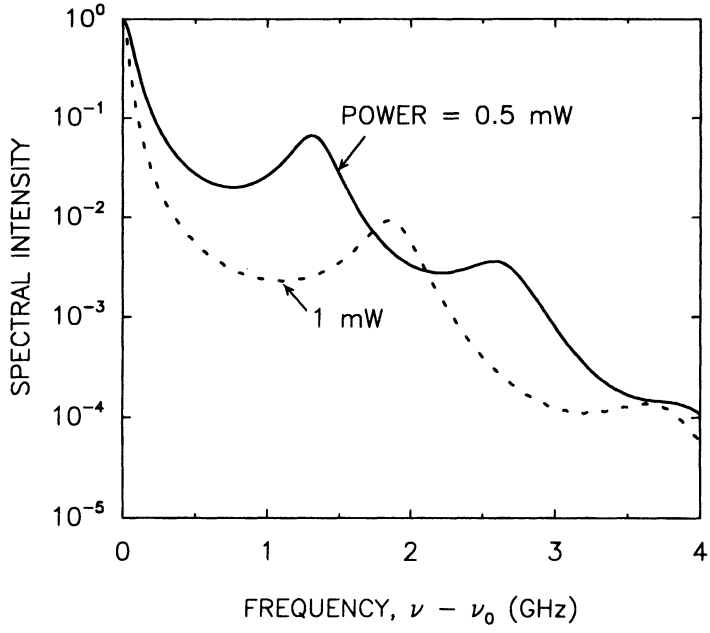


Fig. 6.14 Spectral line shape calculated for a single-mode semiconductor laser operating at 0.5-mW and 1-mW output power. Satellite peaks resulting from relaxation oscillations are clearly seen because of the logarithmic power scale.

of relaxation oscillations, which are in turn affected by the laser power. At a power level of a few milliwatts the spectrum is dominated by the central peak whose width determines the line width of the laser mode.

An approximate expression for the power spectrum can be obtained by using Eq. (6.5.39) in (6.5.47) to yield

$$\langle(\Delta_\tau\phi)^2\rangle = \frac{1}{\pi}\int_{-\infty}^{\infty} S_{\dot{\phi}}(\omega)\frac{(1-\cos\omega\tau)}{\omega^2}\,d\omega \tag{6.5.50}$$

where the frequency-noise spectral density $S_{\dot{\phi}}(\omega)$ is given by Eq. (6.5.40). The origin of the side peaks in the power spectrum is a direct consequence of the relaxation-oscillation-induced enhancement of $S_{\dot{\phi}}(\omega)$ at $\omega = \Omega_R$. If we neglect this enhancement by assuming that $S_{\dot{\phi}}(\omega)$ is constant and replace it by its value at $\omega = 0$, the integration is readily performed. The result is

$$\langle(\Delta_\tau\phi)^2\rangle = \tau S_{\dot{\phi}}(0). \tag{6.5.51}$$

Using Eqs. (6.5.43), (6.5.45), and (6.5.51), the power spectrum is found to be a simple Lorentzian centered at $\omega = \omega_0$ and having an FWHM of $\Delta\omega = S_{\dot{\phi}}(0)$. Using $\Delta\omega = 2\pi\Delta f$, the line width Δf is given by

$$\begin{aligned}\Delta f &= \frac{1}{2\pi} S_{\dot{\phi}}(0) \\ &= \frac{R_{sp}(1 + \beta_c^2)}{4\pi P} \\ &= (1 + \beta_c^2)\Delta f_0 \end{aligned} \tag{6.5.52}$$

where we have used Eq. (6.5.40) and the fact that $\Gamma_R \ll \Omega_R$. This expression shows that the line width of a semiconductor laser is enhanced by a factor of $1 + \beta_c^2$ compared with its value Δf_0 expected on the basis of the modified Schawlow-Townes formula.[91] Henry[77] was the first to point out the significance of this enhancement factor in an attempt to reconcile the experimental results[75] with theory. The enhancement factor was present in the earlier work[68,91] on laser line width, but had been thought to be unimportant since $\beta_c \ll 1$ for lasers other than semiconductor lasers.

Physically the two terms in the last line of Eq. (6.5.52) can be interpreted to arise from instantaneous and delayed phase fluctuations resulting from each spontaneous-emission event.[77] The contribution Δf_0 is due to instantaneous phase change governed by the Langevin-noise source $F_\phi(t)$ in Eq. (6.5.10). The contribution $\beta_c^2 \Delta f_0$ is due to the δN term in Eq. (6.5.10). Its origin lies in the following sequence of events. Each spontaneously emitted photon changes the laser power, which changes the gain (or equivalently the carrier population); this in turn affects the refractive index and consequently the optical phase. The resulting delayed phase fluctuation is affected by relaxation oscillations and leads to satellite peaks at multiples of Ω_R as well as a broadening of the central peak by $1 + \beta_c^2$.

Equation (6.5.52) shows that the line width Δf depends only on β_c, R_{sp}, and P. If we choose β_c equal to 5 and use values of R_{sp} and P from Table 6.2, we find that $\Delta f \cong 70$ MHz at a power level of 1 mW. This provides an order-of-magnitude estimate since both β_c and R_{sp} are likely to vary from device to device. As the laser power increases, the line width decreases as P^{-1} in accordance with Eq. (6.5.52). Such an inverse dependence is observed experimentally at low power levels (< 10 mW) for most semiconductor lasers. However, often the line width is found to saturate to a value in the range 1–10 MHz at a power level above 10 mW.[75,83] Figure 6.15 shows such a line-width-saturation behavior for several 1.55-μm DFB lasers[98] (see Chapter 7 for a discussion of DFB lasers). It also shows that the line width

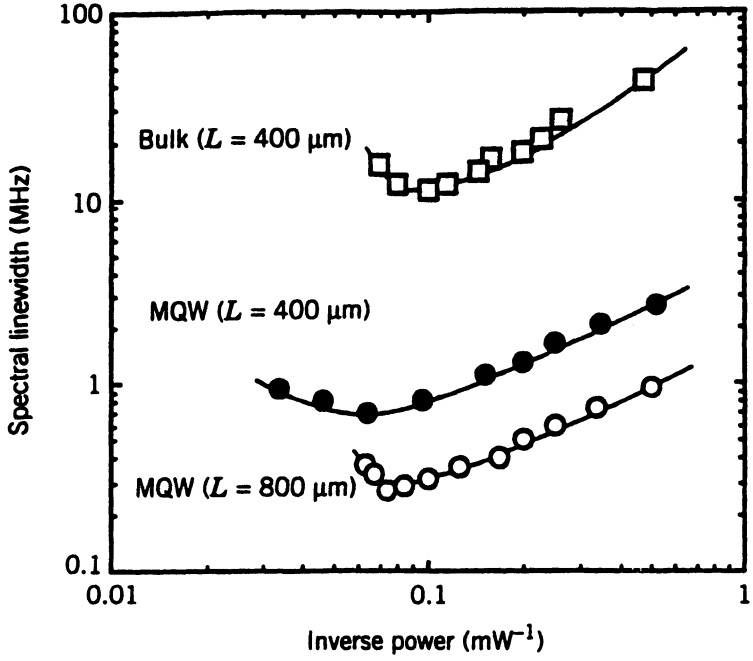


Fig. 6.15 Spectral line width measured as a function of power for two MQW DFB lasers of 400- and 800-μm cavity lengths. The results for a conventional DFB laser with a relatively thick active layer (curve marked bulk) are also shown for comparison. In all cases the line width saturates at a certain power level and begins to rebroaden with a further increase in the output power. (After Ref. 98 © 1991 IEEE)

can be considerably reduced by using a multiquantum-well (MQW) design for the DFB laser (MQW design is discussed in Chapter 9). The reduction is due to a smaller value of the parameter β_c realized by such a design. The line width can also be reduced by increasing the cavity length L, since R_{sp} decreases and P increases at a given output power as L is increased. Although not obvious from Eq. (6.5.52), Δf can be shown to vary as L^{-2} when the length dependence of R_{sp} and P is incorporated. As seen in Fig. 6.15, Δf is reduced by about a factor of 4 when the cavity length is doubled. The 800-μm-long MQW DFB laser is found to exhibit a line width as small as 270 kHz at a power output of 13.5 mW.[98] It should be stressed, however, that the line width of most semiconductor lasers is typically in the range 10–100 MHz.

Figure 6.15 shows that as the laser power increases, the line width not only saturates but begins to rebroaden. Several mechanisms have been invoked to explain such a behavior; a few among them are side-mode

interaction,[84,85] $1/f$-noise,[86] spatial-hole burning,[88,89] nonlinear gain,[90] and current noise.[99] The influence of weak side modes on the line width is of particular importance from the practical standpoint. It can be studied by considering the multimode rate equations (6.5.29) and (6.5.30) together with the phase equation (6.5.10). In the specific case of a single side mode it is possible to calculate the frequency-noise spectrum and the line shape by using Eqs. (6.5.14), (6.5.32), and (6.5.33). It turns out that cross-saturation, governed by the parameters θ_{12} and θ_{21} in Eqs. (6.5.37) and (6.5.38), plays an important role and can lead to considerable spectral broadening even when the MSR exceeds 20 dB. The issue of laser line width is of considerable importance for lightwave systems. A narrow-line-width optical source is required for coherent communication systems. Considerable effort has been directed toward developing semiconductor lasers that not only operate in a single longitudinal mode but whose wavelength can also be tuned over a considerable range while maintaining a narrow line width (~ 1 MHz or less).

6.6 MODULATION RESPONSE

One of the important advantages of semiconductor lasers is that they can be directly modulated; i.e., one can readily obtain short optical pulses useful for optical communications by modulating the device current. The modulation response of semiconductor lasers has been studied from the early days.[3–7] Because of an intrinsic resonance of the device it was found that the response peaks at the relaxation-oscillation frequency Ω_R, and the modulation efficiency sharply drops for modulation frequencies ω_m greater than Ω_R. Experimentally, however, the resonance peak is often less pronounced than the one predicted by the single-mode rate equations.[3,4] Mechanisms such as spontaneous emission, carrier diffusion, and spectral-hole burning have been proposed to account for the suppression of the relaxation-oscillation peak. The initial small-signal analysis[3] has been extended to include the nonlinear effects that become important under large-signal modulation.[100–102] Considerable attention has recently been paid to increase the modulation bandwidth of semiconductor lasers.[103–113]

A characteristic feature of semiconductor lasers is that intensity or amplitude modulation (AM) leads simultaneously to phase or frequency modulation (FM). The interdependence between AM and FM under direct current modulation is governed by the line-width enhancement factor β_c encountered in the previous section. It has its origin in the index change that invariably occurs when the optical gain changes in response to variations in the carrier population. As a result of simultaneously occurring AM and FM, the mode frequency of a directly modulated semiconductor laser shifts

periodically during each modulation cycle.[114] This phenomenon, called frequency chirping,[115–123] has received considerable attention recently as it is often a limiting factor in the performance of 1.55-μm optical communication systems. On the other hand, the same physical phenomenon can be used for direct frequency modulation,[124,125] a useful technique for coherent optical communication systems.

6.6.1 Small-Signal Analysis

A basic understanding of the modulation characteristics can be developed by using the single-mode rate equations (6.5.1)–(6.5.3). Without the noise sources they become

$$\dot{P} = (G - \gamma)P + R_{sp} \tag{6.6.1}$$

$$\dot{N} = I/q - \gamma_e N - GP \tag{6.6.2}$$

$$\dot{\phi} = -(\omega_0 - \omega_{th}) + \tfrac{1}{2}\beta_c(G - \gamma). \tag{6.6.3}$$

Under direct-current modulation, the device current I is time-dependent and consists of two parts:

$$I(t) = I_b + I_m(t). \tag{6.6.4}$$

In the absence of modulation, the laser operates continuously at the bias level I_b. The corresponding steady-state values P, N, and ϕ are obtained by setting all time derivatives to zero in Eqs. (6.6.1)–(6.6.3). The effect of modulation current $I_m(t)$ is to introduce deviations $\delta P(t)$, $\delta N(t)$, and $\delta\phi(t)$, which vary periodically at the modulation frequency ω_m. In the small-signal analysis $I_m(t)$ is assumed to be small enough such that the deviation from the steady state remains small at all times [$\delta P(t) \ll P$, $\delta N(t) \ll N$, and $\delta\phi(t) \ll \phi$]. A criterion for its validity is generally expressed using the depth of modulation defined as

$$m = \frac{(\delta P)_{max}}{P} = \frac{[I_m(t)]_{max}}{I_b - I_{th}}. \tag{6.6.5}$$

The small-signal analysis is valid only if $m \ll 1$. Assuming its validity, the rate equations for δP, δN, and $\delta\phi$ can be linearized using Eqs. (6.6.1)–(6.6.4). We then obtain

$$\delta\dot{P} = -\Gamma_P \delta P + (G_N P)\delta N \tag{6.6.6}$$

$$\delta\dot{N} = -\Gamma_N \delta N - G\delta P + I_m(t)/q \tag{6.6.7}$$

$$\delta\dot{\phi} = \tfrac{1}{2}\beta_c G_N \delta N. \tag{6.6.8}$$

These equations are similar to Eqs. (6.5.8)–(6.5.10) and have been simplified using $\partial R_{sp}/\partial N \ll G_N P$ and $G_P P \ll G$. They are readily solved in the frequency domain using Fourier analysis. The solutions are

$$\delta\tilde{P}(\omega) = \frac{G_N P \tilde{I}_m(\omega)/q}{(\Omega_R + \omega - i\Gamma_R)(\Omega_R - \omega + i\Gamma_R)} \tag{6.6.9}$$

$$\delta\tilde{N}(\omega) = \frac{(\Gamma_P + i\omega)(\tilde{I}_m(\omega)/q)}{(\Omega_R + \omega - i\Gamma_R)(\Omega_R - \omega + i\Gamma_R)} \tag{6.6.10}$$

$$\delta\tilde{\phi}(\omega) = \frac{\beta_c}{2i\omega}[G_N \delta\tilde{N}(\omega)] \tag{6.6.11}$$

where

$$\tilde{I}_m(\omega) = \int_{-\infty}^{\infty} I_m(t) \exp(-i\omega t)\, dt \tag{6.6.12}$$

is the Fourier transform of the modulation current and a similar relation exists for $\delta\tilde{P}$, $\delta\tilde{N}$, and $\delta\tilde{\phi}$. The decay rate Γ_R and the frequency Ω_R of the relaxation oscillations are given by Eqs. (6.4.21) and (6.4.23). Equations (6.6.9)–(6.6.11) show that the modulation current changes the photon and carrier populations inside the active region, which in turn affect the optical phase. However, because of intrinsic laser resonance, the modulation response is frequency-dependent.

6.6.2 Intensity Modulation

The preceding small-signal analysis can be used for an arbitrary form of $I_m(t)$. The results are particularly simple for the case of sinusoidal modulation, which is often used experimentally as well. Using

$$I_m(t) = I_P \sin(\omega_m t) \tag{6.6.13}$$

in Eq. (6.6.12), we obtain

$$\tilde{I}_m(\omega) = -i\pi I_P[\delta(\omega - \omega_m) - \delta(\omega + \omega_m)] \tag{6.6.14}$$

where the modulation frequency $\nu_m = \omega_m/2\pi$ and I_P is the peak value of the modulation current. Using Eq. (6.6.14) in Eq. (6.6.9) and taking the inverse Fourier transform, we obtain the modulated power

$$\delta P(t) = \delta P_0 \sin(\omega_m t + \theta_P) \tag{6.6.15}$$

where

$$\delta P_0 = \frac{G_N P I_P / q}{[(\omega_m^2 - \Omega_R^2 - \Gamma_R^2)^2 + 4\omega_m^2 \Gamma_R^2]^{1/2}} \tag{6.6.16}$$

and

$$\theta_P = \tan^{-1}\left(\frac{2\Gamma_R \omega_m}{\omega_m^2 - \Omega_R^2 - \Gamma_R^2}\right). \tag{6.6.17}$$

We find that the photon population (or equivalently the output power) varies sinusoidally with a phase shift θ_P and with the peak amplitude δP_0. Both δP_0 and θ_P vary with the modulation frequency ω_m. Figure 6.16 shows the frequency dependence of the modulation response for several values of Γ_R and Ω_R. The response is relatively flat when $\omega_m \ll \Omega_R$, peaks in the vicinity of Ω_R, and then drops sharply for $\omega_m > \Omega_R$, indicating that the laser is no longer able to respond at such high modulation speeds. The peak height decreases with an increase in Γ_R. Equation (6.6.16) can be used to obtain the peak height given by

$$\frac{\delta P_0(\omega_m = \Omega_R)}{\delta P_0(\omega_m = 0)} \cong \frac{\Omega_R}{2\Gamma_R} = \frac{\pi \nu_R}{\Gamma_R} \tag{6.6.18}$$

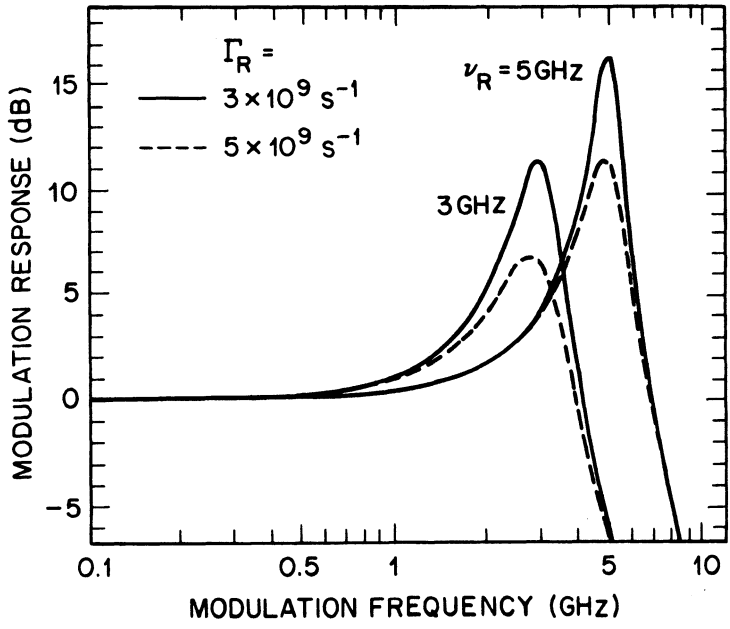


Fig. 6.16 Calculated small-signal modulation response showing the relaxation-oscillation peak at ν_R and its height dependence on the decay rate Γ_R.

where we assumed that $\Gamma_R \ll \Omega_R$ and

$$\nu_R = \Omega_R/2\pi.$$

The peak value of the modulated power in the flat region ($\omega_m \ll \Omega_R$) can also be obtained using Eq. (6.6.16) and is given by the remarkably simple expression

$$\delta P_0 = \tau_P(I_P/q) \tag{6.6.19}$$

where we used Eq. (6.4.23) and $G \cong \tau_P^{-1}$. This equation shows that the modulation response is mainly governed by the photon lifetime τ_P.

From a practical viewpoint the quantity of interest is the modulation bandwidth ν_B, which indicates the frequency range over which the laser responds to the current modulation. It is usually defined as the frequency at which the modulation response has dropped by 3 dB relative to its low-frequency or dc value. Figure 6.16 suggests that under ideal conditions, ν_B exceeds the relaxation-oscillation frequency, and ν_R provides a reasonable estimate of it. In practice, however, ν_B may be significantly lower than ν_R if the electrical parasitics associated with a specific device structure lead to a premature roll-off in the modulation response. This decrease in the response occurs when an increasing fraction of the applied modulation current is bypassed outside the active region with an increase in the modulation frequency. The parasitic roll-off is of concern for some buried-heterostructure lasers employing a current-blocking junction and requires design improvements so as to reduce the parasitic capacitance.

Figure 6.17 shows the effect of electrical parasitics on the modulation response of a 1.3-μm etched-mesa buried-heterostructure (EMBH) laser ($I_{th} \cong 30$ mA) for three bias currents.[108] Even though the relaxation-oscillation frequency increases with an increase in the bias current, the modulation bandwidth $\nu_B \ll \nu_R$ and actually decreases with a higher biasing of the laser. To model the effect of electrical parasitics, circuit models that incorporate all the relevant features of the rate equations have been developed.[101] The continuous curves in Fig. 6.17 were obtained using one such circuit model and appear to explain the data reasonly well.[108] As one may expect, the parasitics are structure-dependent and vary from one laser structure to another. In general, weakly index-guided lasers such as those using the ridge waveguide structure suffer less from the parasitic problem.[107,108] Recently considerable attention has been paid[106–113] to increasing the modulation bandwidth of semiconductor lasers. At room temperature ($T = 20°C$) the modulation bandwidth of 24 GHz under CW operation was obtained in 1990 for a 1.3-μm InGaAsP laser.[113]

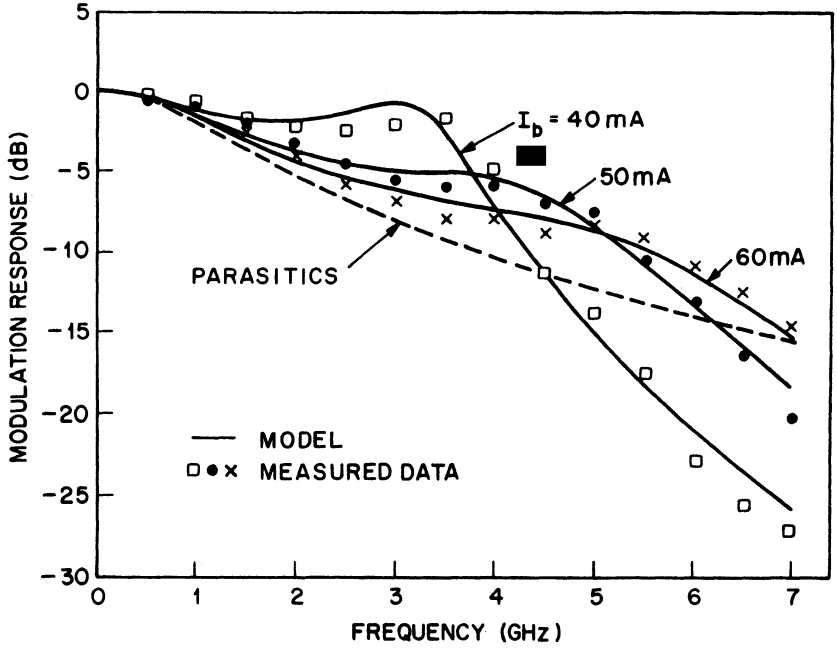


Fig. 6.17 Measured small-signal modulation response of a buried-heterostructure laser at three bias levels. The dashed line shows the parasitic roll-off expected in the absence of relaxation oscillations. (After Ref. 108 © 1984 IEEE)

A comparison of Fig. 6.16 and 6.17 shows that, apart from the parasitic effects, the height of the observed relaxation-oscillation peak is considerably suppressed in Fig. 6.17. This suggests that the decay rate Γ_R is larger compared to the value estimated in Table 6.2 and that it increases with an increase in the bias power. If we use Eqs. (6.4.16) and (6.4.17) to obtain $\Gamma_R = (\Gamma_N + \Gamma_P)/2$, we find that Γ_N is carrier-dominated and varies little with power. By contrast, the dominant contribution to Γ_P comes from the power-dependent gain-suppression term $G_P P$. Several mechanisms such as carrier diffusion, carrier heating, and spectral-hole burning may contribute to gain suppression and can be accounted for by increasing the absolute value of G_P.

Sinusoidal current modulation, although often used, does not cover the practical case where rectangular pulses of duration T are used to transmit the information at a bit rate of $B = T^{-1}$. If we allow for the finite rise and fall times, a realistic description of a single pulse is provided by

$$I_m(t) = I_P\big(H(t)[1 - \exp(-t/\tau)] - H(t-\tau)\{1 - \exp[-(t-T)/\tau]\}\big) \qquad (6.6.20)$$

where the Heaviside step function $H(t) = 1$ for $t \geq 0$, $H(t) = 0$ for $t < 0$, and I_P is the peak value of the modulation current. If we define the rise time τ_r as the time during which the current changes from 10 to 90% of its peak value, τ_r is related to the exponential decay time τ by $\tau_r = (\ln 9)\tau \cong 2.2\tau$. For simplicity we have assumed that the rise and fall times are equal. A typical value of τ_r is ~ 100 ps. Using Eq. (6.6.20) and (6.6.12), the Fourier components of the modulation currents are

$$\tilde{I}_m(\omega) = I_P[1 - \exp(-i\omega\tau)]\left(\frac{1}{i\omega} - \frac{1}{i\omega + 1/\tau}\right) \tag{6.6.21}$$

The term proportional to $(i\omega + 1/\tau)^{-1}$ does not contribute when $\tau = 0$ and is responsible for the effects arising from the finite rise and fall times associated with the modulation circuitry. The modulation response is readily obtained by substituting $\tilde{I}_m(\omega)$ from Eq. (6.6.21) into Eq. (6.6.9) and taking the inverse Fourier transform. The frequency integration can be carried out using the method of contour integration. The result is[126]

$$\delta P(t) = \frac{G_N P I_P}{q}\{H(t)[V(t, 0) - V(t, \tau^{-1})] - H(t - T)[V(t - T, 0) - V(t - T, \tau^{-1})]\} \tag{6.6.22}$$

where

$$V(t, \eta) = \frac{\exp(-\eta t)}{[\Omega_R^2 + (\Gamma_R - \eta)^2]} + \frac{\exp(-\Gamma t)\sin(\Omega_R t + \theta_P)}{\Omega_R[\Omega_R^2 + (\Gamma_R - \eta)^2]^{1/2}} \tag{6.6.23}$$

and the phase shift θ_P is given by

$$\theta_P = -\tan^{-1}\left(\frac{\Gamma - \eta}{\Omega_R}\right). \tag{6.6.24}$$

Clearly the response to the square-wave modulation has qualitatively different features when compared with the case of sinusoidal modulation. The physical origin of the four terms in Eq. (6.6.22) is self-evident. The terms proportional to $H(t)$ and the terms proportional to $H(t - T)$ correspond respectively to the leading and trailing edges of the current pulse. In each case the term in the form of $V(t, \tau^{-1})$ represents the effect of a finite rise time and vanishes when $\tau = 0$. Equation (6.6.23) shows that under square-wave modulation the optical pulse exhibits damped oscillations during turn-on and turn-off whose frequency and damping time are governed by relaxation

oscillations. Furthermore, these oscillations are also affected by the current-pulse rise or fall time. In comparing the results of sinusoidal and square-wave modulations, it should be noted that for an on-off sequence of pulses of duration T, the effective modulation frequency $\nu_m = (2T)^{-1}$.

Equations (6.6.16) and (6.6.22) show that the modulation response of a semiconductor laser is strongly affected by the frequency Ω_R and the damping rate Γ_R of relaxation oscillations. The derivation of Ω_R and Γ_R in Sec. 6.4.3 is based on a truncated Taylor-series expansion of $G(N, P)$ [see Eq. (6.4.18)]. For a more accurate description of modulation response at high powers one should use the exact functional form of $G(N, P)$. A simple approach assumes, in analogy with a two-level atomic system, that $G(N, P)$ is given by[110]

$$G(N, P) = \frac{G_N(N - N_0)}{1 + P/P_s} \tag{6.6.25}$$

where P_s is the saturation photon number. In general, semiconductor lasers cannot be modeled as a two-level system. For a single-mode semiconductor laser the density-matrix equations can be solved approximately to yield the following expression:[120]

$$G(N, P) = \frac{G_N(N - N_0)}{\sqrt{1 + P/P_s}}. \tag{6.6.26}$$

An advantage of the density-matrix approach is that it provides an expression that relates P_s to the fundamental material parameters through the relation[112]

$$P_s = \frac{\epsilon_0 \bar{\mu} \mu_g V \hbar}{\Gamma \omega_0 d^2 \tau_{in}(\tau_c + \tau_v)} \tag{6.6.27}$$

where ϵ_0 is the vacuum permittivity, d is the dipole moment, τ_{in} is the dipole relaxation time, and τ_c and τ_v are the carrier scattering time for the conduction and valence bands respectively ($\tau_{in} \cong 0.1$ ps, $\tau_c = 0.2$–0.3 ps, $\tau_v = 0.05$–0.1 ps).

One can perform the small-signal analysis of Sec. 6.6.1 by using $G(N, P)$ from Eq. (6.6.26) in Eqs. (6.6.1) and (6.6.2). The modulation response for the case of sinusoidal current modulation is still given by an equation similar to Eq. (6.6.16) but with the following expressions for Ω_R and Γ_R:[112]

$$\Omega_R = \left[G_L G_N P \frac{1 + p/2}{(1 + p)^2} - \tfrac{1}{4}(\Gamma_P - \Gamma_N)^2 \right]^{1/2} \tag{6.6.28}$$

and

$$\Gamma_R = (\Gamma_N + \Gamma_P)/2 \tag{6.6.29}$$

$$\Gamma_N = \gamma_e + N\frac{\partial \gamma_e}{\partial N} + \frac{G_N P}{(1+p)^{1/2}} \tag{6.6.30}$$

$$\Gamma_P = \frac{R_{sp}}{P} + \frac{G_L}{2}\frac{P}{(1+p)^{3/2}} \tag{6.6.31}$$

where $G_L = G_N(N - N_0)$ is the linear part of the gain and p is a dimensionless parameter defined as $p = P/P_s$. At low operating powers such that $p \ll 1$, Ω_R varies as $\sqrt{P}$, and Eq. (6.6.28) reduces to the result given in Eq. (6.4.23). At high powers, Ω_R begins to saturate. The damping rate Γ_R peaks near $p = 2$ and then decreases with a further increase in p.

The quantity of interest from a practical standpoint is not the relaxation-oscillation frequency but the 3-dB bandwidth ν_B, defined as the modulation frequency at which the modulation response drops by a factor of 2 from its zero-frequency value. By using Eq. (6.6.16), ν_B is found to be related to the relaxation-oscillation parameters by the relation

$$\nu_B = \frac{1}{2\pi}\{\Omega_R^2 - \Gamma_R^2 + 2[\Omega_R^2(\Omega_R^2 + \Gamma_R^2) + \Gamma_R^4]^{1/2}\}^{1/2}. \tag{6.6.32}$$

Figure 6.18 shows the variation of ν_B as a function of p for several values of $G_N P_s$ with $\tau_p = 1$ ps. The modulation bandwidth ν_B saturates at high powers, and the limiting value depends on the value of $G_N P_s$. The limiting value is obtained from (6.6.32) and is given by

$$\nu_{B,\max} = (3G_N P_s/8\pi^2\tau_p)^{1/2} \tag{6.6.33}$$

where $\Gamma_R \ll \Omega_R$ was assumed. Equation (6.6.33) can be written in terms of the device and material parameters by using $G_N = \Gamma v_g a/V$ and P_s from Eq. (6.6.27). The result is

$$\nu_{B,\max} = \sqrt{\frac{3\epsilon_0 \hbar \bar{\mu} a}{16\pi^3 d^2 \tau_p \tau_{in}(\tau_c + \tau_v)}}. \tag{6.6.34}$$

The most crucial parameter is the differential gain coefficient a that depends on the density of states and can be enhanced by using a quantum-well structure. If we use $a = 2 \times 10^{-16}$ cm^2 and typical parameter values for 1.55-μm InGaAsP laser, $\nu_{B,\max}$ is found to be about 32 GHz. For a quantum-well laser a is enhanced by a factor of nearly 2, and $\nu_{B,\max}$ should

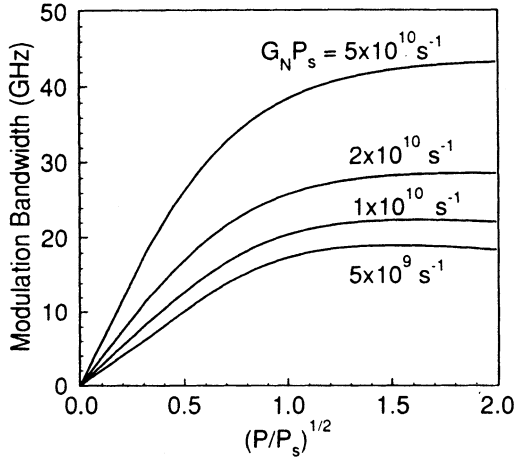


Fig. 6.18 Modulation bandwidth plotted as a function of $(P/P_s)^{1/2}$ for several values of the product $G_N P_s$. Saturation of the modulation bandwidth is due to the nonlinear gain.

increase by $\sqrt{2}$ if other parameters remain the same. The main conclusion is that the nonlinear gain provides a fundamental mechanism that limits the modulation response of semiconductor lasers.

6.6.3 Frequency Chirping

We have seen that modulation of the device current leads to laser-power modulation that is useful for transmitting information as a sequence of optical pulses. However, modulation of the current also affects the optical frequency. This is seen clearly in Eq. (6.6.8), which emphasizes that carriers influence the optical phase. If the steady-state frequency $\nu_0 = \omega_0/2\pi$, this frequency shifts by an amount

$$\delta\nu(t) = \delta\dot{\phi}(t)/2\pi \tag{6.6.35}$$

during modulation. The phenomenon of the dynamic shift of the lasing frequency is referred to as *frequency chirping* [or wavelength chirping, since $\delta\lambda = (-\lambda^2/c)\delta\nu$]. An expression for the chirp $\delta\nu(t)$ can be obtained using the Fourier relation

$$\delta\dot{\phi}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} i\omega\tilde{\phi}(\omega)\exp(i\omega t)\,d\omega. \tag{6.6.36}$$

Using Eqs. (6.6.11), (6.6.35), and (6.6.36), we obtain

$$\delta\nu(t) = \frac{\beta_c G_N}{8\pi^2 q} \int_{-\infty}^{\infty} \frac{[\Gamma_P + i\omega]\tilde{I}_m(\omega) \exp(i\omega t)\, d\omega}{(\Omega_R + \omega - i\Gamma_R)(\Omega_R - \omega + i\Gamma_R)}. \tag{6.6.37}$$

We note that the chirp is directly proportional to the line-width enhancement factor β_c and has its origin in the carrier-induced index change that accompanies any gain change in semiconductor lasers.

Equation (6.6.37) can be used to obtain the frequency chirp $\delta\nu(t)$ for an arbitrary time dependence of the modulation current $I_m(t)$. For the case of sinusoidal modulation, $I_m(t)$ is given by Eq. (6.6.13). The use of Eq. (6.6.14) in Eq. (6.6.37) then leads to

$$\delta\nu(t) = \delta\nu_0 \sin(\omega_m t + \theta_c) \tag{6.6.38}$$

where

$$\delta\nu_0 = \frac{\beta_c I_P G_N}{4\pi q} \left(\frac{\omega_m^2 + \Gamma_P^2}{(\omega_m^2 - \Omega_R^2 - \Gamma_R^2)^2 + (2\omega_m \Gamma_R)^2} \right)^{1/2} \tag{6.6.39}$$

and

$$\theta_c = \tan^{-1}\left(\frac{\omega_m}{\Gamma_P}\right) + \tan^{-1}\left(\frac{2\Gamma_R \omega_m}{\omega_m^2 - \Omega_R^2 - \Gamma_R^2}\right). \tag{6.6.40}$$

Equation (6.6.38) shows that the mode frequency shifts periodically in response to current modulation and that $\delta\nu_0$ is the maximum shift. Similar to the power response, $\delta\nu_0$ peaks at the relaxation-oscillation frequency. If we use the parameter values given in Tables 6.1 and 6.2 for a typical buried-heterostructure laser, $\delta\nu_0 \cong 2$ GHz/mA when $\omega_m = \Omega_R/2$. At a given ω_m, however, the chirp decreases with an increase in the bias level. The dependence of the frequency chirp on various parameters under sinusoidal modulation has been investigated, and the experimental results in the small-signal regime are in agreement with the rate-equation analysis.[114] The case of square-wave modulation can be treated using Eq. (6.6.21) in Eq. (6.6.37) and is useful to understand the effects of the shape of the current pulse on frequency chirping.[120,126]

The phenomenon of frequency chirping under current modulation has attracted considerable attention.[115–123] It manifests as a broadening of the CW line width associated with a single longitudinal mode when the time-averaged optical spectrum is recorded using a spectrometer. Figure 6.19 shows the experimentally observed spectra of a 1.3-μm InGaAsP laser under sinusoidal modulation, when $\nu_m = 100$ MHz, for several values of the

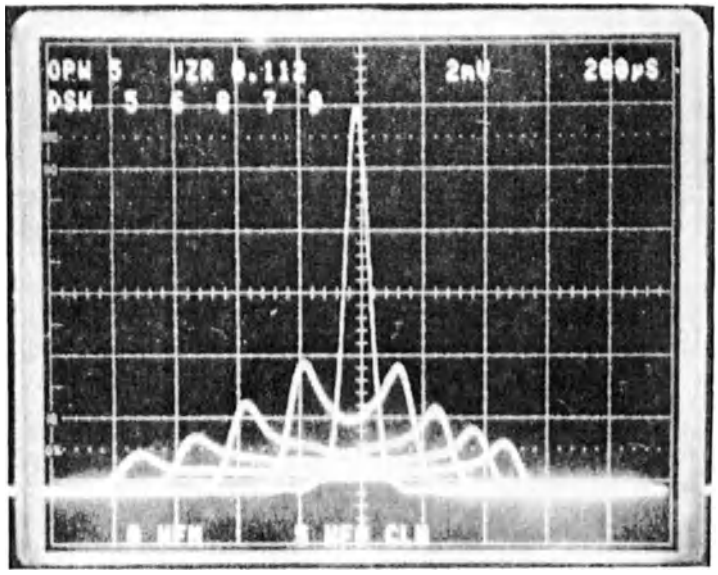


Fig. 6.19 Time-averaged power spectra of a 1.3-μm InGaAsP laser under sinusoidal modulation at 100 MHz. Spectrum broadens with an increase in the modulation current due to frequency chirping. The horizontal scale is 0.5 Å per division. (After Ref. 116)

modulation current I_p. In the absence of modulation (the narrowest spectrum in Fig. 6.19) the observed line width is limited by the spectrometer resolution ($\sim$0.02 nm), and the actual CW line width is typically less than $\sim$100 MHz. Under modulation, the spectrum develops an asymmetric double-peaked profile whose width is in the gigahertz range and increases with an increase in the amplitude I_p of the modulation current. The calculated chirp $\delta\nu_0$ is related to the width of the dynamically broadened optical spectrum; according to Eq. (6.6.39), $\delta\nu_0$ should vary linearly with I_p. Figure 6.20 shows the measured chirp as a function of I_p for several laser structures and verifies the predicted linear increase.

We note from Fig. 6.20 that the chirp depends on the device structure. To account for the structure-dependent variation, consider the limit where $\omega_m \ll \Omega_R$ in Eq. (6.6.39). If we use $\Gamma_R \ll \Omega_R$, and $\Omega_R^2 \cong GG_NP$ from Eq. (6.4.23), we obtain the simplified expression

$$\delta\nu_0 = \frac{\beta_c I_P}{4\pi qGP}(\omega_m^2 + \Gamma_P^2)^{1/2}. \tag{6.6.41}$$

The largest chirp in Fig. 6.20 occurs for a gain-guided device. This can be

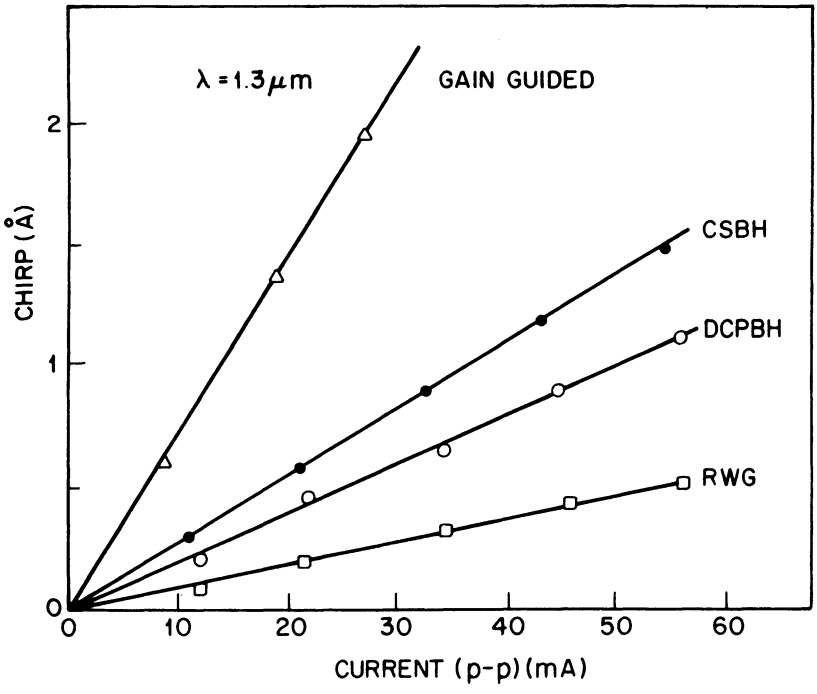


Fig. 6.20 Measured chirp as a function of the peak-to-peak value of the modulation current for several laser structures (discussed in Chapter 5) at 1.3 μm. (After Ref. 116)

understood by noting that R_{sp}, and hence Γ_P [see Eq. (6.4.16)], is generally larger for a gain-guided laser. Using $G \cong \tau_P^{-1}$, the chirp at a given bias power P depends mainly on two device parameters β_c and τ_P. From Eq. (6.2.16) the photon lifetime τ_P depends on the internal loss that may vary somewhat with structure. In particular, τ_P may be shorter for ridge waveguide lasers if α_{int} is larger because of diffraction losses. The line-width enhancement factor β_c may also vary from device to device because it is strongly wavelength-dependent. Such variations of β_c and τ_P can account for the measured chirp differences in index-guided lasers. It should be noted, however, that in Fig. 6.20 the comparison is being made at low modulation frequencies ($\nu_m = 100$ MHz) that are well below the relaxation-oscillation frequency. At high modulation frequencies in the gigahertz range, weakly index-guided lasers such as the ridge waveguide laser generally show stronger chirping simply because they have lower values of Ω_R at the same bias power.

The frequency chirp also varies with the device wavelength. Figure 6.20 shows the data obtained for 1.3-μm InGaAsP lasers. Similar measurements

for 1.55-μm lasers show that the chirp is larger typically by about a factor of 2 when the comparison is made at a constant value of the modulation current and the bias power. This may be attributed in part to the line-width enhancement factor β_c, which may be larger at 1.55 μm. The frequency chirping of 1.55-μm lasers at high bit rates is of major concern since, in combination with fiber dispersion, chirping degrades system performance considerably.[115]

Several methods[119,120] have been proposed to reduce the extent of frequency chirping. One scheme uses injection locking of the modulating laser by a CW semiconductor laser.[119] The external injection provides a mode-pulling mechanism, reducing the dynamic frequency shift or the chirp of the modulating laser. Another scheme requires a careful tailoring of the current-pulse shape.[120] A small current step in the leading edge of the current pulse is applied to reduce the amplitude of relaxation oscillations, thus reducing the chirp significantly.[123] These chirp-reduction schemes are particularly useful for single-frequency lasers (see Chapter 7).

We have seen that the chirp arises from periodic oscillations of the carrier and photon populations in response to modulation of the current. Clearly the power spectrum would be different at different times during one modulation cycle. Time-resolved power spectra show[115] that the laser frequency briefly shifts upwards during turn-on (leading edge of the current pulse) and downwards during turn-off (trailing edge of the current pulse). A theoretical description of the time-resolved or the time-averaged spectra such as those shown in Fig. 6.19 requires a generalization of the phase noise discussed in Sec. 6.5.3 to include the effect of current modulation.[127,128] Such a generalization is readily carried out in the small-signal regime. The time-averaged power spectrum can be calculated using Eq. (6.5.43) and includes the effect of frequency chirping as well as the effect of phase diffusion due to spontaneous emission.[128] Figure 6.21 shows the calculated power spectra for four modulation currents at a modulation frequency of 1 GHz. The central mode corresponds to the single longitudinal mode supported by the laser in the absence of modulation. The effect of modulation is to generate sidebands at multiples of the modulation frequency on both sides of the optical line. These are the well-known FM sidebands[125] and arise since the current modulation leads to both AM and FM. When the power spectrum is recorded using a spectrometer (see Fig. 6.19), the FM sidebands are often unresolved due to a limited resolution and the observed spectrum corresponds to the envelope of the calculated power spectra. Figure 6.21 shows that the envelope is asymmetric and double-peaked, which agrees with experimental observations. The asymmetry results from the simultaneously occurring AM and FM under current modulation. The chirp-induced fine structure of the FM sidebands can be observed using an FP

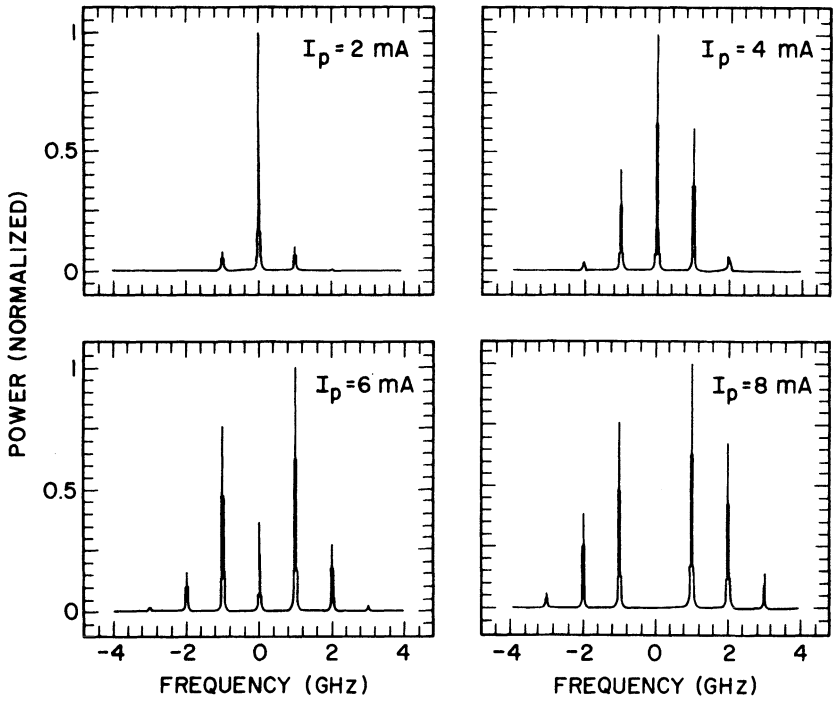


Fig. 6.21 Calculated power spectra at a sinusoidal modulation of 1 GHz showing the generation of FM sidebands with an increase in the modulation current.

interferometer[125,129] or by increasing the modulation frequency so that their separation exceeds the spectrometer resolution.[130]

6.6.4 Large-Signal Modulation

The small-signal analysis of the modulation response, although useful for predicting the parameter dependence, is not necessarily valid under practical conditions. For use in optical communication systems, semiconductor lasers are generally biased below or close to threshold and the modulation depth m defined by Eq. (6.6.5) does not satisfy the validity criterion that $m \ll 1$. The nonlinear effects then play an important role, and a numerical analysis of the rate equtions becomes a necessity. In this section we briefly describe the new qualitative features that may arise under large-signal modulation.

Numerical simulations indicate that, when the modulation depth increases, the power does not vary sinusoidally under sinusoidal modulation. Figure 6.22 shows[131] the optical pulses for several modulation depths obtained for

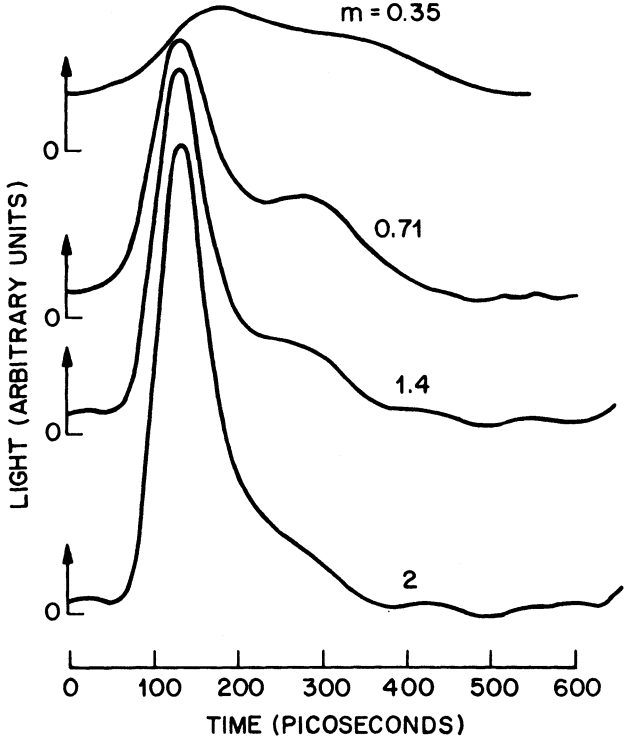


Fig. 6.22 Large-signal response of a buried-heterostructure 1.3-μm laser sinusoidally modulating at 1.73 GHz for several depths of modulation. (After Ref. 131 © 1985 IEEE)

a 1.3-μm buried-heterostructure laser modulated sinusoidally at 1.73 GHz. The power response is not sinusoidal even at $m = 0.35$. Under large-signal modulation the optical response takes the form of a narrow pulse followed by "ringing," which is related to damped relaxation oscillations. The main pulse corresponds to the first peak of the relaxation oscillations. As the modulation depth increases, the subsequent peaks are strongly suppressed and the output takes the form of a short optical pulse. In this way large-signal modulation can be used to generate short optical pulses. This technique is referred to as gain switching and is discussed in Sec. 6.6.5.

For lightwave-system applications semiconductor lasers are used to convert a digital electrical bit stream (containing "1" and "0" bits in random order) into an optical replica with as little distortion as possible. The laser is biased close to threshold so as not to turn off the laser completely during "0" bits. Figure 6.23 shows the optical pulse obtained by integrating the rate equations numerically when a 500-ps rectangular current pulse is applied

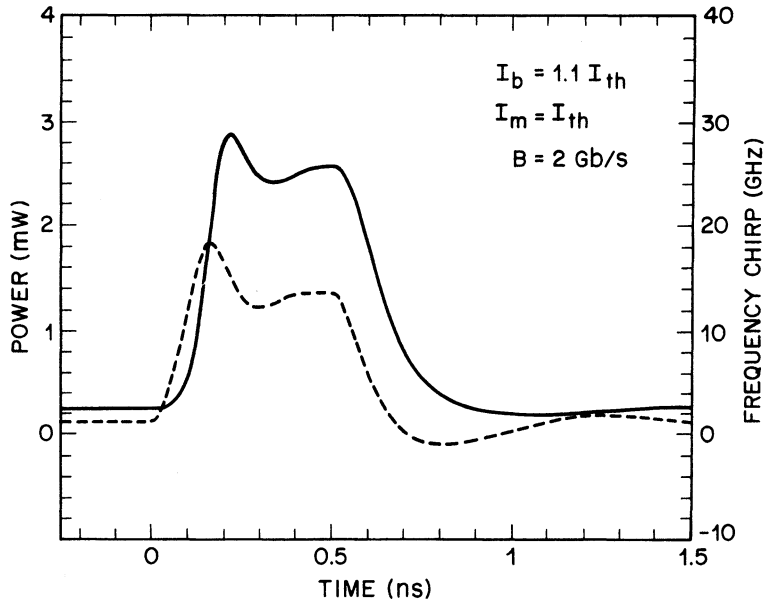


Fig. 6.23 Calculated shape of the optical pulse when a semiconductor laser is modulated at 2 Gb/s by using rectangular 500-ps current pulses. Frequency chirp imposed on the pulse is shown by the dashed curve.

at $t = 0$ and the laser is biased 10% above threshold. The optical pulse has finite rise and fall times and exhibits features related to relaxation oscillations. Even though the optical pulse is far from being an exact replica of the electrical pulse, it can be used to transmit information over optical fibers.

Similar to the case of small-signal modulation, power variations in the light output are accompanied by phase variations, resulting in considerable frequency chirping.[132–136] The dashed curve in Fig. 6.23 shows the frequency chirp imposed on the pulse through current modulation. The laser frequency shifts toward the high-frequency (blue) side near the leading edge, and then shifts back toward the red side near the trailing edge. Such a frequency chirp is harmful for optical communication systems since it leads to considerable broadening of optical pulses during their propagation in the anomalous-dispersion regime of optical fibers. The physical origin of chirp is related to the amplitude-phase coupling in semiconductor lasers governed by the line-width enhancement factor β_c. Semiconductor lasers with low values of β_c are desirable for lightwave-system applications. This parameter is generally lower for multiquantum-well (MQW) lasers. It is reduced further in strained MQW semiconductor lasers.

Considerable attention has been paid to obtaining an analytic expression for the frequency chirp $\delta\nu_0$ that is valid for large-signal modulation.[132–135] Using Eqs. (6.6.16) and (6.6.39), the chirp for the case of small-signal modulation can be related[137] to the modulation depth ($m = \delta P_0/P$) by

$$\delta\nu_0 = \tfrac{1}{2}\beta_c \nu_m m \tag{6.6.42}$$

where we assumed that $\omega_m = 2\pi\nu_m \gg \Gamma_p$. Equation (6.6.42) shows that the chirp varies linearly with the modulation depth m as long as $m \ll 1$. This expression is not applicable when $m \gtrsim 1$. One way to obtain the instantaneous frequency shift during large-signal modulation is to use the phase equation (6.6.3) with $\omega_0 = \omega_{th}$ and note that

$$\delta\nu(t) = \frac{\dot{\phi}(t)}{2\pi} = \frac{\beta_c}{4\pi}(G - \gamma). \tag{6.6.43}$$

If we now use Eq. (6.6.1) to eliminate $G - \gamma$ and add the contribution of the gain nonlinearities, we obtain

$$\delta\nu(t) = \frac{\beta_c}{4\pi}\left(\frac{1}{P}\frac{\partial P}{\partial t} - \frac{R_{sp}}{P} + G_P P\right) \tag{6.6.44}$$

where the last two terms correspond to the dc shift (related to the frequency offset between the on and off power levels during modulation) arising from spontaneous emission and gain suppression. The first term corresponds to the dynamic frequency shift or the transient chirp. Both the transient and dc chirps have been observed experimentally through time-resolved spectral measurements.[115] Since $\delta\nu_0$ represents the maximum transient frequency shift, for the case of large-signal modulation it is given by

$$\delta\nu_0 = \frac{\beta_c}{4\pi}\left(\frac{1}{P}\frac{\partial P}{\partial t}\right)_{\max}. \tag{6.6.45}$$

Equation (6.6.45) can be used to estimate the frequency chirp $\delta\nu_0$ for specific pulse profiles. Even though the exact value depends on details of the pulse shape,[134] an order-of-magnitude estimate of $\delta\nu_0$ can be obtained using a Gaussian pulse shape[135] $\exp(-t^2/T^2)$ and gives $\delta\nu_0 \cong \beta_c/2\pi T$. In contrast to Eq. (6.6.42), $\delta\nu_0$ becomes independent of the modulation depth under large-signal modulation ($m \gtrsim 1$). It has been suggested[118] that the ratio of the frequency chirp to the modulated power is a useful measure of the

modulation response of semiconductor lasers. This ratio varies linearly with the modulation frequency.

6.6.5 Ultrashort Pulse Generation

Direct current modulation of semiconductor lasers generally produces optical pulses whose width is about the same as the applied current pulse. This technique can produce optical pulses as short as 100 ps by modulating the laser at 10 Gb/s, but it becomes increasingly difficult to obtain shorter pulses simply because current pulses shorter than 100 ps are not readily available. Many applications of semiconductor lasers require ultrashort optical pulses of width a few picoseconds or shorter. An example is provided by the fifth generation of fiber-optic communication systems[138] based on the concept of optical solitons.[139] These systems require semiconductor lasers capable of emitting optical pulses of duration 10–20 ps at high repetition rates in the range 2–10 GHz. The techniques of gain switching and mode locking have been used to generate ultrashort optical pulses. This section briefly reviews these two techniques.

Gain Switching. Gain switching, as the name suggests, consists of switching the optical gain rapidly from a low value to a high value so that the semiconductor laser, biased well below threshold, suddenly finds itself well above threshold. After an initial turn-on delay (see Sec. 6.4.2) the optical field builds up rapidly during its first relaxation-oscillation cycle, and drives the laser below threshold through the process of gain saturation. If duration of the current pulse is chosen such that the current is turned off before formation of the second relaxation-oscillation peak, the output consists of a short optical pulse of 10–20 ps duration even though the current pulse may be longer than 100 ps. By applying the current pulse repetitively, a picosecond pulse train can be generated at a repetition rate as high as a few gigahertz. Gain switching was studied extensively during the 1980s.[140–157] Starting in 1988 it was used to demonstrate soliton propagation in optical fibers.[158–163]

The main qualitative features of gain switching can be understood by solving the single-mode rate equations numerically. Figure 6.24 shows the pulse shape and the frequency chirp for the case in which the laser is biased below threshold ($I_b/I_{th} = 0.8$) and switched 3 times above threshold by applying a current pulse of 200-ps duration. The effect of finite rise and fall times associated with the current pulse is included by using a super-Gaussian pulse shape such that

$$I(t) = I_b + I_p \exp[-(t/T_p)^{2T_p/T_r}] \tag{6.6.46}$$

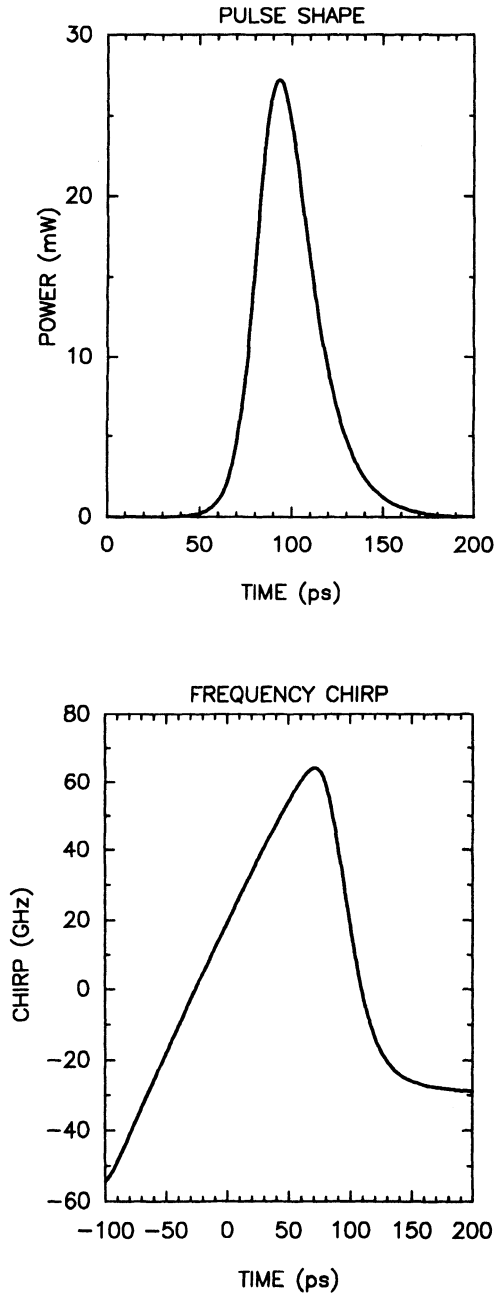


Fig. 6.24 Pulse shape (top) and frequency chirp for a semiconductor laser biased below threshold ($I_b/I_{th} = 0.8$) and gain-switched at 3 times threshold by applying 200-ps current pulses.

where the rise time T_r was chosen to be 10% of the pulse duration T_p. The optical pulse is about 30-ps wide under such conditions. Numerical simulations show that the pulse width depends sensitively on many operating parameters such that the bias level I_b, the peak value I_p, current-pulse width T_p, the rise time T_r, and the photon lifetime τ_p. The last parameter changes with the cavity length and causes the pulse width to depend on the laser length. Pulse widths shorter than 10 ps can be realized by optimizing the operating parameters.[153] The experimental results support the rate-equation model well. Optical pulses as short as 15 ps were generated in 1981 by using a 60-μm-long semiconductor laser.[142] In a later experiment,[151] the pulse width was reduced to 6.7 ps for a 100-μm laser by optimizing other device parameters. The use of a multiquantum-well (MQW) active region helps to reduce the pulse width even further. In a GaAs MQW laser of 155-μm length the pulse width was reduced to below 1.8 ps.[150] The physical reason behind pulse shortening for MQW lasers is related to the enhanced value of the differential gain constant a that leads to relatively large values of the parameter G_N in Eq. (6.4.26).

A shortcoming of the gain-switching method is that optical pulses are considerably chirped because of the amplitude-phase coupling governed by the line-width enhancement factor. This shortcoming, however, can be used to advantage if the pulse is propagated through an optical fiber of optimized length chosen such that it experiences normal (positive) group-velocity dispersion. The frequency chirp imposed on the pulse during gain switching manifests through a frequency shift of the leading and trailing edges toward blue and red sides, respectively (see Fig. 6.24). Since red-shifted frequency components travel faster than blue-shifted components in the normal-dispersion regime, the trailing edge is advanced and the leading edge is retarded during propagation inside the fiber, and the output pulse is compressed.[139] With suitable optimization, this technique can compress the pulse by a factor of 4–5. Indeed, pulses shorter than 4 ps have been obtained by compressing 15–20-ps pulses commonly obtained by the technique of gain switching. Compressed optical pulses are also nearly chirp-free and can be used in the design of soliton communication systems operating near 1.5 μm.[158–163]

The gain-switching technique in combination with a suitable pulse-compression technique can be used to produce pulses as short as 1–2 ps. In one experiment[157] 17-ps optical pulses were obtained by gain switching a 1.3-μm semiconductor laser. They were compressed to about 2.5 ps by launching them in a 500-m-long dispersion-shifted fiber exhibiting normal dispersion of 32 ps/(km·nm). The compressed pulse was amplified by using semiconductor laser amplifiers which broadened it to about 5.5 ps. However, the amplified pulse could be compressed to 1.26 ps by passing it through a

1.1-km fiber exhibiting anomalous dispersion at 1.3 μm. The physical mechanism behind the second stage of pulse compression is self-phase modulation[164] occurring inside the amplifier as a result of gain saturation.[165] Clearly gain switching is a useful technique to generate picosecond pulses from a semiconductor laser.

Mode Locking. The mode-locking technique is well known in laser literature[166] and is commonly used to obtain ultrashort optical pulses (as short as 11 fs) by using solid-state and dye lasers. The basic idea behind mode locking is quite simple.[166] A large number of longitudinal modes can be excited simultaneously in lasers with a broad gain spectrum. In the absence of a definite phase relationship among different longitudinal modes, the laser operates continuously simply because the interference effects due to mode beating are washed out. However, if the phase difference between neighboring modes can be forced to lock to a fixed value, beating among the longitudinal modes forces the laser to emit short pulses at a repetition rate equal to the mode spacing (one pulse on every round trip inside the laser cavity). The gain bandwidth sets the ultimate limit on the pulse width since the pulse spectrum cannot be broader than the gain spectrum. In practice, pulse width depends not only on the gain bandwidth, but also on many parameters associated with the specific mode-locking technique. Two commonly used mode-locking techniques are known as active and passive depending on whether an external signal at the pulse-repetition rate is applied or not (consult Ref. 166 for details on mode locking). Since the gain bandwidth of semiconductor lasers is quite large (about 10 THz) pulses shorter than 100 fs are possible in principle. For this reason, mode locking in semiconductor lasers has been extensively studied by using active, passive, or hybrid techniques.[167–190] This section reviews briefly the progress made in this direction.

Active mode-locking of semiconductor lasers requires modulation of the applied current at a frequency equal to the longitudinal-mode spacing. However, because of a relatively small cavity length (0.2–0.5 mm) the required frequency ($\approx$100 GHz) is so large that current sources are not readily available. A commonly used technique to reduce the modulation frequency is to extend the laser cavity by placing an external mirror or a grating next to an antireflection-coated laser facet (see Chapter 8). Such external-cavity semiconductor lasers can be actively mode-locked relatively easily and provide pulse trains at repetition rates of a few GHz depending on the cavity length. The pulse width can vary over a wide range with typical values in the range 10–50 ps. With a proper design, pulses shorter than 1 ps can be obtained, but they are often accompanied by one or more trailing pulses whose spacing is governed by the round-trip time in the semi-

conductor-laser cavity.[177] Their presence is related to the residual reflectivity of the laser facet facing the external cavity; reflectivities as small as 10^{-6} are needed to eliminate these trailing satellite pulses.[188]

Passive mode locking generally requires a fast saturable absorber.[166] Subpicosecond optical pulses by this technique were obtained as early as 1981.[172] An MQW device whose carrier lifetime is artificially shortened through creation of defects (e.g. by proton bombardment) is sometimes used as a saturable absorber.[176] The use of an MQW saturable absorber in the external-cavity configuration together with a grating-pair pulse compressor has provided 0.46-ps mode-locked pulses at 300-MHz repetition rate.[179] The laser current was also modulated at the pulse-repetition rate. Such a scheme is often referred to as hybrid mode locking since it makes use of both active and passive techniques. For lightwave-system applications it is desirable to use a monolithically integrated version of the external-cavity semiconductor laser. Considerable progress has been made in this direction.[181–184] Mode-locked pulses as short as 1.2 ps were obtained by using hybrid mode locking of a monolithic device.[181] The colliding-pulse mode-locking (CPM) technique, similar to that used for dye lasers, has also been implemented in a monolithic implementation of the ring-cavity geometry.[182] It produced in 1990 1.4-ps pulses at a repetition rate of 32.6 GHz. Further refinements of this technique have generated subpicosecond pulses at a repetition rate as high as 300 GHz by using passive mode locking.[189] Mode-locked semiconductor lasers are commonly used as a pulse source for soliton communication systems.[138] They are likely to find many other applications as a compact source of ultrashort optical pulses at a repetition rate that is much higher than that achieved in solid-state and dye lasers.

6.7 EXTERNAL OPTICAL FEEDBACK

So far in this chapter we have considered the emission characteristics of a semiconductor laser operating in isolation. In practice, however, a small portion of the emitted light is inevitably fed back into the laser cavity owing to parasitic reflections that may occur at a surface outside the cavity. In optical communication systems, such unintentional reflections may occur at the near end or the far end of the fiber link. It is often observed that even a relatively small amount of external optical feedback can significantly affect the performance of a semiconductor laser. Considerable attention has therefore been given to investigating the effects of external optical feedback on the static, dynamic, spectral, noise, and modulation characteristics of the semiconductor laser.[193–238] In this section we briefly describe these effects using the single-mode rate equations modified to account for the weak external feedback.

6.7.1 Modified Rate Equations

For a theoretical analysis of the optical feedback effects on laser performance, the rate equations of Sec. 6.2 need to be modified. In the case of strong feedback, both the laser cavity and the external cavity (formed by the laser facet and the external reflecting surface) should be considered. This case is discussed in Chapter 8 in relation to coupled-cavity single-frequency lasers. When the feedback is relatively weak, it is possible to consider the rate equations for the laser cavity alone. However, the field equation (6.2.9) should be modified by adding a term that includes the contribution of the feedback.[197] The situation is similar to the case of external injection. Using Eqs. (6.2.8), (6.2.9), and (6.2.17), we obtain

$$\frac{dE}{dt} = i(\omega_0 - \omega_{th})E(t) + \tfrac{1}{2}(G - \gamma)(1 - i\beta_c)E(t) + \kappa E(t - \tau)\exp(i\omega_0\tau) \tag{6.7.1}$$

where the last term accounts for the reflection feedback and is generally obtained[197] using the concept of an effective facet reflectivity (see Sec. 8.3.2). In Eq. (6.7.1), ω_0 is the mode frequency in the presence of feedback whereas ω_{th} is the mode frequency of the solitary laser operating near threshold.

A simple way to obtain an expression for κ is to consider the field at the laser facet facing the external cavity. During each round trip, the field immediately after reflection is

$$E'(t) = (R_m)^{1/2}E(t) + (1 - R_m)(f_{ext})^{1/2}E(t - \tau)\exp(i\omega_0\tau) \tag{6.7.2}$$

where R_m is the facet reflectivity, f_{ext} is the fraction of output power reflected back into the laser cavity, τ is the external-cavity round-trip time, and $\theta = \omega_0\tau$ is the phase shift that occurs during a single round trip. Multiple reflections in the external cavity are neglected since $f_{ext} \ll 1$ for the case of weak feedback. Equation (6.7.2) shows that during each round trip inside the laser cavity, a small fraction $(1 - R_m)(f_{ext}/R_m)^{1/2}$ of the intracavity field reenters the laser cavity after some delay and phase shift. Since this happens at a rate of τ_L^{-1}, the coefficient κ is

$$\kappa = \frac{(1 - R_m)}{\tau_L}\left(\frac{f_{ext}}{R_m}\right)^{1/2} \tag{6.7.3}$$

where τ_L is the laser-cavity round-trip time. The feedback fraction f_{ext} in Eq. (6.7.3) should include all coupling, absorption, and diffraction losses that

may occur during the round trip in the external cavity. In the case of reflections occurring at the far end of a fiber of length L_f, the feedback fraction is given by

$$f_{\text{ext}} = \eta_c^2 \exp(-2\alpha_f L_f) R_f \tag{6.7.4}$$

where η_c is the power-coupling efficiency, α_f is the absorption coefficient, and R_f is the reflectivity of the fiber-air interface.

Equation (6.7.1) can be used to obtain the rate equations for the amplitude and phase of the intracavity field. The carrier rate equation (6.2.20) remains unchanged. In terms of the photon and carrier populations P and N and the optical phase ϕ, the modified rate equations are

$$\dot{P} = (G - \gamma)P(t) + R_{\text{sp}} + 2\kappa[P(t)P(t-\tau)]^{1/2} \cos[\omega_0 \tau + \phi(t) - \phi(t-\tau)] \tag{6.7.5}$$

$$\dot{N} = I/q - \gamma_e N(t) - GP(t) \tag{6.7.6}$$

$$\dot{\phi} = (\omega_0 - \omega_{\text{th}}) + \tfrac{1}{2}\beta_c(G - \gamma) - \kappa\left[\frac{P(t-\tau)}{P(t)}\right]^{1/2} \sin[\omega_0 \tau + \phi(t) - \phi(t-\tau)] \tag{6.7.7}$$

where, as in Sec. 6.2, the spontaneous-emission contribution has been included through R_{sp}. The effect of reflection feedback is included through the parameter κ given by Eq. (6.7.3). When $\kappa = 0$, Eqs. (6.7.5)–(6.7.7) reduce to the single-mode rate equations of a solitary laser considered in previous sections. The difference-differential form of the modified rate equations implies that the external-reflection feedback would affect significantly the dynamic behavior of a semiconductor laser.

Before discussing the feedback-induced features, let us consider what amount of minimum external feedback is necessary in order to influence the device behavior in a significant way. A rough estimate can be obtained by comparing the various terms in Eq. (6.7.5) and requiring that the contribution of the feedback term proportional to κ exceed that of spontaneous emission R_{sp}. In the steady state, this requirement implies that $\kappa P \gg R_{\text{sp}}$, or after using Eq. (6.7.3),

$$f_{\text{ext}} \gg (R_{\text{sp}} \tau_L / P)^2 \tag{6.7.8}$$

where we have set $R_m/(1 - R_m)^2$ equal to 1 in this order-of-magnitude calculation after using $R_m \cong 0.32$. The condition (6.7.8) shows the important device parameters that determine the sensitivity of a semiconductor laser

to external parasitic reflections; the smaller the right-hand side, the more sensitive is the laser to external reflections. In particular, (i) index-guided lasers are more sensitive than gain-guided lasers because of a smaller R_{sp}; (ii) short-cavity lasers are more sensitive because of a smaller τ_L; and (iii) reflection sensitivity increases with an increase in the output power. The condition (6.7.8) can be written in other equivalent forms by noting that $\tau_L^{-1} = \Delta\nu$ (the longitudinal-mode spacing) and from Eq. (6.5.52) $R_{sp}/P \sim \Delta f = \tau_c^{-1}$, where Δf is the solitary-laser CW line width and τ_c is the corresponding coherence time. We then obtain

$$f_{ext} \gg \left(\frac{\tau_L}{\tau_c}\right)^2 \quad \text{or} \quad f_{ext} \gg \left(\frac{\Delta f}{\Delta \nu}\right)^2. \tag{6.7.9}$$

These inequalities show that single-frequency lasers, oscillating predominantly in a single longitudinal mode, are more sensitive to external reflections because of their longer coherence time, or equivalently because of their smaller spectral width.

6.7.2 Steady-State Behavior

The steady-state behavior of a semiconductor laser can be readily analyzed using Eqs. (6.7.5)–(6.7.7) after assuming that P, N, and ϕ are time-independent. For simplicity we neglect R_{sp}, which mainly affects the threshold sharpness as discussed in Sec. 6.3.1. In the absence of reflection feedback ($\kappa = 0$), the solution is

$$\bar{G} = \gamma \qquad \bar{\omega}_0 = \omega_{th} \qquad \bar{P} = (I - I_{th})/q\gamma \tag{6.7.10}$$

where the threshold current $I_{th} = q\gamma_e N$. If ΔG and $\Delta\omega$ represent the gain and frequency shifts introduced by the feedback, they are given by

$$\Delta G = G - \bar{G} = -2\kappa \cos(\omega_0 \tau) \tag{6.7.11}$$

$$\Delta\omega = \omega_0 - \omega_{th} = \kappa[\sin(\omega_0\tau) + \beta_c \cos(\omega_0\tau)]. \tag{6.7.12}$$

The corresponding output power can be obtained using Eq. (6.7.6) and is given by

$$P = (I - I_{th})/[q(\gamma + \Delta G)]. \tag{6.7.13}$$

The output power may increase or decrease with feedback depending on whether ΔG is negative or positive.

For given values of κ and τ, Eqs. (6.7.11)–(6.7.13) can be used to obtain the steady-state characteristics. The device behavior, however, depends on whether the frequency equation (6.7.12) allows single or multiple solutions. It can be readily verified that a single solution is obtained when the feedback parameter

$$C = \kappa\tau(1 + \beta_c^2)^{1/2} < 1. \tag{6.7.14}$$

Using Eqs. (6.7.3) and (6.7.14) we note that the value of C depends on both the feedback fraction f_{ext} and the external-cavity optical length L_{ext} ($\tau = 2L_{ext}/c$). Let us first consider the case of a short external cavity such that $C < 1$.

The frequency shift $\Delta\omega$ can be obtained by solving Eq. (6.7.12). Equations (6.7.11) and (6.7.13) then yield the gain change ΔG and the corresponding mode power P. Both the output power P and the frequency shift $\Delta\omega$ can however be changed by changing τ through variations in the external cavity length. The situation can be visualized by treating the round-trip phase shift ($\theta = \omega_0\tau$) occurring in the external cavity as a continuous parameter. Its parametric variation results in an ellipse in the $\Delta\omega$–ΔG plane, as shown in Fig. 6.25. As θ is varied by changing τ, both the mode frequency and the mode power change accordingly. The periodic variation of P and ω_0 with the external-cavity length has been observed experimentally.[195,197] The relative power change depends on the feedback fraction. If we use Eqs. (6.7.3) and (6.7.11), it is easy to see that $|\Delta G/G|_{max} \cong (f_{ext})^{1/2}$; i.e., a 1% feedback can change the mode gain and the power each by 10%. This result again emphasizes the sensitivity of semiconductor lasers to external reflections.

The effect of external reflections on the light-current (L-I) characteristics is to introduce periodic ripples. Lang and Kobayashi[197] have discussed the L-I curves and the current-induced spectral shifts in the presence of feedback. When the feedback is relatively strong, a hysteresis is observed in both the oscillation frequency and the L-I curve. These experimental results can be explained by taking into account the thermal variation of the refractive index. As the current is increased, the temperature rise under CW operation shifts the longitudinal-mode frequency of the laser cavity, which affects the stability of the lasing mode.

When the feedback conditions are such that the parameter C defined by Eq. (6.7.14) exceeds unity, Eq. (6.7.12) can be satisfied for multiple values of ω_0 while τ is kept fixed. The number of allowed stable solutions increases as C increases. The frequency difference between the two neighboring solutions is approximately given by τ^{-1}, or the longitudinal-mode spacing of the external cavity. These multiple solutions are referred[202] to as the *external-cavity modes* (XCM), which appear in the vicinity of the *laser-cavity mode*

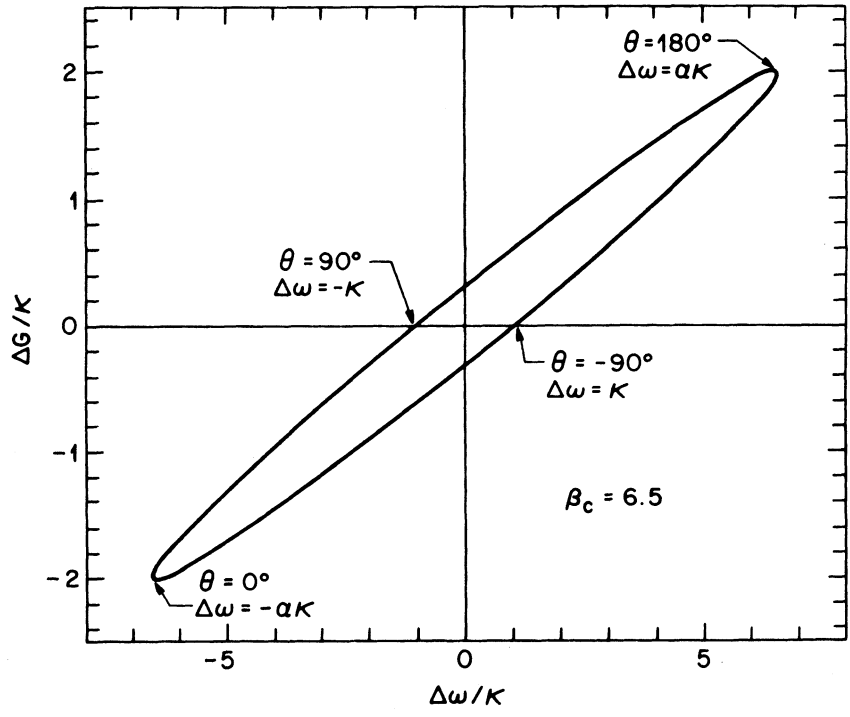


Fig. 6.25 Gain change ΔG versus frequency change $\Delta\omega$ for the possible steady states under external feedback. The laser modes are points that move along the ellipse as the feedback parameters change. (After Ref. 219 © 1986 IEEE)

(LCM), the dominant longitudinal mode of the solitary laser. Figure 6.26 shows schematically the longitudinal-mode spectra of semiconductor laser with and without feedback. With reference to Fig. 6.25, the multiple solutions of Eqs. (6.7.11) and (6.7.12) are discrete points on the lower limb of the ellipse. The solution closest to the bottom of the ellipse corresponds to the lowest threshold mode, which becomes the dominant mode in Fig. 6.26. Other solutions correspond to XCMs that have smaller amplitudes because of the mode selectivity provided by the gain margin (different ΔG values in Fig. 6.25).

Feedback-induced XCMs have been observed and studied in detail.[196,202] Their number and amplitudes depend on the external-cavity length L_{ext} as well as on the feedback fraction f_{ext}. They are readily observed for relatively long external cavities ($L_{ext} \gtrsim 1$ m) since the feedback parameter $C \gg 1$ in that case, even for weak feedback ($f_{ext} \lesssim 10^{-3}$). In particular, the feedback from the far end of an optical fiber can generate a larger number

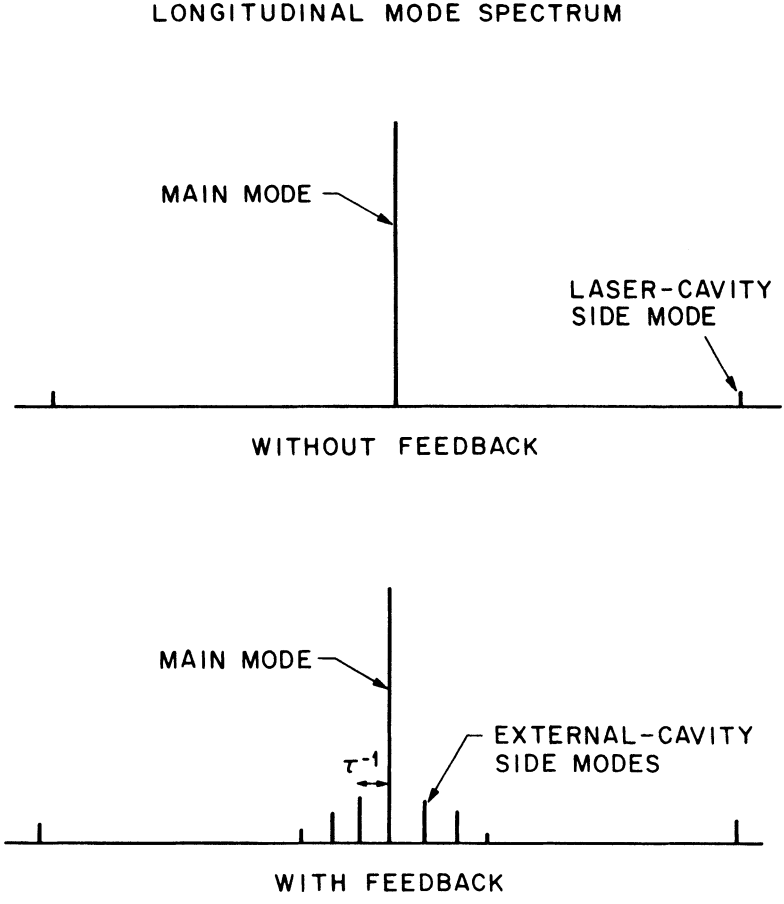


Fig. 6.26 Schematic illustration of external-gravity side modes induced by reflection feedback.

of XCMs in the power spectrum of a semiconductor laser.[196] Both the parameter C and the feedback fraction f_{ext} are shown in Fig. 6.27 as a function of the fiber length for three values of the fiber loss after using Eqs. (6.7.4) and (6.7.14). The feedback parameter $C \gg 1$ for fiber lengths in a wide range of 1 m to 100 km. Clearly, far-end reflections can generate additional XCMs under typical operating conditions of an optical communication system. Among other effects, the presence of far-end reflections increases laser-intensity noise and can lead to an increase in the bit error rate.[220]

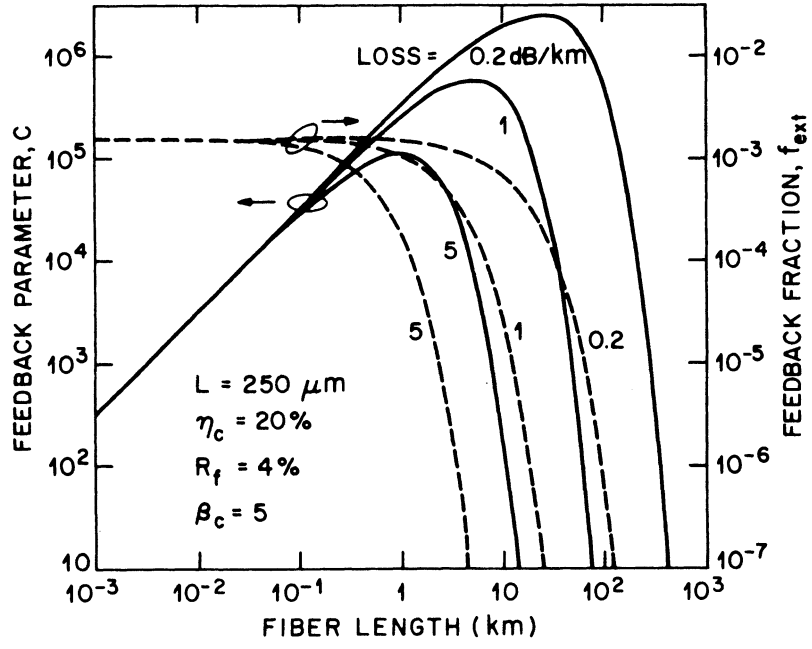


Fig. 6.27 Variation of the feedback parameter C (solid curves) and the feedback fraction f_{ext} (dashed curves) with the fiber length. In both cases curves are shown for three different values of fiber loss. (After Ref. 220 © 1986 IEEE)

6.7.3 Dynamic Behavior

The previous section discussed the steady-state solutions of the modified rate equations (6.7.5)–(6.7.7). However, a stationary solution may not be dynamically stable; i.e., a small deviation from the steady state may grow indefinitely. As in Sec. 6.4.3, the dynamic stability is examined through small-signal analysis of the modified rate equations. In contrast to the relaxation-oscillation analysis presented there, the optical phase ϕ now plays a critical role since the external feedback provides a reference phase (similar to the case of injection locking). In the small-signal analysis, the variables P, N, and ϕ are perturbed by a small amount around their steady-state values and Eqs. (6.7.5)–(6.7.7) are linearized in terms of small perturbations δP, δN, and $\delta\phi$. The resulting set of homogeneous linear equations can be readily solved after assuming that perturbations vary in time as $\exp(zt)$. The procedure is similar to that of Sec. 6.4.3 except for the complications arising because the rate equations take the difference-differential form in the presence of reflection feedback. The decay or growth rate z of perturbations is obtained

by the transcendental equation[197,213]

$$(z + \Gamma_N)[(z + \kappa_c\xi)(z + \kappa_c\xi + \Gamma_P) + \kappa_s^2\xi^2] + GG_NP[z + (\kappa_c - \beta_c\kappa_s)\xi] = 0 \quad (6.7.15)$$

where

$$\kappa_c = \kappa\cos(\omega_0\tau) \qquad \kappa_s = \kappa\sin(\omega_0\tau) \quad (6.7.16)$$

$$\xi(z) = 1 - \exp(-z\tau) \quad (6.7.17)$$

and other parameters have been defined in Sec. 6.4.3.

A steady-state solution represents a dynamically stable state of the system if all roots of Eq. (6.7.15) have a negative real part. In general the roots are obtained numerically for a given set of parameters. However, in the case of weak feedback such that $\kappa\tau \ll 1$, the solutions of Eq. (6.7.15) satisfy $|\tau z| \ll 1$, and one can replace $\xi(z)$ with τz in Eq. (6.7.15). The resulting polynomial has three roots: $z = 0$ and $z = -\Gamma_R \pm i\Omega_R$, where Γ_R and Ω_R represent, respectively, the decay rate and the frequency of relaxation oscillations given by

$$\Gamma_R = \frac{1}{2}\left[\Gamma_N + \frac{(1 + \kappa_c\tau)}{(1 + \kappa_c\tau)^2 + \kappa_s^2\tau^2}\Gamma_P\right] \quad (6.7.18)$$

$$\Omega_R \cong \bar{\Omega}_R\left[\frac{1 + (\kappa_c - \beta_c\kappa_s)\tau}{(1 + \kappa_c\tau)^2 + \kappa_s^2\tau^2}\right]^{1/2} \quad (6.7.19)$$

where $\bar{\Omega}_R = (GG_NP)^{1/2}$ and is the relaxation-oscillation frequency of the solitary laser.

Equations (6.7.18) and (6.7.19) show that the external feedback affects both the frequency and the decay rate of relaxation oscillations. Under certain feedback conditions, Ω_R^2 can become negative, making z real and positive. The CW state is then unstable and the laser is likely to exhibit self-pulsing. Using Eqs. (6.7.16) and (6.7.19), a necessary condition for stability is

$$1 + C\cos(\omega_0\tau + \tan^{-1}\beta_c) > 0 \quad (6.7.20)$$

where the feedback parameter C is given by Eq. (6.7.14). When $C > 1$, the inequality in Eq. (6.7.20) is violated for some phases θ, making the corresponding steady-state solution unstable. For large values of C, however, it is necessary to use Eq. (6.7.15) for the linear stability analysis.[213] The device behavior is then sensitive to the exact value of τ, and small changes

can enhance or damp relaxation oscillations. In general, for a given external-cavity length, the steady state becomes unstable for a range of feedback levels f_{ext} and phases θ. The semiconductor laser would then exhibit feedback-induced pulsations. By contrast, under certain conditions the external feedback can stabilize the output of a self-pulsing semiconductor laser.[239,240]

Recently a new kind of reflection-induced instability has been observed[216,217] wherein the semiconductor laser exhibits pulsations with a period of about $5–10\tau$. Furthermore, during each pulsation, ripples of width τ are superimposed on the pulse. It is found[219] to be necessary to go beyond the preceding linear stability analysis in order to explain the pulsation behavior. The physical mechanism behind this instability can be understood by referring to Fig. 6.25. In the absence of external reflections, the operating point is at the origin ($\Delta\omega = \Delta G = 0$). In the presence of reflections, the steady state corresponds to operation at the bottom of the ellipse. The approach to the steady state requires traveling along the lower portion of the ellipse. However, as the steady-state point is approached, an instability due to carrier-induced phase fluctuations shifts the operating point from the bottom to the top of the ellipse, resulting in a drop of the output power.[219] The whole process then repeats itself. The frequency chirp associated with self-pulsations manifests as a substantial broadening of the CW line width that has been observed experimentally.[217,218]

6.7.4 Noise Characteristics

The external feedback affects the intensity noise as well as phase noise. The intensity-noise spectrum at high frequencies is found to consist of a large number of equally spaced peaks[196,215] whose frequency separation coincides exactly with the external-cavity mode spacing τ^{-1}. Temkin et al.[217] have observed an additional low-frequency peak that is associated with the instability described in the previous section. The high-frequency noise peaks can be explained[215] using the small-signal analysis of the modified rate equations after adding the Langevin noise sources and following the procedure of Sec. 6.5.2. Explanation of the low-frequency noise requires consideration of relatively large fluctuations.[219]

The effect of external reflections on the phase noise has attracted attention, since it can lead to considerable narrowing or broadening of the line width of a single longitudinal mode[201–203] depending on the relative phase of the returned light. The problem has been analyzed theoretically[207–210] using the modified rate equations. Following an analysis similar to that given in Sec. 6.5.3, the line width in the presence of external feedback is approximately

given by[210]

$$(\Delta f)_{\text{ref}} \cong \frac{\Delta f}{[1 + C\cos(\omega_0\tau + \phi_R)]^2} \tag{6.7.21}$$

where $\phi_R = \tan^{-1}(\beta_c)$ and Δf is the solitary-laser line width given by Eq. (6.5.52). Equation (6.7.21) shows that, depending upon the feedback phase $\omega_0\tau$, the laser line width can broaden or narrow. The maximum narrowing occurs when $\omega_0\tau + \phi_R = 2m\pi$, and the reduction factor in that case is $(1 + C)^2$. Clearly the line width can be reduced considerably for large values of C if the external-cavity length is fine-tuned to adjust the phase $\omega_0\tau$. Reductions by more than three orders of magnitude have been observed.[241] As discussed earlier, when $C > 1$ there are multiple stable steady states over which the laser can be made to oscillate by changing the external-cavity length.[214] In the vicinity of mode hops, the line width can vary by a large amount with a small change in τ.[207] For $C < 1$, only one mode oscillates for all values of the phase. Considerable line broadening can occur in this case in a certain range of phase values, as seen clearly by an inspection of Eq. (6.7.21).

The effect of external feedback on the phase noise at high frequencies has also been discussed.[209] The phase noise is considerably enhanced in the vicinity of the relaxation-oscillation frequency Ω_R. This results in enhanced side peaks when $\omega = \omega_0 \pm \Omega_R$ in the optical spectrum of the semiconductor laser. Figure 6.28 shows the experimentally observed power spectrum when a small fraction ($f_{\text{ext}} \cong 10^{-3}$) of the power is reflected back from the far end of a 7.5-km-long fiber. The side peaks are displaced from the center-line by 4.88 GHz, which corresponds to the relaxation-oscillation frequency of the laser. These side peaks are present even for the case of a solitary laser, but their amplitude is usually below 1% of the central peak.[79]

Physically one can interpret the growth of the side peaks in the optical spectrum (see Fig. 6.28) as unclamping of the relaxation oscillations such that the laser power is self-modulated at the relaxation-oscillation frequency. In other words, feedback destabilizes the CW state and makes the laser output periodic. With a further increase in the feedback, the amplitude of the side peaks (AM sidebands) grows, and the optical spectrum becomes broader. When the external feedback exceeds a critical level, the optical spectrum becomes extremely broad with a line width in the range of a few gigahertz. This extreme broadening of the laser line width is referred to as coherence collapse,[218] since it implies that the semiconductor laser has a much shorter coherence time than the value expected for the solitary laser in the absence of the feedback. The phenomenon of coherence collapse has attracted considerable attention[222–238] since it represents an example of

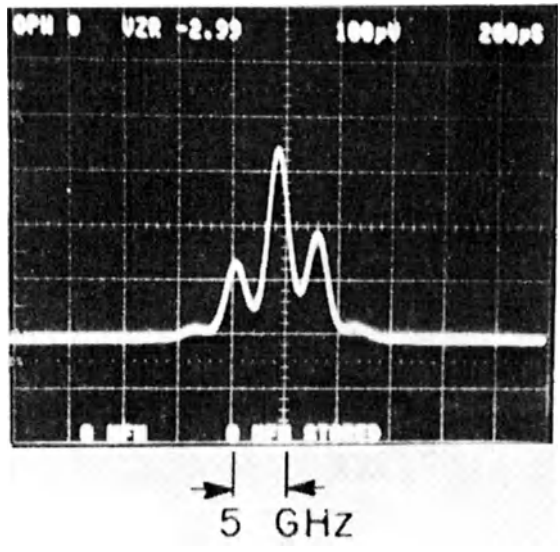


Fig. 6.28 Optical spectrum of a single longitudinal mode of a semiconductor laser in the presence of external feedback. The side peaks at multiples of the relaxation-oscillation frequency of 4.88 GHz are considerably enhanced because of feedback-induced phase noise. (After Ref. 160)

feedback-induced chaos in semiconductor lasers resulting from destabilization of the CW or self-pulsing state. The presence of spontaneous-emission fluctuations makes it difficult to distinguish between the chaotic and stochastic dynamics, but the evidence of chaos has been seen in several experiments.[236–238] Numerical solutions of the rate equations [Eqs. (6.7.5)–(6.7.7)] clearly predict the onset of feedback-induced chaos when the feedback exceeds a critical level.[223,227] Considerable experimental and theoretical work has been carried out to understand the chaotic dynamics in semiconductor lasers.[222–238] From a practical standpoint, the coherence-collapse region should be avoided in the design of coherent communication systems requiring narrow-line-width semiconductor lasers.[138]

The effect of external feedback on the RIN has also been studied.[227,242–248] In the low-feedback regime in which the steady state remains stable, an analytic expression for RIN can be obtained.[242] In general, RIN can increase considerably with feedback. Similar to the case of laser line width, RIN is sensitive to the feedback phase [the parameter $\omega_0\tau$ in Eq. (6.7.21)] which can cause variations in RIN by as much as 10–20 dB. In fact, RIN becomes minimum under the same conditions in which the line width is minimum.[248] In practice, however, the feedback phase is not easy to control, and RIN often increases significantly with feedback. The feedback-induced

enhancement of RIN is of considerable concern in the design of optical data-storage systems where the unintended feedback from the optical disk can severely degrade the system performance.[244–247] A technique, known as high-frequency injection,[245] solves this problem by modulating the laser current at frequencies much higher than the data rate. However, RIN depends on many parameters such as modulation frequency, modulation depth, and round-trip time in the external cavity. In general, RIN decreases with an increase in the modulation current, and the decrease is larger for longer external-cavity lengths.[211] It can increase abruptly above a certain modulation depth whose exact value depends, among other things, on the modulation frequency and the external-cavity length.[212] The use of the high-frequency injection technique requires optimization of several operating parameters which generally vary from laser to laser.[247]

It is clear that parasitic reflections in an optical communication system can degrade system performance.[220] The problem of parasitic reflections becomes severe for coherent systems requiring semiconductor lasers with a stable amplitude and phase. The use of isolators between the laser and the fiber may then become a necessity. The amount of required isolation is governed by Eqs. (6.7.4) and (6.7.9), and may exceed 50 dB ($f_{\text{ext}} \lesssim 10^{-5}$).[221]

PROBLEMS

6.1 Determine the number of photons inside the cavity of a 1.55-μm semiconductor laser with 30% facet reflectivities. Assume a group index of 4.0. The laser is operating continuously at an output power of 5 mW.

6.2 A 1.3-μm semiconductor laser has its facet coated such that $R_1 = 90\%$ and $R_2 = 1\%$. Determine the average number of photons stored in the cavity when the power emitted from the antireflection-coated facet is 5 mW.

6.3 Show that Eq. (6.2.9) follows from the wave equation as outlined in Sec. 6.2. Complete the missing steps in its derivation.

6.4 Use the single-mode rate equations to obtain analytic expressions of P and N in the steady state by treating γ_e as a constant. Neglect spontaneous emission for simplicity.

6.5 Repeat Problem 6.4 by using $R_{\text{sp}} = n_{\text{sp}} G$ for the spontaneous-emission term in the photon rate equation, Eq. (6.2.14). Plot the quantities P and N as a function of I/I_{th} for $n_{\text{sp}} = 0$, 1, and 2.

6.6 Prove by using the multimode rate equations that the longitudinal-mode spectrum of a semiconductor laser is Lorentzian when the gain spectrum can be approximated by a parabola near the gain peak.

6.7 Derive analytic expressions for the frequency Ω_R and the decay rate Γ_R

of relaxation oscillations for a single-mode semiconductor laser by assuming that the gain G varies with N and P as

$$G(N, P) = \frac{G_N(N - N_0)}{1 + P/P_s}.$$

The parameter P_s is referred to as the intraband saturation photon number.

6.8 Repeat Problem 6.7 for the case in which G varies as

$$G(N, P) = \frac{G_N(N - N_0)}{\sqrt{1 + P/P_s}}.$$

Compare the variation of Ω_R and Γ_R with P in the two cases.

6.9 Complete the derivation of Eq. (6.5.19) for the RIN spectrum. Plot the low-frequency RIN as a function of P by setting $\omega = 0$ and using the appropriate parameter values from Tables 6.1 and 6.2. Show that RIN decreases as P^{-3} at low powers, but changes over to P^{-1} at high powers.

6.10 Use the RIN given in Eq. (6.5.19) to obtain an expression for the signal-to-noise ratio (SNR) of the laser output. Plot the SNR as a function of P by using the same parameter values as in Problem 6.9.

6.11 Use the expression for $G(N, P)$ given in Problem 6.8 and derive an expression for the SNR. What happens to the SNR in the limit of large output power? Consult Ref. 92 if necessary.

6.12 Show that Eq. (6.5.48) for the phase variance follows from Eq. (6.5.47) by evaluating the integral. Use it to prove that the spectral line shape is Lorentzian with an FWHM given by Eq. (6.5.52) if the contribution of relaxation oscillations is neglected.

6.13 Derive an expression for the line width of a single-mode semiconductor laser when $G(N, P)$ varies with N and P as

$$G(N, P) = \frac{G_N(N - N_0)}{\sqrt{1 + P/P_s}}.$$

Show that the line width saturates to a constant value when $P \gg P_s$. Consult Ref. 112 if necessary.

6.14 Study the effect of $1/f$-noise on the spectral line shape by adding a term A/ω to the frequency noise spectrum $S_{\dot{\phi}}(\omega)$ in Eq. (6.5.50). Prove that the line width saturates to a constant value determined by the parameter A in the large-power limit. What is the line shape in this limit?

6.15 Use the small-signal modulation response given by Eq. (6.6.16) to obtain an analytic expression of the 3-dB modulation bandwidth in terms of the relaxation-oscillation parameters Ω_R and Γ_R.

6.16 Study the effect of nonlinear gain on the 3-dB modulation bandwidth for the case

$$G(N, P) = \frac{G_N(N - N_0)}{1 + P/P_s}$$

by using the results of Problem 6.7. How do the results change when $G(N, P)$ is replaced by the one given in Problem 6.8?

6.17 Use Eq. (6.6.32) for the 3-dB modulation bandwidth ν_B to discuss how one can design semiconductor lasers with large values of ν_B.

6.18 Discuss the origin of frequency chirping during direct modulation of semiconductor lasers. Derive an expression for the maximum frequency shift when the laser current is modulated sinusoidally.

6.19 Show that the chirp-to-power ratio (CPR), defined as the ratio of modulated power δP_0 to the frequency chirp $\delta\nu_0$, takes a constant value $\beta_c|G_p|/4\pi$ in the limit $\omega_m \ll \Gamma_P$. Find an expression for the CPR in the opposite limit $\omega_m \gg \Gamma_P$. How can you determine the two important parameters β_c and G_P by measuring the CPR?

6.20 Write a computer program capable of solving the single-mode rate equations for an arbitrary form of the injection current $I(t)$. Use it to show the approach to steady state when the current is changed abruptly at time $t = 0$ from its below-threshold value $I_{th}/2$ to $2I_{th}$. Repeat your calculations for several values of the nonlinear-gain parameter G_p and discuss your results.

6.21 Describe the technique of gain switching. Use the computer program developed in Problem 6.20 to study the dependence of gain-switched pulses on the bias level.

6.22 Derive Eq. (6.7.21) by solving Eqs. (6.7.5)–(6.7.7) after adding the Langevin noise terms. Consult Ref. 210 if necessary.

REFERENCES

1. Statz, H., and G. A. deMars, p. 530 in *Quantum Electronics*, ed. C. H. Townes. New York: Columbia University Press, 1960.
2. Statz, H., C. L. Tang, and J. M. Lavine. *J. Appl. Phys.* **35,** 2581 (1964).
3. Ikegami, T., and Y. Suematsu. *Proc. IEEE* **55,** 122 (1967); *Electron. Commun. Jpn.* **51-B,** 51 (1968); **53-B,** 69 (1970).
4. Paoli, T. L., and J. E. Ripper. *Proc. IEEE* **58,** 1457 (1970).
5. Adams, M. J. *Opto-Electron.* **5,** 201 (1973).
6. Salathé, R. P., C. Voumard, and H. Weber. *Opto-Electron.* **6,** 451 (1974).

7. Bores, P. M., M. T. Vlaardingerbroek, and M. Danielsen. *Electron. Lett.* **11,** 206 (1975).
8. Kressel, H., and J. K. Butler. *Semiconductor Lasers and Heterojunction LEDs.* New York: Academic Press, 1977.
9. Thompson, G. H. B. *Physics of Semiconductor Laser Devices.* Chichester: John Wiley & Sons, 1980.
10. Arnold, G., P. Russer, and K. Petermann. Chap. 7 in *Semiconductor Devices for Optical Communication*, ed. H. Kressel. Berlin: Springer-Verlag, 1982.
11. Henry, C. H. Chap. 3 in *Semiconductors and Semimetals*, Vol. 22, Part B, ed. W. T. Tsang. New York: Academic Press, 1985; *J. Lightwave Technol.* **LT-4,** 288 (1986).
12. Lau, K. Y., and A. Yariv. Chap. 2 in *Semiconductors and Semimetals*, Vol. 22, Part B, ed. W. T. Tsang. New York: Academic Press, 1985.
13. Helkey, R. J., P. A. Morton, and J. E. Bowers. *Opt. Lett.* **15,** 112 (1990).
14. Petermann, K. *Laser Diode Modulation and Noise.* Dordrecht: Kluwer Academic Pub., 1991.
15. Agrawal, G. P., and G. R. Gray. *SPIE Proc.* **1497,** 444 (1991).
16. Suematsu, Y., and K. Furuya. *Electron. Commun. Jpn.* **E-60,** 467 (1977).
17. Petermann, K. *IEEE J. Quantum Electron.* **QE-15,** 566 (1979).
18. Streifer, W., D. R. Scifres, and R. D. Burnham. *Electron. Lett.* **17,** 933 (1981).
19. Patzak, E. *Electron. Lett.* **18,** 278 (1982).
20. Marcuse, D. *Electron. Lett.* **18,** 920 (1982).
21. Yariv, A., and S. Margalit. *IEEE J. Quantum Electron.* **QE-18,** 1831 (1982).
22. Elsässer, W., and E. O. Göbel. *Electron. Lett.* **19,** 335 (1983).
23. Arnaud, J. *Electron. Lett.* **19,** 688 (1983).
24. Agrawal, G. P. *J. Opt. Soc. Am. B.* **1,** 406 (1984).
25. Newstein, M. *IEEE J. Quantum Electron.* **QE-20,** 1270 (1984).
26. Lee, T. P., C. A. Burrus, D. Marcuse, A. G. Dentai, and R. J. Nelson. *Electron. Lett.* **18,** 902 (1982).
27. Gordon, E. I. *Bell Syst. Tech. J.* **43,** 507 (1964).
28. Wilcox, J. Z., and L. W. Casperson. *J. Appl. Phys.* **56,** 57 (1984).
29. Casperson, L. W. *J. Appl. Phys.* **46,** 5194 (1975).
30. Suematsu, Y., S. Akiba, and T. Hong. *IEEE J. Quantum Electron.* **QE-13,** 596 (1977).
31. Renner, D., and J. E. Carroll. *Electron. Lett.* **14,** 779 (1978).
32. Streifer, W., D. R. Scifres, and R. D. Burnham. *IEEE J. Quantum Electron.* **QE-18,** 1918 (1982).
33. Pietzch, J., and T. Kamiya. *Jpn. J. Appl. Phys.*, Part 2, **21,** L633 (1982).
34. Lee, T. P., C. A. Burrus, J. A. Copeland, A. G. Dentai, and D. Marcuse. *IEEE J. Quantum Electron.* **QE-18,** 1101 (1982).
35. Streifer, W., D. R. Scifres, and R. D. Burnham. *Appl. Phys. Lett.* **40,** 305 (1982).
36. Stegmüller, B. *Appl. Phys. Lett.* **42,** 15 (1983).
37. Danielmeyer, H. G. *J. Appl. Phys.* **42,** 3125 (1971).
38. Streifer, W., R. D. Burnham, and D. R. Scifres. *IEEE J. Quantum Electron.* **QE-13,** 403 (1977).
39. Copeland, J. A. *IEEE J. Quantum Electron.* **QE-16,** 721 (1980).
40. Zee, B. *IEEE J. Quantum Electron.* **QE-14,** 727 (1978).
41. Manning, J., R. Olshansky, D. M. Fye, and W. Powazinik. *Electron. Lett.* **21,** 497 (1985).
42. Yamada, M., and Y. Suematsu. *IEEE J. Quantum Electron.* **QE-15,** 743 (1979); *J. Appl. Phys.* **52,** 2653 (1981).
43. Kazarinov, R. F., C. H. Henry, and R. A. Logan. *J. Appl. Phys.* **53,** 4631 (1982).
44. Agrawal, G. P. *IEEE J. Quantum Electron.* **QE-23,** 860 (1987).
45. Ogasawara, N., and R. Ito. *Jpn. J. Appl. Phys.* **27,** 607 & 615 (1988).
46. Yamada, M. *J. Appl. Phys.* **66,** 81 (1989).
47. Agrawal, G. P. *Opt. Lett.* **12,** 260 (1987); *Appl. Phys. Lett.* **51,** 302 (1987); *J. Opt. Soc. Am. B* **5,** 147 (1988).

48. Adams, M. J., and M. Osinski. *IEE Proc. H.* **129,** 271 (1982).
49. Marcuse, D., and T. P. Lee. *IEEE J. Quantum Electron.* **QE-19,** 1397 (1983).
50. Lau, K. Y., C. Harder, and A. Yariv. *IEEE J. Quantum Electron.* **QE-20,** 71 (1984).
51. M. Demokan, and A. Nacaroglu. *IEEE J. Quantum Electron.* **QE-20,** 1016 (1984).
52. Konnerth, K., and C. Lanza. *Appl. Phys. Lett.* **4,** 120 (1964).
53. Ripper, J. E. *J. Appl. Phys.* **43,** 1762 (1972).
54. Chinone, N., R. Ito, and O. Nakada. *IEEE J. Quantum Electron.* **QE-10,** 81 (1974).
55. Dixon, R. W., and W. B. Joyce. *J. Appl. Phys.* **50,** 4591 (1979).
56. Dutta, N. K., R. L. Hartman, and W. T. Tsang. *IEEE J. Quantum Electron.* **QE-19,** 1613 (1983).
57. Olshansky, R., C. B. Su, J. Manning, and W. Powazinik. *IEEE J. Quantum Electron.* **QE-20,** 838 (1984).
58. Van der Ziel, J. P., J. L. Mertz, and T. L. Paoli. *J. Appl. Phys.* **50,** 4620 (1979); see other references cited therein for previous work.
59. Arnold, G., and K. Petermann. *Opt. Quantum Electron.* **10,** 311 (1978).
60. Dixon, R. W., and W. B. Joyce. *IEEE J. Quantum Electron.* **QE-15,** 470 (1979).
61. Haaki, B. W. *J. Appl. Phys.* **50,** 5630 (1979).
62. Henry, C. H. *J. Appl. Phys.* **51,** 3051 (1980).
63. Van der Ziel, J. P. *IEEE J. Quantum Electron.* **QE-17,** 60 (1981).
64. Furuya, K., Y. Suematsu, Y. Sakakibara, and M. Yamada. *Electron. Commun. Jpn.* **E-62,** 241 (1979).
65. Adams, M. J., and M. Osinski. *Electron. Lett.* **19,** 627 (1983).
66. McCumber, D. E. *Phys. Rev.* **141,** 306 (1966).
67. Lax, M. *IEEE J. Quantum Electron.* **QE-3,** 37 (1967).
68. Haug, H. *Phys. Rev.* **184,** 338 (1969); Haug, H., and H. Haken, *Z. Phys.* **204,** 262 (1967).
69. Morgan, D. J., and M. J. Adams. *Phys. Status Solidi A* **11,** 243 (1972).
70. Paoli, T. L. *IEEE J. Quantum Electron.* **QE-11,** 276 (1975).
71. Jäckel, H., and G. Guekos. *Opt. Quantum Electron.* **9,** 233 (1977); Jäckel, H. PhD. Thesis (1980).
72. Ito, T., S. Machida, K. Nawata, and T. Ikegami. *IEEE J. Quantum Electron.* **QE-13,** 574 (1977).
73. Paoli, T. L. *IEEE J. Quantum Electron.* **QE-15,** 807 (1979).
74. Osterwalder, J. M., and B. J. Rickett. *Proc. IEEE* **67,** 1671 (1979).
75. Fleming, M. W., and A. Mooradian. *Appl. Phys. Lett.* **38,** 511 (1981).
76. Yamamoto, Y., T. Mukai, and S. Saito. *Electron. Lett.* **17,** 327 (1981).
77. Henry, C. H. *IEEE J. Quantum Electron.* **QE-18,** 259 (1982); **QE-19,** 1391 (1983); *J. Lightwave Technol.* **LT-4,** 298 (1986).
78. Daino, B., P. Spano, M. Tamburrini, and S. Piazzolla. *IEEE J. Quantum Electron.* **QE-19,** 266 (1983).
79. Vahala, K., C. Harder, and A. Yariv. *Appl. Phys. Lett.* **42,** 211 (1983); Vahala, K., and A. Yariv, *IEEE J. Quantum Electron.* **QE-19,** 1096 (1983); **QE-19,** 1102 (1983).
80. Schimpe, R. *Z. Phys. B* **52,** 289 (1983); Schimpe, R., B. Stegmüler, and W. Harth. *Electron. Lett.* **20,** 206 (1984).
81. Yamamoto, Y. *IEEE J. Quantum Electron.* **QE-19,** 34 (1983); Yamamoto, Y., S. Saito, and T. Mukai. *IEEE J. Quantum Electron.* **QE-19,** 47 (1983); Mukai, T., and Y. Yamamoto. *Electron. Lett.* **20,** 29 (1984).
82. Kikuchi, K., T. Okoshi, and R. Arata. *Electron. Lett.* **20,** 353 (1984); Kikuchi, K., and T. Okoshi. *Electron. Lett.* **20,** 1044 (1984).
83. Elsässer, W., and E. O. Göbel. *Appl. Phys. Lett.* **45,** 353 (1984); *IEEE J. Quantum Electron.* **QE-21,** 687 (1985); Elsässer, W. *Appl. Phys. Lett.* **44,** 1126 (1984).

84. Krüger, U., and K. Petermann. *IEEE J. Quantum Electron.* **QE-24,** 750 (1988); **QE-26,** 2058 (1990).
85. Miller, S. E. *IEEE J. Quantum Electron.* **QE-24,** 750 (1988); **QE-24,** 1873 (1988); **QE-25,** 1771 (1989).
86. Kikuchi, K. *Electron. Lett.* **24,** 1001 (1988); *IEEE J. Quantum Electron.* **QE-25,** 684 (1989).
87. Arnaud, J. *Electron. Lett.* **22,** 538 (1986); **24,** 116 (1988); *Opt. Lett.* **13,** 728 (1988).
88. Duan, G.-H., P. Gallion, and G. Debarge. *IEEE J. Quantum Electron.* **QE-26,** 32 (1990).
89. Tromborg, B., H. Olesen, and X. Pan. *IEEE J. Quantum Electron.* **QE-27,** 178 (1991).
90. Agrawal, G. P. *SPIE Proc.* **1376,** 224 (1991); Gray, G. R., and G. P. Agrawal. *IEEE Photon. Technol. Lett.* **3,** 204 (1991).
91. Lax, M. *Rev. Mod. Phys.* **32,** 25 (1960); **38,** 541 (1966); *Phys. Rev.* **160,** 290 (1967); *Phys. Rev.* **157,** 231 (1967); Lax, M., and W. H. Louisell. *Phys. Rev.* **185,** 568 (1969).
92. Agrawal, G. P. *Electron. Lett.* **27,** 232 (1991).
93. Ogawa, K. *IEEE J. Quantum Electron.* **QE-18,** 849 (1982).
94. Liu, P.-L., and K. Ogawa, *J. Lightwave Technol.* **LT-2,** 44 (1984).
95. Henry, C. H., P. S. Henry, and M. Lax. *J. Lightwave Technol.* **LT-2,** 209 (1984).
96. Marcuse, D. *IEEE J. Quantum Electron.* **QE-20,** 1139 (1984); **QE-20,** 1148 (1984).
97. Agrawal, G. P. *Phys. Rev. A,* **37,** 2488 (1988).
98. Aoki, M., K. Uomi, T. Tsuchiya, S. Sasaki, M. Okai, and N. Chinone. *IEEE J. Quantum Electron.* **QE-27,** 1782 (1991).
99. Agrawal, G. P., and R. Roy. *Phys. Rev. A* **37,** 2495 (1988).
100. Harth, W. *Electron. Lett.* **9,** 532 (1973).
101. Tucker, R. S. *IEE Proc. I,* **128,** 180 (1981); *J. Lightwave Technol.* **LT-3,** 1180 (1985).
102. Agrawal, G. P. *Appl. Phys. Lett.* **57,** 1 (1990).
103. Russer, P., and G. Arnold. *IEEE Trans. Microwave Theory Tech.* **MTT-30,** 1809 (1982).
104. Seeds, A. J., and J. R. Forrest. *IEE Proc. I,* **129,** 275 (1982).
105. Hagimoto, K., N. Ohta, and K. Nakagawa, *Electron. Lett.* **18,** 796 (1982).
106. Lau, K. Y., N. Bar-Chaim, I. Ury, C. Harder, and A. Yariv. *Appl. Phys. Lett.* **43,** 1 (1983); Lau, K. Y., C. Harder, and A. Yariv. *Appl. Phys. Lett.* **44,** 273 (1984).
107. Linke, R. A. *J. Lightwave Technol.* **LT-2,** 40 (1984).
108. Tucker, R. S., and I. P. Kaminow. *J. Lightwave Technol.* **LT-2,** 385 (1984).
109. Lau, K. Y., and A. Yariv. *IEEE J. Quantum Electron.* **QE-21,** 121 (1985); *Appl. Phys. Lett.* **46,** 326 (1985).
110. Bowers, J. E., B. R. Hemenway, A. H. Gnauck, and D. P. Wilt. *IEEE J. Quantum Electron.* **QE-22,** 833 (1986).
111. Olshansky, R., P. Hill, V. Lanzisera, and W. Powazinik. *IEEE J. Quantum Electron.* **QE-23,** 1410 (1987).
112. Agrawal, G. P. *IEEE J. Quantum Electron.* **QE-26,** 1901 (1990).
113. Meland, E., R. Holstrom, J. Schlafer, R. B. Lauer, and W. Powazinik. *Electron. Lett.* **26,** 1827 (1990).
114. Kishino, K., S. Aoki, and Y. Suematsu. *IEEE J. Quantum Electron.* **QE-18,** 343 (1982); see other references cited therein for previous work.
115. Linke, R. A. *Electron. Lett.* **20,** 472 (1984); *IEEE J. Quantum Electron.* **QE-21,** 593 (1985).
116. Dutta, N. K., N. A. Olsson, L. A. Koszi, P. Besomi, and R. B. Wilson. *J. Appl. Phys.* **56,** 2167 (1984).
117. Agrawal, G. P., N. A. Olsson, and N. K. Dutta. *Appl. Phys. Lett.* **45,** 119 (1984).
118. Koch, T. L., and J. E. Bowers. *Electron. Lett.* **20,** 1038 (1984).
119. Lin, C., J. K. Anderson, and F. Mengel. *Electron. Lett.* **21,** 80 (1985).
120. Olshansky, R., and D. Fye. *Electron. Lett.* **20,** 928 (1984).
121. Frisch, D. A., and I. D. Henning. *Electron. Lett.* **20,** 631 (1984).

122. Agrawal, G. P. *Opt. Lett.* **1,** 10 (1985).
123. Bickers, L., and L. D. Westbrook. *Electron. Lett.* **21,** 103 (1985).
124. Osterwalder, J. M., and B. J. Rickett. *IEEE J. Quantum Electron.* **QE-16,** 250 (1980).
125. Kobayashi, S., Y. Yamamoto, M. Ito, and T. Kimura. *IEEE J. Quantum Electron.* **QE-18,** 582 (1982).
126. Shen, T. M., and G. P. Agrawal, *J. Lightwave Technol.* **LT-4,** 497 (1986).
127. Olesen, H., and G. Jacobson. *IEEE J. Quantum Electron.* **QE-18,** 2069 (1982).
128. Agrawal, G. P. *IEEE J. Quantum Electron.* **QE-21,** 680 (1985).
129. Yoshikuni, Y., T. Matsuoka, G. Motosugi, and N. Yamanoka. *Appl. Phys. Lett.* **45,** 820 (1984).
130. Lin, C., G. Eisenstein, C. A. Burrus, and R. S. Tucker. *Appl. Phys. Lett.* **46,** 12 (1985).
131. Hakki, B. W., F. Bosch, S. Lumish, and N. R. Dietrich. *J. Lightwave Technol.* **LT-3,** 1193 (1985).
132. Koch, T. L., and R. A. Linke. *Appl. Phys. Lett.* **48,** 613 (1986).
133. Tsushima, H., and Y. Suematsu. *Trans. Inst. Electron. Commun. Eng. Jpn.* **E67,** 480 (1984).
134. Buus, J. *Electron. Lett.* **21,** 129 (1985).
135. Koyama, F., and Y. Suematsu. *IEEE J. Quantum Electron.* **QE-21,** 292 (1985).
136. Tucker, R. S. *J. Lightwave Technol.* **LT-3,** 1180 (1985).
137. Harder, C., K. Vahala, and A. Yariv. *Appl. Phys. Lett.* **42,** 328 (1983).
138. Agrawal, G. P. *Fiber-Optic Communication Systems.* New York: Wiley, 1992.
139. Agrawal, G. P. *Nonlinear Fiber Optics.* San Diego: Academic Press, 1989.
140. Lin, C., P. L. Liu, T. C. Damen, D. J. Eilenberger, and R. L. Hartman. *Electron. Lett.* **16,** 600 (1980).
141. AuYeung, J. *Appl. Phys. Lett.* **38,** 308 (1981); *IEEE J. Quantum Electron.* **QE-17,** 398 (1981).
142. Aspin, G. J., J. E. Carroll, and R. G. Plumb. *Appl. Phys. Lett.* **39,** 860 (1981); *IEE Proc. I,* **29,** 283 (1982).
143. Ho, P. T., pp. 11–31 in *Picosecond Optoelectronic Devices,* ed. by C. H. Lee. New York: Academic Press, 1984.
144. Otsuka, K., pp. 34–72 in *Picosecond Optoelectronic Devices,* ed. by C. H. Lee. New York: Academic Press, 1984.
145. Schöll, E., D. Bimberg, H. Schuhmacher, and P. Landsberg. *IEEE J. Quantum Electron.* **QE-20,** 394 (1984).
146. Demokan, M. S., and A. Nacaroglu. *IEEE J. Quantum Electron.* **QE-20,** 1016 (1984).
147. White, I. H., D. F. G. Gallagher, M. Osinski, and D. Bowley. *Electron. Lett.* **21,** 197 (1985).
148. Lin, C., and J. E. Bowers. *Electron. Lett.* **21,** 1200 (1985).
149. Bimberg, D., K. Ketterer, E. H. Böttcher, and E. Schöll. *Int. J. Electron.* **60,** 23 (1986).
150. Arakawa, Y., T. Sogawa, M. Nishioka, M. Tanaka, and H. Sakaki. *Appl. Phys. Lett.* **51,** 1295 (1987).
151. Lau, K. Y. *Appl. Phys. Lett.* **52,** 257 (1988); *J. Lightwave Technol.* **7,** 400 (1989).
152. Paulus, P., R. Langenhorst, and D. Jäger. *IEEE J. Quantum Electron.* **QE-24,** 1519 (1988).
153. Liu, H. F., M. Fukazawa, Y. Kawai, and T. Kamiya. *IEEE J. Quantum Electron.* **QE-25,** 1417 (1989).
154. Hawkins, R. T. *Electron. Lett.* **26,** 292 (1990).
155. Chusseau, L., J. M. Xie, L. Duvillaret, J. M. Lourtioz, A. Accard, and J. P. Hebert. *Electron. Lett.* **26,** 1085 (1990).
156. Sogawa, T., and Y. Arakawa. *J. Appl. Phys.* **67,** 2675 (1990); *IEEE J. Quantum Electron.* **QE-27,** 1648 (1991).
157. Liu, H. F., Y. Ogawa, S. Oshiba, and T. Nonaka. *IEEE J. Quantum Electron.* **QE-27,** 1655 (1991).
158. Iwatsuki, K., A. Takada, and M. Saruwatari. *Electron. Lett.* **24,** 1572 (1988).

159. Iwatsuki, K., A. Takada, S. Nishi, and M. Saruwatari. *Electron. Lett.* **25,** 1003 (1989).
160. Iwatsuki, K., S. Nishi, M. Saruwatari, and M. Shimuzu. *Electron. Lett.* **26,** 1 (1990).
161. Nakazawa, M., K. Suzuki, and Y. Kimura. *Opt. Lett.* **15,** 588 (1990); *IEEE Photon. Technol. Lett.* **2,** 216 (1990).
162. Nakazawa, M., K. Suzuki, E. Yamada, and Y. Kimura. *Electron. Lett.* **26,** 1592 (1990).
163. Iwatsuki, K., K. Suzuki, S. Nishi, M. Saruwatari, and K. Nakazawa. *IEEE Photon. Technol. Lett.* **2,** 905 (1990).
164. Olsson, N. A., and G. P. Agrawal. *Appl. Phys. Lett.* **55,** 13 (1989).
165. Agrawal, G. P., and N. A. Olsson. *Opt. Lett.* **14,** 500 (1989); *IEEE J. Quantum Electron.* **QE-25,** 2297 (1989).
166. Siegman, A. E. *Lasers.* Mill Valley: University Science Books, 1986.
167. Ho, P. T., L. A. Glasser, E. P. Eppen, and H. A. Haus. *Appl. Phys. Lett.* **33,** 241 (1978); Ho, P. T. *Electron. Lett.* **5,** 526 (1979).
168. Ito, H., H. Yokoyama, S. Murata, and H. Inaba. *Electron. Lett.* **5,** 738 (1979).
169. Holbrook, M. B., W. Sleat, and D. Bradley. *Appl. Phys. Lett.* **37,** 59 (1980).
170. E. P. Ippen, D. J. Eilenberger, and R. W. Dixon. *Appl. Phys. Lett.* **37,** 267 (1980).
171. Haus, H. A. *J. Appl. Phys.* **51,** 4042 (1980); *Jpn. J. Appl. Phys.* **20,** 1007 (1981).
172. van der Ziel, J. P. Vol. 22B in *Semiconductors and Semimetals,* ed. by W. T. Tsang. New York: Academic Press, 1985.
173. Lau, K. Y., and A. Yariv. *Appl. Phys. Lett.* **46,** 326 (1985).
174. Lau, K. Y., and I. Ury. *Appl. Phys. Lett.* **46,** 1117 (1985).
175. Tucker, R. S., S. K. Korotky, G. Eisenstein, U. Koren, L. W. Stulz, and J. J. Veselka. *Electron. Lett.* **21,** 239 (1985).
176. Silberberg, Y., and P. W. Smith. *IEEE J. Quantum Electron.* **QE-22,** 759 (1986).
177. Bowers, J. E., P. A. Morton, A. Mar, and S. W. Corzine. *IEEE J. Quantum Electron.* **QE-25,** 1426 (1989).
178. Vasil'ev, P. P., and A. B. Sergeev. *Electron. Lett.* **25,** 1049 (1989).
179. Delfyett, P. J., C. H. Lee, L. T. Florez, N. G. Stoffel, T. J. Gmitter, N. C. Andreadakis, G. A. Alphonse, and J. C. Connolly. *Opt. Lett.* **15,** 1371 (1990).
180. Derickson, D. J., R. J. Helkey, A. Mar, P. A. Morton, and J. E. Bowers. *Appl. Phys. Lett.* **56,** 7 (1990).
181. Morton, P. A., J. E. Bowers, L. A. Koszi, M. Soler, J. Lopata, and D. P. Wilt. *Appl. Phys. Lett.* **56,** 111 (1990).
182. Wu, M. C., Y. K. Chen, T. Tanbun-Ek, R. A. Logan, M. A. Chin, and G. Raybon. *Appl. Phys. Lett.* **57,** 759 (1990).
183. May, P. G., and M. Bierbaum. *IEEE Photon. Technol. Lett.* **3,** 296 (1991).
184. Sanders, S., A. Yariv, J. Paslaski, J. E. Ungar, and H. A. Zarem. *Appl. Phys. Lett.* **58,** 681 (1991).
185. Zarrabi, J. H., E. L. Portnoi, and A. V. Chelnokov. *Appl. Phys. Lett.* **59,** 1526 (1991).
186. Derickson, D. J., P. A. Morton, J. E. Bowers, and R. L. Thornton. *Appl. Phys. Lett.* **59,** 3372 (1991).
187. Nager, R., D. Abraham, N. Tessler, A. Fraenkel, G. Eisenstein, E. P. Ippen, U. Koren, and G. Raybon. *Opt. Lett.* **16,** 1750 (1991).
188. M. Schell, A. G. Weber, E. Schöll, and D. Bimberg. *IEEE J. Quantum Electron.* **QE-27,** 1661 (1991).
189. Chen, Y. K., M. C. Wu, T. Tanbun-Ek, R. A. Logan, and M. A. Chin. *Appl. Phys. Lett.* **58,** 1253 (1991); *IEEE Photon. Technol. Lett.* **3,** 971 (1991).
190. Lau, K. Y., and J. Paslaski. *IEEE Photon. Technol. Lett.* **3,** 974 (1991).
191. Hsu, K., C. M. Verber, and R. Roy. *Appl. Phys. Lett.* **60,** 307 (1992).
192. Jiang, W. B., R. Mirin, and J. E. Bowers. *Appl. Phys. Lett.* **60,** 677 (1992).

193. Salathé, R. P. *Appl. Phys.* **20,** 1 (1979); see other references cited therein for previous work.
194. Hirota, O., and Y. Suematsu. *IEEE J. Quantum Electron.* **QE-15,** 142 (1979).
195. Kanada, T., and K. Nawata. *IEEE J. Quantum Electron.* **QE-15,** 559 (1979).
196. Ikushima, I., and M. Maeda. *IEEE J. Quantum Electron.* **QE-15,** 844 (1979); **QE-14,** 331 (1978).
197. Lang, R., and K. Kobayashi. *IEEE J. Quantum Electron.* **QE-16,** 347 (1980).
198. Lau, K. Y., L. Figueroa, and A. Yariv. *IEEE J. Quantum Electron.* **QE-16,** 1329 (1980).
199. Olsson, A., and C. L. Tang. *IEEE J. Quantum Electron.* **QE-17,** 1320 (1981).
200. Dandridge, A., and R. O. Miles. *Electron. Lett.* **17,** 273 (1981).
201. Saito, S., and Y. Yamamoto. *Electron. Lett.* **17,** 325 (1981); Nilsson, O., S. Saito, and Y. Yamamoto. *Electron. Lett.* **17,** 589 (1981).
202. Goldberg, L., H. F. Taylor, A. Dandridge, J. F. Weller, and R. O. Miles. *IEEE J. Quantum Electron.* **QE-18,** 555 (1982).
203. Kikuchi, K., and T. Okoshi. *Electron. Lett.* **18,** 10 (1982).
204. Petermann, K., and G. Arnold. *IEEE J. Quantum Electron.* **QE-18,** 543 (1982).
205. Fye, D. M. *IEEE J. Quantum Electron.* **QE-18,** 1675 (1982).
206. Osmundsen, J. H., and N. Gade. *IEEE J. Quantum Electron.* **QE-19,** 465 (1983).
207. Patzak, E., H. Olesen, A. Sugimura, S. Saito, and T. Mukai. *Electron. Lett.* **19,** 938 (1983).
208. Suris, R. A., and A. A. Tager. *Sov. J. Quantum Electron.* **14,** 21 (1984).
209. Spano, P., S. Piazolla, and M. Tamburrini. *IEEE J. Quantum Electron.* **QE-20,** 350 (1984).
210. Agrawal, G. P. *IEEE J. Quantum Electron.* **QE-20,** 468 (1984).
211. Stubkjaer, K. E., and M. B. Small. *IEEE J. Quantum Electron.* **QE-20,** 472 (1984).
212. Fujita, T., S. Ishizuka, K. Fujito, H. Serizawa, and H. Sato. *IEEE J. Quantum Electron.* **QE-20,** 492 (1984).
213. Tromborg, B., J. H. Osmundsen, and H. Olesen, *IEEE J. Quantum Electron.* **QE-20,** 1023 (1984).
214. Acket, G. A., D. Lenstra, A. J. den Boef, and B. H. Verbeek. *IEEE J. Quantum Electron.* **QE-20,** 1163 (1984).
215. Agrawal, G. P., N. A. Olsson, and N. K. Dutta. *Appl. Phys. Lett.* **45,** 597 (1984).
216. Smiyan, B. I., D. P. Tregub, and B. B. Elenkrig. *Sov. Phys.-Tech. Phys.* **29,** 115 (1984).
217. Temkin, H., N. A. Olsson, J. H. Abeles, R. A. Logan, and M. B. Panish. *IEEE J. Quantum Electron.* **QE-22,** 286 (1986).
218. Lenstra, D., B. H. Verbeek, and A. J. den Boef. *IEEE J. Quantum Electron.* **QE-21,** 674 (1985).
219. Henry, C. H., and R. F. Kazarinov. *IEEE J. Quantum Electron.* **QE-22,** 294 (1986).
220. Agrawal, G. P., and T. M. Shen. *J. Lightwave Technol.* **LT-4,** 58 (1986).
221. Tkach, R. W., and A. R. Chraplyvy. *Electron. Lett.* **21,** 1081 (1985); *J. Lightwave Technol.* **LT-4,** 1655 (1986).
222. Cho, Y., and T. Umeda. *Opt. Commun.* **59,** 131 (1986).
223. Olesen, H., J. H. Osmundsen, and B. Tromborg. *IEEE J. Quantum Electron.* **QE-22,** 762 (1986).
224. Tromborg, B., H. Olesen, X. Pan, and S. Saito. *IEEE J. Quantum Electron.* **QE-23,** 1875 (1987).
225. Mork, J., B. Tromborg, and P. L. Christiansen. *IEEE J. Quantum Electron.* **QE-24,** 123 (1988).
226. Dante, G. C., P. S. Durkin, K. A. Wilson, and C. E. Moeller. *IEEE J. Quantum Electron.* **QE-24,** 2441 (1988).
227. Schunk, N., and K. Petermann. *IEEE J. Quantum Electron.* **QE-24,** 1242 (1988); *Electron. Lett.* **25,** 63 (1989); *IEEE Photon. Technol. Lett.* **1,** 49 (1989).
228. Sacher, J., W. Elsässer, and E. O. Göbel. *Phys. Rev. Lett.* **63,** 2224 (1989).

229. Tromborg, B., and J. Mork. *IEEE J. Quantum Electron.* **QE-26,** 642 (1990).
230. Tromborg, B., and J. Mork. *IEEE Photon. Technol. Lett.* **2,** 549 (1990).
231. Mork, J., J. Mark, and B. Tromborg. *Phys. Rev. Lett.* **65,** 1999 (1990).
232. Cohen, J. S., F. Wittgrefe, M. D. Hoogerland, and J. P. Woerdman. *IEEE J. Quantum Electron.* **QE-26,** 982 (1990).
233. Park, J. D., D. S. Seo, and J. G. McInerney. *IEEE J. Quantum Electron.* **QE-26,** 1353 (1990).
234. Lenstra, D. *Opt. Commun.* **81,** 209 (1991).
235. Wang, J., and K. Petermann. *IEEE J. Quantum Electron.* **QE-27,** 3 (1991).
236. Sacher, J., W. Elsässer, and E. O. Göbel. *IEEE J. Quantum Electron.* **QE-27,** 373 (1991).
237. Mork, J., B. Tromborg, and J. Mark. *IEEE J. Quantum Electron.* **QE-28,** 93 (1992).
238. Sacher, J., D. Baums, P. Panknin, W. Elsässer, and E. O. Göbel. *Phys. Rev.* **45,** 1893 (1992).
239. Dayment, J. C., and K. D. Chik. *J. Appl. Phys.* **51,** 5252 (1980).
240. Morozov, V. N., A. S. Semenov, A. B. Sergeev, and I. A. Skopin. *Sov. J. Quantum Electron.* **12,** 1513 (1982).
241. Wyatt, R., and W. J. Devlin. *Electron. Lett.* **19,** 110 (1983).
242. Agrawal, G. P., N. A. Olsson, and N. K. Dutta. *Appl. Phys. Lett.* **45,** 597 (1984).
243. Sato, H., T. Fujita, and K. Fujito. *IEEE J. Quantum Electron.* **QE-21,** 46 (1985).
244. Kanada, T. *Trans. IECE Jpn.* **68,** 180 (1985).
245. Arimoto, A., M. Ojima, N. Chinone, A. Oishi, T. Gotoh, and N. Ohnuki. *Appl. Opt.* **25,** 1398 (1986).
246. Ojima, M., A. Arimoto, N. Chinone, T. Gotoh, and K. Aiki. *Appl. Opt.* **25,** 1404 (1986).
247. Gage, E. C., and S. Beckens. *SPIE Proc. on Optical Data Storage, vol. 1316.* Bellingham: SPIE, 1990.
248. Port, M., and K. J. Ebeling. *IEEE J. Quantum Electron.* **QE-26,** 449 (1990).

Chapter 7
DISTRIBUTED-FEEDBACK SEMICONDUCTOR LASERS

7.1 INTRODUCTION

We have seen in Chapter 6 that a conventional semiconductor laser does not emit light in a single longitudinal mode. In general, the mode closest to the gain peak is most intense, and a few percent of the output power is carried by other longitudinal modes lying close to the gain peak. Furthermore, even when these side modes are reasonably suppressed under CW operation, their power content increases significantly when the laser is pulsed rapidly. Direct modulation of a semiconductor laser at frequencies in the gigahertz range is commonly employed in optical communication systems. In the presence of chromatic dispersion, the unwanted side modes limit the information transmission rate by reducing the fiber bandwidth. It is therefore desirable to devise means for a semiconductor laser to emit light predominantly in a single longitudinal mode even under high-speed modulation. Such lasers are referred to as single-frequency or single-longitudinal-mode lasers and in view of their potential application in optical communication systems were studied extensively during the 1980s.[1] They were used in commercial lightwave systems by 1990.

In conventional FP-type semiconductor lasers, the feedback is provided by facet reflections whose magnitude remains the same for all longitudinal modes. The only longitudinal-mode discrimination in such a laser is provided by the gain spectrum itself. However, since the gain spectrum is usually much wider than the longitudinal-mode spacing, the resulting mode discrimination is poor. One way of improving the mode selectivity is to make the feedback frequency-dependent so that the cavity loss is different for different longitudinal modes. Two mechanisms have been found useful in this respect and are known as the distributed feedback and the coupled-cavity mechanisms. Distributed-feedback (DFB) lasers are described in this chapter. Chapter 8 is devoted to coupled-cavity lasers.

As the name implies, the feedback necessary for the lasing action in a DFB laser is not localized at the cavity facets but is distributed throughout the cavity length. This is achieved through the use of a grating etched so that the thickness of one layer (participating in the heterostructure) varies periodically along the cavity length. The resulting periodic perturbation of the refractive index provides feedback by means of backward Bragg scattering, which couples the forward- and backward-propagating waves. Mode selectivity of the DFB mechanism results from the Bragg condition. According to the Bragg condition, coherent coupling between counterpropagating waves occurs only for wavelengths such that the grating period $\Lambda = m\lambda_m/2$ where λ_m is the wavelength inside the laser medium and the integer m is the order of Bragg diffraction induced by the grating.[2] By choosing Λ appropriately, such a device can be made to provide distributed feedback only at selected wavelengths.

Kogelnik and Shank[3] were the first to observe the lasing action in a periodic structure that utilized the DFB mechanism. Since then, DFB semiconductor lasers have attracted considerable attention both experimentally[4–32] and theoretically.[33–60] Although most of the early work[4–14] related to GaAs lasers, the need for a single-frequency semiconductor laser operating at the

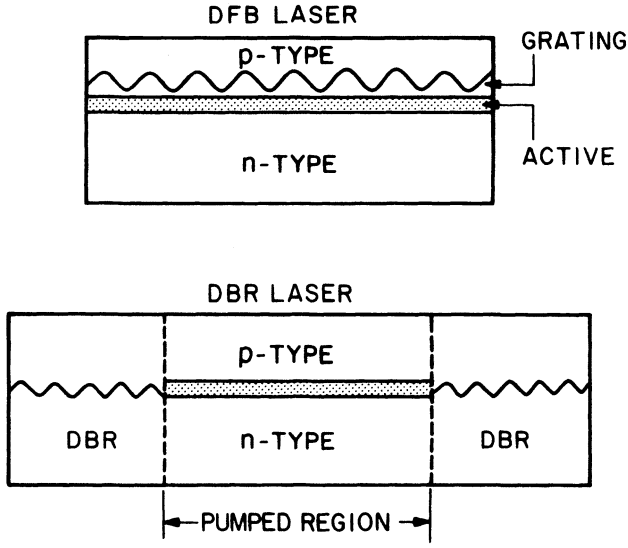


Fig. 7.1 Schematic illustration of distributed-feedback (DFB) and distributed Bragg reflector (DBR) semiconductor lasers. Different refractive indices on opposite sides of the grating result in a periodic index perturbation that is responsible for the distributed feedback. Shaded area shows the active region of the device.

minimum-fiber-loss wavelength of 1.55 μm has resulted in the development of InGaAsP DFB lasers.[15–32] From the viewpoint of device operation, semiconductor lasers employing distributed feedback can be classified into two broad categories: DFB lasers and *distributed Bragg reflector* (DBR) lasers.[39] These are shown schematically in Fig. 7.1. In DBR lasers,[10,11] the grating is etched near the cavity ends and distributed feedback does not take place in the central active region. The unpumped corrugated end regions act as effective mirrors whose reflectivity is of DFB origin and is therefore wavelength-dependent.[24]

This chapter is organized as follows. In Sec. 7.2 we briefly discuss the fabrication details of the DFB structures used for InGaAsP lasers. Section 7.3 deals with the theory of DFB lasers. Starting with the coupled-wave equations, we obtain the wavelength and the threshold gain of the DFB longitudinal modes for the general case in which both the etched grating and the cleaved facets participate in providing the feedback necessary for the lasing action. The threshold gain of the lowest-loss mode is used to discuss such light-current characteristics as the threshold current and the differential quantum efficiency. Section 7.4 describes the performance characteristics of a DFB laser, considering both CW and pulsed modes of operation. Section 7.5 is devoted to DBR lasers. Finally, in Sec. 7.6 we discuss recent transmission experiments employing DFB semiconductor lasers.

7.2 DFB LASER STRUCTURES

In Chapter 5 we discussed various semiconductor laser structures that are commonly used. Any of these structures can be employed to make a DFB semiconductor laser by etching a grating onto one of the layers. The direct etching of the active layer is generally not preferred since it can increase the rate of nonradiative recombination by introducing defects in the active region. This would affect the device performance by producing a higher threshold current.[6] The grating is therefore etched onto one of the cladding layers. Since only the evanescent field associated with the fundamental transverse mode interacts with the grating, it is clear that the exact location of the grating with respect to the active layer and the corrugation depth are critical in determining the effectiveness of the grating. The grating period Λ is determined by the device wavelength in the medium and the order of Bragg diffraction used for the distributed feedback. The Bragg condition for the mth-order coupling between the forward- and backward-propagating waves is[2]

$$\Lambda = m\lambda/2\bar{\mu} \tag{7.2.1}$$

where $\bar{\mu}$ is the effective-mode index and $\lambda/\bar{\mu}$ is the wavelength inside the

medium. For a 1.55-μm InGaAsP laser, $\Lambda \cong 0.23$ μm if we use $m = 1$ (first-order grating), and a typical value $\bar{\mu} \cong 3.4$. This value doubles if a second-order grating is used. Both first-order and second-order gratings are employed to fabricate InGaAsP DFB lasers.

Two techniques have been used for the formation of a grating with submicrometer periodicity. In the holographic technique, optical interference is used to form a fringe pattern on the photoresist deposited on the wafer surface. Two optical beams obtained from an ultraviolet laser (e.g., an HeCd laser) are made to interfere on the wafer; the grating period can be adjusted by changing the angle between the two beams. In the electron-beam lithographic technique, an electron beam scans the wafer surface and "writes" the desired pattern on the electron-beam resist. Both methods generally use chemical etching to produce the grating corrugations, with the patterned resist acting as a mask. Once the grating has been etched on the substrate or on an epitaxial layer, the wafer can be processed in the usual way to obtain a specific laser structure (see Chapter 5).

The choice of whether to use the upper or the lower cladding layer for forming the grating requires careful consideration. In practice the device wavelength λ may vary from wafer to wafer because of variations in the location of the gain peak that are due to variations in alloy composition. The placement of the grating on the upper cladding layer has the advantage that the grating period Λ can be adjusted after identifying the effective-mode index $\bar{\mu}$ and the wavelength λ corresponding to the gain peak. At the same time, however, this choice has the disadvantage that it requires an additional epitaxial step.

As discussed in Sec. 2.5, semiconductor lasers can be classified into three broad categories based on the lateral-mode guiding mechanism: gain-guided, weakly index-guided, and strongly index-guided. Although the initial work[4–10] on GaAs DFB lasers related to gain-guided structures, gain-guided InGaAsP DFB lasers are unattractive; many undesirable characteristics of gain-guided lasers worsen in the long-wavelength regime (see Sec. 5.3). Strongly index-guided devices such as the buried-heterostructure type[12] are most suitable for distributed feedback and are commonly used to make 1.55-μm InGaAsP DFB lasers.[17–27].

Figure 7.2 shows schematics and a photomicrograph of a 1.5-μm InGaAsP double-channel planar buried-heterostructure (DCPBH) laser. These DFB lasers have attracted considerable attention[21] at both 1.3-μm and 1.55-μm wavelengths. Their performance is nearly comparable to the conventional FP lasers as far as characteristics such as threshold current, output power, and high-temperature performance are concerned. At the same time, the use of the DFB mechanism leads to a high degree of longitudinal-mode selectivity, and side modes of such lasers are often suppressed as much as

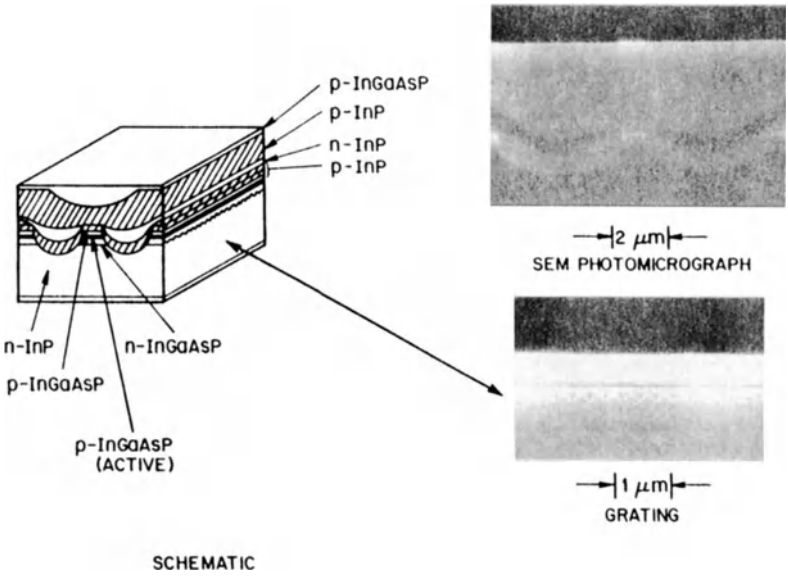


Fig. 7.2 Schematic illustration of an InGaAsP DCPBH DFB semiconductor laser. The top and bottom SEM photomicrographs show the planar active layer and the etched grating, respectively. The second-order grating was formed on the substrate.

30–35 dB relative to the main lasing mode. In Fig. 7.3 we have shown schematically a 1.55-μm InGaAsP ridge waveguide DFB laser.[27] These lasers also exhibit a side-mode suppression ratio exceeding 30–35 dB. Their threshold current is generally higher than the DCPBH-DFB laser; this is mainly due to a larger mode volume and to some extent because of current spreading and carrier diffusion occurring in any weakly index-guided structure.

7.3 THEORY

To understand the operating characteristics of a DFB laser, it is first necessary to consider wave propagation in periodic structures. Grating-induced dielectric perturbation leads to a coupling between the forward- and backward-propagating waves associated with a particular laser mode. Two equivalent approaches have been used to analyze the device behavior. In the coupled-wave approach,[33,36] a set of two equations corresponding to the counterpropagating forward and backward waves is solved subject to the specific boundary conditions applicable to a given device. In an alternative but equivalent approach,[39] one first obtains the Bloch-type

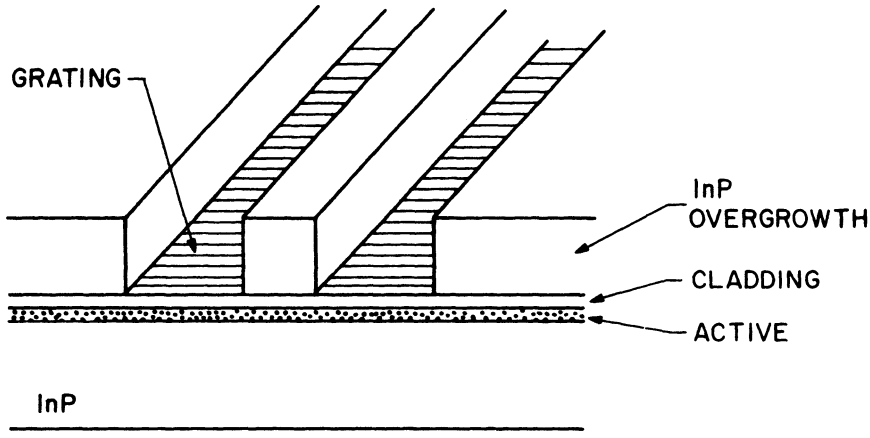


Fig. 7.3 Schematic illustration of an InGaAsP ridge waveguide (DFB) laser. The *p*-type upper cladding layer provides the grating corrugation. (After Ref. 27)

eigensolutions by assuming an infinitely long structure; one then applies these solutions to a specific device. In both approaches the objective is to obtain the threshold gain and the longitudinal-mode spectrum of a DFB laser. We follow an approach that combines the essential features of both. Starting from the wave equation, we derive the coupled-wave equations whose general solution, without reference to any particular boundary conditions, is then used to construct Bloch-type eigensolutions. These are then used to obtain the DFB longitudinal modes and their threshold gains for the general case where both the grating and the cleaved facets contribute to the feedback.

7.3.1 Coupled-Wave Equations

The starting point of the analysis is the time-independent wave equation [Eq. (2.2.19)]

$$\nabla^2 \mathbf{E} + \epsilon(x, y, z) k_0^2 \mathbf{E} = 0 \tag{7.3.1}$$

where $k_0 = \omega/c$ and ω is the mode frequency. The main difference from the waveguide-mode theory presented in Sec. 2.5 is that in the grating region, ϵ is a periodic function of z. It is useful to write

$$\epsilon(x, y, z) = \bar{\epsilon}(x, y) + \Delta\epsilon(x, y, z) \tag{7.3.2}$$

where $\bar{\epsilon}(x, y)$ is the average value of ϵ and the dielectric perturbation $\Delta\epsilon$ is nonzero only over the grating region whose thickness is equal to the corrugation depth. In the absence of a grating ($\Delta\epsilon = 0$), the general solution of Eq. (7.3.1) takes the form

$$\mathbf{E}(x, y, z) = \hat{\mathbf{x}}U(x, y)[E_f \exp(i\beta z) + E_b \exp(-i\beta z)] \tag{7.3.3}$$

where $\hat{\mathbf{x}}$ is the unit vector (along the junction plane for the TE mode). The field distribution corresponding to a specific waveguide mode is obtained by solving

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + [\bar{\epsilon}(x, y)k_0^2 - \beta^2]U = 0 \tag{7.3.4}$$

for a given device structure, where β is the mode-propagation constant. We assume that the device supports only the fundamental waveguide mode and the corresponding $U(x, y)$ and β have been obtained following the analysis of Sec. 2.5. The mode-propagation constant β is complex:

$$\beta = \bar{\mu}k_0 - i\bar{\alpha}/2 \tag{7.3.5}$$

where $\bar{\mu}$ is the effective-mode index and $\bar{\alpha}$ is the mode-gain coefficient given by

$$\bar{\alpha} = \Gamma g - \alpha_{\text{int}} \tag{7.3.6}$$

Γ is the confinement factor, g is the active-region gain, and α_{int} accounts for the internal loss. These relations are identical to Eqs. (2.5.37) and (2.5.48) except that in those equations $\bar{\alpha}$ represents the mode-absorption coefficient rather than the mode-gain coefficient.

In the presence of dielectric perturbation $\Delta\epsilon$, the amplitudes E_f and E_b of the forward and backward waves becomes z-dependent because a periodic structure produces Bragg diffraction. A major simplification occurs if we assume that the spatial distribution $U(x, y)$ is unaffected by the weak perturbation $\Delta\epsilon$. We substitute Eq. (7.3.3) in (7.3.1) and allow for slow axial variations of E_f and E_b. Multiplying the resulting equation with $U(x, y)$ and integrating over x and y, we obtain

$$\frac{dE_f}{dz}\exp(i\beta z) - \frac{dE_b}{dz}\exp(-i\beta z) = \frac{ik_0^2}{2\beta V}\iint \Delta\epsilon(x, y, z)U^2(x, y) \times [E_f \exp(i\beta z) + E_b \exp(-i\beta z)]\, dx\, dy \tag{7.3.7}$$

where $V = \int\int U^2(x, y)\, dx\, dy$ and is a normalization constant. Since $\Delta\epsilon$ is periodic in z with the grating period Λ, it can be expanded in a Fourier series to

$$\Delta\epsilon(x, y, z) = \sum_{l \neq 0} \Delta\epsilon_l(x, y) \exp[i(2\pi/\Lambda)lz]. \tag{7.3.8}$$

We substitute Eq. (7.3.8) into (7.3.7) and equate the coefficients of $\exp(\pm i\beta z)$ on both sides. We assume for simplicity that $\Delta\epsilon$ is real. Collecting only the terms that are approximately phase-matched, we obtain

$$\frac{dE_f}{dz} = i\kappa E_b \exp(-2i\Delta\beta z) \tag{7.3.9}$$

$$\frac{dE_b}{dz} = -i\kappa^* E_f \exp(2i\Delta\beta z) \tag{7.3.10}$$

where

$$\Delta\beta = \beta - m\pi/\Lambda = \beta - \beta_0 \tag{7.3.11}$$

is the phase mismatch and was assumed to be smallest for the term $l = m$ in Eq. (7.3.8). The contribution of other terms has been neglected because of their large phase mismatch. The coupling coefficient κ is given by

$$\kappa = \frac{k_0^2}{2\beta} \frac{\displaystyle\int\int \Delta\epsilon_m(x, y) U^2(x, y)\, dx\, dy}{\displaystyle\int\int U^2(x, y)\, dx\, dy}. \tag{7.3.12}$$

It shows that the Fourier component $\Delta\epsilon_m$, for which the Bragg condition ($\beta \cong m\pi/\Lambda$) is approximately satisfied, couples the forward- and backward-propagating waves. The magnitude of κ is governed by the weighted average of $\Delta\epsilon_m$ with the mode intensity over the grating region. In general, κ may be complex even when $\Delta\epsilon$ is real. This is the case, for example, for gain-guided lasers[53] where $U(x, y)$ is complex because of the wave-front curvature. The small imaginary part of β in Eq. (7.3.5) also makes κ complex. However, this is a minor effect, and κ can be assumed to be real for index-guided lasers by replacing β with $\bar{\mu}k_0$ in Eq. (7.3.12).

Instead of using complex β as the propagation constant in Eq. (7.3.3), it is more convenient to use the Bragg wave number $\beta_0 = m\pi/\Lambda$ and write the

axially varying field as

$$E(z) = A(z)\exp(i\beta_0 z) + B(z)\exp(-i\beta_0 z) \tag{7.3.13}$$

where

$$A = E_f \exp(i\Delta\beta z) \qquad \text{and} \qquad B = E_b \exp(-i\Delta\beta z). \tag{7.3.14}$$

The coupled-wave equations for A and B then become

$$\frac{dA}{dz} = i\Delta\beta A + i\kappa B \tag{7.3.15}$$

$$-\frac{dB}{dz} = i\Delta\beta B + i\kappa A \tag{7.3.16}$$

where we have used Eqs. (7.3.9)–(7.3.11). A general solution of these equations takes the form

$$A(z) = A_1 \exp(iqz) + A_2 \exp(-iqz) \tag{7.3.17}$$

$$B(z) = B_1 \exp(iqz) + B_2 \exp(-iqz) \tag{7.3.18}$$

where q is the complex wave number to be determined from the boundary conditions. The constants A_1, A_2, B_1, and B_2 are interdependent. If we substitute the general solution into Eqs. (7.3.15) and (7.3.16) and equate the coefficients of $\exp(\pm iqz)$, we obtain

$$(q - \Delta\beta)A_1 = \kappa B_1 \qquad (q + \Delta\beta)B_1 = -\kappa A_1 \tag{7.3.19}$$

$$(q - \Delta\beta)B_2 = \kappa A_2 \qquad (q + \Delta\beta)A_2 = -\kappa B_2. \tag{7.3.20}$$

These relations are satisfied with nonzero values of A_1, A_2, B_1, and B_2 if the possible values of q obey the following dispersion relation (obtained by setting the determinant of the coefficient matrix to be zero):

$$q = \pm[(\Delta\beta)^2 - \kappa^2]^{1/2}. \tag{7.3.21}$$

The plus and minus signs correspond to forward- and backward-propagating waves, respectively. Furthermore, we can define the DFB reflection coefficient as

$$r(q) = \frac{q - \Delta\beta}{\kappa} = -\frac{\kappa}{q + \Delta\beta} \tag{7.3.22}$$

and eliminate A_2 and B_1 in Eqs. (7.3.17) and (7.3.18) in favor of $r(q)$. The general solution of the coupled-wave equations then becomes

$$A(z) = A_1 \exp(iqz) + r(q)B_2 \exp(-iqz) \tag{7.3.23}$$

$$B(z) = B_2 \exp(-iqz) + r(q)A_1 \exp(iqz). \tag{7.3.24}$$

Since $r = 0$ if $\kappa = 0$, it is evident that $r(q)$ represents the fraction of the forward-wave amplitude that is reflected back toward the backward wave and vice versa. The sign ambiguity in the expression (7.3.21) can be resolved by choosing the sign of q such that $|r(q)| \leq 1$. Mode selectivity of DFB lasers stems, as will become clear, from the dependence of r on q. The eigenvalue q and the ratio B_2/A_1 are determined by applying the appropriate boundary conditions at the laser facets.

Before considering a finite-length DFB laser, it is instructive to discuss wave propagation in an infinite-length periodic structure. The dispersion relation (7.3.21) governs the possible q values in such a medium. In general, q is complex since $\Delta\beta$ is complex in the presence of gain or loss associated with the medium. If we neglect $\bar{\alpha}$ in Eq. (7.3.5), $\Delta\beta = \beta - \beta_0$ and is real. According to Eq. (7.3.21), q is purely imaginary when $|\Delta\beta| \leq \kappa$, and the medium cannot support a propagating wave. Consequently, an infinite periodic structure has a stop band; i.e., only waves of certain frequencies for which the propagation constant β satisfies the inequality $|\beta - \beta_0| > \kappa$ can propagate inside the medium. It will be seen later that the stop band exists even for a finite DFB laser, although under specific conditions the device may support a "gap mode" that lies inside the stop band.

An infinite periodic structure is conceptually similar to a one-dimensional crystal, and the concept of stop band is analogous to the band gap occurring in a crystal. The analogy can be extended further by noting that the field $E(z)$ in Eq. (7.3.13) is in the form of a Bloch wave. This is seen more clearly if we substitute A and B from Eqs. (7.3.23) and (7.3.24) into Eq. (7.3.13) and rearrange the terms to obtain

$$\begin{aligned} E(z) = {} & A_1[\exp(i\beta_0 z) + r(q)\exp(-i\beta_0 z)]\exp(iqz) \\ & + B_2[\exp(-i\beta_0 z) + r(q)\exp(i\beta_0 z)]\exp(-iqz). \end{aligned} \tag{7.3.25}$$

The expression inside each pair of square brackets is a periodic function of z with periodicity Λ. The Bloch-type eigenmodes are expected in view of the periodic variation of the dielectric constant and could have been obtained directly by applying Floquet's theorem to Eq. (7.3.1).[39]

7.3.2 Longitudinal Modes and Threshold Gain

Threshold Condition. For a finite-length DFB laser, the boundary conditions at the facets are satisfied only for discrete values of q. The real and imaginary parts of q give, respectively, the DFB longitudinal-mode frequency and the corresponding threshold gain. Consider a laser of length L. Since L is not necessarily an exact multiple of the grating period Λ, the last period of the grating close to the facet is generally not complete. Even though no significant distributed feedback occurs over these incomplete grating petiods, the phase shift in this region plays an important role in determining DFB laser characteristics and should be properly accounted for. A simple way to do this is to assume that the effective-facet-reflection coefficient is complex, i.e.,

$$r_j = R_j^{1/2} \exp(i\phi_j) \qquad j = 1, 2 \tag{7.3.26}$$

where R_j is the facet reflectivity and ϕ_j is the round-trip phase shift encountered by the field in traveling the extra distance (a fraction of Λ) corresponding to the last incomplete grating period. In applying the boundary conditions, one can assume that the mirrors are located at the last complete grating period. The constant phase shift of about 180° occurring during facet reflections can also be incorporated into ϕ_j. A characteristic feature of a DFB semiconductor laser is that the phase shifts ϕ_1 and ϕ_2 are expected to vary from device to device since the relative distance between the cleaved facet and the last complete grating period (a small fraction of 1 μm) is uncontrollable at present except through the use of controlled etching or coating of individual facets.

The boundary conditions at the two facets are

$$A(0) = r_1 B(0) \qquad \text{and} \qquad B(L) = r_2 A(L). \tag{7.3.27}$$

If we use them in the general solution given by Eqs. (7.3.23) and (7.3.24), we obtain the following two homogeneous equations for the unknown constants A_1 and B_2:

$$(r_1 - r)B_2 - (1 - rr_1)A_1 = 0 \tag{7.3.28}$$

$$(r_2 - r)\exp(2iqL)A_1 - (1 - rr_2)B_2 = 0. \tag{7.3.29}$$

These equations have a consistent nontrivial solution only for values of q

satisfying the eigenvalue equation[42,57]

$$\left(\frac{r_1 - r}{1 - rr_1}\right)\left(\frac{r_2 - r}{1 - rr_2}\right)\exp(2iqL) = 1. \tag{7.3.30}$$

This is the threshold condition for DFB lasers. It is similar to the threshold condition obtained for an FP laser (see Sec. 2.3) and reduces to the latter if the DFB contribution is neglected by setting r to zero. On the other hand, if facet reflectivities are neglected by setting r_1 and r_2 to zero, we obtain

$$r^2 \exp(2iqL) = 1 \tag{7.3.31}$$

implying that r is the effective-reflection coefficient for a DFB laser.

The DFB eigenvalues q are obtained from Eq. (7.3.30) after noting that r as given by Eq. (7.3.22) also depends on q. Each eigenvalue q is complex and can be used to calculate

$$\Delta\beta = \pm(q^2 + \kappa^2)^{1/2} \tag{7.3.32}$$

after using Eq. (7.3.21). From Eqs. (7.3.5) and (7.3.11) the real and imaginary parts of the relation

$$\Delta\beta = \delta - i\bar{\alpha}/2 \tag{7.3.33}$$

yield the mode detuning

$$\delta = \bar{\mu}k_0 - \beta_0 = \frac{-2\pi\mu_g}{\lambda^2}\Delta\lambda \tag{7.3.34}$$

and the threshold gain $\bar{\alpha}$. In Eq. (7.3.34), $\Delta\lambda$ is the deviation of the mode wavelength from the Bragg wavelength ($\lambda_0 = 2\bar{\mu}\Lambda/m$), and care is taken to account for the chromatic dispersion (wavelength variation of $\bar{\mu}$). Note that $\bar{\alpha}$ is the power gain. In previous work[33,42] the amplitude gain ($\alpha = \bar{\alpha}/2$) has often been used. The threshold material gain is determined by using Eq. (7.3.6) and is given by

$$g_{\text{th}} = \frac{\bar{\alpha} + \alpha_{\text{int}}}{\Gamma}. \tag{7.3.35}$$

The values of δ corresponding to different eigenvalues indicate how far that longitudinal mode is displaced from the Bragg wavelength λ_0. According to Eqs. (7.3.32)–(7.3.34), the DFB longitudinal modes are situated

symmetrically with respect to λ_0. In contrast to FP lasers, however, different modes have different threshold gains as determined by the eigenvalue equation (7.3.30) and by Eq. (7.3.35).

The physical interpretation of the threshold condition (7.3.35) is clear. The mode gain $\bar{\alpha}$ can be thought of as the gain required to overcome both the distributed and the facet losses; and it plays the same role as the mirror loss

$$\alpha_m = \frac{1}{2L} \ln\left(\frac{1}{R_1 R_2}\right) \tag{7.3.36}$$

for an FP laser. The thresholds for DFB and FP lasers can thus be compared through $\bar{\alpha}$ and α_m provided the internal loss α_{int} is assumed to be identical for both kinds of lasers. Using cleaved-facet reflectivities $R_1 = R_2 \cong 0.32$, we obtain $\alpha_m L \cong 1.1$, and a DFB laser would have a lower threshold than an FP laser if $\bar{\alpha}L < 1.1$ for the lowest-threshold mode.

Nonreflecting Facets. Numerical solution of the eigenvalue equation (7.3.30) is generally required to obtain the longitudinal-mode frequencies and their respective threshold gains. For some physical insight, it is instructive to consider the ideal case in which the facet reflectivities R_1 and R_2 are set to zero. Experimentally it can be realized by etching, coating, or burying the laser facets.[22,25] Equations (7.3.22) and (7.3.31) can be combined to yield the transcendental equation[33]

$$\kappa \sin(qL) = \pm iq. \tag{7.3.37}$$

The DFB modes are obtained by solving this equation for a given value of the coupling coefficient κ. To make the results independent of the cavity length L, it is useful to consider the dimensionless parameter κL and plot $\bar{\alpha}L$ versus δL.

Figure 7.4 shows the mode gain $\bar{\alpha}L$ versus the mode detuning δL for the three modes closest to the Bragg wavelength λ_0. Since the spectrum is symmetric with respect to λ_0, only half of the spectrum is shown. Each solid curve is obtained by varying κL. For a given value of κL, the intersections of the corresponding dashed line with the solid lines provide the values for mode gain and mode frequency. In particular, the mode closest to the Bragg wavelength ($m = 1$) has the lowest threshold gain and would reach the threshold first. The gain difference between the modes $m = 2$ and $m = 1$ is a measure of the mode selectivity and determines the extent to which other side modes are suppressed. The horizontal dashed line in Fig. 7.4 corresponds to an FP laser (which has identical gain for all modes)

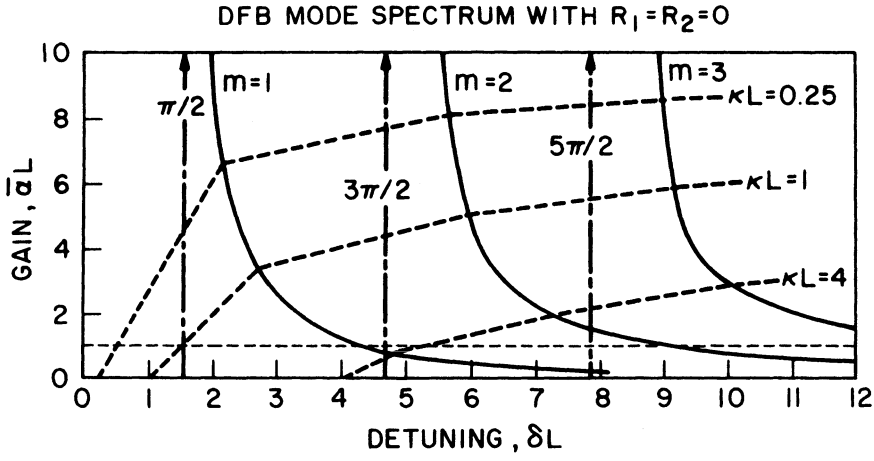


Fig. 7.4 Normalized threshold gain $\bar{\alpha}L$ and detuning δL (from the Bragg wavelength) of three longitudinal modes of a DFB laser with no end reflections ($R_1 = R_2 = 0$). The coupling coefficient κ varies continuously along the solid curves. The intersection of each curve with a dashed line for a given κL provides the values of $\bar{\alpha}L$ and δL. The horizontal dashed line shows the threshold gain of an FP laser with a cleaved-facet reflectivity of 0.32. Vertical lines show mode positions in the limit as κL approaches 0. (After Ref. 33)

and shows that the DFB mechanism would lead to a lower threshold only if $\kappa L > 3.2$.

A noteworthy feature of a DFB laser with nonreflecting facets is that, because of the symmetric nature of the DFB mode spectrum, the two modes closest to the Bragg frequency (one on each side) have the lowest threshold gain and would lase simultaneously. The frequency or wavelength separation between these two degenerate modes corresponds to the stop band discussed earlier. Simultaneous oscillation of the two DFB modes separated by the stop band has been observed.[22,25] The width of the stop band is related to the coupling coefficient κ and can be used for its estimation. This can be seen clearly in the low-gain limit.[33] When the trivial solution $q = 0$ of Eq. (7.3.37) is used in Eqs. (7.3.32) and (7.3.33), we find that $\delta \cong \pm\kappa$ for the mode with lowest threshold gain. Thus in the low-gain limit, the width of the stop band is approximately 2κ.

The eigenvalue equation (7.3.31) can be solved approximately in an analytic form for the high-gain case[33] that is likely to occur under weak coupling, i.e., when $\kappa L \ll 1$. Using Eqs. (7.3.21) and (7.3.22), $q \cong \Delta\beta$ and the DFB reflection coefficient $r \cong -\kappa/(2\Delta\beta)$. With the help of Eqs. (7.3.31) and (7.3.33), the eigenvalue equation becomes

$$\kappa^2 \exp[(2i\delta + \bar{\alpha})L] = -(2i\delta + \bar{\alpha})^2. \tag{7.3.38}$$

Equating the modulus and the phase on the two sides, we obtain

$$\bar{\alpha}^2 + 4\delta^2 = \kappa^2 \exp(\bar{\alpha}L) \tag{7.3.39}$$

$$\delta L = (m - \tfrac{1}{2})\pi + \tan^{-1}\left(\frac{2\delta}{\bar{\alpha}}\right) \tag{7.3.40}$$

where the integer m labels various longitudinal modes (see Fig. 7.4). If we neglect the last term in Eq. (7.3.40) and use Eq. (7.3.34), the longitudinal-mode spacing is given by

$$\Delta\lambda = \lambda^2/(2\mu_g L). \tag{7.3.41}$$

This value is identical to that obtained for FP lasers in Sec. 2.3. Mode selectivity of DFB lasers is evident from Eq. (7.3.39). As δ increases, $\bar{\alpha}$ increases for a fixed value of κ, and higher-order modes require higher threshold gains.

Cleaved Facets. From a practical viewpoint it is important to consider the effect of facet reflections on the DFB mode spectrum.[42] This is accomplished by using Eq. (7.3.26) in (7.3.30) and solving for the eigenvalue q. As mentioned before, the phases ϕ_1 and ϕ_2 are arbitrary and would vary from device to device. The threshold characteristics of DFB lasers with cleaved facets are therefore susceptible to considerable variations due to uncontrollable phase shifts. Such variations have been observed experimentally.[25,26] In the following discussion we examine a few representative cases to show the new qualitative features arising from facet reflections.

We first consider the case where $R_1 = R_2 = 0.32$ in order to include reflections from both cleaved facets. This is the case commonly encountered. Figure 7.5 shows the mode gain $\bar{\alpha}L$ and the detuning δL for six DFB modes as ϕ_1 is varied over its 2π range for a fixed value of ϕ_2. Two curves for different values of κL are shown. The horizontal dashed line shows the FP laser threshold in the absence of distributed feedback ($\kappa = 0$). The open circles denote the location of DFB modes for $\phi_1 = 0$. As ϕ_1 increases, the modes shift to the left on the curve. Their position for the specific case where $\phi_1 = \pi$ is marked by closed circles. Arrows denote the modes for the intermediate case where $\phi_1 = \pi/2$.

Several conclusions can be drawn from Fig. 7.5. It is evident that facet reflections break the gain degeneracy that occurs in the absence of reflections for the two modes separated by the stop band. Which of the two modes has the lower threshold gain depends on the relative values of phases ϕ_1 and ϕ_2. For the case where $\phi_2 = \pi/2$ shown in Fig. 7.5, the mode on the low-frequency or long-wavelength side with respect to the Bragg wavelength

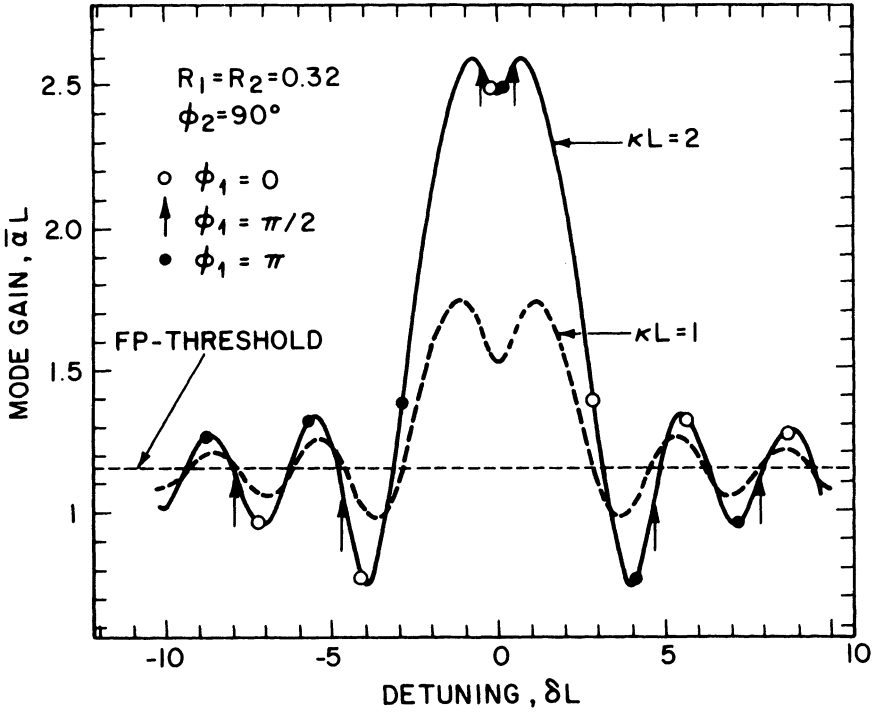


Fig. 7.5 Normalized threshold gain $\bar{\alpha}L$ and detuning δL of the six longitudinal modes of a DFB laser with two cleaved facets. Continuous curves are obtained by changing ϕ_1 over its entire range from 0 to 2π. DFB modes for three specific phases ($\phi_1 = 0$, $\pi/2$, and π) are marked on the solid curve. The horizontal dashed line shows the threshold gain of an FP laser with cleaved facets (no distributed feedback).

(negative value of δ) has a lower gain for $-\pi/2 \leq \phi_1 \leq \pi/2$, while the reverse occurs (positive values of δ) for $\pi/2 \leq \phi_1 \leq 3\pi/2$. The mode spectrum is degenerate for $\phi_1 = \pi/2$ and $\phi_1 = 3\pi/2$. Arrows in Fig. 7.5 mark the location of degenerate modes for $\phi_1 = \pi/2$. Another feature arising from facet reflections is the existence of a high-gain mode inside the stop band.[42]

It is of interest to compare the threshold of the DFB mode that has the lowest threshold gain with the threshold of an FP laser. Figure 7.5 shows that the DFB threshold is always lower than the FP threshold and that the gain margin increases with the coupling coefficient. It should be stressed, however, that the answer to the question of whether the DFB or FP mode would lase in a given device also depends on the material-gain spectral profile. The DFB mode would lase if the gain peak occurred in the vicinity of the Bragg wavelength. When the gain peak and Bragg wavelength are significantly far

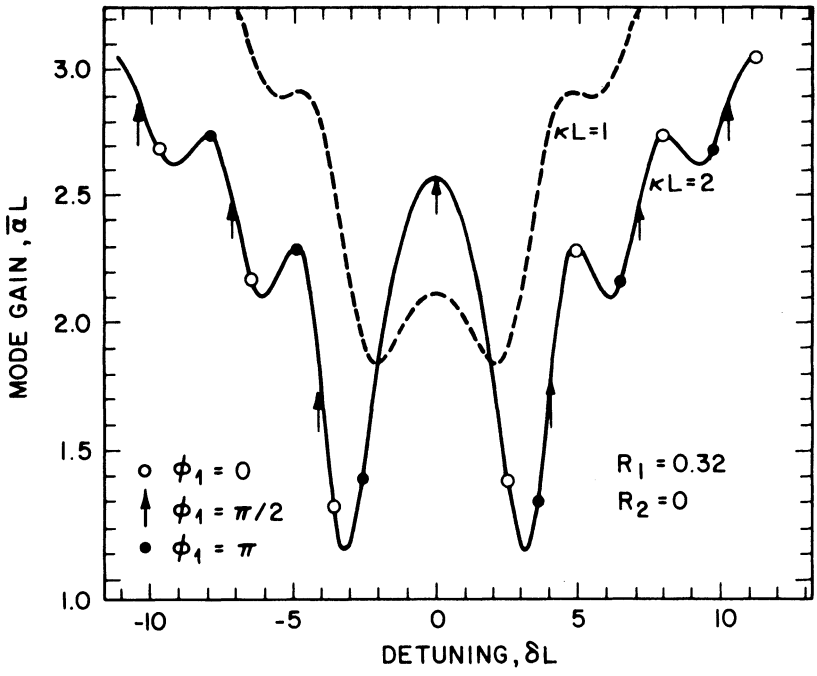


Fig. 7.6 Same plot as for Fig. 7.5 except that the case of a DFB laser with one cleaved facet and one nonreflecting facet is considered.

apart, an FP mode close to the gain peak might have an overall lower threshold gain and reach threshold first.

The case of one reflecting facet and one nonreflecting facet has also drawn considerable attention.[22,25] It is shown in Fig. 7.6 for the case where $R_2 = 0$ and $R_1 = 0.32$. Similar to Fig. 7.5, the phase ϕ_1 is varied over its entire 2π range. The qualitative features are similar to the case of two cleaved facets. The most significant difference occurs for the higher-order DFB modes, which have much higher thresholds compared with those shown in Fig. 7.5. For a given κL, the threshold of the lowest-gain mode is also higher when compared with the case of two cleaved facets. A comparison of Figs. 7.5 and 7.6 shows that the effect of reducing the reflectivity of one facet is to improve the mode selectivity at the expense of a threshold increase.

It is clear from Figs. 7.5 and 7.6 that facet reflections give rise to a mode in the stop band.[42] This mode generally has a high threshold and is sometimes referred to as the *gap mode*. The reason for the high threshold is that the gap mode has a considerable amplitude only in the vicinity of the reflecting laser facet. However, it is possible to design lasers in which the gap

mode has the lowest threshold. For example, this can be accomplished by making one facet high-reflecting and the other facet low-reflecting.[58] If R_1 in Fig. 7.6 is increased to 0.9, one finds that the peak at $\delta = 0$ is replaced by a broad minimum.

From a conceptual point of view, such a laser is essentially a folded version of a DFB laser with two low-reflecting ends and in which the grating is phase-shifted in the center. Such phase-shifted gratings were originally proposed by Haus and Shank[47] and drew considerable attention[61–66] during the 1980s in relation to 1.55-μm InGaAsP DFB lasers. The optimum phase shift is $\pi/2$, for which the gap mode occurs exactly at the Bragg wavelength ($\delta = 0$). Several different approaches have been followed to introduce the required phase shift into the middle of the laser cavity. In one approach, electron-beam lithography was used to change the phase at the central groove.[61] Simultaneous holographic exposure of positive and negative photoresists has also been used for this purpose.[65] Alternatively, a small uncorrugated section (typical length of ~ 10 μm) is introduced into the center such that the optical phase changes by $\pi/2$ during wave propagation over this region.[64] In another approach, the phase shift is introduced indirectly by using a nonuniform stripe width.[62,63] A different active-region width over a central region (typical length of ~ 60 μm) changes the effective mode index, which in turn changes the optical phase. All of these schemes have been used to demonstrate lasing action near the Bragg wavelength.

7.3.3 Coupling Coefficient

We have seen in the previous section that the threshold gain of a DFB laser depends on the magnitude of the coupling coefficient κ given by Eq. (7.3.12). The numerical value of κ depends on such grating parameters as shape, depth, and period of the corrugation. Furthermore, since the coupling occurs as a result of the interaction between the grating and the evanescent part of the transverse mode, κ is also affected by the thickness and composition of the active and cladding layers. The evaluation of κ for GaAs lasers has been discussed extensively,[13,45] and the techniques used there have been applied for InGaAsP lasers as well.[20,57] In this section we discuss briefly how κ varies with various device parameters.

We assume that waveguide dimensions are chosen such that the DFB semiconductor laser supports only the fundamental lateral and transverse modes. Then $U(x, y) = \psi(x)\phi(y)$ in Eq. (7.3.12), where $\psi(x)$ and $\phi(y)$ are the corresponding field distributions introduced in Sec. 2.5 of Chapter 2. If we assume that the grating is laterally uniform so that the dielectric perturbation $\Delta\epsilon_m$ is independent of x, the integration over x in Eq. (7.3.12) is readily performed and we obtain

$$\kappa = \frac{k_0}{2\bar{\mu}} \frac{\int_{-\infty}^{\infty} \Delta\epsilon_m(y)\phi^2(y)\,dy}{\int_{-\infty}^{\infty} \phi^2(y)\,dy} \tag{7.3.42}$$

where we used $\beta = \bar{\mu}k_0$ and $\bar{\mu}$ is the effective-mode index. In some devices the corrugation depth may be laterally nonuniform. For example, this may happen in a weakly index-guided ridge-waveguide-type device if the grating corrugation is partially filled during the ridge overgrowth.[30] Equation (7.3.42) is then used separately in the two regions to obtain κ_1 and κ_2, and κ is obtained by a weighted average:[67]

$$\kappa = \Gamma_L \kappa_1 + (1 - \Gamma_L)\kappa_2 \tag{7.3.43}$$

where Γ_L is the lateral confinement factor. This equation can be easily obtained using the general expression (7.3.12) for κ.

In Eq. (7.3.42) the Fourier coefficient $\Delta\epsilon_m$ for the mth-order Bragg diffraction is obtained using Eq. (7.3.8) and is given by

$$\Delta\epsilon_m(y) = \frac{1}{\Lambda} \int_0^{\Lambda} \Delta\epsilon(y, z) \exp\left(\frac{-2\pi i m z}{\Lambda}\right) dz \tag{7.3.44}$$

where Λ is the corrugation period. In general, κ is evaluated numerically using Eqs. (7.3.42) and (7.3.44) for a specific corrugation profile $\Delta\epsilon(y, z)$. A simple expression for κ can be obtained for rectangular-shaped grooves. Consider the geometry shown in Fig. 7.7 for a typical DFB semiconductor

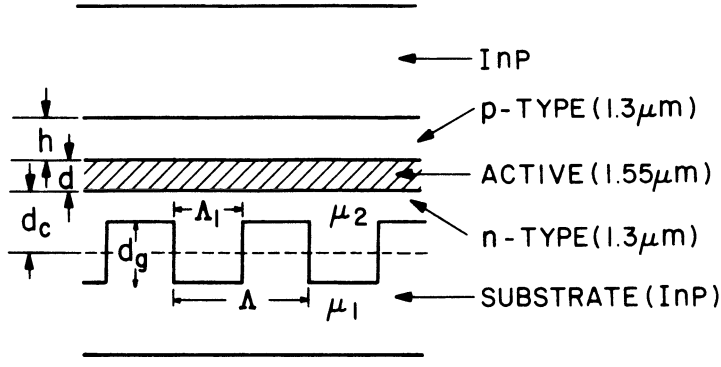


Fig. 7.7 Schematic illustration of the geometry and notation used for different epitaxial layers of a DFB semiconductor laser.

laser where d is the active-layer thickness, h is the thickness of the p-type cladding layer, d_g is the corrugation depth, and d_c is the average thickness of the n-type cladding layer. We note that $\Delta\epsilon(y, z) = \mu_2^2 - \mu_1^2$ when $d_c - d_g/2 \leq y \leq d_c + d_g/2$ and $-\Lambda_1/2 \leq z \leq \Lambda_1/2$. The origin of z is chosen at the center of the groove of width Λ_1. Integration in Eq. (7.3.44) leads to

$$\Delta\epsilon_m(y) = (\mu_2^2 - \mu_1^2)\frac{\sin(\pi m\Lambda_1/\Lambda)}{\pi m} \tag{7.3.45}$$

inside the grating region and $\Delta\epsilon_m(y) = 0$ outside. Using this value in Eq. (7.3.42), we obtain

$$\kappa \cong k_0 \Delta\mu\Gamma_g \frac{\sin(\pi m\Lambda_1/\Lambda)}{\pi m} \tag{7.3.46}$$

where we assumed that $\mu_1 + \mu_2 \cong 2\bar{\mu}$. Here

$$\Gamma_g = \frac{\displaystyle\int_{d_c - d_g/2}^{d_c + d_g/2} \phi^2(y)\, dy}{\displaystyle\int_{-\infty}^{\infty} \phi^2(y)\, dy} \tag{7.3.47}$$

and is the fraction of the mode energy inside the grating region. Further, $\Delta\mu = \mu_2 - \mu_1$ and is the index difference on the two sides of the grating. Several features of Eq. (7.3.46) are noteworthy. For the first-order grating ($m = 1$), κ is maximum for a symmetric grating such that $\Lambda_1 = \Lambda/2$. However, a symmetric second-order grating ($m = 2$) leads to $\kappa = 0$. The maximum coupling for $m = 2$ occurs when $\Lambda_1 = \Lambda/4$ or $3\Lambda/4$.

To evaluate Γ_g, it is necessary to solve the five-layer slab-waveguide problem (see Fig. 7.7). This is done by using the method discussed in Sec. 2.5 that requires the matching of ϕ and $d\phi/dy$ at various interfaces. Figure 7.8 shows[20] the calculated values of κ for the TE mode of a 1.55-μm InGaAsP laser with 1.3-μm cladding layers using $d = 0.1$ μm, $d_c = 0.1$ μm, and $h = 0.1$ μm. The corrugation depth d_g is varied. For comparison, the results of a sinusoidal grating are shown by dashed lines. The case of a symmetric grating is considered ($\Lambda_1 = \Lambda/2$) and $\kappa = 0$ for a second-order rectangular grating. However, κ can become appreciable even when $m = 2$ if other corrugation shapes are employed. Figure 7.8 shows that a coupling coefficient of about 100 cm^{-1} can be achieved for a first-order grating with corrugation depths as small as 0.05 μm. However, κ is expected to be smaller

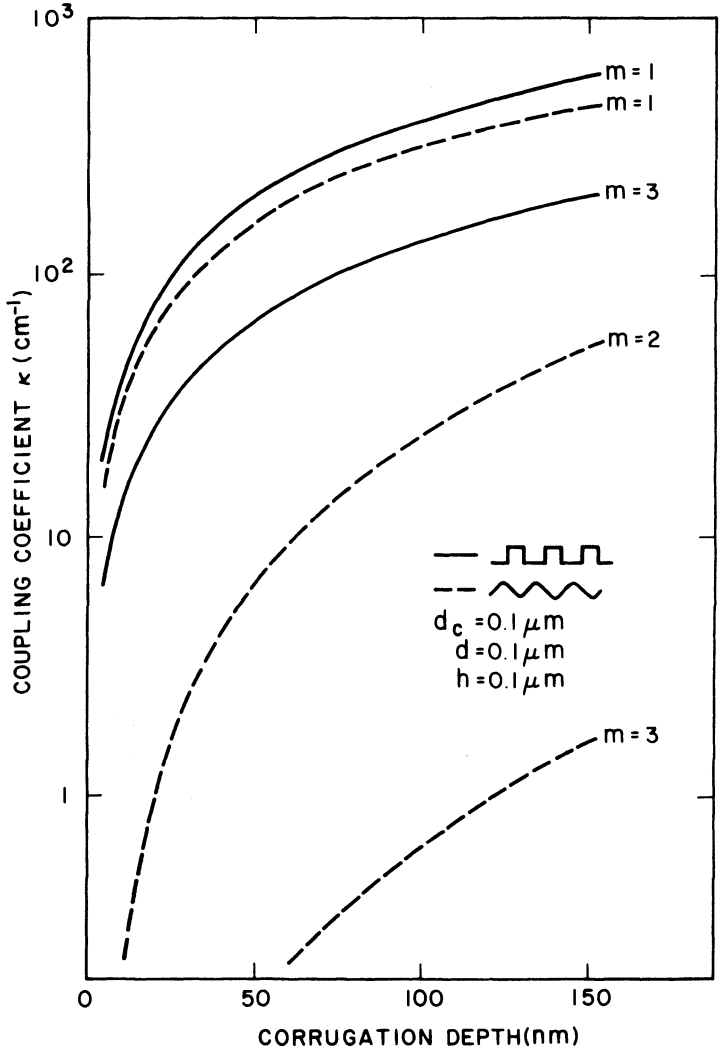


Fig. 7.8 Calculated coupling coefficient κ as a function of corrugation depth d_g for square (solid) and sinusoidal (dashed) gratings. The order of Bragg diffraction is denoted by m. (After Ref. 20 © 1982 IEEE)

for a second-order grating and values in the range of 40–50 cm^{-1} can be obtained using $d_g \cong 0.15\ \mu m$. For a 250-μm-long laser, one then obtains $\kappa L \cong 1$ and, as Figs. 7.5 and 7.6 show, such values of κ are sufficient to operate a DFB laser with a reasonable mode discrimination.

7.3.4 Threshold Behavior

The required threshold gain of a DFB laser given by Eq. (7.3.35) can be used to obtain the threshold current density J_{th}. If we assume that the material gain varies linearly with the carrier density, i.e.,

$$g(n) = a(n - n_0) \tag{7.3.48}$$

then the threshold value of the carrier density obtained using Eqs. (7.3.35) and (7.3.48) is given by

$$n_{th} = n_0 + \frac{\bar{\alpha} + \alpha_{int}}{a\Gamma}. \tag{7.3.49}$$

Here a is the gain coefficient and n_0 is the transparency value of the carrier density corresponding to the onset of population inversion. If carrier diffusion is ignored, the threshold current density can be obtained using the analysis of Sec. 2.6 and is given by

$$J_{th} = qdn_{th}(A_{nr} + Bn_{th} + Cn_{th}^2) \tag{7.3.50}$$

where A_{nr}, B, and C take into account various radiative and nonradiative recombination mechanisms.

To evaluate J_{th}, consider the device structure shown schematically in Fig. 7.7 for a 1.55-μm InGaAsP laser. The coupling coefficient κ for this structure is first calculated using the procedure described in Sec. 7.3.3. For the given value of κL, one can then obtain the DFB mode spectrum as discussed in Sec. 7.3.2. This provides us with the value of $\bar{\alpha}$ for the lowest-gain mode that would reach threshold first. The confinement factor Γ can be calculated for the specific structure shown in Fig. 7.7 by solving the five-layer waveguide problem. Equations (7.3.49) and (7.3.50) are then used to evaluate J_{th}.

The same method can be followed for either the TE or the TM mode after making use of the appropriate boundary conditions (see Sec. 2.5). For a DFB laser, it should be noted that both Γ and κ are different for TE and TM modes. In particular, while Γ is generally smaller for the TM mode, the coupling coefficient can be larger since a larger fraction of the mode (due to weaker mode confinement) interacts with the grating; i.e., Γ_g in Eq. (7.3.46) is larger. This suggests that the threshold margin between the TE and TM modes is likely to be smaller when compared with that of FP lasers. The threshold current density for an FP laser can also be obtained using Eq. (7.3.49) and (7.3.50) after replacing $\bar{\alpha}$ by α_m given by Eq. (7.3.36).

To illustrate the threshold behavior, we consider the device shown in

Fig. 7.7 and assume that a rectangular second-order grating of 0.1-μm thickness such that $\Lambda_1 = 0.3\Lambda$ is etched on the InP substrate. The cladding-layer thicknesses h and d_c are each 0.15 μm. The other device parameters are as follows: $L = 250$ μm, $\alpha_{int} = 40\ cm^{-1}$, $a = 2.5 \times 10^{-16}\ cm^2$, $n_0 = 1 \times 10^{18}\ cm^{-3}$, $A_{nr} = 6.6 \times 10^7\ s^{-1}$, $B = 0.9 \times 10^{-10}\ cm^3/s$, and $C = 4 \times 10^{-29}\ cm^6/s$. We also assume that the DFB laser has negligible facet reflectivities. Figure 7.9 shows the variation of J_{th} with the active-layer thickness for the TE and TM modes of the DFB semiconductor laser.[68] For comparison, the corresponding threshold curves for a similar FP semiconductor laser are also shown. For the FP laser, the facet loss α_m was calculated using $R_m = 0.36$ and $R_m = 0.28$ for TE and TM modes, respectively, while all other parameters were taken to be the same.

The threshold current density J_{th} in Fig. 7.9 is minimum at an optimum value of d; as d decreases J_{th} increases sharply, since Γ reduces substantially due to poor mode confinement. This qualitative behavior is identical for DFB and FP lasers, although the minima occur at different values of d. From a practical point of view, the important quantity is the threshold margin ΔJ_{th} between the TE and TM modes. Devices with a high value of ΔJ_{th} are preferred because the TE mode power at which the TM mode starts to lase (leading to a kink in the L-I curve) is pushed upward. Figure 7.9 shows that ΔJ_{th} is considerably smaller for a DFB laser in comparison to that of an FB laser. Furthermore, ΔJ_{th} varies with d. For a 0.2-μm-thick active layer, ΔJ_{th} is so small that the TE and TM modes would lase almost simultaneously. By contrast, a choice of $d \cong 0.1$ μm provides a significant margin and lowers the TE-mode threshold. The main point to note is that the design of a good DFB laser requires consideration and optimization of a large number of device parameters.

In the preceding discussion we did not include the effect of internal stress[69–72] that is invariably present in actual heterostructure devices. It can be incorporated through the stress dependence of the material gain g. In the presence of a compressive stress S normal to the junction plane, the gain $g(n)$ given by Eq. (7.3.48) decreases for the TE mode and increases for the TM mode. It can be approximated by[69,71]

$$g_{TE}(n) = a(n - n_0) - \gamma_1 S \tag{7.3.51a}$$

$$g_{TM}(n) = a(n - n_0) + 2\gamma_2 S \tag{7.3.51b}$$

where the constants γ_1 and γ_2 are related to the elastic tensor components. The physical reason for this gain difference is related to the splitting of the light-hole and heavy-hole valence bands as the uniaxial stress lowers the crystal symmetry from cubic to tetragonal.[73] Figure 7.10 shows the effect of stress on the threshold margin ΔJ_{th} with the same parameter values

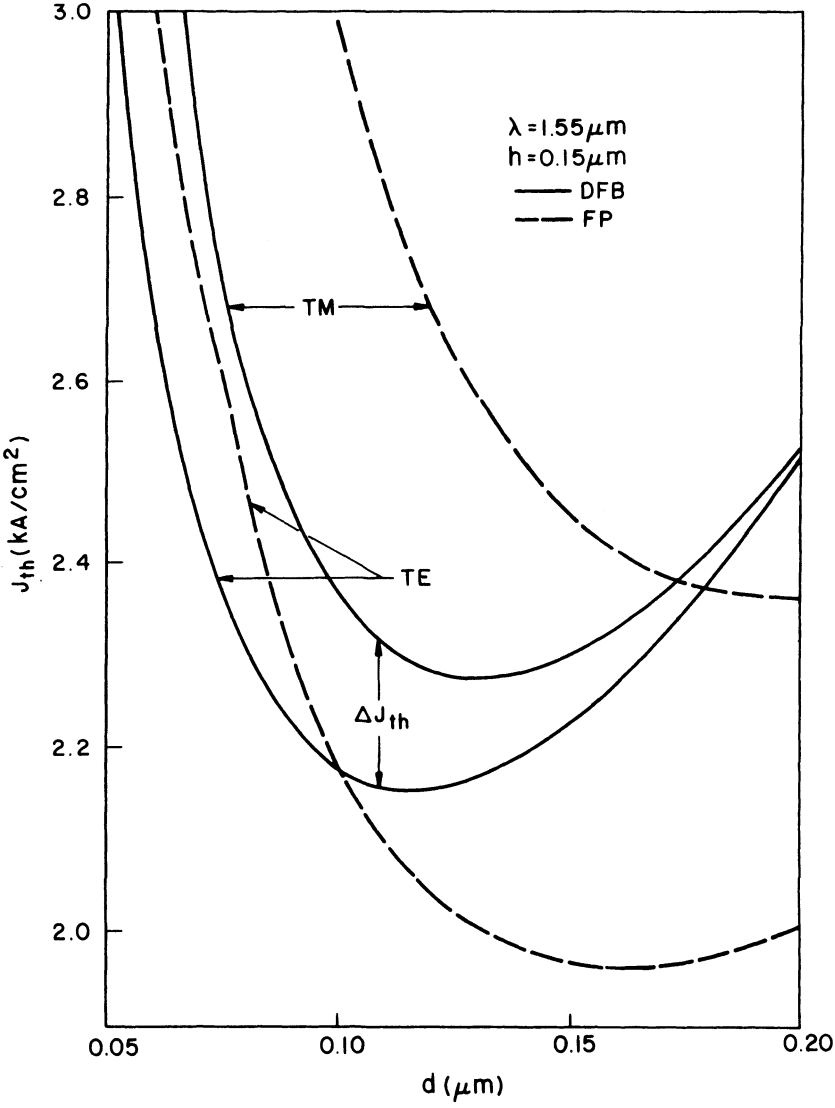


Fig. 7.9 Calculated threshold current densities (solid curves) for the TE and TM modes in a DFB semiconductor laser. For comparison, dashed curves show the corresponding behavior of a Fabry-Perot laser. (After Ref. 68)

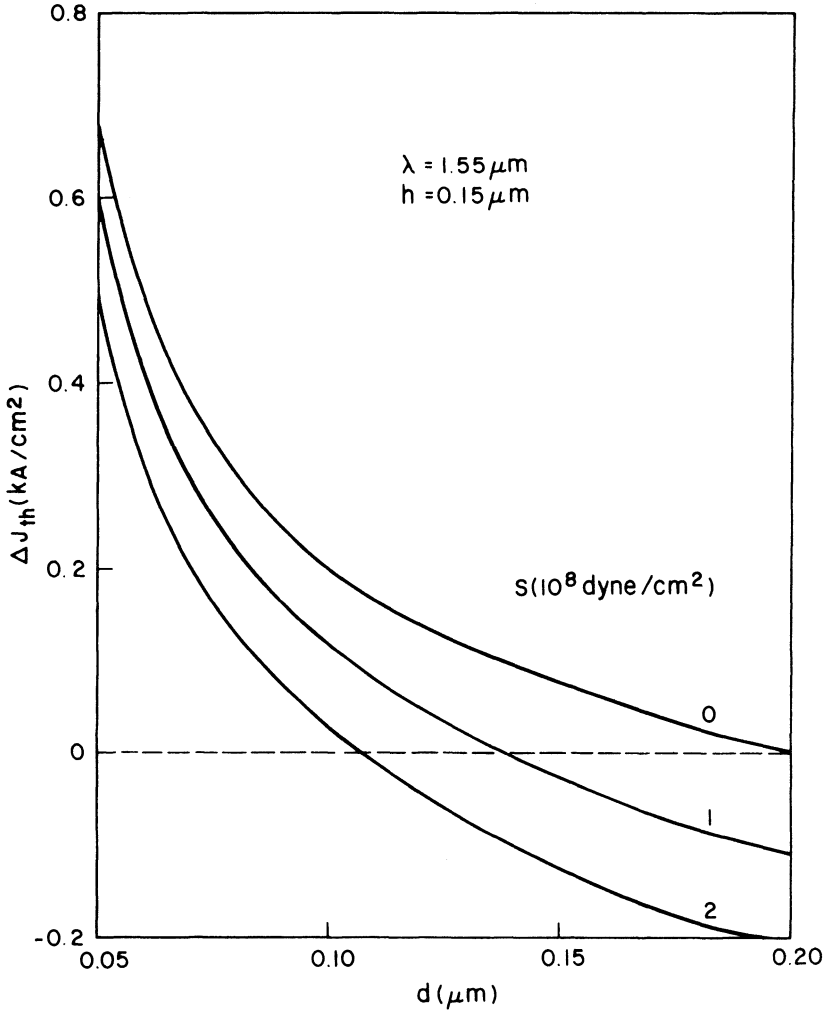


Fig. 7.10 Effect of stress S on the threshold difference ΔJ_{th} between the onsets of TM and TE modes. For values of the active-layer thickness d above which $\Delta J_{th} < 0$, the TM mode reaches threshold first. (After Ref. 68)

used for Fig. 7.9. We assumed that $\gamma_1 = \gamma_2 = \gamma/3$ and that $\gamma = 2 \times 10^{-7}$ cm/dyne.[71] As expected, the threshold margin decreases in the presence of stress. Furthermore, above a critical value of the active-layer thickness, the TM mode has a lower threshold and thus would lase first, before the TE mode. This behavior has been observed in some DFB lasers.

7.3.5 Light-Current Characteristics

Similar to an FP semiconductor laser, the light-current (L-I) characteristics of a DFB laser yield information about the threshold current and the differential quantum efficiency η_d that is related to the slope of the L-I curve. In the previous section we have seen that the threshold current density J_{th} has similar magnitudes for FP and DFB lasers. The difference between the two kinds of devices is governed by the relative magnitudes of feedback losses $\bar{\alpha}$ and α_m. For a 250-μm-long FP laser with 32% facet reflectivities, $\alpha_m \cong 45\,\text{cm}^{-1}$ [see Eq. (7.3.36)]. For a DFB laser, the magnitude of $\bar{\alpha}$ depends on the coupling coefficient κ as well as on the facet reflectivities. For a laser with nonreflecting facets, values of κL greater than 3.2 are necessary to make $\bar{\alpha}$ less than or equal to α_m (see Fig. 7.4). However, for cleaved-facet DFB lasers, small values of κ can lead to the situation where $\bar{\alpha} < \alpha_m$ (see Fig. 7.5). In contrast to FP lasers, however, the DFB threshold (or $\bar{\alpha}$) is sensitive to phase shifts at the facets and can vary significantly from device to device.

The differential quantum efficiency η_d of a DFB laser can also be obtained using an expression similar to that derived in Sec. 2.6 for an FP laser and is given by[74]

$$\eta_d = \eta_i \frac{\bar{\alpha}}{\bar{\alpha} + \alpha_{\text{int}}} \tag{7.3.52}$$

where η_i is the internal quantum efficiency and is nearly 100% when stimulated emission dominates in the above-threshold regime. The internal loss α_{int} incorporates losses from all possible mechanisms such as free-carrier absorption and scattering at the heterostructure interfaces. For a DFB laser, two additional mechanisms may contribute to increase α_{int} and decrease η_d relative to their values for an FP laser. These are (i) scattering losses due to grating imperfections and (ii) radiation losses from periodic structures. The latter mechanism does not occur for a first-order grating but may contribute a loss of $\sim 5\,\text{cm}^{-1}$ in the case of a second-order grating since first-order Bragg diffraction then radiates perpendicular to the junction plane.[48,57] Although numerical values of $\bar{\alpha}$ anf α_{int} are generally device-dependent, η_d for a DFB laser is expected to be lower than that of an FP laser when both lasers have similar threshold densities.

An interesting aspect of DFB lasers is that the differential quantum efficiency (or the slope of the L-I curves) can be different for the light emitted by the two facets even when the device is apparently axially symmetric. This feature is related to the phase shifts ϕ_1 and ϕ_2 introduced in Eq. (7.3.26) to account for the relative positions of the facet and the last complete grating corrugation. Equation (7.3.52) is not helpful for obtaining individual facet

efficiencies η_1 and η_2. Furthermore, its derivation implicitly assumes that the internal power is axially uniform. To overcome these deficiencies, we use the general definition that the differential quantum efficiency for each facet is the power transmitted through that facet divided by the total power generated by stimulated emission throughout the cavity length.[57] Using Eq. (7.3.13) for the intracavity field, the differential quantum efficiency for the facet at $z = 0$ becomes

$$\eta_1 = \frac{(1 - R_1)|B(0)|^2}{(\bar{\alpha} + \alpha_{\text{int}}) \displaystyle\int_0^L |E(z)|^2 \, dz}. \tag{7.3.53}$$

A similar expression is used for η_2 after replacing the numerator by $(1 - R_2)|A(L)|^2$. Equations (7.3.27)–(7.3.30) can be used to obtain η_1 and η_2 for a DFB laser with arbitrary facet reflectivities R_1 and R_2 and phase shifts ϕ_1 and ϕ_2. The effect of facet phases on the differential quantum efficiencies has been discussed.[57,75] The individual efficiencies η_1 and η_2 as well as the total efficiency ($\eta_d = \eta_1 + \eta_2$) vary significantly with the phases ϕ_1 and ϕ_2. The range of variation depends on the coupling coefficient κ. For typical values for κL of ~ 1, one can expect a 20% variation in the differential quantum efficiency from device to device. For a laser with one cleaved facet and one nonreflecting facet, powers emitted by the two facets differ considerably; their ratio can be in the range of 2–5 depending on the coupling coefficient and the corrugation phase at the cleaved facet.[75]

7.4 PERFORMANCE

In Sec. 7.2 we briefly discussed the heterostructures used to make a DFB semiconductor laser. The purpose of this section is to describe their performance by considering the CW and pulsed modes of operation. The features to be considered are the L-I curves and their temperature dependence, the longitudinal-mode spectra and the extent of side-mode suppression, and the modulation-response characteristics such as modulation bandwidth and frequency chirp. Chapters 5 and 6 discussed these features for an FP semiconductor laser. We shall refer to them in comparing the performances of FP and DFB semiconductor lasers. We consider the CW and the modulation characteristics in Secs. 7.4.1 and 7.4.2 respectively. The performance of phase-shifted DFB lasers and multiquantum-well DFB lasers is then described in the next two subsections. Section 7.4.5 is devoted to a new class of DFB lasers referred to as gain-coupled DFB lasers.

7.4.1 CW Operation

The L-I curves at different temperatures characterize the CW performance of a semiconductor laser and are useful for indicating how the threshold current I_{th} and the differential quantum efficiency η_d degrade as the device temperature increases. Figure 7.11 shows the L-I curves of a 1.55-μm DFB laser with the DCPBH structure (see Fig. 7.2). This device has a first-order grating ($\Lambda = 240$ nm) and the corrugation depth after regrowth is estimated to be 30 nm.[28] The cavity length is about 300 μm with both facets cleaved. Figure 7.11 should be compared with Fig. 5.26 obtained for an FP-type DCPBH laser.

Several features of the L-I curves are noteworthy. At room temperature the threshold current is about 30 mA and the slope is nearly constant up to a power level of 10 mW from each facet. The two-facet differential quantum

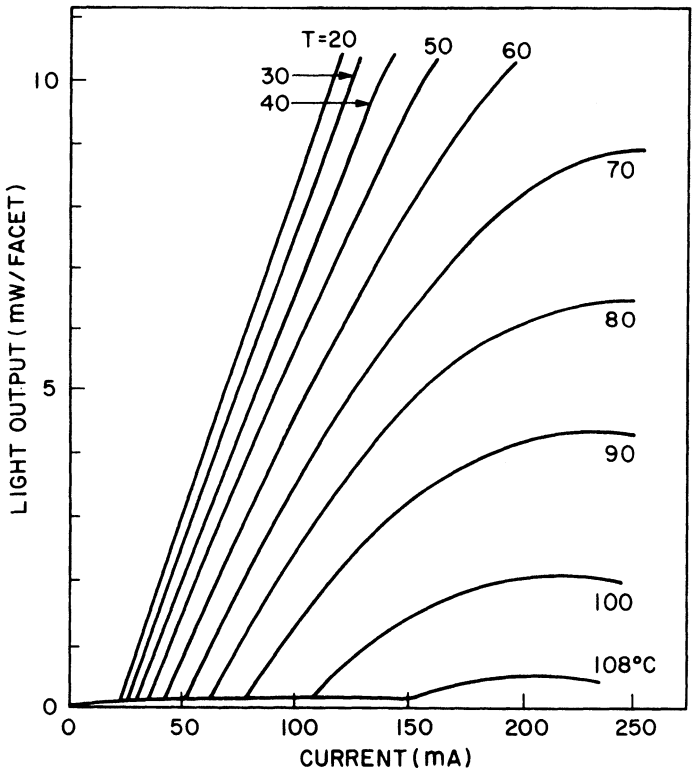


Fig. 7.11 Temperature dependence of the light-current characteristics for a 1.55-μm InGaAsP DCPBH DFB semiconductor laser. (After Ref. 28)

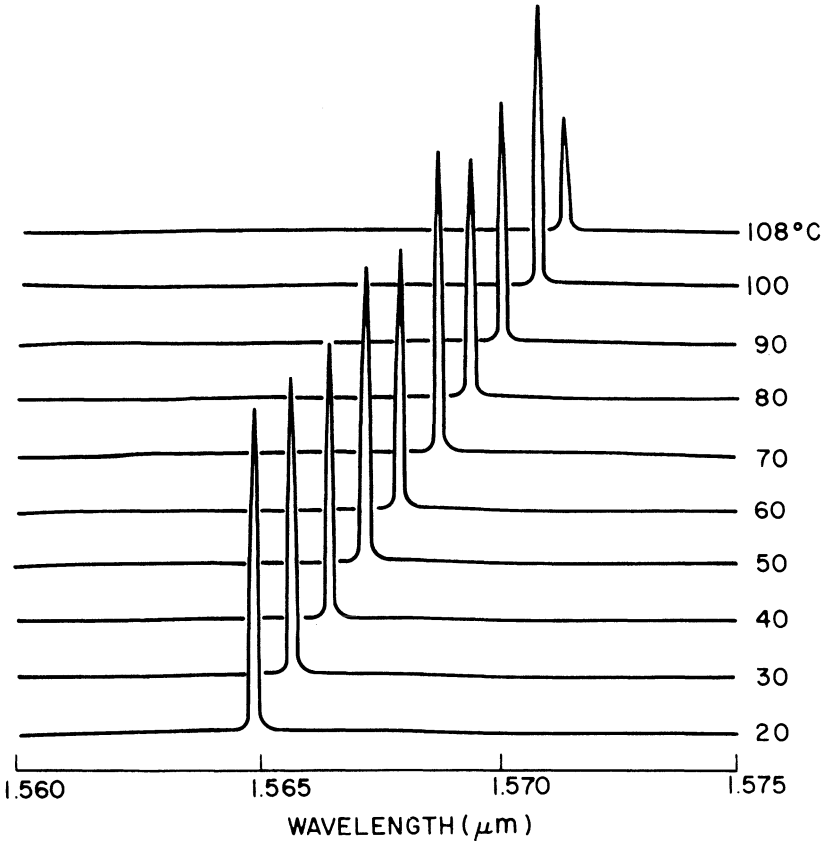


Fig. 7.12 Longitudinal-mode spectra of the DCPBH DFB laser shown in Fig. 7.11 at several temperatures. The laser maintains the same longitudinal mode while its wavelength shifts continuously at a rate of 0.09 nm/°C. (After Ref. 28)

efficiency η_d at room temperature is typically 30–40% for these devices.[28] From the temperature dependence of the threshold current, the characteristic temperature T_0 (see Sec. 3.6) is estimated to be 67 K at room temperature. The performance of these DFB lasers is comparable to DCPBH FP lasers as far as L-I curves are concerned.

However, DFB lasers are expected to outperform conventional FP lasers with resepct to their spectral purity. This is shown in Fig. 7.12 where the longitudinal-mode spectra at $I = 1.5I_{th}$ are shown at several temperatures for the device of Fig. 7.11. The DFB laser maintains the same longitudinal mode in the entire temperature range of 20–108°C. By contrast, an FP

semiconductor laser later would exhibit several mode jumps over this temperature range because of the temperature-induced shift of the gain peak. The stability of the DFB longitudinal mode is due to the built-in grating whose period determines the lasing wavelength [see Eq. (7.2.1)].

In Fig. 7.12 the wavelength changes slightly with temperature at a rate $d\lambda/dT \cong 0.09$ nm/°C. This happens because the mode index $\bar{\mu}$ in Eq. (7.2.1) varies with temperature.[76] However, the wavelength shift for a DFB laser is considerably smaller compared with that of an FP laser, where $d\lambda/dT \cong 0.5$ nm/°C. The calculated value of $d\lambda/dT$ for the DFB laser is about 0.1 nm/°C after including the temperature dependence of the carrier-induced contribution to the refractive index[75] and is in agreement with the experimental value.

As discussed in Chapter 6, the mode-suppression ratio (MSR), defined as the ratio of the power of the main mode to that of the most intense side mode, plays an important role in characterizing the spectral purity. For the FP laser, an MSR greater than 50 is difficult to achieve; typical values fall in the 10–20 range. For the DFB laser, MSR values of ~1,000 (30 dB) are readily obtained at power levels of a few milliwatts. For example, the MSR for the laser in Fig. 6.12 was over 35 dB at a 5-mW output level.[31]

To understand the origin of such a high degree of spectral purity in DFB lasers, we consider the multimode analysis of Sec. 6.3.2. If we use Eq. (6.3.9) to obtain the powers P_0 and P_1 for the main mode and the most intense side mode, respectively, the MSR is given by

$$\text{MSR} = \frac{P_0}{P_1} = \frac{\gamma_1 - G_1}{\gamma_0 - G_0} \tag{7.4.1}$$

where

$$\gamma_i = v_g(\bar{\alpha}_i + \alpha_{\text{int}}) \tag{7.4.2}$$

is the photon decay rate and

$$G_i = \Gamma v_g g(n, \omega_i) \tag{7.4.3}$$

is the net rate of stimulated emission for the mode lasing at the frequency ω_i ($i = 0$). In contrast to an FP laser, the photon decay rate is different for different modes because of the frequency-selective feedback provided by the grating. This is evident in Figs. 7.4–7.6, where the mode-gain coefficient $\bar{\alpha}$ is lowest for the main mode. If we define $\Delta\gamma$ and ΔG as $\Delta\gamma = \gamma_1 - \gamma_0$ and $\Delta G = G_0 - G_1$, Eq. (7.4.1) becomes

$$\text{MSR} = 1 + \left(\frac{\Delta\gamma + \Delta G}{\gamma_0 \delta}\right) \tag{7.4.4}$$

where $\delta = 1 - G_0/\gamma_0$ and is a small dimensionless parameter whose value is $\sim 10^{-4}$ and decreases with an increase in the main-mode power [see Eq. 6.3.11)]. For an FP laser, $\Delta\gamma = 0$ and ΔG alone, arising from the gain roll-off, provides mode discrimination. However, $\Delta G/G_0$ for the two adjacent longitudinal modes is typically less than 0.5% and the MSR $\lesssim 50$. For a DFB laser the relative change $\Delta\gamma/\gamma_0$ in the photon decay rates can be made $\sim 10\%$ with a proper design, and MSR values of $\sim 1{,}000$ can be achieved. The required loss margin depends on δ and is 4–5 cm^{-1} when $\delta = 5 \times 10^{-5}$.

The performance of the DCPBH DFB lasers shown in Figs. 7.11 and 7.12 was achieved with the use of a first-order grating. The use of a second-order grating can also lead to high-performance DFB lasers.[26] However, since the coupling coefficient is generally smaller for a second-order grating (see Fig. 7.8), the mode discrimination between the lowest-gain DFB mode and the FP mode (closest to the gain peak) is likely to be poor unless the gain peak is made to nearly coincide with the Bragg wavelength. This can be understood from Eq. (7.4.4) by noting that $\Delta G < 0$ in the case of gain-peak deviation from Bragg resonance. Furthermore, since the gain peak shifts with the drive current and the temperature, such DFB lasers may exhibit high MSRs only over a limited current and temperature range. Their performance can be improved by making one or both facets low-reflecting; this increases the losses for the FP modes, which results in making $\Delta\gamma$ in Eq. (7.4.4) larger. Reduced facet reflectivity can be achieved by etching[25] or burying[22] the facet or by using an antireflection coating.[26] The use of a low-reflectivity facet, however, increases the device threshold, as can be seen by comparing Figs. 7.5 and 7.6. High-performance DCPBH lasers with a second-order grating, one cleaved facet, and one low-reflectivity facet have been demonstrated[27] at both 1.3-μm and 1.5-μm wavelengths. The threshold current is typically 50–60 mA. At the 1.3-μm wavelength, a relatively high power (over 55 mW) at room temperature has been achieved under CW operation, and the laser exhibited CW operation up to a temperature of 105°C while maintaining the single longitudinal mode. At the 1.55-μm wavelength, the corresponding reported values[26] of power and temperature are 20 mW and 75°C. Comparable performance has also been achieved using other buried-heterostructure-type lasers.[22,25]

The performance of DFB lasers is quite sensitive to the facet reflectivities.[58–60] As discussed in Sec. 7.3.2 and shown in Figs. 7.5 and 7.6, the loss margin $\Delta\gamma$, or the difference $\Delta\alpha$ in the threshold gain of the main and side modes, depends not only on the facet reflectivities but also on the grating phases ϕ_1 and ϕ_2 at the two facets. In particular, the gain margin $\Delta\alpha$ becomes quite small for some specific phase combinations, indicating that those DFB lasers will not oscillate in a single longitudinal mode. Since ϕ_1 and ϕ_2 can vary randomly from device to device, the performance of DFB lasers can be

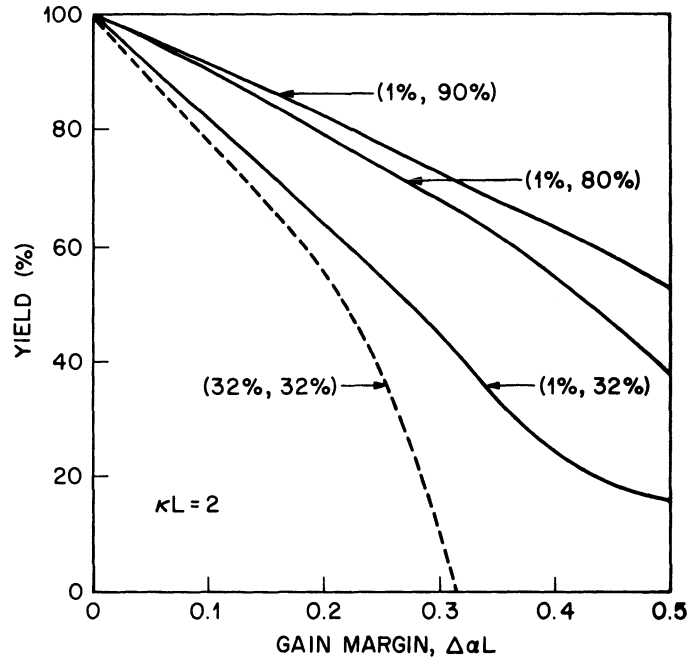


Fig. 7.13 Yield calculated for DFB lasers as a function of the gain margin $\Delta\alpha L$ for several combinations of the facet reflectivities (given in parentheses).

characterized only in a statistical sense. The "yield" of acceptable devices can be estimated numerically by finding the fraction of phase combinations for which $\Delta\alpha$ exceeds a certain value.[77,78] Figure 7.13 shows the numerically calculated yield as a function of $\Delta\alpha L$ for several combinations of the facet reflectivities R_1 and R_1 (given in parentheses) for a DFB laser with $\kappa L = 2$.[1] Typically $\Delta\alpha L$ should exceed 0.2 for the satisfactory performance of a DFB laser. The dashed curve in Fig. 7.13 shows that the yield of acceptable devices for this value of $\Delta\alpha L$ is below 50% for a laser with uncoated facets ($R_1 = R_2 = 32\%$). However, it can be made to exceed 80% by a proper combination of the facet reflectivities.

Many other laser characteristics such as the threshold current J_{th}, the differential quantum efficiency η_d, and the laser line width, also depend on the facet reflectivities and the grating phases.[77,78] Yield curves similar to those shown in Fig. 7.13 can be generated by considering the phase combinations that satisfy not only the gain-margin condition but also preset conditions on J_{th}, and η_d. The numerical results show[78] that it is possible to achieve an overall yield of 80–90% by a proper choice of values for the

parameters κL, R_1, and R_2. In particular, a combination of low-reflection ($\sim 1\%$) and high-reflection (80–90%) coatings on the laser facets has provided DFB lasers of high output powers with high yield.[79]

7.4.2 Modulation Performance

For their application in high-bit-rate optical communication systems, the performance of DFB semiconductor lasers under high-frequency direct modulation is of significant interest. Conventional FP lasers suffer from one drawback: they become essentially multimode (MSR ~ 1) under direct modulation even when the side modes are reasonably suppressed during CW operation. As discussed in Chapter 1, unless dispersion-shifted fibers are used, the large source bandwidth, combined with fiber dispersion at the 1.55-μm wavelength, severely limits the allowable bit rate for long-distance optical transmission through silica fibers. DFB semiconductor lasers are expected to overcome this limitation by virtue of their design.

Figure 7.14 shows the measured powers in the main mode and the most intense side mode of a buried-heterostructure laser with a second-order grating and a tilted (nonreflecting) facet.[25] Solid curves were obtained under a sinusoidal modulation of 500 MHz with a peak-to-peak amplitude of 14 mA. Dashed curves show for comparison the mode powers under CW operation. The DFB laser maintains an MSR of about 30 dB in a large current range. The lowest value of MSR (~ 16 dB) occurs at the 67-mA bias current corresponding to a 100% depth of modulation. Similar behavior has been observed for other DFB structures. A properly designed, state-of-the-art DFB semiconductor laser can maintain the single-longitudinal-mode operation with an MSR greater than 30 dB under direct modulation in the gigahertz range. It is this property that makes the DFB laser an ideal candidate for high-bit-rate optical communication systems.

The multimode rate equations (6.2.25) and (6.2.26) of Chapter 6 can be used to simulate the longitudinal-mode spectra of DFB lasers under transient conditions after using Eq. (7.4.2) for the photon decay rates. The numerical results show[56] that coupling-coefficient values as small as $\kappa L = 0.15$ significantly reduce the side-mode power at the first relaxation-oscillation peak when compared with the FP case ($\kappa = 0$), where the main and side modes are almost equally intense (see Figs. 6.7 and 6.8). The side mode remains suppressed by more than 20 dB at all times when $\kappa L = 1$. These results suggest that for a 250-μm-long laser a coupling coefficient κ of 40 cm^{-1} is large enough to provide dynamic side-mode suppression. Such values of κ are readily achieved with both first- and second-order gratings. Numerical simulations using the multimode rate equations under 2 Gb/s direct modulations also show that an MSR of 30 dB is maintained as long

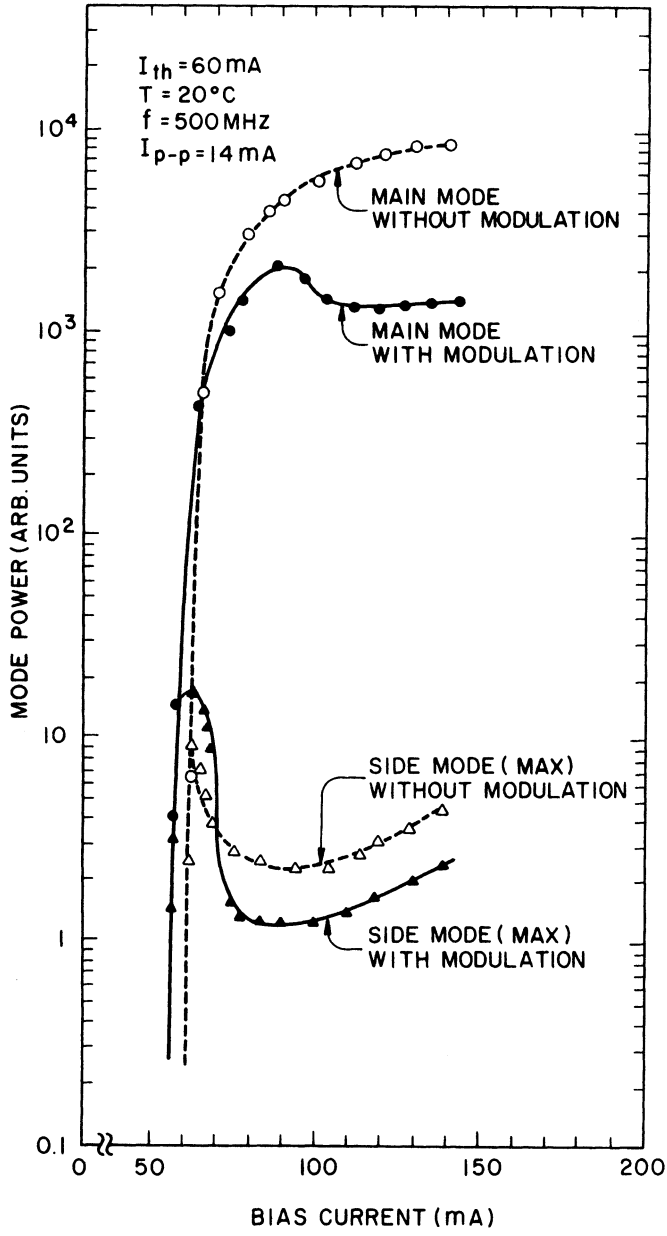


Fig. 7.14 Measured optical power in the main mode and the most intense side mode of a DFB laser as a function of the bias current with (solid curve) and without (dashed curve) modulation. An MSR of about 30 dB is maintained for bias currents greater than 70 mA. (After Ref. 25 © 1984 IEEE)

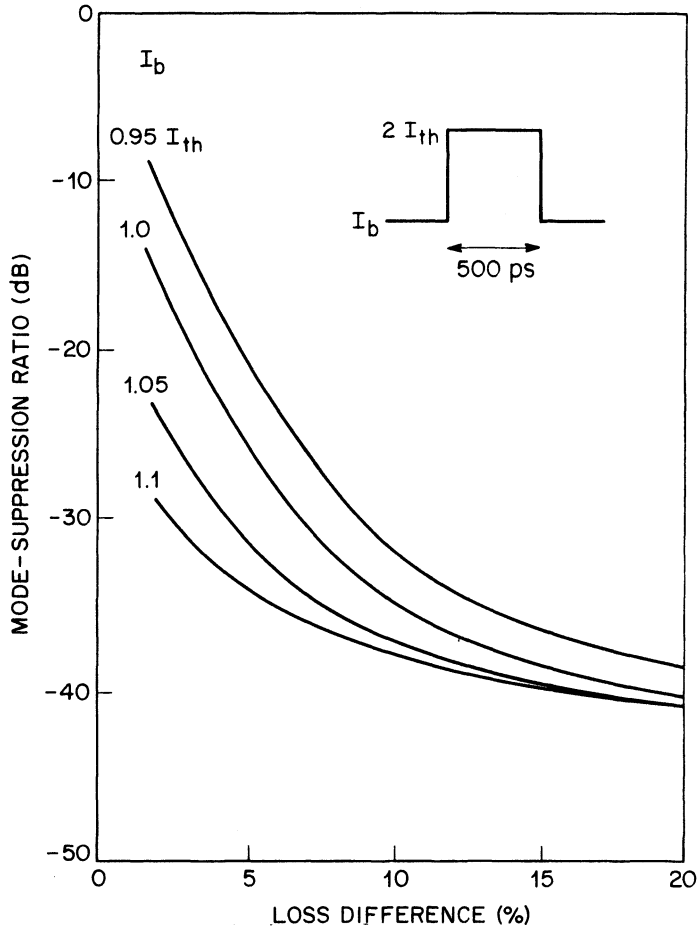


Fig. 7.15 Dependence of the MSR on the relative loss difference between the main and side modes for several bias levels I_b. The inset shows the current pulse corresponding to a bit rate of 2 Gb/s. (After Ref. 60)

as the relative loss difference [$\Delta\gamma/\gamma_0$ in Eq. (7.4.4)] is greater than $\sim 10\%$. Similar to the experimental data of Fig. 7.14, the MSR depends on the bias level. Figure 7.15 shows the MSR calculated by solving the rate equations numerically as a function of the relative loss difference $\Delta\gamma/\gamma_0$ at several bias levels. The semiconductor laser is modulated at 2 Gb/s by applying 500-ps current pulses. The MSR is defined as the ratio of the pulse energies for the main and side modes. In general, the modulation performance of a DFB

laser is superior in the sense of maintaining high MSR when the laser is biased slightly above threshold.

As discussed in Chapter 6, the direct modulation also leads to wavelength chirping that shifts the laser wavelength over a wide range during each modulation cycle. This range is usually referred to as the chirp and is typically ~ 0.1 nm under practical conditions.[25–31] The wavelength chirp leads to a significant dispersion penalty in optical communication systems and is often the limiting factor in their performance. The origin of the chirp relates to the carrier-induced index change accompanying the modulation-induced gain variations. Since carrier-induced index changes are about the same for FP and DFB lasers operating under identical conditions, the wavelength chirp for DFB lasers is expected to be comparable to that observed for FP lasers. Measurements of the chirp for DFB lasers have shown that this is indeed the case.[25–31] However, the amount of chirp is dependent to some extent on the device structure (see Fig. 6.15). With a proper design, DFB lasers with a chirp of only 0.4 nm under 2-GHz modulation were realized as early as 1985.[80]

Another important characteristic that should be considered is the modulation bandwidth ν_B. Under ideal conditions it is determined by the relaxation-oscillation frequency ν_R, which can be increased by increasing the bias power P_b since $\nu_R \propto P_b^{1/2}$ (see Sec. 6.3). However, in practice electrical parasitics can decrease the modulation response by more than 3 dB for frequencies well below ν_R, thereby significantly reducing the modulation bandwidth. The bandwidth of early DFB lasers was often limited to below 5 GHz because of electrical parasitics. Advances in the packaging technology resulted in considerable improvement after 1985. A DFB laser exhibited a modulation bandwidth of 14 GHz in 1987. Figure 7.16 shows the modulation response of this laser at several bias levels.[81] The 3-dB bandwidth exceeds 14 GHz when the laser is biased at $I_b = 7.69 I_{\text{th}}$. A modulation bandwidth of 17 GHz was realized in 1989 by using a phase-shifted DFB laser.[82] Since phase-shifted DFB lasers have several advantages over the conventional design, we turn to their discussion in the next subsection.

7.4.3 Phase-Shifted DFB Lasers

As discussed in Sec. 7.3.2, DFB lasers designed such that they experience an extra phase shift in the middle of the laser cavity have a low-threshold mode (the so-called gap mode) whose frequency lies inside the stop band associated with conventional DFB lasers. In the case of 90° or $\pi/2$ phase shift, the mode lies at the center of the stop band and its wavelength coincides exactly with the Bragg wavelength. At the same time, the gain margin between this mode and the neighboring modes is so large ($>20\%$) that side modes of such a

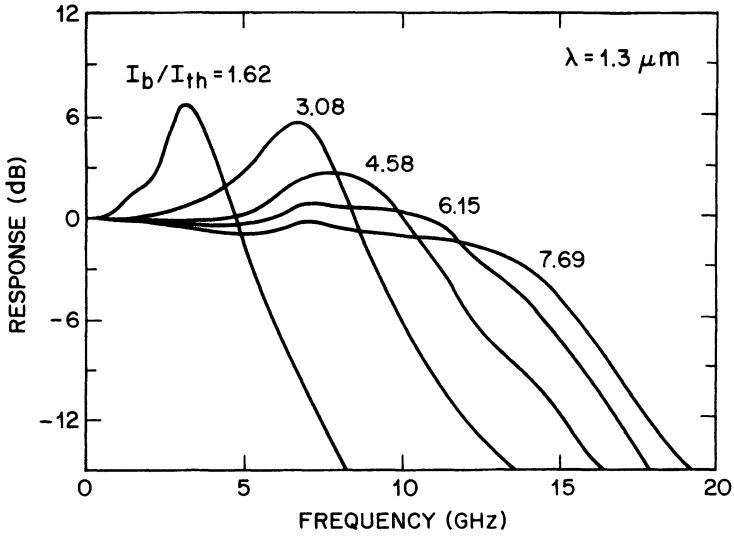


Fig. 7.16 Small-signal modulation response of a high-speed DFB laser at several bias levels. (After Ref. 81 © 1987 IEEE)

laser are likely to remain suppressed (by more than 30 dB) even when the laser is directly modulated at high speeds in the GHz range. Such phase-shifted DFB lasers are referred to as $\lambda/4$-shifted DFB lasers since the $\pi/2$ phase shift is realized by shifting the grating by a quarter wavelength. Several schemes were proposed around 1984–85 to realize such a phase shift during the grating fabrication (see Sec. 7.3.2).[61–65] Since then, phase-shifted DFB lasers have been extensively studied.[83–93] This section reviews the performance issues associated with such devices.

The advantages of phase-shifted DFB lasers can be understood by considering the threshold characteristics such as the gain of the lowest-loss mode, the gain margin $\Delta\alpha$ between the main and the side modes, and the detuning of various modes from the Bragg wavelength. These quantities can be calculated by solving the coupled-mode equations, Eqs. (7.3.15) and (7.3.16), with the appropriate boundary conditions. The main difference from the theory of Sec. 7.3.1 is that the parameter $\Delta\beta$ is no longer constant along the entire cavity length because of the abrupt phase shift occurring in the middle of the cavity. Since $\Delta\beta$ is constant in the two sections on each side of the phase-shift region, a simple approach applies the analytic solution of Sec. 7.3.1 to each section and connects the two solutions in the middle through the phase shift. Such an approach is referred to as the matrix method[94–96] since propagation through each section is governed by a 2×2 matrix.

It can be readily extended to study devices with multiple phase shifts or multisection devices with different grating parameters in each section.[95]

In the case of $\lambda/4$-shifted DFB lasers with negligible facet reflectivities the lowest-loss mode is found to be located exactly at the Bragg wavelength. The gain margin $\Delta\alpha$ depends on the coupling coefficient κ, but is much larger than that for conventional DFB lasers. For example, $\Delta\alpha L$ exceeds 1 when $\kappa L = 2$. It is this property of phase-shifted DFB lasers that makes them attractive since the MSR can remain quite large even under high-speed modulation for such large values of the gain margin. However, similar to the case of conventional DFB lasers, the gain margin is very sensitive to the facet reflectivities. Generally speaking, both facet reflectivities should be quite small ($\ll 1\%$) and the coupling coefficient should not be too high before the yield of $\lambda/4$-shifted DFB lasers is at an acceptable level.[89]

There are two main problems associated with $\lambda/4$-shifted DFB lasers. First, as discussed in Sec. 7.3.4, the gain margin between the TE and TM modes is generally small in DFB lasers compared with FP lasers. This problem becomes especially severe for $\lambda/4$-shifted DFB lasers since the loss differential resulting from different facet reflectivities of TE and TM modes is absent.[97] Phase-shifted DFB lasers need to be carefully designed to enhance the gain margin between TE and TM modes.

The second problem results from a phenomenon referred to as longitudinal spatial-hole burning.[89] The mode intensity in $\lambda/4$-shifted DFB lasers is highly nonuniform and peaks in the center of the cavity where the phase-shift region is located. Because of enhanced stimulated emission occurring near the cavity center, the gain, or equivalently the carrier density, is depleted more in the center than near the cavity ends, as if a spatial hole were burnt in the axial profile of the carrier density. Since the refractive index of the optical mode depends negatively on the carrier density, a lower carrier density implies an increase in the local refractive index. This change translates into an extra power-dependent phase shift near the cavity center that adds to the built-in phase shift. Since the gain margin depends on the amount of phase shift and is maximum for a shift of $\pi/2$, it degrades with an increase in the output power. Eventually the side mode becomes strong enough that the laser no longer qualifies as a single-longitudinal-mode laser.[89] The extent of degradation depends on nonuniformity of the axial intensity profile which in turn depends on the coupling coefficient. Numerical results show that longitudinal spatial-hole burning is particularly severe for large values of κL. In particular, the intensity profile is least nonuniform for $\kappa L = 1.25$.[89] For this reason, $\lambda/4$-shifted DFB lasers often employ low balues of κL in the range 1–1.5.

Several schemes can be used to counteract the impact of spatial-hole burning in phase-shifted DFB lasers. For example, the laser can be designed with a nonoptimum built-in phase shift $\phi_{sh} < \pi/2$. Even though the gain

margin is reduced near threshold, spatial-hole burning initially improves the gain margin before it becomes so severe that the gain margin nearly vanishes.[89] The net result is to enhance the range of single-mode operation. A similar effect can be achieved by injecting the current nonuniformly such that the carrier density at threshold is higher in the center than near the cavity edges.[90] Spatial-hole burning burns the hole, but the laser can be operated to higher powers before its impact becomes too severe.

Another technique consists of designing DFB lasers with multiple phase-shift regions.[98–101] The number, the location, and the amount of phase shift then provide additional control over the extent of spatial-hole burning since they affect the intensity distribution of the lasing mode. Figure 7.17 compares the axial distribution of mode intensity when the number of phase-shift regions N_{sh} varies from 1 to 3. The phase shift in each region is assumed to be the same, and the phase-shift regions are located symmetrically with equal spacing. Thus, phase shifts occur with a spacing of $L/(N_{sh} + 1)$ for a cavity of length L. The amount of phase shift ϕ_{sh} is optimized for each value of N_{sh} to obtain the maximum gain margin ($\phi_{sh} = 90°$, 50°, and 65° for $N_{sh} = 1$, 2, and 3, respectively). In all cases the mode intensity peaks at the location of the phase-shift regions. However, the intensity distribution is considerably

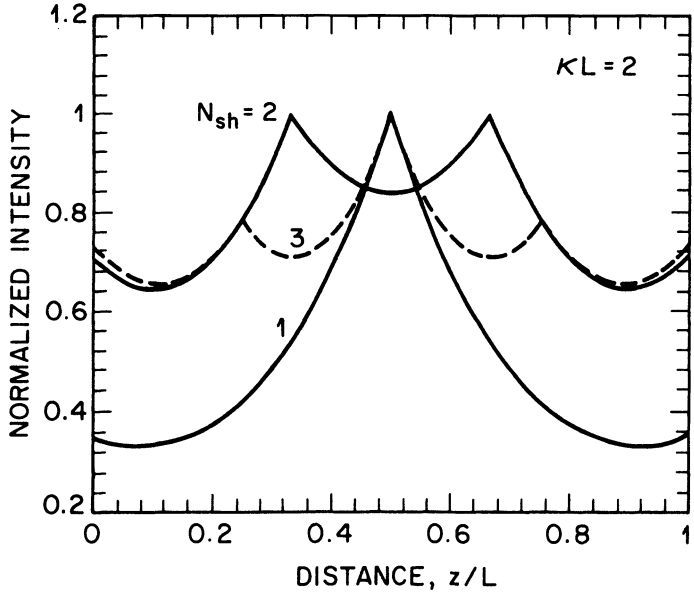


Fig. 7.17 Axial distribution of the mode intensity for DFB lasers with 1, 2, and 3 phase-shift regions located where the intensity peaks. (After Ref. 98)

more uniform for $N_{sh} = 2$ and 3 compared with $N_{sh} = 1$, resulting in reduced spatial-hole burning. Spatial-hole burning can also be reduced by making the coupling coefficient nonuniform such that it is smallest in the central portion of the cavity.[102] In another approach, the corrugation pitch of the grating is made larger in the central region of the cavity so that the phase shift is distributed over a wider region. The length of the central region and the pitch difference are chosen such that the optical field experiences a phase shift of $\pi/2$. Such lasers are called corrugation-pitch-modulated DFB lasers.[103]

7.4.4 Multiquantum-Well DFB Lasers

The use of a multiquantum-well (MQW) design (see Chapter 9) has several advantages over conventional semiconductor lasers with a single, relatively thick ($\sim$0.1 μm) active layer.[104] It is thus natural to consider its use for DFB lasers as well. Even though DFB lasers with an MQW active region were fabricated as early as 1985 by using LPE,[105] their development had to wait until the use of the MOVPE and MBE growth techniques became practical. Such DFB lasers are called MQW DFB lasers and have been extensively studied since 1988.[106–119] This section briefly reviews their performance.

The advantages of the MQW structure stem from a fundamental change in the density of states associated with the conduction and valence bands of a thin layer (thickness $\sim$10 nm) because of the quantum-size effects (see Chapter 9). Both the gain and the refractive index are affected by changes in the density of states.[104] In particular, the differential gain, quantified by the parameter a, is larger by about a factor of 2 while, at the same time, the line-width enhancement factor β_c is smaller by about a factor of 2. As discussed in Chapters 6 and 9, these changes favorably affect many laser characteristics such as the threshold current, modulation bandwidth, frequency chirp, and the laser line width.

The use of an MQW structure in a phase-shifted DFB laser has considerably improved the performance of DFB lasers operating at 1.55 μm. Most of the improvements occur because of a lower value of the line-width enhancement factor β_c in MQW lasers compared with the conventional DFB lasers. In the case of a phase-shifted MQW DFB laser, lower values of β_c reduce the degradation caused by the spatial-hole burning,[115] a major benefit as far as the yield issue is concerned. This benefit can be understood by noting that the reduction in the carrier density near the phase-shift region caused by spatial-hole burning results in a smaller change in the refractive index (or the phase shift) in MQW DFB lasers simply because the index change is proportional to β_c. As a result, values of the MSR as large as 40–50 dB can be maintained over a large current range.[113] Such lasers can provide

20–30 mW of output power with a high MSR while maintaining a line width below 1 MHz.[116] In fact, line widths as small as 170 kHz have been obtained by using lasers with a 1.2-mm cavity length.[114] Figure 6.15 shows the line width of an 800-μm MQW DFB laser as a function of the output power. This laser attains a 270-kHz line width at an output power of only 13.5 mW.[115] MQW DFB lasers also show excellent modulation characteristics. The small-signal bandwidth can attain a value as large as 14 GHz with an optimization of the number and thickness of quantum wells.[113] Such lasers can be modulated at bit rates as high as 10 Gb/s and exhibit ultralow values of the frequency chirp because of relatively low values of β_c. Strained-layer MQW DFB lasers have also been developed since a proper amount of strain can further lower the line-width enhancement factor β_c and result in even better performance.[118] MQW DFB lasers have been used in many transmission experiments. These experiments show the potential application of such lasers in 1.55-μm lightwave systems operating at 10 Gb/s.

7.4.5 Gain-Coupled DFB Lasers

Distributed feedback in most DFB lasers is generated by introducing a periodic variation of the modal refractive index. However, distributed feedback can also occur when the modal gain varies periodically along the cavity length. Such lasers are called gain-coupled DFB lasers since the coupling between the forward- and backward-propagating waves is provided by the gain in place of the refractive index. Their properties were analyzed by Kogelnik and Shank[33] in 1972. Attempts were made soon after to fabricate gain-coupled DFB lasers by etching the active layer to provide a periodic modulation of its thickness.[7] Such attempts were not very successful because of the defects introduced during the etching process. Interest in gain-coupled DFB lasers revived in the late 1980s because of the yield problems associated with index-coupled DFB lasers (see Fig. 7.13). Extensive studies began and have shown there are many properties that make gain-coupled DFB lasers superior to index-coupled DFB lasers.[120–132] This section reviews their performance.

The coupled-mode equations of Sec. 7.3.1 can be used to understand the performance advantages of gain-coupled DFB lasers.[120–123] The main difference is that the coupling coefficient κ is no longer real, but becomes complex. It can be written as $\kappa = \kappa_r + i\kappa_i$, where κ_r and κ_i represent the contributions of index and gain gratings, respectively. In the case of pure gain coupling, $\kappa_r = 0$, and κ is imaginary. The lowest-threshold mode of such a laser lies exactly at the center of the stop band associated with index-coupled DFB lasers.[33] The gain margin $\Delta\alpha L$ depends on the value of $\kappa_i L$ and exceeds 0.2 even for its relatively low values ($\kappa_i L > 0.5$). This feature is

similar to the case of phase-shifted DFB lasers. However, in contrast with phase-shifted DFB lasers the gain margin is not very sensitive to facet reflections or to the grating phases at the facets.[120] As a result, a relatively high yield can be obtained even for lasers whose facets are left as cleaved. The impact of spatial-hole burning is also reduced in gain-coupling DFB lasers.[123] These advantages are maintained even in the case of partial gain coupling, a case commonly encountered in practice, as long κ_r is properly optimized (in the range 1.5–2). With a proper design, the yield of gain-coupled DFB lasers can exceed 80% even after taking into account the effect of spatial-hole burning.[123]

The simplest way to induce gain coupling in a DFB laser is to insert an absorbing layer between the transparent cladding layer and the substrate (see Fig. 7.7) before the grating is etched. The absorbing layer is partially etched away during the grating formation, resulting in an internal absorption that has the grating periodicity. Indeed, this was the approach adopted for the demonstration of a gain-coupled DFB laser in 1989.[120] Periodic modulation of the internal loss is equivalent to gain modulation along the cavity length since the gain has to be higher in the high-loss region to overcome the cavity losses at threshold. Such a gain grating is invariably accompanied by an index grating since the structure is basically the same as the one used for index-coupled DFB lasers. The coupling coefficient κ is not only complex but is generally dominated by the index coupling ($\kappa_r \gg \kappa_i$). The experimental data revealed that even small values of κ_i improve the device performance significantly.[120]

It is possible to design a DFB laser in which the mode coupling is almost completely provided by the gain ($\kappa_r \ll \kappa_i$).[122] The basic idea is to grow the cladding and active layers on top of the etched grating in such a way that the thickness of both layers is spatially modulated along the cavity length, resulting in two gratings. The bottom grating provides only index coupling ($\kappa_i = 0$), but the top grating formed by thickness variations of the active layer has a complex coupling coefficient because both the gain and the index are periodically modulated. Furthermore, the two index gratings are not in phase, making it possible nearly to cancel their contributions with a proper design. Such a DFB laser has almost pure gain coupling and has a number of attractive features.[122] Gain-coupled 1.55-μm DFB lasers with a corrugated active layer have been found to have excellent operating characteristics[127] with threshold currents as low as 12 mA and an MSR as high as 55 dB. The same technique can also be used to produce gain-coupled single-quantum-well lasers by growing an ultrathin active layer whose thickness is periodic along the cavity length.[128]

The main advantage of gain-coupled DFB lasers is that their performance is relatively unaffected by the facet reflections, and they can be used without

requiring an antireflection coating at the facets. The insensitivity to facet reflections results in a high yield since the grating phases at the facets do not play an important role in determining the laser performance. Such lasers are also relatively immune to the external feedback.[124] This insensitivity to the facet or external reflections can be understood qualitatively by noting that the laser can adjust its axial intensity distribution to take advantage of the gain grating. A gain-coupled DFB laser uses its gain most effectively if the intensity is high in the high-gain regions and low in the low-gain regions. Thus, the intensity distribution is determined by the gain grating in contrast with index-coupled lasers where facet reflections and phase-shift locations determine the intensity distributions. For this reason, spatial-hole burning is less important in gain-coupled DFB lasers. The effective line-width enhancement factor β_c is also affected by the gain grating and can be made quite small under certain conditions.[130,131] These properties of gain-coupled DFB lasers make them an ideal device for many applications in the field of lightwave technology.[132]

7.5 DBR LASERS

As mentioned in Sec. 7.1, the DBR laser provides an alternative scheme in which the frequency dependence of the distributed-feedback mechanism is utilized to select a single longitudinal mode of an FP cavity. In contrast to DFB lasers, the grating in a DBR laser is etched outside the active region (see Fig. 7.1). In effect, a DBR laser is an FP laser whose mirror reflectivity varies with wavelength; lasing occurs at the wavelength for which the reflectivity is maximum. In this section we briefly describe the operating principle and characteristics of DBR lasers. For a detailed discussion we refer to Ref. 133.

7.5.1 Design Issues

Historically DBR lasers developed[9–11] in parallel with DFB lasers. A problem inherent in the DBR laser is that when the unpumped active material is used to etch the grating at both ends, optical losses inside the DBR region are high and the resulting DBR reflectivity is poor. The problem of material loss can be overcome by using a material for the distributed Bragg reflector that is relatively transparent at the laser wavelength.[10] For InGaAsP lasers the InP substrate or a quaternary cladding layer (with a band gap higher than that of the active layer) can be used for this purpose. However, in a this case the distributed Bragg reflector and the active region form two separate waveguides, and transfer of the optical mode between them leads invariably to coupling losses that reduce the effective reflectivity of the DBR.

The design of a DBR laser involves minimization of the coupling losses, and many coupling schemes have been used for this purpose.[24] If C_0 is the power coupling efficiency between the DBR and the active waveguide, the effective amplitude-reflection coefficient of the DBR is given by

$$r_{\text{eff}} = C_0 r_g \tag{7.5.1}$$

where r_g is the amplitude-reflection coefficient of the DBR at the junction of the two waveguides. The subscript g emphasizes the grating origin of the reflectivity. The power coupling efficiency appears in Eq. (7.5.1) since coupling losses occur twice during each reflection.

7.5.2 Theory

An expression for r_g can be obtained using the analysis of Sec. 7.3. Consider a distributed Bragg reflector of length $\bar{L}$. A forward-propagating wave is incident at its surface where $z = 0$ and excites the counterpropagating waves $A(z)$ and $B(z)$ inside the reflector. The coupled-wave equations (7.3.15) and (7.3.16) govern their propagation, and the general solution is given by Eqs. (7.3.23) and (7.3.24). The amplitude-reflection coefficient r_g is obtained from

$$r_g = \frac{B(0)}{A(0)} = \frac{B_2 + r(q)A_1}{A_1 + r(q)B_2} \tag{7.5.2}$$

where q and $r(q)$ are given by Eqs. (7.3.21) and (7.3.22). If we use the boundary conditions $B(\bar{L}) = 0$ at the other end of the reflector, we obtain from Eq. (7.3.24) the relation

$$B_2 = r(q)A_1 \exp(2iq\bar{L}). \tag{7.5.3}$$

Using Eq. (7.5.3) in Eq. (7.5.2), we obtain

$$r_g = \frac{r(q)[1 - \exp(2iq\bar{L})]}{1 - r^2(q)\exp(2iq\bar{L})}. \tag{7.5.4}$$

This expression can be written in terms of κ and $\Delta\beta$ using Eq. (7.3.22); after some simplifications, it becomes

$$r_g = |r_g| \exp(i\phi) = \frac{i\kappa \sin(q\bar{L})}{q\cos(q\bar{L}) - i\Delta\beta \sin(q\bar{L})} \tag{7.5.5}$$

where, similar to Eq. (7.3.33),

$$\Delta\beta = \delta + i\alpha_g/2. \tag{7.5.6}$$

The imaginary part of $\Delta\beta$ is positive to account for the material losses inside the DBR medium, and α_g is the corresponding power-absorption coefficient. The parameter δ given by Eq. (7.3.34) accounts for the detuning of the laser wavelength from the Bragg wavelength. Figure 7.18 shows the wavelength dependence of the power reflectivity $|r_g|^2$ and the phase ϕ of a reflector such that $\alpha_g \bar{L} = 0.1$. The reflectivity is maximum at the Bragg wavelength ($\delta = 0$), and a 50% reflectivity can be obtained even at the relatively small value for κL of 1. Of course, the coupling loss C_0 in Eq. (7.5.1) would reduce the effective DBR reflectivity by C_0^2.

The threshold gain and the longitudinal modes of a DBR laser can be obtained in a manner similar to that discussed for FP lasers in Sec. 2.3. We assume for simplicity that the two DBRs are identical and can be described through the same effective amplitude-reflection coefficient r_{eff}. If we set the net change in the field amplitude after one round trip to unity, the threshold condition, similar to Eq. (2.3.6), becomes

$$(r_{\text{eff}})^2 \exp(2i\beta L) = 1 \tag{7.5.7}$$

where

$$\beta = \bar{\mu} k_0 - i\bar{\alpha}/2 \tag{7.5.8}$$

and is the mode-propagation constant. Equating the modulus and the phase on the two sides of Eq. (7.5.7), we obtain

$$C_0^2 |r_g|^2 \exp(\bar{\alpha} L) = 1 \tag{7.5.9}$$

$$\bar{\mu} k_0 L + \phi = m\pi \tag{7.5.10}$$

where we have used Eq. (7.5.1) and ϕ is the phase of r_g. As in the case of an FP laser, it is convenient to define the DBR loss as

$$\alpha_{\text{DBR}} = \frac{1}{L} \ln \frac{1}{C_0^2 |r_g|^2}. \tag{7.5.11}$$

Equation (7.5.9) then simply becomes $\bar{\alpha} = \alpha_{\text{DBR}}$. If we use Eq. (7.3.6) for $\bar{\alpha}$, the material gain at threshold is given by

$$g_{\text{th}} = (\alpha_{\text{DBR}} + \alpha_{\text{int}})/\Gamma. \tag{7.5.12}$$

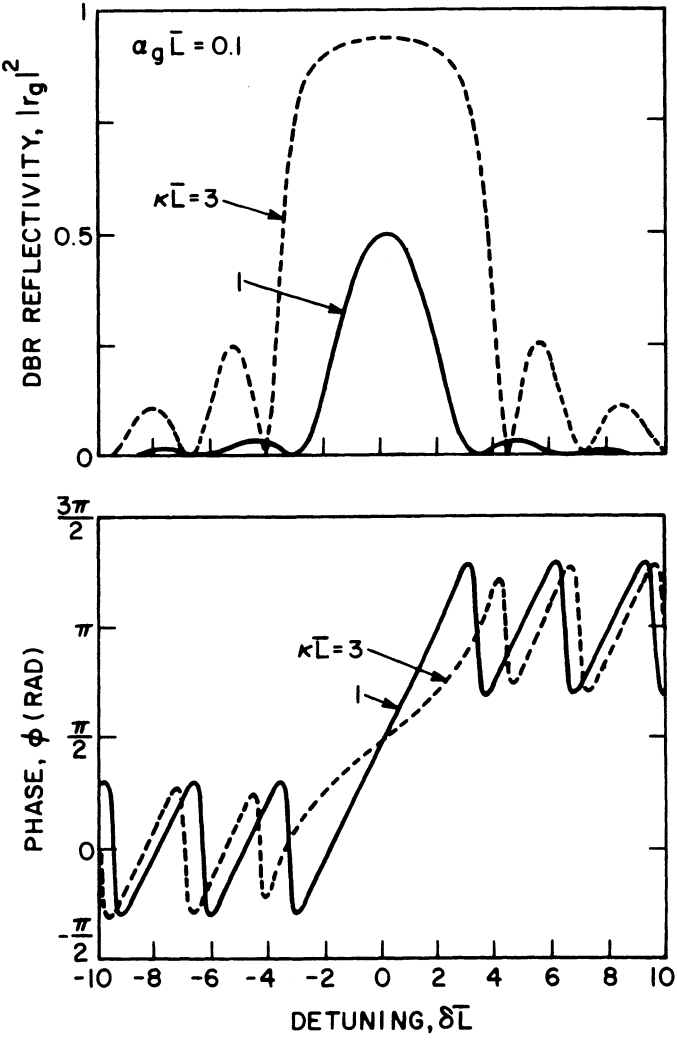


Fig. 7.18 Power reflectivity and phase of a distributed Bragg reflector as a function of detuning from the Bragg wavelength for two values of κL. (After Ref. 133)

The phase equation (7.5.10) determines the longitudinal modes of a DBR laser. In contrast to FP lasers, however, the modes are not equally spaced since ϕ depends on the detuning δ of the mode wavelength from the Bragg wavelength. Similarly, the DBR loss α_{DBR} is different for different longitudinal modes and the lowest-threshold mode occurs at the wavelength for which α_{DBR} is the smallest.

Figure 7.19 shows the longitudinal modes and their respective threshold gains for a DBR laser.[24] Equation (7.5.5) was used to obtain $|r_g|$ and ϕ as a function of the detuning δ. The normalized DBR loss $\alpha_{DBR}L$ was then obtained using Eq. (7.5.11). In obtaining Fig. 7.19, we choose C_0 equal to 1 after assuming no coupling losses. The lowest-threshold mode occurs close

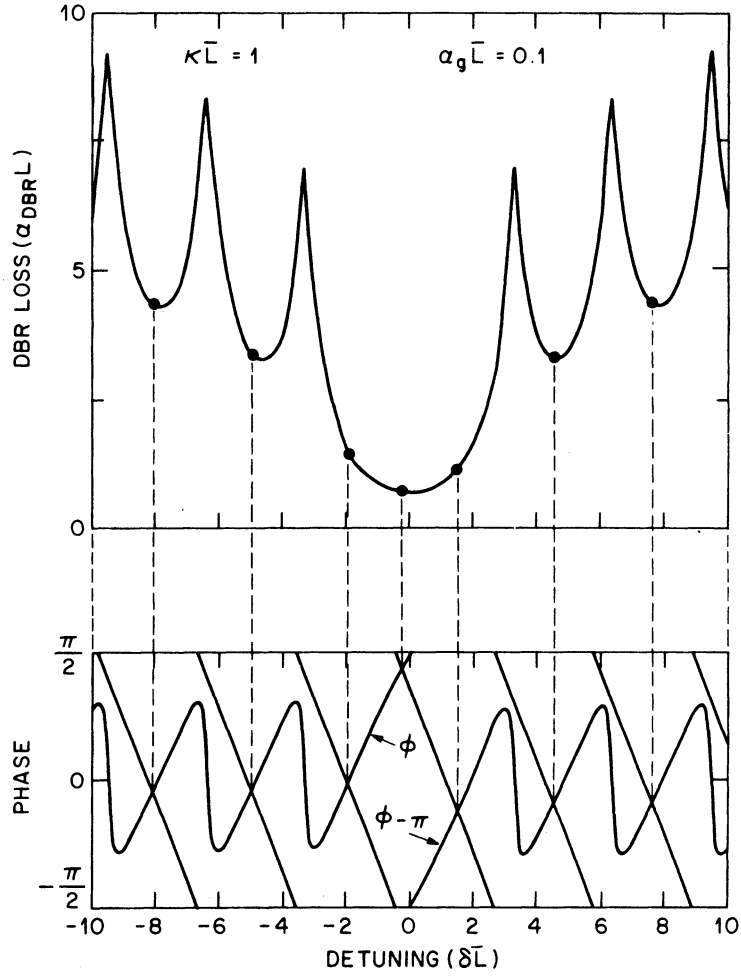


Fig. 7.19 Longitudinal modes and their respective threshold gains for a DBR laser. Variation of the amplitude-reflection coefficient of the distributed Bragg reflector with the detuning $\delta\bar{L}$ is used to obtain the phase and the DBR loss. The bottom part shows graphical solutions (points of intersection) of the phase condition; various diagonal lines correspond to different integer values of m in Eq. (7.5.10). Solid dots in the top part show the DBR loss, or equivalently the threshold gain, for various DBR modes. (After Ref. 24 © 1983 IEEE)

to the Bragg wavelength since the DBR reflectivity peaks at that wavelength. In contrast to DFB lasers, DBR lasers do not exhibit a stop band in their longitudinal-mode spectra. This is because no distributed feedback occurs inside the DBR laser cavity. Conceptually, a DBR laser is just an FP laser whose end mirrors exhibit frequency-dependent reflectivity.

7.5.3 Emission Characteristics

DBR semiconductor lasers operating at 1.55 μm were developed during the 1980s in parallel with DFB lasers.[133–153] A buried-heterostructure design is used in order to reduce the threshold current and to maintain single-mode operation in the lateral and transverse directions. The first-order gratings are generally formed in the DBR regions with a coupling coefficient κ of $\sim 100\ \text{cm}^{-1}$.

The emission characteristics of DBR lasers are similar to those of DFB lasers.[24] Relative to DFB lasers, the threshold current of a DBR laser is generally high because of coupling losses, which increase α_{DBR} given by Eq. (7.5.11). For the same reason, the differential quantum efficiency is also relatively low. Several different designs have been used[24] to reduce the coupling losses in DBR lasers, and threshold currents below 100 mA are readily achieved.

The performance of a DBR laser is comparable to that of a DFB semiconductor laser as far as the spectral and dynamic properties are concerned. The longitudinal mode closest to the Bragg wavelength has the lowest threshold gain (see Fig. 7.18) and becomes the dominant mode. Because of the significant gain margin ($\sim 8\ \text{cm}^{-1}$), the other longitudinal modes are suppressed by about 30 dB relative to the main mode. These side modes remain suppressed under high-speed modulation even though the mode suppression ratio degrades somewhat as the modulation depth increases. In a structure, referred to as a bundle-integrated-guide DBR laser,[134] side modes remained suppressed by more than 30 dB for modulation frequencies as high as 2 GHz while, at the same time, the frequency chirp was below 0.2 nm. The spectral line width under CW operation is also quite small for DBR lasers. In one device, a line width of 3.2 MHz was measured at a relatively low output power of 1.5 mW.[138] The line width of a DBR laser depends on the design details such as whether one or two Bragg reflectors are used and whether the Bragg reflectors are adjacent to the active region or separated by a passive region. A line width of 560 kHz was realized in one DBR laser whose structure was properly optimized.[139]

A new type of DBR laser has been developed to improve the performance of DBR lasers. This structure, referred to as the distributed reflector (DR) laser,[140–142] is a combination of DFB and DBR structures since two separate

gratings of different pitches are fabricated in the active and passive sections of such a device. The main advantages of this structure are (i) almost 100% coupling efficiency between the active and passive sections, (ii) a high differential quantum efficiency because of a low internal loss, and (iii) a low chirp under direct modulation because of a reduction in the effective line-width enhancement factor. Threshold currents in the range 30–35 mA and differential quantum efficiencies in the range 15–20% have been realized in such lasers.[141] The effective line-width enhancement factor is found to be smaller by about a factor of 2, resulting in low chirp under direct modulation and a low line width under CW operation. However, it is difficult to obtain line widths below 1 MHz. The use of an MQW active region has resulted in a line width of 743 kHz in one device that also exhibited an MSR in excess of 40 dB.[142] DBR lasers with an MSR as high as 58 dB have been fabricated by using the technique of chemical beam epitaxy.[146] In another record performance, the use of a strained-layer quantum-well InGaAs active layer has resulted in 0.98-μm DBR lasers capable of operating at power levels up to 42 mW with a differential quantum efficiency as high as 47%.[147] DBR lasers operating in the visible region near 0.6 μm have also been fabricated by using a strained InGaP quantum well as the active layer.[148] In a new class of DBR lasers the light is emitted in a direction perpendicular to the wafer surface. Such DBR lasers are known as surface-emitting lasers and are discussed in Chapter 10.

In a different approach, hybrid-type DBR lasers have been developed.[149–153] These lasers are coupled-cavity type (see Chapter 8) in the sense that an external Bragg reflector is coupled to a conventional multimode semiconductor laser. In one scheme, a silicon-chip Bragg reflector is used to provide the distributed feedback.[149] Line widths below 0.5 MHz have been obtained by such a hybrid approach, and even smaller line widths are possible.[150] Such DBR lasers are also capable of producing picosecond optical pulses through mode locking.[151] In another approach the grating is etched directly onto a fiber that is coupled to the multi-mode laser.[152] Such a fiber-DBR laser has been used to produce 18.5-ps pulses at a repetition rate of 2.37 GHz.[153] The main disadvantage of such hybrid DBR lasers is related to their not having a monolithic design. It is possible to design a monolithic extended-cavity DBR laser by etching the grating directly onto a passive waveguide formed on the same substrate used for fabrication of the laser. In one 1.55-μm device, the 0.6-mm long Bragg mirror was separated from the 0.7-mm active region by a 4.2-mm passive waveguide, resulting in a 5.5-mm-long DBR laser.[144] All three components were formed on the same InP substrate to preserve the monolithic nature of the device. This laser produced 20-ps optical pulses at an 8.1-GHz repetition rate through active mode locking. In another implementation,[145] a 1.2-mm-long MQW DBR

laser provided a 200-kHz line width at 17-mW operating power. The laser has a low threshold current (8 mA) and could be operated up to 28 mW. Coherent and multichannel communication systems[154] often require semiconductor lasers which not only operate in a single longitudinal mode with a narrow line width but whose wavelength can also be tuned over as wide a range as possible. Next to be discussed are multisection DFB and DBR lasers, which have been developed to meet these requirements.

7.6 TUNABLE SEMICONDUCTOR LASERS

Tunable semiconductor lasers are useful for many applications. Examples of their applications in fiber-optic communication systems[154] are (i) wavelength-division multiplexed lightwave systems where optical signals at many distinct wavelengths are simultaneously modulated and transmitted through a fiber and (ii) coherent communication systems which require wavelength matching between the local oscillator and the transmitter laser. The simplest kind of tunable semiconductor lasers make use of an external cavity formed by placing a grating at some distance from a multimode laser. Such external-cavity semiconductor lasers (see Sec. 8.4.5) can be tuned over a wide range (up to 80 nm) by simply rotating the grating.[155] However, their use in practical lightwave systems is limited because of their nonmonolithic design. For this reason considerable effort was spent during the 1980s to develop tunable multisection DFB and DBR lasers.[156–177] This section reviews the performance of such devices.

The basic idea behind the tunability of DFB and DBR lasers is simple. The operating wavelength of such lasers is determined by the etched grating through the Bragg wavelength given by $\lambda_0 = 2\bar{\mu}\Lambda/m$, where $\bar{\mu}$ is the mode index, Λ is the grating period, and m is the diffraction order of the grating [see Eq. (7.2.1)]. Even through the grating period L is fixed during grating fabrication, the laser wavelength can be changed by changing the mode index $\bar{\mu}$. A simple way to change the mode index is to change the material index by injecting current into the grating region, since the refractive index depends on the injected carrier density. In general, the refractive index decreases with an increase in the carrier density, resulting in a shift of the Bragg wavelength (and hence the laser wavelength) toward shorter wavelengths with an increase in the injected current.

Several schemes can be used to design tunable DFB and DBR devices. For example, a DBR laser with one Bragg reflector, designed such that different currents can be injected into the active and Bragg sections allows wavelength tuning by varying the current in the passive Bragg section. However, such two-section lasers allow only discrete tuning. Continuous tuning is achieved by using a multisection DBR laser shown schematically in

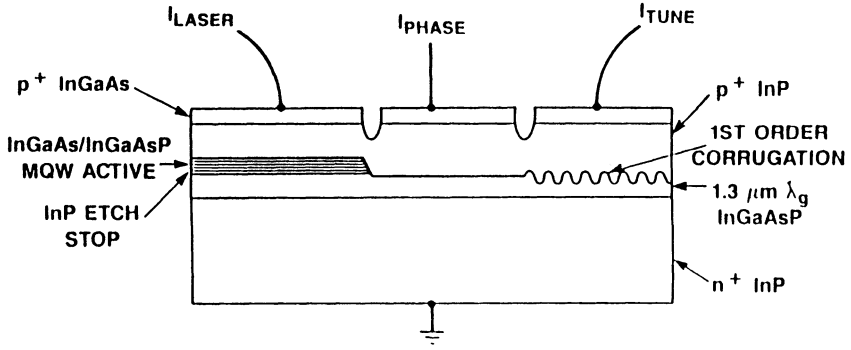


Fig. 7.20 Schematic illustration of a multisection DBR laser. The three sections are coupled by the thick waveguide layer below the MQW active layer. (After Ref. 168 © 1990 IEEE)

Fig. 7.20. It consists of three sections, referred to as the active section, the phase-control section, and the Bragg section. Each section can be biased independently by injecting different amounts of current. The current injected into the Bragg section is used to change the Bragg wavelength through carrier-induced changes in the refractive index. The current injected into the phase-control section is used to change the phase of the feedback from the DBR through carrier-induced index changes in that section. The laser wavelength can be tuned continuously over a range 5–10 nm by controlling the currents injected in the three sections.[157–165] The performance of a three-section DBR laser with 80-mm-long phase-control section is shown in Fig. 7.21 where the measured device wavelength is plotted as a function of the phase-section current.[168] The current in the Bragg section is also adjusted by a small amount to maximize the sweep range. This laser is tunable continuously over more than 6 nm by this technique.

It is possible to extend the tuning range by reverse biasing the Bragg section. A tuning range of 22 nm was obtained in a three-section DBR laser[172] by changing the current from -120 to 120 mA, where the negative and positive currents correspond to reverse and forward biasing of the Bragg section. The physical mechanism behind reverse-biased tuning is local heating of the Bragg section. Since the refractive index increases with an increase in the temperature, the laser wavelength shifts toward the red side, in contrast with the forward biasing that results in a blue shift because of the carrier-induced reduction in the refractive index. A disadvantage of the thermal-tuning mechanism is that it is relatively slow, i.e., wavelength changes occur at a time scale of ~ 100 μs compared with the case of forward biasing where carrier-induced tuning occurs at a time scale of ~ 1 ns. It is nonetheless fast enough for some applications.[172]

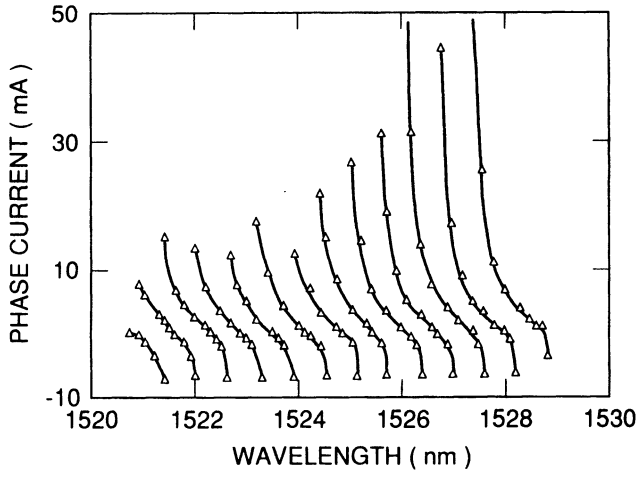


Fig. 7.21 Wavelength tuning characteristics of a multisection DBR laser. Measured wavelength is plotted as a function of current in the phase-control section. The current in the Bragg section is also adjusted by a small amount to maximize the sweep range. (After Ref. 168 © 1990 IEEE)

For applications in coherent communication systems the laser should exhibit a narrow line width. Considerable theoretical work has been done to understand the performance of multisection DBR lasers.[166,169,174] A formalism based on a Green's-function approach[178] is generally used to calculate the line width since it allows inclusion of the effects of spatial inhomogeneities and spatial-hole burning.[169,174] The results show that the effective line-width enhancement factor of a tunable DBR laser depends on many device parameters and can change with the currents applied to the phase-control and Bragg sections.[174] As a result, the line width is expected to change considerably with wavelength tuning. Figure 7.22 shows the line width as a function of wavelength for the same laser whose tuning curves are shown in Fig. 7.21.[168] Although the line width is quite small (in the range 5–10 MHz), it becomes quite large just before the laser jumps to a neighboring mode. The reason behind this change in line width is related to a reduction in the Q factor of the cavity. It should be noted that the current in the active section was kept constant in obtaining the data shown in Fig. 7.22. A part of the increase in the line width is due to a decrease in the output power with tuning. Line-width variations are somewhat smaller if the tunable DBR laser is operated at a constant output power. Modulation characteristics of multisection DFB and DBR lasers have also been studied. A useful property is their frequency modulation (FM) response. This can be quite uniform over a wide frequency range. Because of their attractive characteristics such

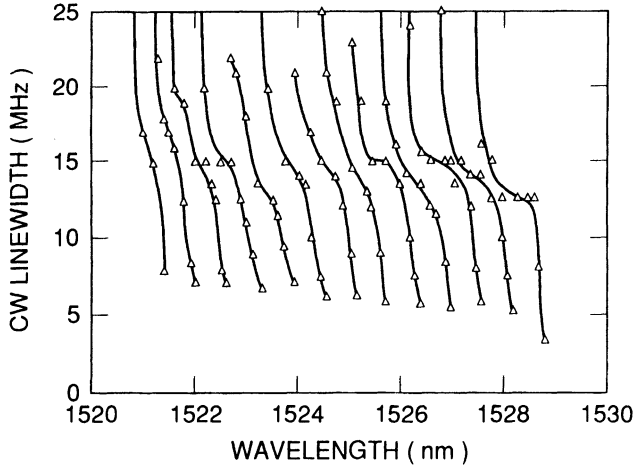


Fig. 7.22 Measured CW line width as a function of wavelength for the multisection DBR laser whose tuning characteristics are shown in Fig. 7.21. (After Ref. 168 © 1990 IEEE)

semiconductor lasers have been used in many transmission experiments. The next section discusses the importance of DFB lasers for lightwave communication systems.

7.7 TRANSMISSION EXPERIMENTS

As mentioned earlier, the main thrust for the development of 1.55-μm InGaAsP DFB lasers comes from their potential application in long-haul, high-bit-rate optical communication systems employing conventional silica fibers. This potential was clearly demonstrated in 1984 in a transmission experiment[178] where a 2-Gb/s signal was transmitted over a 130-km-long fiber with an error rate less than 10^{-9}. Figure 7.23 shows the experimental setup. The single-mode fiber loss at the laser wavelength of 1.54 μm was 0.212 dB/km, while its dispersion at this wavelength amounted to 19 ps/(nm·km). The receiver employed an avalanche photodiode with a high-impedance GaAs field-effect-transistor preamplifier.

The DFB laser used in this experiment[178] was a vapor-phase-transported buried-heterostructure InGaAsP laser operating at 1.54 μm. Its temperature was held at 16°C, and the average dc bias current was 60.3 mA. The laser was modulated at 2 Gb/s with a pseudorandom nonreturn-to-zero (NRZ) bit pattern (with a length of $2^{15} - 1$ bits). The peak-to-peak amplitude of the modulation current was 32 mA. The DFB laser maintained an MSR value of 440 under modulation, assuring suppression of the side modes under

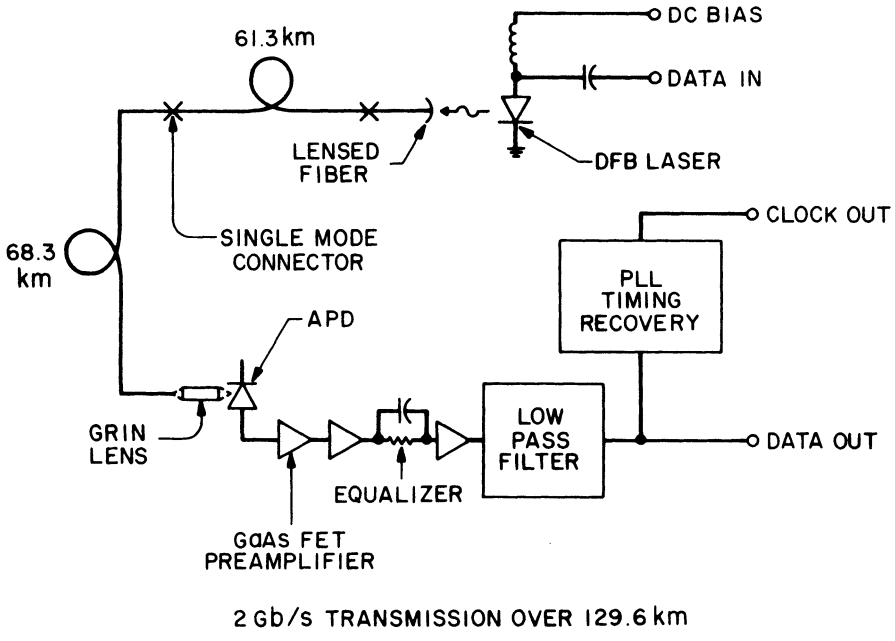


Fig. 7.23 Schematic of a transmission experiment used to demonstrate the application of 1.55-μm DFB semiconductor lasers in high-bit-rate long-distance optical communication systems. (After Ref. 178)

dynamic conditions. A separate measurement of the peak-to-peak wavelength chirp during modulation with 1-ns pulses gave a result of 0.04 nm. The *laser extinction ratio*, defined as the power ratio in the on and off states, was 3.3.

Figure 7.24 shows the bit-error-rate curves of the experiment.[178] (The dBm unit on the abscissa expresses the received power in decibels using 1 mW as the reference level; e.g., 1 mW = 0 dBm and 1 μW = −30 dBm.) The average power launched into the single-mode fiber was −3.8 dBm. The left curve in Fig. 7.24 illustrates the receiver performance without a fiber and with a laser extinction ratio of 20:1. It provides a reference for the power penalty that may occur due to chirp or low extinction ratio. The right-most curve, for a 10-km fiber, shows that a power penalty of 5 dB occurred because of a low extinction ratio of 3.3:1 used for the experiment. This extinction ratio was found to be optimum from the viewpoint of wavelength chirping. The middle curve, for a fiber length of 130 km, shows that the bit error rate was 8×10^{-10} at a received power level of −32 dBm. A surprising aspect of the experiment is the negative dispersion penalty observed; i.e., reception was 0.6 dB better for the 130-km fiber than for the 10-km fiber. This is believed

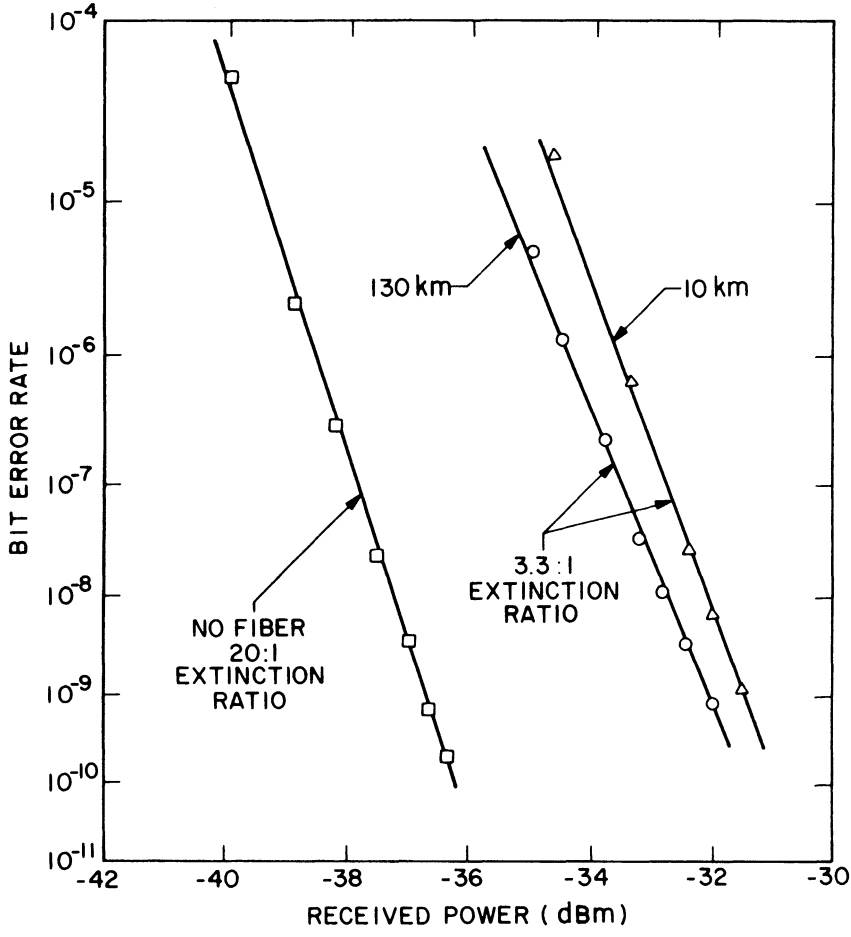


Fig. 7.24 Bit-error-rate curves for the 130-km transmission experiment at 2 Gb/s. The curve on the left serves as a reference and shows that a power penalty of about 4 dB occurring for a 130-km fiber can be attributed to the low extinction ratio used in the experiment. (After Ref. 178)

to be due to an imperfect initial equalization by the receiver circuitry (see Fig. 7.23); the chirp appears to aid in equalizing the system response slightly.

The results of Fig. 7.24 indicated in 1984 that DFB lasers can be used in 1.55-μm optical communication systems operating at a bit rate of 2 Gb/s and can provide a bit rate–distance (BL) product of 260 (Gb/s)·km. Since then, the performance of DFB lasers has improved considerably, and these

improvements have resulted in a better system performance. Several transmission experiments in 1989 showed that a BL product of up to 1 (Tb/s)·km can be realized by operating the lightwave system near the zero-dispersion wavelength of optical fibers.[179–181] In one experimental[179] a $\lambda/4$-shifted MQW DFB laser operating near 1.55-μm was used to transmit the 10-Gb/s signal over 100 km of the dispersion-shifted fiber. In a similar experiment[180] the bit rate was extended to 11 Gb/s, but the signal could be transmitted only over 81 km. Another experiment used a 1.3-μm DFB laser to transmit a 16-Gb/s signal over 64.5 km of conventional (nondispersion-shifted) fiber.[181] Signal transmission at 1.55 μm in conventional fibers is severely limited by the large chromatic dispersion. An experiment in 1991 used the technique of optical dispersion equalization in combination with a fiber preamplifier to transmit the 8-Gb/s signal over 130 km of conventional fiber.[182] Of course, the performance can be further improved by using dispersion-shifted fibers. Indeed, transmission over 220 km at 10 Gb/s was demonstrated in a later experiment by using the dispersion-shifted fiber.[183]

As mentioned earlier, DFB lasers are particularly useful for coherent communication systems that permit improved receiver sensitivity through homodyne or heterodyne detection techniques.[154] Many coherent transmission experiments[184–188] have shown the usefulness of DFB lasers in the design of such systems. Signal transmission at 4 Gb/s over 160 km of the conventional fiber was demonstrated in 1990 by using heterodyne detection with frequency-shift keying (FSK) or differential phase-shift keying (PSK).[184] The transmitter laser was a directly modulated DFB laser while a tunable DBR laser was used as the local oscillator. A similar experiment transmitted signal over 167 km at 4 Gb/s by using homodyne detection with PSK.[185] The bit rate was extended to 10 Gb/s by using a DFB laser whose output was phase-modulated through a $LiNbO_3$ external modulator.[186] The bit rate of coherent communication systems can be increased by transmitting several channels simultaneously. In one experiment[187] 4 channels, each operating at 2.5 Gb/s, were transmitted over 150 km by using MQW DFB lasers whose wavelengths were spaced by only $\sim$0.1 nm to provide a 15-GHz channel spacing. The effective BL product for this experiment was 1.5 (Tb/s)·km. Another experiment transmitted 100 channels over 50 km of fiber.[188] The channel spacing was 10 GHz at a relatively low bit rate of 622 Mb/s. The system capacity for this experiment exceeded 3 (Tb/s)·km. Such experiments would have been impossible without the development of high-performance DFB lasers.

Many experiments have been performed to demonstrate the usefulness of DFB lasers in local-area networks which require broadcasting and distribution of a large number of channels.[189–192] In one 16-channel experiment[189] the bit rate was 2 Gb/s. In another,[190] the bit rate was relatively low

($B = 622$ Mb/s), but 100 channels were distributed by using a 128×128 broadcast star. The main limitation of such systems results from large distribution losses. The distribution-loss problem can be overcome by the use of optical amplifiers. In one experiment[191] two erbium-doped fiber amplifiers were used to demonstrate a broadcast network capable of reaching 39.5 million subscribers. The experiment demonstrated the concept by using 12 DFB lasers with a wavelength spacing of 2 nm. Each laser could be modulated at 2.2 Gb/s. The whole system was capable of transmitting 384 digital video channels to 39.5 million subscribers. This concept was later extended to demonstrate a wavelength-division-multiplexed (WDM) broadcast network capable of reaching 43.8 million subscribers within a 507-km range while operating at 39.81 Gb/s.[192]

DFB lasers have also found applications in cable television (CATV) systems where a large number of video channels are distributed using the technique of subcarrier multiplexing (SCM).[154] Relative intensity noise (RIN) of semiconductor lasers drastically affects the performance of such systems; DFB lasers are preferred since they generally have a lower RIN than multimode semiconductor lasers. Several laboratory experiments have demonstrated the potential of analog SCM lightwave systems for video transmission over optical fibers. In one AM experiment, 42 video channels with 6-MHz channel spacing were transmitted by using microwave subcarriers in the 50–400 MHz frequency range.[193] Another experiment demonstrated transmission of eight FM channels over 100 km by directly modulating a 1.537-μm DFB laser.[194] Yet another experiment[195] used digital modulation for 20 FSK video channels, operating at 100 Mb/s with 200-MHz channel spacing. In an interesting demonstration of the usefulness of DFB lasers,[196] 16 DFB lasers with wavelength spacing of 2 nm in the 1.55-μm band were modulated with 100 analog FM video channels and six 622-Mb/s digital channels. Video channels were multiplexed by using the SCM technique such that one DFB laser carried 10 FM video channels over the bandwidth 300–700 MHz. Therefore 10 DFB lasers were used for 100 FM video channels and the remaining 6 lasers for 6 digital channels each operating at 622 Mm/s. The combined signal was split into 4096 parts by using a single erbium-doped fiber amplifier capable of providing 17–24-dB gain over the 34-nm spectral range.

DFB semiconductor lasers in combination with erbium-doped fiber amplifiers (EDFAs) have revolutionized the design and performance of lightwave systems.[154] An experiment in 1990 demonstrated transmission over 904 km at 1.2 Gb/s by using 12 EDFAs.[197] It was followed by a record-breaking coherent transmission experiment[198] in which the 2.5-Gb/s signal was transmitted over 2,223 km by using 25 EDFAs placed at approximately 80-km intervals, offering a total gain of more than 440 dB. The effective BL

product for this experiment exceeded 5.5 (Tb/s)·km, as the 2.5-Gb/s signal is transmitted without regeneration. Even larger values of BL are possible in principle. Indeed, an effective transmission distance of 21,000 km at the bit rate of 2.4 Gb/s and of 14,300 km at 5 Gb/s [BL = 71 (Tb/s)·km] was demonstrated in a laboratory experiment by using the recirculating fiber-loop configuration and by operating the transmission system very close to the zero-dispersion wavelength of the fiber to minimize the effects of chromatic dispersion.[199] Such performance may be difficult to realize under field conditions, but the experiment clearly demonstrates the potential capacity of lightwave systems that use DFB lasers together with cascaded optical amplifiers.

The detrimental effect of fiber dispersion can almost be eliminated in soliton communication systems.[154] Soliton transmission over 15,000 km at 5 Gb/s and over 11,000 km at 10 Gb/s has been demonstrated by using the recirculating fiber-loop configuration.[200] This experiment used an external-cavity mode-locked laser. In another scheme, a gain-switched DFB laser was used to generate the picosecond soliton bit stream. Signal transmission over 1,020 km at 20 Gb/s was realized in a laboratory experiment by this method.[201] The highest bit rate of 32 Gb/s was demonstrated in 1992 in another soliton experiment, but the transmission distance was only 90 km.[202] It is difficult to realize bit rates in excess of 10 Gb/s through nonsoliton lightwave systems. Transmission over 1,111 km has been realized, but only at a relatively low bit rate of 2.5 Gb/s.[203] However, several channels can be multiplexed to obtain a high bit rate. In one nonsoliton experiment[204] four 5-Gb/s channels were multiplexed to obtain a 20-Gb/s transmission rate, and the signal was transmitted over 137 km without any in-line amplifiers. It should be evident from these transmission experiments that DFB semiconductor lasers play an important role in the design of high-performance fiber-optic communication systems. They are likely to find many other applications in optical systems that require the use of a compact optical source with a narrow spectrum.

PROBLEMS

7.1 Describe how a grating can force a semiconductor laser to operate in a single longitudinal mode. What is the difference between DFB and DBR lasers?

7.2 Explain how thickness variations of a cladding layer create an index grating in a DFB laser. How are such thickness variations produced in practice? What is the grating period for a first-order grating for lasers operating at 0.85 and 1.55 μm?

7.3 Use the coupled-wave equations to prove that an infinite-periodic structure has a stop band centered at the Bragg wavelength such that light can propagate only if its wavelength lies outside the stop band. Determine the wavelength range $\Delta\lambda$ of the stop band in terms of the coupling coefficient κ.

7.4 Derive the threshold condition for a DFB laser with finite facet reflectivities. Show that it reduces to the threshold condition obtained in Sec. 2.3 for FP lasers in the limit $\kappa = 0$.

7.5 Use the coupled-wave equations to obtain the longitudinal modes and their threshold gains for a DFB laser of length L with nonreflecting facets ($R_1 = R_2 = 0$). For what values of the coupling coefficient κ is the threshold gain of such a laser lower than that of an FP laser with 32% facet reflectivities?

7.6 Solve Eq. (7.3.30) numerically to obtain a figure similar to Fig. 7.6 for the case $R_1 = 0.1\%$ and $R_2 = 80\%$. Use $\phi_1 = 0$ and vary ϕ_2 over its entire range from 0 to 2π.

7.7 Derive Eq. (7.3.46) for the coupling coefficient κ starting from its definition given by Eq. (7.3.42).

7.8 Study the dependence of κ on the grating depth for the case of a sinsuoidal grating.

7.9 Why are the coupling coefficients different for the TE and TM modes? Derive an expression for κ in the case of TM-mode operation of a DFB laser.

7.10 The performance of a DFB laser is very sensitive to the grating phases ϕ_1 and ϕ_2 at the facets. Write a computer program and generate the yield curves shown in Fig. 7.13.

7.11 Use the two-mode rate equations to study the dependence of the MSR on the loss difference between the main and side modes. Generate curves similar to those shown in Fig. 7.15 for the case of 5-Gb/s modulation.

7.12 What are the advantages and disadvantages of a phase-shifted DFB laser over the conventional DFB laser?

7.13 Derive the threshold condition of a phase-shifted DFB laser having a phase shift ϕ_{sh} exactly at the cavity center. Calculate the gain margin as a function of ϕ_{sh} by assuming nonreflecting facets ($R_1 = R_2 = 0$) for a laser with $\kappa L = 1$ and 2.

7.14 Obtain the frequencies and threshold gains of the two modes located closest to the Bragg frequency for a gain-coupled DFB laser using coupled-wave equations with $\kappa = i\kappa_g$.

7.15 Repeat Problem 7.14 for the case of a complex coupling coefficient given by $\kappa = \kappa_i + i\kappa_g$. Plot the gain margin as a function of κ_g/κ_i for $|\kappa|L = 1$ and 2.

REFERENCES

1. Agrawal, G. P. Vol. 26 in *Progress in Optics*, ed. by E. Wolf. Amsterdam: North-Holland Pub. Co., 1988.
2. See, for example, Marcuse, D., *Light Transmission Optics*. New York: Van Nostrand Reinhold, 1982, Chap. 2.
3. Kogelnik, H., and C. V. Shank. *Appl. Phys. Lett.* **18,** 152 (1971).
4. Nakamura, M., A. Yariv, H. W. Yen, S. Somekh, and H. L. Garvin. *Appl. Phys. Lett.* **22,** 515 (1973); **23,** 224 (1973).
5. Shank, C. V., R. V. Schmidt, and B. I. Miller. *Appl. Phys. Lett.* **25,** 200 (1974).
6. Scifres, D. R., R. D. Burnham, and W. Streifer. *Appl. Phys. Lett.* **25,** 203 (1974).
7. Nakamura, M., K. Aiki, J. Umeda, A. Yariv, H. W. Yen, and T. Morikawa. *Appl. Phys. Lett.* **25,** 478 (1974).
8. Anderson, D. B., R. R. August, and J. E. Coker. *Appl. Opt.* **13,** 2742 (1974).
9. Casey, H. C., Jr., S. Somekh, and M. Ilegems. *Appl. Phys. Lett.* **27,** 142 (1975).
10. Reinhart, F. K., R. A. Logan, and C. V. Shank. *Appl. Phys. Lett.* **27,** 45 (1975).
11. Tsang, W. T., and S. Wang. *Appl. Phys. Lett.* **28,** 596 (1976).
12. Burnham, R. D., D. R. Scifres, and W. Streifer. *Appl. Phys. Lett.* **29,** 287 (1976).
13. Aiki, K., M. Nakamura, and J. Umeda. *IEEE J. Quantum Electron.* **QE-12,** 597 (1976).
14. Suematsu, Y., M. Yamamoto, and Y. Unno. *IEEE J. Quantum Electron.* **QE-13,** 646 (1977).
15. Kawanishi, H., Y. Suematsu, Y. Itaya, and S. Arai. *Jpn. J. Appl. Phys.* **17,** 1439 (1978); *IEEE J. Quantum Electron.* **QE-15,** 701 (1979).
16. Doi, A., T. Fukuzawa, M. Nakamura, R. Ito, and K. Aiki. *Appl. Phys. Lett.* **35,** 441 (1979).
17. Utaka, K., K. Kobayashi, K. Kishino, and Y. Suematsu. *Electron. Lett.* **16,** 455 (1980).
18. Mikami, O. *Jpn. J. Appl. Phys.* **20,** L488 (1981).
19. Utaka, K., S. Akiba, K. Sakai, and Y. Matsushima. *Electron. Lett.* **17,** 961 (1981).
20. Sakai, K., K. Utaka, S. Akiba, and Y. Matsushima. *IEEE J. Quantum Electron.* **QE-18,** 272 (1982).
21. Itaya, Y., T. Matsuoka, Y. Nakano, Y. Suzuki, K. Kuroiwa, and T. Ikegami. *Electron. Lett.* **18,** 1006 (1982).
22. Akiba, S., K. Utaka, K. Satai, and Y. Matsushima. *IEEE J. Quantum Electron.* **QE-19,** 1052 (1983); *IEEE J. Quantum Electron.* **QE-20,** 236 (1984).
23. Kitamura, M., M. Seki, M. Yamaguchi, I. Mito, K. E. Kobayashi, and K. O. Kobayashi. *Electron. Lett.* **19,** 840 (1983).
24. Suematsu, Y., S. Arai, and K. Kishino. *J. Lightwave Technol.* **LT-1,** 161 (1983).
25. Itaya, Y., T. Matsuoka, K. Kuroiwa, and T. Ikegami. *IEEE J. Quantum Electron.* **QE-20,** 230 (1984); Itaya, Y., K. Wakita, G. Motosugi, and T. Ikegami. *IEEE J. Quantum Electron.* **QE-21,** 527 (1985).
26. Kitamura, M., M. Yamaguchi, S. Murata, I. Mito, and K. Kobayashi. *J. Lightwave Technol.* **LT-2,** 363 (1984) *Electron. Lett.* **20,** 595 (1984).
27. Temkin, H., G. J. Dolan, N. A. Olsson, C. H. Henry, R. A. Logan, R. F. Kazarinov, and L. F. Johnson. *Electron. Lett.* **45,** 1178 (1984).
28. Kitamura, M., M. Yamaguchi, S. Murata, I. Mito, and K. Kobayashi. *Electron. Lett.* **20,** 595 (1984).
29. Oishi, M., M. Nakano, Y. Itaya, K. Sato, and Y. Imamura. *IEEE J. Quantum Electron.* **QE-23,** 822 (1987).
30. Itaya, Y., H. Saito, G. Motosugi, and Y. Tohmori. *IEEE J. Quantum Electron.* **QE-23,** 828 (1987).
31. Takigawa, S., T. Uno, M. Kume, K. Hamada, N. Yoshikawa, H. Shimizu, and G. Kano. *IEEE J. Quantum Electron.* **QE-25,** 1489 (1989).

32. Ketelsen, L. J. P., I. Hoshini, and D. A. Ackerman. *IEEE J. Quantum Electron.* **QE-27,** 957 (1991); **27,** 965 (1991).
33. Kogelnik, H., and C. V. Shank. *J. Appl. Phys.* **43,** 2327 (1972).
34. Wang, S. *J. Appl. Phys.* **44,** 767 (1973).
35. DeWames, R. E., and W. F. Hall. *Appl. Phys. Lett.* **23,** 28 (1973).
36. Yariv, A. *IEEE J. Quantum Electron.* **QE-9,** 919 (1973).
37. Subert, R. *J. Appl. Phys.* **45,** 209 (1974).
38. Chinn, S. R., and P. L. Kelly. *Opt. Commun.* **10,** 123 (1974).
39. Wang, S. *IEEE J. Quantum Electron.* **QE-10,** 413 (1974); *Wave Electron.* **1,** 31 (1974).
40. Nakamura, M., and A. Yariv. *Opt. Commun.* **11,** 18 (1974).
41. Iga, K., and K. Kawabata. *Jpn. J. Appl. Phys.* **14,** 427 (1975).
42. Streifer, W., R. D. Burnham, and D. R. Scrifres. *IEEE J. Quantum Electron.* **QE-11,** 154 (1975).
43. Elachi, C., and G. Evans. *J. Opt. Soc. Am.* **65,** 404 (1975).
44. Okuda, M., and K. Kubo. *Opt. Commun.* **13,** 363 (1975).
45. Streifer, W., D. R. Scifres, and R. D. Burnham. *IEEE J. Quantum Electron.* **QE-11,** 867 (1975); *IEEE J. Quantum Electron.* **QE-12,** 74 (1976).
46. Hill, K. O., and A. Watanabe. *Appl. Opt.* **14,** 950 (1975); Hill, K. O., and R. I. MacDonald. *IEEE J. Quantum Electron.* **QE-12,** 716 (1976).
47. Haus, H. A., and C. V. Shank. *IEEE J. Quantum Electron.* **QE-12,** 532 (1976).
48. Streifer, W., R. D. Burnham, and D. R. Scifres. *IEEE J. Quantum Electron.* **QE-12,** 737 (1976).
49. Streifer, W., D. R. Scifres, and R. D. Burnham. *IEEE J. Quantum Electron.* **QE-13,** 134 (1977).
50. Wang, S. *IEEE J. Quantum Electron.* **QE-13,** 176 (1977).
51. Yariv, A., and M. Nakamura. *IEEE J. Quantum Electron.* **QE-13,** 233 (1977).
52. Sargent III, M., W. H. Swantner, and J. D. Thomas. *IEEE J. Quantum Electron.* **QE-16,** 465 (1980).
53. Kapon, E., A. Hardy, and A. Katzir. *IEEE J. Quantum Electron.* **QE-18,** 66 (1982); *J. Opt. Soc. Am.* **73,** 45 (1983).
54. Lyakhov, G. A., and N. V. Suyazov. *Sov. J. Quantum Electron.* **13,** 1034 (1983).
55. Hardy, A. *IEEE J. Quantum Electron.* **QE-20,** 1132 (1984).
56. Marcuse, D. *IEEE J. Quantum Electron.* **QE-21,** 161 (1985).
57. Kazarinov, R. F., and C. H. Henry. *IEEE J. Quantum Electron.* **QE-31,** 138 (1985).
58. Henry, C. H. *IEEE J. Quantum Electron.* **QE-21,** 1913 (1985).
59. Buus, J. *Electron. Lett.* **21,** 179 (1985).
60. Matsuoka, T., Y. Yoshikuni, and G. Motosugi. *Electron. Lett.* **21,** 1152 (1985).
61. Sekartedjo, K., N. Eda, K. Furuya, Y. Suematsu, F. Koyama, and T. Tanbun-Ek. *Electron. Lett.* **20,** 80 (1984).
62. Tada, K., Y. Nakano, and A. Ushirokawa. *Electron. Lett.* **20,** 82 (1984).
63. Soda, H., K. Wakao, H. Sudo, T. Tanahashi, and H. Imai. *Electron. Lett.* **20,** 1016 (1984).
64. Koentjoro, S., B. Broberg, K. Furuya, and Y. Suematsu. *Jpn. J. Appl. Phys.* Part 2 **23,** L791 (1984); Broberg, B., S. Koentjoro, K. Furuya, and Y. Suematsu. *Appl. Phys. Lett.* **47,** 4 (1985).
65. Utaka, K., S. Akiba, K. Sakai, and Y. Matsushima. *Electron. Lett.* **20,** 1008 (1984); *Electron. Lett.* **21,** 367 (1985).
66. McCall, S. L., and P. M. Platzman. *IEEE J. Quantum Electron.* **QE-21,** 1899 (1985).
67. Agrawal, G. P., and N. K. Dutta. *IEEE J. Quantum Electron.* **QE-21,** 534 (1985).
68. Agrawal, G. P., and N. K. Dutta. *Appl. Phys. Lett.* **46,** 213 (1985).
69. Patel, N. B., J. E. Ripper, and P. Brosson. *IEEE J. Quantum Electron.* **QE-9,** 338 (1973).

70. Kalandarishvili, K. G., S. Yu. Karpov, V. I. Kuchinskii, M. N. Mizerov, E. L. Portnoi, and V. B. Smirnitskii. *Sov. Phys.-Tech. Phys.* **28,** 959 (1983).
71. Dutta, N. K. *J. Appl. Phys.* **55,** 285 (1984); Dutta, N. K., and D. C. Craft. *J. Appl. Phys.* **56,** 65 (1984).
72. Chen, Y. C., and J. M. Liu. *Appl. Phys. Lett.* **45,** 731 (1984); Liu, J. M., and Y. C. Chen. *IEEE J. Quantum Electron.* **QE-21,** 271 (1985).
73. Bir, G. L., and G. E. Pikus. *Symmetry and Strain Induced Effects in Semiconductors.* New York: John Wiley & Sons, 1974.
74. Thompson, G. H. B. *Physics of Semiconductor Laser Devices.* Chichester: John Wiley & Sons, 1980.
75. Akiba, S., K. Utaka, K. Sakai, and Y. Matsushima. *Jpn. J. Appl. Phys.* Part 1 **23,** 1054 (1984).
76. Akiba, S., K. Utaka, K. Sakai, and Y. Matsushima. *Jpn. J. Appl. Phys.* Part 1 **21,** 1736 (1982).
77. Agrawal, G. P., N. K. Dutta, and P. J. Anthony. *Appl. Phys. Lett.* **48,** 457 (1986).
78. Glinski, J., and T. Makano. *IEEE J. Quantum Electron.* **QE-23,** 849 (1987).
79. Yoshita, J., Y. Itaya, Y. Naguchi, T. Matsuoka, and Y. Nakano. *Electron. Lett.* **22,** 327 (1986).
80. Koch, T. L., T. J. Bridges, E. G. Burkhardt, P. J. Corvini, L. A. Coldren, R. A. Logan, L. F. Johnson, R. F. Kazarinov, R. T. Yen, and D. P. Wilt. *J. Appl. Phys.* **47,** 12 (1985).
81. Ishikawa, H., H. Soda, K. Wakao, K. Kihara, K. Kamite, Y. Kotaki, M. Matsuda, H. Sudo, S. Yamakoshi, S. Isozumi, and H. Imai. *J. Lightwave Technol.* **LT-5,** 848 (1987).
82. Uomi, K., H. Nakano, and N. Chinone. *Electron. Lett.* **25,** 668 (1989).
83. Utaka, K., S. Akiba, K. Sakai, and Y. Matsushima. *IEEE J. Quantum Electron.* **QE-22,** 1042 (1986).
84. Soda, H., H. Ishikawa, and H. Imai. *Electron. Lett.* **22,** 1047 (1986); *IEEE J. Quantum Electron.* **QE-23,** 598 (1987).
85. Akiba, S., M. Usami, and K. Utaka. *J. Lightwave Technol.* **LT-5,** 1564 (1987).
86. Usami, M., S. Akiba, and K. Utaka. *IEEE J. Quantum Electron.* **QE-23,** 815 (1987).
87. Okuda, H., Y. Hirayama, H. Furuyama, J. Kinoshita, and M. Nakamura. *IEEE J. Quantum Electron.* **QE-23,** 843 (1987).
88. Usami, M., S. Akiba, K. Utaka, and Y. Matsushima. *IEE Proc. J,* **135,** 289 (1988).
89. Soda, H., Y. Kotaki, H. Sudo, H. Ishikawa, S. Yamakoshi, and H. Imai. *IEEE J. Quantum Electron.* **QE-23,** 804 (1987).
90. Usami, M., and S. Akiba. *IEEE J. Quantum Electron.* **QE-25,** 1245 (1989).
91. Whiteaway, J. E. A., G. H. B. Thompson, A. J. Collar, and C. J. Armistead. *IEEE J. Quantum Electron.* **QE-25,** 1261 (1989).
92. Okai, M., S. Tsuji, and N. Chinone. *IEEE J. Quantum Electron.* **QE-25,** 1314 (1989); S. Tsuji, M. Okai, N. Chinone, and M. M. Choy. *IEEE J. Quantum Electron.* **QE-25,** 1333 (1989).
93. Hillmer, H., S. Hansmann, and H. Burkhard. *Appl. Phys. Lett.* **57,** 534 (1990); *IEEE J. Quantum Electron.* **QE-27,** 1753 (1991).
94. Yamada, M., and K. Sakuda. *Appl. Opt.* **26,** 3474 (1987).
95. Agrawal, G. P., and A. H. Bobeck. *IEEE J. Quantum Electron.* **QE-24,** 2407 (1988).
96. Mols, P. P. G., P. I. Kundersma, and I. A. F. Baele. *IEEE J. Quantum Electron.* **QE-25,** 1303 (1989).
97. Akiba, S., M. Usami, and K. Utaka. *Trans. IECE* **E-69,** 389 (1986).
98. Agrawal, G. P., J. E. Geusic, and P. J. Anthony. *Appl. Phys. Lett.* **53,** 178 (1988).
99. Ogita, S., Y. Kotaki, H. Ishikawa, and H. Imai. *Electron. Lett.* **24,** 731 (1988); Ogita, S., Y. Kotaki, M. Matsuda, Y. Kuwahara, and H. Ishikawa. *Electron. Lett.* **25,** 629 (1989).
100. Kinoshita, J. I., and K. Matsumoto. *IEEE J. Quantum Electron.* **QE-25,** 1324 (1989).

101. Whiteaway, J. E. A., B. Garrett, G. H. B. Thompson, A. J. Collar, C. J. Armistead, and M. J. Fice. *IEEE J. Quantum Electron.* **QE-28,** 1277 (1992).
102. Morthier, G., K. David, P. Vankwikelberge, and R. Baets. *IEEE Photon. Technol. Lett.* **2,** 388 (1990).
103. Okai, M., N. Chinone, H. Taira, and T. Harada. *IEEE Photon. Technol. Lett.* **1,** 200 (1989).
104. Arakawa, Y., and A. Yariv. *IEEE J. Quantum Electron.* **QE-21,** 1666 (1985); **QE-22,** 1887 (1986).
105. Dutta, N. K., T. Wessel, N. A. Olsson, R. A. Logan, L. A. Koszi, and R. Yen. *Appl. Phys. Lett.* **46,** 525 (1985).
106. Kitamura, M., S. Takano, N. Henmi, T. Sasaki, H. Yamada, Y. Shinohara, H. Hasumi, and I. Mito. *Electron. Lett.* **24,** 1045 (1988).
107. Sasai, Y., J. Ohya, and M. Ogura. *IEEE J. Quantum Electron.* **QE-25,** 662 (1989).
108. Uomi, K., S. Sasaki, T. Tsuchiya, H. Nakano, and N. Chinone. *IEEE Photon. Technol. Lett.* **2,** 229 (1990).
109. Kitamura, M., H. Yamazaki, T. Sasaki, N. Kida, H. Hasumi, and I. Mito. *IEEE Photon. Technol. Lett.* **2,** 310 (1990).
110. Wang, S. J., L. J. P. Ketelson, V. R. McCrary, Y. Twu, S. G. Napholtz, and W. Werner. *IEEE Photon. Technol. Lett.* **2,** 775 (1990).
111. Okai, M., T. Tsuchiya, K. Uomi, N. Chinone, and T. Harada. *IEEE Photon. Technol. Lett.* **2,** 529 (1990); **3,** 427 (1991).
112. Miller, L. M., J. T. Verdeyen, J. J. Coleman, R. P. Bryan, J. J. Alwan, K. J. Beernink, J. S. Hughes, and T. M. Cockerill. *IEEE Photon. Technol. Lett.* **3,** 6 (1991).
113. Uomi, K., T. Tsuchiya, H. Nakano, M. Aoki, M. Suzuki, and N. Chinone. *IEEE J. Quantum Electron.* **QE-27,** 1705 (1991).
114. Okai, M., T. Tsuchiya, K. Uomi, N. Chinone, and T. Harada. *IEEE J. Quantum Electron.* **QE-27,** 1767 (1991).
115. Aoki, M., K. Uomi, T. Tsuchiya, S. Sasaki, M. Okai, and N. Chinone. *IEEE Photon. Technol. Lett.* **2,** 617 (1990); *IEEE J. Quantum Electron.* **QE-27,** 1782 (1991).
116. Kitamura, M., H. Yamazaki, T. Sasaki, N. Kida, and I. Mito. *IEEE Photon. Technol. Lett.* **3,** 305 (1991).
117. Uomi, K., M. Aoki, T. Tsuchiya, M. Suzuki, and N. Chinone. *IEEE Photon. Technol. Lett.* **3,** 493 (1991).
118. Kano, F., Y. Yoshikuni, M. Fukuda, and J. Yoshida. *IEEE Photon. Technol. Lett.* **3,** 877 (1991).
119. Okai, M., T. Tsuchiya; A. Takai, and N. Chinone. *IEEE Photon. Technol. Lett.* **4,** 526 (1992).
120. Nakano, Y., Y. Luo, and K. Takada. *Appl. Phys. Lett.* **55,** 1606 (1989).
121. Morthier, G., P. Vankwikelberge, K. David, and R. Baets. *IEEE Photon. Technol. Lett.* **2,** 170 (1990); David, K., G. Morthier, P. Vankwikelberge, and R. Baets. *Electron. Lett.* **26,** 238 (1990).
122. Luo, Y., Y. Nakano, K. Tada, T. Inoue, H. Hosomatsu, and H. Iwaoka. *Appl. Phys. Lett.* **56,** 1620 (1990); *IEEE J. Quantum Electron.* **QE-27,** 1725 (1991).
123. David, K., G. Morthier, P. Vankwikelberge, R. G. Baets, T. Wolf, and B. Borchert. *IEEE J. Quantum Electron.* **QE-27,** 1714 (1991).
124. Nakano, Y., Y. Deguchi, K. Ikeda, Y. Luo, and K. Tada. *IEEE J. Quantum Electron.* **QE-27,** 1732 (1991).
125. Luo, Y., R. Takahashi, Y. Nakano, K. Tada, T. Kamiya, H. Hosomatsu, and H. Iwaoka. *Appl. Phys. Lett.* **59,** 37 (1991).
126. Borchert, B., K. David, B. Stegmüller, R. Gessner, M. Beschorner, D. Sacher, and G. Franz. *IEEE Photon. Technol. Lett.* **3,** 955 (1991).

127. Inoue, T., S. Nakajima, Y. Luo, T. Oki, H. Iwaoka, Y. Nakano, and K. Tada. *IEEE Photon. Technol. Lett.* **3,** 958 (1991).
128. Luo, Y., H. L. Cao, M. Dobashi, H. Hosomatsu, Y. Nakano, and K. Tada. *IEEE Photon. Technol. Lett.* **3,** 1052 (1991).
129. Tsang, W. T., F. S. Choa, M. C. Wu, Y. K. Chen, R. A. Logan, A. M. Sergent, and C. A. Burrus. *IEEE Photon. Technol. Lett.* **4,** 212 (1992).
130. Olofsson, L., and T. G. Brown. *Appl. Phys. Lett.* **57,** 2773 (1990); *IEEE J. Quantum Electron.* **QE-28,** 1450 (1992).
131. Kudo, K., S. Arai, and K. Komori. Digest 13th Int. Semiconductor Laser Conf., Takamatsu, Japan, Sept., 1992.
132. Nakajima, S., T. Inoue, Y. Luo, T. Oki, Y. Nakano, K. Tada, R. Takahashi, and T. Kamiya. Digest 13th Int. Semiconductor Laser Conf., Takamatsu, Japan, Sept., 1992.
133. Suematsu, Y., K. Kishino, S. Arai, and F. Koyoma. Vol. 22, Part B, Chap. 4 in *Semiconductors and Semimetals*, ed. by W. T. Tsang. New York: Academic Press, 1985.
134. Posadas, L. S., K. Komori, Y. Tohmori, S. Arai, and Y. Suematsu. *IEEE J. Quantum Electron.* **QE-23,** 796 (1987).
135. Tohmori, Y., and M. Oishi. *Jpn. J. Appl. Phys.* **27,** L693 (1988).
136. Wang, M. T., K. Komori, K. S. Lee, S. Arai, and Y. Suematsu. *Jpn. J. Appl. Phys.* **27,** L1313 (1988).
137. Pellegrino, S., K. Komori, H. Suzuki, K. S. Lee, S. Arai, Y. Suematsu, and M. Aoki. *Electron. Lett.* **24,** 435 (1988).
138. Takahashi, M., Y. Michitsuji, M. Yoshimura, Y. Yamazoe, H. Nishizawa, and T. Sugimoto. *IEEE J. Quantum Electron.* **QE-25,** 1280 (1989).
139. Kano, F., M. Nakano, M. Kukuda, Y. Kondo, Y. Tohmori, and K. Oe. *Electron. Lett.* **25,** 709 (1989).
140. Komori, K., S. Arai, Y. Suematsu, I. Arima, and M. Aoki. *IEEE J. Quantum Electron.* **QE-25,** 1235 (1989).
141. Shim, J.-I., K. Komori, S. Arai, I. Arima, Y. Suematsu, and R. Somchai. *IEEE J. Quantum Electron.* **QE-27,** 1736 (1991).
142. Shim, J.-I., T. Nomaguchi, K. Kudo, and S. Arai. *Electron. Lett.* **28,** 741 (1992).
143. Chien, M., U. Koren, T. L. Koch, B. I. Miller, M. Oron, M. G. Young, and J. L. Demiguel. *IEEE Photon. Technol. Lett.* **3,** 418 (1991).
144. Hansen, P. B., G. Raybon, U. Koren, B. I. Miller, M. G. Young, M. Chien, C. A. Burrus, and R. C. Alferness. *IEEE Photon. Technol. Lett.* **3,** 212 (1991).
145. Matsui, Y., T. Kunii, H. Horikawa, and T. Kamijoh. *IEEE Photon. Technol. Lett.* **3,** 424 (1991).
146. F. S. Choa, W. T. Tsang, R. A. Logan, R. P. Gnall, U. Koren, T. L. Koch, C. A. Burrus, M. C. Wu, Y. K. Chen, P. F. Sciortino, A. M. Sergent, and P. J. Corvini. *Electron. Lett.* **28,** 1001 (1992).
147. O'Brien, S., R. Parke, D. F. Welch, D. Mehuys, and D. Scifres. *Electron. Lett.* **28,** 1272 (1992).
148. Jang, D.-H., Y. Kaneko, and K. Kishino. *Electron. Lett.* **28,** 428 (1992).
149. Olsson, N. A., C. H. Henry, R. F. Kazarinov, H. J. Lee, and K. J. Orlowsky. *IEEE J. Quantum Electron.* **QE-24,** 143 (1988).
150. Agrawal, G. P., and C. H. Henry. *IEEE J. Quantum Electron.* **QE-24,** 134 (1988).
151. Morton, P. A., R. Adar, R. C. Kistler, C. H. Henry, T. Tanbun-Ek, R. A. Logan, D. L. Coblentz, A. M. Sergent, and K. W. Wecht. *Appl. Phys. Lett.* **58,** 1239 (1991).
152. Bird, D. M., J. R. Armitage, R. Kashyap, R. M. A. Fatah, and K. H. Cameron. *Electron. Lett.* **27,** 1115 (1991).

153. Morton, P. A., V. Mizrahi, S. G. Kosinski, L. F. Mollenauer, T. Tanbun-Ek, R. A. Logan, D. L. Coblentz, A. M. Sergent, and K. W. Wecht. *Electron. Lett.* **28,** 561 (1992).
154. Agrawal, G. P. *Fibre-Optic Communication Systems.* New York: Wiley, 1992.
155. Wyatt, R., and W. J. Devlin. *Electron. Lett.* **19,** 110 (1983).
156. Tohmori, Y., Y. Suematsu, H. Tsushima, and S. Arai. *Electron. Lett.* **19,** 656 (1983).
157. Yamaguchi, M., M. Kitamura, S. Murata, I. Mito, and K. Kobayashi. *Electron. Lett.* **21,** 63 (1985).
158. Westbrook, L. D., A. W. Nelson, P. J. Fiddyment, and V. J. Collins. *Electron. Lett.* **21,** 957 (1985).
159. Yoshikuni, Y., K. Oe, G. Motosugi, and T. Matsuoka. *Electron. Lett.* **22,** 1153 (1986).
160. Kotaki, K., M. Matsuda, M. Yano, H. Ishikawa, and H. Imai. *Electron. Lett.* **23,** 325 (1987).
161. Murata, S., I. Mito, and K. Kobayashi. *Electron. Lett.* **23,** 403 (1983); *IEEE J. Quantum Electron.* **QE-23,** 835 (1987); *Electron. Lett.* **24,** 577 (1988).
162. Coldren, L. A., and S. W. Corzine. *IEEE J. Quantum Electron.* **QE-23,** 903 (1987).
163. Kotaki, Y., M. Matsuda, H. Ishikawa, and H. Imai. *Electron. Lett.* **24,** 503 (1988).
164. Kobayashi, K., and I. Mito. *J. Lightwave Technol.* **LT-6,** 1623 (1988).
165. Koch, T. L., U. Koren, and B. J. Miller. *Appl. Phys. Lett.* **53,** 1036 (1988).
166. Pan, X., H. Olesen, and B. Tromborg. *IEEE J. Quantum Electron.* **QE-24,** 2432 (1988); **25,** 1254 (1989).
167. Kotaki, Y., and H. Ishikawa. *IEEE J. Quantum Electron.* **QE-25,** 1340 (1989).
168. Koch, T. L., and U. Koren. *J. Lightwave Technol.* **LT-8,** 274 (1990).
169. Duan, G. H., P. Gallion, and G. Debarge. *IEEE J. Quantum Electron.* **QE-26,** 32 (1990).
170. Woodward, S. L., T. L. Koch, and U. Koren. *IEEE Photon. Technol. Lett.* **2,** 104 (1990).
171. Woodward, S. L., I. M. I. Habbab, T. L. Koch, and U. Koren. *IEEE Photon. Technol. Lett.* **2,** 854 (1990).
172. Oberg, M., S. Nilsson, T. Klinga, and P. Pjala. *IEEE Photon. Technol. Lett.* **3,** 299 (1991).
173. Kotaki, Y., and H. Ishikawa. *IEE Proc. J,* **138,** 171 (1991).
174. Tromborg, B., H. Olesen, and X. Pan. *IEEE J. Quantum Electron.* **QE-26,** 178 (1991).
175. Kuznetsov, M. *IEEE J. Quantum Electron.* **QE-26,** 668 (1991).
176. Suzuki, A. *IEEE J. Quantum Electron.* **QE-26,** 2109 (1991).
177. Tsigopulos, A., T. Sphicopoulos, I. Orfanos, and S. Pantelis. *IEEE J. Quantum Electron.* **QE-27,** 415 (1992).
178. Kasper, B. L., R. A. Linke, K. L. Walker, L. G. Cohen, T. L. Koch, T. J. Bridges, E. G. Burkhardt, R. A. Logan, R. W. Dawson, and J. C. Campbell. *Proc. Eur. Conf. Opt. Commun.,* Stuttgart, Aug. 1984.
179. Fujita, S., M. Kitamura, T. Torikai, N. Henmi, H. Yamada, T. Suzaki, I. Takano, and M. Shikada. *Electron. Lett.* **25,** 702 (1989).
180. Gimlett, J. L., M. Z. Iqbal, J. Young, L. Curtis, R. Spicer, and N. K. Cheung. *Electron. Lett.* **25,** 596 (1989).
181. Gnauck, A. H., R. M. Jopson, J. D. Evankow, C. A. Burrus, S.-J. Wang, N. K. Dutta, and H. M. Presby. *Electron. Lett.* **25,** 1695 (1989).
182. Gnauck, A. H., C. R. Giles, L. J. Cimini, Jr., J. Sone, L. W. Stulz, S. K. Korotky, and J. J. Veselka. *IEEE Photon. Technol. Lett.* **3,** 1147 (1991).
183. Gnauck, A. H., and C. R. Giles. *IEEE Photon. Technol. Lett.* **4,** 80 (1992).
184. Gnauck, A. H., K. C. Reichmann, J. M. Kahn, S. K. Korotky, J. J. Veselka, and T. L. Koch. *IEEE Photon. Technol. Lett.* **2,** 908 (1990).
185. Kahn, J. M., A. H. Gnauck, J. J. Veselka, S. K. Korotky, and B. L. Kasper. *IEEE Photon. Technol. Lett.* **2,** 285 (1990).
186. Norimatsu, S., K. Iwashita, and K. Noguchi. *Electron. Lett.* **26,** 648 (1990).

187. Yamazaki, S., T. Ono, H. Shimizu, M. Kitamura, and K. Emura. *IEEE Photon. Technol. Lett.* **2,** 914 (1990).
188. Toba, H., K. Oda, K. Nakanishi, N. Shibata, K. Nosu, N. Takato, and K. Sato. *Electron. Lett.* **26,** 376 (1990).
189. Lin, C., H. Kobrinki, A. Frenkel, and C. A. Brackett. *Electron. Lett.* **24,** 1215 (1988).
190. Toba, H., K. Oda, K. Nakanishi, N. Shibata, K. Nosu, N. Takato, and M. Fukuda. *J. Lightwave Technol.* **LT-8,** 1396 (1990).
191. Hill, A. M., R. Wyatt, J. F. Massicott, K. J. Blyth, D. S. Forrester, R. A. Lobbett, P. J. Smith, and D. B. Payne. *Electron. Lett.* **26,** 1882 (1990).
192. Forrester, D. S., A. M. Hill, R. A. Lobbett, R. Wyatt, and S. F. Carter. *Electron. Lett.* **27,** 2051 (1991).
193. Darcie, T., and G. Bodeep. *IEEE Trans. Microwave Theory Tech.* **38,** 524 (1990).
194. Rosher, P. A., and S. C. Fenning. *Electron. Lett.* **26,** 534 (1990).
195. Hill, P. M., and R. Olshansky. *Electron. Lett.* **24,** 892 (1988).
196. Way, W. I., S. S. Wagner, M. M. Choy, C. Lin, R. C. Menendez, H. Tohme, A. Yi-Yan, A. C. von Lehman, R. E. Spicer, M. Andrejco, M. Saifi, and H. L. Lemberg. *IEEE Photon. Technol. Lett.* **2,** 665 (1990).
197. Edagawa, N., Y. Toshida, H. Taga, S. Yamamoto, K. Mochizuchi, and H. Wakabayashi. *Electron. Lett.* **26,** 66 (1990).
198. Saito, S., T. Imai, and T. Ito. *J. Lightwave Technol.* **LT-9,** 161 (1991).
199. Bergano, N. S., J. Aspell, C. R. Davidson, P. R. Trischitta, B. M. Nyman, and F. W. Kerfoot. *Electron. Lett.* **27,** 1889 (1991).
200. Mollenauer, L. F., E. Lichtman, G. T. Harvey, M. J. Neubelt, and B. M. Nyman. *Electron. Lett.* **28,** 792 (1992).
201. Nakazawa, M., K. Suzuki, and E. Yamada. *Electron. Lett.* **28,** 1046 (1992).
202. Andrekson, P. A., N. A. Olsson, M. Haner, J. R. Simpson, T. Tanbun-Ek, R. A. Logan, D. Coblentz, H. M. Presby, and K. W. Wecht. *IEEE Photon. Technol. Lett.* **4,** 76 (1992).
203. Gabla, P. M., J. O. Frourd, E. Leclerc, S. Gauchard, and V. Havard. *IEEE Photon. Technol. Lett.* **4,** 717 (1992).
204. Spirit, D. M., G. E. Wickens, T. Widdowson, G. R. Walker, D. L. Williams, and L. C. Blank. *Electron. Lett.* **28,** 1218 (1992).

Chapter 8
COUPLED-CAVITY SEMICONDUCTOR LASERS

8.1 INTRODUCTION

In the previous chapter it was seen that the distributed-feedback mechanism can provide single-frequency semiconductor lasers with a high degree of side-mode suppression. The operating wavelength is relatively unaffected by external perturbations since it is determined by the spatial period of a permanently etched grating. Although wavelength stability is an attractive feature of distributed-feedback lasers, it is achieved at the expense of tunability. For some applications it is desirable to have a semiconductor laser whose wavelength can be tuned over a wide range, at least discretely. Coupled-cavity semiconductor lasers have the potential of offering mode selectivity together with wavelength tunability.

The basic concept of using the coupled-cavity scheme for longitudinal-mode selection is known from the early work[1,2] on gas lasers and has been widely used for mode selection.[3] In the case of semiconductor lasers, considerable work was carried out by coupling the semiconductor laser to an external cavity.[4] *Monolithically integrated coupled-cavity devices* were considered occasionally[5–8] in relation to their specific properties such as optical bistability, optical amplification, and longitudinal-mode selectivity. Such devices attracted wide attention[9–28] during the 1980s in an attempt to obtain single-frequency semiconductor lasers that are useful for optical communications in the 1.55-μm wavelength region. At the same time, the interest in semiconductor lasers coupled to a small-length external cavity has also revived.[29–40] To distinguish them from monolithic coupled-cavity devices, such lasers are often referred to as *external-cavity* semiconductor lasers. Considerable work[41–60] has been carried out to understand the theory behind their operation.

The mechanism of mode selectivity in coupled-cavity lasers can be understood by referring to Fig. 8.1. The effect of feedback from the external cavity can be modeled through an effective wavelength-dependent reflectivity of the facet facing the external cavity. As a result, the cavity loss is different

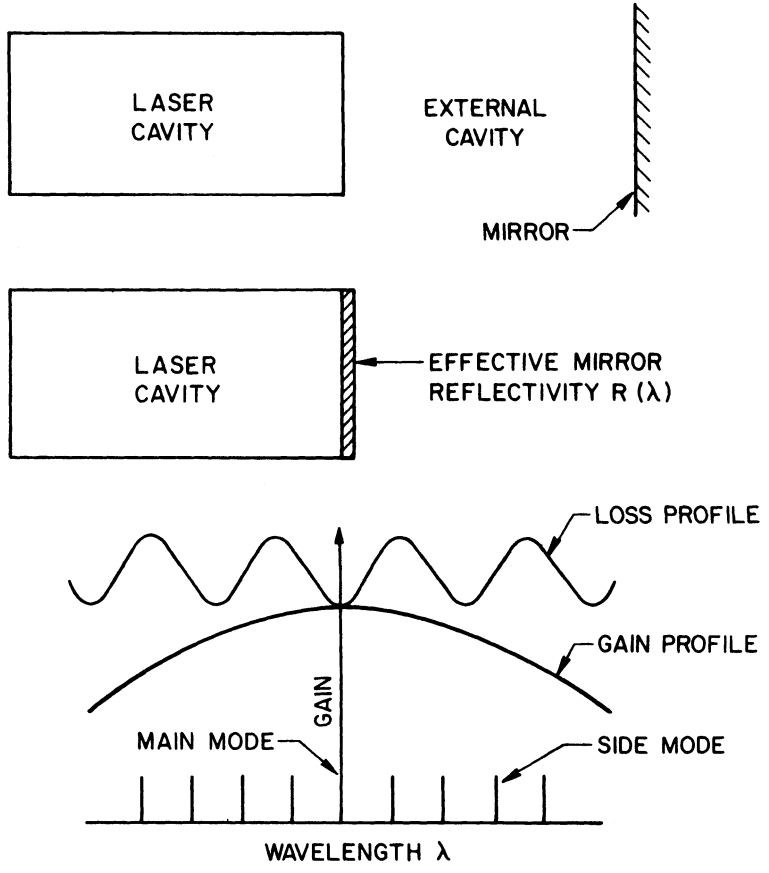


Fig. 8.1 Schematic illustration of longitudinal-mode selectivity in a coupled-cavity laser. Top drawing shows the two cavities. The effect of an external cavity is to make the effective reflectivity of the facet wavelength dependent (middle drawing). Bottom drawing shows the resulting periodic loss profile superimposed on the gain profile. The main mode is the FP mode closest to the gain peak with the lowest loss. Other FP modes become side modes suppressed because of their relatively higher losses.

for different Fabry-Perot (FP) modes of the laser cavity. In general the loss profile is periodic as shown schematically in the bottom part of Fig. 8.1. The mode selected by the coupled-cavity device is the FP mode that has the lowest cavity loss and is closest to the peak of the gain profile. Because of the periodic nature of the loss profile, other FP modes with relatively low cavity losses may exist. Such modes are discriminated by the gain roll-off because of their large separation from each other. The power carried by these

side modes is typically 20–30 dB below that of the main mode occurring in the vicinity of the gain peak.

In this chapter we consider coupled-cavity semiconductor lasers and their performance. In Sec. 8.2 we briefly discuss various coupled-cavity schemes; these can be broadly classified as active-passive and active-active, depending on whether the second cavity can be externally pumped or not. Section 8.3 deals with the theoretical aspects of coupled-cavity devices. Using a simple model for the intercavity coupling, one can determine the longitudinal modes of the coupled system and their respective threshold gains. Particular attention is paid to the qualitative differences occurring when the cavity lengths are comparable and when they differ significantly from each other. The modulation performance is discussed using single-mode rate equations generalized to account for the intercavity coupling. In particular, we discuss the effect of coupling on the frequency chirp. Experimental performance of coupled-cavity lasers is considered in Sec. 8.4, where we discuss longitudinal-mode selection, wavelength tuning, and modulation characteristics. To emphasize the application of coupled-cavity lasers in optical communications, we briefly present the results of a few transmission experiments. Finally, Sec. 8.5 considers diverse applications of the active-active type of coupled-cavity devices.

8.2 COUPLED-CAVITY SCHEMES

As mentioned in Sec. 8.1, coupled-cavity semiconductor lasers can be classified into two broad categories. We shall refer to them as *active-passive* and *active-active* schemes, depending on whether the second cavity (sometimes referred to as the "controller") remains unpumped or can be pumped to provide gain. The devices in the latter category are also called *three-terminal* devices since three electrical contacts are used to pump the two optically coupled but electrically isolated cavity sections. Figures 8.2 and 8.3 show a specific example of each kind of device.

In the active-passive scheme, the semiconductor laser is coupled to an external cavity that is unpumped and plays a passive role. In the simplest design, a plane[61] or a spherical[62] mirror is placed at a short distance from one of the laser facets; the facet may have an antireflection coating to increase the coupling between the two cavity sections. The spherical-mirror geometry for InGaAsP lasers has attracted considerable attention[29,36] and has been used in a transmission experiment at 1.52 μm.[63] Several variations of the plane-mirror geometry have been adopted for InGaAsP lasers. In the short coupled-cavity scheme,[33] a short laser cavity (50–80 μm) is coupled to a short external cavity (30–80 μm). The gold-coated facet of a semiconductor chip acts as a plane mirror. In another scheme[34,38] the external cavity consists of a graded-index fiber lens to provide improved coupling and to avoid

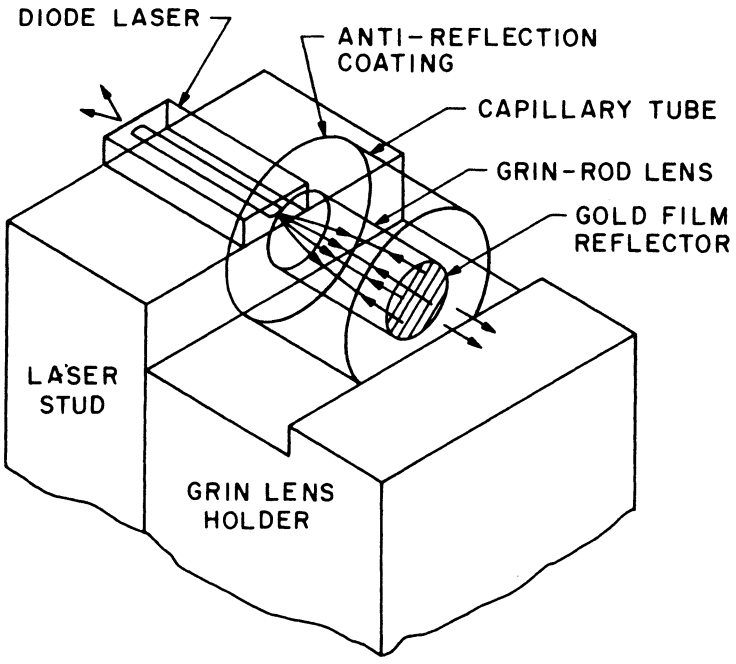


Fig. 8.2 Schematic illustration of an active-passive coupled-cavity laser. In the scheme shown, the light is focused on the highly reflecting back facet of a graded-index fiber rod (GRINROD) that acts as an external cavity. (After Ref. 38)

diffraction losses. A short rod ($\sim$100–200 μm) of graded-index multimode fiber with one end gold-coated is used for this purpose (see Fig. 8.2). Etching techniques have also been used[32,40] to make monolithic coupled-cavity devices with integrated and passive sections.

In the active-active scheme shown in Fig. 8.3 both sections can be independently pumped, giving an additional degree of freedom that can be used to control the behavior of the device. A natural choice is to use identical active materials for both cavities. Furthermore, alignment between the active regions is automatically achieved if the two cavity sections are created by forming a gap in a conventional semiconductor laser. Cleaving[6,7,12] and etching[9] techniques have been used for this purpose. In another scheme[17] the interferometric property of a coupled-cavity device is obtained by bending the active region: although the cavities are not physically separated, the p contact can be separated into two parts by chemical etching to accomplish separate pumping. Whatever the technique employed, the qualitative behavior of such three-terminal coupled-cavity lasers is similar with respect to mode selectivity and wavelength tunability.

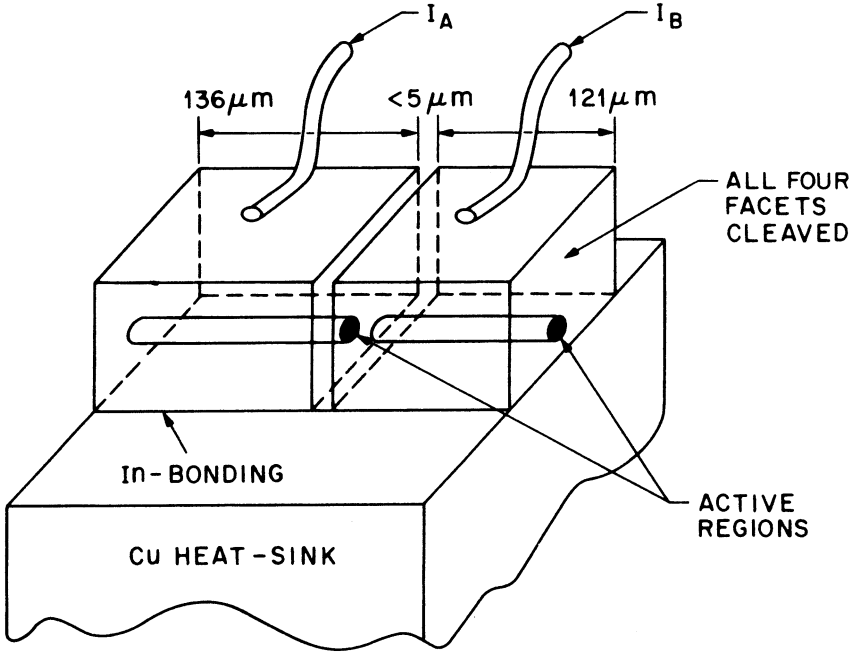


Fig. 8.3 Schematic illustration of an active-active coupled-cavity device. A cleaved-coupled-cavity (C^3) laser is shown. In this three-terminal device, changing the currents I_A and I_B adjusts the optical length of each cavity through refractive-index alterations. (After Ref. 13)

The advantage of the cleaving technique lies in providing a device whose four facets are mirrorlike and parallel to each other. The cleaved-facet reflectivity (about 32%) allows reasonable coupling between the two cavity sections as long as the gap is not too wide. Coupled-cavity lasers made by the cleaving technique are sometimes referred to as *cleaved-coupled-cavity*, or C^3 lasers (see Fig. 8.3). Their operating characteristics have been extensively studied.[28] We shall refer to C^3 lasers while considering three-terminal devices after keeping in mind that similar behavior would occur for other kinds of active-active devices.

The mechanism of mode selectivity is the same for both active-active and active-passive devices and can be understood from Fig. 8.1. The main difference between the two kinds of coupled-cavity lasers lies in the external means used to shift the FP modes of the two cavities. In the active-passive case, the modes of the external cavity can be shifted by changing its length or its temperature. By contrast, the active-active case offers the possibility of electronic shifting since the current in the two cavities can be independently

controlled. If one of the cavities is operated below threshold, a change in its drive current significantly changes the carrier density inside the active region. Since the refractive index of a semiconductor laser changes along with the carrier density, the longitudinal modes shift with a change in the drive current, and different FP modes of the laser cavity can be selected.

In the design of a coupled-cavity laser, the cavity lengths L_1 and L_2 are adjustable to some extent. The performance of such lasers depends on the relative optical lengths $\mu_1 L_1$ and $\mu_2 L_2$ of the two cavities, where μ_1 and μ_2 are the mode indices. Two cases should be distinguished depending on whether the optical lengths are similar ($\mu_1 L_1 \cong \mu_2 L_2$) or differ significantly ($\mu_1 L_1 \gg \mu_2 L_2$); they are referred to as the long-long and long-short geometries, respectively. Although both geometries are capable of mode selection and wavelength tuning, they differ in the one important aspect of wavelength stability. For the long-long geometry, even a relatively small shift of the FP mode can lead to "mode hopping." By contrast, a shift of about one mode spacing is required to achieve mode hopping in the long-short geometry. Clearly, unintentional mode hopping because of temperature and current fluctuations is less likely to occur in the long-short case. By the same token, wavelength tuning would require larger current or temperature changes for a long-short than for a long-long device.

8.3 THEORY

A theoretical analysis of the coupled-cavity semiconductor lasers requires simultaneous consideration of the gain and loss in the two cavities after taking into account their mutual optical feedback. In contrast to a single-cavity laser for which the facet loss is wavelength-independent (and hence the same for all FP modes), in a coupled-cavity laser the effective facet loss becomes different for different FP modes. One objective of the theory is to find the longitudinal modes of the coupled system and their respective gains[41–48] required to reach threshold. The threshold gains can then be used in the rate equations to study the transient behavior[44,46] and the modulation response.[54,59] For three-terminal devices the gain and index variations in both cavities should be considered with respect to individual current changes in the two coupled cavities.

A general analysis of coupled-cavity semiconductor lasers is extremely complicated, so it is necessary to make several simplifying assumptions. We assume that the field distribution associated with the fundamental lateral and transverse mode of the waveguide is unaffected by the intercavity coupling and that only axial propagation in each cavity needs to be considered. This reduces the problem to one dimension. Note, however, that the coupling depends upon the mode width and other related parameters,

and mode-conversion losses occur[46–48] as the optical mode leaves cavity 1, diffracts in the air inside the gap, and then reenters cavity 2. Such losses would be incorporated through an effective gap loss α_g.

In this section we introduce the concept of the coupling constant and then obtain the longitudinal modes and their respective threshold gains for the coupled system. The gain margin between the lowest-threshold mode and the mode with the second-lowest threshold gain determines the extent of side-mode suppression. Particular attention is paid to the modulation characteristics, and under certain conditions the intercavity coupling is found to reduce the frequency chirp.

8.3.1 Coupling Constant

The first step in the analysis is to determine the extent of coupling between the two cavities.[41–49] We consider the general case applicable for both active-active and active-passive devices. Figure 8.4 shows the geometry and notation. The coupling between the cavities is governed by an air gap of width L_g. The air gap itself forms a third FP cavity, and the intercavity coupling is affected by the loss and phase shift experienced by the optical field while traversing the gap. In the scattering-matrix approach[10,64] the fields in the two cavities are related by

$$\begin{pmatrix} E_1' \\ E_2' \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \end{pmatrix}. \tag{8.3.1}$$

The scattering matrix elements can be obtained by considering multiple

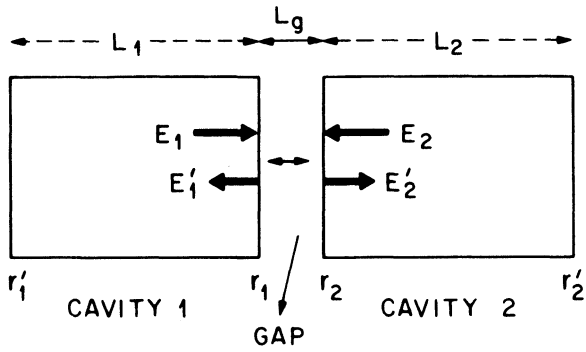


Fig. 8.4 Geometry and notation used in the theoretical description ocf a coupled-cavity semiconductor laser. Notation is explained in the text. In the general case, reflectivities r at the four facets may differ from each other.

reflections inside the gap. More explicitly, the gap is treated as an FP cavity on which the field E_1 is incident from the left (Fig. 8.4) and the reflected field E_1' and the transmitted field E_2' are obtained in terms of E_1 after considering multiple round trips inside the cavity. The procedure yields $S_{11} = E_1'/E_1$ and $S_{21} = E_2'/E_1$. The same method is used to obtain S_{22} and S_{12} after assuming that only the field E_2 is incident on the FP cavity from the right. The result is[10]

$$S_{11} = r_1 - \frac{r_2(1 - r_1^2)t_g}{1 - r_1 r_2 t_g} \tag{8.3.2}$$

$$S_{22} = r_2 - \frac{r_1(1 - r_2^2)t_g}{1 - r_1 r_2 t_g} \tag{8.3.3}$$

$$S_{12} = S_{21} = \frac{[t_g(1 - r_1^2)(1 - r_2^2)]^{1/2}}{1 - r_1 r_2 t_g} \tag{8.3.4}$$

where

$$t_g = \exp(2i\beta_g L_g) = \exp(2ik_0 L_g)\exp(-\alpha_g L_g). \tag{8.3.5}$$

Here $\beta_g = k_0 + i\alpha_g/2$ and accounts for the phase shift and the loss inside the gap. Further, r_1 and r_2 are the amplitude-reflection coefficients at the two facets forming the gap; i.e., $r_n = (\mu_n - 1)/(\mu_n + 1)$, where μ_1 and μ_2 are the mode indices in the two cavities. Physically the gap has been replaced by an interface whose effective reflection coefficients are S_{11} and S_{22}, while S_{12} and S_{21} are the effective-transmission coefficients from cavity 1 to cavity 2 and vice versa. It is useful to define a complex coupling parameter

$$\tilde{C} = C\exp(i\theta) = \left(\frac{S_{12}S_{21}}{S_{11}S_{22}}\right)^{1/2} \tag{8.3.6}$$

where C governs the strength of mutual coupling and θ is the coupling phase, which will be seen later to play an important role.

The magnitudes of C and θ depend on a large number of device parameters. The simplest case occurs for a semiconductor laser coupled to an external mirror. In this case the coupling element is just the laser-air interface. Since $L_g = 0$, $t_g = 1$. Furthermore $\mu_2 = 1$, and therefore $r_2 = 0$. Using these values in Eqs. (8.3.2)–(8.3.6), we find that $C = (1 - r_1^2)^{1/2}/r_1$ and $\theta = \pi/2$. For a C^3 laser, the two cavities have nearly equal indices of refraction, and therefore $r_1 = r_2 = r$. Using Eq. (8.3.2)–(8.3.6), we now obtain

$$C\exp(i\theta) = \left(\frac{1 - r^2}{r}\right)\left(\frac{t_g^{1/2}}{1 - t_g}\right). \tag{8.3.7}$$

The coupling depends on t_g, and both C and θ vary with the gap width L_g. The evaluation of C and θ requires a knowledge of the gap loss α_g. These losses arise mainly from diffraction spreading of the beam[10,48] inside the gap and have been estimated[10] using a simple diffraction-spreading model perpendicular to the junction plane. Figure 8.5 shows[47] the variation of C and θ with the gap width using this model after taking $|r|^2 = 0.31$ as the cleaved-facet reflectivity. Both C and θ vary considerably with small

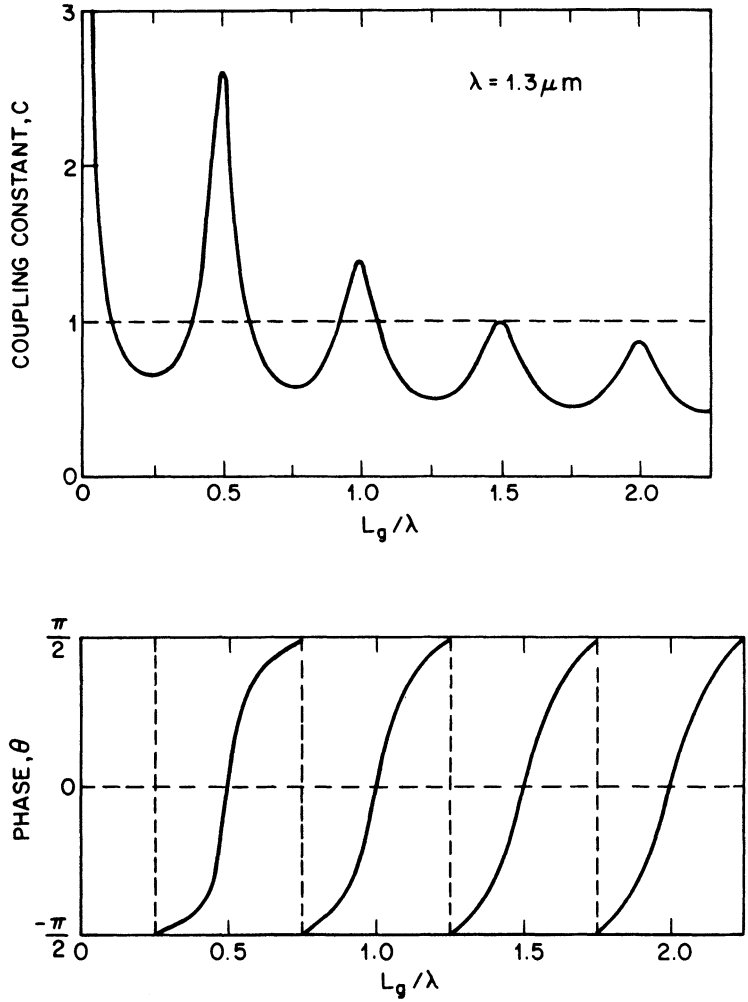


Fig. 8.5 Dependence of the coupling constant C and the coupling phase θ on the gap width L_g for a laser oscillating at the wavelength λ. (After Ref. 47)

changes in L_g. The in-phase coupling ($\theta = 0$) occurs whenever $L_g = m\lambda/2$ (m is an integer), and C also goes through a maximum for that value of L_g. In practice, however, the gap width may vary from device to device. The corresponding large variations occurring in C and θ imply that the performance of C^3 lasers would also be device-dependent.

8.3.2 Longitudinal Modes and Threshold Gain

In this section we obtain an eigenvalue equation whose solutions yield the wavelength and threshold gain associated with the longitudinal modes of the coupled system. A simple way to do this[47] is to consider the relationship between E_1 and E_1' using Fig. 8.4. The field E_1' results from reflection of E_1 and transmission of E_2 and is given by [see Eq. (8.3.1)]

$$E_1' = S_{11}E_1 + S_{12}E_2. \tag{8.3.8}$$

On the other hand, the round trip through cavity 1 results in the relation

$$E_1 = r_1' \exp(2i\beta_1 L_1)E_1' = r_1' t_1 E_1' \tag{8.3.9}$$

where the complex propagation constant

$$\beta_n = \mu_n k_0 - i\bar{\alpha}_n/2 \qquad (n = 1, 2) \tag{8.3.10}$$

governs wave propagation in nth cavity, $k_0 = 2\pi/\lambda$, λ is the device wavelength, and α_n is the mode gain. The definition of β_n is similar to Eq. (7.3.5) of Chapter 7 with the difference that both cavities should be considered and the mode-propagation constant is in general different in each cavity. As in Sec. 7.3.1, the mode gain is related to the material gain g_n by

$$\bar{\alpha}_n = \Gamma g_n - \alpha_n^{\text{int}} \tag{8.3.11}$$

where Γ is the confinement factor and α_n^{int} is the internal loss. For an active-active device, g_1 and g_2 can be individually varied by changing the current passing through each cavity section. For an active-passive device, $g_2 = 0$ and $\bar{\alpha}_2$ accounts for the absorption loss in the passive cavity.

Equations (8.3.8) and (8.3.9) can be combined to obtain the relation

$$(1 - r_1' t_1 S_{11})E_1 = r_1' t_1 S_{12} E_2. \tag{8.3.12}$$

Similar considerations for cavity 2 lead to

$$(1 - r_2' t_2 S_{22})E_2 = r_2' t_2 S_{21} E_1. \tag{8.3.13}$$

These two homogeneous equations have nontrivial solutions only if the secular condition

$$(1 - r_1' t_1 S_{11})(1 - r_2' t_2 S_{22}) = r_1' r_2' t_1 t_2 S_{12} S_{21} \tag{8.3.14}$$

is satisfied. Equation (8.3.14) is the desired eigenvalue equation for the coupled system and has been extensively studied.[45–48] In the absence of coupling, $S_{12} = 0$, $S_{nn} = r_n$, and we recover the threshold condition for uncoupled cavities. Note that

$$t_n = \exp(2i\beta_n L_n) = \exp(2i\mu_n k_0 L_n) \exp(-\bar{\alpha}_n L_n) \tag{8.3.15}$$

incorporates the phase shift and the gain (or loss) experienced by the optical field during a round trip in each cavity. The loss and the phase shift inside the gap are included through t_g and S_{ij}, as given by Eqs. (8.3.2)–(8.3.5). The eigenvalue equation (8.3.14) is applicable for all kinds of coupled-cavity devices with arbitrary reflectivities at four interfaces (see Fig. 8.4).

Before proceeding, it is useful to introduce the concept of an effective mirror reflectivity.[45,48] In many cases the role of one cavity (say cavity 2) is to provide a control through which a single FP mode of cavity 1 is selected. The effect of cavity 2 on mode selectivity can be treated by an effective reflectivity for the laser facet facing cavity 2. Equation (8.3.14) can be written in the equivalent form of

$$(1 - r_1' R_{\text{eff}} t_1) = 0 \tag{8.3.16}$$

where the effective reflectivity

$$R_{\text{eff}} = S_{11} + \frac{r_2' t_2 S_{12} S_{21}}{1 - r_2' t_2 S_{22}} = S_{11}\left(1 + \frac{\tilde{C}^2 f_2}{1 - f_2}\right) \tag{8.3.17}$$

and $f_2 = r_2' t_2 S_{22}$ and is the fraction of the amplitude coupled back into the laser cavity after a round trip in cavity 2. Equation (8.3.16) suggests that as far as mode selectivity is concerned, the coupled-cavity laser is equivalent to a single-cavity laser with facet reflection coefficients r_1' and R_{eff}. Although Eq. (8.3.16) is formally exact, its practical utility is limited to the case of weak coupling so that a change in t_1 (through operating conditions of cavity 1) does not affect R_{eff} through a change in the feedback fraction f_2. This is often the case for active-passive devices. In the case of active-active devices, the effective reflectivity concept is valid when cavity 2 is biased below threshold.

To illustrate the extent of mode selectivity offered by the coupled-cavity

mechanism, we consider solutions of the eigenvalue equation (8.3.14) for a specific C^3-type active-active device for which $r_1 = r_1' = r_2 = r_2' \cong 0.56$. We assume that cavity 2 is biased below threshold such that $\bar{\alpha}_2 = 0$ (in the absence of coupling) and that it does not change significantly with coupling. Equation (8.3.14) is used to obtain $\bar{\alpha}_1$, and the wavelength $\lambda = 2\pi/k_0$ corresponding to various longitudinal modes. Figure 8.6 shows the longitudinal modes and their respective gains for a specific gap width ($L_g = 1.55$ μm) and for two sets of cavity lengths L_1 and L_2 corresponding to long-long and long-short geometries. Since the concept of effective reflectivity is approximately valid, $|R_{\text{eff}}|$ versus λ is also shown. The effect of the second cavity

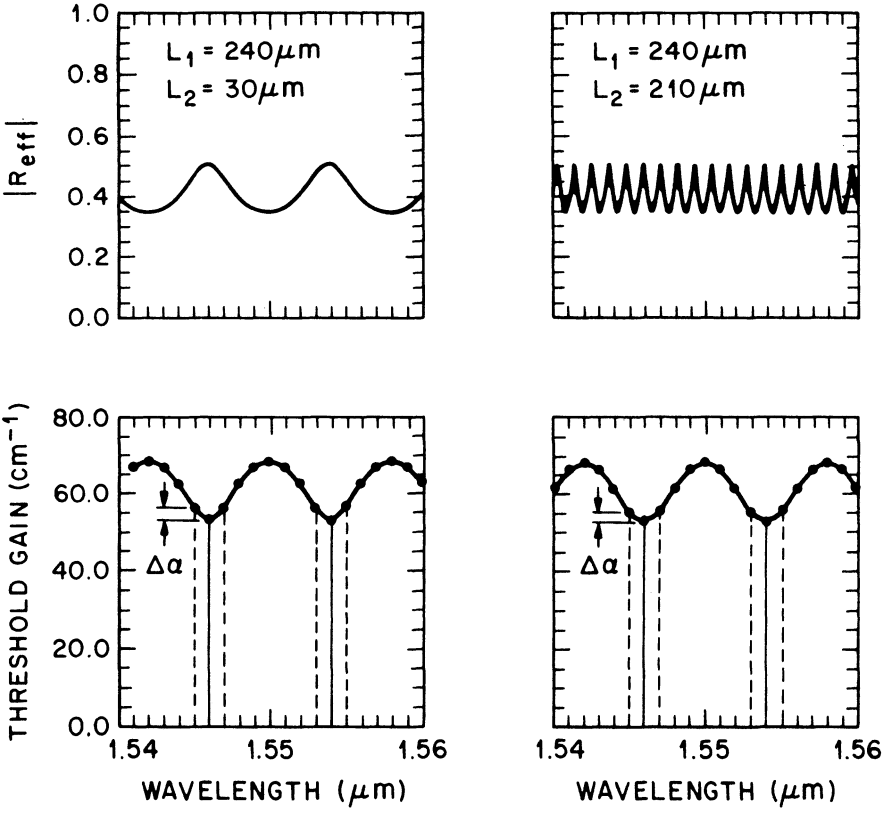


Fig. 8.6 Mode discrimination in coupled-cavity lasers employing long-short (left column) and long-long (right column) geometries. In each case the upper figure shows the wavelength variation of the effective laser-facet reflectivity, while the lower figure shows the longitudinal modes (marked by solid dots) and their respective threshold gains. The gain margin $\Delta\alpha$ between the dominant mode (solid vertical lines) and the next mode (dashed vertical lines) is a measure of mode selectivity and determines the mode suppression ratio. (After Ref. 46)

is to modulate the effective reflectivity, and the minimum threshold gain is required for modes for which R_{eff} is maximum. In the absence of coupling, the same threshold gain ($\bar{\alpha}_1 = 48\ \text{cm}^{-1}$) is required for all modes. With coupling, the gain difference $\Delta\alpha$ between the lowest-gain mode and the neighboring mode provides mode discrimination and leads to side-mode suppression. Figure 8.6 is drawn for the optimum case ($L_g = \lambda$) using $S_{11} = S_{22} = 0.409$ and $S_{12} = S_{21} = 0.371$; from Eq. (8.3.6) these two relations give $C = 0.907$ and $\theta = 0$).

As mentioned before, C and θ vary considerably with the gap width L_g. The effect of gap width on the threshold gains has been extensively studied.[43–48] For a lossy gap the best mode selectivity occurs for the case of in-phase coupling ($\theta = 0$), which requires[46,47] that L_g be an integer multiple of $\lambda/2$. The worst case ($\theta = \pi/2$) occurs wherever L_g is an odd multiple of $\lambda/4$ (see Fig. 8.5). For a lossless gap the situation is reversed; i.e., the best and worst selectivity occur for $\theta = \pi/2$ and $\theta = 0$, respectively.[44,51] This is the case for external-cavity devices where a single semiconductor-air interface acts as the coupling element. As discussed in Sec. 8.3.1, $\theta = \pi/2$ for such devices. Reasonable mode discrimination can thus be expected and has been observed experimentally.[29]

A comparison of long-long and long-short devices in Fig. 8.6 reveals an interesting feature. Even though the wavelength variation of $|R_{\text{eff}}|$ is much faster in the long-long case, the two devices behave similarly as far as the threshold gains are concerned. In Fig. 8.6 in particular, the mode pattern repeats after eight modes ($M = 8$). For the long-short geometry ($L_1 \gg L_2$), the repeat mode is determined by the ratio of the optical lengths in each cavity ($M \cong \mu_1 L_1/\mu_2 L_2$). For the long-long geometry, by contrast, $M \cong \mu_1 L_1/(\mu_1 L_1 - \mu_2 L_2)$. For a given coupled-cavity device the repeat mode is thus determined by $\mu_2 L_2$ or $\mu_1 L_1 - \mu_2 L_2$ depending on which quantity is smaller. However, the long-long and long-short devices, differ with respect to tunability and stability of the mode wavelength.

We briefly consider the general case in which both cavities are close to or above threshold and the gain and index variations in one cavity influence the behavior of the other. The eigenvalue equation (8.3.14) can be used to obtain the threshold gains $\bar{\alpha}_1$ and $\bar{\alpha}_2$ for a given mode wavelength λ. Because of the intercavity coupling, the mode gains $\bar{\alpha}_1$ and $\bar{\alpha}_2$ are not pinned in the above-threshold regime, but are related for each mode in the form of a curve in the $\bar{\alpha}_1$-$\bar{\alpha}_2$ plane. The numerical results show[46,47] that as the currents in the two sections are varied, the laser wavelength can be tuned discretely by wavelength hopping from one mode to another. The current range over which the laser maintains the same longitudinal mode, however, depends on the gap width as well as on the cavity lengths. This range is considerably larger for long-short devices ($L_1 \gg L_2$) compared to those for which

$L_1 \cong L_2$. The same behavior holds when mode hopping occurs because of temperature changes.[25] The long-short geometry is generally preferable if the objective is to obtain a laser with a stable wavelength.[23]

A detailed theoretical understanding for three-terminal C^3 devices requires that the mode thresholds for various longitudinal modes be known as a function of the currents I_1 and I_2 passing through the two cavity sections. This behavior has been studied[47] using a simple model for the gain and index variations with the current. Single-mode operation above threshold is described by zones in the I_1-I_2 plane, and mode hopping occurs at the zone boundaries. However, these zones are extremely sensitive to the coupling phase θ; optimum performance for C^3 lasers is expected to occur for in-phase coupling ($\theta = 0$). For other values of θ, the steady-state operation can become unstable or display bistable behavior.

8.3.3 Side-Mode Suppression

The mode suppression ratio (MSR), defined as the ratio of the power in the main mode to that of the most intense side mode, is often used to express the extent of mode selectivity for coupled-cavity devices. It can be related to the threshold gains in a similar way as was done for DFB lasers in Sec. 7.4.1. Using Eqs. (7.4.1)–(7.4.4), the MSR can be written in the form

$$\text{MSR} = 1 + \frac{\Delta\alpha + \Gamma\Delta g}{(\bar{\alpha}_1 + \alpha_1^{\text{int}})\delta} \tag{8.3.18}$$

where $\Delta\alpha$ is the threshold-gain margin (see Fig. 8.6), Γ is the confinement factor, Δg is the gain roll-off from its peak value, and $\bar{\alpha}_1 + \alpha_1^{\text{int}}$ is the total mode gain for the lowest-loss mode. The parameter δ is a small dimensionless number and decreases with an increase in the main-mode power [see Eq. (6.3.11)]. Typically, $\delta \cong 5 \times 10^{-5}$ at a power level of few milliwatts.

Mode discrimination in coupled-cavity lasers can be understood using Eq. (8.3.18) and Fig. 8.6. Lasing occurs for the lowest-loss mode closest to the gain peak. The next mode sees practically no gain roll-off ($\Delta g \cong 0$) since the longitudinal-mode spacing is very small relative to the gain bandwidth. However, the cavity-loss discrimination ($\Delta\alpha \cong 4$–$5\ \text{cm}^{-1}$) can provide MSR values of ~ 30 dB. As we move away from the lasing mode, $\Delta\alpha$ decreases and Δg increases. Nonetheless, their sum can be large enough to provide sufficient mode discrimination. For the next pair of coincident modes, $\Delta\alpha = 0$ and mode discrimination is completely determined by the gain roll-off. However, one can expect $\Gamma\Delta g$ to be $\sim 5\ \text{cm}^{-1}$ if the repeat mode is more than 4 or 5 mode spacings away from the gain peak.

This discussion applies for CW operation of a coupled-cavity device. For

single-cavity lasers the MSR degrades severely under transient conditions, as discussed in Sec. 6.4.1. The multimode rate equations used there can be applied for coupled-cavity lasers after incorporating different photon lifetimes for different modes. Similar to the case of DFB lasers, side modes remain suppressed by ~ 30 dB even under transient conditions if $\Delta\alpha \cong 8\text{–}10\ \text{cm}^{-1}$. Because of their capability for dynamic side-mode suppression, coupled-cavity lasers are useful as single-frequency sources for optical communication systems employing dispersive fibers at the 1.55-μm wavelength.

8.3.4 Modulation Response

For their application in optical fiber communications, semiconductor lasers are directly modulated to provide pulsed output. Important features such as modulation bandwidth and frequency chirp were discussed in Chapter 6 for single-cavity lasers. In this section we consider the effect of intercavity coupling on the modulation response of coupled cavity lasers. Three-terminal devices such as the C^3 laser offer an additional degree of control since the bias and modulation currents in each cavity can be individually adjusted. It has been observed experimentally[27,58] that the chirp for C^3 lasers can be reduced, typically by a factor of 2, through proper adjustment of the device currents. Chirp reduction is also expected to occur for external-cavity devices,[54,65] even though the passive cavity remains unbiased.

Although one generally should consider multimode rate equations,[44,46] modulation analysis is considerably simplified by following the dynamics of the dominant single-longitudinal mode. This is justified since practically useful devices must maintain the same longitudinal mode with sufficient side-mode suppression (greater than ~ 20 dB) throughout modulation. Such weak side modes do not significantly affect the modulation response of the main mode. However, one should generalize the single-cavity, single-mode rate equation of Chapter 6 to incorporate the feedback owing to intercavity coupling. This generalization can be carried out using Eq. (8.3.8), which shows that on every reflection a relative fraction $S_{12}E_2/S_{11}E_1$ of the field in cavity 2 enters cavity 1. Since this fraction is complex, both the power and the phase in cavity 1 are affected. Adding the additional term representing the field entering from the neighboring cavity to Eq. (6.2.9) and following the analysis of Sec. 6.2 with the above-mentioned modification, we obtain the generalized rate equations[54,59]

$$\dot{P}_j = (G_j - \gamma_j)P_j + R_{sp}(N_j) + \kappa_j \cos(\theta \pm \phi) \tag{8.3.19}$$

$$\dot{N}_j = I_j/q - \gamma_e(N_j)N_j - G_jP_j \tag{8.3.20}$$

$$\dot{\phi}_j = -\frac{\mu_j}{\mu_{gj}}(\omega - \Omega_j) + \tfrac{1}{2}\beta_c(G_j - \gamma_j) + \frac{\kappa_j}{2P_j}\sin(\theta \pm \phi) \tag{8.3.21}$$

where $\phi = \phi_1 - \phi_2$ and is the relative phase between the two cavities. The feedback rate

$$\kappa_j = \frac{c}{\mu_{gj} L_j} (P_1 P_2)^{1/2} C \tag{8.3.22}$$

where the coupling constant C and the coupling phase θ are defined by Eq. (8.3.6). The other symbols have the same meanings as in Chapter 6 except for the subscript j that has been added to distinguish the two cavities.

Consider the case of a C^3 laser under direct sinusoidal modulation at the frequency ν_m. In the general case, a fraction m_s of the modulation current is applied to cavity 2. The device currents in that case are

$$I_1(t) = I_1^b + (1 - m_s) I_m \sin(2\pi \nu_m t) \tag{8.3.23}$$

$$I_2(t) = I_2^b + m_s I_m \sin(2\pi \nu_m t + \theta_m) \tag{8.3.24}$$

where I_1^b and I_2^b are the bias levels, I_m is the peak value of the modulation current, and θ_m is the phase shift between the modulation current applied to the two cavities.

Equations (8.3.19)–(8.3.24) can be used to model the modulation response of a C^3 laser of arbitrary cavity lengths. The large-signal regime requires a numerical solution of the rate equations. However, a qualitative understanding can be developed by considering the small-signal solution of the generalized rate equations. In this approach, P_j, N_j, and ϕ_j are expanded around their steady-state values corresponding to the bias levels, and Eqs. (8.3.19)–(8.3.21) are linearized in terms of the small deviations δP_j, δN_j, and $\delta \phi_j$, which vary sinusoidally in response to direct modulation. Two quantities of practical interest are the normalized power response

$$H(\nu_m) = \left| \frac{\delta P_1(\nu_m)}{\delta P_1(0)} \right| \tag{8.3.25}$$

and the frequency chirp per milliampere of the modulation current

$$\delta \nu = \left| \frac{\delta \dot{\phi}_1}{2\pi I_m} \right| \tag{8.3.26}$$

for laser cavity 1. The dependence of $H(\nu_m)$ and $\delta \nu$ on various device parameters has been considered for both active-passive[54,65] and active-active[55,59] schemes.

To demonstrate the effect of intercavity coupling, we consider a specific

C^3 laser such that $L_1 = 200\ \mu m$ and $L_2 = 50\ \mu m$, and we assume in-phase coupling ($\theta = 0$). Figure 8.7 shows the power response and the chirp when both sections are biased at 1.5 times above threshold and the modulation current is split equally with no phase shift ($m_s = 0.5$, $\theta_m = 0$). The power response is similar to the single-cavity case and peaks at the relaxation-oscillation frequency ν_R. However, the frequency ν_R increases with an increase in the coupling constant C. This can be understood by noting that the effect of coupling is to decrease the laser threshold and that at the fixed bias current the output power and hence ν_R increase.

The most notable feature of the coupled-cavity laser is the chirp reduction occurring at high modulation frequencies. Figure 8.7 shows that when the coupling is strong, the chirp initially decreases before peaking at the relaxation-oscillation frequency ν_R. Thus, depending on the modulation frequency, the chirp may be higher or lower relative to its single-cavity value (dashed curve). A reduction by a factor of 2 is predicted when $\nu_m = 1$ GHz and $C = 1$. Figure 8.5 shows that such values of C can be achieved for narrow-gap C^3 devices. The chirp reduction has also been observed experimentally.[27,58] The origin of chirp reduction can be understood by noting that the second and third terms in Eq. (8.3.21) contribute to chirp, whereas the third term is absent for single-cavity lasers. Depending on operating conditions, the two contributions may interfere to increase or decrease the total chirp.

The chirp increase for C^3 lasers at low modulation frequencies is another attractive feature that can be used for analog frequency modulation with negligible spurious intensity modulation. Frequency-excursion rates in the range of 1–10 GHz/mA have been experimentally observed[14] and are in agreement with the theoretical results shown in Fig. 8.7. The numerical value for a given device depends, among other things, on the coupling constant C and the coupling phase θ.

Figure 8.7 is drawn for a specific set of device parameters. Both $H(\nu_m)$ and $\delta\nu$ change with the bias currents, with the modulation splitting (fraction m_s, phase shift θ_m), and with the coupling parameters C and θ. Refs. 54 and 59 give a detailed modulation-response analysis of active-passive and active-active devices. Qualitatively speaking, the in-phase coupling is desirable to minimize chirp for C^3 devices. Since the in-phase condition is also optimum from the viewpoint of mode stability and dynamic side-mode suppression, the C^3 device for which the gap is nearly an integer multiple of $\lambda/2$ is expected to provide the best performance.

8.4 OPERATING CHARACTERISTICS

In this section we present the experimental results for coupled-cavity lasers with particular attention paid to active-active C^3 devices. The properties

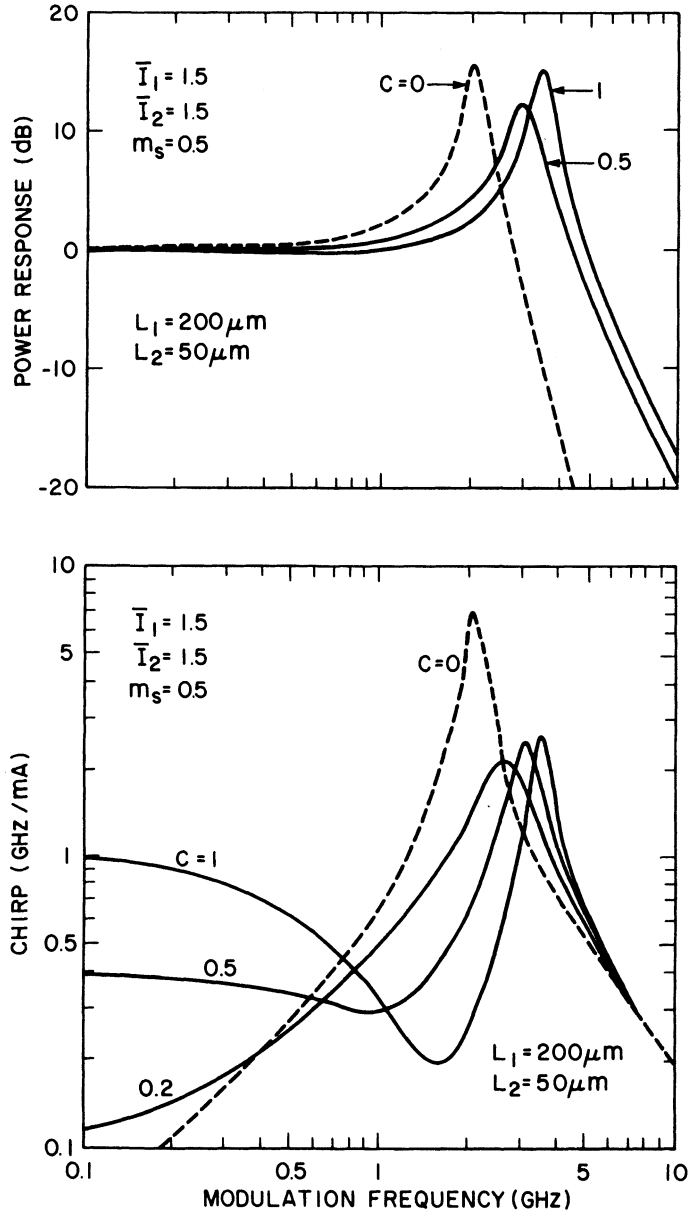


Fig. 8.7 Effect of intercavity coupling (C = coupling constant) on the modulation characteristics of a three-terminal coupled-cavity laser when both sections are biased equally above threshold and the modulation current is split equally between them. The optimum value ($\theta = 0$) of the coupling phase is assumed. (After Ref. 59 © 1985 IEEE)

considered are longitudinal-mode selectivity under CW operation, wavelength tuning with current adjustment, domains of single-mode operation under modulation, and the frequency chirp. A few transmission experiments employing C^3 lasers are also briefly described. Finally, we discuss the performance of external-cavity semiconductor lasers and their potential applications in optical fiber communications.

8.4.1 Longitudinal-Mode Control

Figure 8.8 shows an example of the L-I curves obtained at room temperature under CW operation of a 1.5-μm CSBH C^3 laser.[12] The inset shows the biasing scheme. I_L and I_R are the bias currents for the left and right sections. When $I_R < 20$ mA (the threshold current of the right section), the L-I curves for the left section are similar to those obtained for single-cavity lasers. An increase in I_R decreases the threshold since the right diode becomes less lossy. However, when $I_R > 20$ mA, the L-I curves show qualitatively different behavior, with a ripple-like structure below threshold. Some ripples or kinks also occur in the above-threshold regime but are more widely separated when compared to those occurring below threshold. Spectral measurements show that these kinks in the L-I curve are associated with the currents at which the laser switches from one longitudinal mode to another one.

The observed behavior can be understood physically as follows. When $I_R = 30$ mA in Fig. 8.8, the right cavity is acting as the laser while the left cavity takes the role of controller. When biased below threshold, the controller acts as a frequency selector. A change in its current changes the carrier density. This in turn changes the refractive index, which results in a shift of the FP modes of the controller with respect to the laser cavity FP modes. A shift of the controller FP modes manifests as a shift of the loss profile in Fig. 8.1. When the shift is large enough so that the lowest loss occurs for the next laser-cavity FP mode, mode hopping occurs. This longitudinal-mode hopping is manifested by the observed ripples in the L-I curve. The discrete mode hops are sometimes referred to as wavelength tuning. During each hop the wavelength shifts by one FP-mode spacing ($\sim$2 nm for a 150-μm-long cavity), and a tuning range of more than 10 nm is readily achieved. The exact number of mode hops is determined by the relative cavity lengths, as discussed in Sec. 8.3.2.

Figure 8.9 shows[12] the longitudinal mode spectra of a CSBH C^3 laser that showed a tuning range of 26 nm through 13 discrete mode hops of about 2 nm each. This is one of the largest tuning ranges achieved with the use of the coupled-cavity scheme. Also shown in the figure is the current change ΔI_L needed for successive mode hops plotted as a function of laser wavelength. Small values of ΔI_L are necessary when the controller is biased

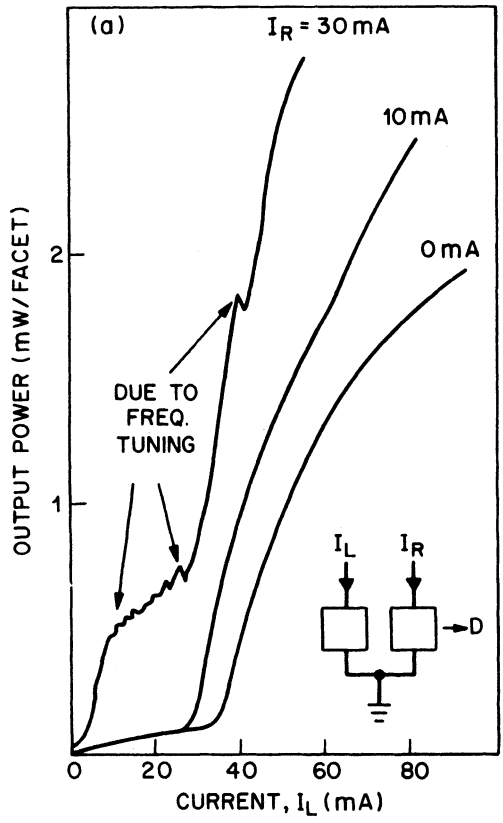


Fig. 8.8 Light-current characteristics of a C^3 laser for different biasing of the right section. Inset shows the biasing scheme and the position of the detector (D). As the right-cavity current I_R increases, the threshold decreases because of reduced absorption in the right cavity. When I_R exceeds the threshold value, the right cavity acts as a laser while the left-cavity acts as controller. The bumps in the I_R = 30 mA curve represent hopping of the laser wavelength to successive longitudinal modes. (After Ref. 12)

considerably below threshold. However, as I_L increases, the value of ΔI_L needed for the next mode hope increases rapidly. This is so because the carrier density begins to saturate as the controller approaches its threshold. Nonetheless, for the C^3 laser shown in Fig. 8.9, the wavelength shifted by about 26 nm when the modulator current was changed by 11.6 mA. This resulted in an average tuning rate of more than 2 nm/mA. The current step ΔI_L required for a mode hop also depends on the relative cavity lengths L_1 and L_2. The long-long geometry ($L_1 \cong L_2$) was used in obtaining Figs. 8.8 and 8.9. For

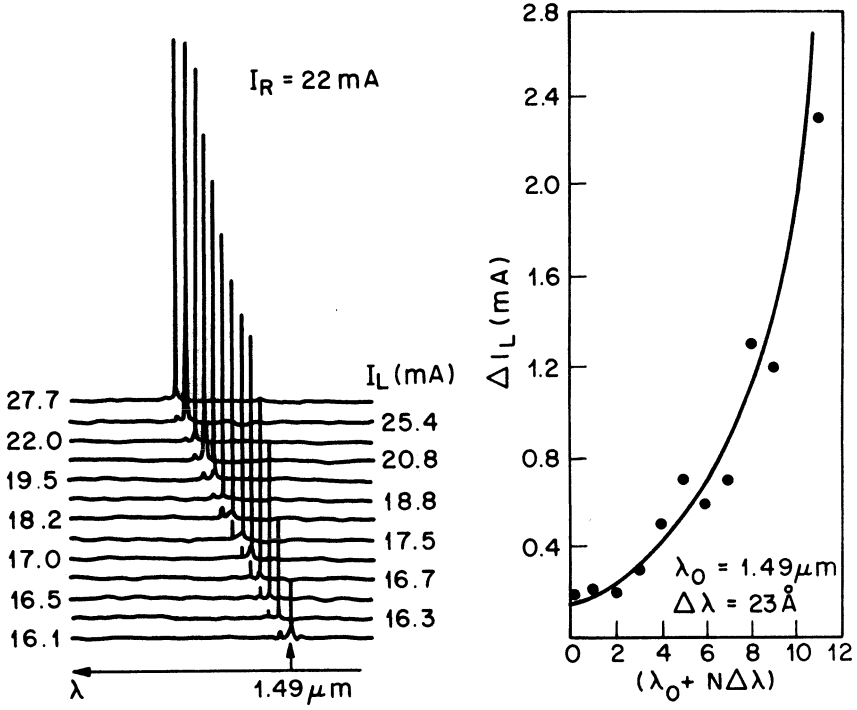


Fig. 8.9 Wavelength tuning of a C^3 laser using the biasing scheme shown in the inset of Fig. 8.8. Successive longitudinal modes of the right cavity (acting as the laser) can be selected by changing the current I_L in the left controller section. The current change ΔI_L as a function of the mode number N is also shown with solid dots. The solid curve shows the trend of ΔI_L increasing with N. (After Ref. 12)

a long-short device, much larger values of ΔI_L than shown in Fig. 8.9 would be required since the FP mode of the short cavity has to shift by one mode spacing of the long cavity for a mode hop to occur.

When the controller is operating above threshold, the carrier density is approximately pinned and its FP modes stop shifting with an increase in I_L. Under ideal conditions no mode hopping is expected. In practice, however, occasional mode hopping occurs because of a partial clamping of the carrier density or because of temperature-induced index variations. Nonetheless, as Fig. 8.8 shows, when both sections of a C^3 laser are biased above threshold, the device is able to maintain the same longitudinal mode over a considerable range of current. This mode of operation is useful from the standpoint of wavelength stability.

8.4.2 Optimum Biasing for Direct Modulation

From the viewpoint of applications to optical fiber communications, it is essential that the laser maintains the same longitudinal mode under high-speed direct modulation while at the same time the side modes remain reasonably suppressed (by more than 20 dB) throughout the modulation cycle.

The L-I curves shown in Fig. 8.8 indicate that a C^3 laser behaves differently depending on whether the controller section is biased below or above threshold. As the currents in the two cavity sections are varied, the C^3 laser has domains of single-longitudinal-mode operation. Figure 8.10 shows these domains relative to the currents applied to each section for a 1.55-μm ridge waveguide C^3 laser.[16] The dotted line corresponds to the threshold. Regions A and B are the tuning regions corresponding to the situation where one section is biased above its lasing threshold while the other section (controller) is biased below threshold all the time. The horizontal lines in region A and

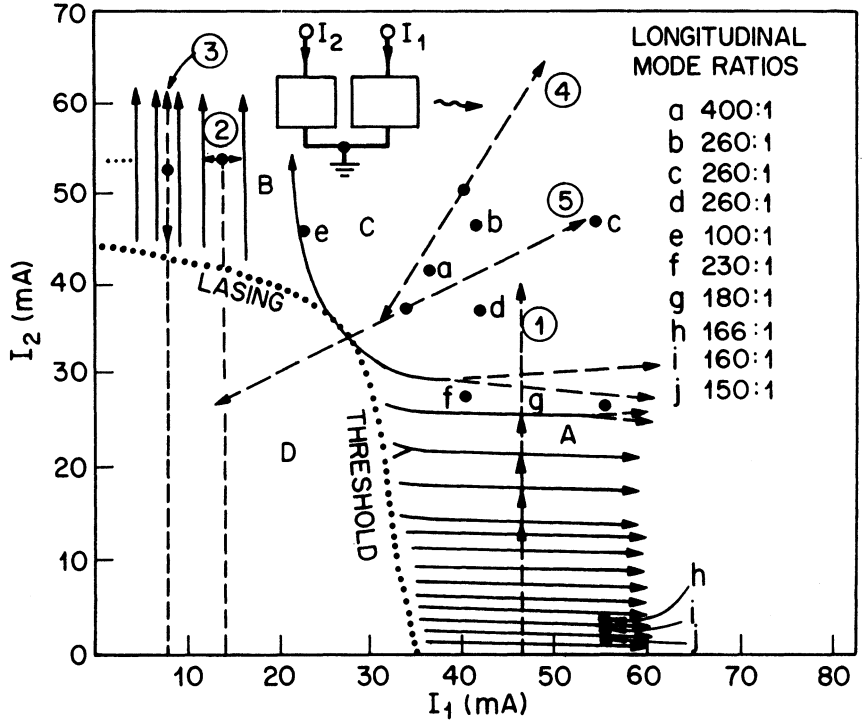


Fig. 8.10 Domains of single-longitudinal-mode operation of a C^3 laser as determined by currents applied to each section. Solid lines with arrows mark the boundaries along which mode hopping occurs. Dashed lines show the operation lines for direct modulation. Mode-suppression ratios at various bias levels are also shown. (After Ref. 16)

the vertical lines in region B are the boundaries along which the laser switches abruptly from one longitudinal mode to the adjacent one. Within each narrow stripe the laser maintains the same longitudinal mode. In the square-shaped region C, both cavities are biased above threshold and the same longitudinal mode is maintained throughout this region. Both sections are below threshold in the region D.

Dashed lines in Fig. 8.10 show the biasing schemes employed[16] for various modes of operation of a C^3 laser. The line 1 corresponds to wavelength tuning. The laser cavity is biased above threshold ($I_1 > I_1^{th}$) and I_2 is varied such that the controller is below threshold. The wavelength-tuning regime has already been discussed and illustrated by Figs. 8.8 and 8.9. Line 2 correponds to analog direct-frequency modulation. Here one section, say cavity 2, is biased above threshold while cavity 1 acts as the controller ($I_1 < I_1^{th}$). A small modulation current is applied to the controller such that the laser operates within the narrow stripe region without mode hops. Carrier-density variations in the controller section leads to the periodic modulation of the laser frequency around its bias value. Modulation amplitudes of ~ 10 GHz/mA are easily accomplished. This should be contrasted with the single-cavity case where frequency modulation is typically less than 1 GHz/mA and, furthermore, is always accompanied by spurious intensity modulation. For a C^3 laser the intensity remains almost constant since the laser bias remains unchanged.

For the intensity modulation commonly employed in optical communication systems, there are two biasing schemes. One can operate in regions A or B of Fig. 8.10 after applying the optimum bias current to the controller (say cavity 1) so that the C^3 laser is operating within one of the narrow stripes at a point where the MSR is highest. The current in cavity 2 is then modulated as shown by the operation line 3. In the other mode of operation, both cavities are biased above threshold so that the laser is operating within the square region C, where it maintains the same longitudinal mode throughout modulation. In order to avoid entering the mode-hopping regions A and B, the modulation current may be split between the two cavities such that the resulting operation line is tilted in the I_1-I_2 plane. This is illustrated in Fig. 8.10 through dashed lines 4 and 5, whose lengths denote the peak-to-peak amplitudes of the modulation current while the solid dot in the center of each line corresponds to the bias level.

The choice between these two biasing schemes for direct intensity modulation depends on several factors. To start with, the side modes should remain reasonably suppressed throughout a modulation cycle. For the C^3 laser of Fig. 8.10, the MSR was measured[16] at different operating points giving the values shown. Although it was slightly higher in region C, MSR values of more than 20 dB were measured in all three regions. Some C^3 lasers have

shown MSR values exceeding 30 dB. The other criterion is related to the mode stability; i.e., the laser should remain at the same longitudinal mode under external perturbations such as temperature fluctuations. In this respect the biasing of a C^3 laser in the region C is preferable simply because the single-mode domain is wider.

For some applications such as wavelength multiplexing, it is desirable to have a semiconductor laser whose wavelength is selected by tuning and then remains stable under modulation. This would require the operation of a C^3 laser in the tuning regime. The wavelength stabilization can then be achieved[28] through a feedback loop that automatically adjusts the bias currents to maximize the MSR. A powerful technique for stabilization and optimum biasing of C^3 lasers consists of plotting the constant power curves in the I_1-I_2 plane.[21] These curves can be used to determine the bias and modulation currents for the two cavities. A feedback loop can provide stable operation over wide ranges of temperature and power.

8.4.3 Frequency Chirp

The mode-discrimination ability of coupled-cavity lasers inherent in their design results in single-frequency lasers that oscillate predominantly in a single longitudinal mode with side modes suppressed by 20–35 dB. Furthermore, in contrast to single-cavity lasers, these side modes remain suppressed even when the coupled-cavity laser is directly modulated at frequencies in the gigahertz range.[12–23] The dynamic spectral width of such a laser is then determined by frequency chirping. Similar to the case of DFB lasers, the frequency chirp is expected to be the limiting factor in the performance of optical communication systems employing coupled-cavity lasers.

Considerable attention has been paid to the measurement[26,27] and minimization[55,58] of frequency chirp in coupled-cavity lasers. The measurements for C^3 lasers operating at 1.3-μm and 1.55-μm wavelengths have shown[27] that the coupled-cavity nature of these lasers not only provides mode selectivity but is also helpful in reducing the chirp. Figure 8.11 shows experimental results[27] where dashed lines show the corresponding behavior for a single-cavity laser. To ensure that the chirp reduction is mainly due to the coupled-cavity mechanism, two lasers were selected from the same wafer and one of them was converted to a C^3 laser by cleaving it approximately in the middle. Both sections were biased above threshold to give about 2 mW of power and were modulated sinusoidally at 100 MHz. The chirp varies almost linearly with the modulation-current (peak-to-peak) amplitude for both single-cavity and coupled-cavity lasers. However, at a given modulation current the chirp for a coupled-cavity laser is typically reduced by a factor of 2 compared with its value for the single-cavity laser.

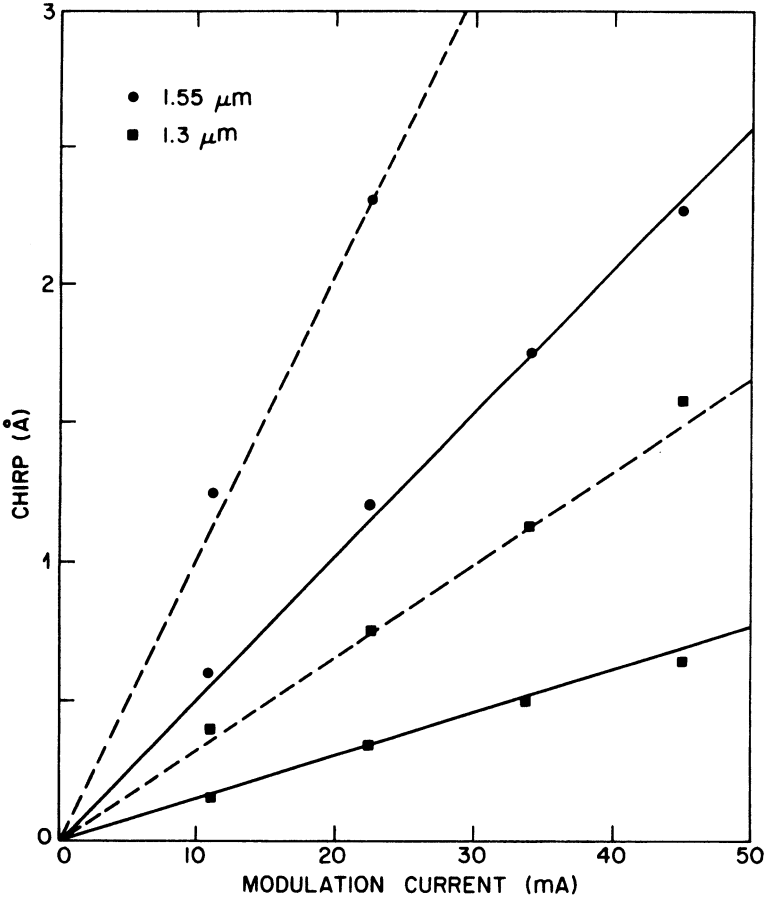


Fig. 8.11 Variation of the chirp with the modulation current (solid lines) for two C^3 lasers oscillating at 1.3-μm and 1.55-μm wavelengths. For comparison, dashed lines show the chirp behavior of a conventional single-cavity semiconductor laser obtained from the same wafer. (After Ref. 27)

The exact amount of chirp reduction for a specific coupled-cavity laser depends on a large number of parameters such as the modulation frequency, the bias levels, and the fractions into which the modulation current is split for the two cavity sections. Some of the parameters can be optimized to obtain the minimum chirp.[55,58] For a long-long C^3 device, both cavity sections are often biased above threshold for the purpose of direct modulation.[16] In this case the bias levels I_1^b, I_2^b, and the modulation-splitting fraction can be optimized for chirp minimization. The small-signal analysis[55,59] for this case reveals that the chirp is minimum when both sections have equal

relative bias levels ($I_1^b/I_1^{th} \cong I_1^b/I_2^{th}$) and when the modulation current is split equally between them. The numerical results (see Fig. 8.7) also show that chirp reduction depends on the strength of intercavity coupling, on the coupling phase, and on the modulation frequency. For a long-short device, chirp measurements have been carried out[58] when the short section (section 2) is biased below threshold. For a given bias I_2^b less than I_2^{th}, the modulation-splitting fraction m_s and the relative phase θ_m [see Eq. (8.3.24)] were adjusted to minimize the chirp. For small values of I_2^b, small values of m_s of about 0.1 were required. However, when I_2^b was increased considerably, the optimum chirp occurred when the modulation current was split equally without phase shift between the two sections.

8.4.4 Transmission Experiments

Similar to the case of DFB lasers, the ability of coupled-cavity semiconductor lasers to operate stably in a single longitudinal mode can be exploited for use in optical communication systems at 1.55 μm where conventional silica fibers have lowest losses but suffer from a relatively large chromatic dispersion of 15–18 ps/(nm·km). Both the active-passive[38,63] (external cavity) and active-active[66–69] schemes have been used for this purpose. In this section we briefly describe the results of transmission experiments employing C^3 lasers.

In one transmission experiment[66] a ridge waveguide C^3 laser at 1.55 μm was used to transmit signals over 161.5 km of single-mode fiber at the bit rate of 420 Mb/s. The experiment set up was similar to that shown in Fig. 7.16. The temperature was held at 20.1°C. The controller section of the C^3 laser was biased below threshold at 12.1 mA. The laser-cavity current was switched between 43.9 and 94.9 mA to generate a pseudorandom bit stream in the NRZ format. The average power launched into the fiber was 0.9 dBm.

The fiber output was coupled to an InGaAs avalanche photodetector with an optimized average gain of 9.1. The receiver sensitivy for a bit error rate of 10^{-9} at 420 Mb/s was −42.6 dBm. An error rate of 5×10^{-10} was obtained after passing the signal through a 162-km-long fiber with an average received power of −41.7 dBm. Figure 8.12 shows the measured bit error rate as a function of the received power for fibers 10 km and 162 km long. A direct comparison shows a dispersion penalty of only about 0.7 dB even though the fiber had a dispersion of 17 ps/(nm·km) at the 1.55-μm wavelength used in the experiment. This penalty may be attributed to the wavelength chirp.

Another transmission experiment[67] used a similar 1.55-μm ridge waveguide C^3 laser to increase the bit rate to 1 Gb/s. A bit error rate of 1×10^{-10} was obtained after transmission through a 101-km-long fiber with an average

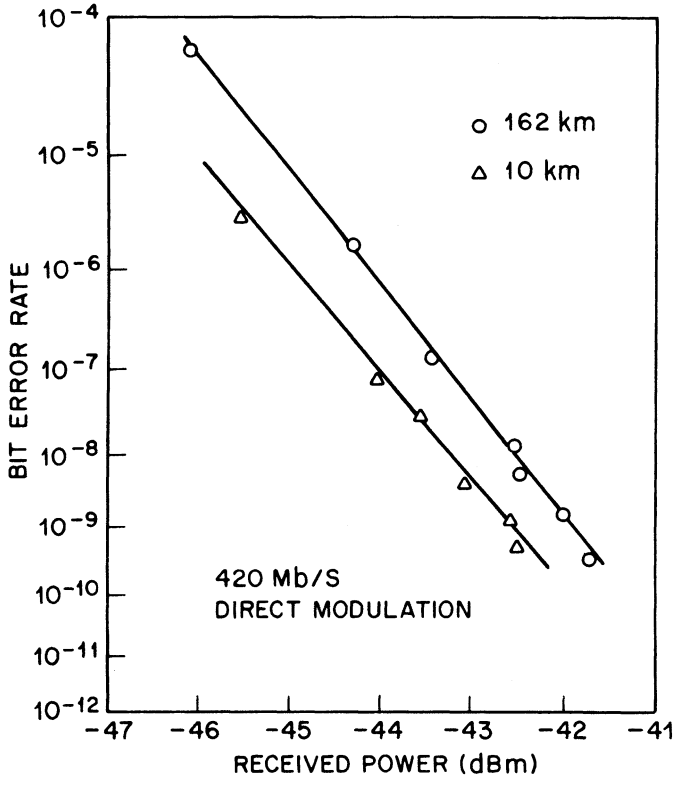


Fig. 8.12 Bit error rate as a function of received power. Results of a transmission experiment employing a C^3 laser modulated at the bit rate of 420 Mb/s over fibers 20 and 162 km long. (After Ref. 66)

received power of -16.2 dBm. The receiver sensitivity was limited because of the use of a PIN-diode photodetector. The use of an avalanche photodetector led to an increase in the fiber length to 120 km at the bit rate of 1 Gb/s.[68] To minimize chirp, it was necessary to operate the laser with an on-off ratio of 3.7:1. Such a low extinction ratio resulted in a power penalty of about 4.5 dB.

In these transmission experiments[67,68] the wavelength chirp was a limiting factor. To increase the bit rate even further, it is necessary to reduce the chirp. One approach eliminated the chirp problem altogether by using an external modulator. Since the semiconductor laser uses CW operation, the modulation-induced line broadening does not occur, and the source bandwidth is just the CW line width of the single longitudinal mode maintained by the device. The potential of this approach was demonstrated in 1985 by a

117-km transmission experiment[69] at the bit rate of 4 Gb/s. A 1.53-μm CSBH C^3 laser was used in the experiment, and the laser exhibited a mode suppression ratio of 5,000. A Ti:$LiNbO_3$ traveling-wave directional coupler switch served as the external modulator with an on-off ratio of 12:1. Figure 8.13 shows the measured bit-error-rate curves with and without fiber. A slight power penalty of about 0.5 dB occurring for the 117-km-long fiber is attributed to the finite dispersion [$\sim$18 ps/(nm·km)] of the fiber

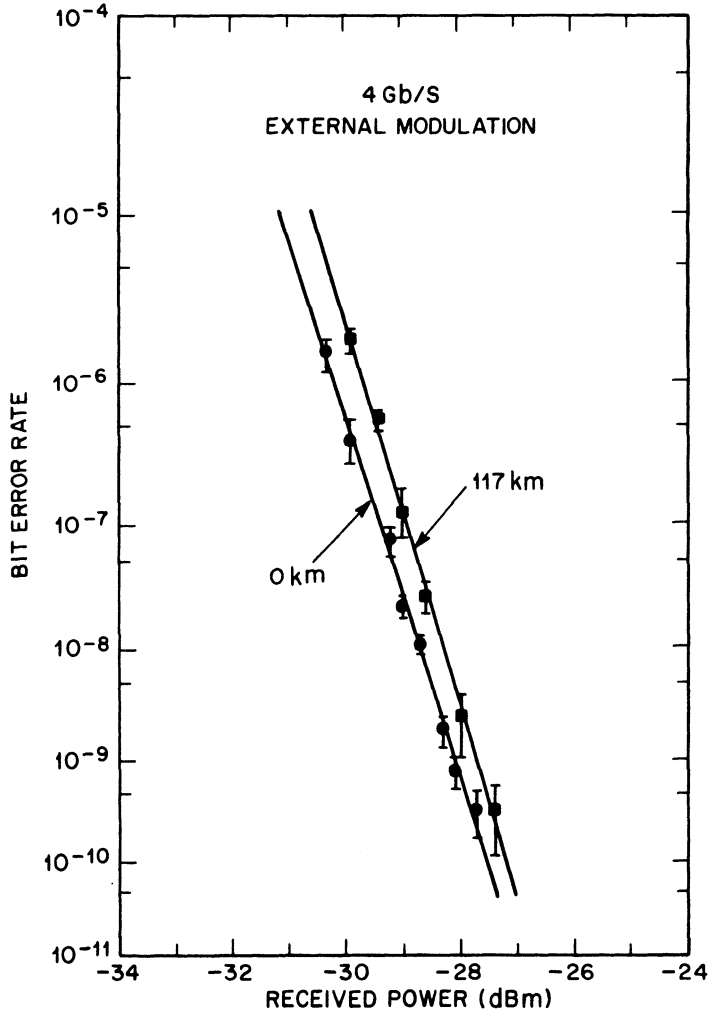


Fig. 8.13 Bit-error-rate curves for a 117-km transmission experiment employing a CW-operated C^3 laser with 4 Gb/s of external modulation. (After Ref. 69)

used in the experiment and represents the intrinsic information-bandwidth limitation.[69]

In spite of their record performance in the 1985 transmission experiment,[69] interest in C^3 lasers declined because they had problems related to mode hopping and stability under field conditions. Tunable DFB and DBR lasers (discussed in Sec. 7.6) were developed to meet the somewhat conflicting requirements of wavelength tuning and stability. Some coupled-cavity semiconductor lasers have continued to be attractive: they are external-cavity lasers belonging to the active-passive class, and the next section discusses their performance.

8.4.5 External-Cavity Semiconductor Lasers

So far we have considered the operating characteristics of active-active devices. As mentioned before, the active-passive scheme, where an external cavity is used to control the longitudinal modes of a conventional semiconductor laser, also has potential applications in optical fiber communications and has attracted considerable attention.[29–40] In this section we briefly describe the performance of such external-cavity semiconductor lasers. Although the capability of longitudinal-mode control through an external cavity has long been known,[4,61] it is only recently that the technique has been applied for InGaAsP lasers oscillating in the 1.3–1.6-μm wavelength range.

In one approach[29] the external cavity was formed by placing a concave spherical reflector at a distance of about 200 μm from one laser facet. The use of confocal geometry resulted in significant coupling even when the laser facet was not antireflection-coated. As discussed in Sec. 8.3, the coupled-cavity device selects the single longitudinal mode closest to the gain peak with the lowest cavity loss. Furthermore, the wavelength tuning can be achieved by changing the external-cavity length, which can be controlled either thermally[29] or electronically[36] (using a piezoelectric transducer in the latter case). For a fixed external-cavity length, tuning can also be achieved by changing the temperature of the laser cavity. A temperature change alters the refractive index resulting in a change in the optical length; this in turn shifts the laser-cavity modes.

The tuning range or the number M of the laser-cavity modes, over which the coupled-cavity laser can be made to hop discretely before returning to its original wavelength, depends on the relative cavity lengths. Similar to the case of C^3 lasers, $M \cong \mu_1 L_1/\mu_2 L_2$ in the long-short geometry, where the external-cavity length $L_2 \ll L_1$ and $\mu_2 \cong 1$ for the air-filled cavity. For a 250-μm-long laser, $M = 6$ if $L_2 = 150$ μm, and a tuning range of 6 nm can be easily achieved.[36] The MSR also depends on the external-cavity length[36]

partly because, for a given radius of curvature of the spherical reflector, the intercavity coupling changes with L_2. However, MSR values of more than 30 dB can be achieved with proper design. The practical utility of spherical-mirror external-cavity design was demonstrated in an early 102-km transmission experiment[63] at a bit rat of 140 Mb/s.

In a slightly different approach[34,38] a graded-index lens (GRINROD) with an antireflection coating at the front end and a reflective coating at the far end was used for the external cavity. The lens was epoxied inside a glass capillary tube, which in turn was held by a copper stud epoxied to a laser heat sink, as shown in Fig. 8.2. The GRINROD cavity was about 160 μm long, leading to the optical length ratio of 4:1 for the laser and external cavities. An advantage of the use of a GRINROD is that its focusing properties ensure significant optical coupling between the two cavities.

The performance of GRINROD external-cavity lasers was demonstrated in a 99-km transmission experiment[38] at the bit rate of 1 Gb/s. A 1.55-μm ridge waveguide laser was coupled to the GRINROD as shown in Fig. 8.2. The coupled-cavity laser exhibited an MSR of better than 30 dB. The average power launched into the fiber was 0.96 dBm and the on-off extinction ratio was 3.5:1. Despite a power penalty of approximately 3 dB caused by the low extinction ratio, a bit error rate of 1×10^{-9} was obtained with a received power of -26.5 dBm.

Short coupled-cavity semiconductor lasers[33,37] have also been shown to maintain the single-longitudinal-mode operation at modulation frequencies lying in the gigahertz range. In this approach a short laser ($L_1 = 50$–80 μm) is coupled to a short ($L_2 = 30$–80 μm) air-filled cavity by simply placing a cleaved semiconductor chip with a gold-coated facet next to the laser on the common copper heat sink. The main advantage of short cavities is that the longitudinal modes are relatively far apart leading to an improved gain roll-off discrimination for the adjacent pairs of coincident modes (see Fig. 8.1).

In these external-cavity schemes, mode selectivity arises from interference between the waves propagating in the two cavities. An additional mode-selective mechanism can be introduced if the feedback from the far end of the external cavity is dispersive, i.e., the reflectivity of the external mirror is wavelength-dependent. This can be achieved by using a grating[39] or a properly designed frequency-selective filter.[35] The use of a grating has the additional advantage that the laser wavelength can be tuned over a considerable range (~ 50 nm) by rotating the grating.

An aspect of coupled-cavity semiconductor lasers that has attracted considerable attention is related to their ability to exhibit a smaller CW line width compared with that of a conventional semiconductor laser.[39,70,71] The line width of a single-cavity semiconductor laser has been discussed in Sec.

6.5.3 and is typically 100 MHz at a power level of 1 mW. For a C^3 laser or a GRINROD external-cavity laser with similar cavity lengths, this number can be reduced by about one order of magnitude.[70] In the case of external-cavity semiconductor lasers, however, the line width can be further reduced by increasing the length of the external cavity. The physical basis for the line-width reduction can be understood by using Eq. (6.5.52). An increase in the total cavity length has the effect of increasing the photon lifetime. This results in a much larger number of intracavity photons at a given output power and, as shown by Eq. (6.5.52) the line width of the single longitudinal mode decreases. Line widths as small as 10 kHz have been demonstrated using few-centimeter-long external cavities.[39]

The performance of external-cavity semiconductor lasers can be considerably improved if a wavelength-selective element such as a grating is used in the external cavity. The simplest scheme discussed earlier makes use of a mechanically ruled grating. However, distributed feedback (see Chapter 7) can also be used to provide a Bragg grating that can be placed at the end of the external cavity.[72–78] In one scheme, the Bragg grating used to provide the distributed feedback is fabricated on a silicon chip.[72–74] A rate-equation approach has been used to model such lasers in which the active section is coupled to a high-Q external cavity.[73,75] It shows that the static frequency chirp is reduced by a factor that depends on the operating conditions. Under the same operating conditions the line width is reduced by the square of that factor.[73] This relationship between the static chirp and the line width under CW operation has been verified experimentally.[72] Line widths below 0.5 MHz have been obtained, and even smaller line widths are possible. The dynamic performance of such external-cavity semiconductor lasers is also considerably improved.[75] They are also capable of producing picosecond optical pulses through mode locking.[76] In another approach the grating is etched directly onto a fiber that is coupled to the multimode laser.[77] Such a fiber-grating laser has been used to produce 18.5-ps pulses at a repetition rate of 2.37 GHz.[78]

External-cavity semiconductor lasers do not usually have a monolithic design, and this is a major limitation. It is possible to design a monolithic external-cavity semiconductor laser by forming the passive waveguide on the same substrate used for fabrication of the active section. In one 1.55-μm device, the 0.6-mm-long Bragg mirror (acting as an integrated external grating) was separated from the 0.7-mm active region by a 4.2-mm passive waveguide, resulting in a 5.5-mm-long external-cavity semiconductor laser.[79] All three components were formed on the same InP substrate to preserve the monolithic nature of the device. This laser produced 20-ps optical pulses at 8.1 GHz repetition rate through active mode locking. In another implementation, a 1.2-mm-long laser provided 200-kHz line width at 17-mW

operating power. The laser has a low current (8 mA) and could be operated up to 28 mW.[80] External-cavity semiconductor lasers have found applications in soliton communication systems.[81] Soliton transmission over 15,000 km at 5 Gb/s and over 11,000 km at 10 Gb/s has been demonstrated by using the recirculating fiber-loop configuration[82] and an external-cavity mode-locked laser. The highest bit rate of 32 Gb/s was demonstrated in 1992 in another soliton experiment that used a mode-locked external-cavity semiconductor laser.[83]

8.5 DIVERSE APPLICATIONS

In Sec. 8.4 the performance of coupled-cavity semiconductor lasers was discussed from the standpoint of their applications in optical fiber communications as a single-frequency optical source. However, such lasers exhibit features that are useful for many other applications as well. This is particularly true for active-active three-terminal devices, and many interesting applications of these have already been proposed and demonstrated.[28] In this section we briefly describe some of the diverse applications of C^3 lasers.

Optical Bistability. The phenomenon of optical bistability has drawn considerable attention[84] for many years. An *optical bistable system* has two stable stationary states for some range of operating conditions and displays hysteresis when switched from one state to another by changing an externally controlled operating parameter. Optical bistability in semiconductor lasers was proposed[5] in 1964 and demonstrated[85] soon thereafter. By applying different currents to different parts of the laser cavity, the laser was effectively (but not physically) separated into two sections. One section acted as a saturable absorber and led to bistable operation. The method of inhomogeneous current injection and the associated bistability and self-pulsing phenomena have attracted considerable attention.[86,87] In the saturable absorption mechanism, bistability occurs close to the lasing threshold over a limited current and temperature range.

C^3 lasers have been found[88,89] to show a different kind of optical bistability, one not based on the phenomenon of saturable absorption. In this scheme, bistability occurs when the two cavities are biased close to or above threshold. Furthermore, depending on the relative bias levels, two kinds of bistable behavior are observed. In one case the total output power in a single longitudinal mode displays bistability and hysteresis.[88,90] In the other case the bistability is associated with a mode hop occurring in the C^3 laser with a change in the controller current. The powers associated

with each longitudinal mode exhibit bistability, while the total power in both modes remains relatively constant.[89]

In Fig. 8.14 we show the bistable behavior for the case in which the C^3 laser maintains the same longitudinal mode. When cavity 2 is unbiased, one observes usual L-I characteristics. However, when $I_2 = 21$ mA and cavity 2 is biased above threshold ($I_2^{th} = 18$ mA), hysteresis is observed in the output of both cavities. The largest hysteresis loop, however, occurs for P_2. The kink-like behavior in P_2 when $I_1 < I_1^{th}$ is due to mode hopping and is similar to that shown in Fig. 8.8. The hysteresis width ΔI_1 depends on I_2 and is ~ 12 mA for $I_2 = 21$ mA. However, it increases to ~ 35 mA when $I_2 = 24.5$ mA. Such large hysteresis widths are useful in many applications.[91]

A theoretical analysis of the bistable behavior seen in Fig. 8.14 can be carried out[90] using the single-mode rate equations (8.3.19)–(8.3.21) generalized to account for the intercavity coupling. Bistability is found to occur most readily when the coupling phase $\theta = \pi/2$. The hysteresis width depends on the coupling constant C, defined by Eq. (8.3.6), and on the relative biasing levels $\bar{I}_n$ [$\bar{I}_n = I_n/I_n^{th}$ $(n = 1, 2)$] of the two cavities. The physical origin of optical bistability can be understood by noting that,

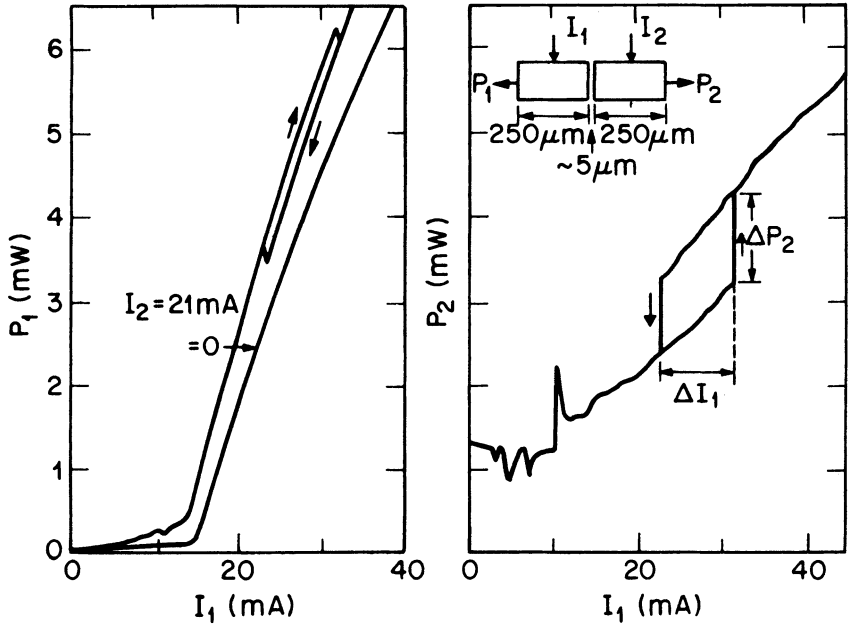


Fig. 8.14 Optical bistability in a C^3 laser for the configuration shown schematically in the inset. Output powers P_1 and P_2 from two facets are shown as functions of I_1 while I_2 is kept fixed. (After Ref. 88)

depending on the relative biasing levels, one cavity is optically pumped at the expense of the other. When $\bar{I}_1 < \bar{I}_2$, cavity 1 is optically pumped so that $P_1 > \bar{P}_1$ and $P_2 > \bar{P}_2$, where $\bar{P}_1$ and $\bar{P}_2$ are the powers expected in the absence of coupling. When $\bar{I}_1 > \bar{I}_2$, the opposite happens. When $\bar{I}_1 \cong \bar{I}_2$, both lasers compete since both are equally above threshold. Depending on the direction in which $\bar{I}_1$ is increased, P_2 is below or above $\bar{P}_2$. This leads to the bistability and hysteresis seen in Fig. 2.14.

We now consider spectral bistability occurring when the C^3 laser is made to switch from one longitudinal mode to the adjacent one by changing current I_1 in the controller section. Figure 8.15 shows the spectrally resolved output power at wavelengths corresponding to the two longitudinal modes.[89] In the bistable region the C^3 device lases in either of two longitudinal modes but never in both simultaneously. Which of the two modes will lase depends on how the operating point is approached. In contrast to the result shown in Fig. 8.14, the total output power in both modes is approximately constant while crossing the bistable region.

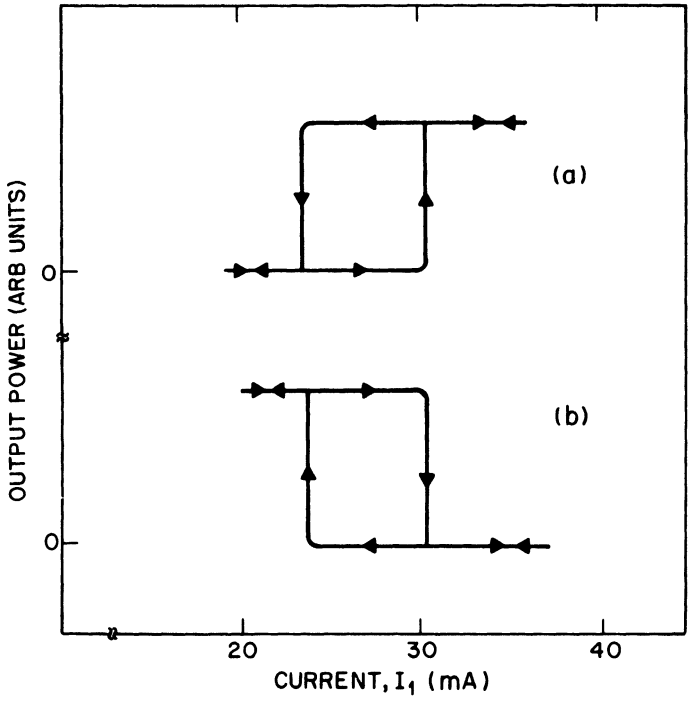


Fig. 8.15 Spectral bistability in a C^3 laser. Output power P_2 when (a) $\lambda = 1.5122$ μm and (b) $\lambda = 1.5102$ μm is shown as a function of I_1. See the inset of Fig. 8.14 for notation. (After Ref. 89)

The physical origin of spectral bistability[47] is different from that for the power bistability considered earlier. When the currents in the two sections of a C^3 laser are varied, there are zones of single-mode operation in the I_1-I_2 plane. Depending on the coupling phase, these zones may or may not overlap. It was found[47] that when the coupling phase θ is negative, the zones overlap resulting in bistability. By contrast, for positive values of θ less than $\pi/2$, there are voids between the current zones, corresponding to unstable regions with no steady-state operation.

Optical bistability of C^3 lasers may be useful for a wide range of signal-processing functions such as optical limiting, optical switching, and optical logic operations.[91] Spectral bistability can also be useful in optical communication systems employing frequency-shift keying. If the laser is biased to operate in the middle of the hysteresis loop (see Fig. 8.15) one or the other wavelength can be selected by applying a small modulation current to one cavity. A basic advantage of such a scheme is the fast switching from one state to another. Since the C^3 laser operates above threshold for both the on and off states, a switching time of less than 1 ns is easily obtained.[89]

Optoelectronic Logic. The ability of a C^3 laser to switch from one longitudinal mode to another when a short current pulse is applied to the controller section can be exploited for use in basic logic operations. In a feasibility demonstration,[92] the two longitudinal modes were selected by applying different current pulses to the two cavity sections. A grating was used to spectrally resolve the output. The dc bias was selected to operate the laser below but close to the threshold. When a current pulse was applied to either of the cavities, the laser operated above threshold at one wavelength, say λ_1. However, when the two pulses were on simultaneously, the laser wavelength changed to a second wavelength, λ_2. By partially overlapping the two pulses, it was shown that the output at λ_1 and λ_2 exhibited characteristics of the basic logic operations. The present device accepts electrical input to produce optical output and should be distinguished from optical bistable devices, where both input and output are in the optical form.[91]

Optical Routing. A related application is to use a C^3 laser as a switching element for routing information among different channels. Again, this application takes advantage of the electronic tuning of C^3 lasers. The route or channel is determined by the wavelength of the optical signal. The number of channels and the corresponding access times are determined by the number of discrete wavelengths over which the C^3 laser can be tuned and how fast the wavelength can be switched. As we saw in Sec. 8.4.1, the number of longitudinal modes over which the C^3 laser can be tuned depends on the

relative cavity lengths; up to 15 discrete modes can be used with a proper design. The switching time depends on the biasing of the wavelength control section and is determined by the carrier lifetime when the controller is biased below threshold. In a feasibility demonstration,[93] switching among four channels was achieved with a switching time of ~ 2 ns.

Mode Locking. Mode locking is often used to obtain short optical pulses from a semiconductor laser.[94–97] The output generally takes the form of a continuous train of short-pulses. For some applications it is desirable to be able to isolate a single pulse or a predetermined number of pulses by a suitable gating mechanism. The use of a C^3 laser for this purpose has been demonstrated.[98] The basic idea is to use one cavity section as a gating diode that blocks the output of the mode-locked semiconductor laser until a current pulse is applied to decrease the absorption. A third cavity section acting as an electrically controlled absorber can also be incorporated.

Figure 8.16 shows schematically the experimental setup for the mode-locking scheme employing such a three-section semiconductor laser. The middle section acts as the semiconductor laser, which is actively mode-locked with a 200-MHz sinusoidal current superimposed on the dc bias. The cavity section between the semiconductor laser and the grating acts as an absorber whose absorption can be controlled by changing the current I_A. The third section, on the other side of the semiconductor laser, is the gating element. When no current is applied to the gate ($I_G = 0$), mode-locked pulses suffer

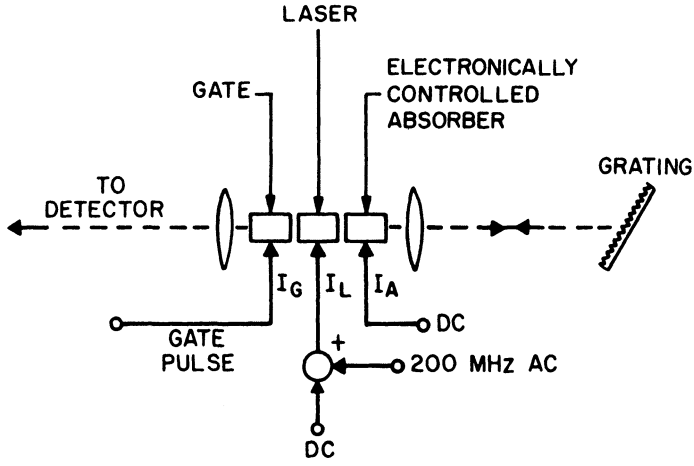


Fig. 8.16 Schematic illustration of a mode-locked C^3 laser with three coupled sections for electronic gating and optical absorption control. (After Ref. 98)

large optical absorption and are absorbed before reaching the detector. For an appropriate value of I_G, the gate becomes optically transparent. By adjusting the duration of the gating pulse, the desired number of mode-locked pulses can be obtained. The width of each pulse can be adjusted by changing the absorber current I_A. It is clear that the mode selectivity of a C^3 laser is not being used in this mode of operation. Since the three sections are formed by cleaving a single laser chip, one obtains self-aligned absorbing sections whose absorption can be electrically controlled.

Integrated Modulator. In an interesting application of three-section coupled-cavity devices, the third section is used as an external modulator that also acts as an amplifier.[99] Figure 8.17 shows the integrated device schematically. The two shorter sections constitute a C^3 laser that operates continuously in a single longitudinal mode. The longer section acts as the amplifier-modulator. Its output facet is antireflection-coated to prevent the amplified signal from being reflected back into the laser cavity. With no current applied to the amplifier section ($I_1 = 0$), the light emitted by the C^3 laser is absorbed inside this section. However, if I_1 is chosen to be sufficiently large, the light can be amplified. By modulating the amplifier current between zero and this value, pulsed output can be obtained while maintaining the single-frequency operation. Experimentally an amplifier gain of 20 dB with an output power of 1 mW has been demonstrated.[99] When used as an external modulator, optical pulses at 1-GHz repetition rate were obtained with a high on-off extinction ratio and an MSR of more than 20 dB.

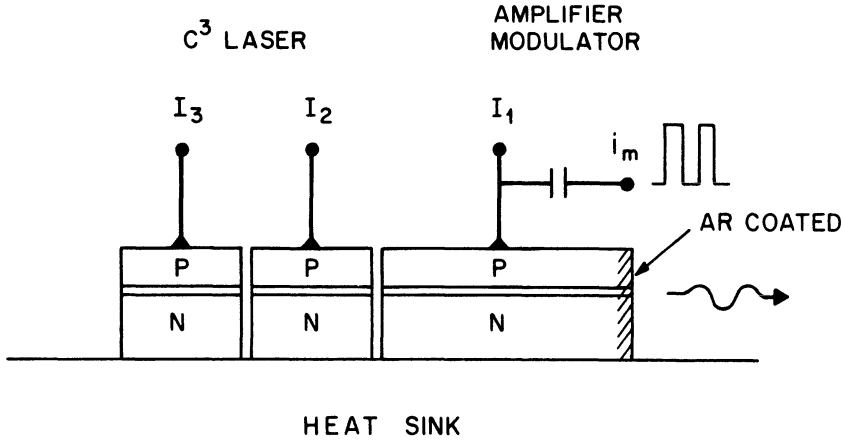


Fig. 8.17 Schematic illustration of a C^3 laser with an integrated external modulator that also acts as an amplifier. (After Ref. 99)

In a related experiment[100] one of the cavity sections of the C^3 laser was used as an intracavity modulator while the other section was modified to obtain Q switching. This was accomplished by pumping only a part of the laser cavity; i.e., its electrode was subdivided into two parts. The unpumped region acts as a saturable absorber and can be used to Q switch the C^3 laser. The intracavity modulator is useful to modulate these Q-switch light pulses after synchronizing the bit rate with pulse repetition rate. The initial experiment has shown the feasibility of this scheme.

In these diverse applications of C^3 lasers, one cavity section is used as a semiconductor laser while the other acts as an absorber whose absorption can be electrically controlled. The chief advantage is that the absorber (or amplifier) is integrated and aligned with the semiconductor laser. With a proper biasing scheme this section can also be used as a detector.

PROBLEMS

8.1 Explain the basic mechanism of mode selection in coupled-cavity semiconductor lasers. What are the main differences between active-active and active-passive coupled-cavity lasers?

8.2 Complete the derivation of Eqs. (8.3.2)–(8.3.4) by considering the scattering matrix appropriate to the geometry of Fig. 8.4. Consult Ref. 10 if necessary.

8.3 Derive the eigenvalue equation, Eq. (8.3.14), for a coupled-cavity laser and discuss how it is used to calculate the frequencies and threshold gains of a coupled-cavity semiconductor laser.

8.4 Explain the concept of the effective reflectivity in coupled-cavity semiconductor laser.

8.5 Complete the derivation of the generalized rate equations, Eqs. (8.3.19)–(8.3.21), for coupled-cavity semiconductor lasers.

8.6 Apply the generalized rate equations to an external-cavity semiconductor laser formed by placing a mirror next to a conventional semiconductor laser. Solve these equations, analytically or numerically, to study the dependence of output power (from the laser end) on the amplitude and phase of the coupling coefficient. Consult Ref. 54 if necessary.

8.7 Use the generalized rate equations to study the small-signal modulation response of an external-cavity semiconductor laser. Show that the frequency chirp at a given modulation frequency is reduced in comparison with a solitary semiconductor laser. Consult Ref. 54 if necessary.

8.8 Discuss the tuning mechanism in a C^3 laser. What are the ultimate limits on the tuning range in such a device?

REFERENCES

1. Kogelnik, H., and C. K. N. Patel. *Proc. IEEE* **50,** 2365 (1962).
2. Kleinman, D. A., and P. P. Kislink. *Bell Syst. Tech. J.* **41,** 453 (1962).
3. Smith, P. W. *Proc. IEEE* **60,** 422 (1972).
4. Salathé, R. P. *Appl. Phys.* **20,** 1 (1978); see other references cited therein for previous work.
5. Lasher, G. J. *Solid-State Electron.* **7,** 707 (1964).
6. Kosonocky, W. F., and R. H. Connely. *IEEE J. Quantum Electron.* **QE-4,** 125 (1968).
7. Allen, L. B., H. G. Koenig, and R. R. Rice. *Proc. SPIE Int. Soc. Opt. Eng.* **157,** 110 (1978).
8. Chang, M. B. and E. Garmire. *IEEE J. Quantum Electron.* **QE-16,** 997 (1980).
9. Coldren, L. A., B. I. Miller, K. Iga, and J. Rentschler. *Appl. Phys. Lett.* **38,** 315 (1981).
10. Coldren, L. A., F. Furuya, B. I. Miller, and J. A. Rentschler. *IEEE J. Quantum Electron.* **QE-18,** 1679 (1982).
11. Wang, S., H. K. Choi, and I. H. A. Fattah. *IEEE J. Quantum Electron.* **QE-18,** 610 (1982).
12. Tsang, W. T., N. A. Olsson, R. A. Linke, and R. A. Logan. *Electron. Lett.* **19,** 415 (1983).
13. Tsang, W. T., N. A. Olsson, and R. A. Logan. *Appl. Phys. Lett.* **42,** 650 (1983); *Electron. Lett.* **19,** 438 (1983).
14. Tsang, W. T., and N. A. Olsson. *Appl. Phys. Lett.* **43,** 527 (1983).
15. Temkin, H., J. P. van der Ziel, R. A. Linke, and R. A. Logan. *Appl. Phys. Lett.* **43,** 723 (1983).
16. Tsang, W. T., N. A. Olsson, R. A. Logan, J. A. Ditzenberger, I. P. Kaminow, and J.-S. Ko. *Appl. Phys. Lett.* **43,** 1003 (1983).
17. Antreasyan, A., and S. Wang. *Appl. Phys. Lett.* **43,** 530 (1983).
18. Fang, Z., and S. Wang. *Appl. Phys. Lett.* **44,** 13 (1984).
19. Coldren, L. A., K. J. Ebeling, J. A. Rentschler, C. A. Burrus, and D. P. Wilt. *Appl. Phys. Lett.* **44,** 368 (1984).
20. Ebeling, K. J., and L. A. Coldren. *Electron. Lett.* **20,** 69 (1984); *Appl. Phys. Lett.* **44,** 735 (1984).
21. Coldren, L. A., K. J. Ebeling, R. G. Swartz, and C. A. Burrus. *Appl. Phys. Lett.* **44,** 169 (1984).
22. Sudbo, A. S. *Electron. Lett.* **20,** 88 (1984).
23. Bowers, J. E., J. E. Bjorkholm, C. A. Burrus, L. A. Coldren, B. R. Hemenway, and D. P. Wilt. *Appl. Phys. Lett.* **44,** 821 (1984).
24. Olsson, N. A., and N. K. Dutta. *Appl. Phys. Lett.* **44,** 840 (1984).
25. Lee, T. P., C. A. Burrus, P.-L. Liu, W. B. Sessa, and R. A. Logan. *IEEE J. Quantum Electron.* **QE-20,** 374 (1984).
26. Olsson, N. A., N. K. Dutta, and K.-Y. Liou. *Electron. Lett.* **20,** 121 (1984).
27. Agrawal, G. P., N. A. Olsson, and N. K. Dutta. *Appl. Phys. Lett.* **45,** 119 (1984).
28. Tsang, W. T., Chap. 5 in *Semiconductors and Semimetals*, Vol. 22, Part B, ed. W. T. Tsang. New York: Academic Press, 1985.
29. Preston, K. R., K. C. Woollard, and K. H. Cameron. *Electron. Lett.* **17,** 931 (1981).
30. Garmire, E., G. Evans, and J. Niesen. *Appl. Phys. Lett.* **39,** 789 (1981).
31. Bessonov, Yu. L., A. P. Bogatov, P. P. Vasil'ev, V. N. Morozov, and A. B. Sergeev. *Sov. J. Quantum Electron.* **12,** 1510 (1982).
32. Choi, H. K., and S. Wang. *Electron. Lett.* **19,** 302 (1983).
33. Lin, C., C. A. Burrus, R. A. Linke, I. P. Kaminow, J. S. Ko, A. G. Dentai, R. A. Logan, and B. I. Miller. *Electron. Lett.* **19,** 561 (1983).
34. Liou, K.-Y. *Electron. Lett.* **19,** 750 (1983).
35. Dutta, N. K., E. I. Gordon, T. M. Shen, P. J. Anthony, and G. Zydzik. *IEEE J. Quantum Electron.* **QE-21,** 559 (1985).
36. van der Ziel, J. P., and R. M. Mikulyak. *IEEE J. Quantum Electron.* **QE-20,** 233 (1984).
37. Lin, C., G. Eisenstein, C. A. Burrus, and R. S. Tucker. *Electron. Lett.* **20,** 843 (1984).

38. Liou, K.-Y., C. A. Burrus, R. A. Linke, I. P. Kaminow, S. W. Granlund, C. B. Swan, and P. Besomi. *Appl. Phys. Lett.* **45,** 729 (1984).
39. Matthews, M. R., K. H. Cameron, R. Wyatt, and W. J. Devlin. *Electron. Lett.* **21,** 115 (1985); Wyatt, R. and W. J. Devlin. *Electron. Lett.* **19,** 110 (1983).
40. Matsuda, K., T. Fujita, J. Ohya, M. Ishino, H. Sato, H. Serizawa, and J. Shibata. *Appl. Phys. Lett.* **46,** 1028 (1985).
41. Spencer, M. B., and W. E. Lamb, Jr. *Phys. Rev. A* **5,** 893 (1972).
42. Ebeling, K. J., and L. A. Coldren. *J. Appl. Phys.* **54,** 2962 (1983).
43. Coldren, L. A., K. J. Ebeling, B. I. Miller, and J. A. Rentschler. *IEEE J. Quantum Electron.* **QE-19,** 1057 (1983).
44. Marcuse, D., and T. P. Lee. *IEEE J. Quantum Electron.* **QE-20,** 166 (1984).
45. Choi, H. K., K. L. Chen, and S. Wang. *IEEE J. Quantum Electron.* **QE-20,** 385 (1984).
46. Coldren, L. A., and T. L. Koch. *IEEE J. Quantum Electron.* **QE-20,** 659 (1984); **QE-20,** 671 (1984).
47. Henry, C. H., and R. F. Kazarinov. *IEEE J. Quantum Electron.* **QE-20,** 733 (1984).
48. Streifer, W., D. Yevick, T. L. Paoli, and R. D. Burnham. *IEEE J. Quantum Electron.* **QE-20,** 754 (1984); *Electron. Lett.* **20,** 553 (1984); *IEEE J. Quantum Electron.* **QE-21,** 539 (1985).
49. Shakir, S. A., and W. W. Chow. *Opt. Lett.* **9,** 202 (1984).
50. Streifer, W., D. Yevick, R. D. Burnham, and T. L. Paoli. *Appl. Phys. Lett.* **45,** 443 (1984); *Appl. Phys. Lett.* **45,** 937 (1984).
51. Buus, J., and M. J. Adams. *Electron. Lett.* **20,** 579 (1984); Adams, M. J., and J. Buus. *IEEE J. Quantum Electron.* **QE-20,** 99 (1984).
52. Vahala, K., and A. Yariv. *Appl. Phys. Lett.* **45,** 501 (1984); Vahala, K., J. Paslaski, and A. Yariv. *Appl. Phys. Lett.* **46,** 1025 (1985).
53. Hammer, J. M. *IEEE J. Quantum Electron.* **QE-20,** 1252 (1984).
54. Agrawal, G. P. *J. Appl. Phys.* **56,** 3110 (1984).
55. Agrawal, G. P. *Opt. Lett.* **10,** 10 (1985).
56. Antreasyan, A., and S. Wang. *IEEE J. Quantum Electron.* **QE-21,** 38 (1985).
57. Marcuse, D. *IEEE J. Quantum Electron.* **QE-21,** 154 (1985).
58. Coldren, L. A., G. D. Boyd, J. E. Bowers, and C. A. Burrus. *Appl. Phys. Lett.* **46,** 125 (1985).
59. Agrawal, G. P. *IEEE J. Quantum Electron.* **QE-21,** 255 (1985).
60. Coldren, L. A., and T. L. Koch. *J. Lightwave Technol.* **LT-2,** 1045 (1984); Koch, T. L., and L. A. Coldren. *J. Appl. Phys.* **57,** 740 (1985).
61. Elissev, P. G., I. Ismailov, and M. A. Manko. *JETP Lett.* **9,** 362 (1969).
62. Voumard, C., R. Salathé, and H. Weber. *Opt. Commun.* **13,** 130 (1975).
63. Cameron, K. H., P. J. Chidgey, and K. R. Preston. *Electron. Lett.* **18,** 650 (1982).
64. Kurokawa, K. *IEEE Trans. Microwave Theory Tech.* **MTT-13,** 194 (1965).
65. Fujita, K., J. Ohya, S. Ishizuka, K. Fuzito, and H. Sato. *Electron. Lett.* **20,** 417 (1984).
66. Kasper, B. L., R. A. Linke, J. C. Campbell, A. G. Dentai, R. S. Vodhanel, P. S. Henry, I. P. Kaminow, and J.-S. Ko. Postdeadline paper presented at the Eur. Conf. on Opt. Commun., Geneva (1983).
67. Linke, R. A., B. L. Kasper, J.-S. Ko, I. P. Kaminow, and R. S. Vodhanel. *Electron. Lett.* **19,** 755 (1983).
68. Linke, R. A., B. L. Kasper, J. C. Campbell, A. G. Dentai, and I. P. Kaminow. *Electron. Lett.* **20,** 489 (1984).
69. Korotky, S. K., G. Eisentein, A. H. Gnauck, B. L. Kasper, J. J. Veselka, R. C. Alferness, L. L. Buhl, C. A. Burrus, T. C. D. Huo, L. W. Stulz, K. C. Nelson, L. G. Cohen, R. D. Dawson, and J. C. Campbell. *J. Lightwave Technol.* **LT-3,** 1027 (1985).
70. Lee, T. P., C. A. Burrus, K.-Y. Liou, N. A. Olsson, R. A. Logan, and D. P. Wilt. *Electron. Lett.* **20,** 1011 (1984); Lee, T. P., C. A. Burrus, and D. P. Wilt. *Electron. Lett.* **21,** 54 (1985).

71. Liou, K.-Y., Y. K. Jhee, C. A. Burrus, K. L. Hall, and P. J. Anthony. *Electron. Lett.* **21,** 933 (1985).
72. Olsson, N. A., C. H. Henry, R. F. Kazarinov, H. J. Lee, and B. H. Johnson. *Appl. Phys. Lett.* **51,** 92 (1987).
73. Henry, C. H., and R. F. Kazarinov. *IEEE J. Quantum Electron.* **QE-23,** 1401 (1987).
74. Olsson, N. A., C. H. Henry, R. F. Kazarinov, H. J. Lee, and K. J. Orlowsky. *IEEE J. Quantum Electron.* **QE-24,** 143 (1988).
75. Agrawal, G. P., and C. H. Henry. *IEEE J. Quantum Electron.* **QE-24,** 134 (1988).
76. Morton, P. A., R. Adar, R. C. Kistler, C. H. Henry, T. Tanbun-Ek, R. A. Logan, D. L. Coblentz, A. M. Sergent, and K. W. Wecht. *Appl. Phys. Lett.* **58,** 1239 (1991).
77. Bird, D. M., J. R. Armitage, R. Kashyap, R. M. A. Fatah, and K. H. Cameron. *Electron. Lett.* **27,** 1115 (1991).
78. Morton, P. A., V. Mizrahi, S. G. Kosinski, L. F. Mollenauer, T. Tanbun-Ek, R. A. Logan, D. L. Coblentz, A. M. Sergent, and K. W. Wecht. *Electron. Lett.* **28,** 561 (1992).
79. Hansen, P. B., G. Raybon, U. Koren, B. I. Miller, M. G. Young, M. Chien, C. A. Burrus, and R. C. Alferness. *IEEE Photon. Technol. Lett.* **3,** 212 (1991).
80. Matsui, Y., T. Kunii, H. Horikawa, and T. Kamijoh. *IEEE Photon. Technol. Lett.* **3,** 424 (1991).
81. Agrawal, G. P. *Fiber-Optic Communication Systems.* New York: Wiley, 1992.
82. Mollenauer, L. F., E. Lichtman, G. T. Harvey, M. J. Neubelt, and B. M. Nyman. *Electron. Lett.* **28,** 792 (1992).
83. Andrekson, P. A., N. A. Olsson, M. Haner, J. R. Simpson, T. Tanbun-Ek, R. A. Logan, D. Coblentz, H. M. Presby, and K. W. Wecht. *IEEE Photon. Technol. Lett.* **4,** 76 (1992).
84. Gibbs, H. M. *Optical Bistability: Controlling Light with Light.* New York: Academic Press, 1985.
85. Nathan, M. I., J. C. Mariance, R. F. Rutz, A. E. Michel, and G. J. Lasher. *J. Appl. Phys.* **36,** 473 (1965).
86. Harder, C., K. Y. Lau, and A. Yariv. *IEEE J. Quantum Electron.* **QE-18,** 1351 (1982).
87. Kawaguchi, H. *IEE Proc. I* **129,** 141 (1982); *Appl. Phys. Lett.* **45,** 1264 (1984).
88. Dutta, N. K., G. P. Agrawal, and M. W. Focht. *Appl. Phys. Lett.* **44,** 30 (1984).
89. Olsson, N. A., W. T. Tsang, and R. A. Logan. *Appl. Phys. Lett.* **44,** 375 (1984).
90. Agrawal, G. P., and N. K. Dutta. *J. Appl. Phys.* **56,** 664 (1984).
91. Smith, P. W. *Opt. Eng.* **19,** 456 (1980); *Bell Syst. Tech. J.* **61,** 1975 (1982).
92. Tsang, W. T., N. A. Olsson, and R. A. Logan. *IEEE J. Quantum Electron.* **QE-19,** 1621 (1983).
93. Olsson, N. A., and W. T. Tsang. *IEEE J. Quantum Electron.* **QE-20,** 332 (1984).
94. Ho, H., H. Yokoyama, and H. Inoba. *Electron. Lett.* **16,** 620 (1980).
95. Holbrook, M. B., W. E. Sleat, and D. J. Bradley. *Appl. Phys. Lett.* **37,** 59 (1980).
96. Ippen, E. P., D. J. Eilenberger, and R. W. Dixon. *Appl. Phys. Lett.* **37,** 267 (1980).
97. Van der Ziel, J. P. *J. Appl. Phys.* **52,** 4435 (1981).
98. Tsang, W. T., N. A. Olsson, and R. A. Logan. *Appl. Phys. Lett.* **43,** 339 (1983).
99. Lee, T. P., C. A. Burrus, G. Eisenstein, W. B. Sessa, and P. Besomi. *Electron. Lett.* **20,** 625 (1984).
100. Lee, T. P., C. A. Burrus, W. B. Sessa, and P. Besomi. *Electron. Lett.* **20,** 1 (1984).

Chapter 9
QUANTUM-WELL SEMICONDUCTOR LASERS

9.1 ENERGY LEVELS

A double-heterostructure laser consists of an active layer sandwiched between two higher-gap cladding layers. The active-layer thickness is typically in the range of 0.1–0.3 μm. In the last few years, double-heterostructure lasers with an active-layer thickness of ~ 10 nm have been fabricated. The carrier (electron or hole) motion normal to the active layer in these structures is restricted. As a result, the kinetic energy of the carriers moving in that direction is quantized into discrete energy levels similar to the well-known quantum-mechanical problem of the one-dimensional potential well, and hence these lasers are called *quantum-well lasers.*

When the thickness of the active region (or any low-gap semiconductor layer confined between higher-gap semiconductors) becomes comparable to the de Broglie wavelength ($\lambda \cong h/p$), quantum-mechanical effects are expected to occur. These effects are observed in the absorption and emission characteristics (including laser action) and transport characteristics (including phenomena such as tunneling).[1–8] The optical characteristics of quantum-well double-heterostructure lasers were initially studied[1] in 1974. Since then, AlGaAs quantum-well lasers have been extensively studied.[2–5] Quantum effects also occur on surface states of a semiconductor or other types of surfaces. These effects have been discussed in several review articles.[6,7]

A carrier (electron or hole) in a double heterostructure is confined within a three-dimensional potential well. The energy levels of such carriers are obtained by separating the system Hamiltonian into three parts, corresponding to the kinetic energies in the x, y, and z directions. When the thickness of the active layer (L_z) is comparable to the de Broglie wavelength, the kinetic energy corresponding to the carrier motion along the z direction is quantized. Along the x and y directions, the energy levels form a continuum of states given by

$$E = (\hbar^2/2m)(k_x^2 + k_y^2) \tag{9.1.1}$$

where m is the effective mass of the carrier and k_x and k_y are the wave-vector components along the x and y directions, respectively. Thus the electrons or holes in a quantum well may be viewed as forming a two-dimensional Fermi gas.

The energy levels in the z direction are obtained by solving the Schrödinger equation for a one-dimensional potential well given by

$$\begin{aligned} E\psi &= -\frac{\hbar^2}{2m}\frac{d^2\psi}{dz^2} && \text{inside the well} \\ & && (0 \le z \le L_z) \\ E\psi &= -\frac{\hbar^2}{2m}\frac{d^2\psi}{dz^2} + V\psi && \text{outside the well} \\ & && (z \ge L_z;\ z \le 0) \end{aligned} \tag{9.1.2}$$

where ψ is the Schrödinger wave function and V is the depth of the potential well. For the limiting case of an infinitely deep well, the energy levels and wave functions are

$$E_n = \frac{\hbar^2}{2m}\left(\frac{n\pi}{L_z}\right)^2 \tag{9.1.3a}$$

$$\psi_n = \begin{cases} A \sin \dfrac{n\pi z}{L_z} & \text{inside the well} \\ 0 & \text{outside the well} \end{cases} \tag{9.1.3b}$$

where n is a positive integer and A is a normalization constant. For very large L_z, Eq. (9.1.3) yields a continuum of states and the system no longer exhibits quantum effects.

For a finite well, the energy levels and wave function can be obtained from Eq. (9.1.2) by using the boundary conditions that ψ and $d\psi/dz$ are continuous at the interfaces $z = 0$ and $z = L_z$. The solution is

$$\psi = \begin{cases} A \exp(k_1 z) & (z \le 0) \\ B \sin(k_2 z + \delta) & (0 \le z \le L_z) \\ C \exp(-k_1 z) & (z \ge L_z) \end{cases} \tag{9.1.4}$$

where

$$k_1 = \left[\frac{2m(V - E)}{\hbar^2}\right]^{1/2} \qquad k_2 = \left(\frac{2mE}{\hbar^2}\right)^{1/2}. \tag{9.1.5}$$

The quantities A, B, C, and δ are constants. Using the boundary conditions stated above, the following eigenvalue equation is obtained:

$$\tan(k_2 L_z) = k_1/k_2. \tag{9.1.6}$$

The eigenvalue equation is numerically solved to yield the energy levels E_n of a particle in a potential well. Figure 9.1 shows the calculated energy levels of the bound states as a function of the well depth.

The potential well for electrons and holes in a double heterostructure depends on the materials involved. The following values for the conduction-band (ΔE_c) and valence-band (ΔE_v) discontinuities are found experimentally[9a] for GaAs-AlGaAs double heterostructures:

$$\Delta E_c/\Delta E = 0.67 \pm 0.01 \tag{9.1.7a}$$

$$\Delta E_v/\Delta E = 0.33 \pm 0.01 \tag{9.1.7b}$$

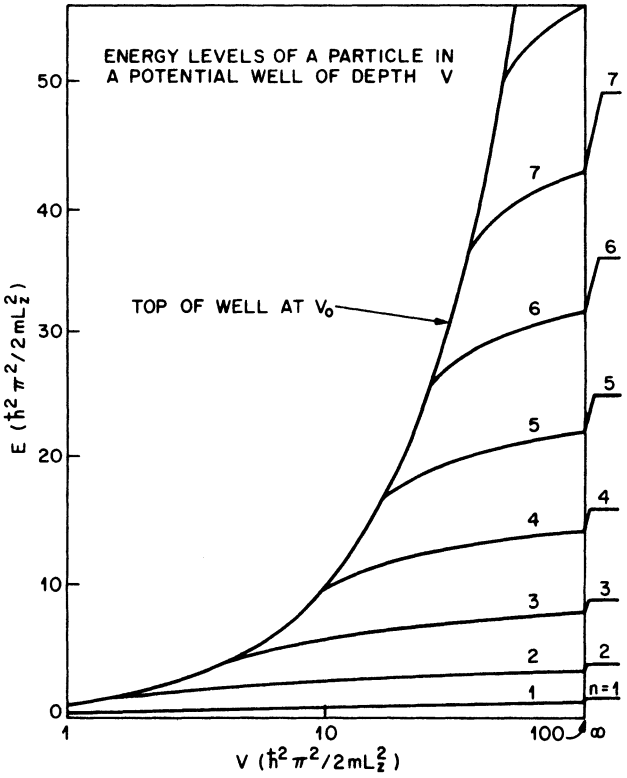


Fig. 9.1 The dependence of the bound-state energy levels E on the depth of the potential well V. The plot is given in reduced energy units, $\hbar^2\pi^2/2mL_z^2$. The curve represents the top of the well ($E = V_0$). Energy levels for an infinite potential well are shown on the extreme right. (After Ref. 8)

where ΔE is the band-gap difference between the confining and active layers. A knowledge of ΔE_c and ΔE_v is necessary in order to accurately calculate the energy levels. For an InGaAsP-InP double heterostructure, the values obtained are[9b]

$$\Delta E_c/\Delta E = 0.39 \pm 0.01 \tag{9.1.8a}$$

$$\Delta E_v/\Delta E = 0.61 \pm 0.01. \tag{9.1.8b}$$

Taking into account discrete states along the z direction and continuous states along the x and y directions, the energy eigenvalues for a particle confined in the quantum well are

$$E(n, k_x, k_y) = E_n + \frac{\hbar^2}{2m_n^*}(k_x^2 + k_y^2) \tag{9.1.9}$$

where E_n is the nth confined-particle energy level for carrier motion normal to the well and m_n^* is the effective mass for this level. Figure 9.2 shows

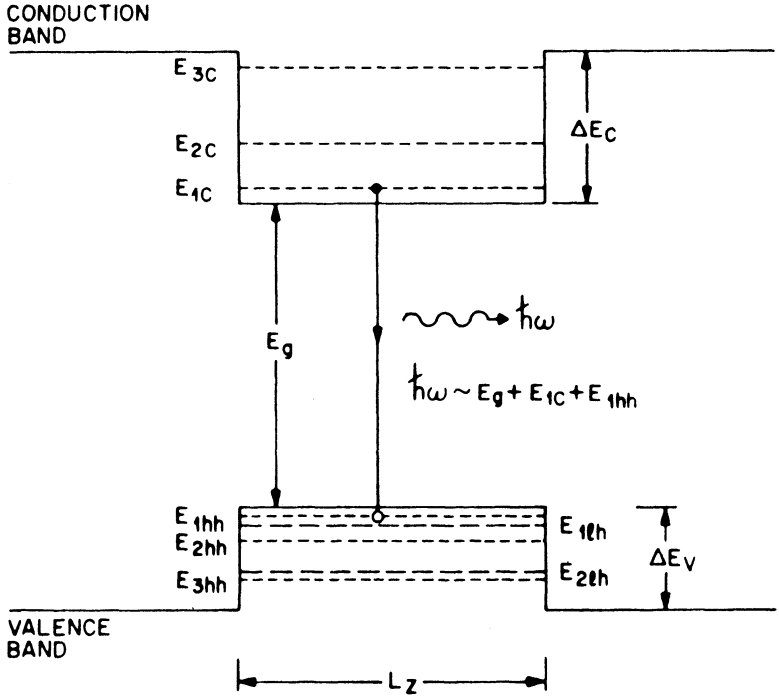


Fig. 9.2 Schematic representation of the confined-particle energy levels of electrons, heavy holes (short dashes), and light holes (long dashes) in a quantum well. (After Ref. 2 © 1980 IEEE)

schematically the energy levels E_n of electrons and holes confined within a quantum well. The confined-particle energy levels E_n are denoted by E_{1c}, E_{2c}, and E_{3c} for electrons; E_{1hh}, E_{2hh}, and E_{3hh} for heavy holes; and E_{1lh} and E_{2lh} for light holes. These quantities can be calculated by solving the eigenvalue equation (9.1.6) for a given potential barrier (ΔE_c or ΔE_v).

9.2 DENSITY OF STATES

In a quantum-well structure, the kinetic energy of confined carriers (electrons or holes) for motion normal to the well (z direction) is quantized into discrete energy levels. This modifies the density of states from the well-known three-dimensional case [see Eq. (3.2.26)]. Using the principle of box quantization for kinetic energies along the x and y directions, the number of electron states per unit area in the x-y plane for the ith subband, within an energy interval dE, is given by

$$D_i(E)\,dE = 2\frac{d^2\mathbf{k}}{(2\pi)^2} \tag{9.2.1}$$

where the factor 2 arises from the two spin states and $\mathbf{k} = (k_x, k_y)$ is the momentum vector. Using the parabolic-band approximation, namely, $E = \hbar^2k^2/2m_{ci}$, and the relation $d^2\mathbf{k} = 2\pi k\,dk$, we obtain

$$D_i = \frac{m_{ci}}{\pi\hbar^2} \tag{9.2.2}$$

where m_{ci} is the effective mass of the electrons in the ith subband of the conduction band. Thus the density of states per unit volume is given by

$$g_{ci} = \frac{D_i}{L_z} = \frac{m_{ci}}{\pi\hbar^2 L_z}. \tag{9.2.3}$$

A similar equation holds for holes in the valence band. The density of states for the three-dimensional case [from Eq. (3.2.26)], rewritten here for the sake of comparison, is

$$\rho_c(E) = 4\pi\left(\frac{2m_c}{h^2}\right)^{3/2} E^{1/2}. \tag{9.2.4}$$

A comparison of Eqs. (9.2.3) and (9.2.4) shows that the density of states in a quantum well is independent of the carrier energy. Figure 9.3 shows the modification of the density of states in a quantum well. This modification

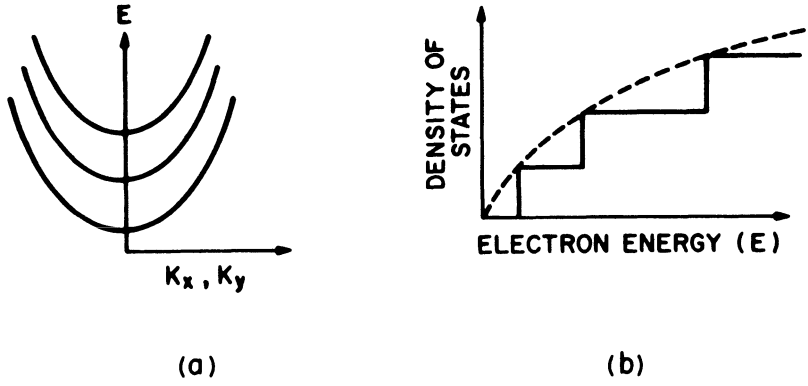


Fig. 9.3 (a) Schematic representation of the two-dimensional parabolic subbands. (b) Schematic representation of the density of states in a quantum well. Dashed curve shows the density of states in a bulk semiconductor.

can significantly alter the recombination rates in a quantum-well double heterostructure compared with those for a regular double heterostructure, as discussed in the next section.

The simple model for the density of states just described neglects the effect of collisions between carriers that would broaden the discrete energy levels in the z direction. A necessary condition for the validity of the model is that the collisional broadening be small relative to the separation between discrete levels. If we use Eq. (9.1.3a) to obtain the level separation $E_2 - E_1$ between the two lowest levels, the condition is

$$\frac{h}{2\pi\tau} \ll \frac{3h^2}{8m_c L_z^2} \tag{9.2.5}$$

where τ is the electron-electron scattering time. The term on the right-side of the equation gives the energy separation between the first two levels for an infinite square well. It is experimentally determined that $\tau < 0.5$ ps for bulk GaAs.[10] Using $\tau = 0.5$ ps and $m_c = 0.071m_0$ (for GaAs), we get from Eq. (9.2.5) that $L_z \ll 110$ nm. This estimate shows that Eq. (9.2.3) for the density of states is a good approximation for a well thickness that is less than ~ 30 nm.

If n is the number of electrons in the conduction band, the Fermi level can be obtained from

$$n = \sum_i \int_{E_{i0}}^{\infty} g_{ci} f_c(E_i)\, dE_i \tag{9.2.6}$$

where

$$f_c(E_i) = \frac{1}{1 + \exp[(E_i - E_{fc})/k_B T]}$$

is the Fermi function, E_{fc} is the quasi-Fermi energy in the conduction band, and E_i is the confined-particle energy level of the ith subband. The summation is performed over all subbands in the conduction band. Since g_{ci} does not depend on E_i, the integration is readily performed to obtain

$$n = k_B T \sum_i g_{ci} \ln\left[1 + \exp\left(\frac{E_{fc} - E_{i0}}{k_B T}\right)\right]. \tag{9.2.7}$$

Using the Boltzmann approximation and assuming that only one subband is occupied, Eq. (9.2.7) becomes

$$n = N_c \exp\left(\frac{E_{fc}}{k_B T}\right) \tag{9.2.8}$$

where $N_c = g_c k_B T$ and E_{fc} is measured from the bottom of the subband at E_{i0}. The Fermi factor under the same approximation takes the form

$$f_c(E) \cong \exp\left(\frac{E_{fc} - E}{k_B T}\right) \cong \frac{n}{N_c} \exp\left(-\frac{E}{k_B T}\right). \tag{9.2.9}$$

Similar relations hold for holes in the valence band.

9.3 EXPERIMENTAL OBSERVATION OF CONFINED STATES

The existence of confined-particle states was demonstrated[1,8] using optical absorption measurements. Figure 9.4 shows a schematic diagram of the sample used. The samples were grown by molecular beam epitaxy and had up to 100 alternate layers of GaAs and AlGaAs (with a total GaAs thickness of ~400 nm) in order to provide sufficient absorption. The measured absorption spectra at 2 K are shown in Fig. 9.5. The direct band gap of GaAs is 1.5192 eV at 2 K, and that of the cladding layer used (AlGaAs) is 1.75 eV. Thus the absorption peaks observed between these two energies are due to confined-particle discrete states and correspond to transitions between the valence band and various subbands within the conduction band.

Temkin et al.[11] have reported the measurements of absorption spectra of a 50-well InGaAs-InP sample with a well thickness of 12.5 nm. Figure 9.6

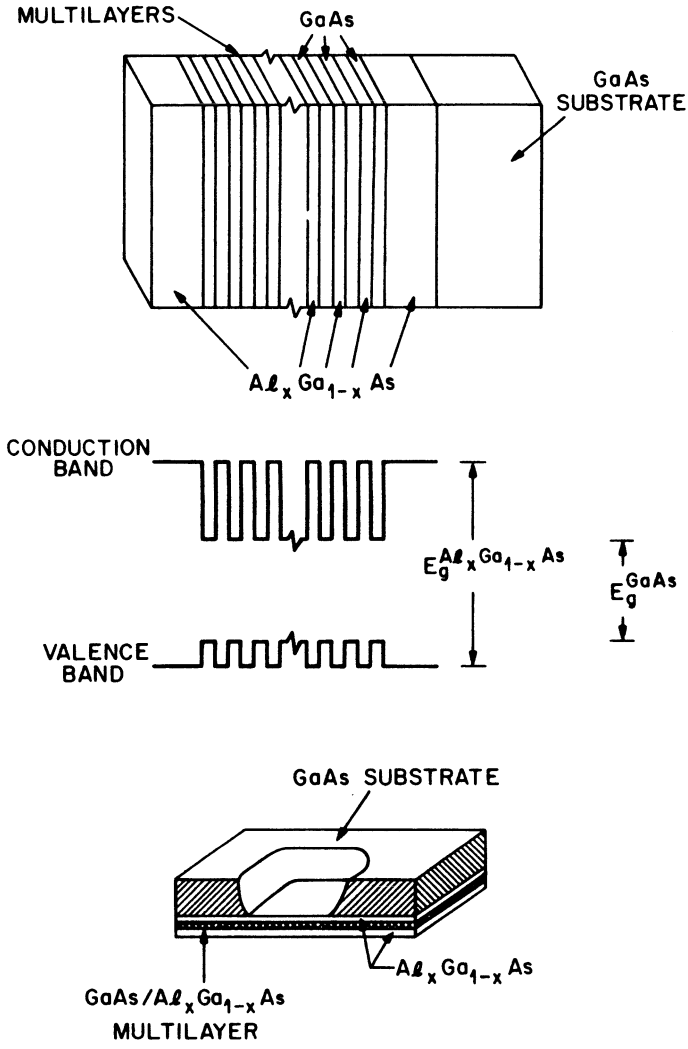


Fig. 9.4 Schematic of the multilayer sample used for optical absorption measurements. Middle drawing shows the variation in the valence- and conduction-band edges caused by the alternating GaAs and AlGaAs layers. (After Ref. 8)

shows the measured data. The sample was prepared by gas-source molecular-beam epitaxy (MBE) (see Sec. 4.5). The transmission dips correspond to absorption peaks caused by transitions between various subbands. The observed shift in the absorption peaks with temperature occurs because band gap varies with temperature.

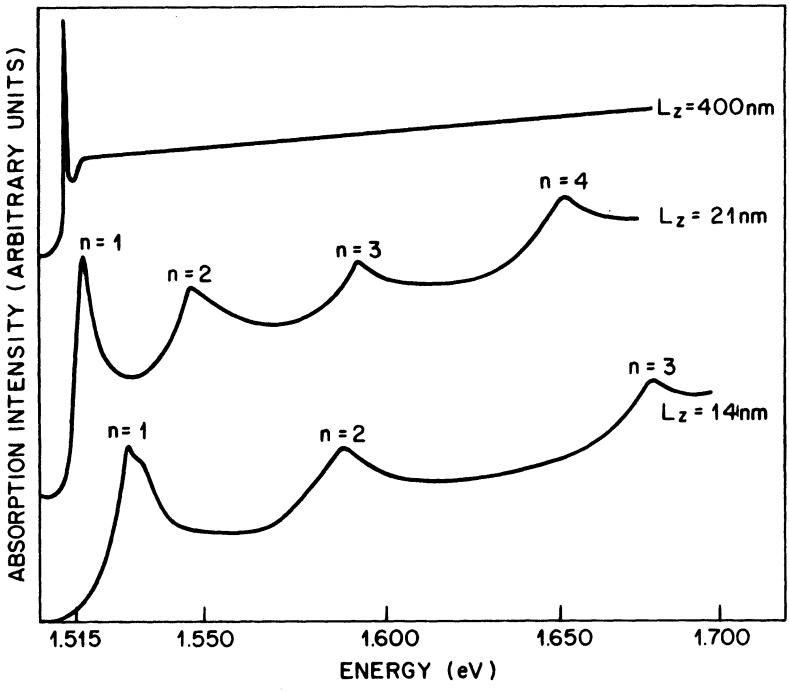


Fig. 9.5 Measured absorption spectra at 2 K of GaAs layers 400-, 21-, and 14-nm thick sandwiched between $Al_{0.2}Ga_{0.8}As$ barrier layers. Multiple absorption peaks correspond to transitions between different subbands arising from quantum-size effects. (After Ref. 8)

The lasing action usually occurs on the transition from E_{1c} to E_{1hh}, i.e., the transition between the lowest conduction-band subband and the highest heavy-hole subband (see Fig. 9.2). One important characteristic of a quantum-well structure is that the lasing wavelength can be easily varied by varying the thickness of the well. This was clearly demonstrated for an InGaAs-InP quantum-well double heterostructure. Figure 9.7 shows the experimental results for the optical spectra with three different well thicknesses obtained by optical pumping of the InGaAs-InP sample using a Nd-YAG laser ($\lambda = 1.06\ \mu m$). The wavelength shift of the main peak corresponds to the change in the lowest energy level of a quantum well with the well thickness.

9.4 RADIATIVE RECOMBINATION

The radiative recombination rate is obtained by calculating the two-dimensional analog of the results presented in Chapter 3 for bulk material.[12–15] A complete calculation of the radiative recombination rate would include

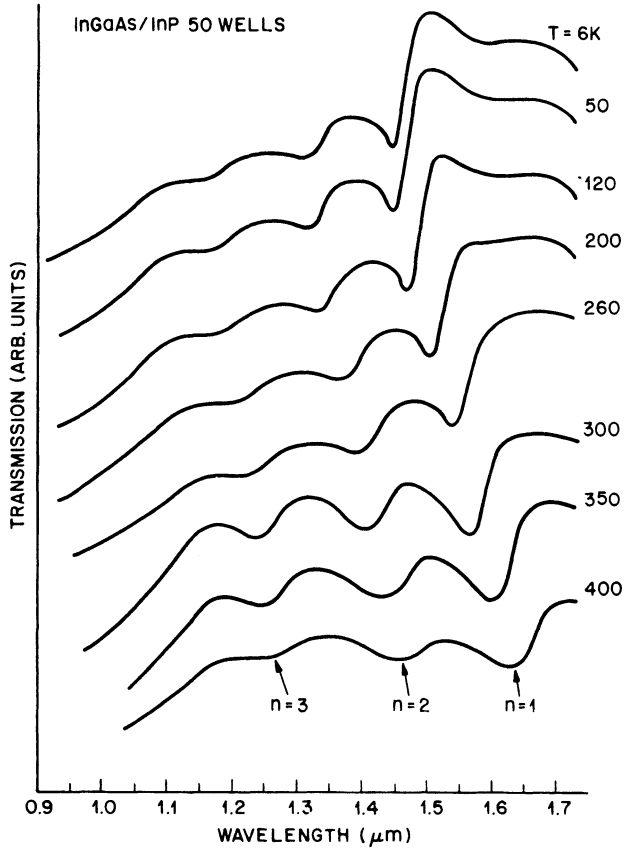


Fig. 9.6 Absorption spectra of a sample with 50 InGaAs wells approximately 12.5-nm thick at various temperatures. (After Ref. 11)

summation over all possible transitions between the various subbands in the conduction and valence bands. Before doing this calculation, we first consider a particular transition, namely, the transition from the lowest level in the conduction band (level 1*c*) to the lowest level in the heavy-hole band (level 1*hh*). This transition has the highest gain (due to Fermi factors) under external excitation (e.g., current injection) and hence is responsible for the lasing action in quantum-well injection lasers.

The spontaneous-emission rate $r_{sp}(E)$ and the absorption coefficient $\alpha(E)$ for electronic transitions between two discrete levels are given by Eqs. (3.2.20) and (3.2.21). As before (Sec. 3.2.2), since the photons carry negligible momentum, two cases are possible: (i) a matrix element obeying the **k**-selection rule in the x and y direction[13] and (ii) a constant matrix element[12] independent

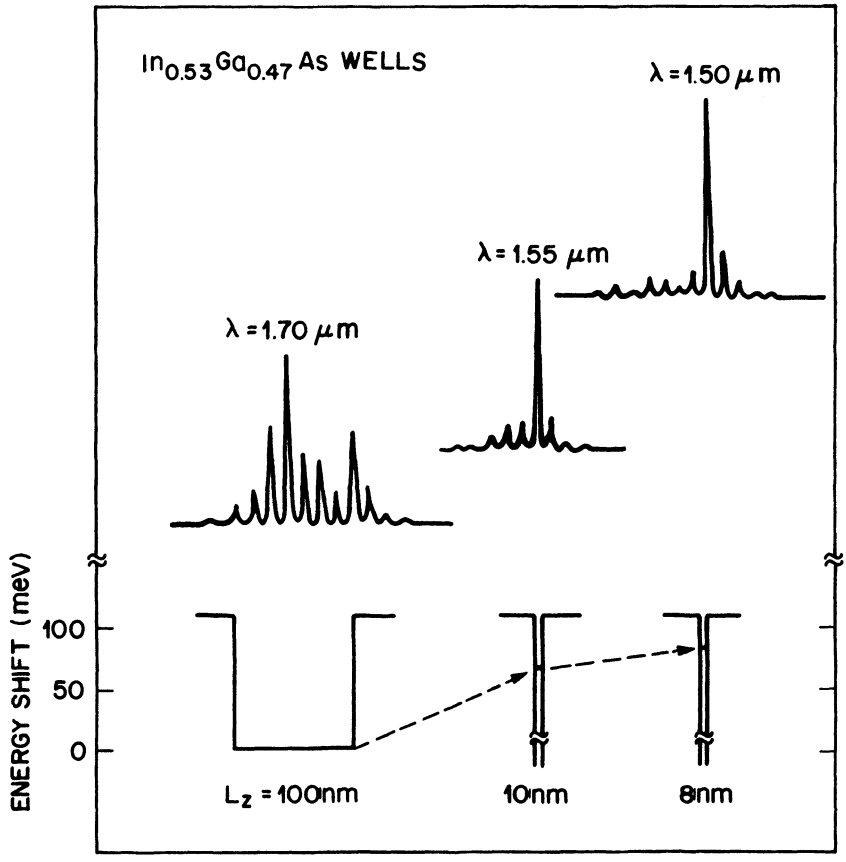


Fig. 9.7 Longitudinal-mode spectra for three well thicknesses of an InGaAs-InP multiquantum-well structure. The wavelength shift corresponds to the energy-level shift of the quantum well (dashed lines). (Courtesy of H. Temkin)

of the wave vector, which can arise from a band-to-impurity transition. We first consider the case of a matrix element obeying the **k**-selection rule.

Consider an area A of the material in the x-y plane so that the volume is given by $V = AL_z$. Using the two-dimensional analog of Eq. (3.2.34), the matrix element $|M_{if}|^2$ is given by

$$|M_{if}|^2 = |M_b|^2 \frac{(2\pi)^2}{A} \delta(\mathbf{k}_c - \mathbf{k}_v)\delta_{nn'}. \tag{9.4.1}$$

The last δ function arises from the selection rule for the confined states in the z direction. The quantity $|M_b|$ is an average matrix element for the

Bloch states of the bands (in the x and y directions). Using the Kane band model, $|M_b|^2$ in bulk semiconductors is given by [see Eq. (3.2.35)]

$$|M_b|^2 = \frac{m_0^2 E_g(E_g + \Delta)}{12m_c(E_g + 2\Delta/3)} = \xi m_0 E_g \tag{9.4.2}$$

where $\xi = 1.3$ for GaAs and lies between 1 and 2 for most semiconductors. We assume that $|M_b|^2$ for a quantum well is given by

$$|M_b|^2 = \xi m_0 E_q \tag{9.4.3}$$

where E_q is the separation between levels $1c$ and $1hh$. The uncertainty in the value of ξ represents the accuracy of our results. Using Eqs. (3.2.24) and (9.4.1) and summing over all the states in the two-dimensional bands, we get

$$r_{sp}(E) = \frac{4\pi\bar{\mu}q^2E}{m_0^2\epsilon_0 h^2c^3}|M_b|^2\frac{(2\pi)^2}{A}2\left(\frac{A}{(2\pi)^2}\right)^2\frac{1}{V}$$
$$\times \iint f_c(E_c)f_v(E_v)\,d^2\mathbf{k}_c\,d^2\mathbf{k}_v\,\delta(\mathbf{k}_c - \mathbf{k}_v)\delta(E_i - E_f - E) \tag{9.4.4}$$

where f_c and f_v denote the Fermi factors for the electron at energy E_c and the hole at energy E_v. The factor 2 arises from the two spin states.

We use the same convention for the electron and hole energy levels as in Chapter 3: the electron energy is measured from the conduction-band edge and is positive into the band, while hole energies are measured positive downward into the valence band from the valence-band edge. The integrals in Eq. (9.4.4) can be evaluated following the procedure of Sec. 3.2 with the following result:

$$r_{sp}(E) = \frac{16\pi^2\bar{\mu}q^2E|M_b|^2m_r}{m_0^2\epsilon_0 h^4c^3L_z}f_c(E_c)f_v(E_v) \tag{9.4.5}$$

with

$$E_c = \frac{m_r}{m_c}(E - E_q) \qquad E_v = \frac{m_r}{m_v}(E - E_q)$$

and

$$m_r = \frac{m_c m_v}{m_c + m_v}.$$

Integrating over all energies E such that $E \geq E_q$, the total spontaneous-emission rate per unit volume is given by

$$R = \frac{16\pi^2 \bar{\mu} q^2 m_r |M_b|^2}{m_0^2 \epsilon_0 h^4 c^3 L_z} I \tag{9.4.6}$$

where

$$I = \int_{E_q}^{\infty} E f_c(E_c) f_v(E_v)\, dE.$$

Following a similar procedure, the expression for the absorption coefficient $\alpha(E)$ [after using Eq. (3.2.21) and summing over the available states] is given by

$$\begin{aligned}\alpha(E) &= \frac{q^2 h}{2\epsilon_0 m_0^2 c \bar{\mu} E} |M_b|^2 \frac{(2\pi)^2 2}{A} \left(\frac{A}{(2\pi)^2}\right)^2 \left(\frac{1}{V}\right) \\ &\quad \times \iint (1 - f_c - f_v)\, d^2\mathbf{k}_c\, d^2\mathbf{k}_v\, \delta(\mathbf{k}_c - \mathbf{k}_v) \delta(E_i - E_f - E) \\ &= \frac{q^2 m_r |M_b|^2}{\epsilon_0 m_0^2 c \hbar \bar{\mu} E L_z} [1 - f_c(E_c) - f_v(E_v)].\end{aligned} \tag{9.4.7}$$

The radiative component of the current or the nominal current density J_n at threshold is given by

$$J_r = J_n = qL_z R. \tag{9.4.8}$$

Constant Matrix Element. In this case, we assume that $|M_b|^2$ in Eq. (9.4.1) is given by

$$|M_b|^2 = |M_c|^2 \delta_{nn'} \tag{9.4.9}$$

where $|M_c|^2$ is independent of the energy and the wave vector of the particles involved. One can now follow the same procedure outlined before and obtain the following expressions for the absorption coefficient and the spontaneous-emission rate:

$$\alpha(E) = \frac{4\pi q^2 m_c m_v |M_c|^2}{\epsilon_0 c m_0^2 \hbar^3 \bar{\mu} E L_z} I_1(E) \tag{9.4.10}$$

$$R = \frac{4\bar{\mu} q^2 m_c m_v}{\pi \epsilon_0 c^3 \hbar^6 m_0^2 L_z} |M_c|^2 \int_{E_q}^{\infty} E I_2(E)\, dE \tag{9.4.11}$$

where

$$I_1(E) = \int_0^{E-E_q} [1 - f_c(E_c) - f_v(E_v)]\, dE_c \tag{9.4.12a}$$

$$I_2(E) = \int_0^{E-E_q} f_c(E_c) f_v(E_v)\, dE_c. \tag{9.4.12b}$$

Optical Gain. Using Eqs. (9.4.6)–(9.4.8), one can obtain the gain as a function of the nominal current density and as a function of the injected carrier density. We first consider the GaAs quantum well and show that the results obtained using the **k**-selection rule agree well with the experimental results for the threshold current density. For longer-wavelength semiconductor lasers (e.g., InGaAsP), additional nonradiative mechanisms such as the Auger effect may be important in determining the threshold current.

Figure 9.8 shows the calculated maximum gain as a function of the injected carrier density and as a function of the nominal current density at various temperatures for a 20-nm-thick undoped GaAs quantum well. The parameters used in the calculation were as follows: $m_c = 0.071m_0$, $m_{hh} = 0.45m_0$, $m_{lh} = 0.081m_0$, $\bar{\mu} = 3.5$, and $\xi = 1.33$. For a 20-nm-thick well, the calculated $E_q - E_g \cong 0.011$ eV. The calculation assumes that the band gap E_g varies with temperature at a rate $dE_g/dT = -0.35$ meV/K. Figure 9.8 shows that the gain increases rapidly with the carrier density n initially and then starts to saturate, leading to a slow increase. Near room temperature the crossover

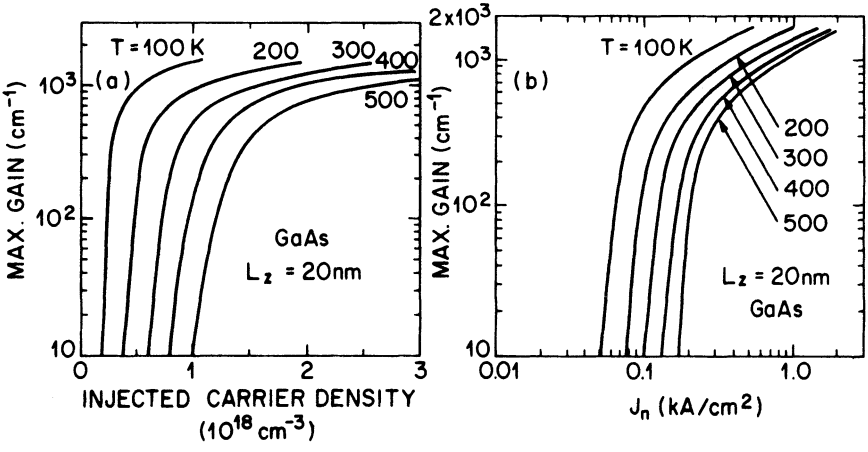


Fig. 9.8 Calculated maximum gain for a GaAs quantum-well structure at various temperatures as a function of (a) the injected carrier density and (b) the nominal current density. (After Ref. 13)

occurs when $n \cong 1 \times 10^{18}\,\text{cm}^{-3}$. The physical reason behind the gain saturation is related to the constant density of states of the lowest subband of the quantum well.

We now calculate the threshold current density and compare the calculated value with the measured results. The threshold gain g_{th} of a laser is given by [see Eqs. (2.3.9) and (2.3.10)]

$$\Gamma g_{\text{th}} = \alpha_{\text{int}} + \frac{1}{L}\ln\left(\frac{1}{R}\right). \tag{9.4.13}$$

The calculated confinement factor Γ for a 25-nm-thick GaAs-$Al_{0.52}Ga_{0.48}As$ quantum well is 0.04. When $\alpha_{\text{int}} = 10\,\text{cm}^{-1}$, $R = 0.3$, and $L = 380\,\mu\text{m}$, we find from Eq. (9.4.13) that $g_{\text{th}} \cong 1{,}000\,\text{cm}^{-1}$. Thus the threshold gain for single quantum-well lasers is an order of magnitude larger than that for regular double-heterostructure lasers. This difference is principally due to a smaller confinement factor for a single quantum well. For the calculated g_{th} of $1{,}000\,\text{cm}^{-1}$, the estimated nominal current density J_n using Fig. 9.8 is $550\,\text{A/cm}^2$. This compares well with the reported threshold current density of $810\,\text{A/cm}^2$ for a 20-nm-wide GaAs quantum-well laser.[16] Equation (9.4.13) shows that the small value of the confinement factor makes the

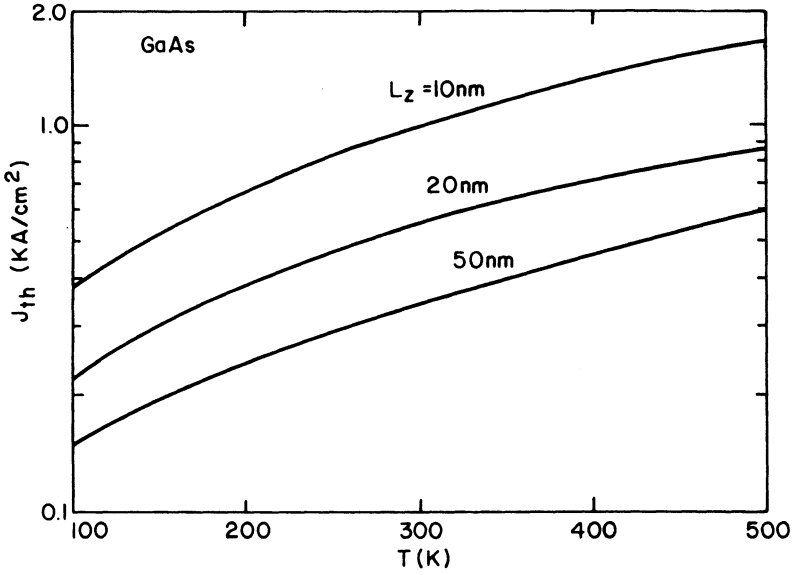


Fig. 9.9 Calculated temperature dependence of the threshold current of a GaAs single quantum-well laser. Three curves for different well thicknesses are shown. (After Ref. 13)

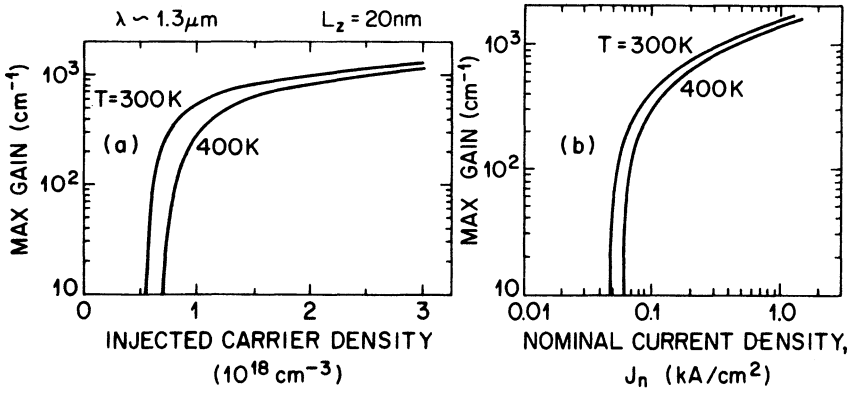


Fig. 9.10 Calculated maximum gain for a 1.3-μm InGaAsP-InP quantum-well laser with a well thickness L_z of 20 nm, as a function of (a) the injected carrier density and (b) the nominal current density. Gain variation at two different temperatures is shown. (After Ref. 19)

threshold gain, and hence the threshold current density, of quantum-well lasers very sensitive to internal losses, which are influenced by defects, nonradiative regions, and unpumped active regions in an injection laser.

It is generally observed that the threshold current of GaAs quantum-well lasers does not increase as rapidly with increasing temperature as does that for regular double-heterostructure lasers ($T_0 \cong 250$–400 K when $T > 270$ K for quantum-well lasers, while $T_0 \cong 200$ K for regular double-heterostructure lasers). Figure 9.9 shows the calculated threshold current as a function of temperature for GaAs-$Al_{0.52}Ga_{0.48}As$ quantum wells of different active-region thicknesses. The calculated T_0 value from Fig. 9.9 is ~ 300 K when $300 \text{ K} < T < 400$ K. The reason for the high T_0 value is a weaker dependence of quasi-Fermi energies on temperature in this essentially two-dimensional system than in bulk semiconductors. Since the condition for gain (at energy E) is given by $E_{fc} + E_{fv} > E - E_g$, fewer additional carriers at threshold are needed at a higher temperature for a quantum-well laser than for a regular double-heterostructure laser.

Figure 9.10 shows the calculated gain as functions of the injected carrier density and the nominal current density at various temperatures for a 1.3-μm InGaAsP-InP quantum-well structure. The parameter values used in the calculation are $m_c = 0.061m_0$, $m_v = 0.45m_0$, and $E_q = 1$ eV. Two curves for a 20-nm-thick well are shown at 300 K and 400 K. Similar to the case of GaAs (Fig. 9.8), the gain saturates at high carrier densities because of the constant density of states of the lowest subband.

Sugimura[14] has calculated the radiative recombination rate in quantum-well structures by taking into account transitions between all subbands (in

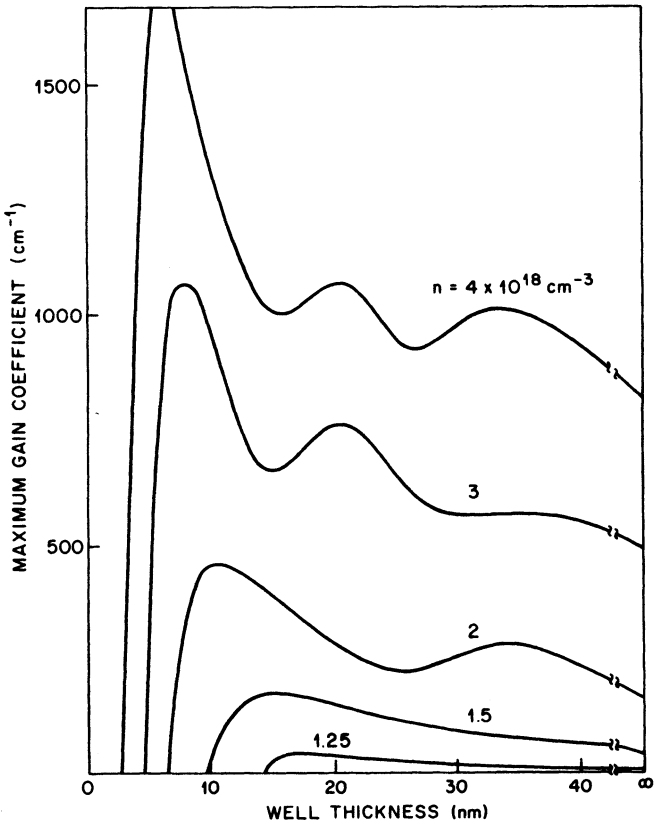


Fig. 9.11 Maximum gain for an undoped InGaAsP-InP ($\lambda \cong 1.07\ \mu m$) single quantum-well laser as a function of the well thickness for several values of injected carrier density. (After Ref. 20 © 1983 IEEE)

the Γ valley) as well as the effect of filling the L valleys. Figure 9.11 shows the dependence of the gain on the well thickness L_z for different injected carrier densities. The undulations in the value of the gain arise from recombinations between electrons in different subbands of the conduction band and the heavy or light holes in the valence band. The steep decrease in the gain coefficient for small well thicknesses arises from the distribution of carriers in the L valley, which reduces the number of carriers in the Γ valley that contribute to optical gain. Holonyak and coworkers[2,3] have postulated the occurrence of phonon-assisted stimulated emission in quantum-well structures in order to explain their emission spectrum. However, the

calculated phonon-assisted gain obtained using LO (longitudinal optical) phonon interactions is smaller than the gain obtained for direct transitions.[17]

9.5 AUGER RECOMBINATION

Auger recombination is an important nonradiative recombination mechanism for low band-gap materials. In Sec. 3.3 we discussed the importance of Auger recombination in determining the threshold current of long-wavelength InGaAsP lasers with relatively thick active regions. The observed high temperature sensitivity of the threshold current of long-wavelength regular double-heterostructure lasers is generally attributed to Auger recombination. In quantum-well structures, the modification of the density of states may reduce the Auger recombination rate significantly to allow long-wavelength quantum-well lasers to have a relatively low temperature sensitivity. We now outline the Auger-rate calculation for quantum-well structures.[18–21]

Figure 9.12 shows the various band-to-band Auger processes. These processes are labeled CCCH, CHHS, and CHHL in analogy with the corresponding processes in bulk semiconductors (see Sec. 3.3). The energy-versus-wavevector diagram is shown for one subband (one state in the z direction) in each of the bands. The dashed curve (for CCCH) shows a second subband in the conduction band. The dashed line shows a process in which the electron makes a transition to a different subband. CCCH involves two electrons and a hole and is dominant in n-type material; CHHS and CHHL are dominant in p-type material. Under high-injection conditions, which are present in semiconductor lasers, all three mechanisms must be considered.

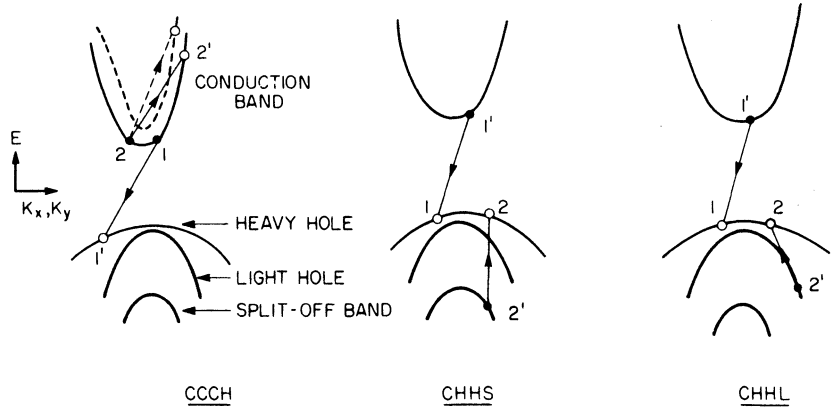


Fig. 9.12 Band-to-band Auger recombination processes shown schematically for a quantum-well laser. Dashed curve shows the possibility of inter-subband scattering.

In this section we derive expressions for the transition probability of these band-to-band Auger processes.

CCCH Process. In Sec. 3.3 we derived an expression for the CCCH Auger rate for bulk semiconductors. The corresponding expression for the quantum-well structure is

$$R_a = \frac{2\pi}{\hbar}\left(\frac{1}{(2\pi)^2}\right)^4 \frac{1}{L_z} \times \iiiint |M_{if}|^2 P(1, 1', 2, 2')\delta(E_i - E_f)\, d^2\mathbf{k}_1\, d^2\mathbf{k}_2\, d^2\mathbf{k}'_1\, d^2\mathbf{k}'_2 \quad (9.5.1)$$

where R_a denotes the Auger rate per unit volume. The quantity $|M_{if}|^2$ denotes the square of the matrix element and $P(1, 1', 2, 2')$ is the difference between the probability of the Auger process shown in Fig. 9.12 and the probability of the inverse process of impact ionization. The factor $1/L_z$ arises from converting the Auger rate per unit area (in two dimensions) to a per unit volume. Equation (9.5.1) is a conversion of Eq. (3.3.1) to two dimensions.

Using the same approximation used in Sec. 3.3, we get $|M_{if}|^2 \cong 4|\bar{M}_{if}|^2$. The quantity $\bar{M}_{if}$ is the matrix element for Coulomb interaction between two electrons and is given by

$$\bar{M}_{if} = \iint \psi'^*_1(\mathbf{r}_1)\psi'^*_2(\mathbf{r}_2)\frac{q^2}{\epsilon|\mathbf{r}_1 - \mathbf{r}_2|}\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\, d^3\mathbf{r}_1\, d^3\mathbf{r}_2 \quad (9.5.2)$$

where $\psi_m(\mathbf{r})$ are the quantum-well wave functions. Free carrier screening has been neglected in Eq. (9.5.2). To evaluate $\bar{M}_{if}$, we use the Fourier transform of the Coulomb interaction:

$$\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} = \frac{1}{(2\pi)^3}\int \frac{4\pi}{\xi^2}\exp[i\boldsymbol{\xi}\cdot(\mathbf{r}_1 - \mathbf{r}_2)]\, d^3\xi \quad (9.5.3)$$

and obtain

$$\bar{M}_{if} = \frac{4\pi q^2}{\epsilon}\frac{1}{(2\pi)^3}\int \frac{1}{\xi^2} I_{1'1}(\xi)I_{2'2}(-\xi)\, d^3\xi \quad (9.5.4)$$

where

$$I_{mn}(\xi) = \int \psi^*_m(\mathbf{r})\psi_n(\mathbf{r})\exp(i\boldsymbol{\xi}\cdot\boldsymbol{\rho})\, d^3\mathbf{r}. \quad (9.5.5)$$

The particle states in a quantum well are represented by the product of a localized Bloch function, a plane-wave term, and a confined-state wave function for the z direction; i.e.,

$$\psi_m(\mathbf{r}) = U_m(\mathbf{r})\phi_m(z)\exp(-i\mathbf{k}_m\cdot\boldsymbol{\rho}) \tag{9.5.6}$$

where $U_m(\mathbf{r})$ is the localized part, $\phi_m(z)$ is the confined-state wave function, $\mathbf{k}_m$ is the two-dimensional wave vector in the plane of the well, and $\boldsymbol{\rho} = (x, y)$ is the two-dimensional position vector. For an infinite well we can use Eq. (9.1.3b) for ψ_n in Eq. (9.5.5) and evaluate $I_{mn}(\boldsymbol{\xi})$. The result is

$$I_{mn}(\boldsymbol{\xi}) = (2\pi)^2\delta(-\mathbf{k}_m + \mathbf{k}_n + \boldsymbol{\xi}_T)F_{mn}G_{mn}(\xi_z) \tag{9.5.7}$$

where

$$\boldsymbol{\xi} = \boldsymbol{\xi}_T + \xi_z\mathbf{z},$$

$$F_{mn} = \int U_m^*(\mathbf{r})U_n(\mathbf{r})\,d^3\mathbf{r}$$

$$G_{mn} = \int \psi_m^*(z)\psi_n(z)\exp(i\xi_z z)\,dz.$$

$\boldsymbol{\xi}_T$ is the component of $\boldsymbol{\xi}$ parallel to the well and the integral F_{mn} is carried out over a unit cell. The matrix element $\bar{M}_{\text{if}}$ is now given by

$$\bar{M}_{\text{if}} = \frac{2\pi q^2}{\epsilon} V\,\delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_1' - \mathbf{k}_2') \tag{9.5.8}$$

where

$$V = 4\pi F_{11'}F_{22'}\int \frac{G_{11'}G_{22'}}{|\mathbf{k}_1 - \mathbf{k}_1'|^2 + \xi_z^2}\,d\xi_z. \tag{9.5.9}$$

In deriving the above, we have used the relation

$$\xi^2 = \xi_T^2 + \xi_z^2 = |\mathbf{k}_1 - \mathbf{k}_1'|^2 + \xi_z^2.$$

Using Eqs. (9.5.1) and (9.5.8), the Auger rate is given by

$$R_a = \frac{8\pi}{\hbar}\left(\frac{1}{4\pi^2}\right)^4\frac{1}{L_z}\left(\frac{2\pi q^2}{\epsilon}\right)^2 I \tag{9.5.10}$$

where

$$I = \iiiint P(1, 1', 2, 2')\, d^2\mathbf{k}_1\, d^2\mathbf{k}'_1\, d^2\mathbf{k}_2\, d^2\mathbf{k}'_2 |V(\mathbf{k}_1, \mathbf{k}'_1)|^2 \times \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}'_1 - \mathbf{k}'_2)\delta(E_i - E_f). \tag{9.5.11}$$

From the requirement of conservation of energy, it follows that

$$\frac{\hbar^2}{2m_c} k_1^2 + \frac{\hbar^2}{2m_c} k_2^2 = \frac{\hbar^2}{2m_v} k_1'^2 - E_q + \frac{\hbar^2}{2m_{ct}} k_2'^2$$

where m_c and m_v are the electron and hole effective masses at the band edge and m_{ct} is the electron effective mass at energy E'_2.

We now evaluate the integral I by following the procedure of Sec. 3.3. Defining the new integration variables

$$\mathbf{h} = \mathbf{k}_1 + \mathbf{k}_2 \qquad \mathbf{j} = \mathbf{k}_1 - \mathbf{k}_2$$

and integrating over them, we obtain

$$I = \frac{m_c}{2\hbar^2} \iint d^2\mathbf{k}'_1\, d^2\mathbf{k}'_2 \int_{-1}^{1} dt\, P(1, 1', 2, 2') |V(\mathbf{k}_1, \mathbf{k}'_1)|^2 |\mathbf{j}_0| \tag{9.5.12}$$

where $t = \cos\theta$ and θ is the angle between the vectors $\mathbf{h}$ and $\mathbf{j}$. The integral is to be evaluated under the condition $j_0^2 \geq 0$, where

$$j_0^2 = 2\left(\mu_c k_2'^2 - \mu k_1'^2 - \frac{2m_c E_q}{\hbar^2}\right) - |\mathbf{k}'_1 + \mathbf{k}'_2|^2$$

$$\mu_c = \frac{m_c}{m_{ct}} \qquad \mu = \frac{m_c}{m_v}.$$

If we define two new variables

$$\mathbf{z}_1 = \mathbf{k}'_1 + \frac{\mathbf{k}'_2}{1 + 2\mu} \qquad \mathbf{z}_2 = \mathbf{k}'_2 \tag{9.5.13}$$

j_0^2 becomes

$$j_0^2 = 2\alpha_c z_2^2 - z_1^2(1 + 2\mu) - \frac{4m_c}{\hbar^2} E_q \tag{9.5.14}$$

where

$$\alpha_c = \mu_c - \frac{\mu}{1 + 2\mu}. \tag{9.5.15}$$

A necessary condition for $j_0^2 \geq 0$ is

$$\alpha_c z_2^2 > \frac{2m_c E_q}{\hbar^2}$$

or, using $E_2' = \hbar^2 z_2^2 / 2m_c$,

$$E_2' > E_T = \mu_c E_q / \alpha_c. \tag{9.5.16}$$

Thus, similar to the three-dimensional case E_2' must reach a threshold energy ($E_T = \mu_c E_q / \alpha_c$) before the Auger process will take place.

The quantity $P(1, 1', 2, 2')$ is the difference between the product of the occupation probabilities of the four states for the Auger process and those for the inverse process of impact ionization. When $E_q \gg k_B T$, the inverse process is relatively unlikely to occur and can be ignored. If we further make the nondegenerate approximation [Eq. (9.2.9)], $P(1, 1', 2, 2')$ is given by

$$P(1, 1', 2, 2') = \frac{n^2 p}{N_c^2 N_v} \exp\left(\frac{-E_2' + E_q}{k_B T}\right) \tag{9.5.17}$$

where n and p are the electron and hole concentrations, respectively, and we have used the energy conservation relation

$$E_1 + E_2 = E_2' - E_q - E_1'. \tag{9.5.18}$$

Since the exponents in Eq. (9.5.17) decreases rapidly as E_2' increases, we can approximate $|V|^2$ in Eq. (9.5.12) by its value at $E_2' = E_T$. Using Eqs. (9.5.13) and (9.5.17) in Eq. (9.5.12), we finally obtain

$$I = \frac{m_c}{\hbar^2} \frac{n^2 p}{N_c^2 N_v} |V_T|^2 \exp\left(\frac{E_q}{k_B T}\right) I_1 \tag{9.5.19}$$

where

$$\begin{aligned} I_1 &= \iint d^2 z_1 \, d^2 z_2 \exp\left(-\frac{\hbar^2 z_2^2}{2m_{ct} k_B T}\right) \\ &= \frac{8\pi^2 \alpha_c}{1 + 2\mu} \left(\frac{2m_{ct} k_B T}{\hbar^2}\right)^2 \exp\left(-\frac{m_c}{m_{ct} \alpha_c} \frac{E_q}{k_B T}\right) \end{aligned} \tag{9.5.20}$$

and $|V_T|^2$ is the value of $|V|^2$ at $E_2' = E_T$. The integral I_1 has been evaluated under the condition $j_0^2 \geq 0$, where j_0 is given by Eq. (9.5.14) and m_{ct} is the electron effective mass at the energy E_T. Using Eqs. (9.5.10) and (9.5.19) we get in the nondegenerate approximation

$$R_a \propto \frac{n^2 p}{N_c^2 N_v} \exp\left(-\frac{\Delta E_c}{k_B T}\right) \tag{9.5.21}$$

where

$$\Delta E_c = E_T - E_q = \frac{m_{ct} E_q}{m_v + 2m_c - m_{ct}}. \tag{9.5.22}$$

Equations (9.5.21) and (9.5.22) show that the Auger rate increases with decreasing band gap and with increasing temperature. For a degenerate electron or hole gas, the integral I in Eq. (9.5.10) can be evaluated using the method of Ref. 21.

CHHS Process. The CHHS Auger rate can be calculated using the same procedure as for CCCH. The expression for the Auger rate per unit volume is still given by Eq. (9.5.10). However, because of different particle states participating in the process, the integral I given by Eq. (9.5.11) is different. Following the analysis of Sec. 3.3 for the CHHS process in bulk semiconductors, it can be shown that the threshold energy for the process is given by [see Eq. (3.3.38)]

$$E_{Ts} = \frac{\hbar^2 k_{2t}'^2}{2m_s} = \frac{2m_v + m_c}{2m_v + m_c - m_s}(E_q - \Delta_1) \tag{9.5.23}$$

where m_s is the mass of the split-off band at energy E_{Ts} and Δ_1 is the energy separation of the split-off band from the top of the heavy-hole band. This minimum energy requirement arises from the conservation laws of energy and momentum that the four particle states involved must satisfy. In the nondegenerate approximation, the integral I can be evaluated in a closed form after using Eq. (9.5.17) in Eq. (9.5.11) and is given by

$$I = \frac{m_c}{\hbar^2} \frac{np^2}{N_c N_v^2} |V_T|^2 \exp\left(\frac{E_q - \Delta_1}{k_B T}\right) I_1 \tag{9.5.24}$$

where

$$I_1 = \frac{8\pi^2 \alpha_s}{(1+\mu/2)} \left(\frac{2m_s k_B T}{\hbar^2}\right)^2 \exp\left[-\frac{m_c}{\alpha_s m_s}(E_q - \Delta_1)\right] \tag{9.5.25}$$

$$\alpha_s = \mu_s - 1 + \frac{1}{1+\mu/2} \tag{9.5.26}$$

$$\mu_s = m_c/m_s \qquad \mu = m_c/m_v. \tag{9.5.27}$$

Using Eqs. (9.5.10) and (9.5.24), the Auger rate is found to vary as (in the nondegenerate approximation)

$$R_a \propto p^2 n \exp\left(-\frac{\Delta E_s}{k_B T}\right) \tag{9.5.28}$$

where

$$\Delta E_s = E_{Ts} - E_q = \frac{m_s(E_q - \Delta_1)}{m_c + 2m_v - m_s}. \tag{9.5.29}$$

This formulation is valid for semiconductors where $E_q > \Delta_1$. Similar to the case of CCCH, the CHHS Auger rate increases with increasing temperature and as the term $E_q - \Delta_1$ decreases.

CHHL Process. The calculation for this process proceeds in an identical fashion as for CHHS. The results given there can be directly applied by setting Δ_1 to 0 and replacing m_s by m_l where m_l is the light-hole mass. The Auger recombination rate per unit volume is given by Eq. (9.5.10), while the threshold energy for the process is [see Eq. (9.5.23)]

$$E_{Tl} = \frac{\hbar^2 k_{2t}'^2}{2m_l} = \frac{2m_v + m_c}{2m_v + m_c - m_l} E_q. \tag{9.5.30}$$

In the nondegenerate approximation I is given by Eqs. (9.5.24) and (9.5.25) with m_s, α_s, and $E_q - \Delta_1$ replaced by m_l, α_l, and E_q, respectively. The Auger rate varies as

$$R_a \propto p^2 n \exp\left(-\frac{\Delta E_l}{k_B T}\right) \tag{9.5.31}$$

where

$$\Delta E_l = \frac{m_l E_q}{m_c + 2m_v - m_l}. \tag{9.5.32}$$

Auger Lifetime. The expressions for the three Auger processes derived here show that the Auger rate increases with decreasing band gap and with increasing temperature. They can be used to estimate the importance of Auger recombination for quantum-well lasers by calculating the Auger lifetime $\tau_A = n/R_a$.

For CCCH the calculated Auger lifetime depends strongly on the value of ΔE_c given by Eq. (9.5.22). For bulk semiconductors, m_{ct} (the conduction band mass at the threshold energy E_T) is obtained using the Kane model, which gives $m_{ct} \cong 1.7m_c$. For a quantum-well laser, the ratio m_{ct}/m_c is not known. For parabolic bands, $m_{ct}/m_c = 1$ and Eq. (9.5.22) reduces to

$$\Delta E_c = \mu E_q/(1 + \mu) \tag{9.5.33}$$

where $\mu = m_c/m_v$. The calculated Auger rate for a given band gap can decrease by more than one order of magnitude if the ratio m_{ct}/m_c increases from 1 to 2.[19]

Under high-injection conditions commonly prevailing in semiconductor lasers, it is often necessary to go beyond the nondegenerate approximation. The Auger rate is then calculated numerically using Eqs. (9.5.10) and (9.5.11). Figure 9.13 shows[20] the calculated Auger lifetime τ_A for a 1.07-μm InGaAsP quantum-well laser as a function of the well thickness L_z for several carrier densities n. The active region is assumed to be undoped. The undulations in the vicinity of $L_z \sim 10$ nm are due to the filling of the higher-order subbands with increasing n. One can estimate the value of the Auger coefficient C that is defined using

$$R_a = n/\tau_A = Cn^3. \tag{9.5.34}$$

If $L_z = 30$ nm, $n = 2 \times 10^{18}$ cm^{-3} (typical values), $C \cong 2.5 \times 10^{-29}$ cm^6/s from Fig. 9.13. However, it should be stressed that because of the uncertainties of both the matrix elements and the effective masses away from the band edge (the ratio m_{ct}/m_c), only order-of-magnitude estimates of the Auger rates can be obtained.

9.6 SINGLE QUANTUM-WELL AND MULTIQUANTUM-WELL LASERS

Quantum-well injection lasers with both single and multiple active layers have been fabricated. Quantum-well lasers with one active region are called *single quantum-well* (SQW) lasers, and those with multiple active regions are called *multiquantum-well* (MQW) lasers. The layers separating the active layers in an MQW structure are called barrier layers. Figure 9.14 schematically shows the energy-band diagram of these laser structures. The MQW lasers

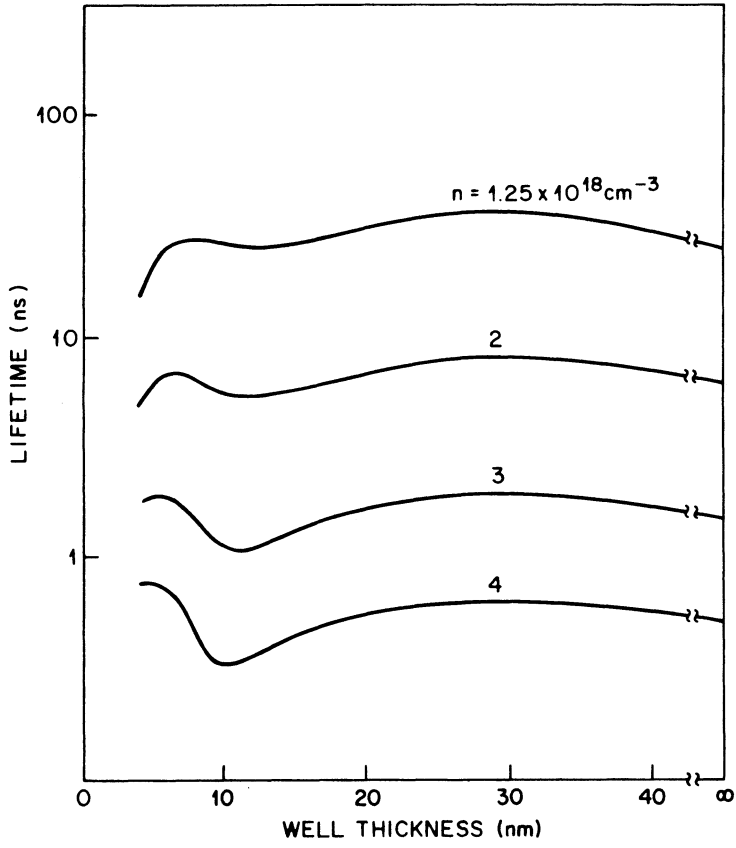


Fig. 9.13 Calculated Auger lifetime for CHHS in an undoped InGaAsP-InP quantum-well laser as a function of the well thickness for several values of the injected carrier density. (After Ref. 20 © 1983 IEEE)

in which the band gap of the barrier layers is different from that of the cladding layers are sometimes referred to as *modified multiquantum-well* lasers.

One of the main differences between the SQW and MQW lasers is that the confinement factor Γ of the optical mode is significantly smaller for the former than that for the latter. This can result in a higher threshold carrier density and a higher threshold current density for SQW lasers when compared with MQW lasers. The confinement factor of an SQW heterostructure can be significantly increased using a graded-index (GRIN) cladding layer (see Fig. 9.14). This allows use of the intrinsic advantage of the quantum-well structure (high gain at low carrier density) without the penalty of a

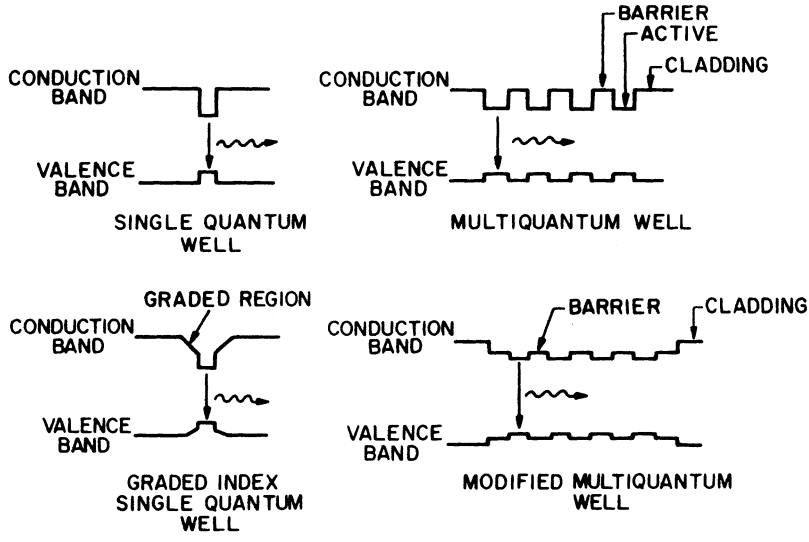


Fig. 9.14 Different single quantum-well and multiquantum-well laser structures shown schematically.

small mode-confinement factor.[22] Threshold current densities as low as 200 A/cm^2 have been reported for a GaAs GRIN SQW laser.[5]

We now discuss the mode-confinement factor in SQW and MQW structures. The confinement factor of the fundamental mode of a double heterostructure was calculated in Sec. 2.5 and is given approximately by Eq. (2.5.32). For a small active-layer thickness, such as that of a single quantum well, the expression for Γ can be further simplified[23] since the normalized waveguide thickness $D \ll 1$. Using Eqs. (2.5.26) and (2.5.32), we then obtain

$$\Gamma \cong 2\pi^2(\mu_a^2 - \mu_c^2)d^2/\lambda_0^2 \tag{9.6.1}$$

where μ_a and μ_c are the refractive indices of the active and cladding layers, respectively; d is the active-layer thickness; and λ_0 is the free-space wavelength. Figure 9.15 shows the calculated Γ obtained using Eq. (9.6.1) for InGaAsP-InP SQW structures emitting at different wavelengths.

The mode confinement in MQW structures can be analyzed using the same procedure as described in Sec. 2.5, except that the electromagnetic wave equation must be solved separately, with appropriate boundary conditions, for each of the active, barrier, and cladding layers. This procedure is quite tedious because of a large number of layers involved. The following simple formula is found[24] to give reasonably accurate results:

$$\Gamma(\text{MQW}) = \gamma \frac{N_a d_a}{N_a d_a + N_b d_b} \tag{9.6.2}$$

where

$$\gamma = 2\pi^2 (N_a d_a + N_b d_b)^2 \frac{\bar{\mu}^2 - \mu_c^2}{\lambda_0^2} \tag{9.6.3}$$

$$\bar{\mu} = \frac{N_a d_a \mu_a + N_b d_b \mu_b}{N_a d_a + N_b d_b}. \tag{9.6.4}$$

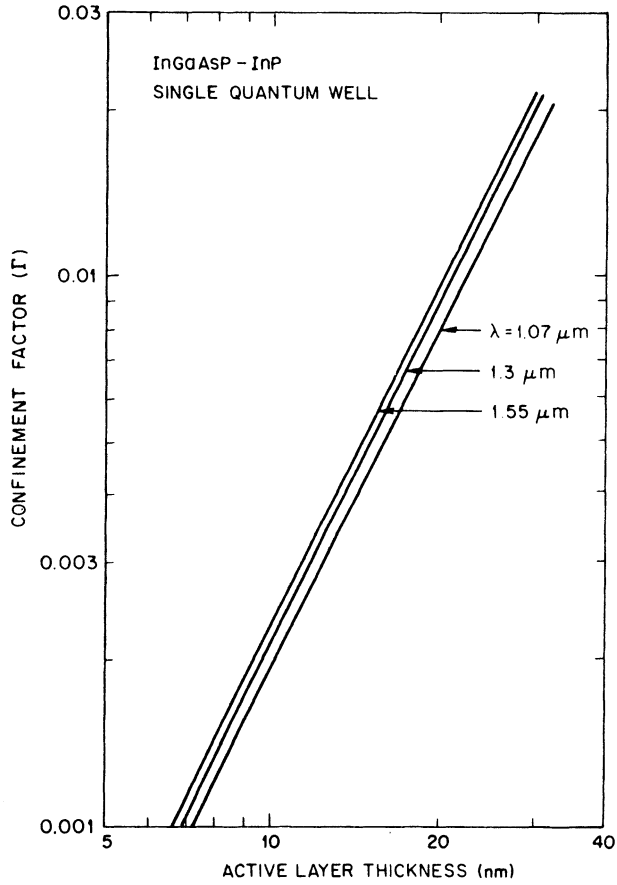


Fig. 9.15 Calculated confinement factor of the fundamental mode as a function of the active-layer thickness for an InGaAsP-InP single quantum-well heterostructure. The refractive indices of the active and cladding layers used in the calculation are: (i) $\lambda = 1.07$ μm, $\mu_a = 3.47$, $\mu_c = 3.27$; (ii) $\lambda = 1.3$ μm, $\mu_a = 3.51$, $\mu_c = 3.22$; and (iii) $\lambda = 1.55$ μm, $\mu_a = 3.54$, $\mu_c = 3.18$.

N_a and N_b are the number of active and barrier layers in the MQW structure, and d_a and d_b (and μ_a and μ_b) are the thicknesses (and refractive indices) of the active and barrier layers, respectively. The form of Eq. (9.6.2) may be understood as follows: $\bar{\mu}$ is the average refractive index of an active region (consisting of all the active and barrier layers) of total thickness $N_a d_a + N_b d_b$, and γ is the confinement factor of the optical mode [same as in Eq. (9.6.1)] in this effective layer of index $\bar{\mu}$. The MQW mode-confinement factor Γ is thus obtained by multiplying γ by the ratio of the total active thickness $N_a d_a$ to the total thickness $N_a d_a + N_b d_b$. For a 1.3-μm InGaAsP-InP multiquantum-well structure with four active-layer wells (each 15 nm thick) and three barrier layers (InP, each 15 nm thick), the calculated $\Gamma \cong 0.2$. This is considerably larger than the typical Γ values for SQW lasers (shown in Fig. 9.15) with active-layer thicknesses in the range of 10–15 nm.

Let us now compare single quantum-well and multiquantum-well lasers for the effect of Auger recombination on the threshold current. The gain at threshold (g_{th}) is given by Eq. (2.3.9), which after using Eqs. (2.3.10) and (2.5.50) becomes

$$\Gamma g_{th} = \Gamma \alpha_a + (1 - \Gamma)\alpha_c + \frac{1}{L} \ln\left(\frac{1}{R}\right). \tag{9.6.5}$$

α_a and α_c are the active and cladding layer losses, L is the length of the laser, and R is the reflectivity of the end mirrors. For an SQW laser, the confinement factor Γ is small and hence the threshold gain is very large ($g_{th} \gtrsim 10^3\ \text{cm}^{-1}$). This makes the carrier density at threshold large: $n_{th} \cong 2.5 \times 10^{18}\ \text{cm}^{-3}$ (see Fig. 9.10). For an MQW laser, the effective confinement factor is large and g_{th} is smaller ($g_{th} \cong 300\ \text{cm}^{-1}$). The carrier density at threshold $n_{th} \cong 8 \times 10^{17}\ \text{cm}^{-3}$ at this gain (see Fig. 9.10). The Auger rate varies as n^3. Thus it follows that the Auger rate in MQW lasers will be significantly smaller (by a factor of 27 in our example) when compared with the Auger rate in SQW lasers. Also, for the same reason, any nonradiative mechanism that depends superlinearly on the carrier density will have a smaller effect in MQW lasers than in SQW lasers.

We next consider the effect of increased optical losses with increasing temperature (e.g., intervalence-band absorption). For SQW lasers, Γ is small and hence the first term on the right-hand side of Eq. (9.6.5) is negligible. Thus the effect of increased intervalence-band absorption in determining the temperature dependence of threshold is smaller in SQW lasers than in MQW lasers. Another point to note is that for $g_{th} \gtrsim 10^3\ \text{cm}^{-1}$ (near the threshold-gain region of SQW lasers) the gain increases slowly with increasing current (see Fig. 9.10). This will make the threshold current of SQW lasers more

sensitive to increased cladding layer losses, which may increase with increasing temperature because free-carrier absorption is larger at higher temperatures.

9.7 MQW LASER RESULTS

A consider amount of experimental work has been reported on quantum-well lasers fabricated using the AlGaAs material.[22–33] The epitaxial layers are grown using molecular beam epitaxy (MBE) or metal-organic chemical vapor deposition (MOVPE) (see Chapter 4). The performance of AlGaAs lasers grown by MBE technique has been reviewed.[25] The results show that AlGaAs quantum-well lasers exhibit a lower threshold current, a somewhat higher differential quantum efficiency, and a weaker temperature dependence of threshold current when compared with regular double-heterostructure lasers.

Some of the results that demonstrate the high performance of AlGaAs quantum-well lasers are: (i) fabrication of GRIN SQW lasers with the lowest reported[5] threshold current density (200 A/cm^2, which compares with values in the range of 0.6–1 kA/cm^2 for regular double-heterostructure lasers); (ii) fabrication of GRIN SQW buried-heterostructure lasers with the lowest reported[28] threshold current (2.5 mA); (iii) fabrication of MQW lasers[4] with the threshold current density of 250 A/cm^2; and (iv) fabrication of laser arrays[33] with MQW active regions that have been operated to very high CW power (more than 2 W). Also, some of the studies relating to the differences between the gain and recombination mechanisms in quantum-well structures and regular double heterostructures have been carried out using the AlGaAs material system. These studies include polarization dependence of gain,[34] spectral characteristics of gain, and measurements of carrier lifetime.[35,36]

Rezek and coworkers[37–39] first reported on the growth of InGaAsP quantum-well structures by liquid-phase epitaxy (LPE). The room-temperature threshold current density of these lasers was ~ 5.2 kA/cm^2 μm. Figure 9.16 shows the measured temperature dependence of the threshold current of a MQW laser with 12 active layers of thickness ~ 16 nm each.

One of the interests in studying InGaAsP quantum-well structures is to examine the temperature dependence of the threshold current. In Sec. 3.6 we showed that the threshold current density as a function of temperature may be written as $J_{th}(T) \cong J_0 \exp(T/T_0)$. For regular double-heterostructure lasers, $T_0 \cong 110$ K when $T < 250$ K and $T_0 \cong 60$ K when $T > 250$ K. The temperature at which T_0 changes is often called the *break-point temperature* (T_b). Figure 9.16 shows that $T_b \cong 300$ K for MQW lasers, which suggests that they have a somewhat lower temperature sensitivity compared with that of regular InGaAsP double-heterostructure lasers, but not as low as that of GaAs double-heterostructure lasers.

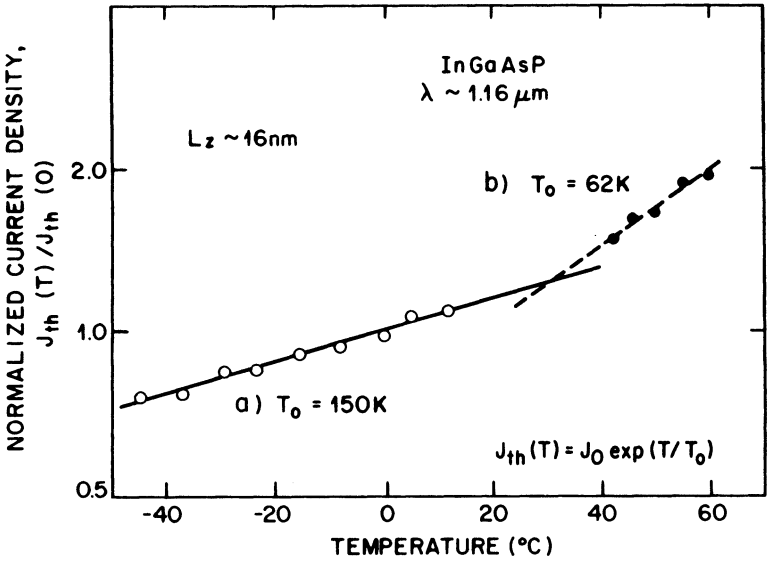


Fig. 9.16 Measured temperature dependence of threshold current of an InGaAsP MQW laser. Solid and dashed lines show the exponential fit below and above the break-point temperature occurring near room temperature. (After Ref. 37)

InGaAsP double-channel planar buried-heterostructure (DCPBH) lasers with a multiquantum-well active layer have also been fabricated.[40,41] The MQW structure was grown by LPE. Two types of MQW structures were studied: (i) with InP barrier layers[40] and (ii) with InGaAsP ($\lambda \cong 1.0\ \mu m$) barrier layers.[41] The MQW active region had four active wells ($\lambda \cong 1.3\ \mu m$, InGaAsP) and three barrier layers. Figure 9.17 shows a schematic cross section of the DCPBH laser and the MQW structure of its active layer. DCPBH lasers with InP barrier layer had room temperature threshold currents in the range of 40–50 mA, while for DCPBH lasers with InGaAsP barrier layers ($\lambda \cong 1.03\ \mu m$), the threshold currents were in the range 20–25 mA. The lower threshold current of devices with a lower barrier height may be due to more uniform current injection. A similar observation was made for AlGaAs MQW lasers.[4,42] Figure 9.18 shows the CW light-current characteristics at different temperatures of an InGaAsP DCPBH MQW laser and should be compared with those shown in Fig. 5.25 for regular double-heterostructure lasers. In both cases the threshold current $I_{th} \cong I_0 \exp(T/T_0)$. However, for MQW lasers the characteristic temperature ($T_0 \cong 100$ K) is higher than for the regular double-heterostructure lasers ($T_0 \cong 60$ K), suggesting that the former are less sensitive to temperature increases.

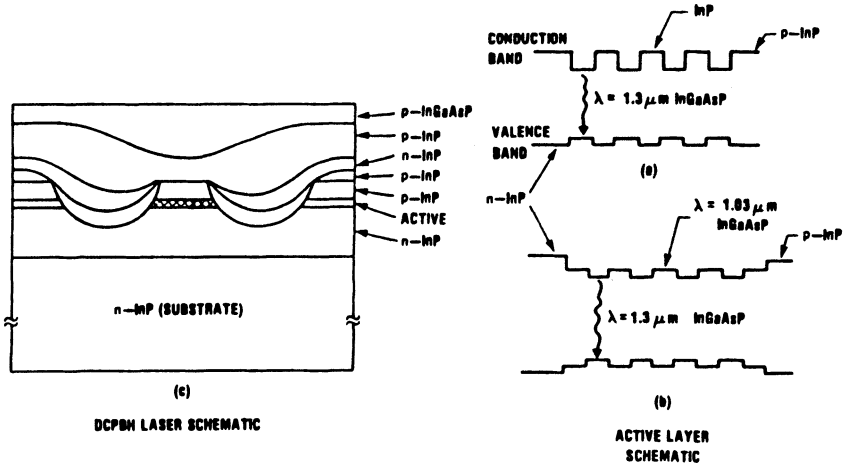


Fig. 9.17 Schematic cross sections of the DCPBH laser and its MQW active region. Details of the active region are shown on the right for (a) InP and (b) 1.03-μm InGaAsP barrier layers. (After Ref. 41)

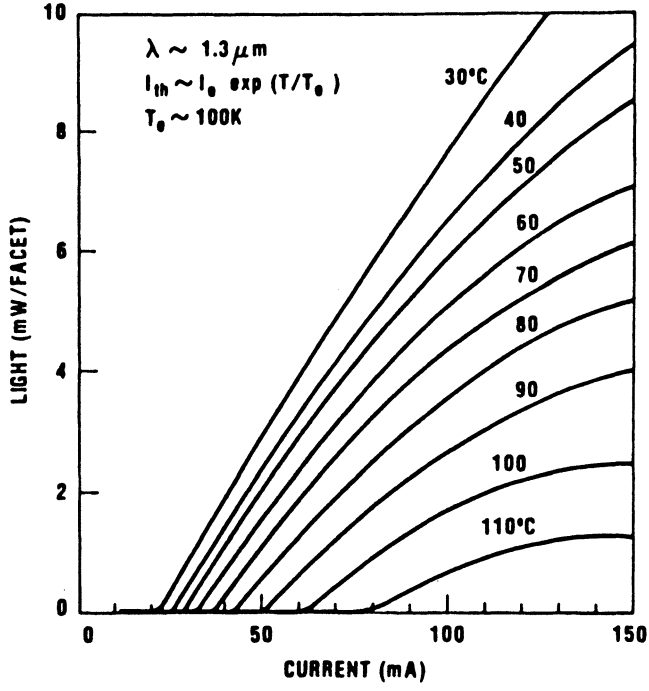


Fig. 9.18 Light-current characteristics of an InGaAsP MQW laser at different temperatures. (After Ref. 41)

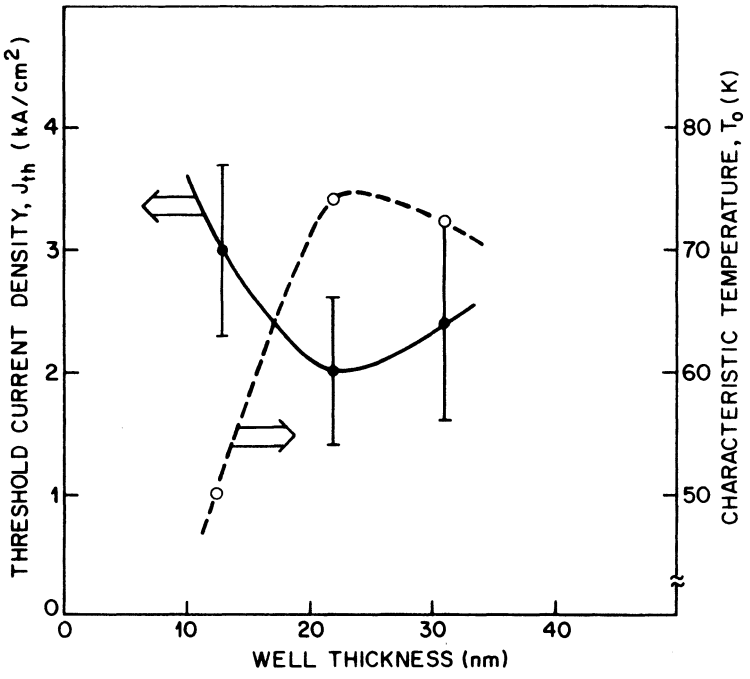


Fig. 9.19 Measured threshold current density J_{th} (solid curve) and characteristic temperature T_0 (dashed curve) as a function of active-layer thickness for a 1.3-μm InGaAsP-InP MQW laser. (After Ref. 43)

InGaAsP-InP MQW lasers emitting near 1.3 μm have also been fabricated using the hydride-transport VPE technique.[43] Figure 9.19 shows the threshold current density measured in these broad-area lasers as a function of well thickness. Also shown in the figure are the measured T_0 values that are comparable to the values reported for regular InGaAsP double-heterostructure lasers. The number and the thickness of quantum wells used in these lasers were such that the total active thickness ($\sim$0.2 μm) was relatively constant for the three devices shown.

Quantum-well lasers emitting near 1.55 μm, the region of lowest optical loss in silica fibers, have also attracted considerable attention.[44–46] Much of the work on InGaAsP/InP MQW lasers has been carried out using the MOVPE and CBE epitaxial growth techniques.[47,48] A separate-confinement heterostructure (SCH) is generally used to increase the confinement of the lasing mode to the active region. It has been shown that the optical loss of the mode can be reduced significantly, and hence high powers can be obtained, by a proper choice of the composition of the materials that form

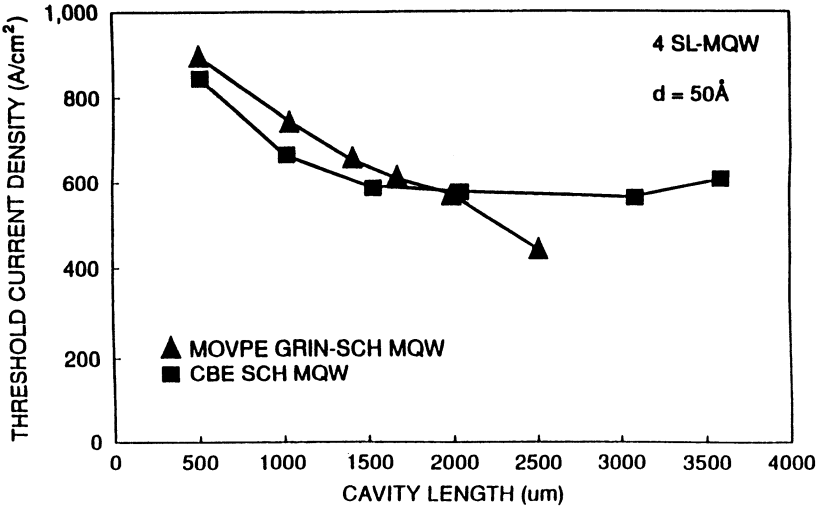


Fig. 9.20 Threshold current density plotted as a function of cavity length for InGaAsP MQW lasers grown by the MOVPE and CBE techniques.

the GRIN region and the barrier layers.[47] The measured threshold current density for InGaAs MQW lasers grown by the MOVPE and CBE growth techniques is shown in Fig. 9.20. The threshold current density is found to decrease with increasing cavity length with the lowest value being 380 A/cm^2 for 2-mm cavity length.

As discussed in Chapter 2, index guiding along the junction plane is necessary in order to control the near-field pattern of a semiconductor laser. The schematic of an index-guided buried heterostructure laser is shown in Fig. 9.21. The active-region width is ~ 2 μm. The laser utilizes Fe-doped InP layers for confinement of the current to the active region and for providing lateral index guiding.

Threshold current of a laser can be reduced by using high-reflectivity facet coatings.[49] For QW lasers the reduction is larger than that for regular double-heterostructures because the transparency current density is lower for QW lasers. The lower transparency current density arises from the two-dimensional density of states in MQWs. The optical gain for a laser with a current density J is given by (Sec. 9.4)

$$g = a(J - J_{tr}) \tag{9.7.1}$$

where a is the gain constant, J_{tr} is the transparency current density. At threshold, gain equals total loss, and the latter is the sum of the of the free

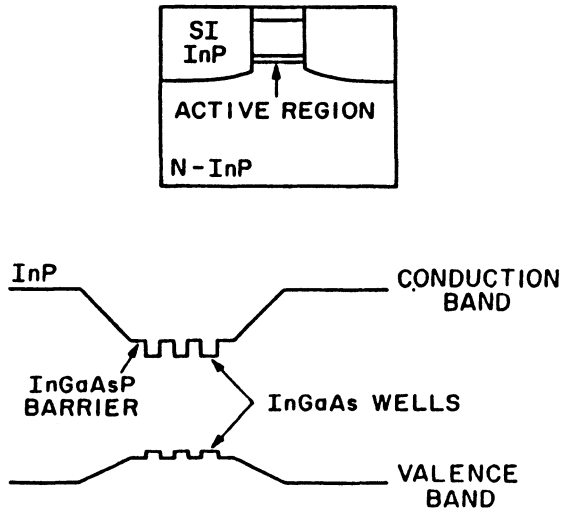


Fig. 9.21 Schematic illustration of an index-guided InGaAsP MQW laser. The band diagram (bottom) shows the graded-index region between the barrier and cladding regions.

carrier and mirror losses, i.e.

$$g = a(J_{th} - J_{tr}) = \alpha + \frac{1}{L}\ln\left(\frac{1}{R_1 R_2}\right) \tag{9.7.2}$$

where J_{th} is the threshold current density, α is the free carrier loss, L is the cavity length and R_1, R_2 are the mirror reflectivities. It follows from Eq. (9.7.2) that J_{th} is given by

$$J_{th} = J_{tr} + \frac{\alpha}{a} + \frac{1}{La}\ln\left(\frac{1}{R_1 R_2}\right). \tag{9.7.3}$$

By increasing R_1, R_2 and decreasing the cavity loss α, J_{th} can be made close to J_{tr}. Typically $J_{tr} = 50\ \text{A/cm}^2$ for QW lasers whereas for DH lasers it is $\sim 500\ \text{A/cm}^2$.

Threshold currents of both GaAs and InGaAsP MQW lasers have been reduced using high-reflectivity facet coatings. Such lasers are important for parallel data link applications. InGaAsP MQW laser emitting near 1.55 μm with a threshold current of 1 mA have been realized.[47] These lasers have 90% reflective coating at the back facet and 70% reflective coating at the front facet. GaAs QW lasers with a threshold current of 0.35 mA have also been obtained.[50] These lasers had 98% reflective coating on both facets.

9.8 MODULATION AND NOISE CHARACTERISTICS

For lightwave system applications, it is important to investigate if the quantum-well nature of a semiconductor laser affects other operating characteristics such as the relaxation-oscillation frequency, the laser line width, and the frequency chirp. These characteristics were discussed in Chapter 6 for regular double-heterostructure lasers. Among various parameters used there (see Table 6.1), the parameters that may be particularly influenced by quantum-size effects are the gain coefficient a, the carrier density at transparency n_0, and the line-width enhancement factor β_c. Attempts have been made[51–53] to analyze the dependence of these parameters on MQW design parameters such as the number and thickness of quantum wells.

The gain parameters a and n_0 for regular double-heterostructure lasers are estimated using the linear-gain model $g(n) = a(n - n_0)$ described in Sec. 2.4. Although this model is reasonably accurate for those lasers (see Fig. 3.8), the gain variation for quantum-well semiconductor lasers is far from linear (see Figs. 9.8 and 9.10); the gain eventually saturates because of the constant density of states of the lowest subband [see Eq. (9.2.3)]. As a result, the gain coefficient defined as $a = dg/dn$ varies with the carrier

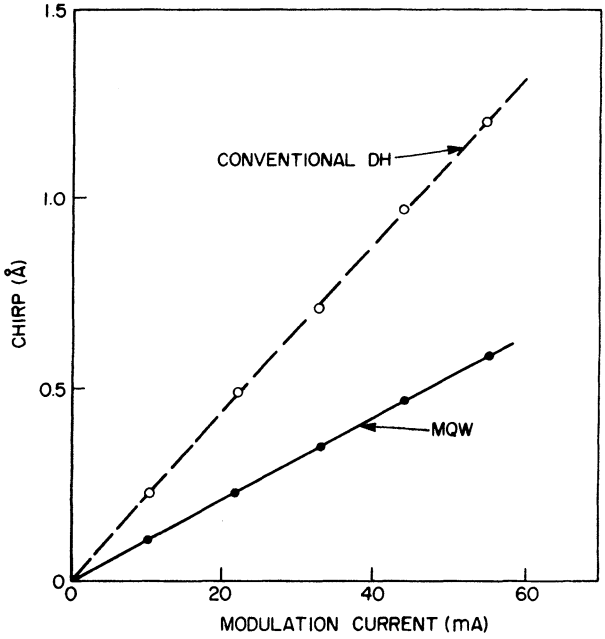


Fig. 9.22 Wavelength chirp as a function of the peak-to-peak modulation current for an MQW laser (solid line) and a regular double-heterostructure laser (dashed line). (After Ref. 40)

density and would be different for different MQW lasers, depending on the structure parameters such as the number and thickness of quantum wells. Calculations[52,53] for SQW and MQW lasers show that with a proper design, the effective value of a can be nearly doubled compared with that of regular semiconductor lasers. Equations (6.4.23) and (6.4.26) then show that, if other parameters are approximately the same, the relaxation-oscillation frequency Ω_R would increase by about 50%. An increase in Ω_R can enhance the modulation bandwidth (Sec. 6.6.2) and decrease the intensity noise (Sec. 6.5.2) at a given output power.

The line-width enhancement factor β_c is defined by Eq. (2.4.6). This equation shows that if the carrier-induced index change is relatively unaffected, an increase in a would lead to a decrease in β_c by the same factor. A lower value of β_c for quantum-well lasers has been predicted theoretically.[51–53] As discussed in Chapter 6, the main effect of this lower value would be to reduce the CW line width [Eq. (6.5.52)] as well as the frequency chirp [Eq. (6.6.45)]. The measurements of the frequency chirp for quantum-well semiconductor lasers shown[40] in Fig. 9.22 suggest that it is smaller by about a factor of 2 compared with the values obtained for regular double-heterostructure lasers. By the same token, the CW line width should reduce by about a factor of 4. These results suggest that properly optimized MQW semiconductor lasers may be useful for applications in optical communication systems.

9.9 STRAINED QUANTUM-WELL LASERS

Quantum-well lasers have also been fabricated using an active layer whose lattice constant differs slightly from that of the substrate and cladding layers. Such lasers are known as strained quantum-well lasers. Over the past few years, strained quantum-well lasers have been extensively investigated all over the world.[54–61] They show many desirable properties such as (i) a very low threshold current density and (ii) a lower line width both under CW operation and under modulation. The origin of the improved device performance lies in the band-structure changes induced by the mismatch-induced strain. Figure 9.23 shows the band structure of a semiconductor under tensile and compressive strains. Strain splits the heavy-hole and the light-hole valence bands at the Γ point of the Brillouin zone where the band gap is minimum in direct band gap semiconductors.

Two material systems have been widely used for strained quantum-well lasers: (i) InGaAs grown over InP by the MOVPE or the CBE growth technique[54–57] and (ii) InGaAs grown over GaAs by the MOVPE or the MBE growth technique.[58–61] The former material system is of importance of a low-chirp semiconductor laser for lightwave system applications, while

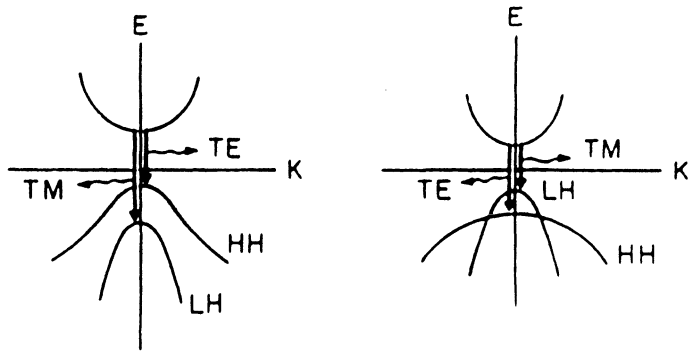


Fig. 9.23 Band structure of a direct band-gap semiconductor under stress. The left and right diagrams show the cases of compressive and tensile strain respectively. The heavy-hole (HH) and light-hole (LH) bands coincide at the center in the absence of strain. Emitted light is TE or TM polarized depending on whether the electron combines with the heavy or light hole.

the latter material system has been used to fabricate high-power lasers emitting near 0.98 μm, a wavelength of interest for pumping erbium-doped fiber amplifiers.

The effects of strain on optical gain and effective mass in semiconductors have been calculated.[62,63] The issues of most practical interest are the relationship between gain and injected carrier density and gain and radiative current density. To illustrate the beneficial effects of strain, a comparison between unstrained and strained QWs is made for the following cases: (i) an 8-nm $GaAs/Al_{0.2}Ga_{0.8}As$ QW which is unstrained and (ii) an 8-nm $In_{0.2}Ga_{0.8}As$-$Al_{0.2}Ga_{0.8}As$ strained quantum well. The results of the calculation[63] are shown in Fig. 9.24. Both the carrier density and the current density needed for a fixed gain are smaller for the strained case than those for the unstrained case. This reduction is partly due to lowering of the effective mass of electrons in the presence of strain.

The alloy $In_{0.53}Ga_{0.47}As$ has the same lattice constant as InP (see Fig. 4.7). Semiconductor lasers with an $In_{0.53}Ga_{0.47}As$ active region have been grown on InP by the MOVPE growth technique. Excellent material quality is also obtained for $In_xGa_{1-x}As$ alloys grown over InP by MOVPE for nonlattice-matched compositions. In this case the laser structure generally consists of one or many $In_xGa_{1-x}As$ quantum-well layers with InGaAsP barrier layers whose composition is lattice-matched to that of InP. For $x < 0.53$, the active layer in these lasers is under tensile stress while for $x > 0.53$, the active layer is under compressive stress.

Superlattice structures of InGaAs-InGaAsP with tensile and compressive stress have been grown by both MOVPE and CBE growth techniques over

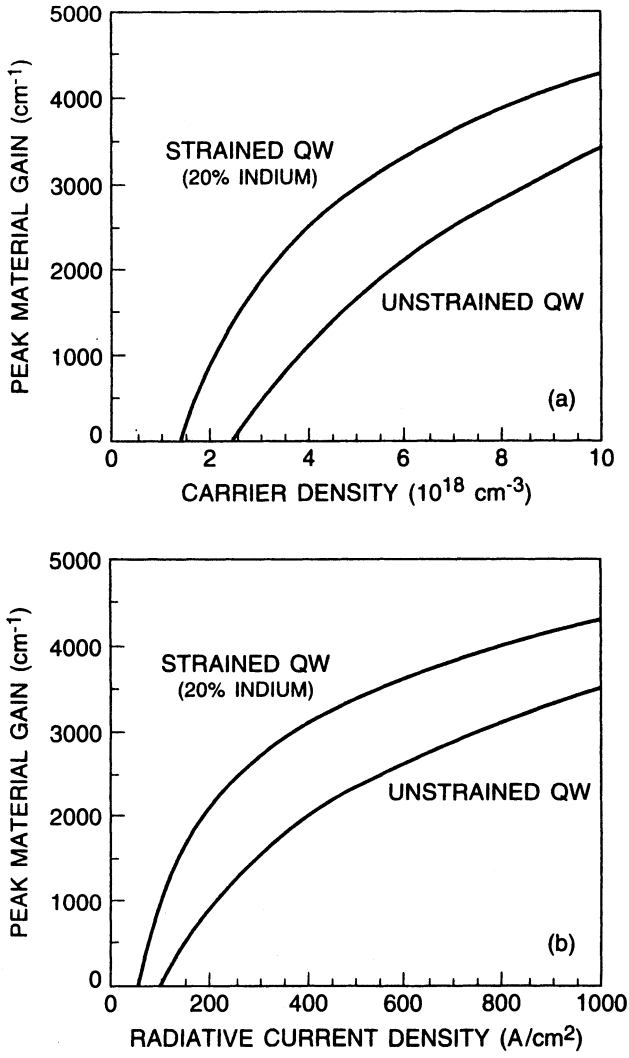


Fig. 9.24 Optical gain calculated as a function of (a) injected carrier density and (b) radiative current density for strained and unstrained quantum-well lasers.

an *n*-type InP substrate. Figure 9.25 shows the broad-area threshold current density as a function of cavity length for strained MQW lasers with four $In_{0.65}Ga_{0.35}As$ quantum wells.[64] The active region in this laser is under 0.8% compressive strain. Also shown for comparison is the threshold current density as a function of cavity length of MQW lattice-matched lasers with

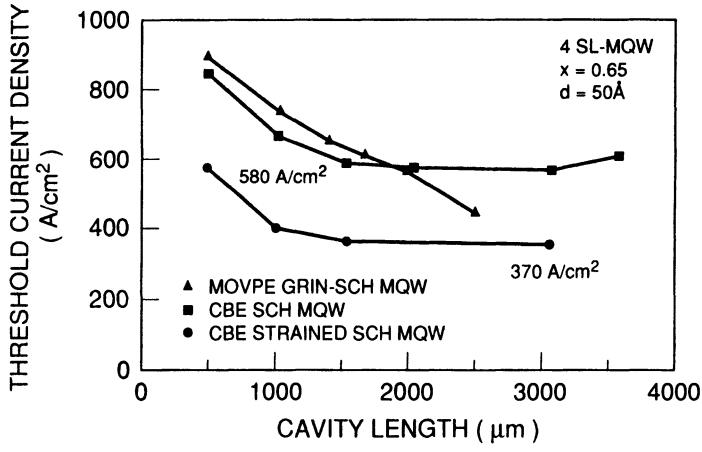


Fig. 9.25 Broad-area threshold current density plotted as a function of cavity length for strained and unstrained InGaAsP MQW lasers. (After Ref. 64)

$In_{0.53}Ga_{0.47}As$ wells. The entire laser structure, apart from the quantum-well composition, is identical for the two cases. The threshold current density is lower for the compressively strained MQW structure than that for the lattice-matched MQW structure.

Buried-heterostructure lasers have been fabricated using compressive- and tensile-strained MQW lasers. The threshold current of these lasers as a function of the In concentration is shown in Fig. 9.26. Lasers with compressive strain have a lower threshold current than that for lasers with tensile strain.[65] This can be explained by splitting of the light hole and heavy hole bands under stress (see Fig. 9.23).[62,66,67]

Strained quantum well lasers fabricated using $In_xGa_{1-x}As$ layers grown over a GaAs substrate have been extensively studied. The lattice constant of InAs is 60.6 pm and that of GaAs is 56.54 pm. The $In_xGa_{1-x}As$ alloy has a lattice constant, between these two values, and to a first approximation it can be assumed to vary linearly with x. Thus, an increase in the In mole fraction x increases the lattice mismatch relative to the GaAs substrate and therefore produces larger compressive strain on the active region.

A typical laser structure grown over the n-type GaAs substrate is shown in Fig. 9.27 for this material system. It consists of an MQW active region with one to four $In_xGa_{1-x}As$ wells separated by GaAs barrier layers. The entire MQW structure is sandwiched between n- and p-type $Al_{0.4}Ga_{0.6}As$ cladding layers, and the p-cladding layer is followed by a p-type GaAs contact layer. Variations of the above structure with different cladding layers or large optical cavity designs have been reported.[68–70] Emission wavelength depends

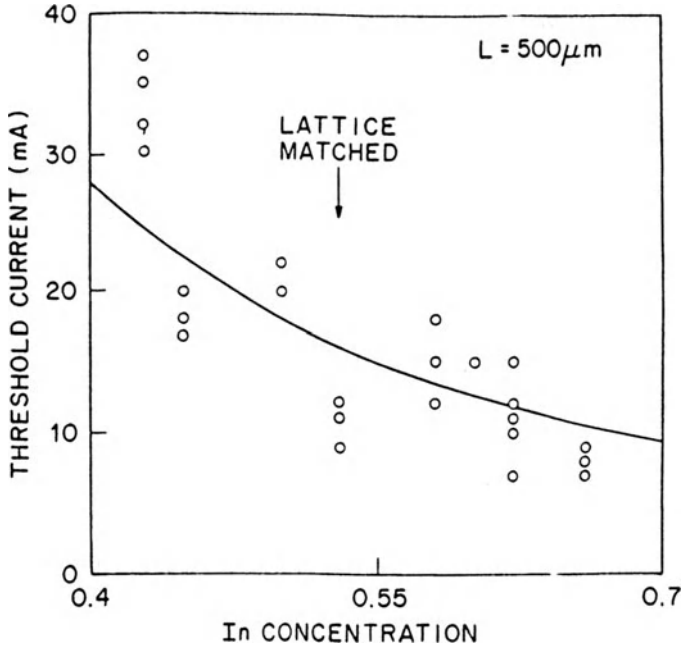


Fig. 9.26 Measured threshold current plotted as a function of In concentration x for several $In_xGa_{1-x}As$ MQW lasers. The arrow indicates the lattice-matched composition ($x = 0.53$). (After Ref. 65 © 1991 IEEE)

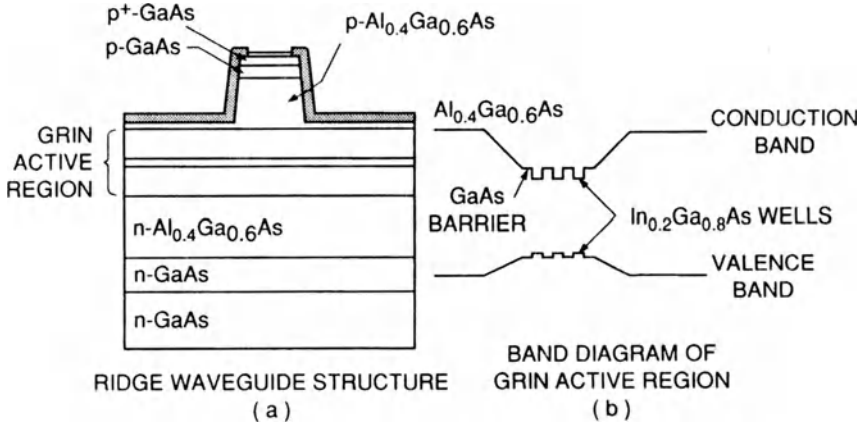


Fig. 9.27 (a) Typical structure for an $In_xGa_{1-x}As$ MQW laser. (b) Band diagram of graded-index active region.

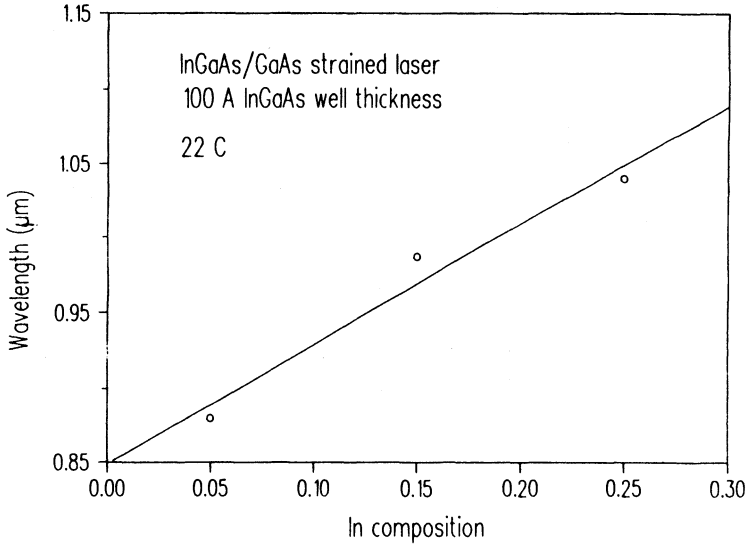


Fig. 9.28 Measured wavelength plotted as a function of In concentration x for several $In_xGa_{1-x}As$ MQW lasers.

on the In composition, x. As x increases, the emission wavelength increases (see Fig. 9.28), and for x larger than a certain value (typically ~ 0.25), the strain is too large to yield high-quality material.

For $x \simeq 0.2$, the emission wavelength is near 0.98 μm, a wavelength region of interest for pumping fiber amplifiers. A threshold current density as low as 47 A/cm^2 has been reported for $In_{0.2}Ga_{0.8}As$-GaAs strained MQW lasers.[71,72] High-power lasers have been fabricated using an $In_{0.2}Ga_{0.8}As$-GaAs MQW active region. Single-mode output powers of greater than 200 mW have been demonstrated using a ridge-waveguide-type laser structure. CW output powers greater than 1 W have been demonstrated using a laser-array geometry.[73] Aging tests indicate high reliability for these strained quantum well lasers.[74]

The frequency chirp of strained and unstrained QW lasers has been measured.[75,76] Strained QW lasers (InGaAs-GaAs) exhibit the lowest chirp (or dynamic line width) under modulation. The lower chirp of strained QW lasers is consistent with a small line-width enhancement factor measured in such devices. The correlation between the measured chirp and line-width enhancement factor for regular double-heterostructure, strained and unstrained QW lasers is shown in Table 9.1. The high efficiency, high power and low chirp of strained and unstrained QW lasers make these devices attractive for lightwave transmission applications.

Table 9.1 Correlation between the measured chirp and line-width enhancement factor for regular double-heterostructure, strained and unstrained QW lasers

	LINE-WIDTH ENHANCEMENT FACTOR	FWHM CHIRP AT 50 mA (1 Gb/s)
DH laser	5.5	0.12 nm
MQW laser ($\lambda = 1.5\ \mu m$)	3.5	0.06 nm
Strained MQW InGaAs-GaAs laser ($\lambda = 1\ \mu m$)	1.0	0.02 nm
Strained MQW InGaAsP-InP laser ($\lambda = 1.5\ \mu m$)	2.0	0.04 nm

PROBLEMS

9.1 What is the basic idea behind quantum-well semiconductor lasers? What is meant by quantum well?

9.2 Consider the motion of an electron inside an active layer of thickness L_z in the direction perpendicular to the layer surface. Find the energy levels for the case in which electrons cannot escape outside the active layer because of an infinite potential barrier.

9.3 Repeat Problem 9.2 for the case of finite potential barrier and show that the energy levels are obtained by solving the eigenvalue equation (9.1.6).

9.4 What is meant by the density of states in a semiconductor? Derive an expression for the density of states of a quantum-well laser and compare it with that obtained for conventional semiconductor lasers.

9.5 A quantum-well semiconductor laser consists of a 10-nm-thick InGaAs active layer sandwiched between two InP cladding layers. Calculate the operating wavelength of such a laser by using Eq. (9.1.6) for the potential barrier.

9.6 An InGaAsP laser operates at 1.55 μm when the active layer is 0.2-μm thick. Calculate the wavelength if the active-layer thickness is reduced to 8 nm.

9.7 Calculate the threshold carrier density of a 1.55-μm InGaAsP laser with 10-nm-thick active layer and 1.3-μm cladding layers. Use Eq. (4.7.4) to obtain the refractive indices and Eq. (9.6.1) to obtain the confinement factor. Assume 32% facet reflectivities, $20\ cm^{-1}$ internal loss, and 250-μm cavity length. The gain can be assumed to vary linearly with the carrier density n as $g = a(n - n_0)$ with $a = 4 \times 10^{-16}\ cm^2$ and $n_0 = 1.0 \times 10^{18}\ cm^{-3}$.

9.8 What is the carrier density if the single quantum well in Problem 9.7 is replaced by a multiquantum-well active region containing five 10-nm-thick active layers spaced by 8-nm-thick barrier layers?

9.9 The differential gain coefficient of a semiconductor laser increases by a factor of 2 when its active-layer thickness is reduced from 100 to 10 nm. What would be the expected increase in its small-signal modulation bandwidth if all other parameters can be assumed to remain the same?

9.10 The line-width enhancement factor of a semiconductor laser decreases by a factor of 2 when its active-layer thickness is reduced from 100 to 10 nm. How would this change affect the laser performance under CW operation and under 5 Gb/s modulation?

REFERENCES

1. Dingle, R., W. Wiegmann, and C. H. Henry. *Phys. Rev. Lett.* **33,** 827 (1974); R. Dingle, and C. H. Henry, U.S. Patent 3,982,207 (Sept. 21, 1976).
2. Holonyak, N., Jr., R. M. Kolbas, R. D. Dupuis, and P. D. Dapkus. *IEEE J. Quantum Electron.* **QE-16,** 170 (1980).
3. Yariv, A. *IEEE Circuits & Dev. Mag.* **5**(6), 25 (1989).
4. Special issue on semiconductor lasers. *IEEE J. Quantum Electron.* **QE-27,** 1317 (1991).
5. Zory, P. S., Editor. *Quantum Well Lasers.* San Diego, California: Academic Press, 1992.
6. Cole, M. W. *Rev. Mod. Phys.* **46,** 451 (1974).
7. Dorda, G. p. 215 in *Festkörperprobleme* [Advances in Solid State Physics], vol. XIII. Braunschweig: Pergamon/Vieweg, 1973.
8. Dingle, R. p. 21 in *Festkörperprobleme* [Advances in Solid State Physics], vol. XV. Braunschweig: Pergamon/Vieweg, 1975.

9a. Okumura, H., S. Misawa, S. Yoshida, and S. Gonda. *Appl. Phys. Lett.* **46,** 377 (1985).

9b. Forrest, S. R., P. H. Schmidt, R. B. Wilson, and M. L. Kaplan. *Appl. Phys. Lett.* **45,** 1199 (1984); *J. Vac. Sci. Technol. B* **4,** 37 (1986).

10. Shank, C. V., R. L. Ford, R. F. Leheny, and J. Shah. *Phys. Rev. Lett.* **42,** 112 (1979).
11. Temkin, H., M. B. Panish, P. M. Petroff, R. A. Hamm, J. M. Vandenberg, and S. Sumski. *Appl. Phys. Lett.* **47,** 394 (1985).
12. Hess, K., B. A. Vojak, N. Holonyak, Jr., R. Chin, and P. D. Dapkus. *Solid-State Electron.* **23,** 585 (1980).
13. Dutta, N. K. *J. Appl. Phys.* **53,** 7211 (1982).
14. Sugimura, A. *IEEE J. Quantum Electron.* **QE-20,** 336 (1984).
15. Asada, M., A. Kameyama, and Y. Suematsu. *IEEE J. Quantum Electron.* **QE-20,** 745 (1984).
16. Tsang, W. T., and J. A. Ditzenberger. *Appl. Phys. Lett.* **39,** 193 (1981).
17. Sugimura, A. *Appl. Phys. Lett.* **43,** 728 (1983).
18. Chiu, L. C., and A. Yariv. *IEEE J. Quantum Electron.* **QE-18,** 1406 (1982).
19. Dutta, N. K. *J. Appl. Phys.* **54,** 1236 (1983).
20. Sugimura, A. *IEEE J. Quantum Electron.* **QE-19,** 932 (1983).
21. Smith, C., R. A. Abram, and M. G. Burt. *J. Phys. C* **16,** L171 (1983).
22. Tsang, W. T. *Appl. Phys. Lett.* **39,** 135 (1981).
23. Dumke, W. P. *IEEE J. Quantum Electron.* **QE-11,** 400 (1975).
24. Streifer, W., D. R. Scifres, and R. D. Burnham. *Appl. Opt.* **18,** 3547 (1979).

25. Tsang, W. T. *IEEE J. Quantum Electron.* **QE-20,** 1119 (1984); see references cited therein.
26. Tsang, W. T. *Appl. Phys. Lett.* **38,** 204 (1981).
27. Woodbridge, K., P. Blood, E. D. Fletcher, and P. J. Hulyer. *Appl. Phys. Lett.* **45,** 16 (1984).
28. Tsang, W. T., R. A. Logan, and J. A. Ditzenberger. *Electron. Lett.* **18,** 845 (1982).
29. Kasemset, D., C. S. Hong, N. B. Patel, and P. D. Dapkus. *Appl. Phys. Lett.* **41,** 912 (1982).
30. Hersee, S. D., M. A. Poisson, M. Baldy, J. P. Duchemin. *Electron. Lett.* **18,** 618 (1982).
31. Dupuis, R. D., R. L. Hartman, and F. R. Nash. *IEEE Electron Device Lett.* **EDL-4,** 286 (1983).
32. Fuzii, T., S. Yamakoshi, K. Nanbu, O. Wada, and S. Hiyamizu. *J. Vac. Sci. Technical. B* **2,** 259 (1984).
33. Scifres, D. R., R. D. Burnham, C. Lindström, W. Streifer, and T. L. Paoli. *Appl. Phys. Lett.* **42,** 645 (1983).
34. Kobayashi, H., H. Iwamura, T. Saku, and K. Otsuka. *Electron. Lett.* **19,** 156 (1983).
35. Dutta, N. K., R. L. Hartman, and W. T. Tsang. *IEEE J. Quantum Electron.* **QE-19,** 1243 (1983); **QE-19,** 1613 (1983).
36. Olsson, N. A., N. K. Dutta, W. T. Tsang, and R. A. Logan. *Electron. Lett.* **20,** 63 (1984).
37. Rezek, E. A., N. Holonyak, Jr., B. A. Vojak, G. E. Stillman, J. A. Rossi, D. L. Keune, and J. D. Fairing. *Appl. Phys. Lett.* **31,** 288 (1977).
38. Rezek, E. A., R. Chin, N. Holonyak, Jr., S. W. Kirchofer, and R. M. Kolbas. *J. Electron. Mater.* **9,** 1 (1980).
39. Rezek, E. A., N. Holonyak, Jr., and B. K. Fuller. *J. Appl. Phys.* **51,** 2402 (1980).
40. Dutta, N. K., S. G. Napholtz, R. Yen, R. L. Brown, T. M. Shen, N. A. Olsson, and D. C. Craft. *Appl. Phys. Lett.* **46,** 19 (1985).
41. Dutta, N. K., S. G. Napholtz, R. Yen, T. Wessel, N. A. Olsson. *Appl. Phys. Lett.* **46,** 1036 (1985).
42. Dutta, N. K. *IEEE J. Quantum Electron.* **QE-19,** 794 (1983).
43. Yanase, T., Y. Kato, I. Mito, M. Yamoyuchi, K. Nishi, K. Kobayashi, and R. Lang. *Electron. Lett.* **19,** 700 (1983).
44. Temkin, H., K. Alavé, W. R. Wagner, T. P. Pearsall, and A. Y. Cho. *Appl. Phys. Lett.* **42,** 845 (1983).
45. Tsang, W. T. *Appl. Phys. Lett.* **44,** 288 (1984).
46. Dutta, N. K., T. Wessel, N. A. Olsson, R. A. Logan, and R. Yen. *Appl. Phys. Lett.* **46,** 525 (1985).
47. Temkin, H., N. K. Dutta, T. Tanbun-Ek, R. A. Logan, and A. M. Sergent. *Appl. Phys. Lett.* **57,** 1610 (1990).
48. Tsang, W. T., L. Yang, M. C. Wu, Y. K. Chen, and A. M. Sergent. *Electron. Lett.* **26,** 2033 (1990).
49. Derry, P. L., A. Yariv, K. Y. Lau, N. Bar-Chaim, K. Lee, and J. Rosenberg. *Appl. Phys. Lett.* **50,** 1773 (1987).
50. Kapon, E., S. Simhony, J. P. Harbison, L. T. Florez, and P. Worland. *Appl. Phys. Lett.* **56,** 1825 (1990).
51. Burt, M. G. *Electron. Lett.* **20,** 27 (1984).
52. Arakawa, Y., K. Vahala, and A. Yariv. *Appl. Phys. Lett.* **45,** 950 (1984).
53. Arawaka, Y., and A. Yariv. *IEEE J. Quantum Electron.* **QE-21,** 1666 (1985).
54. Thijs, P. J. A., L. F. Tiemeijer, P. I. Kuindersma, J. J. M. Binsma, and T. van Dongen. *IEEE J. Quantum Electron.* **QE-27,** 1426 (1991).
55. Thijs, P. J. A., and T. van Dongen. *Electron. Lett.* **25,** 1735 (1989).
56. Temkin, H., T. Tanbun-Ek, and R. A. Logan. *Appl. Phys. Lett.* **56,** 1210 (1990).
57. Koren, U., M. Oron, M. G., B. I. Miller, J. L. DeMiguel, G. Raybon, and M. Chien. *Electron. Lett.* **26,** 465 (1990).

58. Laidig, W. D., Y. F. Lin, and P. J. Caldwell. *J. Appl. Phys.* **57,** 33 (1985).
59. Fischer, S. E., D. Fekete, G. B. Feak, and J. M. Ballantyne. *Appl. Phys. Lett.* **50,** 714 (1987).
60. Bour, D. P., D. B. Gilbert, L. Elbaum, and M. G. Harvey. *Appl. Phys. Lett.* **53,** 2371 (1988).
61. Beernik, K. J., P. K. York, and J. J. Coleman. *Appl. Phys. Lett.* **55,** 2582 (1989).
62. Loehr, J. P., and J. Singh. *IEEE J. Quantum Electron.* **QE-27,** 708 (1991).
63. Corzine, S. W., R. Yan, and L. A. Coldren. In *Quantum Well Lasers,* ed. by P. S. Zory. San Diego: Academic Press, 1992.
64. Tsang, W. T., L. Yang, M. C. Wu, Y. K. Chen, and A. M. Sergent. *Electron. Lett.* **26,** 2033 (1990).
65. Temkin, H., T. Tanbun-Ek, R. A. Logan, D. A. Cebula, and A. M. Sergent. *IEEE Photon. Technol. Lett.* **3,** 100 (1991).
66. Adams, A. R. *Electron. Lett.* **22,** 249 (1986).
67. Yablonovitch, E., and E. O. Kane. *J. Lightwave Technol.* **LT-4,** 504 (1986).
68. Wu, M. C., Y. K. Chen, M. Hong, J. P. Mannaerts, M. A. Chin, and A. M. Sergent. *Appl. Phys. Lett.* **59,** 1046 (1991).
69. Dutta, N. K., J. Lopata, P. R. Berger, D. L. Sivco, and A. Y. Cho. *Electron. Lett.* **27,** 680 (1991).
70. Kuo, J. M., M. C. Wu, Y. K. Chen, and M. A. Chin. *Appl. Phys. Lett.* **59,** 2781 (1991).
71. Chand, N., E. E. Becker, J. P. van der Ziel, S. N. G. Chu, and N. K. Dutta. *Appl. Phys. Lett.* **58,** 1704 (1991).
72. Choi, H. K., and C. A. Wang. *Appl. Phys. Lett.* **57,** 321 (1990).
73. Dutta, N. K., J. D. Wynn, J. Lopata, D. L. Sivco, and A. Y. Cho. *Electron. Lett.* **26,** 1816 (1990).
74. Okayasu, M., M. Fukuda, M., T. Takeshita, and S. Vehara. *IEEE Photon. Technol. Lett.* **2,** 689 (1990).
75. Dutta, N. K., J. D. Wynn, D. L. Sivco, and A. Y. Cho. *Appl. Phys. Lett.* **56,** 2293 (1990).
76. Kano, F., Y. Yoshikuni, M. Fukuda, and J. Yoshida. *IEEE Photon. Technol. Lett.* **3,** 877 (1991).

Chapter 10
SURFACE-EMITTING LASERS

10.1 INTRODUCTION

Semiconductor lasers described in the previous chapters have cleaved facets that form the optical cavity. The facets are perpendicular to the surface of the wafer and light is emitted parallel to the surface of the wafer. For many applications requiring a two-dimensional laser array or monolithic integration of lasers with electronic components (e.g., optical interconnects), it is desirable to have the laser output normal to the surface of the wafer. Such lasers are known as surface-emitting lasers (SEL). There is a class of surface-emitting lasers which have their optical cavity normal to the surface of the wafer.[1–8] These devices are known as vertical-cavity surface-emitting lasers (VCSEL) in order to distinguish them from other surface emitters such as those described in Sec. 5.8. In this chapter the fabrication and performance characteristics of these vertical-cavity types of surface-emitting lasers are described.

In the late 1970s, Soda et al.[1] reported on an SEL fabricated using the InGaAsP material system. Their device structure is shown in Fig. 10.1. The surfaces of the wafer form the Fabry-Perot (FP) cavity of the laser. Fabrication of this device involves the growth of a double heterostructure on an *n*-type InP substrate. A circular contact is made on the *p*-side using an SiO_2 mask. The substrate side is polished making sure that it is parallel to the epitaxial layers, and ring electrodes (using an Au-Sn alloy) are deposited on the *n*-side. The lasers had a threshold current density of ~ 11 kA/cm^2 at 77 K and operated to output powers of several milliwatts.

Compared to this early work in the late 1970s, the advances in surface-emitting laser fabrication technology have been phenomenal. Present state-of-the-art SELs have threshold currents of < 1 mA at room temperature and can be modulated at high bit rates. Although there are many differences between the early work (Fig. 10.1) and the recent devices, one key difference is the development of very high reflectivity ($> 99\%$) mirrors by using Bragg

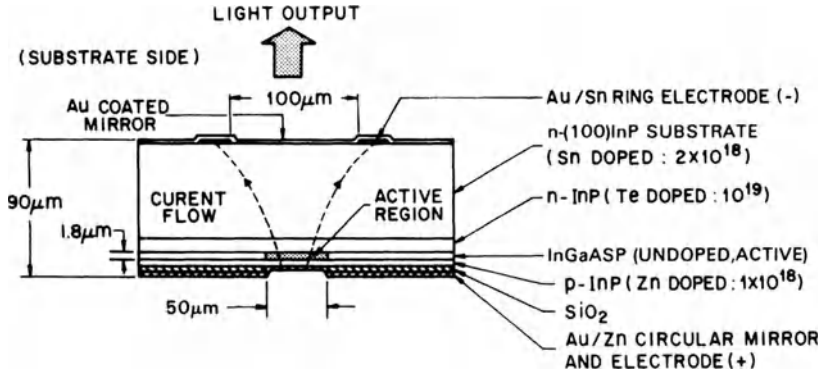


Fig. 10.1 Schematic cross section of a surface-emitting laser. Surfaces of the wafer form a Fabry-Perot cavity. (After Ref. 1)

reflectors composed of multilayer semiconductor stacks. The following section discusses this technology.

10.2 MIRROR REFLECTIVITY

A generic SEL structure utilizing multiple semiconductor layers to form a Bragg reflector is shown in Fig. 10.2. The active region is sandwiched between the *n*- and *p*-type cladding layers which are themselves sandwiched between the two *n*- and *p*-type Bragg mirrors. The Bragg mirrors consist of alternating layers of low-index and high-index materials. The thicknesses of each layer is equal to one quarter of the wavelength of light in that medium. Such periodic quarter-wave-thick layers can have a very high reflectivity. For normal incidence, the reflectivity is given by[9]

$$R = \left[\frac{1 - \dfrac{\mu_l}{\mu_1}\left(\dfrac{\mu_2}{\mu_3}\right)^{2N}}{1 + \dfrac{\mu_l}{\mu_1}\left(\dfrac{\mu_2}{\mu_3}\right)^{2N}} \right]^2 \tag{10.2.1}$$

where μ_2 and μ_3 are the refractive indices of the alternating layer pairs, μ_l and μ_1 are the refractive indices of the medium on the transmitted and incident sides of the DBR mirror, and N is the number of pairs. As N increases, R increases. Also for a given N, R is larger if the ratio μ_2/μ_3 is smaller.

For an SEL to have a threshold current density comparable to that of an

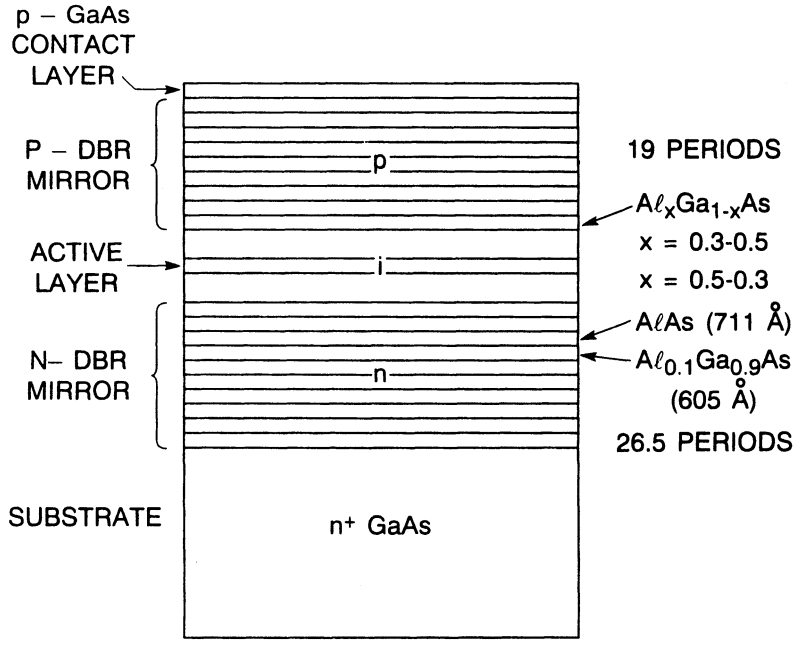


Fig. 10.2 Schematic illustration of a generic SEL structure utilizing distributed Bragg mirrors formed by using multiple semiconductor layers. DBR pairs consist of AlAs (71.1-nm thick) and $Al_{0.1}Ga_{0.9}As$ (60.5-nm thick) alternate layers. The active layer could be either a quantum well or have a thickness similar to a regular double-heterostructure laser.

edge-emitting laser, the threshold gains must be comparable for the two devices. The threshold gain of an edge-emitting laser is $\sim 100\,cm^{-1}$ (see Sec. 2.3). For an SEL with an active-layer thickness of 0.1 μm, this value corresponds to a single-pass gain of ~1%. Thus, for the SEL device to lase with a threshold current density comparable to that of an edge emitter, the mirror reflectivities must be >99% to ensure that cavity losses are smaller than the gain available during a single pass.

Instead of alternating low- and high-index layers of semiconductors, SELs have also been fabricated using alternating low- and high-index layers of dielectric materials such as SiO_2-TiO_2 or Si-SiO_2. These dielectric materials are useful for SELs emitting near 1.3 μm or 1.55 μm (made by using the InGaAsP material system), because it is more difficult to fabricate semiconductor stacks for the InGaAsP material system compared with the AlGaAs material system.

Central to the fabrication of low-threshold SELs is the ability to fabricate high-reflectivity mirrors. The scanning electron photomicrograph of an SEL

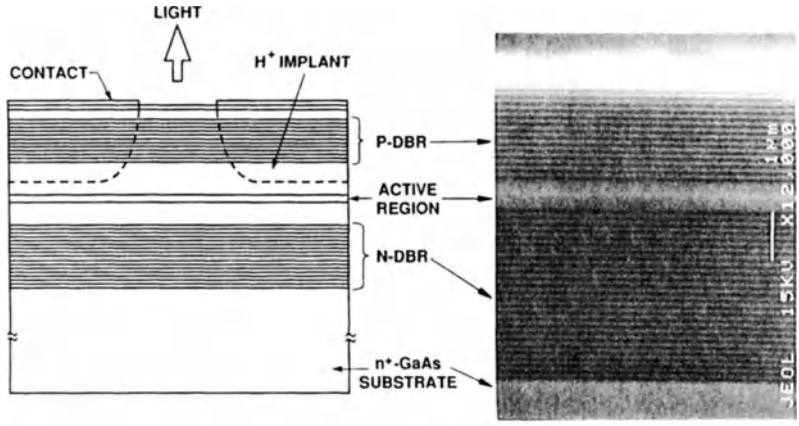


Fig. 10.3 Scanning electron photomicrograph of a GaAs-AlGaAs SEL wafer.

structure is shown in Fig. 10.3. It consists of alternating layers (22 pairs) of n-$Al_{0.9}Ga_{0.1}As$ and n-$Al_{0.1}Ga_{0.9}As$ grown over an n-GaAs substrate followed by the GaAs active region sandwiched between n- and p-type $Al_{0.3}Ga_{0.4}As$ cladding layers. The top cladding layer is further followed by 15 pairs of p-$Al_{0.9}Ga_{0.1}As$ layers and p-$Al_{0.1}Ga_{0.9}As$ layers and a thin ($\sim$20 nm) p-GaAs contact layer. The two stacks of alternating layers on each side of the active layer form the mirrors of the laser cavity. The thickness of each layer in the stack equals $\lambda/4$, where λ is the wavelength of light in the medium.

As the number of layer pairs increases, the effective reflectivity of the stack increases. The calculated peak reflectivity at 0.85 μm of an alternating stack of $Al_{0.9}Ga_{0.1}As$ and $Al_{0.1}Ga_{0.9}As$ layers as a function of the number of pairs is shown in Fig. 10.4. The reflectivity spectrum of an SEL structure is shown in Fig. 10.5. The reflectivity is >99% over a 10-nm band. The drop in reflectivity in the middle of the band is due to an FP mode.

The number of pairs needed to fabricate a high-reflectivity mirror depends on the refractive index of layers of a pair. For large index differences fewer pairs are needed. For example, in the case of alternating layers of ZnS and CaF_2 for which the index difference is 0.9, only six pairs are needed for a reflectivity of 99%. By contrast, for an InP-InGaAsP ($\lambda \sim 1.3$ μm) layer pair for which the index difference is 0.3, more than 40 pairs are needed to achieve a reflectivity of 99%.

While the large index difference between the alternating semiconductor layers is useful for high reflectivity, the accompanying energy-gap difference

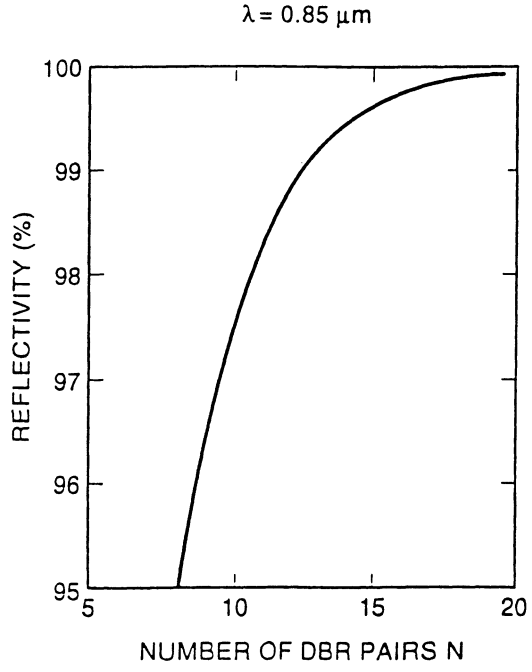


Fig. 10.4 Calculated reflectivity of an AlAs-$Al_{0.1}Ga_{0.9}As$ multilayer semiconductor Bragg reflector as a function of the number of pairs.

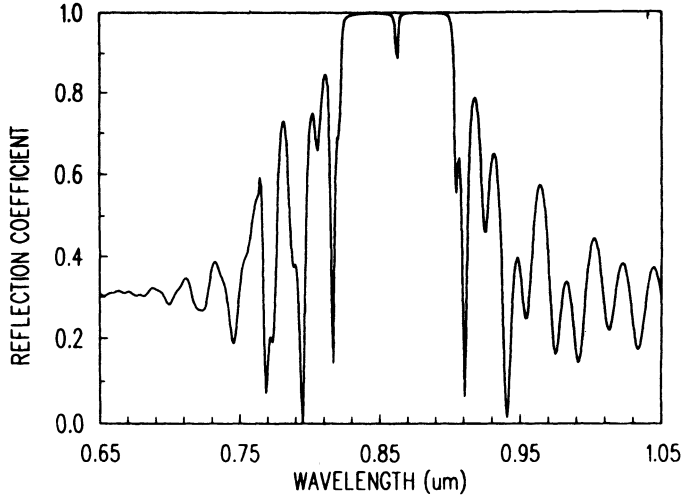


Fig. 10.5 Reflectivity of an SEL wafer consisting of top and bottom DBR mirrors similar to that shown in Fig. 10.3.

results in potential barriers at heterointerfaces. These potential barriers impede the flow of carriers through a Bragg-reflector mirror stack, resulting in a high-resistance device. It turns out that the use of graded heterobarrier interfaces in place of an abrupt interface reduces the series resistance significantly without compromising the reflectivity.[8]

10.3 GaAs-AlGaAs AND InGaAs-GaAs SURFACE-EMITTING LASERS

Much of the work on surface-emitting lasers has been carried out using the GaAs-AlGaAs material system. The GaAs SELs ($\lambda \sim 0.85$) have either a thick (~ 0.1 μm) or thin (quantum-well) active region. The InGaAs SELs ($\lambda \sim 1$ μm) grown over a GaAs substrate generally have a multiquantum-well active region. Several types of lasers have been fabricated by using this system. They generally differ in the way in which high-reflectivity mirrors and current confinement are achieved.

Five principle structures used for SEL fabrication are (i) etched-mesa structure,[10] (ii) ion-implanted structure,[11] (iii) dielectric isolated structure,[12] (iv) buried heterostructure,[13] and (v) metallic-reflector structure.[14] Schematics of these devices are shown in Fig. 10.6 for the case of a GaAs active region. Since the substrate is absorbing in this case, light is generaly emitted from the top. It can also be emitted from the bottom by using a structure in which the substrate near the emitting region has been etched away. In the case of InGaAs lasers ($\lambda \sim 1$ μm) grown over a GaAs substrate, light is not absorbed by the substrate, and can therefore be collected from the substrate side (Fig. 10.7). Jewell and coworkers[10] first reported excellent performance for such lasers using the etched-mesa structure with an InGaAs active region. Etched mesas are typically a few micrometers in diameter (3 to 5 μm) which allows fabrication of a large number of lasers on a single substrate. Critical to the fabrication of this device is mesa etching with damage-free surfaces. Jewell and coworkers used a reactive-ion etching process for their devices.[3] However, for practical use a suitable bonding scheme is required. A polyimide can be used to fill up the region around the etched mesas,[15] to allow a practical bonding scheme and subsequent CW and high-speed testing.

In addition to very high reflectivity mirrors, the key element necessary for the fabrication of good etched-mesa SELs is the use of an etching process that produces surfaces with very little nonradiative recombination. Threshold currents of 0.7 mA and 2 mA have been reported for InGaAs-GaAs ($\lambda \sim 1$ μm)[5] and GaAs-AlGaAs ($\lambda \sim 0.85$ μm)[16] etched-mesa lasers respectively. The L-I characteristics of a low-threshold etched-mesa-type SEL are shown in Fig. 10.8. Modulation bandwidths in excess of 2 GHz have been realized for this structure.

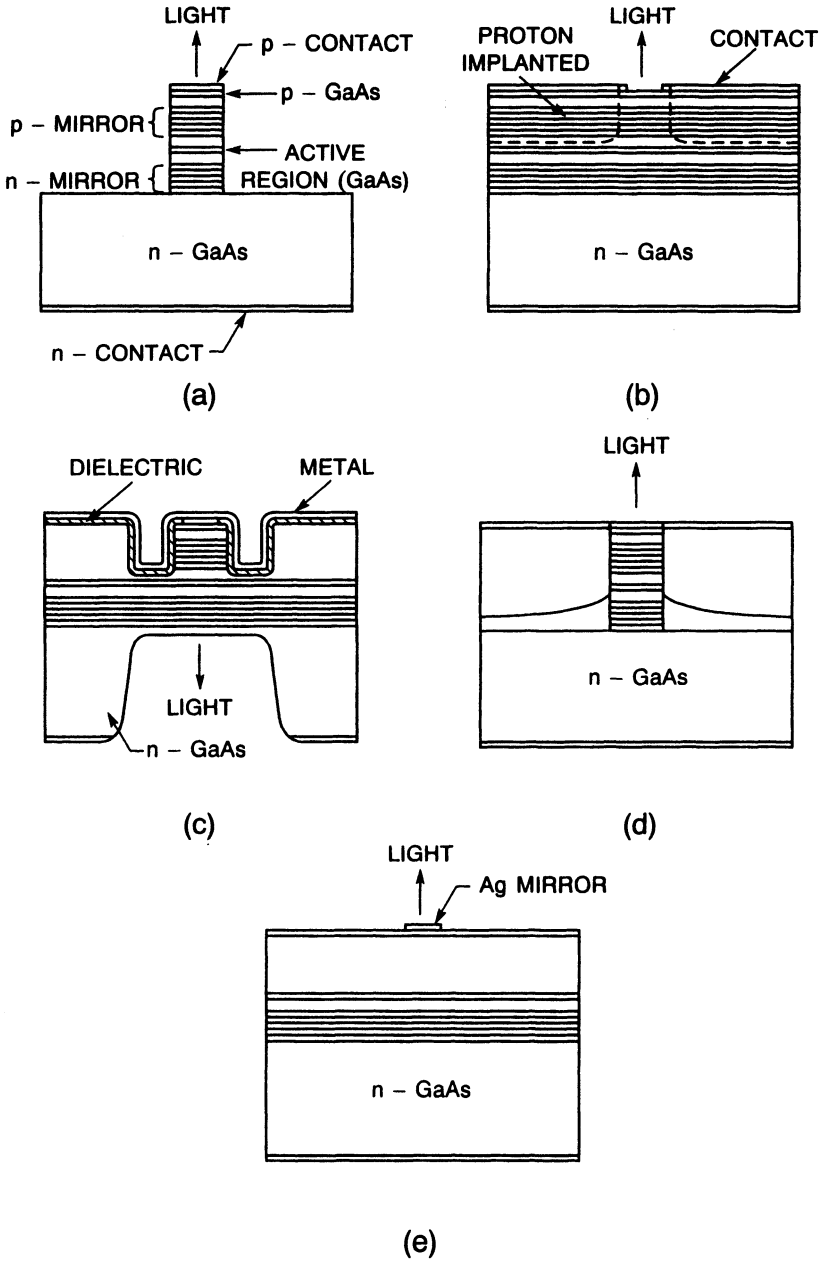


Fig. 10.6 GaAs-AlGaAs SEL structures: (a) etched-mesa structure, (b) ion-implanted top-emitting structure, (c) dielectric isolated structure, (d) buried heterostructure, and (e) metallic-reflector structure..

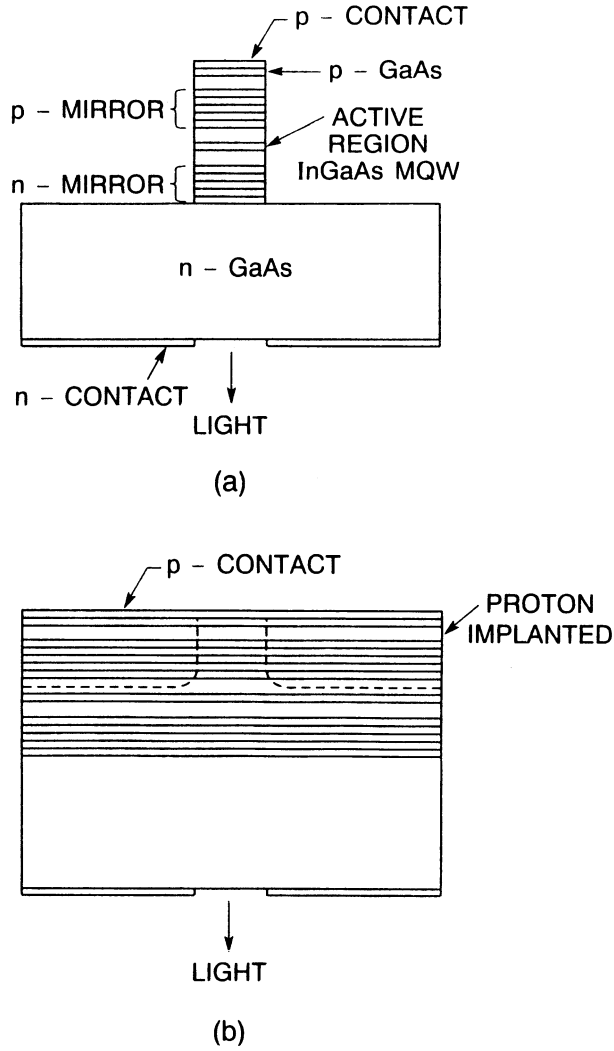


Fig. 10.7 InGaAs-GaAs SEL structures. The *n*-GaAs substrate is transparent to the emitted light. Hence the light can be collected from the substrate side. (a) Etched-mesa structure and (b) ion-implanted structure.

Top-emitting GaAs ($\lambda \sim 0.85$ μm) SELs of the type shown in Fig. 10.6b have been studied extensively. Proton implantation is generally used to produce a region of high resistivity around a 10-μm diameter opening. This scheme reduces current spreading, producing a region of high gain at the center of the opening where the lasing action takes place. Typically, laser

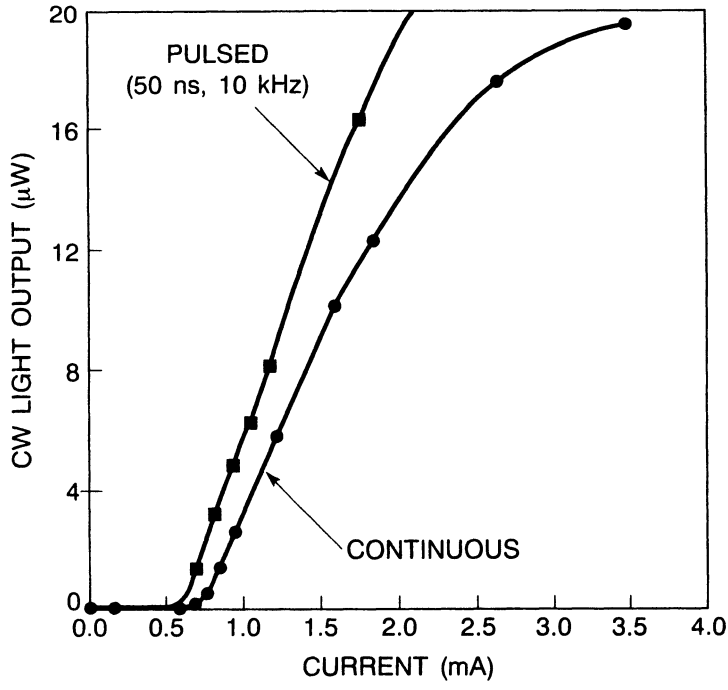


Fig. 10.8 L-I characteristics of a low-threshold InGaAs-GaAs SEL. Calculated top and bottom mirror reflectivities of this structure were 0.99998 and 0.9998 respectively. The light is emitted from the bottom. (After Ref. 5 © 1991 IEEE)

structures are designed to have 99.9% reflectivity for the substrate-side multilayer mirror and 99–99.5% reflectivity for the top mirror. These lasers have operated over a wide temperature range. CW and pulsed L-I characteristics of top-emitting GaAs SELs at different temperatures are shown in Fig. 10.9. Also shown is the single-wavelength output spectrum of the laser output showing a single longitudinal mode with a mode-suppression ratio in excess of 30 dB. The single-mode behavior arises from the short FP cavity. The cavity length is so short (~ 1 μm) that only one longitudinal cavity mode fits within the gain spectrum. Under pulsed operation both the output power and the operating temperature are higher, indicating the importance of heat dissipation in SELs. The rollover in the L-I curve of SELs is primarily due to heating which arises from a high series resistance (typically $> 100\ \Omega$) and a voltage drop at threshold of these devices.[17,18]

For some SELs, the threshold current under pulsed operation is found to be higher than that for CW operation. This effect has been explained by heating-induced self-focusing of the optical mode which results in higher

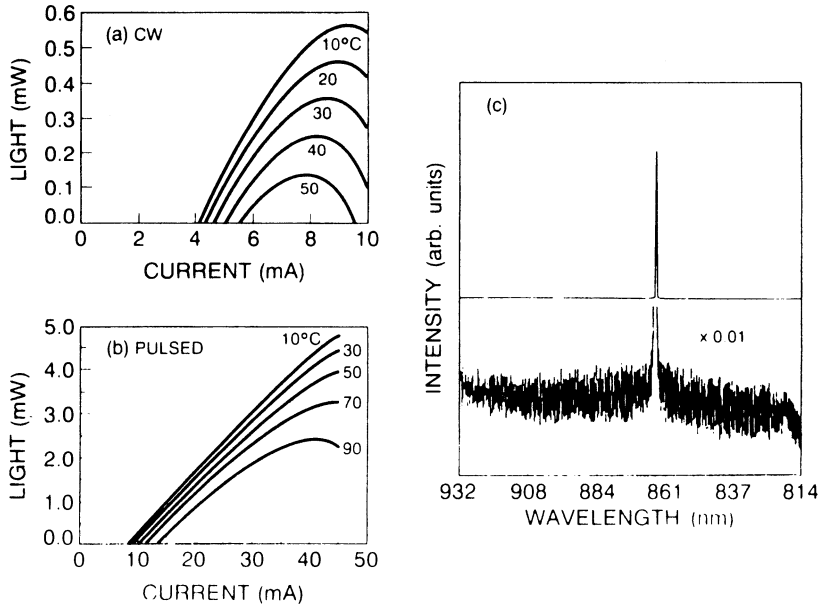


Fig. 10.9 L-I characteristics under (a) CW operation and (b) pulse operation, and (c) spectrum of a proton-implanted GaAs-AlGaAs surface-emitting laser. The higher output power of this device compared to that shown in Fig. 10.8 is due to the lower reflectivity of the output mirror. (After Ref. 17 © 1991 IEEE)

modal gain at a given current for CW operation than that for pulsed operation.[19]

The ion-implanted planar SELs can be modulated at high speeds. The measured small-signal modulation response at different output powers is shown in Fig. 10.10. As discussed in Sec. 6.6, the bandwidth is found to be proportional to the square root of the output power. Under gain-switched conditions pulses as short as 22 ps have been obtained.[20]

A variant of the SEL structures makes use of a combination of metallic reflector and a multilayer Bragg mirror on the epitaxial side for achieving high reflectivity.[21] These devices have the advantage of somewhat lower resistance and are more tolerant of epitaxial growth variations simply because fewer layers are needed to form the mirror. Pulsed operation using a silver metal film as a reflector (without any multilayer Bragg mirror) on the *p*-side has also been achieved.[13]

In a modification of the SEL structure, the quantum-well gain regions are located at the maximum of the standing-wave pattern in the FP cavity. This structure is known as the resonant periodic-gain structure.[22,23] It allows for the maximum coupling between the gain region and the optical field, resulting in a very low threshold for such devices.

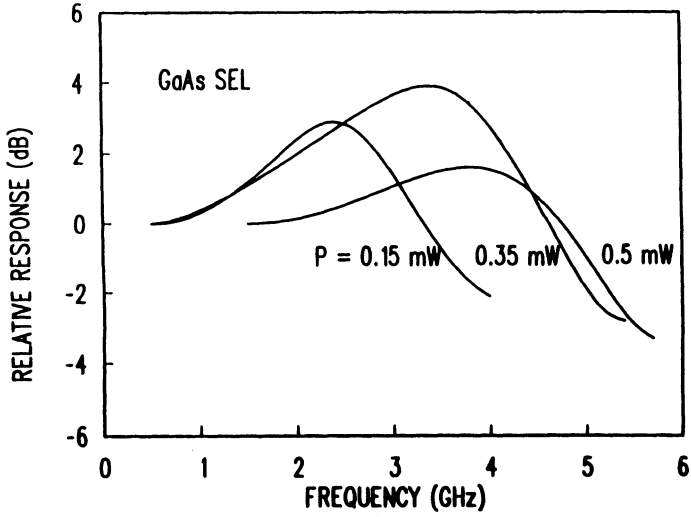


Fig. 10.10 Small-signal modulation characteristics of a proton-implanted GaAs-AlGaAs SEL.

10.4 InGaAsP-InP SURFACE-EMITTING LASERS

As mentioned previously, the refractive-index difference between InGaAsP ($\lambda \sim 1.3$ μm) and InP layers is smaller than in GaAs SELs, hence InGaAsP-InP SELs use more pairs (typically 40–50) to produce a high-reflectivity

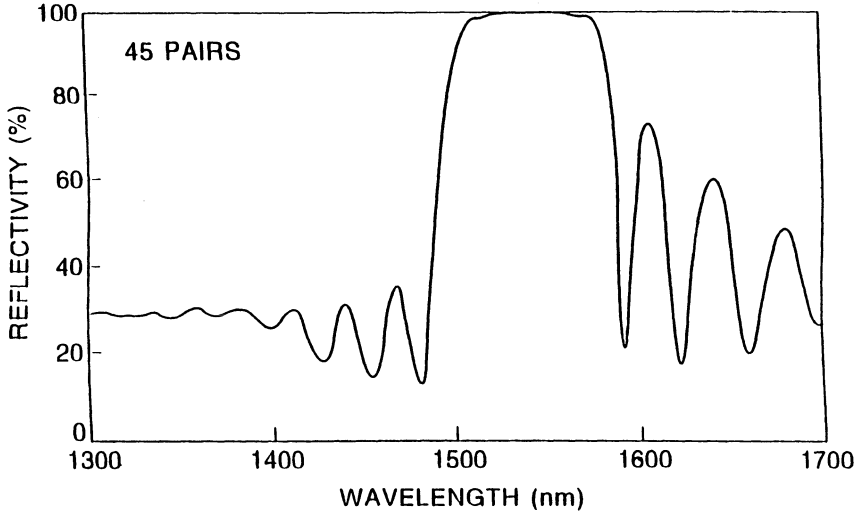


Fig. 10.11 Measured reflectivity spectrum of a 45-pair InGaAsP (121.6-nm thick, $\lambda_g = 1.45$ μm) and InP (134.8 nm) DBR stack. (After Ref. 26)

(>99%) mirror. Such mirror stacks have been grown by chemical-beam epitaxy (CBE), metal-organic chemical vapor deposition (MOVPE), and molecular-beam epitaxy (MBE) growth techniques.[24,25] The measured reflectivity of a 45-pair quarter-wave stack grown by CBE is shown in Fig. 10.11. The mirror stack was designed for maximum reflectivity around 1.6 μm.[26] Such mirror stacks along with a dielectric Si-SiO_2 stack on the *p*-side have been used to fabricate InGaAsP SELs.[27] The structure and the

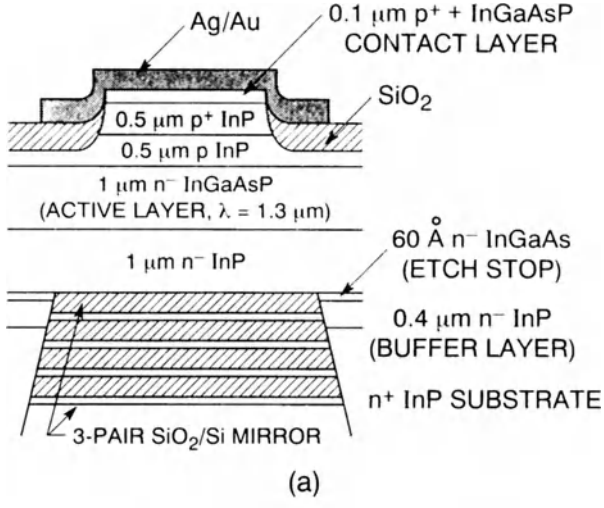


Fig. 10.12 (a) Structure and (b) L-I characteristics of an InGaAsP-InP SEL. (After Ref. 27)

L-I characteristics of such a device are shown in Fig. 10.12. The threshold current of these lasers is generally higher than that for cleaved-facet lasers. As mirror fabrication technology develops further, a reduction in threshold current is likely to follow.

10.5 LASER ARRAYS

Edge-emitting phase-locked laser arrays have been extensively studied for high-power operation (see Sec. 5.7). An often-encountered problem in the performance of an edge-emitting laser array is that the far field along the junction plane has two lobes when neighboring emitters oscillate out of phase. This out-of-phase oscillation is favored for gain-guided arrays because there is no gain between emitters.

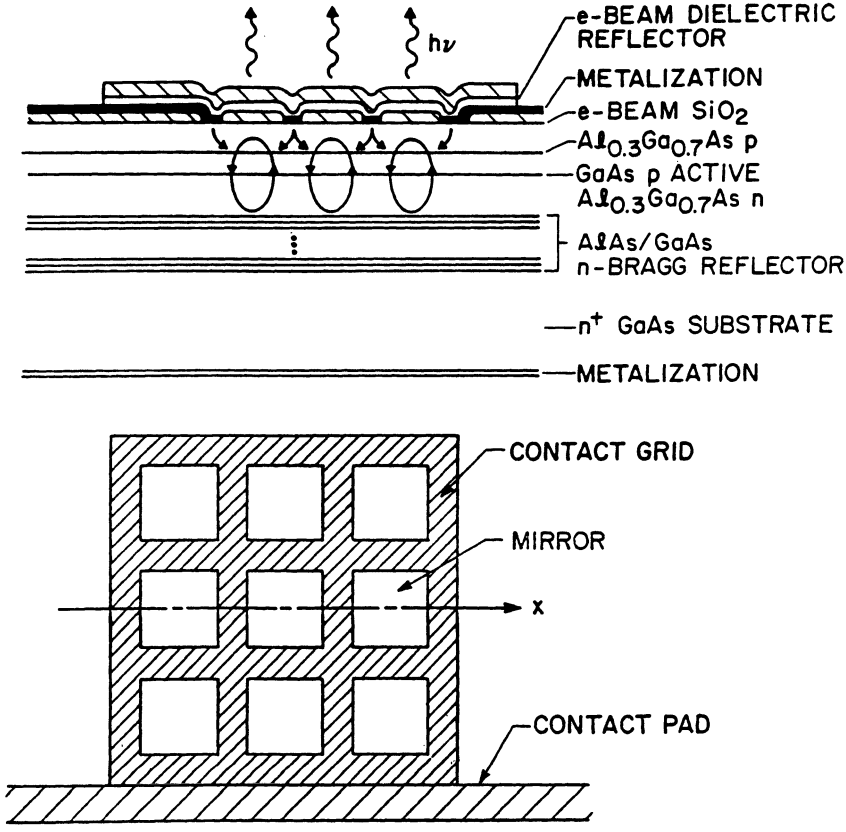


Fig. 10.13 Schematic of an SEL array. (After Ref. 28)

It is possible to design surface-emitting laser arrays such that the neighboring emitters oscillate in phase by design,[28,29] and hence the device has a single-lobe far field. In such a device (see Fig. 10.13), current flows in the region between the mirrors and hence the gain is maximum between the mirrors which favors the in-phase operation.

For many applications of SEL arrays, each laser in the array should be biased and controlled individually. Independently addressable laser arrays have been fabricated.[30] It is even possible to design multiwavelength SEL arrays such that each laser operates at a slightly different wavelength using a growth technique where the Bragg wavelength of mirror stacks is varied across a wafer. Such laser arrays have been used in transmission experiments.[31]

PROBLEMS

10.1 What techniques can be used to deflect the output of a conventional semiconductor laser normal to the surface of the wafer? Use diagrams as necessary.

10.2 Sketch the design of a vertical-cavity surface-emitting laser. What are the implications of an extremely short cavity length (~ 1 μm)?

10.3 Vertical-cavity surface-emitting lasers require Bragg mirrors of very high reflectivity ($> 99\%$). Why is this so? What happens if the reflectivity is only 80%.

10.4 The spontaneous-emission factor β_{sp} of vertical-cavity surface-emitting lasers is considerably enhanced because the microcavity of such lasers supports a much smaller number of modes. Use Eq. (6.2.22) to discuss the implications of this effect on the noise and modulation characteristics of such lasers. Do you expect an enhancement in the spontaneous-emission rate R_{sp}?

10.5 What is the value of the spontaneous-emission factor β_{sp} if the cavity of a surface-emitting laser is designed to support a single mode? When does such a laser reach threshold?

REFERENCES

1. Soda, H., K. Iga, C. Kitahara, and Y. Suematsu. *Jpn. J. Appl. Phys.* **18,** 2329 (1979).
2. Iga, K., F. Koyama, and S. Kinoshita. *IEEE J. Quantum Electron.* **QE-24,** 1845 (1988).
3. Jewell, J. L., J. P. Harbison, A. Scherer, Y. H. Lee, and L. T. Florez. *IEEE J. Quantum Electron.* **QE-27,** 1332 (1991).
4. Chang-Hasnain, C. J., M. W. Maeda, N. G. Stoffel, J. P. Harbison, and L. T. Florez. *Electron. Lett.* **26,** 940 (1990).
5. Geels, R. S., S. W. Corzine, and L. A. Coldren. *IEEE J. Quantum Electron.* **QE-27,** 1359 (1991).

6. Geels, R. S., and L. A. Coldren. *Appl. Phys. Lett.* **57,** 1605 (1990).
7. Tai, K., G. Hasnain, J. D. Wynn, R. J. Fischer, Y. H. Wang, B. E. Weir, J. Gamelin, and A. Y. Cho. *Electron. Lett.* **26,** 1628 (1990).
8. Tai, K., L. Yang, Y. H. Wang, J. D. Wynn, and A. Y. Cho. *Appl. Phys. Lett.* **56,** 2496 (1990).
9. Born, M., and E. Wolf. Sec. 1.6.5, p. 69 in *Principles of Optics.* New York: Pergamon Press, 1977.
10. Jewell, J. L., A. Scherer, S. L. McCall, Y. H. Lee, S. J. Walker, J. P. Harbison, and L. T. Florez. *Electron. Lett.* **25,** 1123 (1989).
11. Lee, Y. H., B. Tell, K. F. Brown-Goebeler, J. L. Jewell, R. E. Leibenguth, M. T. Asom, G. Livescu, L. Luther, and V. D. Mattera. *Electron. Lett.* **26,** 1308 (1990).
12. Tai, K., R. J. Fischer, C. W. Seabury, N. A. Olsson, D. T. C. Huo, Y. Ota, and A. Y. Cho. *Appl. Phys. Lett.* **55,** 2473 (1989).
13. Ibaraki, A., K. Kawashima, K. Furusawa, T. Ishikawa, T. Yamayachi, and T. Niina. *Jpn. J. Appl. Phys.* **28,** L667 (1989).
14. Schubert, E. F., L. W. Tu, R. F. Kopf, G. J. Zydzik, and D. G. Deppe. *Appl. Phys. Lett.* **57,** 117 (1990).
15. Geels, R. S., S. W. Corzine, J. W. Scott, D. B. Young, and L. A. Coldren. *IEEE Photon. Technol. Lett.* **2,** 234 (1990).
16. Choquette, K., G. Hasnain, Y. H. Wang, J. D. Wynn, R. S. Freund, A. Y. Cho, and R. E. Leibenguth. *IEEE Photon. Technol. Lett.* **3,** 859 (1991).
17. Hasnain, G., K. Tai, L. Yang, Y. H. Wang, R. J. Fischer, J. D. Wynn, B. E. Weir, N. K. Dutta, and A. Y. Cho. *IEEE J. Quantum Electron.* **QE-27,** 1377 (1991).
18. Gourley, P. L., S. K. Lyo, T. M. Brennan, B. E. Hammons, C. P. Schaus, and S. Sun. *Appl. Phys. Lett.* **55,** 2698 (1989).
19. Dutta, N. K., L. W. Tu, G. J. Zydzik, G. Hasnain, Y. H. Wang, and A. Y. Cho. *Electron. Lett.* **27,** 208 (1991).
20. Hasnain, G. K. Tai, N. K. Dutta, Y. H. Wang, J. D. Wynn, B. E. Weir, and A. Y. Cho. *Electron. Lett.* **27,** 915 (1991).
21. Hasnain, G., K. Tai, J. D. Wynn, Y. H. Wang, R. J. Fischer, M. Hong, B. E. Weir, G. J. Zydzik, J. P. Mannaerts, J. Gamelin, and A. Y. Cho. *Electron. Lett.* **26,** 1590 (1990).
22. Raja, M. Y. A., S. R. J. Brueck, M. Osinski, C. F. Schaus, J. G. McInerney, T. M. Brennan, and B. E. Hammons. *IEEE J. Quantum Electron.* **QE-25,** 1500 (1989).
23. Corzine, S. W., R. S. Geels, J. W. Scott, R. H. Yan, and L. A. Coldren. *IEEE J. Quantum Electron.* **QE-25,** 1513 (1989).
24. Tai, K., S. L. McCall, S. N. G. Chu, and W. T. Tsang. *Appl. Phys. Lett.* **51,** 826 (1987).
25. Tai, K., K. F. Huang, R. J. Fischer, and A. Y. Cho. *Electron. Lett.* **25,** 1159 (1989).
26. Choa, F. S., K. Tai, W. T. Tsang, and S. N. G. Chu. *Appl. Phys. Lett.* **22,** 2820 (1991).
27. Yang, L., M. C. Wu, K. Tai, T. Tanbun-Ek, and R. A. Logan. *Appl. Phys. Lett.* **56,** 889 (1990).
28. Deppe, D. G., J. P. van der Ziel, N. Chand, G. J. Zydzik, and S. N. G. Chu. *Appl. Phys. Lett.* **56,** 740 (1990).
29. van der Ziel, J. P., D. G. Deppe, N. Chand, G. J. Zydzik, and S. N. G. Chu. *IEEE J. Quantum Electron.* **QE-26,** 1873 (1990).
30. Chang-Hasnain, C. J., J. P. Harbison, C. E. Zah, M. W. Maeda, L. T. Florez, N. G. Stoffel, and T. P. Lee. *IEEE J. Quantum Electron.* **QE-27,** 1368 (1991).
31. Maeda, M. W., C. Chang-Hasnain, A. von Lehman, H. Izadpanah, C. Lin, M. Z. Iqbal, L. Florez, and J. Harbison. *IEEE Photon. Technol. Lett.* **3,** 863 (1991).

Chapter 11
OPTICAL AMPLIFIERS

11.1 INTRODUCTION

Optical amplifier, as the name implies, is a device that amplifies an input optical signal. The amplification factor or gain can be higher than 1,000 (>30 dB) in some devices. There are two principal types of optical amplifier: the semiconductor-laser amplifier (SLA), and the fiber amplifier. In an SLA, light is amplified when it propagates through a semiconductor medium fabricated in the form of a waveguide. In a fiber amplifier, light is amplified when it travels through a fiber doped with rare-earth ions (such as Nd^+, Er^+, etc.). SLAs are typically less than 1 mm in length whereas fiber amplifiers are typically 1–100 m in length. The operating principals, design, fabrication and performance characteristics of SLAs and fiber amplifiers are described in this chapter.

In a lightwave transmission system, as the optical signal travels through the fiber, it weakens and gets distorted. Regenerators are used to restore the optical pulses to their original form. Figure 11.1a shows the block diagram of a typical lightwave regenerator. Its main components are an optical receiver, an optical transmitter, and electronic timing and decision circuits. Optical amplifiers can nearly restore the original optical pulses. This increases their transmission distance without using conventional regenerators. An example of a semiconductor amplifier that functions as a regenerator is shown schematically in Fig. 11.1b. Semiconductor amplifiers need external current to produce gain and fiber amplifiers need pump lasers for the same purpose. Because of its simplicity, an optical amplifier is an attractive alternative for a new lightwave system.

11.2 GENERAL CONCEPTS

Optical amplifiers amplify incident light through stimulated emission, the same mechanism used by lasers. Indeed, an optical amplifier is nothing but

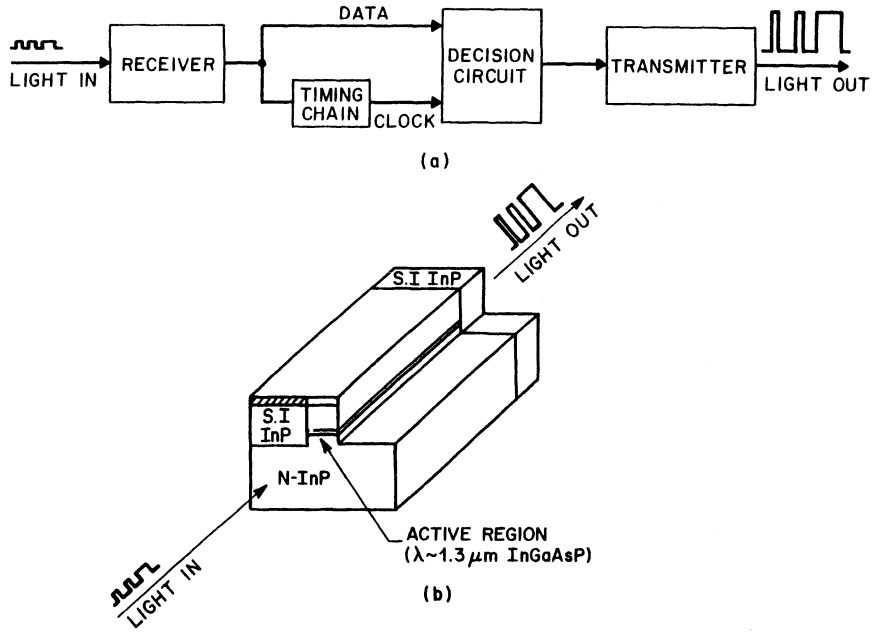


Fig. 11.1 (a) Block diagram of a lightwave regenerator. (b) Schematic of a semiconductor laser amplifier.

a laser without feedback. Its main ingredient is the optical gain realized when the amplifier is pumped (optically or electrically) to achieve population inversion. In general, the optical gain depends on the frequency (or wavelength) of the incident signal, and on the local beam intensity at any point inside the amplifier. Details of the frequency and intensity dependence of the optical gain depend on the amplifier medium. To illustrate the general concepts, let us consider the case in which the gain medium is modeled as a homogeneously broadened two-level system. The *gain coefficient* of such a medium can be written as

$$g(\omega) = \frac{g_0}{1 + (\omega - \omega_0)^2 T_2^2 + P/P_s} \tag{11.2.1}$$

where g_0 is the peak value of the gain determined by the pumping level of the amplifier, ω is the optical frequency of the incident signal, ω_0 is the atomic transition frequency, and P is the optical power of the signal being amplified. The *saturation power* P_s depends on the gain-medium parameters such as the fluorescence time T_1 and the transition cross section; its

expression for different kinds of amplifiers will be given in the following sections. The parameter T_2 in Eq. (11.2.1) is known as the dipole relaxation time and is typically quite small (0.1 ps–1 ps). The fluorescence time T_1 is also called the population relaxation time and varies in the range 100 ps–10 ms depending on the gain medium used to make the amplifier. Equation (11.2.1) can be used to discuss important characteristics of optical amplifiers, such as the gain bandwidth, amplification factor, and the output saturation power. We begin by considering the case in which $P/P_s \ll 1$ throughout the amplifier. This is referred to as the *unsaturated region*, as the gain remains unsaturated during amplification.

11.2.1 Gain Spectrum and Bandwidth

By neglecting the term P/P_s in Eq. (11.2.1), the gain coefficient is given by

$$g(\omega) = \frac{g_0}{1 + (\omega - \omega_0)^2 T_2^2}. \tag{11.2.2}$$

This equation shows that the gain is maximum when the incident frequency ω coincides with the atomic transition frequency ω_0. The gain reduction for $\omega \neq \omega_0$ is governed by a Lorentzian profile that is a characteristic of homogeneously broadened two-level systems. As discussed later, the gain spectrum of actual amplifiers can deviate considerably from the Lorentzian profile. The *gain bandwidth* is defined as the full width at half maximum (FWHM) of the gain spectrum $g(\omega)$. For the Lorentzian spectrum the gain bandwidth is given by $\Delta\omega_g = 2/T_2$ or by

$$\Delta\nu_g = \frac{\Delta\omega_g}{2\pi} = \frac{1}{\pi T_2}. \tag{11.2.3}$$

As an example, $\Delta\nu_g \sim 3$ THz for semiconductor laser amplifiers for which $T_2 \sim 0.1$ ps. Amplifiers with a relatively large bandwidth are preferred for optical communication systems, since the gain is then nearly constant over the entire bandwidth of even a multichannel signal.

A related concept of *amplifier bandwidth* is commonly used in place of the gain bandwidth. The difference becomes clear when one considers the *amplifier gain* G, also known as the *amplification factor* and defined as

$$G = P_{\text{out}}/P_{\text{in}} \tag{11.2.4}$$

where P_{in} and P_{out} are the input and output powers of the continuous-wave

(CW) signal being amplified. We can obtain an expression for G by using

$$\frac{dP}{dz} = gP \tag{11.2.5}$$

where $P(z)$ is the optical power at a distance z from the input end. A straightforward integration with the initial condition $P(0) = P_{in}$ shows that the signal power grows exponentially as

$$P(z) = P_{in} \exp(gz). \tag{11.2.6}$$

By noting that $P(L) = P_{out}$ and using Eq. (11.2.4), the amplification factor for an amplifier of length L is given by

$$G(\omega) = \exp[g(\omega)L] \tag{11.2.7}$$

where the frequency dependence of both G and g is shown explicitly. Both the amplifier gain $G(\omega)$ and the gain coefficient $g(\omega)$ are maximum when $\omega = \omega_0$ and decrease with the signal detuning $\omega - \omega_0$. However, $G(\omega)$ decreases much faster than $g(\omega)$ because of the exponential dependence of G on g. The amplifier bandwidth $\Delta\nu_A$ is defined as the FWHM of $G(\omega)$ and is related to the gain bandwidth $\Delta\nu_g$ as

$$\Delta\nu_A = \Delta\nu_g \left(\frac{\ln 2}{g_0 L - \ln 2} \right). \tag{11.2.8}$$

As expected, the amplifier bandwidth is smaller than the gain bandwidth, and the difference depends on the amplifier gain itself.

11.2.2 Gain Saturation

The origin of gain saturation lies in the power dependence of the gain coefficient in Eq. (11.2.1). When $P \ll P_s$, $g(\omega)$ reduces to Eq. (11.2.2) and is referred to as the *small-signal gain* since the incident signal power should be small and remain small during amplification. Since g is reduced when P becomes comparable to P_s, the amplification factor G is also expected to decrease. To simplify the discussion, let us consider the case in which incident signal frequency is exactly tuned to the atomic transition frequency ω_0 to maximize the small-signal gain. The detuning effects can be incorporated in a straightforward manner: by substituting g from Eq. (11.2.1) in Eq. (11.2.5),

we obtain

$$\frac{dP}{dz} = \frac{g_0 P}{1 + P/P_s}. \tag{11.2.9}$$

This equation can be easily integrated over the amplifier length. By using the initial condition $P(0) = P_{in}$ together with $P(L) = P_{out} = GP_{in}$, we obtain the following implicit relation for the large-signal amplifier gain:

$$G = G_0 \exp\left(-\frac{G-1}{G}\frac{P_{out}}{P_s}\right) \tag{11.2.10}$$

where $G_0 = \exp(g_0 L)$ is the unsaturated value of the amplification factor ($P_{out} \ll P_s$).

Equation (11.2.10) shows that the amplification factor G decreases from its unsaturated value G_0 when P_{out} becomes comparable to P_s. A quantity of practical interest is the *output saturation power* P_{out}^s, defined as the output power for which the amplifier gain G is reduced by a factor of 2 (or by 3 dB) from its unsaturated value G_0. By using $G = G_0/2$ in Eq. (11.2.10), P_{out}^s is given by

$$P_{out}^s = \frac{G_0 \ln 2}{G_0 - 2} P_s. \tag{11.2.11}$$

P_{out}^s is smaller than P_s by about 30%. Indeed, by noting that $G_0 \gg 2$ in practice ($G_0 = 1{,}000$ for 30 dB amplifier gain) $P_{out}^s \approx (\ln 2)P_s \approx 0.69P_s$. As seen by Eq. (11.2.11), P_{out}^s becomes nearly independent of G_0 for $G_0 > 20$ dB.

11.2.3 Amplifier Noise

All amplifiers degrade the signal-to-noise ratio (SNR) of the amplified signal because of spontaneous emission that adds to the signal during its amplification. The SNR degradation is quantified through a parameter F_n, called the *amplifier noise figure* in analogy with the electronic amplifiers, and defined as

$$F_n = \frac{(\text{SNR})_{in}}{(\text{SNR})_{out}} \tag{11.2.12}$$

where SNR refers to the electrical power generated when the signal is converted to electrical current by using a photodetector. In general, F_n would depend on several detector parameters which govern the shot noise and

thermal noise associated with the detector. One can obtain a simple expression for F_n by considering an ideal detector whose performance is limited by shot noise only.

Consider an amplifier with the amplification factor G so that the output power is related to the input power by $P_{out} = GP_{in}$. The SNR of the input signal is given by

$$(\text{SNR})_{in} = \frac{\langle I \rangle^2}{\sigma_s^2} = \frac{(\text{RP}_{in})^2}{2q(\text{RP}_{in})\Delta f} = \frac{P_{in}}{2h\nu\Delta f} \tag{11.2.13}$$

where $\langle I \rangle = \text{RP}_{in}$ is the average photocurrent, $R = q/h\nu$ is the responsivity of an ideal photodetector with unit quantum efficiency, and

$$\sigma_s^2 = 2q(\text{RP}_{in})\Delta f \tag{11.2.14}$$

is the shot noise. Here Δf is the detector bandwidth. To evaluate the SNR of the amplified signal, one should add the contribution of spontaneous emission to the receiver noise.

The spectral density of spontaneous-emission-induced noise is nearly constant (white noise) and can be written as

$$S_{sp}(\nu) = (G - 1)n_{sp}h\nu \tag{11.2.15}$$

where ν is the optical frequency. The parameter n_{sp} is called the *spontaneous-emission factor* or population-inversion factor. Its value is 1 for amplifiers with complete population inversion (all atoms in the excited case), but becomes >1 when the population inversion is incomplete. For a two-level system,

$$n_{sp} = N_2/(N_2 - N_1) \tag{11.2.16}$$

where N_1 and N_2 are the atomic populations for the ground and excited states respectively. The effect of spontaneous emission is to add fluctuations to the amplified power which are converted to current fluctuations during the photodetection process.

It turns out that the dominant contribution to the receiver noise comes from beating of spontaneous emission with the signal itself. This beating phenomenon is similar to heterodyne detection in the sense that spontaneously emitted radiation mixes coherently with the amplified signal at the photodetector and produces a heterodyne component of the photocurrent. The variance of the photocurrent can then be written as

$$\sigma^2 = 2q(\text{RGP}_{in})\Delta f + 4(\text{GRP}_{in})(\text{RS}_{sp})\Delta f \tag{11.2.17}$$

where the first term is due to shot noise and the second term results from signal-spontaneous emission beating. All other contributions to the receiver noise have been neglected for simplicity. The SNR of the amplified signal is thus given by

$$(\text{SNR})_{\text{out}} = \frac{\langle I \rangle^2}{\sigma^2} = \frac{(\text{RGP}_{\text{in}})^2}{\sigma^2} \approx \frac{\text{GP}_{\text{in}}}{4S_{\text{sp}}\Delta f} \tag{11.2.18}$$

where the last relation was obtained by neglecting the first term in Eq. (11.2.17) and is valid for $G \gg 1$.

The amplifier noise figure can now be obtained by substituting Eqs. (11.2.13) and (11.2.18) in Eq. (11.2.12). If we also use Eq. (11.2.15) for S_{sp}, we obtain

$$F_{\text{n}} = 2n_{\text{sp}}(G - 1)/G \approx 2n_{\text{sp}}. \tag{11.2.19}$$

This equation shows that the SNR of the amplified signal is degraded by a factor of 2 (or 3 dB) even for an ideal amplifier for which $n_{\text{sp}} = 1$. For most practical amplifiers, F_{n} exceeds 3 dB and can be as large as 6–8 dB. For its application in optical communication systems, an optical amplifier should have F_{n} as low as possible.

11.2.4 Amplifier Applications

Optical amplifiers can serve several purposes in the design of fiber-optic communication systems. As already mentioned in the chapter's introduction, an important application for long-haul systems is in replacing electronic regenerators with optical amplifiers. Such a replacement can be carried out as long as the system performance is not limited by the cumulative effects of dispersion and spontaneous emission. The use of optical amplifiers is particularly attractive for multichannel lightwave systems, since electronic regeneration requires demultiplexing of channels before each channel signal is regenerated using separate receiver and transmitters, a rather costly procedure. Optical amplifiers can amplify all channels simultaneously. When optical amplifiers are used to replace electronic regenerators, they are called *in-line amplifiers.*

Another way to use optical amplifiers is to increase the transmitter power by placing an amplifier just after the transmitter. Such amplifiers are called *power amplifiers* or boosters, as their main purpose is to boost the transmitted power. A power amplifier increases the transmission distance by 10–100 km depending on the amplifier gain and the fiber loss. Transmission distance can also be increased by putting an amplifier just before the receiver

to boost the received power. Such amplifiers are called *preamplifiers.* Another application of optical amplifiers is to use them for compensating distribution losses in local-area networks. Distribution losses often limit the number of nodes in a network, particularly in the case of bus topology. Further applications use optical amplifiers in photonic switching systems.

11.3 SEMICONDUCTOR LASER AMPLIFIERS

The SLA is a device very similar to a semiconductor laser. Hence its operating principle, fabrication and design are also similar. The basic laser structure can be used to study light amplification. When the injection current is below threshold, the laser acts as an optical amplifier for incident light waves, and above threshold it undergoes oscillation. Initial optical amplifier studies were carried out on GaAs homostructure devices in the mid 1960s.[1–3] Extensive work on AlGaAs laser amplifiers was carried out in the 1980s.[4–7] These amplifiers used an index guiding structure and are therefore closer to a practical device.

SLAs can be classified into two categories: the Fabry-Perot (FP) amplifier, and the traveling-wave (TW) amplifier. An FP amplifier has considerable reflectivity at the input and output ends, resulting in resonant amplification between the end mirrors. Thus, an FP amplifier exhibits very large gain at wavelengths corresponding to the longitudinal modes of the FP cavity. The TW amplifier, by contrast, has negligible reflectivity at each end, resulting in signal amplification during a single pass. The optical gain spectrum of a TW amplifier is quite broad and corresponds to that of the semiconductor gain medium. Most practical TW amplifiers exhibit some small ripple in the gain spectrum arising from residual facet reflectivities. TW amplifiers are more suitable for system applications. Therefore much effort has been devoted over the last few years to fabricate amplifiers with very low facet reflectivities. Such amplifier structures either utilize special low-reflectivity dielectric coatings, or have tilted or buried facets. Fabrication and performance of these devices are described later.

Much of the experimental work on semiconductor optical amplifiers has been carried out using the InGaAsP material system with the optical gain centered around 1.3 μm or 1.55 μm.[6–24] The interest in these wavelength regions is primarily due to low loss and low dispersion of silica fibers which are extensively used as the transmission medium for the optical fiber transmission systems.

The amplifiers used in lightwave system applications, either as preamplifiers in front of a receiver or as in line amplifiers as a replacement of regenerators, must also exhibit equal optical gain for all polarizations of the input light.

In general, the optical gain in a waveguide is polarization dependent although the material gain is independent of polarization for bulk semiconductors. This arises from unequal mode confinement factors for light polarized parallel to the junction plane (TE mode) and that for light polarized perpendicular to the junction plane (TM mode). For thick active regions, the confinement factors of the TE and TM mode are nearly equal. Hence the gain difference between the TE and TM modes is smaller for amplifiers with a thick active region. Figure 11.2 shows the calculated gain difference as a function of cavity length for different active layer thicknesses.[6] In general, the gain difference increases with increasing cavity length since the overall gain increases.

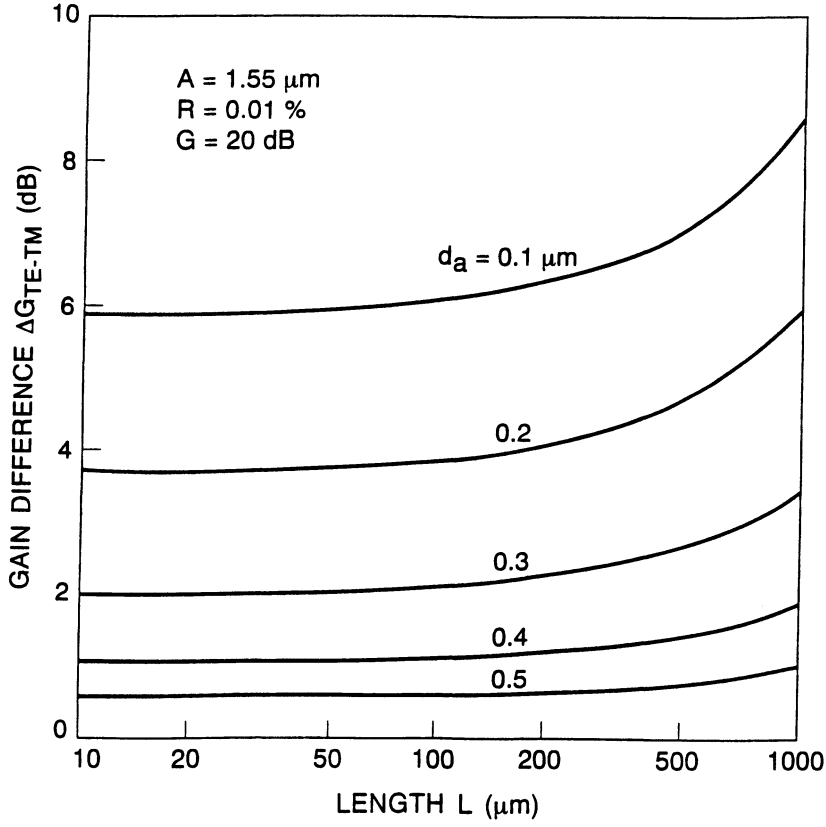


Fig. 11.2 Gain difference between the TE and TM modes of a semiconductor laser amplifier plotted as a function of the device length for several active-layer thicknesses.

11.3.1 Impact of Facet Reflectivity

In an ideal TW amplifier the optical beam should not experience any reflection at the facets. In practice, facets with antireflection (AR) coatings generally exhibit some residual reflectivity and form an optical cavity. An example of the transmission characteristics of a TW amplifier with antireflection-coated cleaved facets is shown in Fig. 11.3. The output exhibits modulations at longitudinal modes of the cavity simply because the optical gain is a few decibels higher at the modes than in between modes.

The phase and amplitude transfer functions of an SLA can be characterized by a change in phase and change in amplitude G. For an amplifier with facet reflectivities R_1, R_2, the gain G is given by[7]

$$G = \frac{(1 - R_1)(1 - R_2)G_s}{(1 + \sqrt{R_1 R_2}\, G_s)^2 + 4\sqrt{R_1 R_2}\, G_s \sin^2 \phi} \tag{11.3.1}$$

where G_s is the single-pass gain, and the phase shift ϕ is given by

$$\phi = \phi_0 + \frac{g_0 L \beta_c}{2}\left(\frac{P}{P + P_s}\right) \tag{11.3.2}$$

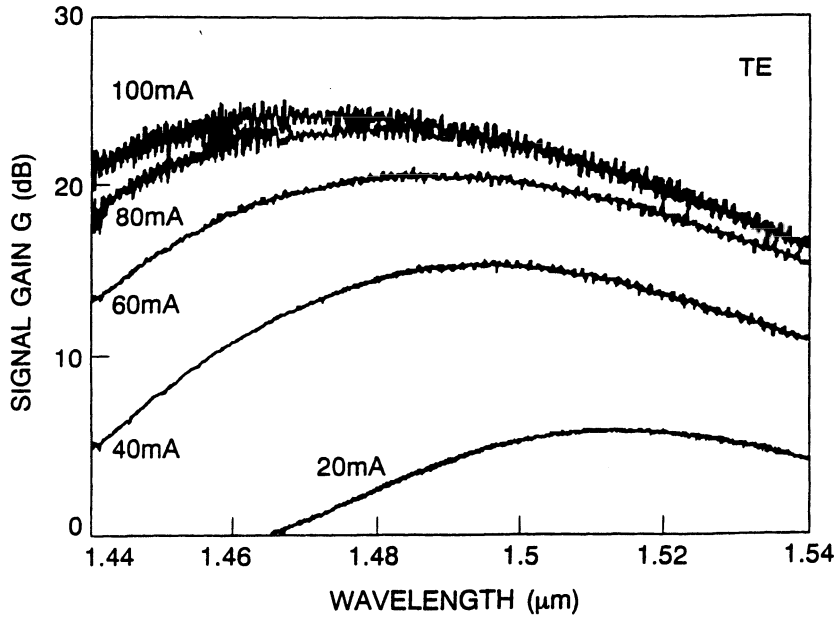


Fig. 11.3 Gain spectrum of a TW-type semiconductor laser amplifier at several current levels. (After Ref. 6 © 1988 IEEE)

where $\phi_0 = 2\pi L\mu/\lambda$ is the nominal phase shift, L is the length of the amplifier, μ is the index, β_c is the line-width enhancement factor, g_0 is the unsaturated gain, and P and P_s are the total internal power and saturation power, respectively.

The $\sin^2 \phi$ term in Eq. (11.3.1) is responsible for the modulation of the output signal at cavity modes. The case of practical interest is one of low reflectivities, i.e., R_1 and $R_2 \leq 10^{-3}$. In this case the effect of residual facet reflectivities appears as small ripples (Fig. 11.3) superimposed on an envelope function. The envelope function is essentially the gain spectrum of the semiconductor material. The peak-to-valley ratio of the intensity ripple is given by[7]

$$V = \left[\frac{1 + \sqrt{R_1 R_2}\, G_s}{1 - \sqrt{R_1 R_2}\, G_s}\right]. \tag{11.3.3}$$

For an ideal TW amplifier both R_1 and $R_2 \to 0$; V equals 1, i.e., no ripple occurs at cavity mode frequencies. The quantity V is plotted as a function of $\sqrt{R_1 R_2}$ (reflectivity) in Fig. 11.4 for different values of gain. A practical amplifier should have $V < 1$ dB. To realize such low values of V, reflectivities

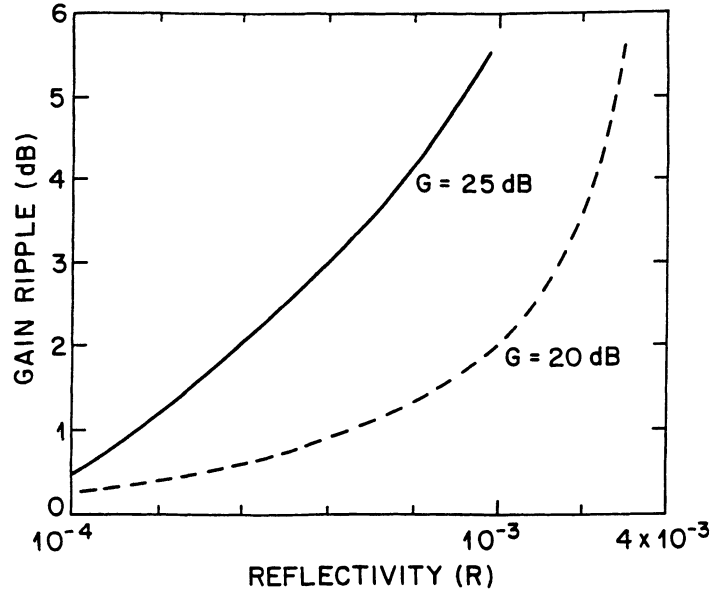


Fig. 11.4 Ripple in the gain spectrum plotted as a function of the effective reflectivity $R = \sqrt{R_1 R_2}$ for two different values of the small-signal gain. (After Ref. 6 © 1988 IEEE)

of $< 10^{-4}$ are needed. Three principal schemes exist for achievement of such low reflectivities. They are (i) ultra-low-reflectivity dielectric-coated amplifiers, (ii) buried-facet amplifiers and (iii) tilted-facet amplifiers.

11.3.2 Amplifier Designs

Low-Reflectivity Coatings. A key factor for good performance characteristics (low gain ripple and low polarization selectivity) for TW amplifiers is very low facet reflectivity. The reflectivity of cleaved facets can be reduced by dielectric coating. For plane waves incident on an air interface from a medium of refractive index n, the reflectivity can be reduced to zero by coating the interface with a dielectric whose refractive index equals $n^{1/2}$ and whose thickness equals $\lambda/4$. However, the fundamental mode propagating in a waveguide is not a plane wave and therefore the above $n^{1/2}$ law only provides a guideline for achieving very low ($\sim 10^{-4}$) facet reflectivity by dielectric coatings. In practice, very low facet reflectivities are obtained by monitoring the amplifier performance during the coating process. The effective reflectivity can then be estimated from the ripple at the FP mode spacings, caused by residual reflectivity, in the spontaneous emission spectrum. The result of such an experiment is shown in Fig. 11.5. The reflectivity is very low ($< 10^{-4}$) only in a small range of wavelengths. Laboratory experiments have been carried out using amplifiers that rely only on low-reflectivity

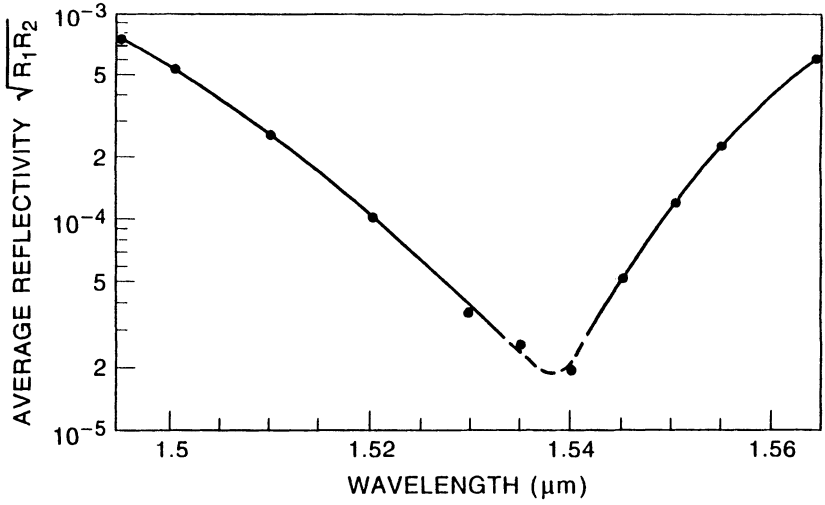


Fig. 11.5 Measured reflectivity of an antireflection-coated facet as a function of wavelength. (Courtesy N. A. Olsson)

coatings for good performance, but thickness is critical. Good antireflection coatings have a limited wavelength range, so alternatives to coatings have been investigated. Several schemes are discussed below.

Buried-Facet Amplifiers. Also known as window devices, the principal feature of buried-facet optical amplifiers relative to AR-coated cleaved-facet devices is a polarization-independent reduction in mode reflectivity. This is because the buried facet gives better control over achieving polarization-independent gain.

A schematic cross section of a buried-facet optical amplifier is shown in Fig. 11.6. Current confinement in this structure is provided by semi-insulating Fe-doped InP layers grown by metal-organic chemical vapor deposition (MOVPE). It is fabricated similar to a laser. The first four layers are shown on a (100)-oriented *n*-type InP substrate by MOVPE. These layers are (i) an *n*-type InP buffer layer, (ii) an undoped InGaAsP ($\lambda \sim 1.55$ μm) active layer, (iii) a *p*-type InP cladding layer, and (iv) a *p*-type InGaAsP ($\lambda \sim 1.3$ μm) layer. Mesas are then etched on the wafer along the [110] direction with 15-μm-wide channels normal to the mesa direction using a SiO_2 mask. The latter is needed for buried-facet formation. Semi-insulating Fe-doped InP layers are then grown around the mesas by MOVPE with the oxide mask in place. The oxide mask and *p*-type InGaAsP layer are removed and a *p*-type InP and *p*-type InGaAsP ($\lambda \sim 1.3$ μm) contact layer is then grown over the entire wafer by the MOVPE growth technique. The wafer is processed using standard methods and cleaved to produce 500-μm-long buried-facet chips with ~7-μm-long buried facets at a each end. Chip facets are then AR-coated using a single-layer film of ZrO_2. Fabrication of cleaved-facet devices follows the same procedure as described above, except that the mesas are continuous with no channels separating them. The latter

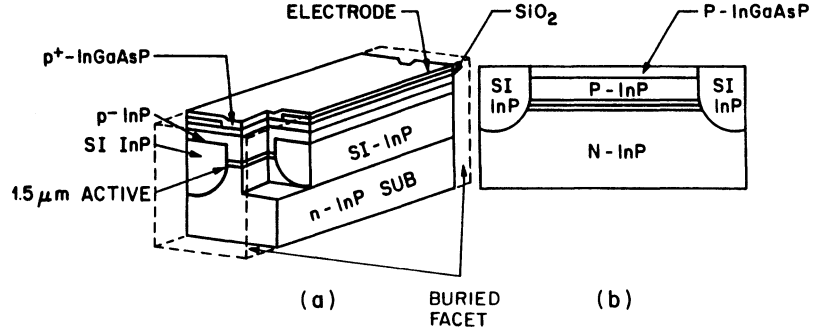


Fig. 11.6 (a) Schematic of a buried-facet optical amplifier. (b) Longitudinal cross section showing the buried facets. (After Ref. 11)

is needed for defining the buried-facet regions. The semi-insulating layer, in both types of devices, provides current confinement and lateral index guiding. For buried-facet devices it also provides the buried-facet region.

The effective reflectivity of a buried facet decreases with increasing separation between the facet and the end of the active region. The effective reflectivity of such a facet can be calculated by using a Gaussian-beam approximation for the propagating optical mode. It is given by[8]

$$R_{\text{eff}} = R/[1 + (2S/kw^2)^2] \tag{11.3.4}$$

where R is the reflectivity of the cleaved facet, S is the length of the buried-facet region, $k = 2\pi/\lambda$, where λ is the optical wavelength in the medium, and w is the spot size at the facet. The calculated reflectivity is plotted in Fig. 11.7 using $w = 0.7$ μm and $R = 0.3$ for an amplifier operating near 1.55 μm. A reflectivity of 10^{-2} can be achieved for a buried-facet length of about 15 μm.

Although increasing the length of the buried-facet region decreases the reflectivity, if the length is too long the beam emerging from the active

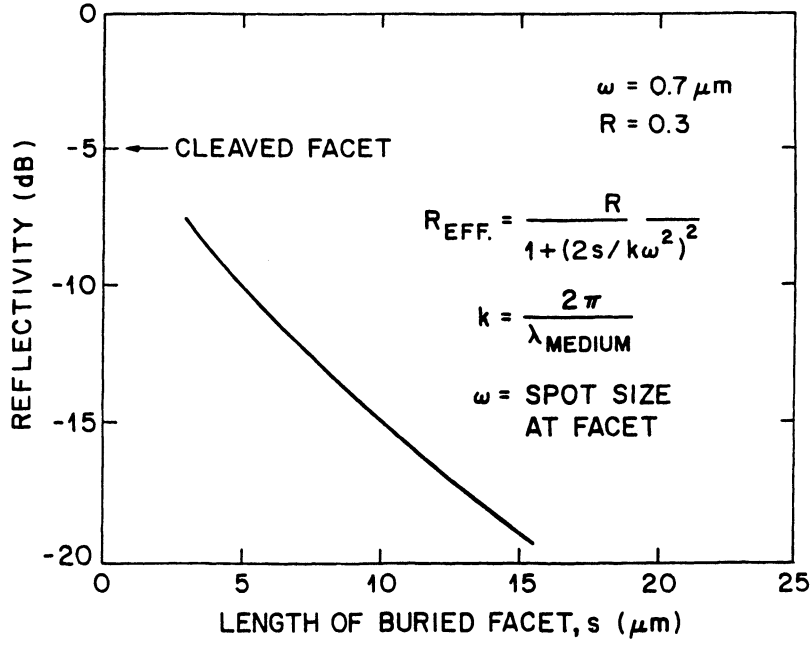


Fig. 11.7 Effective reflectivity of a buried facet calculated as a function of the buried-region thickness. (After Ref. 11)

region will strike the top metalized surface, producing multiple peaks in the far-field pattern, a feature not desirable for coupling into a single-mode fiber. The beam waist w of a Gaussian beam after traveling a distance z is given by the equation[9]

$$w^2(z) = w_0^2\left[1 + \left(\frac{\lambda z}{\pi w_0^2}\right)^2\right] \tag{11.3.5}$$

where w_0 is the spot size at the beam waist and λ is the wavelength in the medium. Since the active region is about 4 μm from the top surface of the chip, it follows from Eq. (11.3.5) that the length of the buried-facet region must be less than 12 μm for single-lobed far-field operation.

The optical gain is determined by injecting light into the amplifier and measuring the output. The internal gain of an amplifier chip as a function of current at two different temperatures is shown in Fig. 11.8.[10] Open circles and squares represent the gain for a linearly polarized incident light with the electric field parallel to the *p-n* junction in the amplifier chip (TE mode). Solid circles represent the measured gain for the TM mode at 40°C.

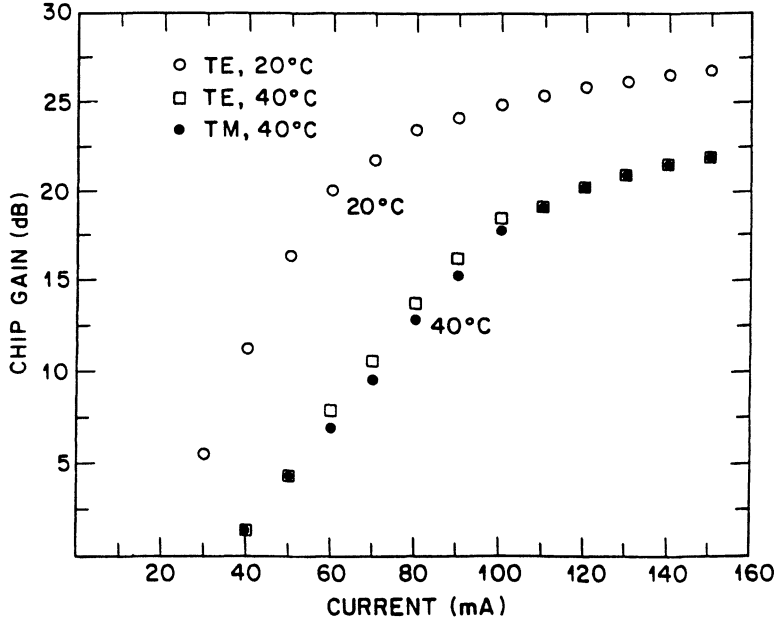


Fig. 11.8 Measured amplifier gain as a function of injection current at 20 and 40°C. Solid circles show the TM-mode gain at 40°C. (After Ref. 10)

Measurements were done for low input power (−40 dBm), so that the observed saturation is not due to gain saturation in the amplifier, but rather to carrier loss caused by Auger recombination. Note that the optical gain for the TE input polarization is nearly equal to the optical gain for TM input polarization. Figure 11.9 shows the measured gain as a function of input wavelength for TE-polarized incident light. The modulation in the gain (gain ripple) with a periodicity of 0.7 nm is due to residual facet

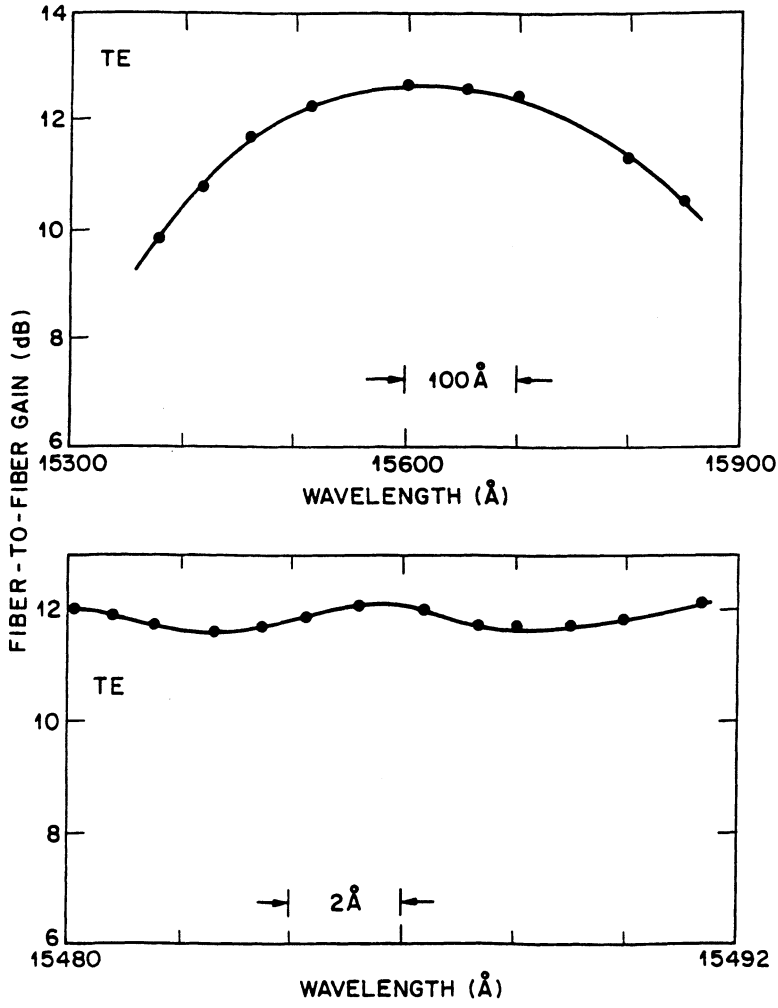


Fig. 11.9 Measured amplifier gain as a function of wavelength for the TE mode. Bottom trace shows variation near the gain peak on a magnified scale.

reflectivities. The measured gain ripple for this device is less than 1 dB. The estimated facet reflectivity from Eq. (11.3.3) by using the measured gain ripple of 0.6 dB at 26 dB internal gain is 9×10^{-5}. The 3-dB bandwidth of the optical gain spectrum is 45 nm for this device. It has been shown that the ripple and polarization dependence of the gain correlate well with the ripple and polarization dependence of the amplified spontaneous-emission spectrum.[11] Measurements of amplified spontaneous emission are much simpler to make than gain measurements, and provide a good estimate of the amplifier performance.

Tilted-Facet Amplifiers. Another way to suppress the resonant modes of the FP cavity is to slant the waveguide (gain region) from the cleaved facet so that the light incident on it internally does not couple back into the waveguide.[12] The process essentially decreases the effective reflectivity of the tilted facet relative to a normally cleaved facet. The reduction in reflectivity as a function of the tilt angle is shown in Fig. 11.10 for the fundamental mode of the waveguide.

A schematic of a tilted-facet optical amplifier is shown in Fig. 11.11. Waveguiding along the junction plane is weaker in this device than that for the

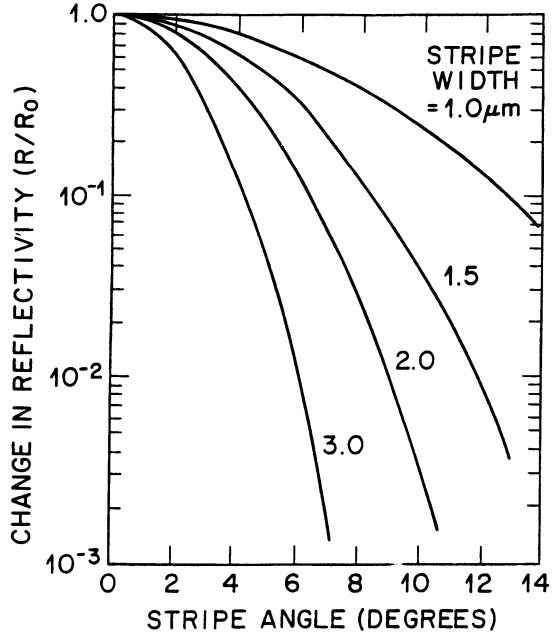


Fig. 11.10 Calculated change in reflectivity as a function of the tilt angle of the facet.

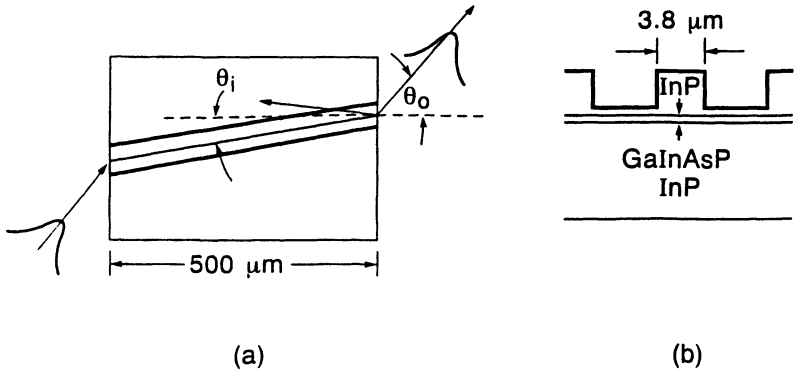


Fig. 11.11 (a) Schematic of a tilted-facet amplifier. (b) Ridge-waveguide structure used for index guiding. (After Ref. 12)

strongly index guided buried-heterostructure device. Weak index guiding for the structure of Fig. 11.11 is provided by a dielectric defined ridge. The fabrication of the device follows a procedure similar to that described previously.

The measured gain as a function of injection current for TM- and TE-polarized light for a tilted facet amplifier is shown in Fig. 11.12. Optical gains as high as 30 dB have been obtained using titled-faced amplifiers. The effective reflectivity of the fundamental mode decreases with increasing tilt of the waveguide. But the effective reflectivity of the higher order modes increases, and this may cause higher-order modes to appear at the output, especially when the ridge widths is large. (Higher-order modes at the output may reduce fiber-coupled power significantly.)

11.3.3 Amplifier Characteristics

The amplification characteristics of SLAs have been discussed extensively.[4–24] As seen in Fig. 11.3, the measured amplifier gain exhibits ripples reflecting the effects of residual facet reflectivities. Gain ripples are negligibly small when the SLA operates in a nearly TW mode. The 3-dB amplifier bandwidth depends on the injected current and is typically 50–60 nm. As discussed in Sec. 11.2.1, this bandwidth reflects the relatively broad gain spectrum, $g(\omega)$, of SLAs. The gain spectrum of semiconductor lasers has been considered in Sec. 3.2.3; the same discussion applies to SLAs. The gain spectrum becomes broader with an increase in the injection current. This feature is clearly seen in Fig. 11.3. Figures 11.8 and 11.12 show how the measured peak gain increases with the injected current.

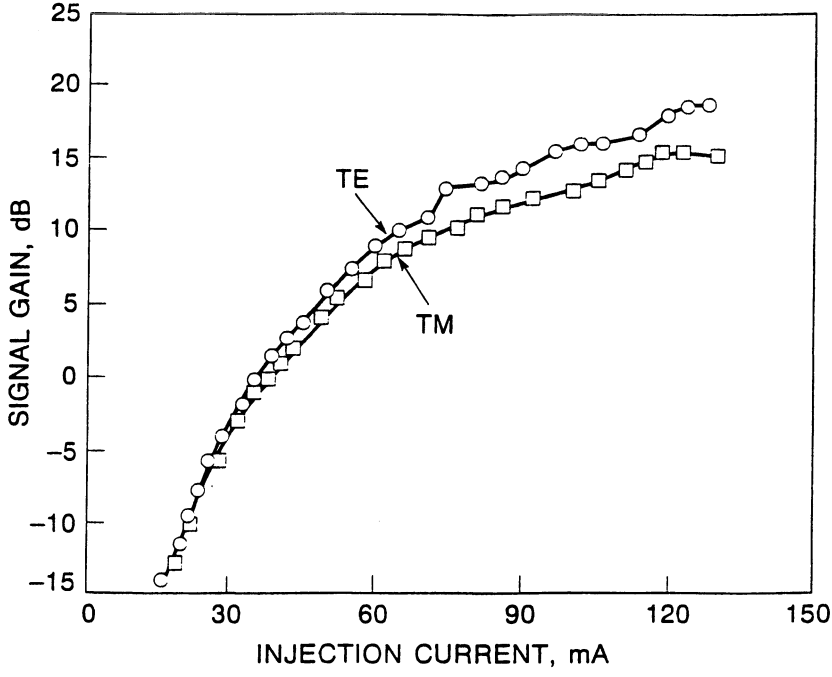


Fig. 11.12 Measured amplifier gain as a function of injection current for a tilted-facet amplifier. (After Ref. 12)

To discuss gain saturation, we consider the peak gain and assume that it increases linearly with the carrier population N (see Sec. 6.3.1). The gain is then given by Eq. (6.3.3) or, equivalently, by

$$g = (\Gamma a/V)(N - N_0) \tag{11.3.6}$$

where Γ is the confinement factor, a is the differential gain coefficient, V is the active volume, and N_0 is the value of N required at transparency. The carrier population N changes with the injection current I and the signal power P according to the rate equation (6.2.20). By expressing the photon number in terms of the optical power, this equation can be written as

$$\frac{dN}{dt} = \frac{I}{q} - \frac{N}{\tau_e} - \frac{a(N - N_0)}{\sigma_m h\nu} P \tag{11.3.7}$$

where τ_e is the carrier lifetime and σ_m is the cross-sectional area of the

waveguide mode. In the case of a CW beam, or pulses much longer than τ_e, the steady-state value of N can be obtained by setting $dN/dt = 0$ in Eq. (11.3.7). When the solution is substituted in Eq. (11.3.6), the optical gain is found to saturate as

$$g = \frac{g_0}{1 + P/P_s} \tag{11.3.8}$$

where the small-signal gain g_0 is given by

$$g_0 = (\Gamma a/V)(I\tau_e/q - N_0) \tag{11.3.9}$$

and the saturation power P_s is defined as

$$P_s = h\nu\sigma_m/(a\tau_e). \tag{11.3.10}$$

A comparison of Eqs. (11.2.1) and (11.3.8) shows that the SLA gain saturates in the same way as that of a two-level system. We can therefore use the analysis of Sec. 11.2.2. In particular, the output saturation power P_{out}^s is given by Eq. (11.2.11) and P_s is given by Eq. (11.3.10). Typical values of P_{out}^s are in the range 5–10 mW.[4]

The noise figure F_n of SLAs is larger than the minimum value of 3 dB for several reasons. The dominant contribution comes from the population-inversion factor n_{sp}, which represents the ratio of spontaneous to net stimulated emission rate.[17] For SLAs, n_{sp} is obtained from Eq. (11.2.16) by replacing N_2 and N_1 by N and N_0, respectively. An additional contribution results from nonresonant internal losses α_{int} (such as free-carrier absorption or scattering loss) which reduce the available gain from g to $g - \alpha_{int}$. By using Eq. (11.2.19) and including this additional contribution, the noise figure can be written as[4]

$$F_n = 2\left(\frac{N}{N - N_0}\right)\left(\frac{g}{g - \alpha_{int}}\right). \tag{11.3.11}$$

Residual facet reflectivities increase F_n by an additional factor that can be approximated by $1 + R_1 G$, where R_1 is the reflectivity of the input facet.[17] In most TW amplifiers $R_1 G \ll 1$, and this contribution can be neglected. Typical values of F_n for SLAs are in the 5–7-dB range.[21]

An undesirable characteristic of SLAs is their polarization sensitivity. The difference in amplifier gain G between TE and TM polarizations is as much as 5–8 dB simply because both Γ and a are different for the two polarizations in the gain expression (8.2.5). This feature makes the amplifier gain dependent

on the polarization state of the input beam, a property undesirable for lightwave system applications where the polarization state changes with propagation along the fiber (unless polarization-preserving fibers are used). Several schemes have been devised to reduce the polarization sensitivity.[25–30] In one scheme, the amplifier is designed such that the width and the thickness of the active region are comparable. A gain difference of less than 1.3 dB between TE and TM polarizations has been realized by making the active layer 0.26-μm thick and 0.4-μm wide.[25] Another scheme makes use of a large-optical-cavity structure; a gain difference of less than 1 dB has been obtained with such a structure.[26]

Several other schemes reduce the polarization sensitivity by using two amplifiers[27] or two passes through the same amplifier.[28] Figure 11.13 shows three such configurations. In Fig. 11.13a the TE-polarized signal in one amplifier becomes TM polarized in the second amplifier, and vice versa. If both amplifiers have identical gain characteristics, the twin-amplifier configuration provides signal gain that is independent of the signal polarization. A drawback of the series configuration is that residual facet reflectivities lead to mutual coupling between the two amplifiers. In the parallel configuration

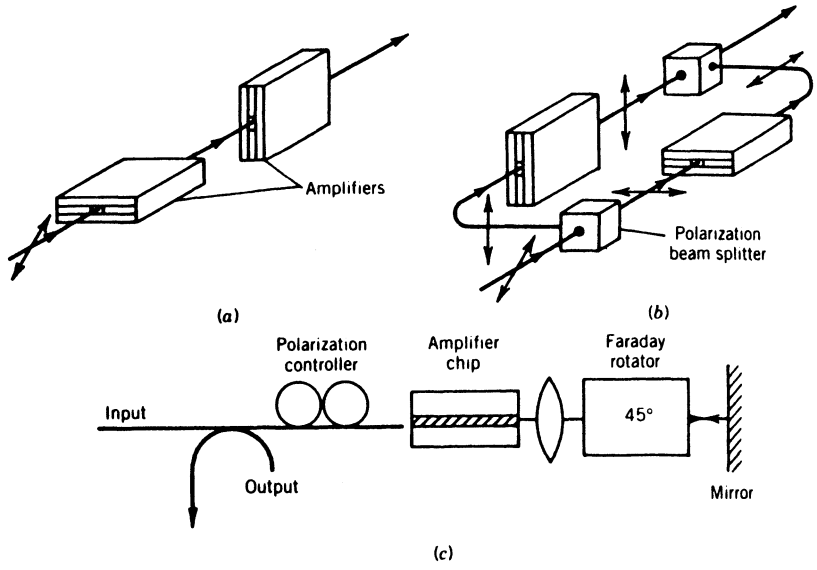


Fig. 11.13 Three configurations used to reduce the polarization sensitivity of semiconductor laser amplifiers: (a) twin amplifiers in series, (b) twin amplifiers in parallel, and (c) double pass through a single amplifier.

shown in Fig. 11.13b the incident signal is split into a TE- and a TM-polarized signal, each of which is amplified by separate amplifiers. The amplifier TE and TM signals are then combined to produce the amplified signal with the same polarization as that of the input beam.[27] The double-pass configuration of Fig. 11.13c passes the signal through the same amplifier twice, but the polarization is rotated by 90° between the two passes.[28] Since the amplified signal propagates in the backward direction, a 3-dB fiber coupler is needed to separate it from the incident signal. In spite of a 6-dB loss occurring at the fiber coupler (3 dB for the input signal and 3 dB for the amplified signal) this configuration provides high gain from a single amplifier, as the same amplifier supplies gain on the two passes.

11.3.4 Multichannel Amplification

One of the advantages of using optical amplifiers is that they can be used to amplify several channels simultaneously as long as the carrier frequencies of multiple channels lie within the amplifier bandwidth. Ideally the signal in each channel should be amplified by the same amount. In practice, several nonlinear phenomena in SLAs induce interchannel cross talk, an undesirable feature that should be minimized for practical lightwave systems. Two such nonlinear phenomena are cross-saturation and four-wave mixing. They both originate from the stimulated recombination term in the carrier rate equation. In the case of multichannel amplification, the power P in Eq. (11.3.7) corresponds to

$$P = \frac{1}{2}\left|\sum_{j=1}^{M} A_j \exp(-i\omega_j t) + \text{c.c.}\right|^2 \tag{11.3.12}$$

where c.c. stands for complex conjugate, M is the number of channels, A_j is the (complex) amplitude, and ω_j is the carrier frequency for the jth channel. Because of the coherent addition of individual channel fields, Eq. (11.3.12) contains a time-dependent term resulting from beating of the signal in different channels, that is,

$$P = \sum_{j=1}^{M} P_j + \sum_{j\neq k}^{M}\sum^{M} 2\sqrt{P_j P_k}\cos(\Omega_{jk} t + \phi_j - \phi_k) \tag{11.3.13}$$

where $A_j = \sqrt{P_j}\exp(i\phi_j)$ is assumed together with $\Omega_{jk} = \omega_j - \omega_k$. When Eq. (11.3.13) is substituted in Eq. (11.3.7), the carrier population is also found to oscillate at the beat frequency Ω_{jk}. Since the gain and the refractive index both depend on N, they are also modulated at the frequency Ω_{jk}; such a

modulation is referred to as the creation of gain and index gratings by the multichannel signal. These gratings induce interchannel cross talk by scattering a part of the signal from one channel to another. This phenomenon can also be viewed as four-wave mixing.[31]

The origin of cross-saturation is also evident from Eq. (11.3.13). The first term on the right side shows that the power P in Eq. (11.3.8) should be replaced by the total power in all channels. Thus, the gain of a specific channel is saturated not only by its own power but also by the power of neighboring channels, a phenomenon known as cross-saturation. Such a cross-saturation has been observed in several experiments.[32–34] It is undesirable for direct-detection or amplitude-shift keying (ASK) coherent systems in which the channel power changes with time depending on the bit pattern. The signal gain of one channel then changes from bit to bit, and the change depends on the bit pattern of neighboring channels. The amplified signal appears to fluctuate more or less randomly. Such fluctuations degrade the effective SNR at the receiver. The cross-saturation-induced interchannel cross talk occurs regardless of the extent of channel spacing. It can only be avoided by operating SLAs in the unsaturated regime. It is also absent for phase-shift keying (PSK) and frequency-shift keying (FSK) coherent systems, since the power in each channel, and hence the total power, remains constant with time.

Interchannel cross talk induced by four-wave mixing, on the other hand, can occur for all multichannel coherent communication systems irrespective of the modulation format used.[35–40] However, in contrast with cross-saturation, it occurs only when the channel spacing is not too large and becomes negligible when channel spacing exceeds 10 GHz simply because carrier population is not able to respond when $\Omega\tau_e \gg 1$, where Ω is the channel spacing. Figure 11.14 shows the gains of two channels for an SLA with 25-dB small-signal gain when the input channel powers are -30 and -25 dBm (1 and 3.2 μW respectively). Dashed lines shows the equal channel gain expected in the absence of four-wave mixing. The gain of the low-frequency channel is enhanced and the gain of the high-frequency channel is reduced as a result of energy transfer initiated by four-wave mixing. The channel gains can differ by more than 5 dB for large input powers and small channel spacing. Such gain variations result in cross talk penalty and degrade the system performance considerably. The cross talk penalty was measured in a system experiment[39] where three channels spaced about 2 GHz apart were transmitted through a 50-km-long fiber by using an FSK coherent scheme. The bit rate was 560 Mb/s. The bit-error rate (BER) degraded severely as a result of four-wave mixing with a large penalty (>2 dB) at 10^{-9}. The performance could be improved by reducing the channel powers or by increasing the channel spacing. A channel spacing of 10 GHz or more is

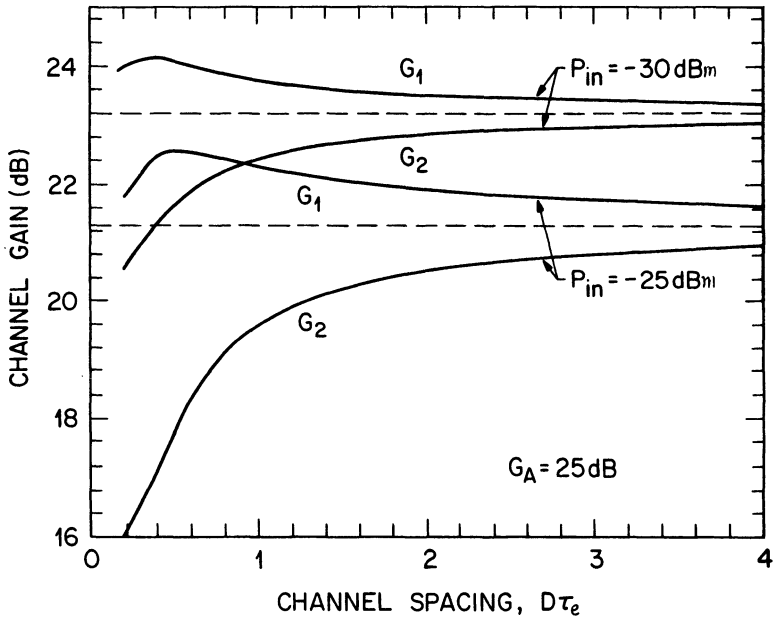


Fig. 11.14 Channel gains G_1 and G_2 as a function of the channel spacing D for two values of the input signal power. Both channels are amplified simultaneously by a semiconductor laser amplifier with 25-dB small-signal gain. Dashed lines show the saturated gain of each channel in the absence of four-wave mixing. (After Ref. 40 © 1990 IEEE)

necessary to reduce the effect of carrier-density modulation and the associated four-wave mixing in SLAs.

11.3.5 Pulse Amplification

The large bandwidth of TW-type SLAs suggest that they are capable of amplifying ultrashort optical pulses (as short as a few picoseconds) without significant pulse distortion. However, when the pulse width τ_p becomes shorter than the carrier lifetime τ_c, gain dynamics play an important role, since both N and g in Eq. (11.3.6) become time dependent. Ultrashort pulse amplification in SLAs has been extensively studied[41–48] both theoretically and experimentally.

The study of pulse propagation in SLAs requires that both spatial and temporal evolution is included by considering a partial differential equation for the amplitude $A(z, t)$ of the pulse envelope. In its simplest form this

equation is given by[44]

$$\frac{\partial A}{\partial z} + \frac{1}{v_g}\frac{\partial A}{\partial t} = \tfrac{1}{2}(1 + i\beta_c)gA \tag{11.3.14}$$

where carrier-induced index changes are included through the line-width enhancement factor β_c. The time dependence of g is governed by Eqs. (11.3.6) and (11.3.7). The two equations can be combined to yield

$$\frac{\partial g}{\partial t} = \frac{g_0 - g}{\tau_e} - \frac{gP}{E_s} \tag{11.3.15}$$

where the saturation energy E_s is defined as

$$E_s = h\nu(\sigma_m/a) \tag{11.3.16}$$

and g_0 is given by Eq. (11.3.9).

Equations (11.3.14) and (11.3.15) govern amplification of optical pulses in SLAs. They can be solved analytically for pulses whose duration is short compared with the carrier lifetime ($\tau_p \ll \tau_e$). The first term on the right-hand side of Eq. (11.3.15) can then be neglected during pulse amplification. By introducing the reduced time $\tau = t - z/v_g$ together with $A = \sqrt{P}\exp(i\phi)$, Eqs. (11.3.14) and (11.3.15) can be written as[44]

$$\frac{\partial P}{\partial z} = g(z, \tau)P(z, \tau) \tag{11.3.17}$$

$$\frac{\partial \phi}{\partial z} = -\tfrac{1}{2}\beta_c g(z, \tau) \tag{11.3.18}$$

$$\frac{\partial g}{\partial \tau} = -\frac{1}{E_s}g(z, \tau)P(z, \tau). \tag{11.3.19}$$

Equation (11.3.17) is identical with Eq. (11.2.5) except that the gain g depends on both z and τ. It can easily be integrated over the amplifier length L to provide the output power as

$$P_{\text{out}}(\tau) = P_{\text{in}}(\tau)\exp[h(\tau)] \tag{11.3.20}$$

where $P_{\text{in}}(\tau)$ is the input power and $h(\tau)$ is the total integrated gain defined as

$$h(\tau) = \int_0^L g(z, \tau)\, dz. \tag{11.3.21}$$

If we integrate Eq. (11.3.19) over the amplifier length after replacing gP by $\partial P/\partial z$, we obtain[42–44]

$$\frac{dh}{d\tau} = -\frac{1}{E_s}[P_{out}(\tau) - P_{in}(\tau)] = -\frac{P_{in}(\tau)}{E_s}(e^h - 1) \tag{11.3.22}$$

where Eq. (11.3.20) was used to relate P_{out} to P_{in}. Equation (11.3.22) can be easily integrated to obtain $h(\tau)$. The quantity of practical interest is the amplification factor $G(\tau)$ related to $h(\tau)$ by $G = \exp(h)$. It is given by

$$G(\tau) = \frac{G_0}{G_0 - (G_0 - 1)\exp[-E_0(\tau)/E_s]} \tag{11.3.23}$$

where G_0 is the unsaturated amplifier gain and

$$E_0(\tau) = \int_{-\infty}^{\tau} P_{in}(\tau)\, d\tau$$

is the partial energy of the input pulse defined such that $E_0(\infty)$ equals the input pulse energy E_{in}.

The solution (11.3.23) shows that the amplifier gain is different for different parts of the pulse. The leading edge experiences the full gain G_0, as the amplifier is not yet saturated. The trailing edge experiences the least gain, since the whole pulse has saturated the amplifier gain. The final value G_f of $G(\tau)$ after passage of the pulse is obtained from Eq. (11.3.23) by replacing $E_0(\tau)$ by E_{in}. The intermediate values of the gain depend on the pulse shape. Figure 11.15 shows the shape dependence of $G(\tau)$ for super-Gaussian input pulses by taking

$$P_{in}(\tau) = P_0 \exp[-(\tau/\tau_p)^{2m}] \tag{11.3.24}$$

where m is the shape parameter. The input pulse is Gaussian for $m = 1$ but becomes nearly rectangular as m increases. For comparison purposes the input energy is held constant for different pulse shapes by choosing $E_{in}/E_s = 0.1$. The shape dependence of the amplification factor $G(\tau)$ implies that the output pulse is distorted when compared with the input pulse, and distortion is itself shape dependent. Figure 11.16 shows the pulse distortion occurring when a super-Gaussian input pulse of energy such that $E_{in}/E_s = 0.1$ is amplified by an SLA with 30-dB small-signal gain. Different curves correspond to different pulse widths τ_0. Pulses shorter than the carrier lifetime τ_e are distorted the most since $G(\tau)$ is the most nonuniform for such

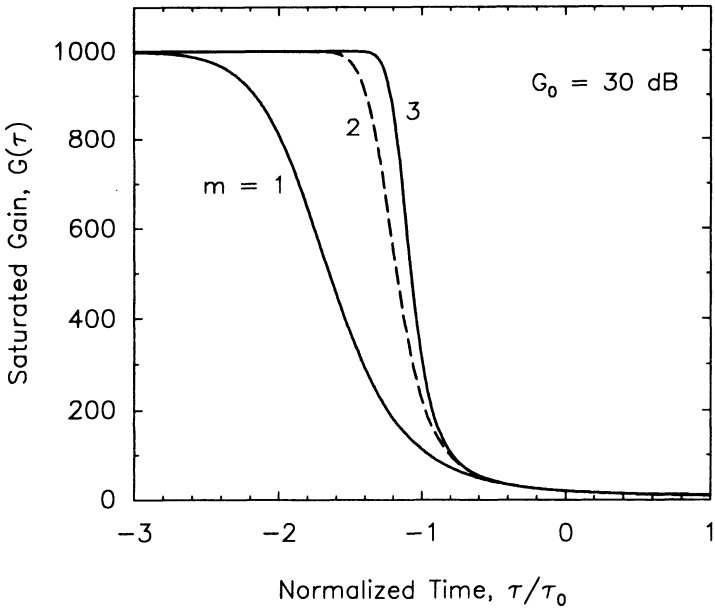


Fig. 11.15 Time-dependent amplification factor for super-Gaussian input pulses of input energy such that $E_{in}/E_s = 0.1$. The unsaturated value G_0 is 30 dB in all cases. The input pulse is Gaussian for $m = 1$ but becomes nearly rectangular as m increases.

pulses. Interestingly enough, the energy gain G_E, defined as $G_E = E_{out}/E_{in}$, does not depend on the pulse shape. It depends only on the initial gain G_0 and the final gain G_f through the relation[44]

$$G_E = \frac{\ln[(G_0 - 1)/(G_f - 1)]}{\ln[(G_0 - 1)/(G_f - 1)] - \ln(G_0/G_f)}. \tag{11.3.25}$$

The above discussion applies to any kind of amplifier. However, SLAs differ in one important aspect from other amplifiers. The difference comes from Eq. (11.3.18), which shows that gain sautration leads to a time-dependent phase shift across the pulse. The total phase shift is found by integrating Eq. (11.3.18) over the amplifier length and is given by

$$\phi(\tau) = -\tfrac{1}{2}\beta_c \int_0^L g(z, \tau)\, dz = -\tfrac{1}{2}\beta_c h(\tau) = -\tfrac{1}{2}\beta_c \ln[G(\tau)]. \tag{11.3.26}$$

Since the pulse modulates its own phase through gain saturation, this

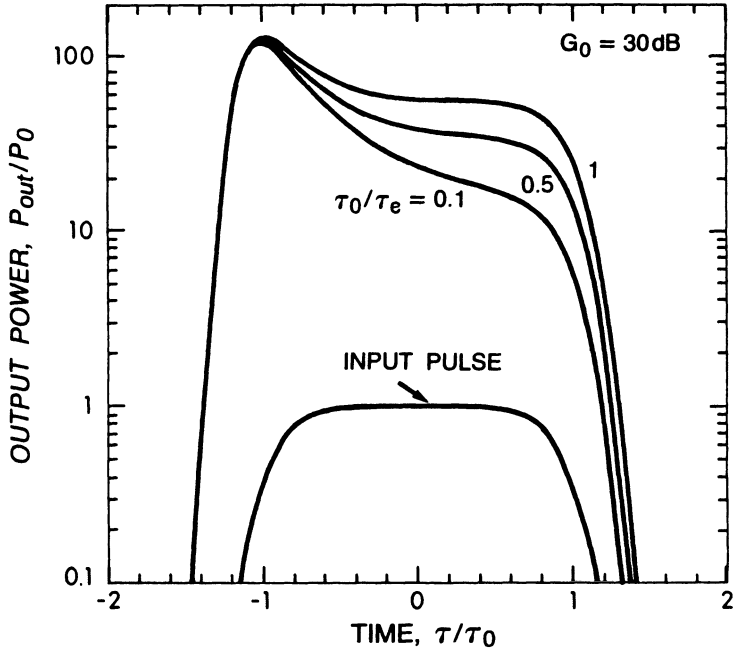


Fig. 11.16 Amplified pulses for three values of τ_0/τ_e when a super-Gaussian input pulse of width τ_0 is amplified by a semiconductor laser amplifier with 30-dB small-signal gain.

phenomenon is referred to as saturation-induced self-phase modulation.[44] Physically, gain saturation leads to temporal variations in the carrier population, which in turn change the refractive index through the line-width enhancement factor β_c. Changes in the refractive index modify the optical phase. The frequency chirp is related to the phase derivative as

$$\Delta\nu_c = -\frac{1}{2\pi}\frac{d\phi}{d\tau} = \frac{\beta_c}{4\pi}\frac{dh}{d\tau} = -\frac{\beta_c P_{in}(\tau)}{4\pi E_s}[G(\tau) - 1] \qquad (11.3.27)$$

where Eq. (11.3.22) was used. Figure 11.17 shows the chirp profiles for several input pulse energies when a Gaussian pulse is amplified in an SLA with 30-dB unsaturated gain. The frequency chirp is larger for more energetic pulses simply because gain saturation sets in earlier for such pulses.

Self-phase modulation and the associated frequency chirp can affect lightwave systems considerably. The spectrum of the amplified pulse becomes considerably broad and contains several peaks of different amplitudes.[42] The dominant peak is shifted toward the red side and is broader than the input

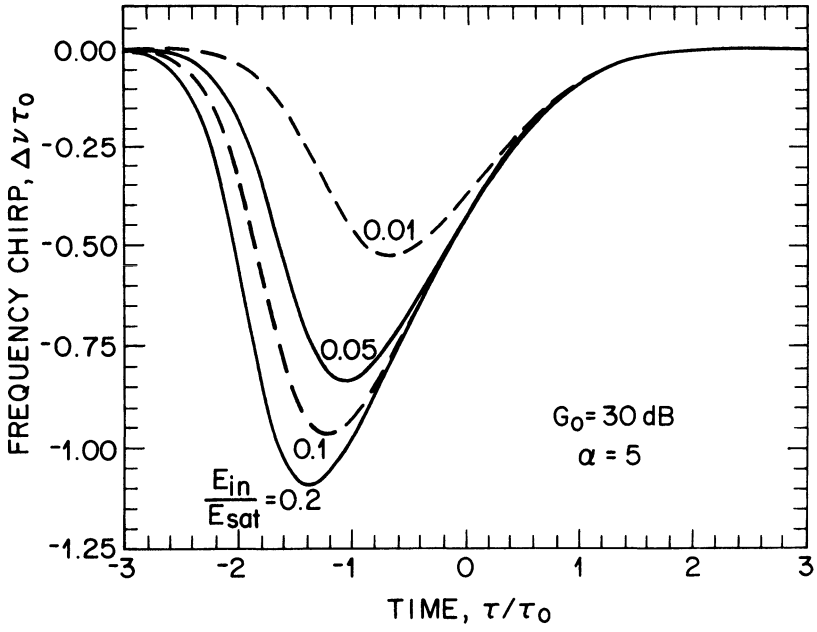


Fig. 11.17 Frequency chirp imposed across the amplified pulse for several values of E_{in}/E_s. A Gaussian input pulse is assumed together with $G_0 = 30$ dB and $\beta_c = 5$. (After Ref. 41)

spectrum. It is also accompanied by one or more satellite peaks. Figure 11.18 shows the expected shape and spectrum of amplified pulses when a Gaussian pulse of energy such that $E_{in}/E_s = 0.1$ is amplified by SLAs. The temporal and spectral changes depend on amplifier gain and are quite significant for $G_0 = 30$ dB. The experiments performed by using picosecond pulses from mode-locked semiconductor lasers[41–44] confirm the expected qualitative behavior. In particular, the spectrum of amplified pulses is found to be shifted toward the red side by 50–100 GHz depending on the amplifier gain.[42] Spectral distortion in combination with the frequency chirp would affect the transmission characteristics when amplified pulses are propagated through optical amplifiers.

It turns out that the frequency chirp imposed by the SLA is opposite in nature compared with that imposed by directly modulated semiconductor lasers. If we also note that the chirp is nearly linear over a considerable portion of the amplified pulse (see Fig. 11.17), it is easy to understand that the amplified pulse would pass through an initial compression stage when it propagates in the anomalous-dispersion region of optical fibers.[24] Figure 11.19 shows the input pulse, the amplified pulse, and the compressed pulse

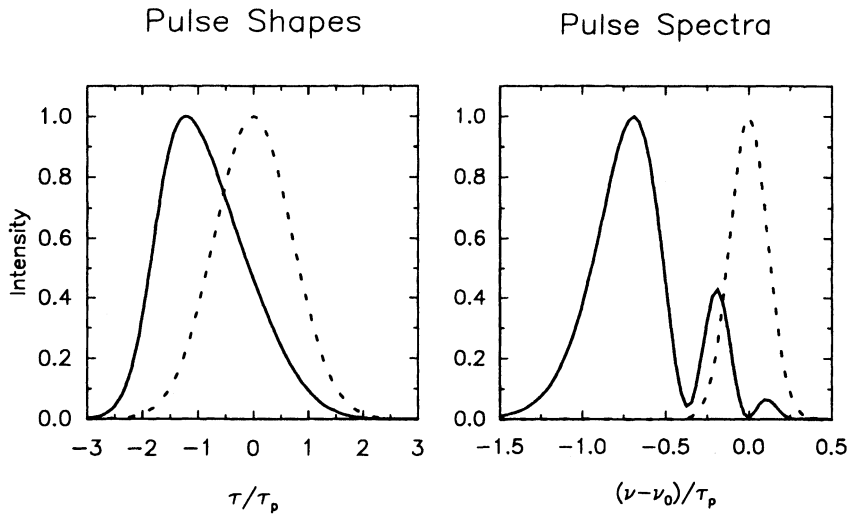


Fig. 11.18 Shape and spectrum of the amplified pulse when a Gaussian input pulse of energy $E_{in}/E_s = 0.1$ is amplified by a semiconductor laser amplifier with $G_0 = 30$ dB and $\beta_c = 5$. The dashed curves show for comparison input shape and spectrum.

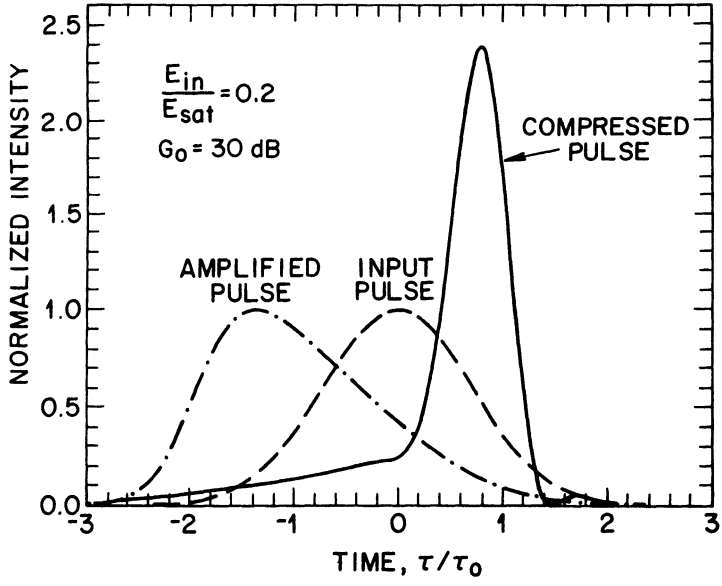


Fig. 11.19 Input, amplified, and compressed pulses when a Gaussian input pulse of width τ_0 is amplified by a semiconductor laser amplifier with 30-dB small-signal gain and then propagated through an optical fiber of optimum length with anomalous dispersion. (After Ref. 41)

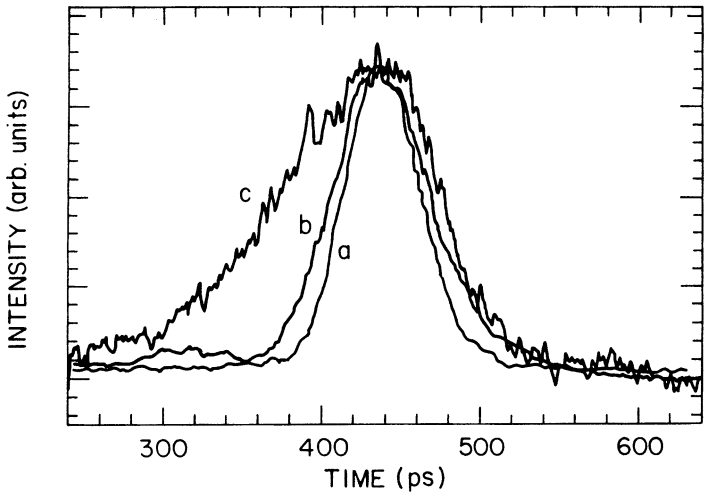


Fig. 11.20 Streak-camera traces of the input pulse (trace a) and the amplified pulse after propagating in a 70-km-long fiber (trace b) for the case $G_0 = 30$ dB. Trace c shows the pulse shape when G_0 is reduced below 10 dB. (After Ref. 44 © 1989 IEEE)

when a Gaussian pulse is first amplified in an SLA with 30-dB small-signal gain and then propagated through an optical fiber of length $0.3L_D$, where L_D is the dispersion length.[41] The pulse is compressed by about a factor of 3 because of the frequency chirp imposed on the amplified pulse by the SLA. Such a compression was observed in an experiment[41] in which 40-ps optical pulses were first amplified in a 1.52-μm SLA and then propagated through 18 km of single-mode fiber with a dispersion of 15 ps/(km·nm). This compression mechanism can be used to design fiber-optic communication systems in which in-line SLAs are used to compensate simultaneously for fiber loss and dispersion by operating SLAs in the saturation region so that they impose frequency chirp on the amplified pulse. The basic concept was demonstrated in 1989 in an experiment[43] in which the signal was transmitted over 70 km at 16 Gb/s by using an SLA. Figure 11.20 shows the pulse shapes measured by a streak camera at the fiber output with and without SLA. In the absence of an SLA or when the SLA was operated in the unsaturated regime, the system was dispersion-limited to the extent that signal could not be transmitted over more than 20 km.

11.3.6 System Applications

Several experiments have demonstrated the potential of SLAs as optical preamplifiers.[49–51] In this application, the signal is optically amplified

before it falls on the receiver. The preamplifier boosts the signal to such a high level that the receiver performance is limited by shot noise rather than by thermal noise. The basic idea is similar to the case of avalanche photodiodes (APDs), which amplify the signal in the electrical domain. However, just as APDs add additional noise, preamplifiers also degrade the SNR through spontaneous-emission noise. The relatively large noise figure of SLAs ($F_n = 5$–6 dB) makes them less attractive as preamplifiers. Nonetheless, they can improve the receiver sensitivity considerably. In an 8-Gb/s transmission experiment, the use of an SLA preamplifier resulted in a 3.7 dB improvement over the best sensitivity achieved by using APD receivers.[51] SLAs can also be used as power amplifiers to boost the transmitter power. It is, however, difficult to achieve powers in excess of 10 mW because of a relatively small value ($\sim$5–7 mW) of the output saturation power.

SLAs have been used as in-line amplifiers in several system experiments.[52–61] In one experiment,[54] signal at 1 Gb/s was transmitted over 313 km by using four cascaded SLAs. This experiment used the conventional IM-DD (intensity modulation with direct detection) technique for signal transmission. In a coherent transmission experiment[55] four cascaded SLAs were used to transmit 420-Mb/s FSK signal over 370 km. In a 140-Mb/s coherent experiment,[58] the transmission distance was extended to 546 km by using 10 cascaded SLAs. Recently, SLAs have been employed to overcome distribution losses in local-area networks.[60,61] In one experiment[61] an SLA was used as a dual-function device. It amplified five FSK channels multiplexed by using the subcarrier multiplexing technique (see Sec. 7.5), improving the power budget by 11 dB. At the same time, the SLA was used to monitor the network performance through a baseband control channel. The 100-Mb/s baseband control signal modulated the carrier density of the amplifier to produce a corresponding electric signal that was used for monitoring.

Although the potential of SLAs for lightwave system applications has been demonstrated, SLAs need to overcome several drawbacks before their use becomes practical. Among the drawbacks are polarization sensitivity, interchannel cross talk, and a large coupling loss. Even though SLAs can have a chip gain as high as 30–35 dB, the usable gain is reduced by 8–10 dB because of a large coupling loss occurring at the input and output ends. Fiber amplifiers are preferable from this standpoint, since their coupling loss (due to fusion splice) is negligible. They are also nearly polarization insensitive and have negligible interchannel cross talk. SLAs would nonetheless find some applications, as they can be monolithically integrated within the transmitter or receiver. Of particular interest are multiquantum-well (MQW) amplifiers discussed in the next section.

11.3.7 Multiquantum-Well Amplifiers

For some applications the polarization independence of gain is not important. An example is a power amplifier following a laser. Since light from the laser is TE polarized, the parameter of interest in this application is high TE gain and high saturation power. Optical amplifiers with a multiquantum-well (MQW) active region (which have strongly polarization dependent gain) satisfy both of these requirements and are therefore ideally suited for this application. The anisotropic nature of the optical transition in a quantum well makes the TE-mode gain of an MQW amplifier much higher (by more than 10 dB) than the TM-mode gain. MQW amplifiers are capable of much higher output power than regular double-heterostructure (DH) amplifiers.

The output power of an amplifier is limited by gain saturation. If P_s is the saturation power in the gain medium, the output saturation power can be approximated by $P_0 = P_s/\Gamma$ where Γ is the confinement factor of the optical mode. For MQW amplifiers whose active region consists of a few (generally 3–4) quantum wells 5–10-nm thick, the confinement factor is considerably smaller than that for a regular DH amplifier. This effect results in high saturation power. However, since the signal gain is given by $G = \exp[(\Gamma g - \alpha)L]$ where g is the material gain, the MQW amplifiers have lower gain than DH amplifiers for the same cavity length.

Gain-saturation characteristics of an optical amplifier are obtained by plotting the measured gain as a function of output power. This is shown in Fig. 11.21 for both a DH and an MQW amplifier. Both devices amplify signals near 1.55 μm. The DH amplifier had a 0.4-μm-thick active region, was 500-μm long and exhibited < 1 dB gain difference between TE and TM polarizations. The MQW amplifier result is shown for the TE mode. It had 4 active-layer wells, each 7-nm thick, and barrier layers also 7-nm thick. Saturation powers as high as 100 mW have been reported for MQW amplifiers. The density of states function for electrons and holes in a quantum well is independent of energy. This feature results in a broad spontaneous-emission spectrum and hence a broad gain spectrum of an MQW amplifier. The measured gain spectrum of a device at two different currents is shown in Fig. 11.22a while Fig. 11.22b shows the corresponding output saturation power.

Since the MQW amplifier is ideally suited for amplifying the output power of a semiconductor laser, it is useful to combine amplifier and laser on a single chip and thereby eliminate coupling losses. MQW amplifiers have been integrated with both distributed-feedback[62] and distributed-Bragg-reflector (DBR)[63] lasers (see Fig. 11.23). A grating provides the frequency-selective feedback resulting in single-mode operation of the laser. MQW layers grown over the InP substrate serve as the active region for both the laser and the

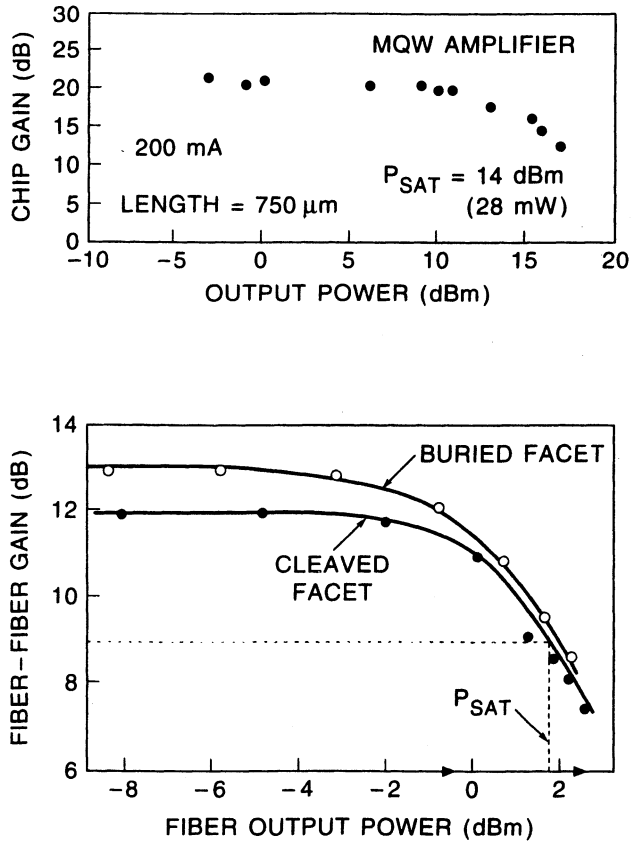


Fig. 11.21 Measured amplifier gain as a function of output power for DH and MQW amplifiers.

amplifiers. The effectiveness of the amplifier can be seen from Fig. 11.24 which shows the L-I characteristics of the laser with the amplifier biased at 170 mA. The slope of the L-I curve (2 mW/mA) is about a factor of 10 higher than that for a typical DBR laser without an amplifier.

11.4 FIBER AMPLIFIERS

Amplification of light in a fiber by the interaction of a pump with a signal can be accomplished in a number of ways. These include nonlinear optical phenomena such as Raman amplification and Brillouin amplification[64–66] or by the stimulated emission from an excited state of a rare-earth ion within

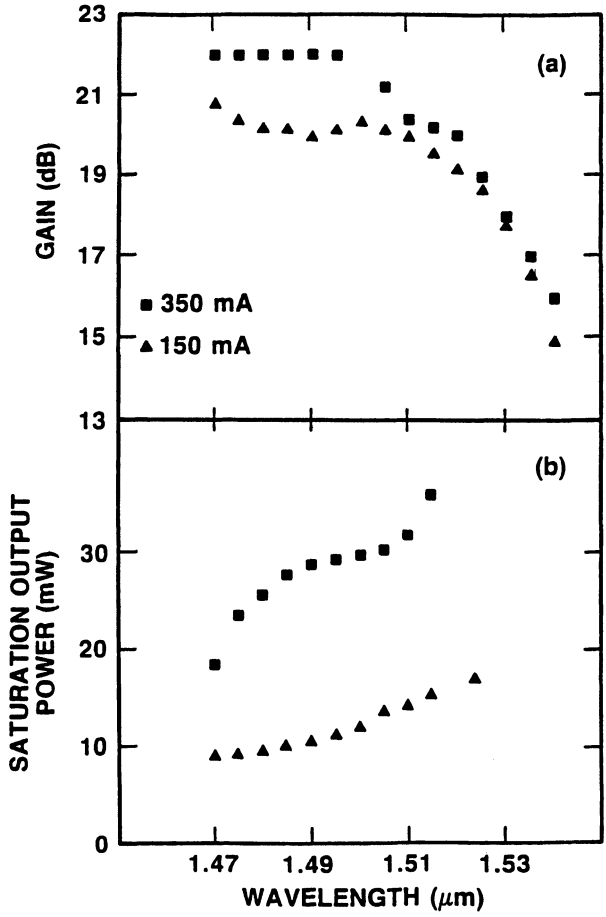


Fig. 11.22 (a) Measured gain spectrum of an MQW amplifier at two injection currents; and (b) the corresponding variation of the output saturation power with wavelength. (After Ref. 62)

the fiber.[24] Nonlinear methods have been shown to be useful, but they are generally less efficient in transferring energy from pump to signal. The method of placing rare-earth ions in the core of an optical fiber as an amplifying medium was first demonstrated in 1964 by Koester and Snitzer.[67] They observed 40 dB of gain at 1.06 μm in a flashlamp-pumped neodymium-doped fiber 1 m in length. Amplification of light in the wavelength region of minimum loss for a silica-based optical fiber (1.5 μm) using transitions of the erbium ion was demonstrated more than twenty years later.[68,69]

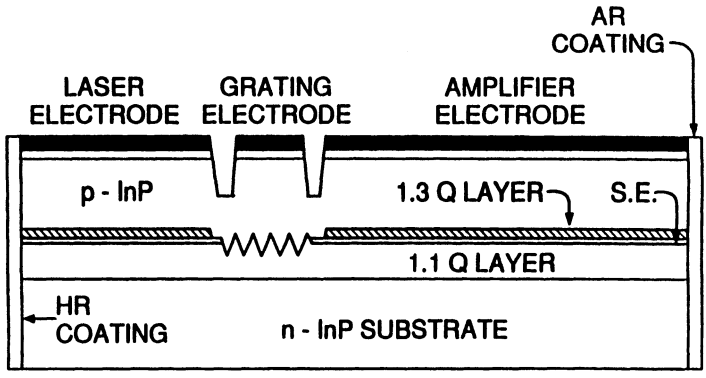


Fig. 11.23 Schematic of a DBR laser integrated with an MQW amplifier. (After Ref. 63)

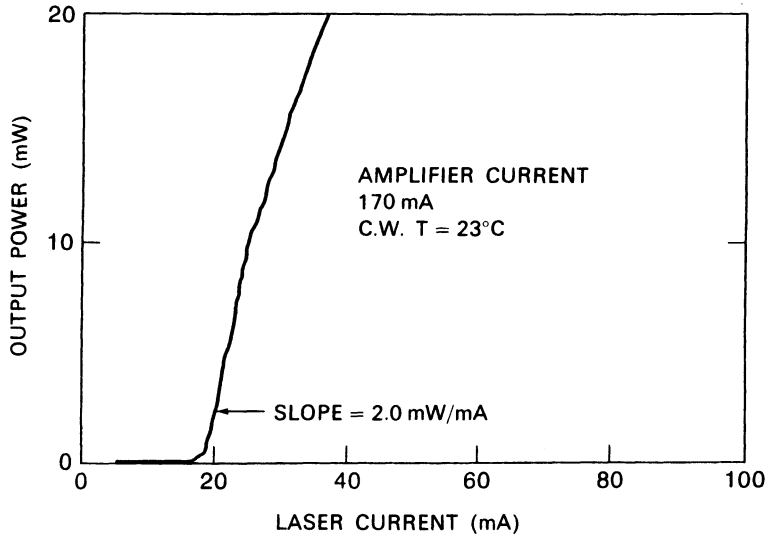


Fig. 11.24 L-I curve of a DBR laser integrated with an MQW amplifier. (After Ref. 63)

11.4.1 Energy Levels

Fundamental to all rare-earth-doped amplifier systems is the ability to invert the population of ions from the ground state to an excited state. The excited state acts as a storage of pump power from which incoming signals may stimulate emission. The pumping schemes are broadly classified as either 3- or 4-level systems as shown in Figure 11.25. Erbium is considered a

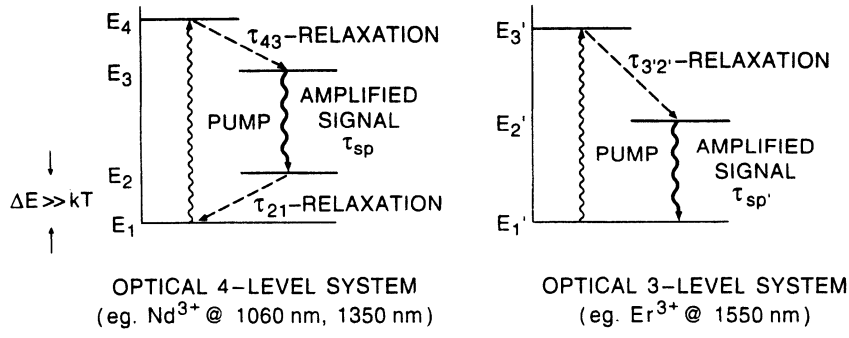


Fig. 11.25 Pumping schemes for amplifiers based on 3- and 4-level atomic systems. Dashed lines indicate that the corresponding transition occurs on a relatively fast time scale.

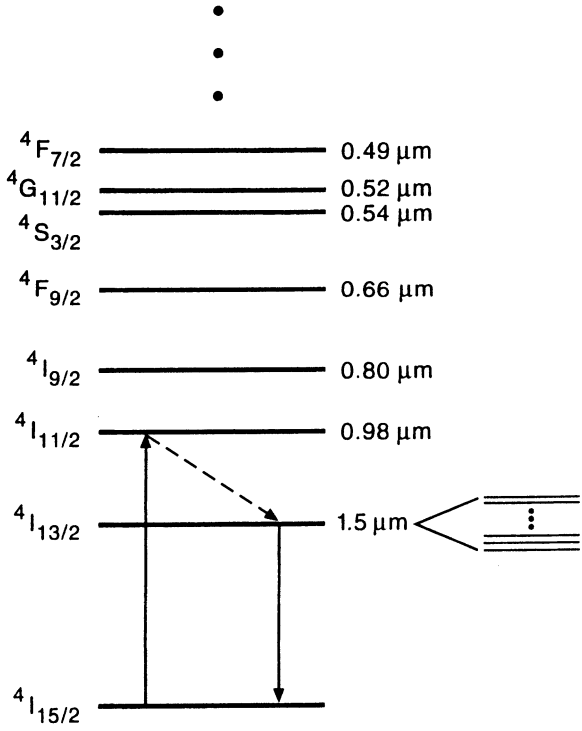


Fig. 11.26 Energy levels of erbium in a silica fiber. Each level consists of many closely spaced sublevels. Wavelengths on the right indicate the spectral region associated with each transition.

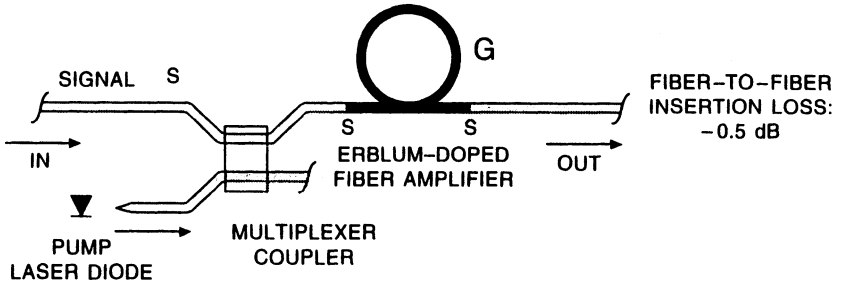


Fig. 11.27 Schematic drawing of an erbium-doped fiber amplifier.

3-level system. Notable in the 3-level system is an absorption at the signal wavelength when the system is not inverted or underpumped.

The energy levels of the Er^{3+} ion in the glass fiber are shown in Fig. 11.26. The local electric field acts as a small perturbation on the erbium ion which results in a splitting of each level of the ion into a number of closely spaced levels (shown for $^4I_{13/2}$). The splitting between these levels is usually much smaller than the energy separation between the discrete levels of the ion. Each of the closely spaced levels are further broadened by their characteristic lifetimes and inhomogeneities in the glass host. This results in the observation of broad absorption and fluorescence spectra. The amplification process involves generation of a sufficiently large number of Er^+ ions in the excited state using an external laser source. For efficient transfer of energy the wavelength of the external source (pump laser) should match one of the energy levels of the Er^+ ion.

A laser emitting near 1,480 nm will populate directly the first excited state. The most commonly used pump lasers at present emit near 1,480 nm and 980 nm. Semiconductors lasers emitting near 1,480 nm and 980 nm are most desirable pump sources. Pumping at 980 nm is believed to be very efficient and also results in less noise. The main elements of a fiber amplifier are schematically shown in Fig. 11.27. The pump power and signal are combined using a wavelength-division multiplexer. High-power semiconductor lasers emitting at 0.98 μm and at 1.48 μm are used as pump sources for fiber amplifiers. Lasers at these wavelengths have been fabricated using InGaAs-GaAs and InGaAsP-InP material systems.[70,71]

11.4.2 Fiber Amplifier Performance

A principal characteristic for a device such as an optical fiber amplifier is optical gain. The measured small-signal gain at two different signal wavelengths as a function of pump power is shown in Fig. 11.28. Note that the

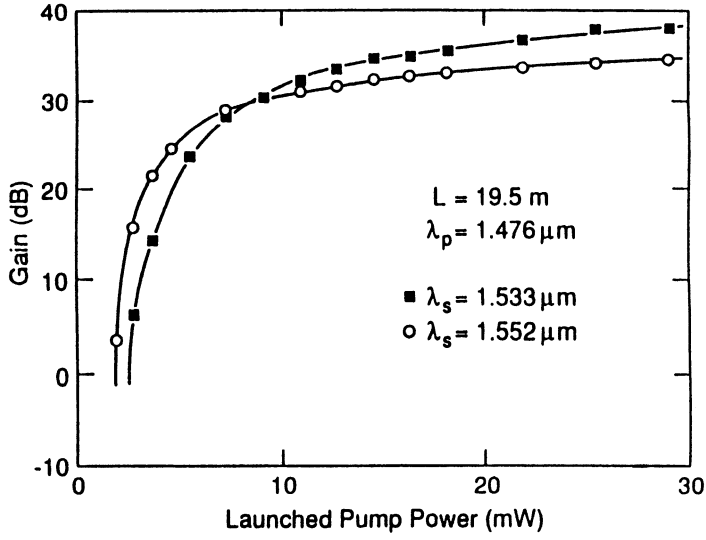


Fig. 11.28 Small-signal amplifier gain measured as a function of pump power for an erbium-doped fiber amplifier. (Courtesy J. Zyskind)

saturation power is quite high relative to that for a semiconductor amplifier. The pump-laser wavelength is 1.476 μm and the fiber length is 19.5 m. The gain increases rapidly at pump powers near threshold and increases slowly at high pump powers, where almost all the erbium ions along the length of the fiber are inverted.

A figure of merit commonly used to describe the amplifier is the slope of the tangent to the gain-versus-pump-power curve (in dB/mW). Large values indicate a low threshold and a steep rise in gain with pumping. Pump efficiencies vary with pumping wavelength and can be as high as 5.9 dB/mW for 1.48-μm pumping, 11.0 dB/mW for 0.98-μm pumping and 1.3 dB/mW for 0.82-μm pumping.

An important characteristic of an amplifier is the noise figure defined by Eq. (11.2.12). The noise figure for a fiber amplifier is given by

$$F_n = \frac{1}{\eta_{in}} \frac{2n_{sp}(G-1) + \dfrac{1}{\eta_{out}} \dfrac{1}{\eta_{det}}}{G} \tag{11.4.1}$$

where η_{in} and η_{out} are the input and output coupling efficiencies, η_{det} is the detector quantum efficiency, G is the gain of the fiber amplifier, and n_{sp} is

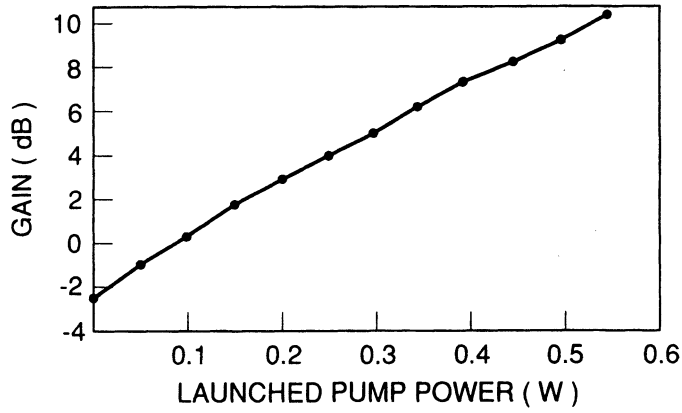


Fig. 11.29 Measured gain of a 1.3-μm fiber amplifier as a function of launched pump power at a wavelength of 1.007 μm. (After Ref. 75)

the spontaneous-emission factor. Under ideal conditions, $\eta_{in} = 1$, $G \gg 1$ and $n_{sp} = 1$ which results in $F \approx 2n_{sp} \approx 3$ dB. The measured values of F_n for 1.48-μm, 0.98-μm, and 0.82-μm pumping are 4.1 dB,[72] 3.2 dB,[73] and 4.0 dB,[74] respectively. The noise figure is lowest for 0.98-μm pumping.

Erbium-doped fiber amplifiers are relatively well developed. Amplifiers operating near 1.3 μm are also being investigated.[75,76] Optical gain has been reported in fluorozirconate fibers doped with 560 ppm of Pr^{3+} and pumped at 1.007 μm. The measured gain at 1.3 μm as a function of the launched pump power is shown in Fig. 11.29. These initial results are promising, even though the pump power is considerably higher than that for an erbium-doped fiber amplifier. Availability of commercial-grade fiber amplifiers at 1.3 μm will have a strong impact on the design of lightwave systems.

PROBLEMS

11.1 The Lorentzian gain profile of a semiconductor laser amplifier has a full width at half maximum of 5 THz. Calculate the amplifier bandwidths when the amplifier is operated to provide 20- and 30-dB gain. Neglect gain saturation.

11.2 A semiconductor laser amplifier can amplify a 1-μW signal to the 1-mW level. What is the output power when a 1-mW signal is incident on the same amplifier? Assume that the saturation power of the optical gain is 10 mW.

11.3 Explain the concept of noise figure for an optical amplifier. Why does the SNR of the amplified signal degrade by 3 dB even for an ideal amplifier.

11.4 A 250 μm-long semiconductor laser is used as a Fabry-Perot amplifier by biasing it below threshold. Calculate the amplifier bandwidth by assuming 32% reflectivity for both facets and 30-dB amplifier gain. The group index $n_g = 4$. How much does the bandwidth change when both facets are coated to reduce the facet reflectivities to 1%?

11.5 Complete the derivation of Eq. (11.3.3) by using Eq. (11.3.1). Use it to calculate the facet reflectivities that would ensure gain ripples of < 1 dB for a semiconductor laser amplifier designed to provide 30-dB small-signal gain. Assume $R_1 = 2R_2$.

11.6 Use Eq. (11.3.7) to derive expressions for the amplifier gain G and the output saturation power for a semiconductor laser amplifier of length L.

11.7 Use Eqs. (11.3.14) and (11.3.15) to derive an expression for $G(\tau)$ when an optical pulse of width $\tau_0 \ll \tau_e$ is amplified by a semiconductor laser amplifier. Explain all assumptions and approximations made during the derivation.

11.8 Complete the derivation of Eq. (11.3.25) for the energy gain.

REFERENCES

1. Coupland, M. J., K. G. Mambleton, and C. Hilsum. *Phys. Lett.* **7,** 231 (1963).
2. Crowe, J. W., and R. M. Craig, Jr. *Appl. Phys. Lett.* **4,** 57 (1964).
3. Kosnocky, W. F., and R. H. Cornely. *IEEE J. Quantum Electron.* **QE-4,** 225 (1968).
4. Saitoh, T., and T. Mukai. Chap. 7 in *Coherence, Amplification and Quantum Effects in Semiconductor Lasers*, ed. by Y. Yamamoto. New York: Wiley, 1991.
5. Nakamura, M., and S. Tsuji. *IEEE J. Quantum Electron.* **QE-17,** 994 (1981).
6. Saitoh, T., and T. Mukai. *J. Lightwave Technol.* **LT-6,** 1656 (1988).
7. O'Mahony, M. *J. Lightwave Technol.* **LT-6,** 531 (1988).
8. Joyce, W. B., and B. C. DeLoach, Jr. *Appl. Opt.* **23,** 4187 (1984).
9. H. Kogelnik, and T. Li. *Proc. IEEE* **54,** 1312 (1966).
10. Lin, M. S., A. B. Piccirilli, Y. Twu, and N. K. Dutta. *Electron. Lett.* **25,** 1378 (1989).
11. Dutta, N. K., M. S. Lin, A. B. Piccirilli, R. L., Brown, and U. K. Chakrabarti. *J. Appl. Phys.* **67,** 3943 (1990).
12. Zah, C. E., J. S. Osinski, C. Caneau, S. G. Menocal, L. A. Reith, J. Salzman, F. K. Shokoohi, and T. P. Lee. *Electron. Lett.* **23,** 990 (1987).
13. Yamamoto, Y. *IEEE J. Quantum Electron.* **QE-16,** 1047 (1980).
14. Luk'yanov, V. N., A. T. Semenov, and S. D. Yakubovich. *Sov. J. Quantum Electron.* **10,** 1432 (1980).
15. Favre, F., L. Jeunhomme, I. Joindot, M. Monerie, and J. C. Simon. *IEEE J. Quantum Electron.* **QE-17,** 897 (1981).
16. Simon, J. C. *Electron. Lett.* **18,** 438 (1982).
17. Mukai, T., Y. Yamamoto, and T. Kimura. *IEEE J. Quantum Electron.* **QE-18,** 1560 (1982).
18. Saitoh, T., and T. Mukai. *IEEE J. Quantum Electron.* **QE-23,** 1010 (1987).
19. Simon, J. C. *J. Lightwave Technol.* **LT-5,** 1286 (1987).
20. Eisenstein, G., B. C. Johnson, and G. Raybon. *Electron. Lett.* **23,** 1020 (1987).
21. Öberg, M. G., and N. A. Olsson. *Electron. Lett.* **24,** 99 (1988).
22. Olsson, N. A. *J. Lightwave Technol.* **LT-7,** 1071 (1989).

23. Fujiwara, M., H. Nishimoto, T. Kajitami, M. Itoh, and S. Suzuki. *J. Lightwave Technol.* **LT-9,** 155 (1991).
24. Agrawal, G. P. Chap. 8 in *Fiber-Optic Communication Systems.* New York: Wiley, 1992.
25. Cha, I., M. Kitamura, H. Honmou, and I. Mito. *Electron. Lett.* **25,** 1241 (1989).
26. Cole, S., D. M. Cooper, W. J. Devlin, A. D. Ellis, D. J. Elton, J. J. Isaak, G. Sherlock, P. C. Spurdens, and W. A. Stallard. *Electron. Lett.* **25,** 314 (1989).
27. Großkopf, G., R. Ludwig, R. G. Waarts, and H. G. Weber. *Electron. Lett.* **23,** 1387 (1987).
28. Olsson, N. A. *Electron. Lett.* **24,** 1075 (1988).
29. Sumida, M. *Electron. Lett.* **26,** 1913 (1989).
30. Koga, M., and T. Mutsumoto. *IEEE Photon. Technol. Lett.* **12,** 431 (1989); *J. Lightwave Technol.* **LT-9,** 284 (1991).
31. Agrawal, G. P. *Opt. Lett.* **12,** 260 (1987); *J. Opt. Soc. Am. B* **5,** 147 (1988).
32. Großkopf, G., R. Ludwig, and H. G. Weber. *Electron. Lett.* **22,** 900 (1986).
33. Mukai, T., K. Inoue, and T. Saitoh. *Electron. Lett.* **23,** 396 (1987).
34. Jopson, R. M., K. L. Hall, G. Eisenstein, G. Raybon, and M. S. Whalen. *Electron. Lett.* **25,** 510 (1987).
35. Agrawal, G. P. *Electron. Lett.* **23,** 1175 (1987).
36. Inoue, K. *Electron. Lett.* **23,** 1293 (1987).
37. Jopson, R. M., T. E. Darcie, K. T. Gayliard, R. T. Ku, R. E. Tench, T. C. Rice, and N. A. Olsson. *Electron. Lett.* **23,** 1394 (1987).
38. Habbab, I. M. I., and G. P. Agrawal. *J. Lightwave Technol.* **LT-7,** 1351 (1989).
39. Ryu, S., K. Mochizuki, and H. Wakabayashi. *J. Lightwave Technol.* **LT-7,** 1525 (1989).
40. Agrawal, G. P., and I. M. I. Habbab. *IEEE J. Quantum Electron.* **QE-26,** 501 (1990).
41. Agrawal, G. P., and N. A. Olsson. *Opt. Lett.* **14,** 500 (1989).
42. Olsson, N. A., and G. P. Agrawal. *Appl. Phys. Lett.* **55,** 13 (1989).
43. Olsson, N. A., G. P. Agrawal, and K. W. Wecht. *Electron. Lett.* **25,** 603 (1989).
44. Agrawal, G. P., and N. A. Olsson. *IEEE J. Quantum Electron.* **QE-25,** 2297 (1989).
45. Marshall, I. W., D. M. Spirit, and M. J. O'Mahony. *Electron. Lett.* **23,** 818 (1987).
46. Wiesenfeld, J. M., G. Eisenstein, R. S. Tucker, G. Raybon, and P. B. Hansen. *Appl. Phys. Lett.* **53,** 1239 (1988).
47. Saitoh, T., H. Itoh, Y. Noguchi, S. Sudo, and T. Mukai. *IEEE Photon. Technol. Lett.* **1,** 297 (1989).
48. Agrawal, G. P. *Electron. Lett.* **27,** 620 (1991); *IEEE J. Quantum Electron.* **QE-27,** 1843 (1991).
49. Olsson, N. A., and P. Garbinski. *Electron. Lett.* **22,** 1114 (1986).
50. Marshall, I. W., and M. J. O'Mahony. *Electron. Lett.* **23,** 1052 (1987).
51. Jopson, R. M., A. H. Gnauck, B. L. Kasper, R. E. Tench, N. A. Olsson, C. A. Burrus, and A. R. Chraplyvy. *Electron. Lett.* **25,** 233 (1989).
52. Marshall, I. W., D. M. Spirit, and M. J. O'Mahony. *Electron. Lett.* **22,** 253 (1986).
53. Westlake, H. J., and M. J. O'Mahony, *Electron. Lett.* **23,** 649 (1987).
54. Olsson, N. A., M. G. Öberg, L. A. Koszi, and G. J. Przybylek. *Electron. Lett.* **24,** 36 (1988).
55. Öberg, M. G., N. A. Olsson, L. A. Koszi, and G. J. Przybylek. *Electron. Lett.* **24,** 38 (1988).
56. Großkopf, G., R. Ludwig, and H. G. Weber. *Electron. Lett.* **24,** 551 (1988).
57. Maylon, D. J., D. J. Elton, J. C. Regnault, S. J. McDonald, W. J. Devlin, K. H. Cameron, D. M. Bird, and W. A. Stallard. *Electron. Lett.* **25,** 235 (1989).
58. Ryu, S., H. Taga, S. Yamamoto, K. Mochizuki, and H. Wakabayashi. *Electron. Lett.* **25,** 1682 (1989).
59. Gabla, P. M., S. Gauchard, and I. Neubauer. *IEEE Photon. Technol. Lett.* **2,** 594 (1990).
60. Way, W. I., C. E. Zah, and T. P. Lee. *IEEE Trans. Microwave Theory Tech.* **38,** 534 (1990).
61. Koai, K. T., R. Olshansky, and P. M. Hill. *IEEE J. Quantum Electron.* **QE-2,** 926 (1990).
62. Dutta, N. K., J. Lopata, R. Logan, and T. Tanbun-Ek. *Appl. Phys. Lett.* **59,** 1676 (1991).

63. Koren, U., B. I. Miller, G. Raybon, M. Oron, M. G. Young, T. L. Koch, J. L. DeMiguel, M. Chien, B. Tell, K. Brown-Goebeler, and C. A. Burrus. *Appl. Phys. Lett.* **57,** 1375 (1990).
64. Aoki, Y. *J. Lightwave Technol.* **LT-6,** 1225 (1988).
65. Olsson, N. A., and J. P. van der Ziel. *Appl. Phys. Lett.* **48,** 1329 (1986).
66. Agrawal, G. P. *Nonlinear Fiber Optics.* Boston: Academic Press, 1989.
67. Koester, C. J., and E. Snitzer. *Appl. Opt.* **3,** 1182 (1964).
68. Mears, R. J., L. Reekie, S. B. Poole, and D. N. Payne. *Electron. Lett.* **22,** 159 (1986).
69. Desurvire, E., J. R. Simpson, and P. C. Becker. *Opt. Lett.* **12,** 11 (1987).
70. Dutta, N. K., J. Lopata, D. L. Sivco, and A. Y. Cho. *Appl. Phys. Lett.* **58,** 1125 (1991).
71. Tanbun-Ek, T., R. A. Logan, N. A. Olsson, H. Temkin, A. M. Sergent, and K. W. Wecht. *Appl. Phys. Lett.* **57,** 224 (1990).
72. Giles, C. R., et al. *IEEE Photon. Technol. Lett.* (1990).
73. Way, W. I., A. C. Von Lehman, M. J. Andrejco, M. A. Saifi, and C. Lin. 1990, Opt. Amp. Mtg., Monteray, CA, Aug., paper TuB3.
74. Kimura, Y., K. Suzuki, and M. Nakagawa. *Electron. Lett.* **27,** 147 (1991).
75. Carter, S. F., D. Szebesta, S. T. Davey, R. Wyatt, M. C. Brierley, and P. W. France. *Electron. Lett.* **27,** 628 (1991).
76. Durteste, Y., M. Monerie, J. Y. Allain, and H. Poignant. *Electron. Lett.* **27,** 626 (1991).
77. Eisenstein, G., U. Koren, G. Raybon, J. M. Wisenfeld, and M. Wegner. *Appl. Phys. Lett.* **57,** 333 (1990).

Chapter 12
PHOTONIC AND OPTOELECTRONIC INTEGRATED CIRCUITS

12.1 INTRODUCTION

During the late 1980s and early 1990s, there has been a significant number of developments in the technology of optical and electronic integration of semiconductor lasers and other related devices on a single chip. These chips allow higher levels of functionality than that achieved with single devices. For example, lasers and electronic drive circuits have been integrated, serving as simple monolithic lightwave transmitters. Similarly, optical detectors have been integrated with amplifier circuits based on field-effect transistors (FETs) or heterojunction bipolar transistors (HBTs). Such integrated devices serve as the front end of a receiver. In addition to being a technological achievement in the area of photonics, optoelectronic integration is expected to bring down the system cost. The name *photonic integrated circuit* (PIC) is generally used when all the integrated components are photonic devices, e.g., lasers, detectors, amplifiers, modulators, and couplers. The name *optoelectronic integrated circuit* (OEIC) is used when the components are a combination of photonic and electronic devices. Several review articles have been published on PICs and OIECs.[1–4] This chapter provides an overview of the technology involved and discusses the properties of some important integrated circuits.

12.2 PHOTONIC INTEGRATED CIRCUITS

12.2.1 Arrays

The simplest of all photonic integrated circuits are one-dimensional arrays of lasers, LEDs, or photodetectors.[5–7] These devices are fabricated exactly the same way as individual devices except the wafers are not scribed to make single-device chips but left in the form of a bar. A schematic of a laser array

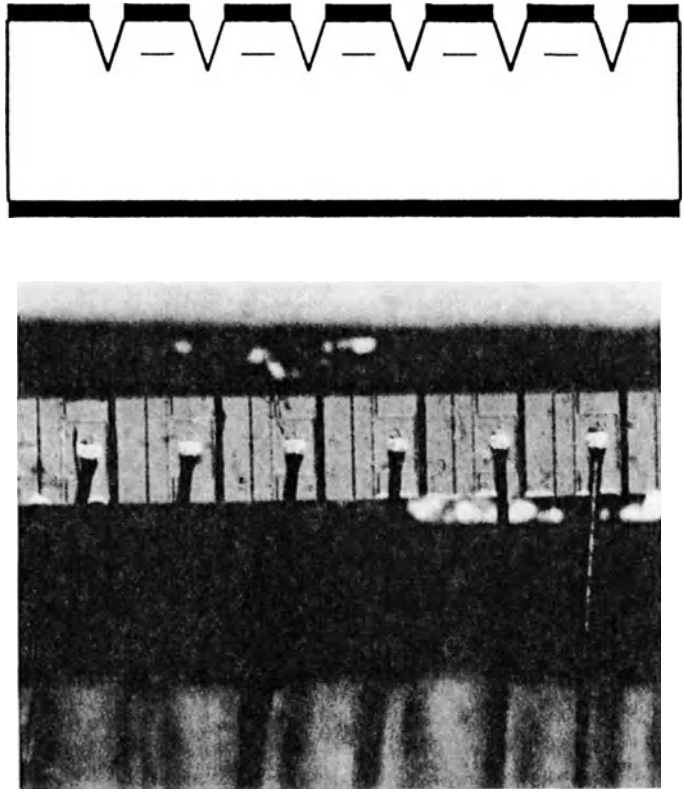


Fig. 12.1 Schematic and photograph of an independently addressable edge-emitting laser array consisting of covered-mesa buried-heterostructure (CMBH) lasers.

is shown in Fig. 12.1. Individual elements are isolated by etching V-shaped grooves between the emitters.

Arrays of lasers or LEDs can be used as sources for dense parallel optical interconnection, a technology finding application in the next generation of computing and switching systems. A necessary requirement for the array technology is low power consumption and the absence of cross talk between the individual elements of the array. The former requirement arises from the limited heat-transfer capacity of the transmitter and receiver module under typical operating conditions and environment. The power consumption problem is considerably reduced if low-threshold lasers are used. Such lasers (threshold current ~ 1 mA) have been fabricated both for the AlGaAs and the InGaAsP material systems. In both cases, the lasers have multiquantum-well (MQW) active regions, short cavity lengths and high-reflectivity facet

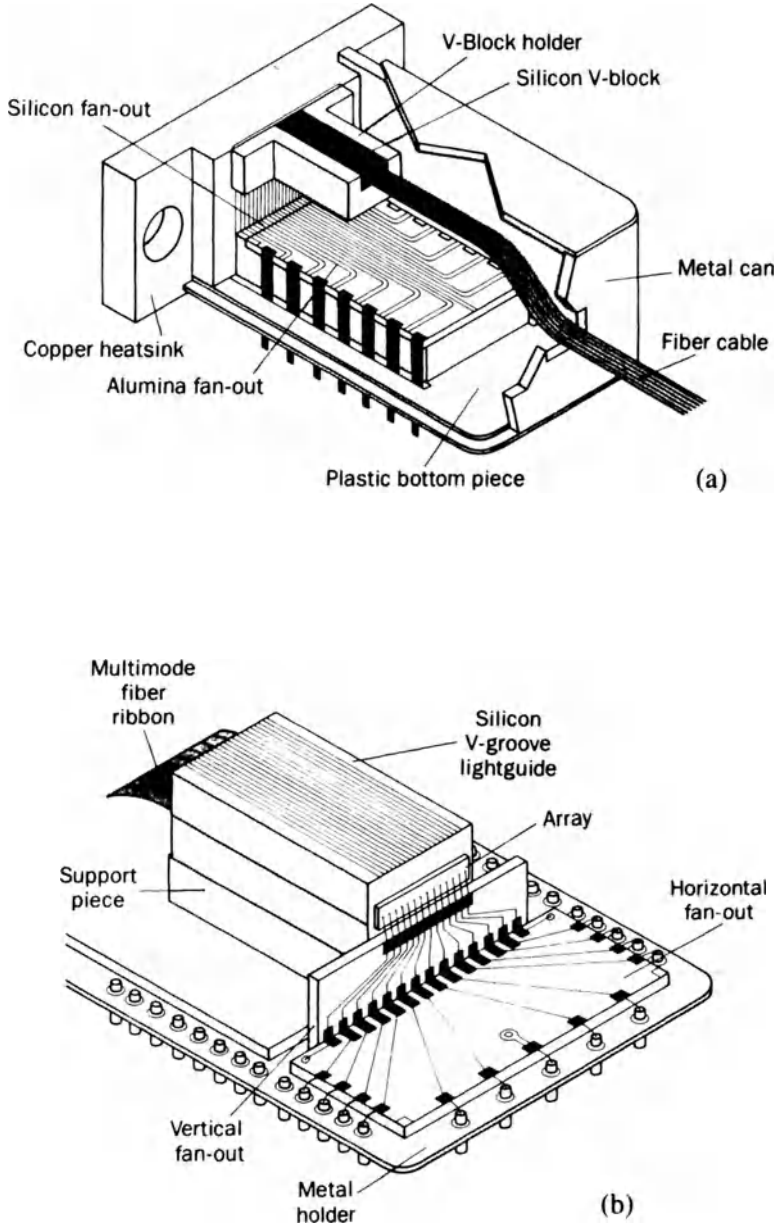


Fig. 12.2 Schematic illustrations of (a) an LED and (b) an edge-emitting laser-array package used in a parallel optical interconnection. The package uses a 60-μm core 1 × 12 multimode fiber bundle which is connectorized to V-grooves in a block of Si. (After Ref. 7)

coatings. Laser and detector arrays have been fabricated such that they exhibit no cross talk up to data rates of 1.4 Gb/s.

LEDs and photodiodes have been packaged into 1 × 12 arrays using a 12-fiber bundle and Si V-grooves for alignment. The fiber bundles terminate in connectors which consist of Si V-grooves with a center-to-center spacing of 250 μm. A schematic of an LED and a laser-array package is shown in Fig. 12.2. For two-dimensional emitter arrays, the surface-emitting laser technology is a natural extension of the surface-emitting LED-array technology and provides higher speed and power levels. Independently addressable two-dimensional arrays of surface-emitting lasers have been reported.

12.2.2 Integrated Laser Detector

Considerable attention has been paid to developing integrated laser-detector structures.[8–13] In many cases, the purpose of the detector is to serve as a monitor for the output power of the laser. The output of the monitor can then be used to stablize the power of the laser, if it drifts due to aging or temperature change, using a feedback circuit. A typical integrated laser-detector structure is shown in Fig. 12.3. The emitting region of the laser and the absorbing region of the photodiode are composed of the same material. The detector monitors the back-facet power of the laser. In this structure the laser has one cleaved and one etched facet.

For DFB lasers, the power outputs of the two facets do not necessarily track. Hence the implicit assumption made in the structure of Fig. 12.3, that the back-facet output is a good measure of the front-facet output, is not necessarily true for a DFB laser. It is possible to fabricate an integrated DFB

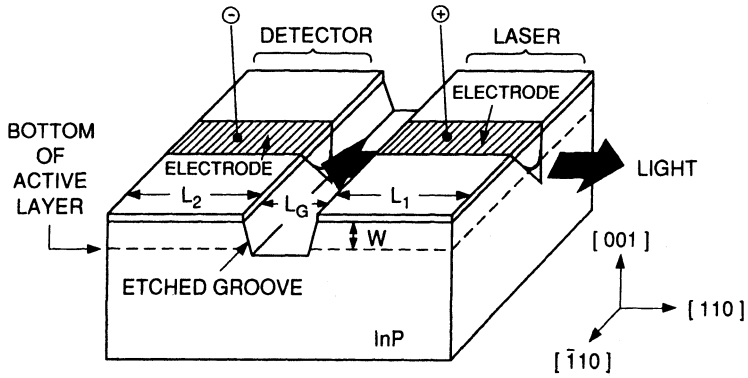


Fig. 12.3 Integrated edge-emitting laser and a back-face monitoring detector. The laser has one etched facet and another cleaved facet. (After Ref. 11)

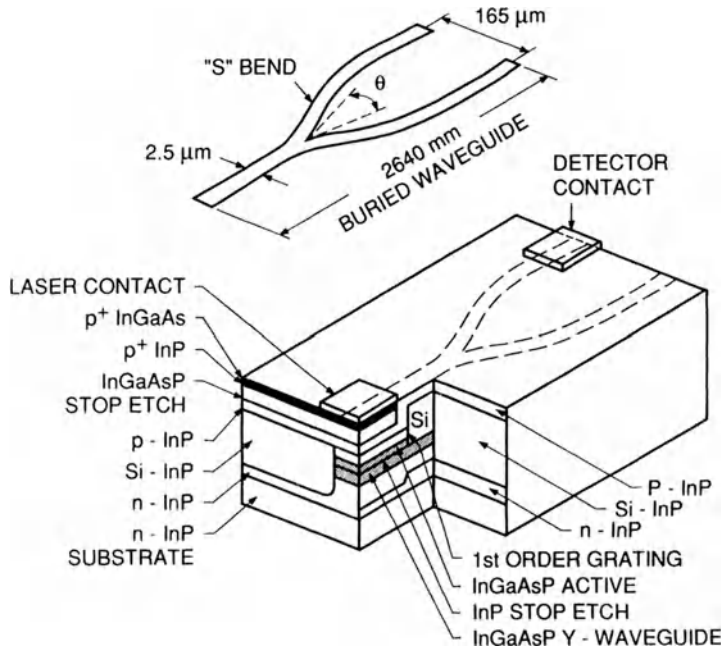


Fig. 12.4 Integrated DFB laser-photodiode structure. The laser and the photodiode are coupled using a Y-branching waveguide. The photodiode monitors the output of the front facet. (After Ref. 12)

laser-detector structure which allows a front-facet monitor.[12] In this structure (see Fig. 12.4), a Y-junction waveguide is fabricated near the end of a DFB laser. One branch of the Y has a monitor photodiode and the second branch serves as the output. The laser output and the monitor photodiode characteristics of this device are shown in Fig. 12.5.

Photodiodes have also been integrated with a surface-emitting laser (SEL). The latter being planar lends itself more easily to integration. The device structure is shown in Fig. 12.6. All layers are grown by the MBE technique. The photodiode mesa is first etched and a self-aligned SEL is then fabricated using ion implantation for current confinement. Since a fraction of the emitted light is absorbed by the photodiode, the photocurrent is a good measure of the light output.

12.2.3 Integrated Laser Modulator

Most commercial lightwave systems use a directly modulated laser diode as the source of information transfer. Under direct modulation, the 3-dB

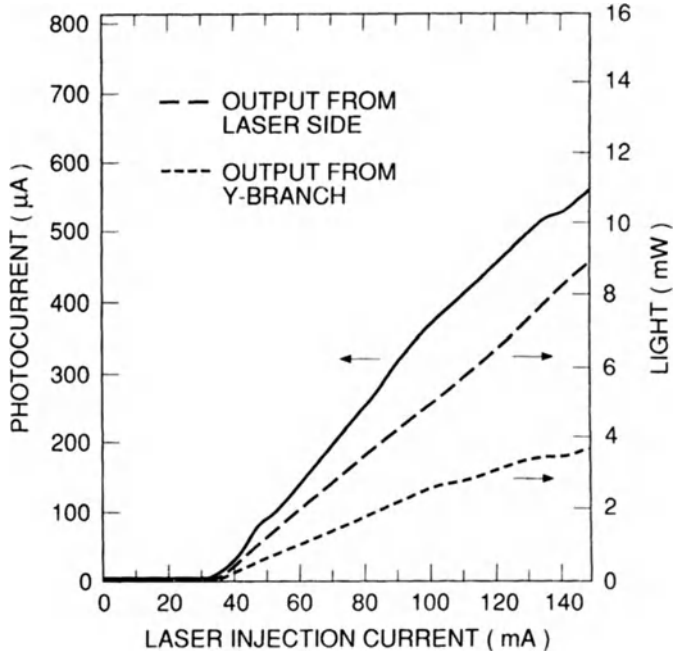


Fig. 12.5 The laser output and the monitor photodiode output plotted as a function of laser current. (After Ref. 12)

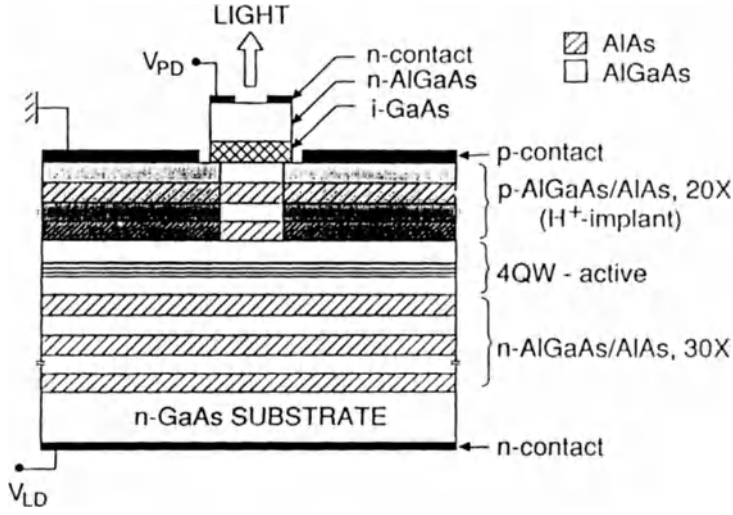


Fig. 12.6 Integrated surface-emitting laser and photodiode. (After Ref. 13)

spectral width of a single-wavelength DFB laser is ~ 0.1 nm. This finite spectral width results in a dispersion penalty due to chromatic dispersion of the fiber and limits the transmission distance. This problem can be partially solved if an external modulator is used to modulate the output of a continuously operating laser. For long-distance transmission, externally modulated lasers are needed. Such externally modulated sources use a CW DFB laser, the output of which is coupled to a $LiNbO_3$ external modulator. It is desirable to have both the laser and the modulator integrated on the same chip.

Although several types of modulation schemes using III–V materials exist, the most common type is the electro-absorption modulator. In this type of modulator, a change in absorption is produced by a change in electric field. The change in absorption at a given wavelength (close to the band gap) occurs due to the Franz-Keldysh effect, according to which the band gap decreases with increasing electric field.

The schematic of a DFB laser integrated with an electro-absorption modulator[14,15] is shown in Fig. 12.7. The fabrication of the device involves first the growth of the layers of the DFB laser on a planar wafer, then etching using a dielectric mask, followed by growth of the layers of the modulator. Note that the band gap of the absorbing layer of the modulator is larger than that of the emitting layer of the laser. This is necessary so that the modulator is nonabsorbing at zero field. With increasing applied field, the band gap of the absorbing layer decreases which decreases the transmission through the modulator.

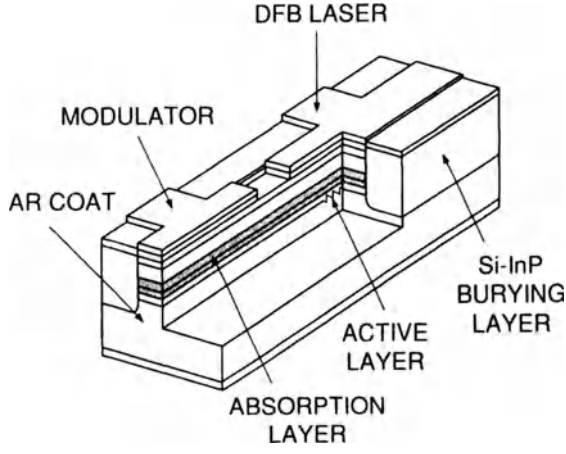


Fig. 12.7 Integrated DFB laser and electro-absorption modulator. (After Ref. 14 © 1987 IEEE)

12.2.4 Integrated Laser Amplifier

Integration schemes in which the modulator is an amplifier as opposed to an absorber have been reported.[16,17] Integrated DFB laser-amplifier structures also have very low spectral width under modulation. The schematic of such a structure is shown in Fig. 12.8. Both the DFB laser and the amplifier have an MQW active region. The laser section has a grating etched on the substrate for frequency-selective feedback. The Fe-doped InP between the laser section and the amplifier section provides good optical coupling and electrical isolation between the sections. For low-chirp operation, the laser is modulated with a small current (~ 5 mA) and the amplifier is CW biased. It is the small modulation current that controls the chirp in this device. The amplifier is used to increase the output power to a level that is obtained with high modulation current.

Another interesting PIC is a series of amplifiers connected in tandem with a laser.[18] A schematic of such a PIC is shown in Fig. 12.9. In this device,

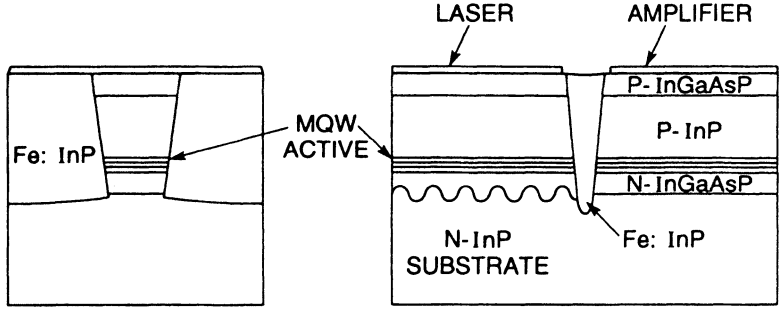


Fig. 12.8 Integrated DFB laser and optical amplifier. (After Ref. 17)

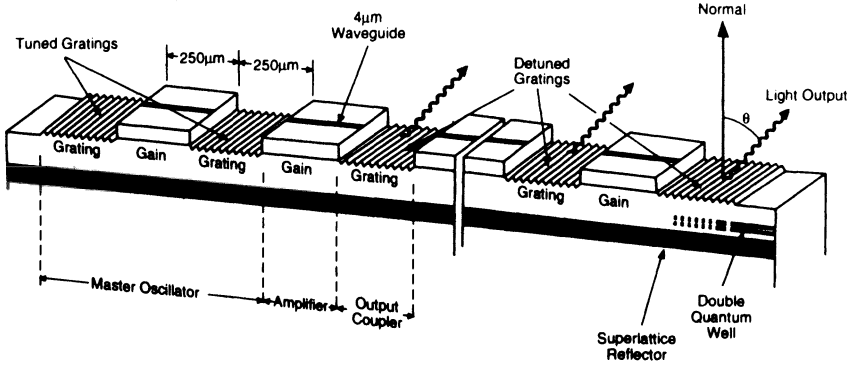


Fig. 12.9 A number of DBR lasers and an amplifier connected in series. (After Ref. 18)

light is emitted normal to the surface of the grating. The first grating-laser combination acts as an oscillator. A fraction of the light propagates along the waveguide and injects into the next amplifier and so on. The device emits in a series of beams normal to the wafer. The emission is in a single wavelength and all the emitters are coherent with respect to each other. These array devices are discussed in Sec. 5.7.

Integration of lasers, amplifiers and DBR gratings of various types has been achieved. Some of these devices emit high power normal to the surface and others emit normal to a cleaved edge.

12.2.5 Heterodyne Receiver

Perhaps the most complicated PIC reported so far is a heterodyne receiver for coherent transmission.[19,20] The essential elements of such a receiver are the local-oscillator laser, the optics for mixing local-oscillator light with incident signal, and the balanced receiver. The receiver PIC fabricated by Koch et al.[20] is shown in Fig. 12.10. The device has a tunable DBR laser, which serves as the local oscillator; several branching waveguides, for mixing and splitting the signal; and two photodiodes, which serve as balanced receivers.

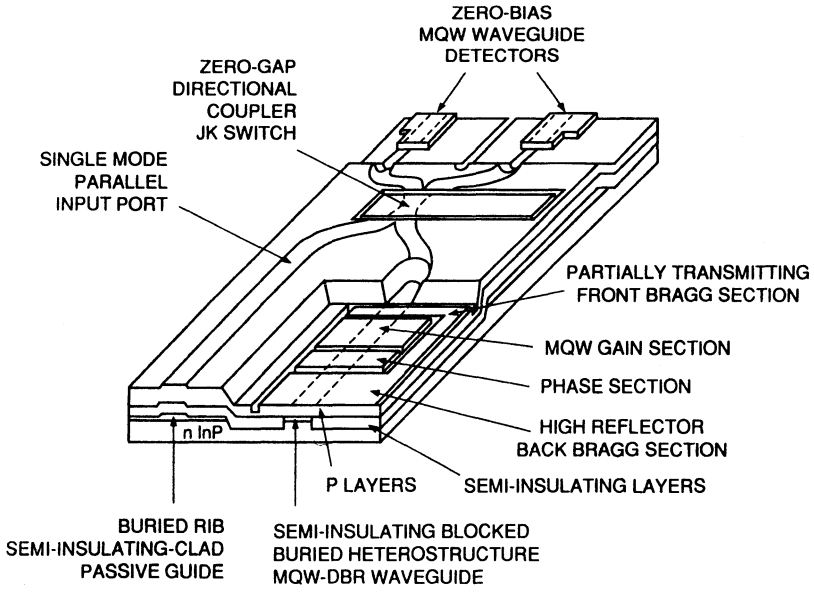


Fig. 12.10 Schematic of a heterodyne receiver photonic integrated circuit. (After Ref. 20)

12.3 OPTOELECTRONIC INTEGRATED CIRCUITS (OEICs)

The first OEIC was reported in the mid 1970s by Yariv and coworkers.[21,22] It was a simple device—a Gunn diode integrated with a laser.[21] Since then, various types of integrated laser-driver and photodiode-amplifier with single and multichannel capability have been reported. So far, most of the OEIC developments have been towards lightwave transmitters and receivers.

12.3.1 Receiver OEICs

The receiver OEIC generally consists of PIN photodiode integrated monolithically with an amplifier.[23–28] The amplifiers using FET- and HBT-based technologies have been fabricated. The first receiver OEIC was an InGaAs PIN photodiode integrated with a single InP-based FET.[23]

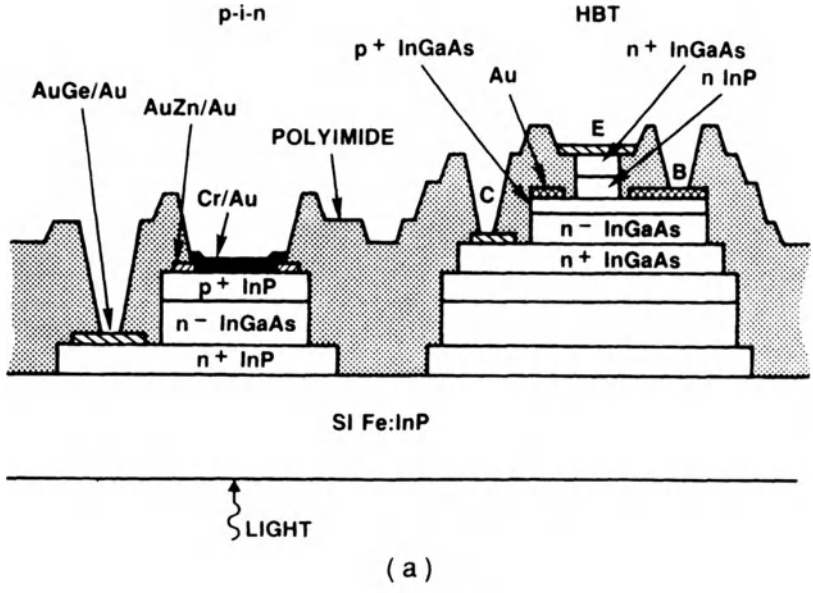
The schematic of a receiver OEIC which consists of an InGaAs PIN photodiode and an HBT-based amplifier is shown in Fig. 12.11. The entire layer structure is grown by the MOVPE growth technique. The detected photocurrent from the PIN photodiode is amplified by transistors T_1 and T_2. Transistor T_3 acts as an output buffer stage.

Arrays of photodiodes and integrated amplifiers for use in optical interconnection have also been reported. The amplifiers used here are FET-based. Figure 12.12 shows the photograph of a PIN-MODFET (modulation-doped FET) array fabricated with a center-to-center spacing of 250 μm. The fabrication involves two epitaxial growths by MBE. First the layers for the InGaAs PIN photodiode are grown by MBE. Then channels are etched on the wafer using a SiO_2 mask and wet chemical etching. The layers for the MODFET are then grown in the channels by MBE with the SiO_2 mask in place. No growth occurs on the top of the SiO_2. The performance of both the MODFET and PIN photodiodes on the regrown structure is comparable to that of single devices.

A measure of the performance of a receiver is the receiver sensitivity. Hence it is important to compare the sensitivity of integrated PIN-amplifier-based receivers to those of their hybrid counterparts. The measured data as a function of bit rate are shown in Fig. 12.13. Note that the hybrid-technology-based receivers have 3–4-dB better sensitivity than the integrated receivers.

12.3.2 Transmitter OEICs

Several types of transmitter OEICs, both using GaAs lasers and InGaAsP lasers, have been reported.[29–34] GaAs OEICs generally use metal-semiconductor field-effect transistors (MESFETs) in their driver circuits. InGaAsP OEICs use HBTs in their driver circuits.



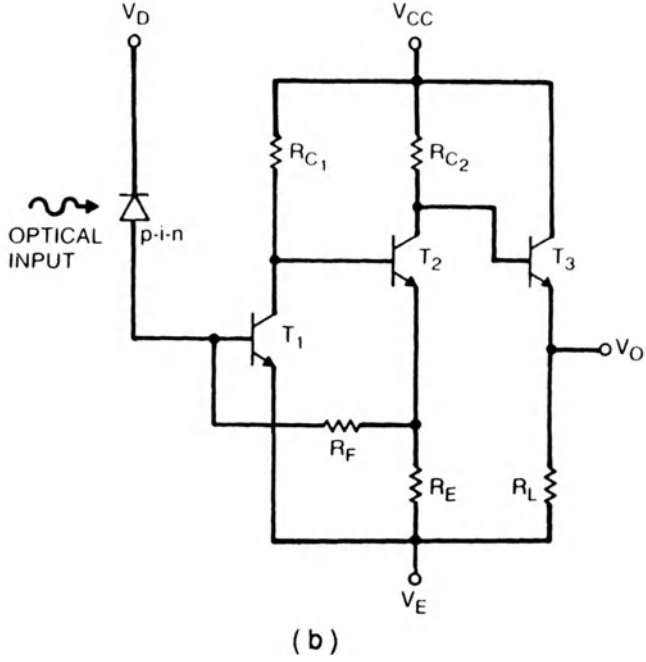


Fig. 12.11 (a) Schematic of a PIN photodiode integrated with HBT amplifiers. (b) Circuit diagram of the monolithic photoreceiver. (After Ref. 25 © 1990 IEEE)

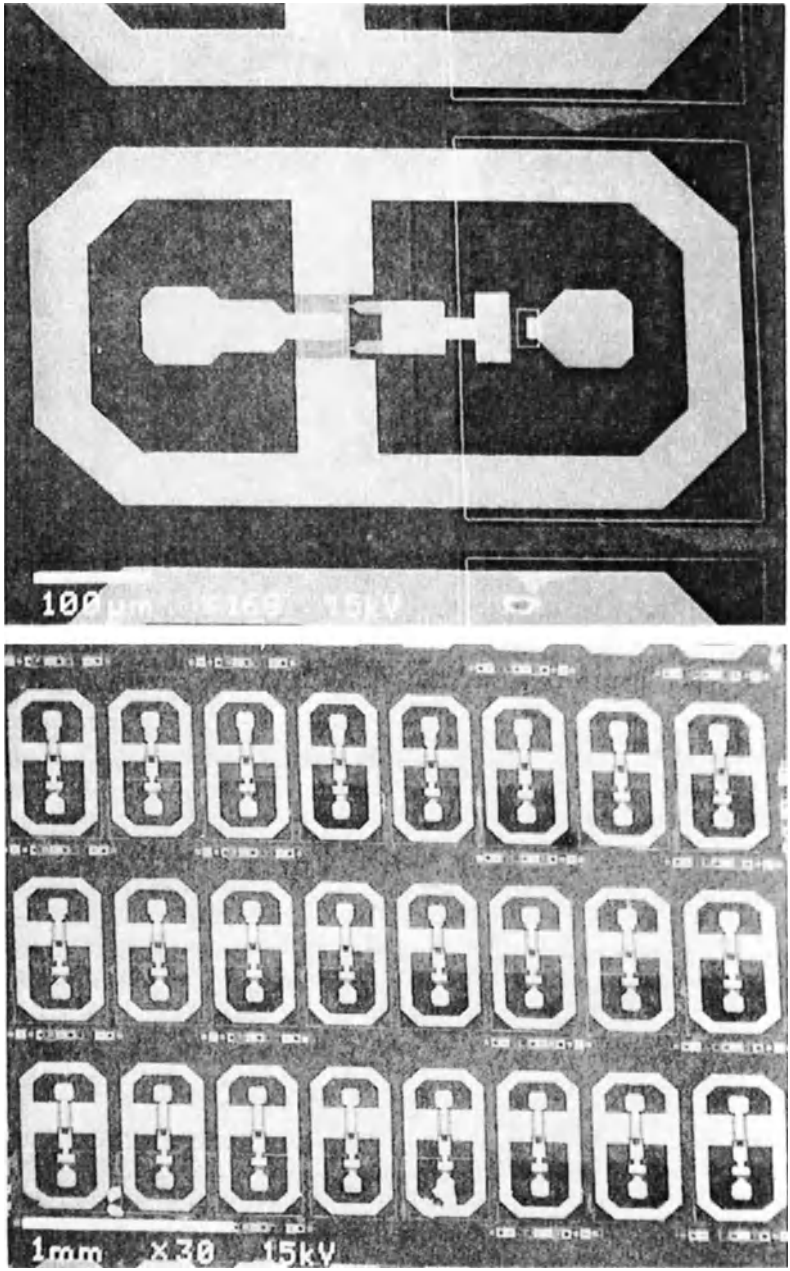


Fig. 12.12 Photomicrograph of a PIN photodiode integrated with a MODFET amplifier. Also shown is the photomicrograph of the fabricated device. (Courtesy P. Berger)

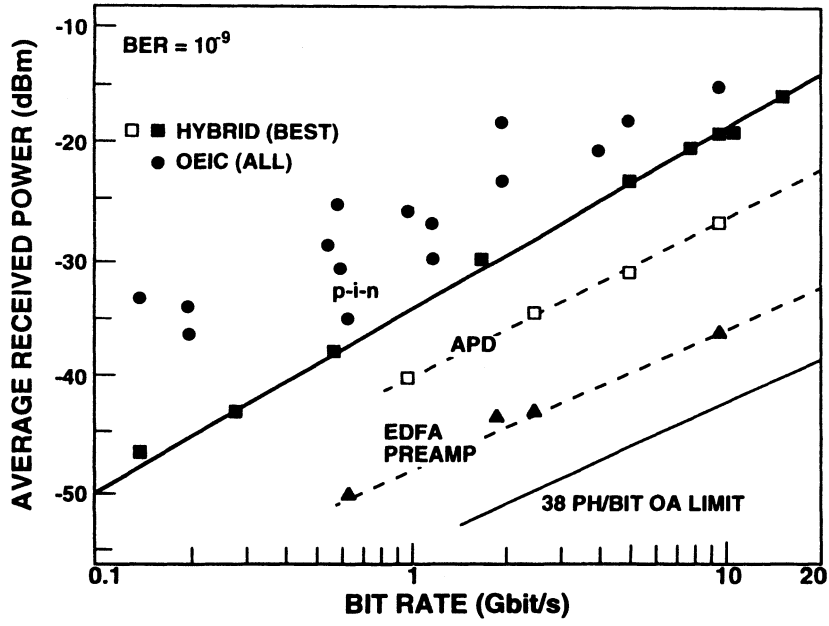


Fig. 12.13 Receiver sensitivity (direct detection) as a function of data rate for integrated and hybrid receivers.

The circuit diagram and the structure of an integrated GaAs transmitter OEIC is shown in Fig. 12.14. The GaAs laser is a single quantum-well laser. The circuit has a waveguide-type photodiode for monitoring the performance of the laser. The transistors T_1 and T_2 are arranged as a differential pair for better stability. The circuit shown in Fig. 12.14 operated at 1 Gb/s. An InGaAsP-laser-based transitter OEIC has also been fabricated.[32] This device structure is shown in Fig. 12.15. It utilizes HBTs for driving the laser and transistors are arranged as a differential pair. The epitaxial layers were grown by MOVPE on an Si-InP substrate. The circuit operated at 5 Gb/s.

12.3.3 Regenerator OEICs

Integration of a light source, a photodiode and an electronic circuit will, in principle, lead to an optical regenerator. Up to now, such monolithic regenerators have behaved as simple optical amplifiers. An example of such a device built on an Si-GaAs substrate is shown in Fig. 12.16. It has 3 FETs and a laser. One of the FETs also serves as a photodiode for the input light. An overall power gain of 10 dB has been achieved for this OEIC. Similar regenerator circuits using InP material systems have also been reported.[35]

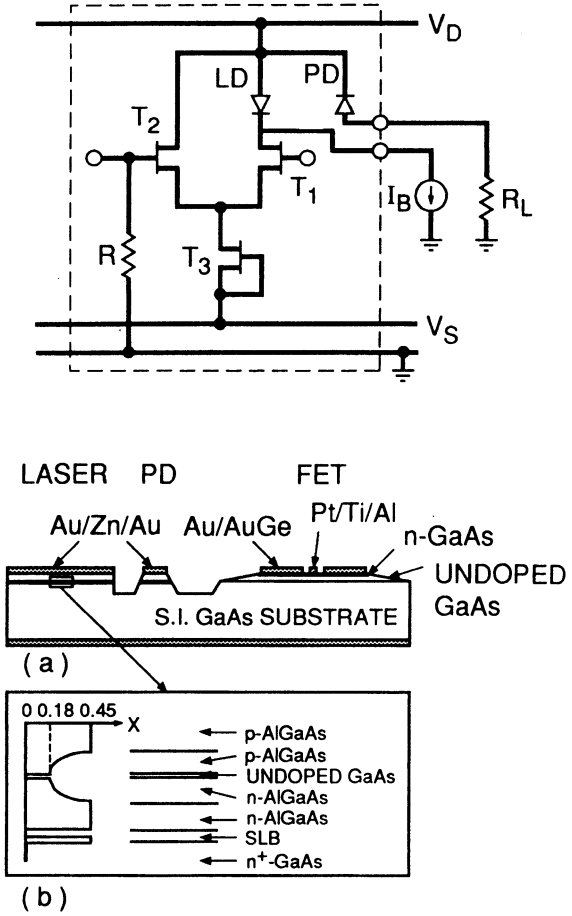


Fig. 12.14 GaAs laser-transmitter OEIC. (After Ref. 29)

12.3.4 Logic OEICs

OEICs which perform some optical logic functions have been fabricated.[36] Many of these OEIC developments are aimed for applications in computing and switching systems where parallel two-dimensional interconnection technology is important. The surface-emitting laser or LED is the source of choice for these two-dimensional interconnection systems.

An example of an optical bistable switch using a surface-emitting laser is shown in Fig. 12.17.[36] The switching operation is performed by a p-n-p-n photothyristor monolithically integrated with the surface-emitting laser. The device has a high contrast (~ 30 dB) between its on and off states.

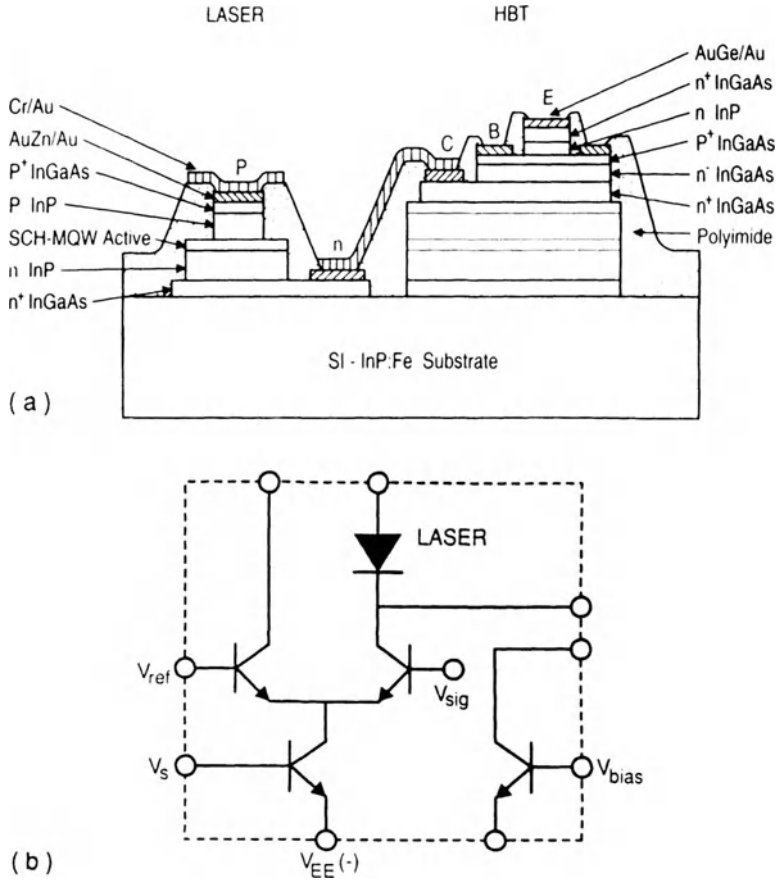


Fig. 12.15 (a) Schematic and (b) circuit diagram of an InGaAsP laser-transmitter OEIC. (After Ref. 32 © 1991 IEEE)

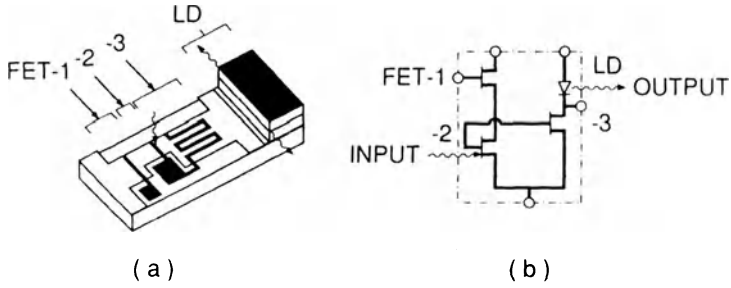


Fig. 12.16 (a) Schematic and (b) circuit diagram of an integrated AlGaAs-GaAs optical regenerator consisting of a laser and three MESFETS. (After Ref. 34)

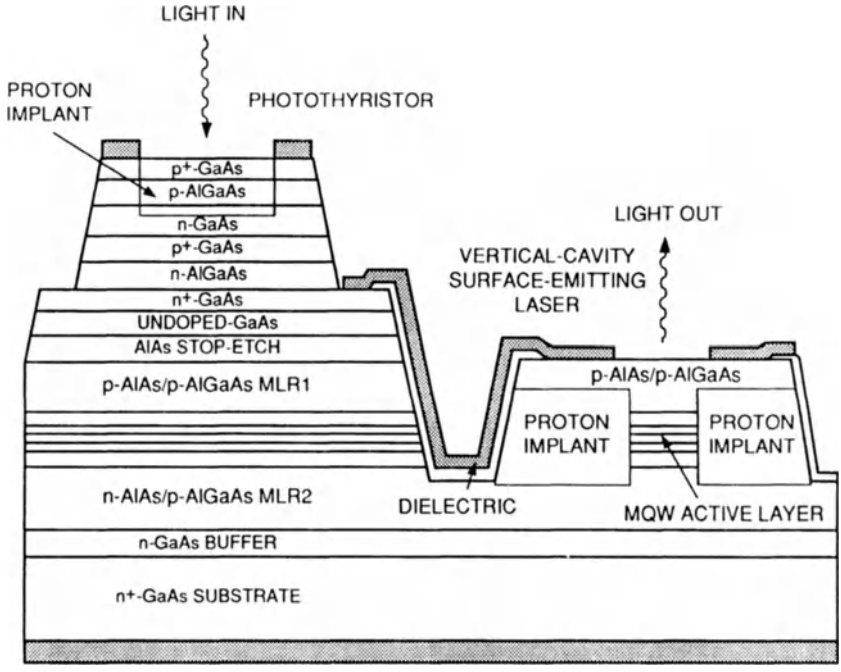


Fig. 12.17 Schematic of a bistable switch which consists of an integrated photothyristor and a surface-emitting laser. (After Ref. 36)

REFERENCES

1. Kobayashi, K. Chap. 15 in *Optical Fiber Telecommunications II*, ed. by S. E. Miller and I. P. Kaminow. San Diego: Academic Press, 1988.
2. Koch, T. L. and U. Koren. *IEEE J. Quantum Electron.* **QE-27,** 641 (1991).
3. Wada, O., T. Sakurai, and T. Nakagami. *IEEE J. Quantum Electron.* **QE-22,** 850 (1986).
4. Forrest, S. R. *J. Lightwave Technol.* **LT-3,** 1248 (1985).
5. Deimel, P. P., J. Cheng, S. R. Forrest, P. H. S. Hu, R. B. Huntington, R. C. Miller, J. R. Potopowicz, D. D. Roccasecca, and C. W. Seabury. *J. Lightwave Technol.* **LT-3,** 988 (1985).
6. Brown, M. G., P. H. S. Hu, D. R. Kaplan, Y. Ota, C. W. Seabury, M. A. Washington, E. E. Becker, J. G. Johnson, M. Koza, and J. R. Potopowicz. *J. Lightwave Technol.* **LT-4,** 283 (1986).
7. Dixon, R. W., and N. K. Dutta. *AT&T Tech. J.* **66,** 73 (1987).
8. Iga, K., and B. I. Miller. *Electron. Lett.* **16,** 830 (1980).
9. Wright, P. D., R. J. Nelson, and R. B. Wilson. *IEEE J. Quantum Electron.* **QE-18,** 249 (1982).
10. Dutta, N. K., J. L. Zilko, A. B. Piccirilli, and R. L. Brown. *Electron. Lett.* **24,** 335 (1988).
11. Koszi, L. A., A. K. Chin, B. P. Segner, T. M. Shen, and N. K. Dutta. *Electron. Lett.* **21,** 1209 (1985).
12. Liou, K. Y., U. Koren, S. Chandrasekhar, T. L. Koch, A. Shahar, C. A. Burrus, and R. P. Gnall. *Appl. Phys. Lett.* **54,** 114 (1989).

13. Hasnain, G., K. Tai, Y. H. Wang, J. D. Wynn, K. D. Choquette, B. E. Weir, N. K. Dutta, and A. Y. Cho. *Electron. Lett.* **27,** 1630 (1991).
14. Suzuki, M., Y. Noda, H. Tanaka, S. Akiba, Y. Kushiro, and H. Isshiti. *J. Lightwave Technol.* **LT-5,** 1277 (1987).
15. Soda, H., M. Furutsu, K. Sato, N. Okazaki, Y. Yamazaki, H. Nishimoto, and H. Ishikawa. *Electron. Lett.* **26,** 9 (1990).
16. Koren, U., B. I. Miller, G. Raybon, M. Oron, M. G. Young, T. L. Koch, J. L. DeMiguel, M. Chien, B. Tell, K. Brown-Goebler, and C. A. Burrus. *Appl. Phys. Lett.* **57,** 1375 (1990).
17. Dutta, N. K., J. Lopata, R. Logan, and T. Tanbun-Ek. *Appl. Phys. Lett.* **59,** 1676 (1991).
18. Welch, D. F., D. Mehuys, R. Parke, R. Waarts, D. Scifres, and W. Streifer. *Electron. Lett.* **17,** 1327 (1990).
19. Takenchi, H., K. Kasaya, Y. Kondo, H. Yasaka, K. Oe, and Y. Imamura. *IEEE Photon. Technol. Lett.* **1,** 398 (1989).
20. Koch, T. L., U. Koren, R. P. Gnall, F. S. Choa, F. Hernandez-Gil, C. A. Burrus, M. G. Young, M. Oron, and B. I. Miller. *Electron. Lett.* **25,** 1621 (1989).
21. Lee, C. P., S. Margalit, I. Ury, and A. Yariv. *Appl. Phys. Lett.* **32,** 806 (1978).
22. Yariv, A. *IEEE Trans. Electron. Dev.* **ED-31,** 1656 (1984).
23. Leheny, R., R. E. Nahory, J. C. DeWinter, R. J. Martin, and E. D. Beebe. *Tech. Digest Int. Electron. Dev. Mtg., Washington, D.C.*, 276 (1981).
24. Chandrasekhar, S., A. H. Gnauck, W. T. Tsang, F. S. Choa, and G. J. Qua. *IEEE Photon. Technol. Lett.* **3,** 823 (1991).
25. Chandrasekhar, S., B. C. Johnson, M. Bonnemason, E. Tokumitsu, A. H. Gnauck, A. G. Dentai, C. H. Joyner, J. S. Perino, G. J. Qua, and E. M. Monberg. *IEEE Photon. Technol. Lett.* **2,** 505 (1990).
26. Chang, G. K., W. P. Hong, J. L. Gimlett, R. Bhat, C. K. Nguyen, G. Sasaki, and J. C. Young. *IEEE Photon. Technol. Lett.* **2,** 197 (1990).
27. Uchida, N., Y. Akahori, M. Ikeda, A. Kohzen, J. I. Yoshida, T. Kokubun, and K. Suto. *IEEE Photon. Technol. Lett.* **3,** 540 (1991).
28. Lai, R., P. K. Bhattacharya, D. Pavlidis, and T. Brook. *Electron. Lett.* **27,** 364 (1991).
29. Kuno, M., T. Sanada, H. Nobuhara, M. Makiuchi, T. Fuji, O. Wada, and T. Sakurai. *Appl. Phys. Lett.* **49,** 1575 (1986).
30. Shibata, J., I. Nakao, Y. Sasai, S. Kimura, N. Hase, and H. Serigawa. *Appl. Phys. Lett.* **45,** 191 (1984).
31. Yamakoshi, S., T. Sanada, O. Wada, T. Fuji, and T. Sakurai. *Electron. Lett.* **19,** 1020 (1983).
32. Liou, K. Y., S. Chandrasekhar, A. G. Dentai, E. C. Burrows, G. J. Qua, C. H. Joyner, and C. S. Burrus. *IEEE Photon. Technol. Lett.* **3,** 928 (1991).
33. Bar-chaim, N., S. Margalit, A. Yariv, and I. Ury. *IEEE Trans. Electron. Dev.* **ED-29,** 1372 (1982).
34. Hata, S., M. Ikeda, S. Kondo, and Y. Noguchi. *Tech. Digest IECE Conf., Japan*, 85 (1985).
35. See for example, Proc. Int. Symp. Adv. in Interconnection and Packaging, vol. 1389, ed. by A. P. DeFonzo. Bellingham, WA: SPIE, 1990.
36. P. Zhou, J. Cheng, S. Z. Sun, C. F. Schaus, C. Hains, D. R. Myers, and G. A. Vaister. *Appl. Phys. Lett.* **59,** 2648 (1991).

Chapter 13
INFRARED AND VISIBLE SEMICONDUCTOR LASERS

13.1 LEAD-SALT LASERS

In previous chapters we paid particular attention to GaAs and InGaAsP semiconductor lasers that emit light in the wavelength range of 0.8–1.6 μm in view of their important applications in optical data storage and optical fiber communications systems. However, longer-wavelength semiconductor lasers have also been of considerable interest and have found applications in molecular spectroscopy among other things, because of their large spectral tuning range. Figure 1.3 showed the wavelength range covered by various material systems. Of particular interest are the lead-salt semiconductor lasers covering a large spectral range (2.5–34 μm) in the far-infrared region. Since they are relatively well developed, the bulk of this chapter is devoted to their study. The last section of this chapter briefly mentions infrared and visible semiconductor lasers fabricated by using other material systems. The interest in lead-salt semiconductor lasers is generated by the wide range of emission wavelengths that can be obtained by varying the composition of the various constituents. Injection lasers of the ternary materials $Pb_{1-x}Sn_xTe$ and $PbS_{1-x}Se_x$ can provide emission wavelengths in the range of 3 μm to ~34 μm. The band gap of these materials also varies considerably with temperature; as a result, a large degree of spectral tuning can be obtained simply by varying the operating temperature. Lead-salt lasers have been used for various spectroscopic studies, including high-resolution spectroscopy and detection of low concentrations of pollutants. Several reviews of the performance of lead-salt diode lasers have been published.[1–5]

13.2 MATERIALS AND PHYSICAL PROPERTIES

Lead-salt lasers, as they are commonly called, usually refer to the ternary alloys $Pb_{1-x}Sn_xTe$, $PbS_{1-x}Se_x$, and $Pb_{1-x}Sn_xSe$ and the quaternary material

Table 13.1 Band Gaps, Lattice Constants, and Effective Masses of PbS, PbSe, PbTe, and SnTe

COMPOUND	BAND GAP (eV) 77 K	BAND GAP (eV) 300 K	LATTICE CONSTANT a (Å)	EFFECTIVE MASS $m_c = m_v$
PbS	0.307	0.41	5.936	$0.22m_0$
PbSe	0.176	0.27	6.124	$0.12m_0$
PbTe	0.217	0.31	6.462	$0.13m_0$
SnTe	—	—	6.303	–

Source: Casey, H. C., Jr., and M. B. Panish. *Heterostructure Lasers*, Part B. New York: Academic Press, 1978, p. 61.

system of $Pb_{1-x}Sn_xSe_yTe_{1-y}$. The various II–VI compounds that make up these ternaries and quaternary are PbS, PbSe, PbTe, SnTe, and SnSe. Table 13.1 lists the band gaps, lattice constants, and effective masses of the first four of these binaries. These four compounds crystallize in the cubic rock-salt structure (similar to NaCl) or in the orthorhombic structure. This allows the formation of $Pb_{1-x}Sn_xTe$ and $PbS_{1-x}Se_x$ for all values of the molar fraction x, whereas $Pb_{1-x}Sn_xSe$ can be fabricated only for lead-rich compositions ($x < 0.4$). Figure 13.1 shows the variation in band gap and lattice constant of four ternaries with respect to their composition. The figure assumes a linear extrapolation between the binary endpoints. The solid line is for the band gap, and the dashed line is for the lattice constant. Most of the lead-salt lasers use $Pb_{1-x}Sn_xTe$, $PbS_{1-x}Sn_x$ or $Pb_{1-x}Sn_xSe$ as the active region. The ternary $PbTe_{1-x}Se_x$, which can also be grown without a large lattice mismatch, is generally used in cladding layers. Also, quaternary PbSnSeTe

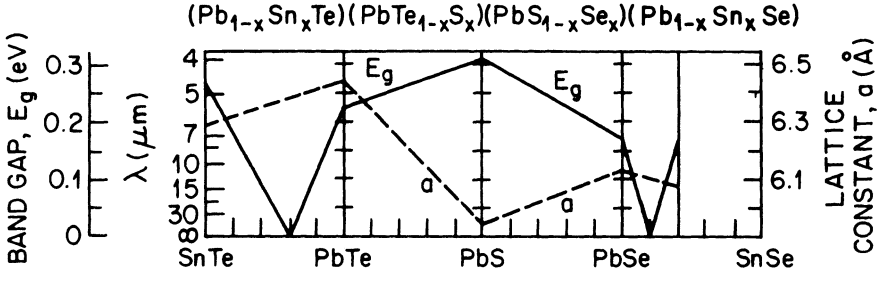


Fig. 13.1 Band gap (solid line) and lattice constant (dashed line) of $Pb_{1-x}Sn_xTe$, $PbTe_{1-x}S_x$, $PbS_{1-x}Se_x$, and $Pb_{1-x}Sn_xSe$ at 77 K as a function of composition x. The lines are drawn assuming linear extrapolation between the endpoints. (After Ref. 5)

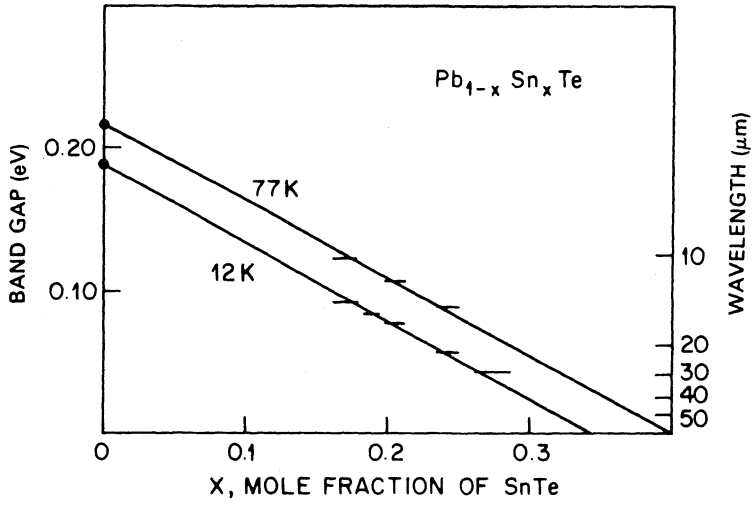


Fig. 13.2 Measured band gap of $Pb_{1-x}Sn_xTe$ as a function of composition at two temperatures. Extrapolated straight lines show that the band gap vanishes at a specific value of the mole fraction x. (After Ref. 6 © 1969 IEEE)

injection lasers on PbTe substrates have been grown using PbTeSe cladding lasers.

Figure 13.1 shows that the band gap of $Pb_{1-x}Sn_xTe$ and $Pb_{1-x}Sn_xSe$ approaches zero at a certain composition. In principle, this should allow the emission of very long wavelengths using these materials. Figure 13.2 shows the measured band gap of $Pb_{1-x}Sn_xTe$ obtained from the laser-emission wavelength as a function of the composition x.[6] For $x < 0.35$, the band gap at 12 K may be represented by[7]

$$E_g \text{ (in eV)} = 0.19 - 0.543x. \tag{13.2.1}$$

Coherent emission from $Pb_{1-x}Sn_xTe$ was first observed by Dimmock et al.[8] Laser emission has been observed in $Pb_{1-x}Sn_xTe$ up to 34 μm at 12 K. The compositional dependence of the band gap of $Pb_{1-x}Sn_xSe$ can be obtained[9] from a measurement of the laser-emission wavelength. Figure 13.3 shows the measured values at 12 K. Another quantity of interest for the fabrication of double-heterostructure laser diodes is the refractive index of the material. Figure 13.4 shows the data[10] for the refractive indices of PbTe, PbSe, and PbS at 77 K, 300 K, and 373 K. Similar values have been reported for the refractive indices of the binaries.[11] The following analytic form for the high-frequency dielectric constant (ϵ_∞) of $Pb_{1-x}Sn_xTe$ and $Pb_{1-x}Sn_xSe$

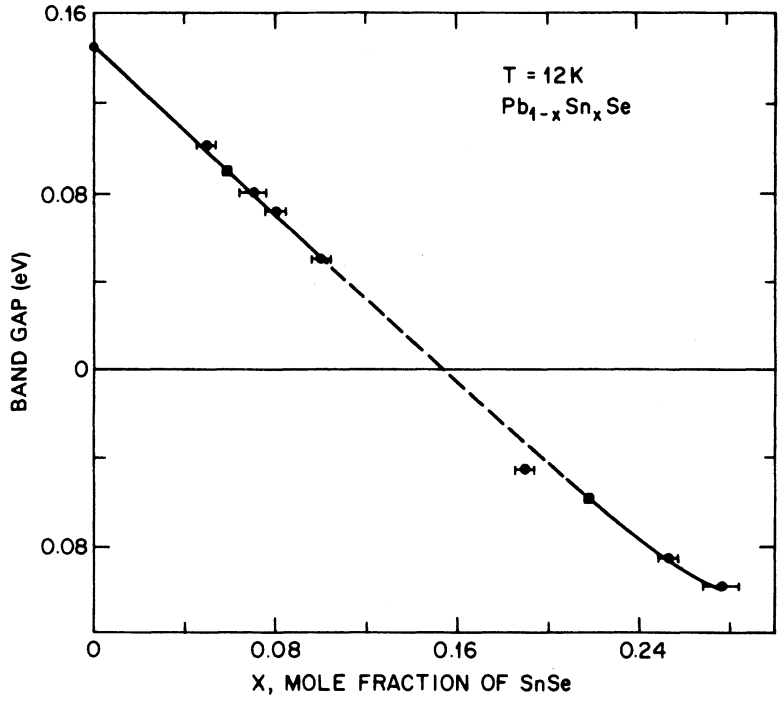


Fig. 13.3 Measured compositional dependence of the band gap of $Pb_{1-x}Sn_xSe$ at 12 K. Solid line shows a linear fit, while the dashed portion indicates band inversion. (After Ref. 9)

(for $0 \leq x \leq 0.2$) has been suggested[5]:

$$\epsilon_\infty = \begin{cases} 35 + 50x & (T = 4.2\text{K}, Pb_{1-x}Sn_xTe) \\ (5.5 + 0.8x)^2 & (T = 100\text{ K}, Pb_{1-x}Sn_xSe). \end{cases} \tag{13.2.2}$$

By varying the composition x, an index difference sufficient for dielectric waveguiding can be obtained between active and cladding layers.

13.3 BAND STRUCTURE

The band structure of the lead-salt narrow-gap semiconductors is considerably different from that of III–V semiconductors (discussed in Chapter 3), where the conduction-band minimum and the valence-band maximum occur at the Γ-point and the bands are isotropic with spherical constant energy surfaces in $\mathbf{k}$ space. In lead-salt compounds there are four equivalent conduction-band minima and valence-band maxima at the L-points of $\mathbf{k}$

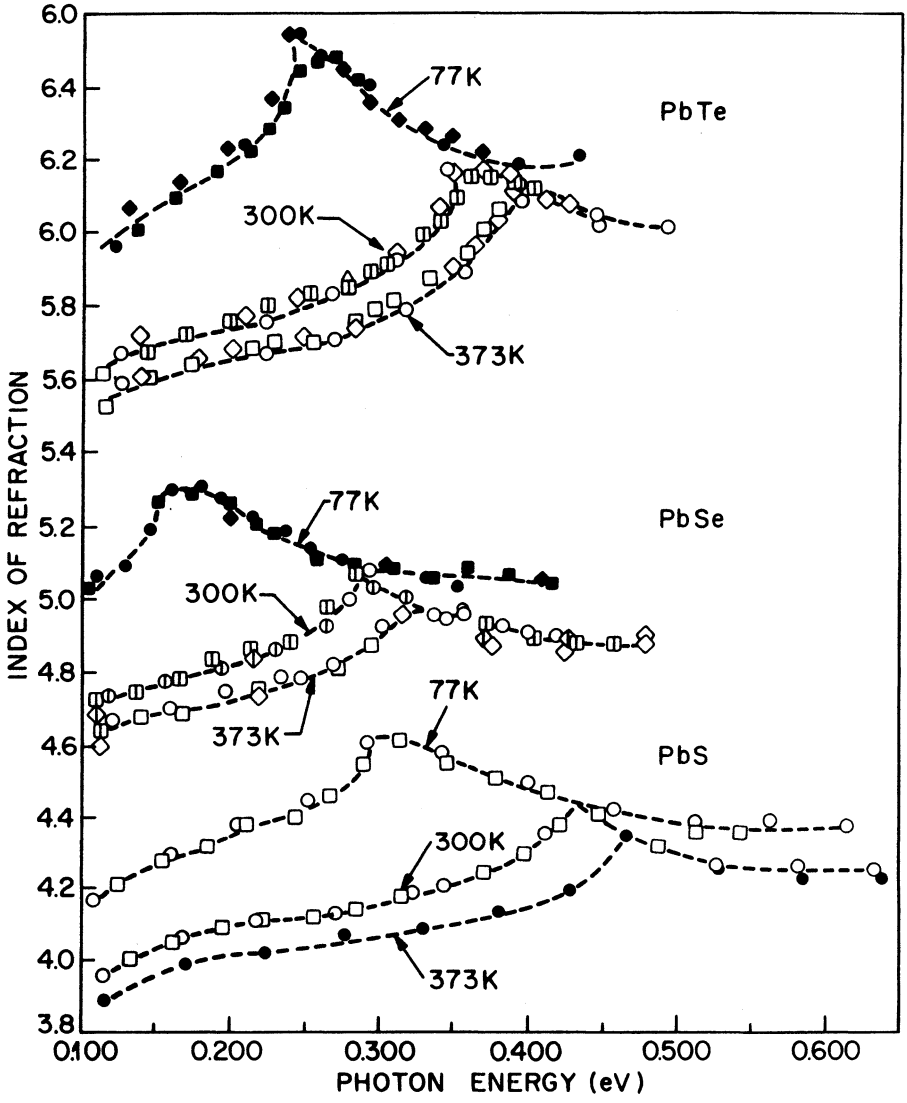


Fig. 13.4 Refractive index of PbTe, PbSe, and PbS at three different temperatures as a function of the photon energy. (After Ref. 10)

space. The minimum of each conduction band occurs at the same point in **k** space as the maximum of the corresponding valence band so that direct transitions (electron-hole recombinations that conserve momentum) are allowed. However, the constant energy surfaces in **k** space are prolate

spheroids characterized by the longitudinal and transverse effective masses m_l and m_t respectively. The ratio m_l/m_t is ~ 10 for PbTe and $Pb_{1-x}Sn_xTe$ and ~ 1.5–2 for $Pb_{1-x}Sn_xSe$ and $PbS_{1-x}Se_x$. Furthermore, conduction and valence bands are mirror images of each other with almost identical effective masses.

We noted previously that with increasing Sn concentration the band gap of $Pb_{1-x}Sn_xTe$, as measured optically, approaches zero (Fig. 13.2). However, tunneling experiments have shown that SnTe is a semiconductor with a band gap of 0.3 eV at 4.2 K. These observations suggest the following model[7] for the band structure of $Pb_{1-x}Sn_xTe$ as its composition varies. The valence-band edge of PbTe is an L_6^+ state and the conduction-band edge is an L_6^- state. With increasing Sn composition, the band gap initially decreases as the states L_6^+ and L_6^- approach each other, reaching zero at some composition; the gap then increases with the L_6^+ state now forming the conduction-band edge and the L_6^- state forming the valence-band edge. This phenomenon is referred to as *band inversion.* Figure 13.5 shows schematically the band inversion in $Pb_{1-x}Sn_xTe$.

The energy–wave-vector relation of the conduction and valence bands of PbTe-type narrow-gap semiconductors in the two-band approximation of Kane[12] is given by

$$E = \pm(\tfrac{1}{4}E_g^2 + p_t^2k_t^2 + p_l^2k_l^2)^{1/2} - \tfrac{1}{2}E_g \qquad (13.3.1)$$

where the + sign is for the conduction band and the − sign is for the valence band. The band gap is E_g, k_t and k_l are the transverse and

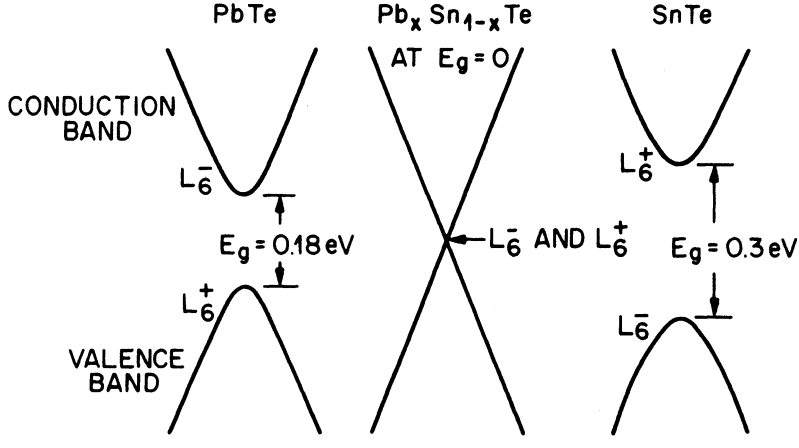


Fig. 13.5 Schematic representation of band inversion in $Pb_{1-x}Sn_xTe$ at 12 K. (After Ref. 7)

longitudinal components of the wave vector, and p_t^2 and p_l^2 are related to the effective masses m_t and m_l. Note that the effective masses of the valence and conduction bands are the same and that the two bands are mirror images of each other. The following approximate dispersion relations near the band edge can be obtained from Eq. (13.3.1) by expanding the square root in a binomial series:

$$E_c(k) = \frac{\hbar^2}{2m_l} k_l^2 + \frac{\hbar^2}{2m_t} k_t^2 \tag{13.3.2a}$$

$$E_v(k) = \frac{\hbar^2}{2m_l} k_l^2 + \frac{\hbar^2}{2m_t} k_t^2 \tag{13.3.2b}$$

where

$$m_{l,t} = \frac{\hbar^2 E_g}{2p_{l,t}^2}$$

and $E_c(k)$ and $E_v(k)$ are measured from the respective conduction- and valence-band edges into the band. The number of electrons per unit volume at a temperature T is given by

$$n = \frac{2\eta}{(2\pi)^3} \int d^3\mathbf{k} f_c(E_c) \tag{13.3.3}$$

where

$$f_c(E_c) = \frac{1}{1 + \exp[(E_c - E_{fc})/k_B T]}$$

is the Fermi factor, η is the number of conduction-band valleys ($\eta = 4$ for PbTe), and E_{fc} is the Fermi energy. A similar equation holds for holes. Using the approximate dispersion relation of Eq. (13.3.2), the carrier density becomes

$$n = N_c \int E^{1/2} f_c(E_c)\, dE_c \tag{13.3.4}$$

where

$$N_c = \eta m_l^{1/2} m_t (2\pi k_B T/h^2)^{3/2}. \tag{13.3.5}$$

Under the Boltzmann or nondegenerate approximation, an equation similar to Eq. (3.2.30) is obtained for both electrons and holes, and the Fermi factors

are

$$f_c(E_c) = \frac{n}{N_c} \exp\left(\frac{-E_c}{k_B T}\right) \tag{13.3.6}$$

$$f_v(E_c) = \frac{p}{N_v} \exp\left(\frac{-E_v}{k_B T}\right) \tag{13.3.7}$$

where $N_c = N_v$ and is given by Eq. (13.3.5).

13.4 OPTICAL GAIN

The optical gain can be calculated following Sec. 3.2. However, the analysis should be modified to take into account the anisotropic nature of the energy bands governed by the different effective masses m_l and m_t in the dispersion relations (13.3.2). The absorption coefficient between two discrete levels given by Eq. (3.2.21) is

$$\alpha(E) = \frac{q^2 h}{2\epsilon_0 m_0^2 c \bar{\mu} E} |M_{if}|^2 \delta(E_i - E_f - E) \tag{13.4.1}$$

where $|M_{if}|^2$ is the momentum-matrix element and E is the photon energy. We assume that the band-to-band transitions conserve momentum (**k**-conserving transitions). The matrix element $|M_{if}|^2$ is then given by [see Eq. (3.2.34)]

$$|M_{if}|^2 = |M_0|^2 \frac{(2\pi)^3}{V} \delta(\mathbf{k}_c - \mathbf{k}_v) \tag{13.4.2}$$

where the δ function assures conservation of momentum and V is the volume of the semiconductor. The matrix element $|M_0|^2$ is related to the interband matrix elements p_l and p_t in Eq. (13.3.1). We assume that the radiation is polarized along the x direction that makes an angle θ with the crystallographic [111] direction. Then

$$|M_0|^2 = (p_l^2 \cos^2\theta + p_t^2 \sin^2\theta)(m_0^2/\hbar^2). \tag{13.4.3}$$

When contributions from all four prolate spheroids are summed and the result is averaged over all directions, we obtain

$$|M_0|^2 = 4(\tfrac{1}{3}p_l^2 + \tfrac{2}{3}p_t^2)\frac{m_0^2}{\hbar^2} \tag{13.4.4}$$

which is independent of the polarization as required by the symmetry of a cubic crystal. The expression for $\alpha(E)$ after summing over the available states is [see Eq. (3.2.39)]

$$\alpha(E) = \frac{q^2 h}{2\epsilon_0 m_0^2 c\bar{\mu} E} |M_0|^2 \frac{(2\pi)^3}{V} 2\left(\frac{V}{(2\pi)^3}\right)^2 \frac{1}{V}$$
$$\times \iint (1 - f_c - f_v)\, d^3\mathbf{k}_c\, d^3\mathbf{k}_v\, \delta(\mathbf{k}_c - \mathbf{k}_v)\delta(E_i - E_f - E). \quad (13.4.5)$$

It can be simplified to become

$$\alpha(E) = \frac{q^2 h |M_0|^2}{\epsilon_0 m_0^2 c\bar{\mu} E(2\pi)^3} [1 - f_c(E_c) - f_v(E_v)] I \quad (13.4.6)$$

where

$$E_c = E_v = (E - E_g)/2$$
$$I = \int \delta(E_i - E_f - E)\, d^3\mathbf{k}_c. \quad (13.4.7)$$

The evaluation of the integral I is carried out using the dispersion relations given by Eq. (13.3.2). The explicit dependence of $E_i - E_f$ on $\mathbf{k}_c = (k_x, k_y, k_z)$ is given by

$$E_i - E_f = (E_g - E) + 2\left[\frac{\hbar^2}{2m_l} k_x^2 + \frac{\hbar^2}{2m_t}(k_y^2 + k_z^2)\right]. \quad (13.4.8)$$

The integration in Eq. (13.4.7) should take into account the effective-mass anisotropy arising from the prolate-spheroid shape of the constant energy surface in the $\mathbf{k}_c$ space. A simple way is to define a new wave vector

$$\mathbf{k}_c' = (k_x', k_y, k_z) \qquad k_x' = k_x(m_t/m_l)^{1/2} \quad (13.4.9)$$

so that the constant energy surface becomes spherical in the new integration variable:

$$E_i - E_f = (E_g - E) + 2(\hbar^2/2m_t)k_c'^2. \quad (13.4.10)$$

The evaluation of I in Eq. (13.4.7) is now straightforward and the result is

$$I = \left(\frac{m_l}{m_t}\right)^{1/2} \int \delta(E_i - E_f)\, d^3\mathbf{k}_c'$$

$$= 2\pi\left(\frac{m_l}{m_t}\right)^{1/2}\left(\frac{2m_t}{\hbar^2}\right)^{3/2}(E - E_g)^{1/2}. \tag{13.4.11}$$

Equations (13.4.6) and (13.4.11) provide the expression for the absorption coefficient in lead-salt semiconductors. Expressions for the total spontaneous-emission rate can be derived in a manner similar to that discussed in Sec. 3.2.2.

Anderson[13] has calculated the gain-current relation in $Pb_{1-x}Sn_xTe$ lasers using the nonparabolic terms in the energy–wave-vector relation for the conduction and valence bands. This model can be used to calculate the optical gain as a function of injected carrier density.[14] Figure 13.6 shows the results for $Pb_{0.82}Sn_{0.18}Te$ ($\lambda \cong 10\ \mu m$) at 77 K. The material is p type, and the four curves correspond to different doping levels. The parameters used in the calculation are from Ref. 15. A lower injected-carrier concentration is needed to obtain the same gain if the active layer is heavily doped. This behavior is similar to that for InGaAsP shown in Fig. 3.10. Figure 13.6 shows that the optical gain g varies approximately linearly with the injected carrier density; i.e.,

$$g = a(n - n_0) \tag{13.4.12}$$

where a is the gain coefficient and n_0 is the injected carrier density at transparency. The approximate validity of Eq. (13.4.12) for lead-salt lasers suggests that the results presented in Chapters 2 and 6 can be applied to these lasers. In particular, the threshold current density can be obtained using Eqs. (2.6.2) and (2.6.3). Note, however, that although the injected carrier density n_0 needed for transparency ($g = 0$ in Fig. 13.6) is lower for higher doping, the threshold current may not be lower. The latter is affected by free-carrier absorption as well as by nonradiative Auger recombination, both of which increase with increased doping.

11.5 AUGER RECOMBINATION

The various Auger recombination mechanisms in a direct-gap semiconductor are discussed in Sec. 3.3. In this section we calculate the Auger lifetime for the CCCH Auger process (see Fig. 3.13) using the band structure appropriate for lead-salt semiconductors. In the nondegenerate approximation, the Auger

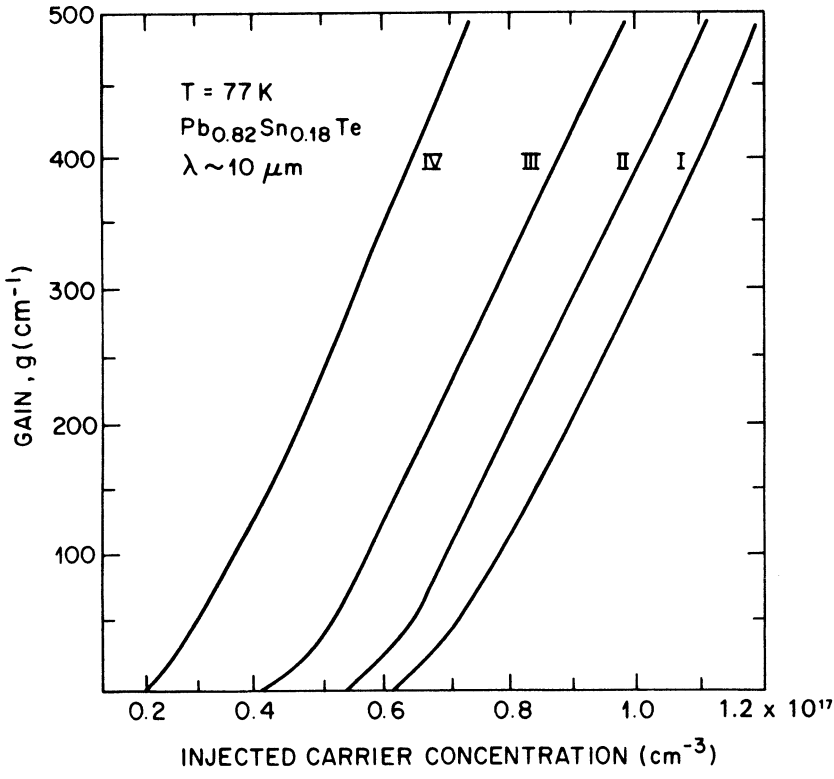


Fig. 13.6 Calculated maximum gain at 77 K as a function of injected carrier concentration in p-type $Pb_{0.82}Sn_{0.18}Te$ ($\lambda \cong 10$ μm) for different doping levels: (I) 4×10^{16} cm^{-3}, (II) 6×10^{16} cm^{-3}, (III) 1×10^{17} cm^{-3}, and (IV) 2×10^{17} cm^{-3}. (After Ref. 14 © 1983 IEEE)

rate R_a for spherical parabolic bands varies as[16] [Eqs. (3.3.26) and (3.3.29) with $m_{c0} = m_c$]

$$R_a \propto n^2 p \exp\left[\frac{-\mu E_g}{(1+\mu)k_B T}\right] \tag{13.5.1}$$

where $\mu = m_c/m_v$, the ratio of the effective masses of the conduction and valence bands. In PbTe semiconductors, four conduction and four valence bands exist at the center of the (111) zone faces. For Auger recombination involving electrons in the same pair of bands, Eq. (13.5.1) still holds for spheroidal energy surfaces so that the Auger rate for such a process is given by

$$R_a \propto n^2 p \exp\left(\frac{-E_g}{2k_B T}\right). \tag{13.5.2}$$

In writing Eq. (13.5.2), we have assumed that $\mu = 1$ because the conduction- and valence-band effective masses are nearly equal for PbTe semiconductors, as shown by Eq. (13.3.2).

Emtage[17] showed that if the interacting electrons in the Auger process belong to different conduction-band valleys, the Auger rate is significantly enhanced and Eq. (13.5.2) takes the form

$$R_a \propto n^2 p \exp\left(\frac{-rE_g}{2k_B T}\right) \tag{13.5.3}$$

where $r = m_t/m_l$, the ratio of the transverse to longitudinal effective masses. The parameter $r \cong 0.1$ for PbTe and $Pb_{1-x}Sn_xTe$ semiconductors. We now generalize this result to take into account nonparabolicity.

Figure 13.7 diagrams the CCCH Auger process with electrons in different conduction-band valleys. Electron 1 is in conduction band a and electron 2 is in conduction band b. After collision, electron 1 transfers to a state 1′ in valence band a and electron 2 transfers to the excited state 2′ in conduction band b. When $r \cong 0.1$, the principal axes of bands a and b can be taken at right angles to each other, as shown in Fig. 13.7b. The electron energies in bands a and b in the parabolic approximation are

$$E_a = \frac{\hbar^2}{2m_l} k_x^2 + \frac{\hbar^2}{2m_t}(k_y^2 + k_z^2) \tag{13.5.4a}$$

$$E_b = \frac{\hbar^2}{2m_l} k_y^2 + \frac{\hbar^2}{2m_t}(k_x^2 + k_z^2). \tag{13.5.4b}$$

The holes are assumed to have the same effective mass as electrons.

The expression for the Auger rate per unit volume is given by [from Eq. (3.3.24)]

$$R_a = \eta(\eta - 1)\left(\frac{1}{(2\pi)^3}\right)^4 \frac{8\pi}{\hbar} \iiiint |M_{if}|^2 P(1, 1', 2, 2') \times \delta(E_i - E_f)\delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_1' - \mathbf{k}_2')\, d^3\mathbf{k}_1\, d^3\mathbf{k}_1'\, d^3\mathbf{k}_2\, d^3\mathbf{k}_2' \tag{13.5.5}$$

where η is the number of equivalent conduction-band and valence-band valleys ($\eta = 4$ for $Pb_{1-x}Sn_xTe$) and $|M_{if}|^2$ is the transition matrix element. The factor $\eta(\eta - 1)$ arises from the number of combinations of two dissimilar pairs of valleys out of η pairs. $P(1, 1', 2, 2')$ is the product of the occupancy probabilities of various states, as in Eq. (3.3.13). Under the nondegenerate approximation, it can be approximated by Eq. (3.3.17);

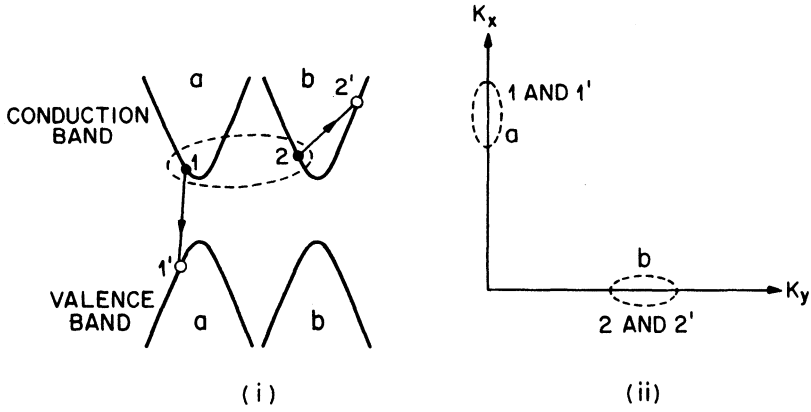


Fig. 13.7 Schematic illustration of (i) the Auger process involving different conduction-band valleys and (iii) possible k_x and k_y values (circled regions) for particles (1, 1') and (2, 2').

i.e.,

$$P(1, 1', 2, 2') = \frac{n^2 p}{N_c^2 N_v} \exp\left(\frac{-E_2' + E_g}{k_B T}\right) \tag{13.5.6}$$

where we have used the conservation of energy ($E_1 + E_2 = E_2' - E_g - E_1'$); where $N_c = N_v$ and is given by Eq. (13.3.5); and where n and p are the total number of electrons and holes, respectively. As discussed in Sec. 3.3, there is a minimum value E_T for which the Auger transition takes place. Since $P(1, 1', 2, 2')$ decreases rapidly for values of E_2' higher than this minimum value, most of the contribution to the integral in Eq. (13.5.5) arises from $E_2' \cong E_T$.

The minimum value of E_2' occurs when all the momentum vectors are aligned in the same direction (for example, along the x axis) in such a way that $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{k}_1'$ are directed in the positive direction and $\mathbf{k}_2'$ is directed in the negative direction, satisfying the following momentum conservation law:

$$k_{2x}' = k_{1x} + k_{2x} + k_{1x}'. \tag{13.5.7}$$

Since states 1 and 1' are in band 1 and states 2 and 2' are in band b, the energy conservation requires that

$$\frac{\hbar^2 k_{1x}^2}{2m_l} + \frac{\hbar^2 k_{2x}^2}{2m_t} = \frac{\hbar^2}{2m_t(E_2')} k_{2x}'^2 - \frac{\hbar^2}{2m_l} k_{1x}'^2 - E_g \tag{13.5.8}$$

where $m_t(E_2')$ is the effective mass at E_2' and is introduced to take into

account nonparabolicity of the bands. Dropping the suffix x and writing $k_1 = ak_1'$, $k_2 = bk_1'$, $r = m_t/m_l$, and $\mu_c = m_t(E_2')/m_t$, Eqs. (13.5.7) and (13.5.8) may be written as

$$k_2' = (a + b + 1)k_1' \tag{13.5.9}$$

$$rk_1^2 + k_2^2 = \frac{1}{\mu_c} k_2'^2 - rk_1'^2 - k_g^2 \tag{13.5.10}$$

with $k_g^2 = 2m_t E_g/\hbar^2$. Eliminating k_1', we obtain

$$k_2'^2 = k_g^2 \frac{(a + b + 1)^2}{(a + b + 1)^2 \mu_c^{-1} - (ra^2 + b^2 + r)}. \tag{13.5.11}$$

The magnitude k_2' can now be minimized with respect to a and b. The values of a and b for minimum k_2' are

$$a = 1 \qquad b = r \tag{13.5.12}$$

and the minimum value is given by

$$k_2'^2 = k_g^2 \frac{(2 + r)}{(2 + r)\mu_c^{-1} - r}. \tag{13.5.13}$$

Hence the minimum value of E_T or E_2', which is also the energy of electron $2'$ for the most probable transition, is given by

$$E_2' = E_T = \frac{1}{\mu_c} \frac{\hbar^2 k_2'^2}{2m_t} = \frac{(2 + r)E_g}{(2 + r) - r\mu_c}. \tag{13.5.14}$$

We can define the energy difference ΔE as

$$\Delta E = E_2' - E_g = \frac{r\mu_c}{(2 + r) - r\mu_c} E_g. \tag{13.5.15}$$

From Eqs. (13.5.5) and (13.5.6), it follows that the Auger rate varies approximately as

$$R \propto n^2 p \exp\left(\frac{-\Delta E}{k_B T}\right) \tag{13.5.16}$$

with ΔE given by (13.5.15). For parabolic bands, $\mu_c = 1$, and Eq. (13.5.16) reduces to Eq. (13.5.3) such that $\Delta E = rE_g/2$. For spherical parabolic bands, $\mu_c = 1$, $r = 1$, and $\Delta E = E_g/2$, which agrees with the result of Ref. 16. These results are valid when $r \ll 1$, which is the case for $Pb_{1-x}Sn_xTe$ ($r \cong 0.1$).

Equation (13.5.16) shows that, under the nondegenerate approximation, the nonradiative Auger recombination rate increases with increasing temperature and is larger for low band-gap semiconductors. Since ΔE appears in the exponent of Eq. (13.5.16), nonparabolicity of the bands [determined by μ_c in Eq. (13.5.15)] can significantly alter the calculated Auger rate. Ziep et al.[18] have calculated the Auger rate in PbSnTe using the Fermi statistics.

An exact calculation of the Auger rates in lead-salt semiconductors is not possible because the band structure away from the band edge is not known accurately. Moreover, only order-of-magnitude estimates of the matrix element are available. Rosman and Katzir[19] have estimated the CCCH Auger lifetime in PbSnTe using Kane's two-band model [Eq. (13.3.1)]. Figure 13.8 shows the calculated Auger lifetime at 77 K as a function of the injected carrier density for p-type $Pb_{0.82}Sn_{0.18}Te$. For an optical gain of $\sim 100\,cm^{-1}$, which is the typical threshold gain for a semiconductor laser, the injected carrier density is $\sim 0.8 \times 10^{17}\,cm^{-3}$ for low-doped material (see curve I in Fig. 13.6). At this injected carrier density, the current lost to Auger recombination (qn/τ_A) at 77 K is 25 A/(cm$^2 \cdot \mu$m) using $\tau_A = 5$ ns from Fig. 13.8. This is a significant fraction of the generally observed threshold current densities which are in the range of 500–1,000 A/(cm$^2 \cdot \mu$m) at 77 K. Since the Auger rate increases with increasing temperature, the current lost to Auger recombination also increases with increasing temperature; hence the Auger effect may be responsible for the observed high temperature dependence of the threshold current of PbSnTe laser diodes. Lead-salt lasers generally cannot operate near room temperature since almost all injected carriers combine nonradiatively and are not available for gain.

13.6 LASER DIODE FABRICATION

Single crystals of $Pb_{1-x}Sn_xTe$ and $Pb_{1-x}Sn_xSe$ have been prepared by conventional crystal growth techniques such as the Bridgman method[20] and the Czochralski method.[21] Vapor-growth techniques, with or without a seed crystal,[22–24] have also been used to grow single crystals of $Pb_{1-x}Sn_xTe$ and $Pb_{1-x}Sn_xSe$. Thallium and bismuth are generally used as p and n dopants, respectively. Single-heterostructure laser diodes have been fabricated from these single crystals using compositional interdiffusion (CID). In the CID technique, $PbS_{1-x}Se_x$ or $Pb_{1-x}Sn_xSe$ is annealed in the presence of PbS or

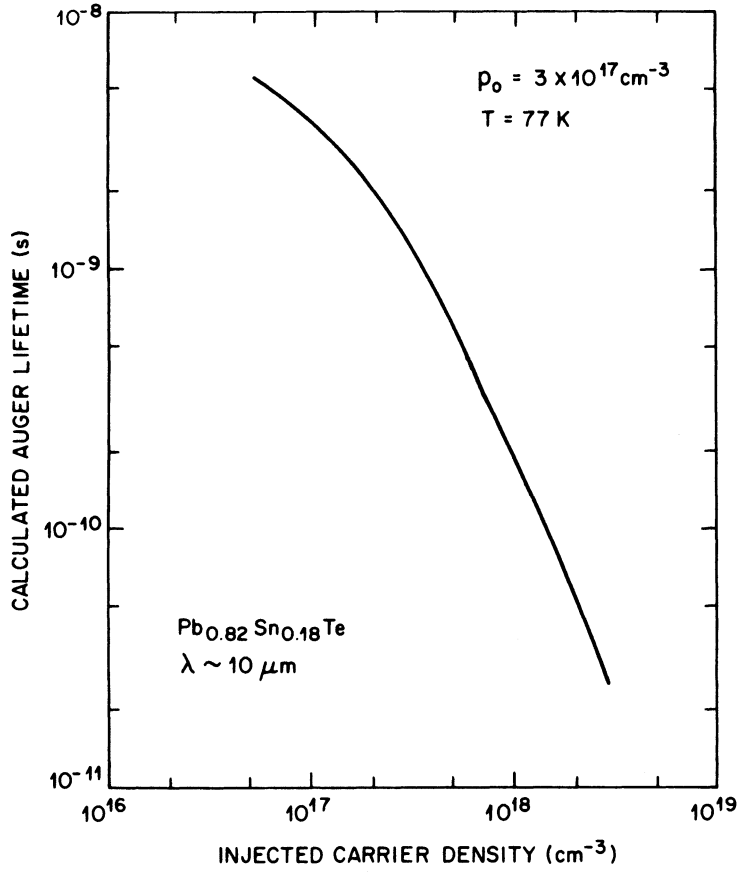


Fig. 13.8 Calculated Auger lifetime as a function of injected carrier density at 77 K for *p*-type PbSnTe ($\lambda \cong 10\ \mu m$) with a doping level of $3 \times 10^{17}\ cm^{-3}$. (After Ref. 19 © 1982 IEEE)

PbSe powder, respectively, resulting in the out-diffusion of Se or Sn from the surface. Single-heterostructure lasers showed excellent tuning characteristics.[25] Double-heterostructure lasers have been fabricated using liquid-phase epitaxy (LPE), vapor-phase epitaxy (VPE), and the molecular-beam epitaxial (MBE) techniques (see Chapter 4). $PbS_{1-x}Se_x$ double-heterostructure lasers have also been fabricated using CID.[26]

Several groups have reported on the growth of $Pb_{1-x}Sn_xTe$ layers by LPE on PbTe substrates.[27–30] The growth occurs from a supersaturated solution consisting of polycrystalline PbTe, Pb, and Sn. The growth temperature is generally between 550 and 600°C. A conventional LPE multibin boat with the substrate on a slider can be used (see Fig. 4.3). In the

$Pb_{0.88}Sn_{0.12}Te$-PbTe lasers fabricated by Groves et al.,[28] (100)-oriented *p*-PbTe, grown by the Bridgman method, was used as the substrate, and stripe-geometry lasers with ~50-μm-wide stripes were made using MgF_2 dielectric mask on the *p*-PbTe substrate. The threshold current density at 77 K was ~4.2 kA/cm^2 at $\lambda \cong 8.35$ μm. Tomasetta and Fonstad[29] also made double-heterostructure lasers of $Pb_{0.82}Sn_{0.18}Te$ with PbTe cladding layers using LPE on *p*-PbTe substrates. For an active-layer thickness of 6 μm, they report a threshold current density of ~1.2 kA/cm^2 at 77 K.

Vapor-phase epitaxy using H_2 transport in an open tube has been used to grow epitaxial layers of PbSnTe on PbTe substrates.[31,32] At a substrate temperature of ~700°C, the growth rate is typically ~0.3 μm/h. Fabrication of single-heterostructure lasers of $Pb_{0.83}Sn_{0.17}Te$-PbTe ($\lambda \cong 12.2$ μm at 77 K) by VPE has been reported.[33]

Walpole et al. have fabricated $Pb_{0.78}Sn_{0.22}Te$ double-heterostructure lasers using MBE.[34,35] Thallium-doped *p*-PbTe was used as the substrate. In MBE, lead-salt alloys in powdered form are commonly used as sources, and the composition of the grown layer is varied by controlling with shutters the rate at which the constituents arrive at a heated substrate. The walls of the growth chamber are cooler than the substrate, which is in direct line with the molecular beam originating from the sources. A variant of this technique is hot-wall vapor deposition (HWVD) which has been extensively used for the growth of $PbS_{1-x}Se_x$ double-heterostructure lasers on *n*-type PbS substrates.[36,37] In HWVD the walls of the growth chamber are hot, so they reflect or diffuse molecules; the substrate is heated independently. The furnace can be designed so that either (i) only the diffused beam from the walls gets deposited or (ii) both the direct and the diffused beams get deposited. PbS-$PbS_{0.6}Se_{0.4}$ ($\lambda \cong 5$ μm) double-heterostructure lasers fabricated using HWVD have threshold current densities of ~400 A/cm^2 at 77 K for an active-layer thickness of 1 μm. The control of the active-layer thickness obtained by MBE and HWVD is considerably better than that obtained by LPE. Further processing of the grown wafer to obtain the injection laser diodes is completed by depositing *p* and *n* contact metals and then cleaving the wafer to produce mirror facets (see Chapter 5). Indium is generally used as the *n* contact material, and Au and Pt are used for *p* contacts.

13.7 LASER PROPERTIES

For a nominal emission wavelength of ~10 μm, the refractive index of a lead-salt compound is about 5, and hence the wavelength in the active layer is ~2 μm. Thus an active-layer thickness in the range of 1–2 μm is desirable for good mode confinement without excessive increase in threshold. From Eqs. (2.6.2) and (2.6.3), the threshold current density J_{th} varies linearly with

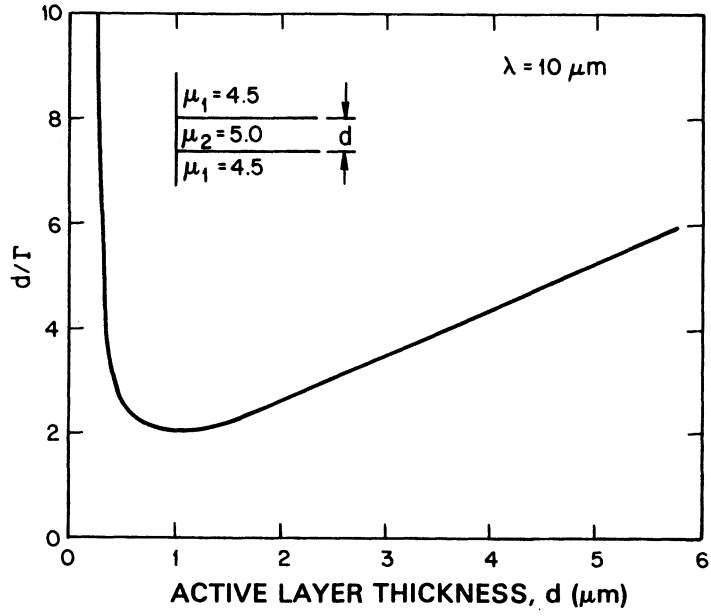


Fig. 13.9 Ratio of the active-layer thickness d to the confinement factor Γ plotted as a function of d for $\lambda = 10\ \mu m$. Inset shows the refractive indices of the heterostructure.

d/Γ, where d is the active-layer thickness and Γ is the confinement factor of the TE mode. Figure 13.9 shows a plot of d/Γ as a function of the active-layer thickness d for $\lambda = 10\ \mu m$. The active-layer and cladding-layer indices are assumed to be 5.0 and 4.5, respectively. Similar to the case of InGaAsP lasers (see Fig. 5.2). J_{th} is minimum for an optimum thickness d. However, note that the active-layer thickness needed is considerably larger for 10-μm lead-salt lasers than for InGaAsP lasers.

The technology for growing the lead-salt system is developing rapidly. As a result, a comparison of different material systems for laser performance is premature at this point. However, the following two general conclusions may be drawn. The longer the emission wavelength, (i) the lower the maximum operating temperature, and (ii) the lower the external differential quantum efficiency and maximum output power. The former may be due to the Auger effect, which is larger at longer wavelengths, and the latter may be due to increased free-carrier absorption ($\alpha_{fc} \propto \lambda^2$) at longer wavelengths. Figure 13.10 shows the measured light-current characteristics of a PbSnTe diode laser at 2 K.[38] The maximum power obtained without any facet coating (curve L1) is 5 mW. This power output can be increased to over 10 mW (curve

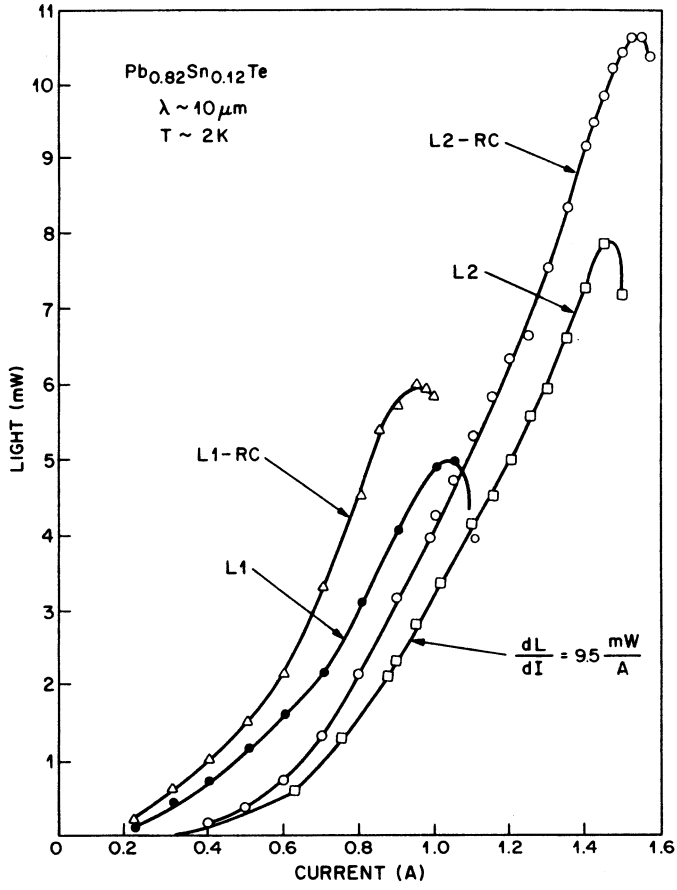


Fig. 13.10 Light-current characteristics of a PbSnTe diode laser with and without antireflection (AR) coating: L1, no coating; L2, one facet AR-coated; L1–RC, one facet totally reflecting; and L2–RC, one faceted AR-coated and the other totally reflecting. The slope efficiency is also indicated. (After Ref. 38)

L2–RC) by using an antireflection (AR) coating on the front facet and a total-reflection coating on the back facet.

The nonradiative recombination mechanisms discussed in Sec. 3.3 determine both the threshold current and its temperature dependence. In contrast to InGaAsP lasers, surface recombination plays an important role in lead-salt lasers because of their relatively poor heterojunction-interface quality. A temperature-independent surface-recombination velocity S increases the

threshold current by

$$\Delta J_{th} = qn_{th}SA \tag{13.7.1}$$

where q is the electron charge, n_{th} is the carrier density at threshold, and A is the surface area. For thick active layers, the relative contribution of interface recombination to the total current is smaller than that for thin layers. Figure 13.11 shows the measured threshold current density as a function of temperature for PbSnTe-PbTe double-heterostructure lasers of different active-layer thicknesses.[39] The low temperature dependence observed for the thin-active-layer device is probably due to interface recombination, which dominates the measured threshold current at low temperatures.[39–42] An interface recombination velocity of $\sim 1 \times 10^5$ cm/s has been measured for PbSnTe-PbTe interfaces at 5 K.[42] Such a high interface recombination velocity can significantly reduce the internal quantum efficiency. The low internal quantum efficiency observed in many lead-salt diode lasers is

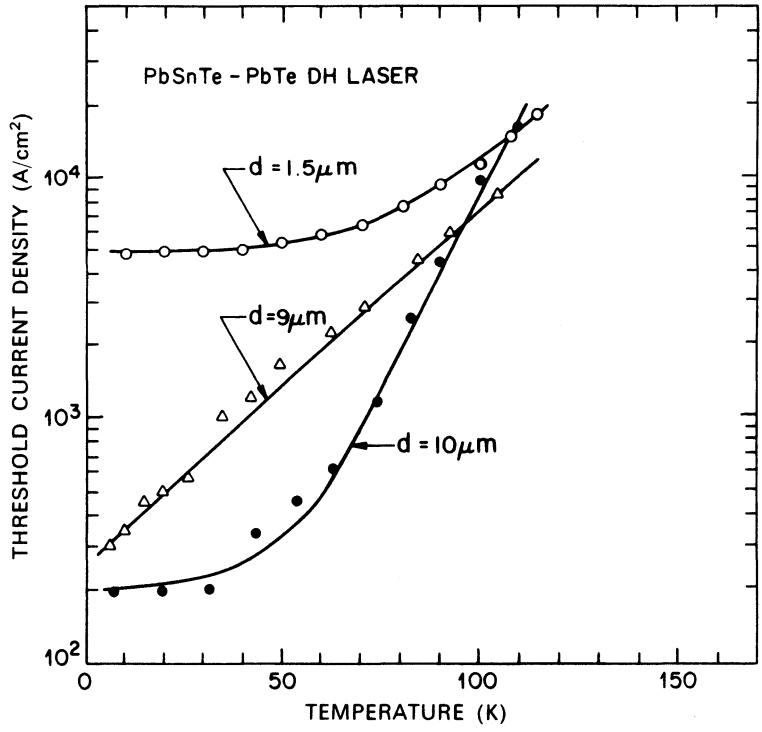


Fig. 13.11 Measured threshold current density as a function of temperature for PbSnTe-PbTe double-heterostructure lasers of different active-layer thicknesses. (After Ref. 39)

probably related to poor interfaces. Since the quality of the interface depends on the material growth and processing techniques used, it is not surprising that differing temperature dependences of the threshold current are observed for different types of diodes. Figure 13.12 shows the measured data for $PbS_{1-x}Se_x$ diodes of different types.

Quaternary PbSnSeTe diode lasers with PbSeTe cladding layers have been

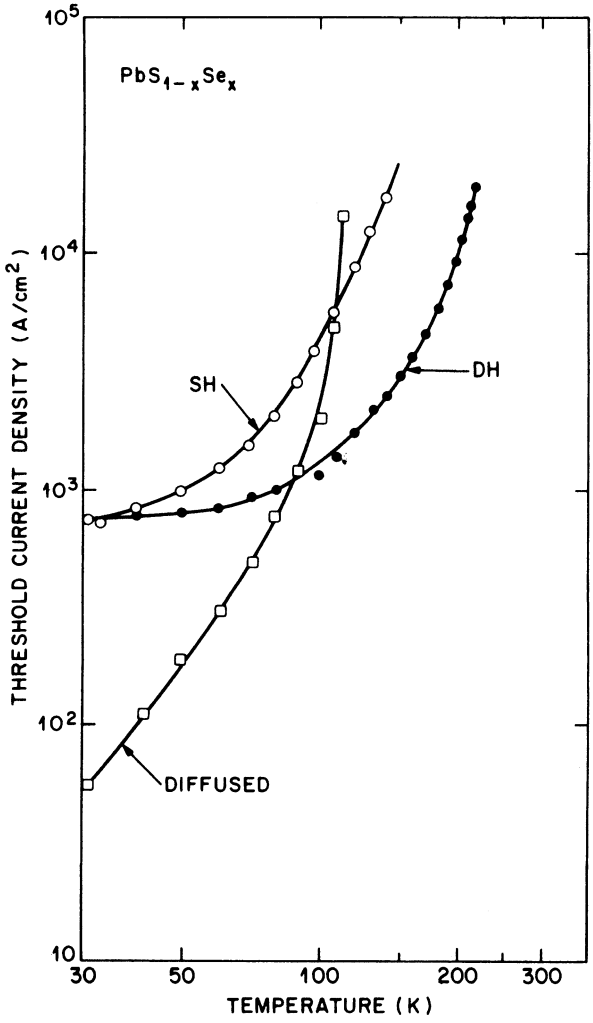


Fig. 13.12 Temperature dependence of threshold current density for different types of $PbS_{1-x}Se_x$ laser diodes. SH, single heterostructure; DH, double heterostructure. (After Ref. 5)

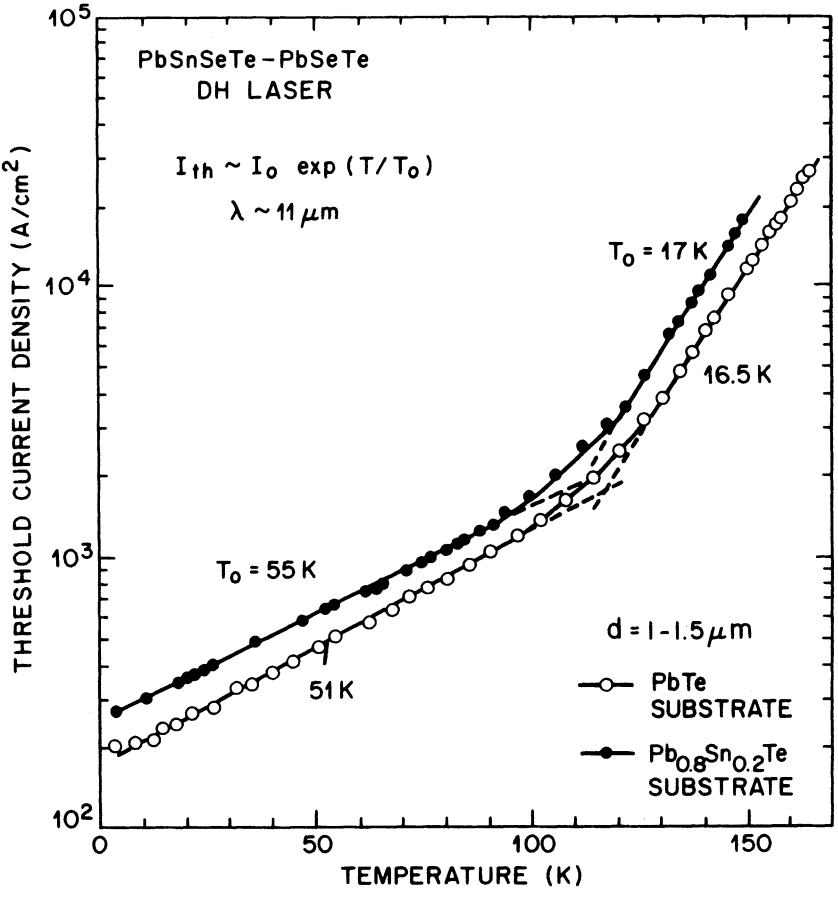


Fig. 13.13 Measured temperature dependence of threshold current of PbSnSeTe diode lasers with PbSeTe cladding layers. Closed and open circles represent active-layer thicknesses of 1.5 μm and 1.2 μm, respectively. (After Ref. 39)

fabricated by LPE.[39] Both PbTe and $Pb_{0.8}Sn_{0.2}Te$ were used as substrates. The PbSeTe layers are used to provide good lattice matching (which should have reduced interface defects) between the substrate and the active region. Figure 13.13 shows the measured temperature dependence of the threshold current for lasers emitting near 10 μm. The active-region thickness of the lasers in Fig. 13.13 falls in the range of 1.0–1.5 μm. The saturation of laser threshold current with a decrease in temperature, similar to that shown in Fig. 13.11 for lasers with a thin active region, is not observed; this suggests that these devices have a reduced interface recombination velocity. The

threshold current as a function of temperature may be written as

$$I_{th} \cong I_0 \exp(T/T_0). \tag{13.7.2}$$

The measured T_0 values in various temperature ranges are also shown in Fig. 13.13.

The output power of the lead-salt family of diode lasers depends on the crystalline quality of the substrate. The maximum output power is observed to be inversely proportional to the dislocation density.[43] Diffused-junction diodes generally emit more power than single- or double-heterostructure diodes.[5] Also, the output powers are higher under pulsed operation than under CW operation.

The emission characteristics of lead-salt lasers can be improved by adopting the fabrication techniques used in the GaAs and InGaAsP material systems. For example, buried-heterostructure PbSnTe lasers have been developed.[44] The PbSnTe active layer with a width of 4 μm and thickness of 1 μm was buried in the PbTeSe confining layers using a two-step LPE procedure (see Chapter 5). These lasers have a threshold current of ~60 mA at 50 K and can provide up to 1.7 mW of single-mode power at a wavelength of ~8 μm. Furthermore, the operating temperature can be increased up to 105 K with 300 μW of output power. Cleaved-coupled-cavity lead-salt lasers have also been developed.[45] Similar to the case of InGaAsP lasers discussed in Chapter 8, these PbSnTe lasers provide an improvement in spectral purity together with a reduction the threshold current. Furthermore, the laser diode can be tuned by adjusting the current in the controller section.

Another development[46] in lead-salt lasers is adoption of the quantum-well structure (see Chapter 9). The motivation was to see if the operating temperature of lead-salt lasers could be increased through quantum-size effects. The stripe-geometry quantum-well lasers used a new quaternary compound, $Pb_{1-x}Eu_xSe_yTe_{1-x}$ (which has previously been used[47] to fabricate regular double-heterostructure lasers). The single-quantum-well (SQW) structure used a PbTe quantum well as the active region and PbEuSeTe for the cladding layers. The quantum well ranged in thickness from 30 to 250 nm. The output power was highest for 30-nm quantum wells, ranging up to 5 mW for multimode emission. The threshold current was found to be significantly reduced because of the quantum-size effects. These SQW lasers could be temperature tuned from $\lambda = 6.45$ μm (at 13 K) to 4.01 μm (at 241 K). (The process of temperature tuning is described in more detail in the next section.) Under CW conditions, the highest operating temperature was 174 K (at 4.41 μm) while pulsed operation could be achieved up to 241 K. These properties suggest that quantum-well lead-salt lasers may prove useful for spectroscopic and fiber optic applications.

13.8 TUNING CHARACTERISTICS

One of the first applications of lead-salt diode lasers was in spectroscopy because of the large tuning range that can be obtained simply by varying the operating temperature.[4] In addition to temperature tuning, pressure tuning and magnetic-field tuning can also be used. In this section we briefly discuss the tuning characteristics of lead-salt lasers.

The band gap of a semiconductor varies with temperature. Figure 13.14 shows the data for three lead-salt binaries. The band gap increases at a rate of ~ 0.5 meV/K, which corresponds to a wavelength change of 10–20 nm/K depending on the wavelength. Clearly, by varying the temperature the emission wavelength can be varied over a large range. Large-range tuning of this nature usually occurs in a discrete fashion; i.e., the laser-emission wavelength jumps from one longitudinal mode to another as the temperature varies. Continuous tuning over a small range can also be obtained because

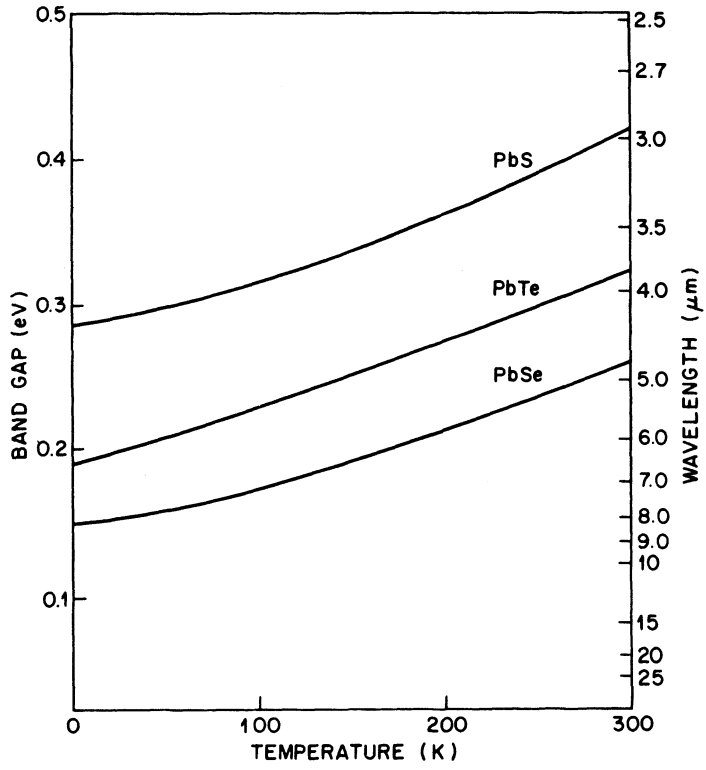


Fig. 13.14 Variation in the band gap of PbS, PbTe, and PbSe as a function of temperature. (After Ref. 4)

the position of the cavity mode itself depends on the temperature through a temperature-dependent refractive index. The wavelengths of the longitudinal modes λ_m are determined by (see Sec. 2.3)

$$m\lambda_m = 2\mu L \tag{13.8.1}$$

where μ is the refractive index, L is the length of the cavity, and m is an integer. With a change in temperature ΔT, μ changes by

$$\Delta\mu = (d\mu/dT)\Delta T. \tag{13.8.2}$$

Hence the mode position changes by

$$\Delta\lambda_m = \frac{(d\mu/dT)}{\mu}\lambda_m \Delta T. \tag{13.8.3}$$

The continuous-tuning rate for typical lead-salt lasers is about one-third the discrete tuning rate with temperature.

The active-region temperature of the diode laser that enters in Eqs. (13.8.2) and (13.8.3) depends on both the ambient temperature and the operating current (through ohmic heating). Thus a fine tuning (which may be sufficient for some spectroscopic applications) can be achieved simply by varying the operating current. Figure 13.15 shows the measured data[48] for $Pb_xSn_{1-x}Te$ diode lasers. Continuous tuning is observed over a small current range followed by successive mode jumps. The measured change in emission wavelength with temperature of various lead-salt diode lasers is shown in Fig. 13.16. The large tuning range is primarily due to an increase in the band gap with increasing temperature.

Large changes in the band gap and hence the emission wavelength of these narrow-gap semiconductors can also be obtained by applying hydrostatic pressure. Figure 13.17 shows the band gap at 77 K as a function of hydrostatic pressure for three binary lead-salt semiconductors. Pressure tuning has been used for spectroscopic applications.[49] Generally a combination of temperature tuning (using current) and pressure tuning can be used. Band-gap tuning also occurs when a uniaxial pressure is applied. Although uniaxial pressure is considerably simpler to apply than hydrostatic pressure, there is a chance that the laser may get crushed.

Magnetic-field tuning can also be used as a laboratory spectroscopic tool. Tuning results from quantization of the valence and conduction bands into the Landau levels, the separation of which varies linearly with the magnetic field. The measured tuning rate[50] of $PbS_{0.82}Se_{0.18}$ diode lasers is in the range

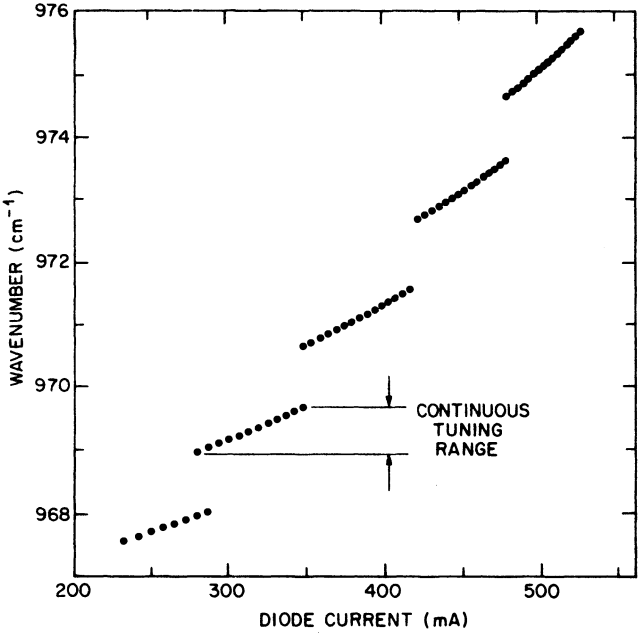


Fig. 13.15 Wavelength tuning of the modes of $Pb_{1-x}Sn_xTe$ diode lasers as a function of injection current. The discrete wavelength jumps correspond to longitudinal-mode hopping. (After Ref. 4)

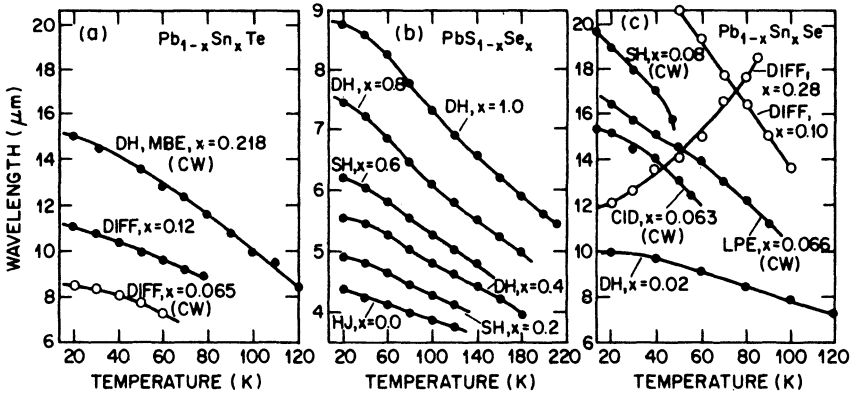


Fig. 13.16 Change in emission wavelength with temperature for different lead-salt diode lasers. Different curves in each case correspond to different structures and compositions: SH, single heterostructure; DH, double heterostructure; MBE, molecular-beam epitaxy; LPE, liquid-phase epitaxy; CID, compositional interdiffusion; HJ, homojunction; and DIFF, diffused junction. (After Ref. 5)

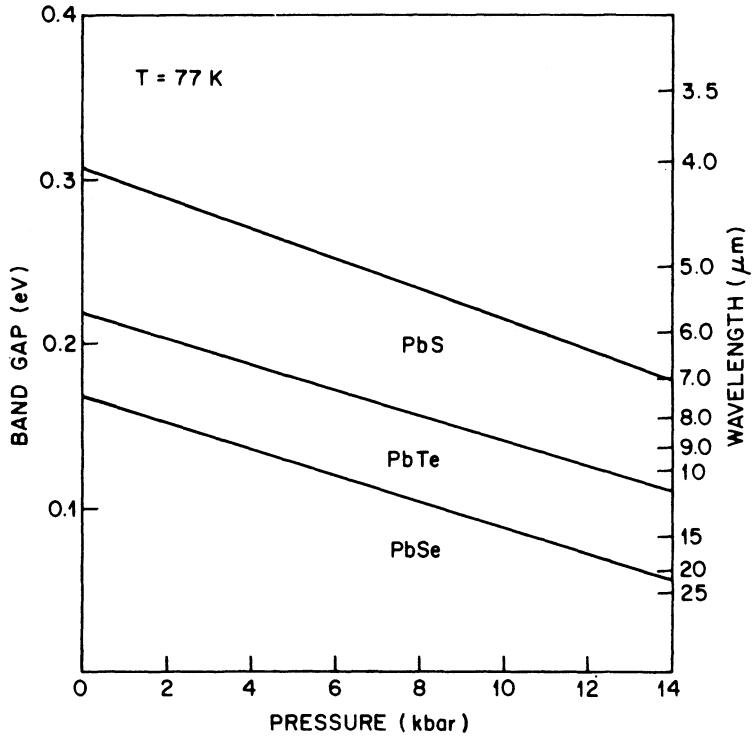


Fig. 13.17 Variation of the band gaps of PbS, PbTe, and PbSe as a function of hydrostatic pressure at 77 K. (After Ref. 4)

of 0.4–2 MHz/G, depending on whether the tuning occurs within one cavity mode or between different cavity modes.

13.9 OTHER MATERIAL SYSTEMS

So far in this chapter we have focused our attention on lead-salt lasers. In this section we briefly mention some of the results on semiconductor lasers obtained using material systems other than the lead salts.

13.9.1 Infrared Semiconductor Lasers

The wavelength of semiconductor lasers has been extended to about 100 μm using single crystals of $Bi_{1-x}Sb_x$ solid solutions.[51] By varying the composition x in the range of 0.08–0.16, the band gap of this material can be varied from 2.7 to 22.6 meV at $T = 4.2$ K. For the composition used in laser

fabrication, the band gap $E_g = 12.3$ meV corresponded to a wavelength of about 100 μm. The active region was 3.7-mm long, 2.7-mm wide, and 0.17-mm thick. When operated at 4.2 K, the threshold for the laser was 1.38 A.

Attempts have been made to increase the wavelength at which room-temperature operation of semiconductor lasers can be realized. These efforts are motivated by the theoretical possibility of an ultra-low-loss fiber at wavelengths exceeding 2 μm.[52] InGaAsP lasers can provide continuous room-temperature operation up to ~1.6 μm. This wavelength can be extended to about 1.8 μm using AlGaAsSb-GaSb double-heterostructure lasers.[53,54] Further increase in the wavelength has been realized[55–63] using the material system InGaAsSb-AlGaAsSb. Using LPE, double-heterostructure lasers were prepared with the quaternary layers lattice-matched to GaSb substrates. Under pulsed operation of these 2.2-μm devices, the threshold current density at room temperature[56] was ~7 kA/cm^2. By 1988 this value had been reduced to 1.5 kA/cm^2 in AlGaAsSb lasers operating continuously at room temperature.[60] The threshold current for a 2.34-μm laser was as low as 80 mA through the use of a ridge-waveguide structure. By 1991 an InGaAsSb semiconductor laser operating at 2.3 μm emitted 900 mW of output power at room temperature under quasi-CW operation; it was necessary to pump it with 200-ns current pulses to avoid heating of the active region.[62] The threshold current density was as low as 1.5 kA/cm^2 with a differential quantum efficiency as high as 50%. The threshold current increased exponentially with temperature with a characteristic temperature of about 50 K. InGaAsSb semiconductor lasers operating at the 3.06-μm wavelength have also been fabricated.[63]

13.9.2 Visible Semiconductor Lasers

AlGaAs semiconductor lasers can be designed to operate at wavelengths as short as 780 nm by using a quantum-well structure. However, many applications (for example, optical data recording and laser printing), require semiconductor lasers operating in the visible region extending from 400–700 nm. Considerable effort was spent in the late 1980s to develop visible semiconductor lasers[64–96] with the initial emphasis on red semiconductor lasers operating in the 600–700-nm range, since such lasers could be used in many applications in place of He-Ne gas lasers. Although the material system GaInAsP-AlGaAs is sometimes used to make red semiconductor lasers, the material of choice for fabrication of such lasers is the quaternary compound InGaAlP formed by replacing a fraction of Ga atoms by Al in the ternary compound $In_{0.5}Ga_{0.5}P$. Its composition is written as $In_{0.5}(Ga_{1-x}Al_x)_{0.5}P$, where x represents the fraction of Al atoms. The bandgap E_g and the

refractive index μ of the quaternary compound InGaAlP vary with the fraction x as

$$E_g = 1.90 + 0.60x \tag{13.9.1}$$

$$\mu = 3.65 - 0.37x \tag{13.9.2}$$

where x is limited to be in the range 0–0.7 since InGaAlP has an indirect band gap for $x > 0.7$. Both active and cladding layers are made of the same compound by changing the fraction x and are grown epitaxially over the GaAs substrate. The fraction x is generally chosen to be 0.7 for the cladding layer in order to maximize the band-gap discontinuity at the heterostructure interface. The active-layer composition depends on the laser wavelength and is typically in the range $x = 0$–0.5.

InGaAlP semiconductor lasers operating continuously at room temperature in the wavelength range 670–680 nm were first demonstrated in 1985 by using InGaP active layers ($x = 0$).[64–66] Since then, the performance of such lasers has improved considerably.[67–76] Output powers as high as 320 mW were obtained in 1990 in a broad-area device.[71] For practical applications it is necessary to obtain the laser output in a single stable spatial mode. Power levels of up to 50 mW were obtained by using an index-guided structure with a thin layer and optimum facet reflectivities.[72] Power levels of up to 80 mW were later obtained by using a window structure.[73] Reliability of InGaAlP semiconductor lasers has also improved considerably.[70] Indeed, 670-nm InGaAlP semiconductor lasers emitting 20 mW of output power in a stable transverse mode were available commercially by 1991.

InGaAlP semiconductor lasers emitting in the wavelength region near 630 nm have been developed for applications where He-Ne lasers are commonly used. It is possible to reduce the operating wavelength of InGaAlP lasers by increasing the Al fraction in the active-layer composition [see Eq. (13.9.1)], and considerable effort was directed[74–76] toward the realization of 630-nm semiconductor lasers during the early 1990s. In one scheme, visible lasers operating at 636 nm were demonstrated by using an InGaAlP active layer (with $x = 0.15$) and a heterobarrier-blocking (HBB) design in which a large band-gap discontinuity is introduced by inserting an InGaP layer between the p-type InGaAlP cladding layer and the p-type GaAs contact layer.[74] Figure 13.18 shows the L-I characteristics of one such laser at several temperatures. The two insets show the optical spectrum and the far-field patterns when the laser operates at 2 mW.[75] The threshold current of such lasers is relatively large ($\sim$100 mA) and they are unable to operate at temperatures above 70–80°C.

In general, the operating characteristics of InGaAlP lasers degrade when the Al composition exceeds 20% ($x > 0.2$) mainly because the band-gap

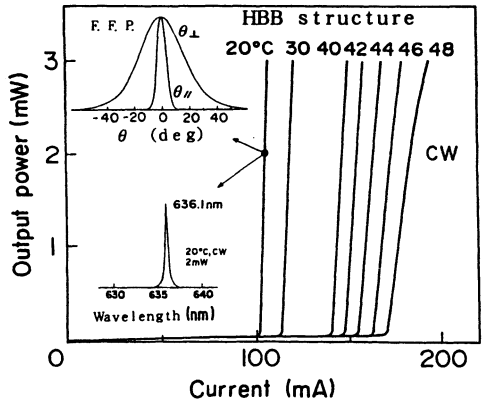


Fig. 13.18 L-I characteristics of a 636-nm visible laser at several temperatures. The two insets show the optical spectrum and the far-field patterns when the laser operates at 2 mW. (After Ref. 75 © 1991 IEEE)

difference between the active and cladding layers becomes so small that carriers can leak over the heterobarrier (see Sec. 3.5.1). Two techniques have been used to overcome this problem. First, the wavelength can be reduced by using the quantum-well structure since the photon energy is higher than the band-gap energy through the quantum confinement effect (see Chapter 9). Second, it was found that the band gap of an InGaAlP active layer can be increased if the layer is epitaxially grown on a misoriented substrate.[76] The band gap is 50–60 meV wider for InGaAlP active layers grown with a misorientation of the (100)-GaAs substrate by 5–7° toward the [011] direction. This increase in the band gap translates into a 20–30 nm reduction in the wavelength and allows operation near 655 nm even with ternary InGaP active layers ($x = 0$). Figure 13.19 shows the measured wavelengths as a function of the Al fraction x for substrates misoriented by 5 and 7° and compares them with conventional lasers fabricated without any misorientation. The nearly linear decrease of wavelength with the Al fraction x implies that the linear relation similar to Eq. (13.9.1) applies even for lasers grown over misoriented substrates. Figure 13.19 shows the L-I characteristics at several temperatures for a 655-nm laser fabricated with the InGaP active layer ($x = 0$) deposited over a 7°-misoriented substrate.[76] The maximum operating temperature of 85°C is comparable with the conventional 670-nm semiconductor lasers. Visible lasers operating near 630 nm can be fabricated by this growth technique by using InGaAlP active layers with only 15% Al content ($x = 0.15$). The threshold current of one such laser operating continuously at room temperature was 90 mA and the laser could be operated up to 6 mW.[76]

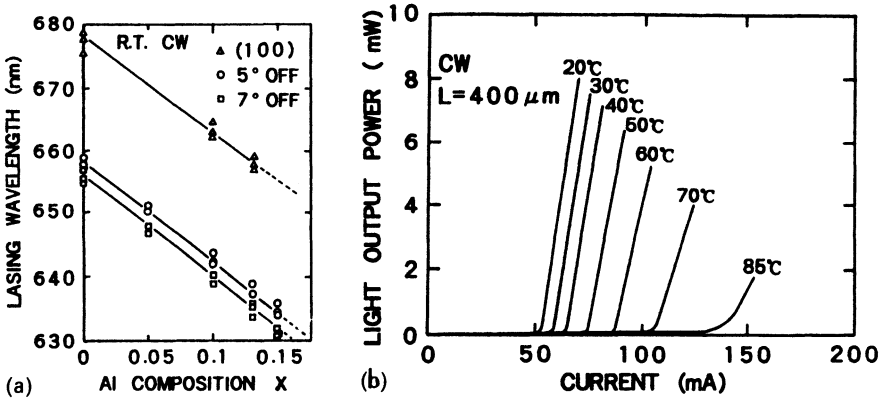


Fig. 13.19 (a) Measured wavelength as a function of the Al fraction x for several visible lasers grown over substrates misoriented by 0, 5, and 7°. (After Ref. 76 © 1991 IEEE)

The performance of InGaAlP lasers can be considerably improved by using the multiquantum-well (MQW) design for the active region.[80–90] The performance is improved even more by using strained MQW lasers. Output powers in excess of 100 mW have been obtained for 670-nm strained MQW lasers.[80] At the same time, such lasers are capable of operating at temperatures as high as 150°C.[81] The performance of 630-nm is also improved by the adoption of the MQW design. The threshold current of one 636-nm laser was 57 mA.[86] It was capable of emitting up to 15 mW of power although single-mode operation was limited to 8 mW. The use of strained MQW InGaAlP lasers has provided output powers of up to 50 mW at 632 nm by using an active region with 4 quantum wells of composition $In_{0.53}Ga_{0.47}P$ ($x = 0$). The laser was capable of providing 30 mW of output power at 50°C.[87] The maximum operating temperature for this laser was more than 100°C since the laser emitted 40 mW even at 80°C. These lasers also appear to have excellent reliability. The aging characteristics of 12 lasers over 5,500 hours of 3-mW operation at 50°C indicated an estimated mean time to failure (see Chapter 14) of more than 10,000 hours.[88] The output power can be increased by using an array design. A 1-cm bar of laser arrays operating at 633 nm provided up to 3 W of output power under CW operation at room temperature.[89]

It is possible to reduce the wavelength further by increasing the misorientation angle of the substrate and/or the Al content of the active layer. The wavelength was reduced to 607 nm in one laser in which a 10-nm-thick InGaP active layer ($x = 0$) was grown over a 15°-misoriented substrate to provide a single strained quantum well.[84] This laser also incorporated 300-μm-long Bragg mirrors on each side of the 300-μm-long active layer,

and therefore acted as a DBR laser (see Chapter 7). It maintained a single-longitudinal mode up to 1.8 times above threshold when operated at 140 K. Even shorter wavelengths can be obtained by such techniques. Yellow and green semiconductor lasers operating at 576 and 555 nm have been demonstrated, but both of them required low-temperature operation.[90,91] In general, room-temperature operation of InGaAlP laser appears to be difficult when the wavelength is reduced considerably below 630 nm.

A totally different approach has been followed to realize visible semiconductor lasers in the blue-green region near 500 nm. Semiconductors such as CdS, CdSe, ZnS, ZnSe belong to the II–VI class and have wide band gaps in excess of 2 eV. Semiconductor lasers made by using these semiconductor materials are expected to emit blue light in the optical region. However, fabrication of such lasers has been hindered by the presence of strain-induced defects during the epitaxial growth process and by the difficulty of making a low-resistance ohmic contact on the epitaxially grown structure. A sustained development effort during the early 1990s has partially solved these problems.[92–96] By 1992, ZnSe-based MQW semiconductor lasers were capable of emitting $\sim$0.1 mW of output power in the blue region near 490 nm provided they were cooled to 77 K.[94] The emission wavelength increased with an increase in the operating temperature, and the laser emitted in the green region near 510 nm for temperatures in excess of 200 K. The threshold current density increased and the differential quantum efficiency declined as the temperature was increased above 77 K. The threshold current density of one laser increased from 0.4 kA/cm^2 at 77 K to 1.5 kA/cm^2 at room temperature. ZnSe lasers generally require the use of the MBE growth technique. The laser structure is grown over a GaAs substrate. The active region consists of MQWs of the ternary compound $Zn_{1-x}Cd_xSe$ (with $x \sim 0.2$). The cladding layers are made of ZnSe or the ternary compound $ZnSe_{1-x}S_x$; the addition of sulfur decreases the laser wavelength by increasing the hole confinement energy.[94] A DFB ZnCdSe laser has also been demonstrated by pumping it optically at room temperature.[95] Further development should make room-temperature operation of electrically pumped ZnSe lasers possible.

PROBLEMS

13.1 The band gap of $Pb_{1-x}Sn_xTe$ at 77 K is found to vary with the mole fraction x as E_g (in eV) $= 0.22 - 0.55x$. At what wavelength would such a lead-salt laser operate when $x = 0.2$?

13.2 Why is it necessary to operate lead-salt lasers at low temperatures? Identify the most important physical process that hinders their operation at room temperature.

13.3 The refractive indices of the active and passive layers of a lead-salt semiconductor laser are 5 and 4.5 respectively. Find the active-layer thickness for which the threshold current would be minimum when the laser is operating at the 12-μm wavelength.

13.4 The band gap of a lead-salt semiconductor laser increases with temperature at a rate of 0.5 meV/K. How much should the temperature be changed to increase the wavelength by 50 nm?

13.5 Determine the active-layer composition for lead-salt lasers designed to operate at 5-, 10-, and 20-μm wavelengths. Suggest a possible composition of the cladding layer in each case.

13.6 Find the active-layer composition for a visible InGaAlP laser designed to operate at 650 nm when the laser structure is grown without substrate misorientation. What should be the cladding layer composition to ensure maximum band-gap discontinuity at the active-cladding interface? What is the index step at the active-cladding interface under these conditions?

13.7 Use the data of Fig. 13.19 to find the active-layer composition for a visible InGaAlP laser designed to operate at 630 nm when the laser structure is grown with 7° substrate misorientation.

REFERENCES

1. Horikoshi, Y., Chap. 3 in *Semiconductors and Semimetals*, Vol. 22, Part C, ed. W. T. Tsang. New York: Academic Press, 1985.
2. Harman, T. C., in *Physics of Semimetals and Narrow Gap Semiconductors*, ed. D. L. Carter and R. T. Bate. New York: Pergamon Press, 1971; Calawa, A. R. *J. Lumin.* **7,** 477 (1973).
3. Hesse, J., and H. Preier, p. 229 in *Festkörperprobleme* [Advances in Solid State Physics], Vol. XV, ed. H. J. Queisser. Braunschweig: Pergamon/Vieweg, 1975.
4. Melngailis, I., and A. Mooradian, in *Laser Applications to Optics and Spectroscopy*, ed. S. F. Jacobs, M. Sargent III, J. F. Scott, and M. O. Scully. Reading, Mass: Addison-Wesley Publishing Co., 1975.
5. Preier, H. *Appl. Phys.* **20,** 189 (1979).
6. Butler, J. F., and T. C. Harman. *IEEE J. Quantum Electron.* **QE-5,** 50 (1969).
7. Harman, T. C., and I. Melngailis, p. 1 in *Applied Solid State Science*, Vol. 4, ed. R. Wolfe. New York: Academic Press, 1974.
8. Dimmock, J. O., I. Melngailis, and A. J. Strauss. *Phys. Rev. Lett.* **16,** 1193 (1966).
9. Harman, T. C., A. R. Calawa, I. Melngailis, and J. O. Dimmock. *Appl. Phys. Lett.* **14,** 333 (1969).
10. Zemel, J. N., J. D. Jensen, and R. B. Schollar: *Phys. Rev. A* **140,** 330 (1965).
11. Dalven, R. *Infrared Phys.* **9,** 141 (1969).
12. Kane, E. O., pp. 13–31 in *Narrow Gap Semiconductors: Physics and Applications*, ed. W. Zawadzki. Berlin: Springer-Verlag, 1980.
13. Anderson, W. W. *IEEE J. Quantum Electron.* **QE-13,** 532 (1977).
14. Rosman, R., and A. Katzir. *IEEE J. Quantum Electron.* **QE-19,** 73 (1983).
15. Appold, G., R. Grisar, G. Bauer, H. Burkhard, R. Ebert, M. Pascher, and H. G. Hafele, presented at 14th Int. Conf. Phys. Semicond., Edinburgh, 1978.

16. Beattie, A. R., and P. T. Landsberg. *Proc. R. Soc. London Ser. A* **249,** 16 (1959).
17. Emtage, P. R. *J. Appl. Phys.* **47,** 2565 (1976).
18. Ziep, O., D. Genzow, M. Mocker, and K. H. Herrmann. *Phys. Status Solidi B* **99,** 129 (1980).
19. Rosman, R., and A. Katzir. *IEEE J. Quantum Electron.* **QE-18,** 814 (1982).
20. Harman, T. C. *J. Nonmetals* **1,** 183 (1973).
21. Wagner, J. W., and R. K. Willardson. *Trans. Mettal. Soc. AIME* **242,** 366 (1968).
22. Lo, W. *J. Electron. Mater.* **6,** 39 (1977).
23. Maier, H., D. R. Daniel, and H. Preier. *J. Cryst. Growth* **35,** 121 (1976).
24. Parker, S. G. *J. Electron. Mater.* **5,** 497 (1976).
25. Linden, K. J., K. W. Nill, and J. F. Butler. *IEEE J. Quantum Electron.* **QE-13,** 720 (1977).
26. Bryant, F. J., and A. Qadeer. *IEE Proc. I* **131,** 113 (1984).
27. Shotov, A. P., O. I. Davarashvili. *Izv. Akad. Nauk SSSR Neorg. Mater.* **13,** 610 (1977).
28. Groves, S. H., K. W. Nill, and A. J. Strauss. *Appl. Phys. Lett.* **25,** 331 (1974).
29. Tomasetta, L. R., and C. G. Fonstad. *Appl. Phys. Lett.* **24,** 567 (1974).
30. Longo, J. T., J. Harris, E. Gertner, and J. Chu. *J. Crys. Growth* **15,** 107 (1972).
31. Rolls, W., R. Lee, and R. J. Eddington. *Solid-State Electron.* **13,** 75 (1970).
32. Bellavance, D. W., and M. R. Johnson. *J. Electron. Mater.* **5,** 363 (1976).
33. Kurbatov, L. N., et al. *Sov. J. Quantum Electron.* **5,** 1137 (1976).
34. Walpole, J. N., A. R. Calawa, T. C. Harman, and S. H. Groves. *Appl. Phys. Lett.* **28,** 552 (1976).
35. Walpole, J. N., A. R. Calawa, S. R. Chinn, S. H. Groves, and T. C. Harman. *Appl. Phys. Lett.* **29,** 307 (1976).
36. Preier, H., M. Bleicher, W. Riedel, and H. Maier. *Appl. Phys. Lett.* **28,** 669 (1976); *J. Appl. Phys.* **47,** 5476 (1976).
37. Preier, H., M. Bleicher, W. Riedel, H. Pfeiffer, and H. Maier. *Appl. Phys.* **12,** 277 (1977).
38. Walpole, J. N., A. R. Calawa, R. W. Ralston, and T. C. Harman. *J. Appl. Phys.* **44,** 2905 (1973).
39. Horikoshi, Y., M. Kawashima, and H. Saito. *Jpn. J. Appl. Phys.* **21,** 77 (1982).
40. Kasemset, D., and C. G. Fonstad. *IEEE J. Quantum Electron.* **QE-15,** 1266 (1979).
41. Tomasetta, L. R., and C. G. Fonstad. *IEEE J. Quantum Electron.* **QE-11,** 384 (1975).
42. Kasemset, D., and C. G. Fonstad. *Appl. Phys. Lett.* **34,** 432 (1979).
43. Lo, W., G. P. Montgomery, Jr., and D. E. Swets. *J. Appl. Phys.* **47,** 267 (1976).
44. Nishijima, Y., K. Ebe, H. Fukuda, K. Shinohara, and K. Murase, presented at Conf. Lasers and Electro-optics, Baltimore, May 1985.
45. Lo, W. *Appl. Phys. Lett.* **44,** 1118 (1984).
46. Partin, D. L., R. F. Majkowski, and D. E. Swets. *J. Vac. Sci. Technol. B* **3,** 576 (1985).
47. Partin, D. L. *Appl. Phys. Lett.* **43,** 996 (1983); **45,** 487 (1984).
48. Hinkley, E. D., K. W. Nill, and F. A. Blum, in *Laser Spectroscopy of Atoms and Molecules,* vol. 2, ed. H. Walther. Heidelberg: Springer-Verlag, 1975.
49. Pine, A. S., C. J. Glassbrenner, and J. A. Kafalas. *IEEE J. Quantum Electron.* **QE-9,** 800 (1973).
50. Nill, K. W., F. A. Blum, A. R. Calawa, and T. C. Harman. *J. Nonmetals* **1,** 211 (1973).
51. Aleksanyan, A. G., R. A. Kazaryan, and A. M. Khachatryan. *Sov. J. Quantum Electron.* **14,** 336 (1984).
52. Harrington, J. A. *Infrared Fiber Optics.* Bellingham, WA: SPIE, 1990.
53. Dolginov, L. M., A. E. Drakin, L. V. Druzhinina, P. G. Eliseev, M. G. Milvidsky, V. A. Skripkin, and B. N. Sverdlov. *IEEE J. Quantum Electron.* **QE-17,** 593 (1981).
54. Takeshima, M. *J. Appl. Phys.* **56,** 2502 (1984).
55. Dutt, B. V., H. Temkin, E. D. Kolb, and W. A. Sunden. *Appl. Phys. Lett.* **47,** 111 (1985).

56. Caneau, C., A. K. Srivastava, A. G. Dentai, J. L. Zyskind, and M. A. Pollack. *Electron. Lett.* **21,** 815 (1985).
57. Chiu, T. H., W. T. Tsang, J. A. Ditzenberger, and J. P. van der Ziel. *Appl. Phys. Lett.* **49,** 1051 (1986).
58. Caneau, T., J. L. Zyskind, J. W. Sulhoff, T. E. Glover, J. Centanni, C. A. Burrus, A. G. Dentai, and M. A. Pollack. *Appl. Phys. Lett.* **51,** 764 (1987).
59. Darkin, A. E., P. G. Eliseev, B. N. Sverdlov, A. E. Bochkarev, L. M. Dolginov, and L. V. Druzhinina. *IEEE J. Quantum Electron.* **QE-23,** 1089 (1987).
60. Bochkarev, A. E., L. M. Dolginov, A. E. Darkin, P. G. Eliseev, and B. N. Sverdlov. *Sov. J. Quantum Electron.* **18,** 1362 (1988).
61. Zyskind, J. L., J. C. DeWinter, C. A. Burrus, J. C. Centanni, and M. A. Pollack. *Electron. Lett.* **25,** 568 (1989).
62. Choi, H. K., and S. J. Eglash. *Appl. Phys. Lett.* **59,** 1165 (1991); *IEEE J. Quantum Electron.* **QE-27,** 1555 (1991).
63. Menna, R. J., D. R. Capewell, R. U. Martinelli, P. K. York, and R. E. Enstrom. *Appl. Phys. Lett.* **59,** 2127 (1991).
64. Kobayashi, K., S. Kwata, A. Gomyo, I. Hino, and T. Suzuki. *Electron. Lett.* **21,** 931 (1985).
65. Ikeda, M., Y. Mori, H. Sato, K. Kaneko, and N. Watanabe. *Appl. Phys. Lett.* **47,** 1027 (1985).
66. Ishikawa, M., Y. Ohba, H. Sugawara, M. Yamamoto, and T. Nakanisi. *Appl. Phys. Lett.* **48,** 207 (1986).
67. Nakono, K., M. Ikeda, A. Toda, and C. Kojima. *Electron. Lett.* **23,** 894 (1987).
68. Shiozawa, H., H. Okuda, M. Ishikawa, M. G. Hatakoshi, and Y. Uematsu. *Electron. Lett.* **24,** 877 (1988).
69. Okuda, H., M. Ishikawa, H. Shiozawa, Y. Watanabe, K. Itaya, K. Nitta, G. Hatakoshi, Y. Kokubun, and Y. Uematsu. *IEEE J. Quantum Electron.* **QE-25,** 1477 (1989).
70. Ishikawa, M., H. Okuda, K. Itaya, H. Shiozawa, and Y. Uematsu. *Jpn. J. Appl. Phys.* **28,** 1615 (1989).
71. Itaya, K., G. Hatakoshi, Y. Watanabe, M. Ishikawa, and Y. Uematsu. *Electron. Lett.* **26,** 214 (1990).
72. Itaya, K., Y. Watanabe, M. Ishikawa, G. Hatakoshi, and Y. Uematsu. *Appl. Phys. Lett.* **56,** 1718 (1990).
73. Ueno, Y., K. Endo, H. Fujii, K. Kobayashi, K. Hara, and T. Yuasa. *Electron. Lett.* **26,** 1726 (1990).
74. Itaya, K., M. Ishikawa, and Y. Uematsu. *Electron. Lett.* **26,** 839 (1990).
75. Hatakoshi, G., K. Itaya, M. Ishikawa, M. Okajima, and Y. Uematsu. *IEEE J. Quantum Electron.* **QE-27,** 1476 (1991).
76. Hamada, H., M. Shono, S. Honda, R. Hiroyama, K. Yodoshi, and T. Yamaguchi. *IEEE J. Quantum Electron.* **QE-27,** 1483 (1991).
77. Itaya, K., M. Ishikawa, G. Hatakoshi, and Y. Uematsu. *IEEE J. Quantum Electron.* **QE-27,** 1496 (1991).
78. Chong, T.-H., and K. Kishino. *IEEE J. Quantum Electron.* **QE-27,** 1501 (1991).
79. Normandin, R., S. Letourneau, F. Chatenoud, and R. L. Williams. *IEEE J. Quantum Electron.* **QE-27,** 1520 (1991).
80. Nitta, K., N. Shimoda, and G. Hatakoshi. *Electron. Lett.* **27,** 1660 (1991).
81. Katsuyama, T., I. Yoshida, J. Shinkai, J. Hashimoto, and H. Hayashi. *Appl. Phys. Lett.* **59,** 3351 (1991).
82. Valster, A., C. J. van der Poel, M. N. Finke, and M. J. B. Boermans. *Electron. Lett.* **28,** 144 (1992).
83. Rennie, J., M. Watanabe, M. Okijima, and G. Hatakoshi. *Electron. Lett.* **28,** 150 (1992).
84. Jang, D.-H., Y. Kaneko, and K. Kishino. *Electron. Lett.* **28,** 428 (1992).

85. Haden, J. M., D. W. Nam, D. F. Welch, J. G. Endriz, and D. R. Scifres. *Electron. Lett.* **28,** 451 (1992).
86. Shono, M., H. Hamada, S. Honda, R. Hiroyama, K. Yodoshi, and T. Yamaguchi. *Electron. Lett.* **28,** 905 (1992).
87. Ueno, Y., H. Fujii, H. Sawano, and K. Endo. *Electron. Lett.* **28,** 860 (1992).
88. Hashimoto, J., T. Katsuyama, J. Shinkai, I. Yoshida, and H. Hayashi. *Electron. Lett.* **28,** 1329 (1992).
89. Geels, R. S., D. P. Bour, D. W. Treat, R. D. Brigans, D. F. Welch, and D. R. Scifres. *Electron. Lett.* **28,** 1043 (1992).
90. Kaneko, Y., A. Kikuchi, A. Nomura, and K. Kishino. *Electron. Lett.* **26,** 657 (1990).
91. Valster, A., M. N. Finke, M. J. B. Boermans, J. M. M. van der Heijden, C. J. G. R. Spreuwenberg, and C. T. H. F. Liedenbaum. IEEE Semiconductor Laser Conf. Paper PD-12, 1990.
92. Hasse, M., J. Qui, J. DePuydt, and H. Cheng. *Appl. Phys. Lett.* **59,** 1272 (1991).
93. Jeon, H., J. Ding, W. Patterson, A. V. Nurmikko, W. Xie, D. Grillo, M. Kobayashi, and R. L. Gunshor. *Appl. Phys. Lett.* **59,** 3619 (1991).
94. Jeon, H., J. Ding, A. V. Nurmikko, W. Xie, D. C. Grillo, M. Kobayashi, R. L. Gunshor, G. C. Hua, and N. Otsuka. *Appl. Phys. Lett.* **60,** 2045 (1992).
95. Ishihara, T., G. Brunthaler, W. Walecki, M. Hagerott, A. V. Nurmikko, N. Samarth, and J. Furdyna. *Appl. Phys. Lett.* **60,** 2460 (1992).
96. Wang, S. Y., I. Hauksson, J. A. Simpson, H. Stewart, S. J. A. Adams, J. M. Wallace, Y. Kawakami, K. A. Prior, and B. C. Cavenett. *Appl. Phys. Lett.* **61,** 506 (1992).

Chapter 14
DEGRADATION AND RELIABILITY

14.1 INTRODUCTION

The performance of semiconductor lasers can degrade during their operation. This degradation is usually characterized by an increase in the threshold current that is often accompanied by a decrease in the external differential quantum efficiency. The dominant mechanism responsible for this degradation is determined by one or several of the fabrication processes including epitaxy, device processing, and bonding. In addition, the degradation rate of lasers processed from a given wafer depends on the operating conditions, namely, the operating temperature and injection current. Although many of the degradation mechanisms are not fully understood, an extensive amount of empirical observations exists in the literature. These observations have allowed the fabrication of InGaAsP laser diodes with an extrapolated median lifetime in excess of 25 years[1] at an operating temperature of 10°C.

Detailed studies of the degradation mechanisms in injection laser diodes have been motivated by the desire to have reasonably accurate estimates of the operating lifetime before using the diodes in practical systems. For many applications the laser sources are expected to operate reliably over a period in excess of ten years. An appropriate reliability assurance procedure becomes necessary, especially for applications such as undersea lightwave transmission systems, where the laser replacement cost is very high.[1,2] Reliability assurance is usually carried out by operating the laser under a high stress (e.g., high temperature) that enhances the degradation rate so that a measurable value can be obtained in an operating time of a few hundred hours. The degradation rate under normal operating conditions can then be obtained from the measured high-temperature degradation rate using the concept of an activation energy. The purpose of this chapter is to discuss (i) degradation mechanisms and (ii) strategies for reliability assurance.

The degradation mechanisms can be separated into three categories. They are: (i) defect formation in the active region; (ii) catastrophic mirror damage

at high power densities; and (iii) degradation of current-confining junctions. It is important to mention at the outset that although reliable semiconductor lasers have been fabricated using the InGaAsP alloy system, the study of degradation mechanisms has been largely descriptive and the results in many cases tentatively interpretative.

Laser diodes are operated at high injected current densities, which create high-energy electrons and holes, thermal gradients, potential for strain fields, and a high nonradiative recombination rate inside the active region. These factors can promote the motion, multiplication, and growth of isolated defects into clusters, which can significantly degrade the performance of lasers.

Catastrophic degradation due to mirror damage usually occurs during pulsed, high-power operation. The surface of a crystal is an imperfect lattice with many dangling bonds that absorb impurities from the air. These absorbed impurities form defect sites that cause excess optical absorption. When the optical intensity at the laser facet exceeds a certain critical value, the localized heat at the mirror facet can be large enough to cause melting and hence destroy the facet.

As discussed in Chapter 5, index-guided buried-heterostructure lasers use reverse-biased junctions for confining the current to the active region. During the aging process under normal operating conditions, defects may be generated in the current-confining junctions, which reduces the ability of these junctions to confine the current to the active region. The observed increase in threshold current may then be due to increased leakage current.

14.2 DEFECT FORMATION IN THE ACTIVE REGION

In this section we discuss observations on defect motion and the development of specific types of defect clusters in InGaAsP. The high density of recombining electrons and holes and a possible presence of strain and thermal gradients can promote defect formation in the active region of the laser. The defect structures that are generally observed are the *dark spot defect* (DSD) and the *dark line defect* (DLD). These defect structures were first observed in AlGaAs lasers and LEDs.[3]

The DLD, as the name suggests, is a region of greatly reduced radiative efficiency of roughly linear form. It was first observed in the active region of an aged AlGaAs proton-stripe double-heterostructure laser.[3] The DLD appears as a dark linear feature crossing the luminescent stripe at 45° in degraded lasers. Since the active stripe is oriented along the ⟨110⟩ direction, the DLDs are oriented along the ⟨100⟩ direction. An extensive literature[3–9] exists on DLDs in AlGaAs lasers, from which the following conclusions can be drawn. These defects are seen to form in regions that do not differ in

radiative efficiency from the surrounding regions prior to the formation of the defect. Optically excited AlGaAs lasers exhibit similar DLD growth.[5] Transmission electron microscopy (TEM) studies have shown that the DLD structure observed in injection lasers and in optically pumped semiconductor lasers is identical and therefore results from the same process.[6] These observations suggest that the phenomenon of DLD growth is intrinsic to the material and is not related to (though may be influenced by) the migration of metals from the contact, and *p-n* junction, or the particular dopant used.

In addition to the DLD, dark spot defects (DSDs) are also observed. The DSD is also a region of reduced radiative efficiency in the active region but lacks the linear aspect of DLDs. DSDs are generally associated with regions of poor epitaxial growth that grow in size or localized defects in the substrate that propagate to the epitaxial layers during device operation.

14.2.1 Experimental Techniques

DLDs and DSDs can be easily observed by means of electroluminescence, photoluminescence, or cathodoluminescence. However, a detailed study requires the use of TEM techniques where the sample preparation is a nontrivial exercise and the quality of the results obtained is determined exclusively by the sample preparation expertise. The defects can also be identified by techniques such as deep-level transient spectroscopy[10] (DLTS), scanning electron microscopy (SEM), and the electron-beam-induced current (EBIC) mode of SEM. We now briefly describe these techniques and the results obtained.

14.2.2 Electroluminescence

The observation of nonluminescent regions or regions of reduced radiative efficiency inside the active region of a semiconductor laser requires the fabrication of a special laser structure. Figure 14.1 schematically shows this laser structure, commonly known as *the window laser*. Typically 20–25 μm wide, the window is formed on the substrate side using photolithographic techniques. Since the InGaAsP laser emits its light with photon energies smaller than the band gap of the InP substrate, spontaneous emission from the active region can be observed directly through the window. The window-laser structure allows continuous monitoring of the luminescence efficiency of the active region and is also compatible with the normal "*p*-down" bonding configuration. Since luminescence occurs as a result of current injection (electrical excitation), it is referred to as electroluminescence (EL).

Very often, degraded lasers show dark regions in the active-stripe EL when

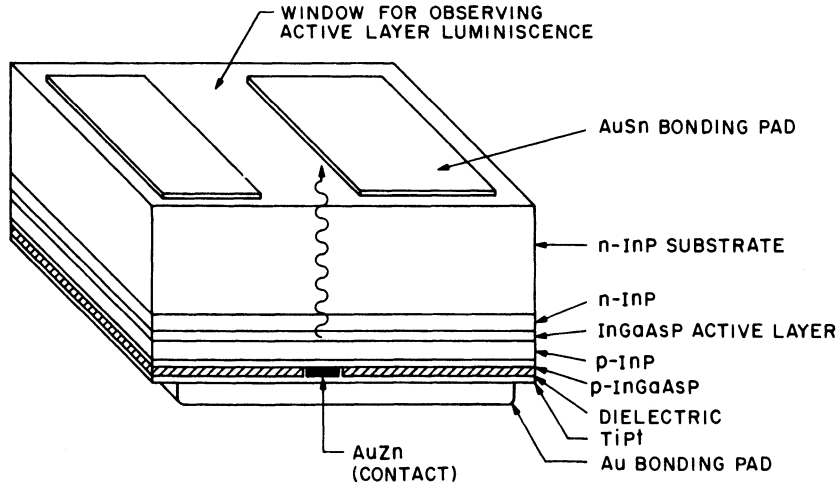


Fig. 14.1 Schematic of a "window laser" structure used for the observation of luminescence from the substrate side. Typically the window is 20–25 μm wide.

observed through the window, as shown in Fig. 14.2. These dark regions sometimes appear in a random fashion without any particular crystallographic orientation. Both DLDs and DSDs have been observed in LEDs[11–14] and in InGaAsP double-heterostructure material.[15,16]

Most of the work related to the observation and generation of DLDs and DSDs under current injection has been reported for LEDs. Although LEDs with extrapolated room-temperature lifetimes in excess of 10^9 h have been reported, DLDs and DSDs do occur in these devices, and when they occur, they indeed limit the lifetime. Yamakoshi et al.[13] have observed generation of the ⟨110⟩-type defects at operating temperatures higher than 170°C. Temkin et al.[14] observed a similar degradation mode under storage aging at 220°C without current injection; Fig. 14.3 shows their results. The ⟨110⟩ DLDs observed in their samples originated from inclusion-like defects that were traced to the first-grown *n*-InP layer and/or to its interfaces with the substrate and active layers. No ⟨100⟩-type DLDs were observed. Additionally, under thermal aging without current injection, no change in the *I-V* characteristic was seen, whereas under aging with current injection, the *I-V* characteristics indicated the presence of leakage current. Ueda and coworkers[11,17] have observed ⟨110⟩-type DLDs and DSDs in aged LEDs. The DLDs corresponded to misfit dislocations and were identified by TEM. X-ray analysis of the DSDs showed that these regions are precipitate-like defects with an excess of In and P. It was suggested that such regions may be formed during epitaxy; for example, thermal decomposition of the substrate

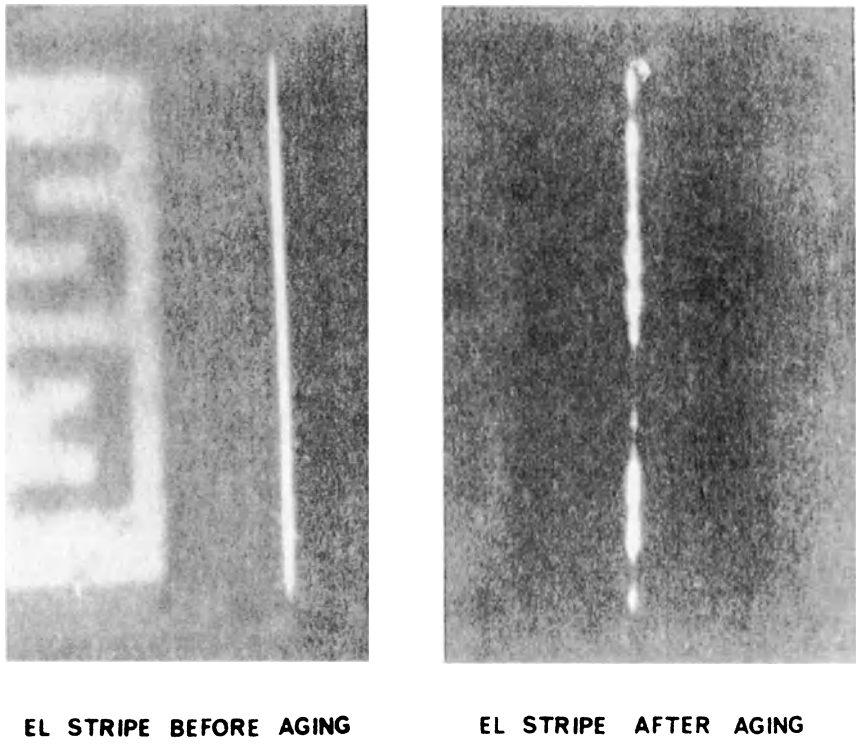


Fig. 14.2 Active-stripe electroluminescence (EL) observed through the window of a degraded laser showing dark regions. Left photograph shows the EL stripe of an undegraded laser. (Courtesy E. J. Flynn)

just prior to the growth of the *n*-InP buffer layer or thermal decomposition at the interface between the *n*-InP layer and the quaternary active layer.

14.2.3 Photoluminescence

Photoluminescence is the luminescence occurring under photoexcitation of the semiconductor. A scanning photoluminescence technique was used in the first report of DLDs observed in InGaAsP.[15] Figure 14.4 shows a schematic diagram of the apparatus used in these studies. Electron-hole pairs can be created at a suitable depth from the surface of the sample by choosing the right excitation wavelength. The apparatus also allows for the measurement of the catastrophic damage threshold. The defects may be created and observed by focusing the beam at a given point of the sample for a suitable length of time.

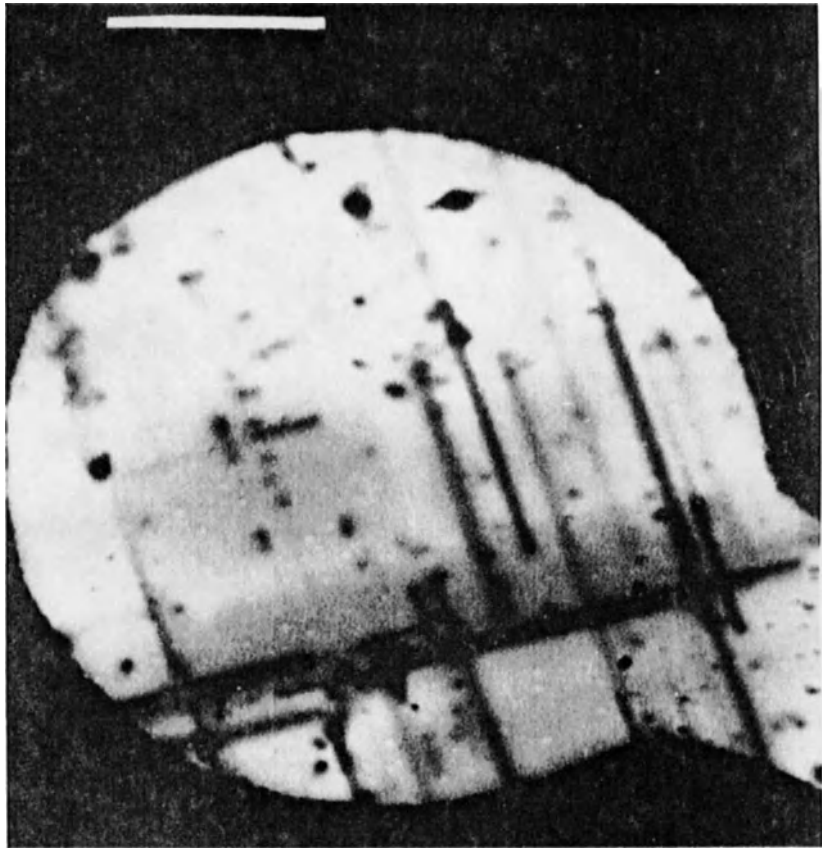


Fig. 14.3 ⟨110⟩ DLDs observed in thermally aged InGaAsP LEDs. The photograph shows the EBIC image. (Courtesy H. Temkin)

Johnston et al.[15] observed DLDs in both single-layer and double-heterostructure InGaAsP. The sample chosen for the experiment had a lattice mismatch of less than 0.1%. Figure 14.5 shows the dependence of the minimum photoexcitation energy density (intensity × exposure time) required for degradation on the active-layer thickness. The observed lower threshold for thick layers is interpreted as resulting from strain in the epitaxial layers increasing with increasing thickness. Dislocations can propagate more easily in the presence of a strain field. The TEM studies of optically degraded samples revealed small-scale defect clusters oriented along ⟨110⟩ in the defective regions. The characteristics of the degraded region were analyzed in detail in Ref. 16. Dislocations clusters oriented along (110) were found

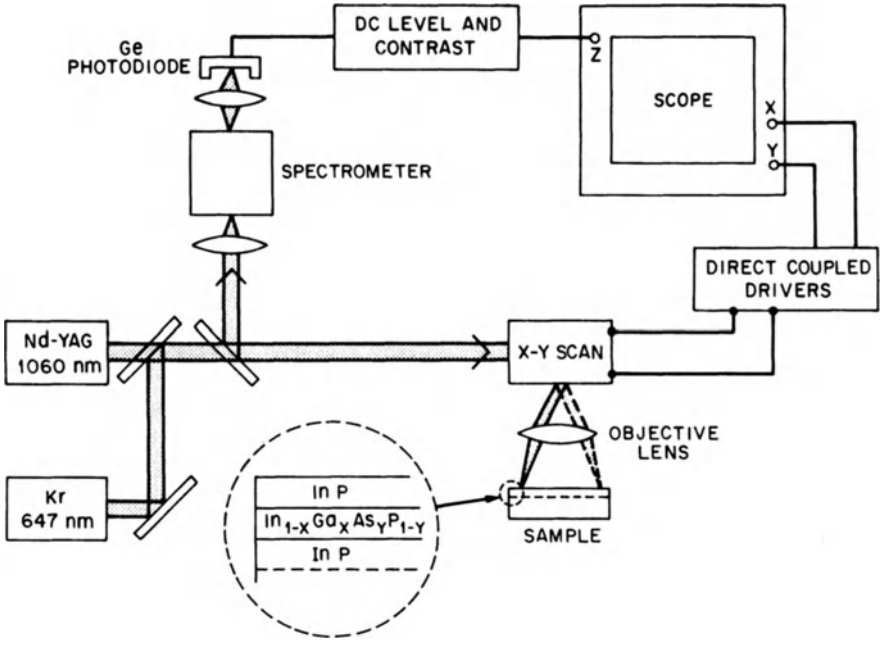


Fig. 14.4 Schematic diagram of the apparatus used for photoluminescence studies and DLD generation in InGaAsP double-heterostructure samples. (After Ref. 15)

in both the thick and thin samples even though the threshold power for degradation differed by several orders of magnitude.

14.2.4 Cathodoluminescence

The cathodoluminescence imaging technique has been extensively used to study dark defects in semiconductors.[18–20] The method involves exciting the semiconductor material with an electron beam typically obtained from a scanning electron microscope. The created electron-hole pairs recombine to emit light. The cathodoluminescence imaging technique involves guiding this radiation (using a light pipe) from the sample to a photomuliplier tube located outside the SEM chamber. Figure 14.6 shows a schematic of a cathodoluminescence imaging system. Figure 14.7 shows examples of spatially resolved images of the light-emitting region in the *p*-InP confining layer of different LEDs. Dark, circular, or oblate features represents examples of DSDs. Chin et al.[20] showed that the migration of gold from the *p* contact is responsible for the DSD formation in LEDs.

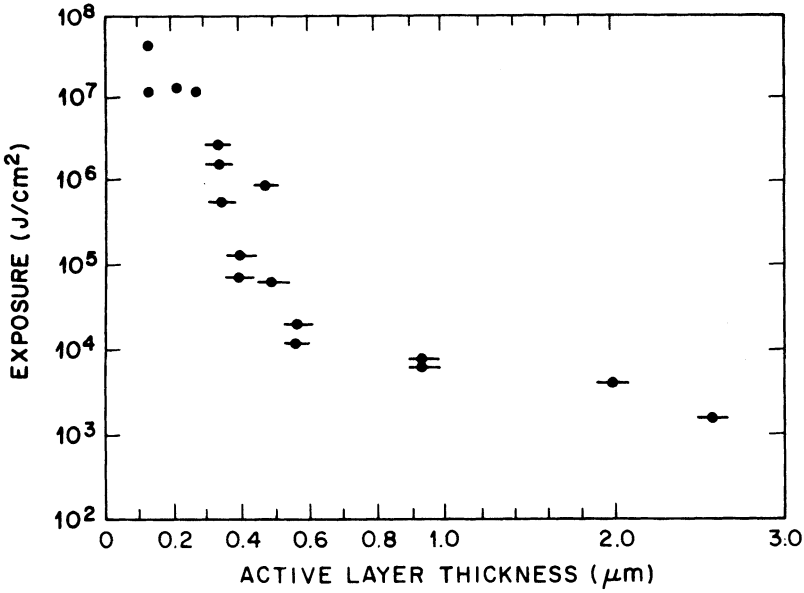


Fig. 14.5 Variation of the degradation-threshold energy density (intensity × exposure time) as a function of InGaAsP layer thickness. (After Ref. 15)

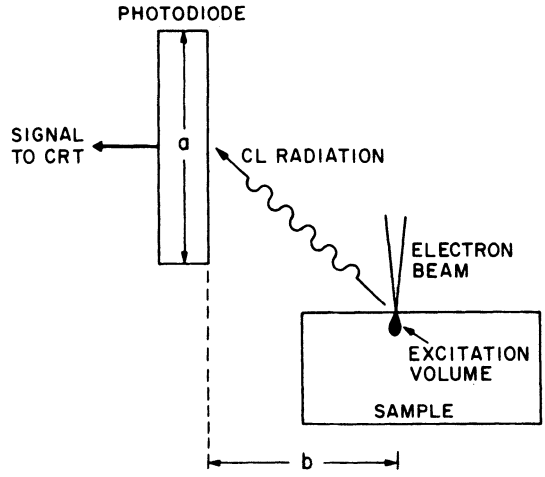


Fig. 14.6 Schematic of a cathodoluminescence (CL) imaging system. Resolution and sensitivity of the system are determined by distances a and b. (After Ref. 20)

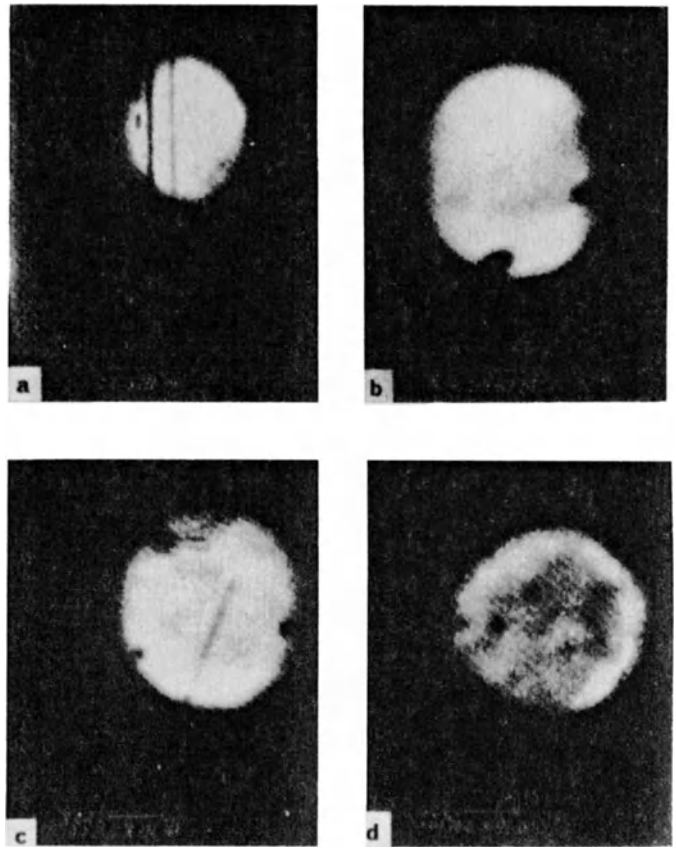


Fig. 14.7 Spatially resolved cathodoluminescence images of the light emitting region in the *p*-InP confining layer of four different LEDs. Dark features represent examples of dark spot defects. (After Ref. 19)

14.2.5 Dark Defects under Accelerated Aging

Accelerated aging techniques, which include high-temperature and high-power operation, are generally used to estimate the usable lifetime of injection lasers under normal operating conditions. The generation rate of DSD and DLDs is enhanced under accelerated aging. When the accelerated aging is done at high temperature, an activation energy can be defined for the generation rate of defects in the active region. The activation energy E_a is defined using

$$t_d = t_0 \exp\left(\frac{E_a}{k_B T}\right) \tag{14.2.1}$$

where t_d is the generation time for the first defect, t_0 is a constant, and T is the operating temperature. Fukuda et al.[21] have measured the generation time of DSDs and DLDs in InGaAsP gain-guided lasers operating at 1.3-μm and 1.55-μm wavelengths. The generation of DSDs and DLDs was observed by electroluminescence using a window-laser configuration. The lasers that did not exhibit DSDs and DLDs operated for a long time without degradation. Figure 14.8 shows the measured pulsed threshold current $I_{th}^{p}(t)$ normalized to the initial value. The increase in threshold current with time was associated with an increase in the number of DSDs and ⟨100⟩ DLDs. Furthermore, Ga- and As-rich regions in the active layer were correlated with the location of the DSDs. A saturation in the number of DSDs and DLDs occurred in about 50 h of aging at 250°C; beyond that time the increase in the threshold current caused by further aging also showed a tendency to saturate. Moreover, samples without any DSDs or DLDs exhibited very little change in threshold even after 3,000 h of aging. It was concluded that dark defects were not created during aging (under the experimental conditions of Ref. 21) and that their origins were already present in samples that exhibited DSDs and DLDs after a short-term aging (less than 100 h) at 250°C.

Fukuda et al.[21] found that the generation time of DSDs and the generation rate of ⟨100⟩ DLDs are weakly dependent on temperature, with activation energies of 0.16 eV and 0.2 eV, respectively. However, they found that the

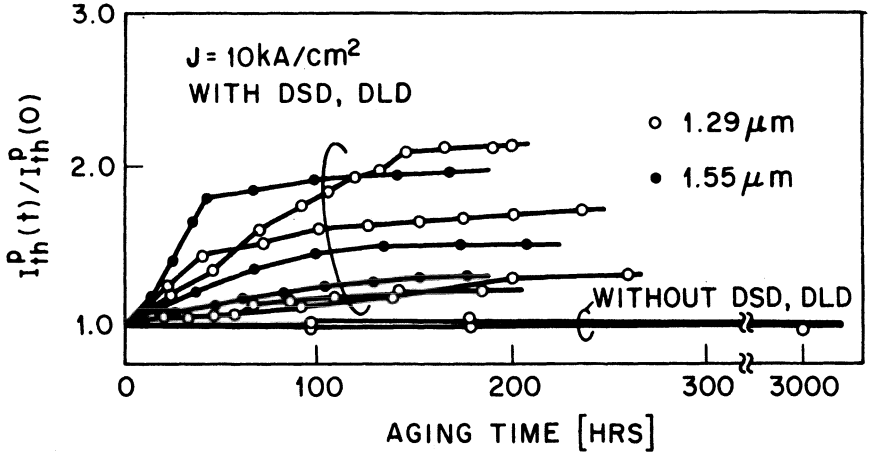


Fig. 14.8 Pulsed threshold current (normalized by its initial value) as a function of aging time. Gain-guided InGaAsP lasers were operated at 250°C with a current density $J = 10\ \text{kA/cm}^2$. (After Ref. 21)

generation time of the first DSD depends strongly on the operating current density. From their experimental result, they derived the following empirical relation:

$$t_d \propto \exp(-AJ^2) \tag{14.2.2}$$

where J is the current density and A is a constant.

The results of Ref. 21 on the aging characteristics of lasers differ considerably from those of Ref. 22 where DSDs and ⟨100⟩ DLDs were observed in a surface-emitting LED. From the experimental observations, the generation time of the first DSD can be empirically expressed by the relation[22]

$$t_d \propto J^{-2} \exp\left(\frac{E_a}{k_B T}\right) \tag{14.2.3}$$

where $E_a = 1.2$ eV. The considerable difference in the activation energies of DSD generation reported by the two groups would suggest that different mechanisms of DSD generation are operative in the two structures. Which mechanism is actually dominant may relate to details of device fabrication, including processing, bonding, and epitaxy.

14.3 CATASTROPHIC DEGRADATION

For the purposes of this section, we define catastrophic degradation as the sudden degradation in the performance characteristics of laser diodes associated with the application of a large current pulse. This phenomenon has been studied in detail in AlGaAs double-heterostructure lasers and epitaxial layers. For AlGsAs lasers the degradation is caused by strong absorption of the stimulated emission at the facet, which causes local heating and subsequent melting of the material near the facet. The degraded-facet region typically exhibits dislocation networks and multiple dislocations loops, which are thought to be generated during the cooling of the molten region.

Catastrophic degradation in InGaAsP-InP double-heterostructure material under intense optical excitation has been studied.[23] The threshold power for catastrophic damage is about one order of magnitude larger in InGaAsP material than in AlGaAs material. Catastrophic degradation of InGaAsP lasers following a large current pulse has also been studied.[24] Generally, no facet degradation, which is characteristic of melting at the active region (similar to AlGaAs lasers[25,26]), has been observed. This may be partly due to a smaller surface recombination rate in InGaAsP than in AlGaAs. In cases where facet degradation has been seen, it extended from the top contact to

the p-InP cladding layer, implying that heating was caused by a large amount of current passing near the facet.

In many cases the degradation may not be catastrophic. Then the increase in the threshold current is accompanied by a decrease in the external differential quantum efficiency and by a "soft" turn-on in the I-V characteristics that is characteristic of a leakage path. This is especially the case for buried-heterostructure lasers that utilize reverse-biased junctions for lateral current confinement. Ueda et al.[24] have measured the current needed to achieve a certain level of degradation as a function of the pulse width for V-grooved substrate buried-heterostructure lasers. Figure 14.9 shows their results. The lasers did not degrade below the current levels denoted by circles; above those points (neighboring crosses) catastrophic degradation did occur. For 0.1-μs current pulses, degradation did not occur up to the maximum available current of 20 A. Figure 14.9 shows that the current required for degradation decreases with increasing pulse width.

The degraded regions of catastrophically degraded diodes were analyzed using selective defect etching and energy-dispersive x-ray spectroscopy;[24] contact metals (e.g., Au) in the core of degraded regions were found. It is suggested that catastrophic degradation may be caused by local degradation of some internal current-confining region that can result in a large localized leakage current and subsequent melting.

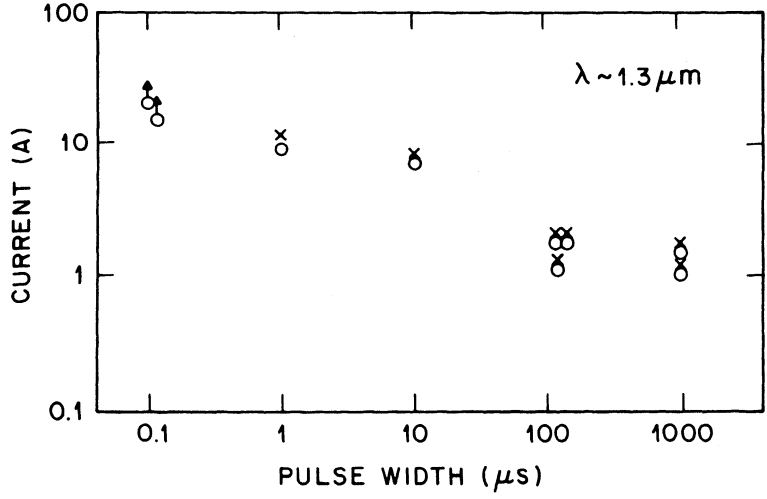


Fig. 14.9 Degradation threshold versus current pulse width for InGaAsP lasers. Each circle-cross pair represents one laser which degraded when the current was at the level denoted by cross (it was not degraded up to circle-level). (After Ref. 24)

14.4 DEGRADATION OF CURRENT-CONFINING JUNCTIONS

As discussed in Sec. 5.5, many high-performance laser structures utilize current-restrictive layers so that most of the injected current flows through the active region. Effective current injection to the active region is necessary in order to obtain low threshold currents and high output powers. A mode of degradation associated with buried-heterostructure lasers is an increase in the leakage current (the current flowing outside the active region) under accelerated aging. This in turn increases the device threshold and decreases the external differential quantum efficiency.

An increase in leakage current usually appears as a soft turn-on in the *I-V* characteristics of the laser. An easily recognizable signature is a "bump" in the $I\,dV/dI$–I characteristics, as shown in Fig. 14.10.[27] We have seen in Sec. 5.6 that such a bump is indicative of a resistive shunt path. An increase in leakage current not only increases the device threshold but also generally decreases the external differential quantum efficiency.

The observations of electron-beam-induced current (EBIC) at laser facets is a useful technique in detecting defects at current-confining junctions. In the buried-heterostructure laser (Fig. 5.25), the *p-n* junctions help in confining the current to the active region. Figure 14.11 shows the aging behavior of three buried-heterostructure lasers along with their EBIC images.[28] The bright regions indicate EBIC signals from the *p-n* junctions, while the dark

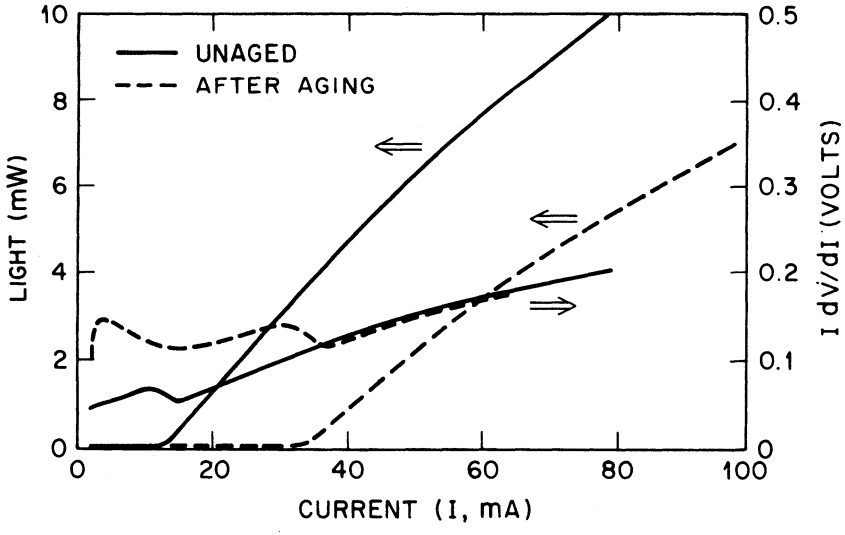


Fig. 14.10 Optical and electrical characteristics of a buried-heterostructure laser before (solid lines) and after (dashed lines) aging. *V* is the voltage across the laser. (Courtesy E. J. Flynn)

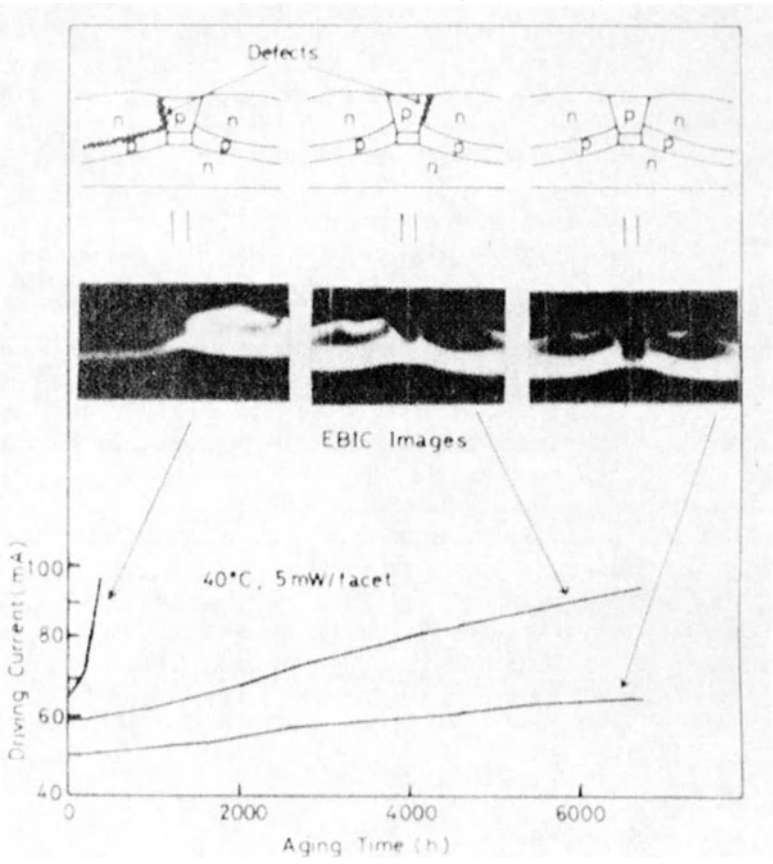


Fig. 14.11 EBIC signals from the facet of three lasers showing different aging behaviors. The absence of an EBIC signal (dark region) near a junction boundary indicates the presence of defects at the *p-n* junction. The defective regions are denoted by crosses in the schematic cross sections. Note the absence of defects for the laser with the lowest degradation rate. (After Ref. 28 © 1983 IEEE)

regions show the presence of defects. The defect can be either the absence of a junction (growth defect) or an electron (or hole) trap generated during aging.

14.5 RELIABILITY ASSURANCE

In some semiconductor laser applications the system design lifetime is long (~20–25 years), and the replacement of components (e.g., lasers) can be probihitively expensive. An example is provided by the repeaters of an

undersea lightwave transmission system where the failure of just a few lasers can cripple the system. Thus it is important to have a strategy for establishing the expected operating lifetime of a semiconductor laser.

Laser failure can occur by several means, including (i) infant mortality (lasers with grown-in defects that degrade very quickly); (ii) chance failures, these could be catastrophic damage due to external factors; and (iii) a gradual degradation mechanism in which some characteristic of laser operation (e.g., the threshold current) changes slowly with time. In the last case an estimate of the operating lifetime can be obtained by monitoring that characteristic parameter. Examples of infant failures are lasers with DSDs or DLDs already present in the active region at the time of fabrication or lasers with defects generated in the buffer layers during epitaxy that quickly propagate to the active region, resulting in a decrease in luminescence or an increase in threshold current.

14.5.1 Stress Aging

Some lasers inhibit an initial rapid degradation after which the operating characteristics of the lasers become stable. Given a population of lasers, it is possible to quickly identify the "stable" lasers using a high-stress test (also known as the purge test).[29,30] The assumption behind the test is that operating the laser under a set of high-stress conditions (e.g., high current, high temperature, or high power) causes weak lasers to fail and possible winners to stabilize. The concept is substantiated by the observations of Ref. 21 (Fig. 14.8), where it was found that in some samples the DLDs and DSDs grew rapidly under stress aging (less than ~ 50 h at 250°C). Associated with DLD and DSD generation was a rapid increase in the threshold current. It was also observed that after an initial increase, the number of DLDs and DSDs saturated.

Similar degradation characteristics have been observed for 1.3-μm InGaAsP buried-heterostructure lasers;[31] the results are shown in Fig. 14.12. The lasers were aged at a constant power output of 5 mW/facet at 60°C. The change in the operating current with the aging time was initially rapid and thereafter it slowed down considerably, similar to the observations of Ref. 21. The initial rapid degradation ("incubation period") took place in the first 500–1,000 h. A high-stress aging test can considerably reduce the aging time needed for the completion of this initial rapid degradation mode. The increase in the operating current after stress aging has also been measured.[1] Figure 14.13 shows the data. The stress conditions required CW operation of the laser at 100°C with a 250-mA current. After 0, 20, and 40 h of stressed aging the operating current for obtaining 3-mW/facet power at 60°C was measured. The lasers were of the buried-heterostructure type. As Fig. 14.13

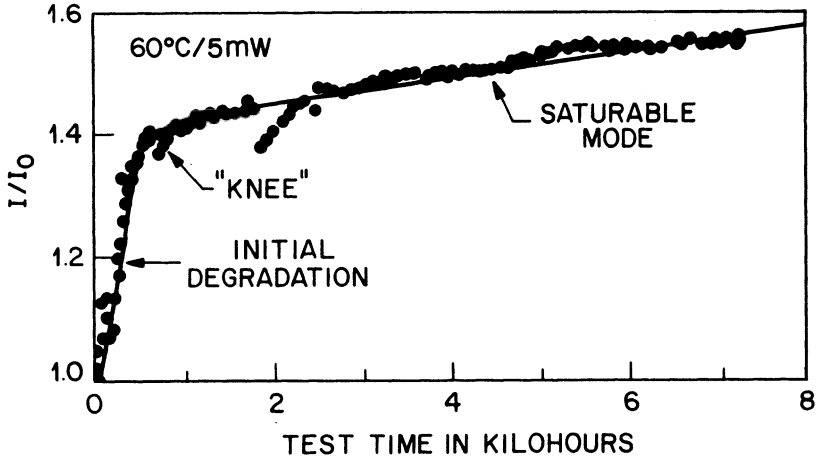


Fig. 14.12 Change in operating current as a function of aging time for a 1.3-μm InGaAsP buried-heterostructure laser aged at 60°C under a constant output power of 5 mW/facet. (After Ref. 31)

shows, some lasers exhibited an increase in operating current before stabilizing (similar to results of Ref. 21) while others exhibited stable characteristics without a significant increase in the opening current. Figures 14.8 and 14.13 show that the high-stress test can be used as a screening procedure to identify robust devices.

Determining the duration and specific conditions for stress aging is critical to the success of this screening procedure. The aim is (i) to identify the weak lasers, which would fail in a short time; (ii) to stabilize the initial high-degradation mode that may be present in some lasers and may give misleading indications of long-term reliability; and (iii) to select lasers whose degradation is governed by a slow acting, long-term mechanism that can be thermally accelerated to allow for a determination of the expected operating lifetime.

14.5.2 Activation Energy

Central to the determination of the expected operating lifetime is the concept of thermally accelerated aging, the validity of which for AlGaAs injection lasers was shown by Hartman and Dixon.[32] The lifetime τ at a temperature T is empirically found to vary as

$$\tau = \tau_0 \exp\left(\frac{E_a}{k_B T}\right) \tag{14.5.1}$$

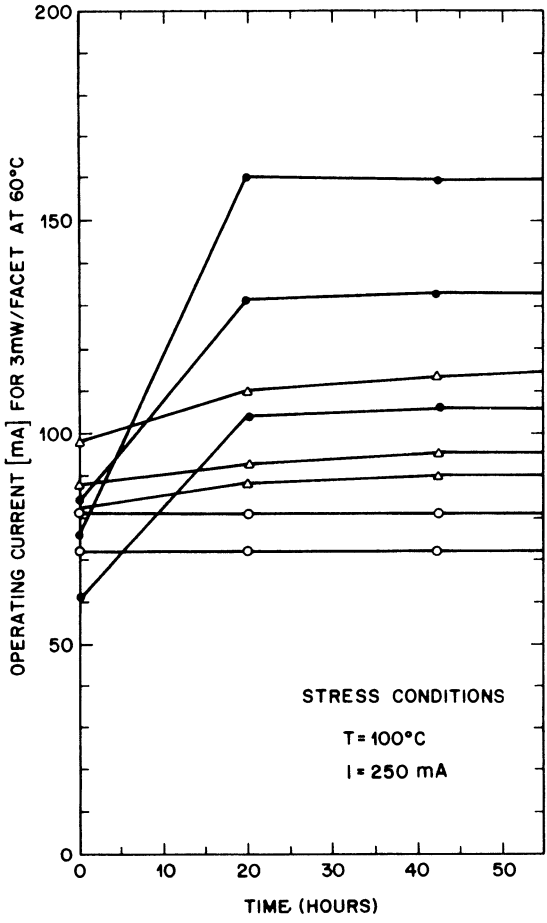


Fig. 14.13 Operating current required for maintaining an output of 3 mW/facet at 60°C as a function of aging time under high-stress conditions ($T = 100°C$ and $I = 250$ mA). (After Ref. 1)

where E_a is the activation energy and τ_0 is a constant. A similar relation also holds for the degradation rate. Equation (14.5.1) shows that by operating the laser at high temperatures (60–80°C), it is possible to establish the expected operating lifetime under normal operating temperatures (typically in the range of 10–25°C). This allows the determination of expected lifetimes of 10–25 years using a high-temperature aging time of $\sim$500–1,000 h. It has been pointed out that the simple Arrhenius-type relationship [Eq. (14.5.1)] may

not be observed under all aging conditions and that the measured activation energy may differ under different aging conditions and temperature ranges.[33]

Measurements of the lifetimes of InGaAsP-InP buried-heterostructure lasers have been reported.[34] Figure 14.14 shows the lifetime values plotted on a log normal chart. The aging rates were measured at 50 and 70°C under a constant power output of 5 mW/facet. The horizontal bars indicate the lifetime estimated from the measured increase in the drive current after 8,500 h of operation. The failure criterion (for Fig. 14.14) was defined as a change ΔI in the drive current of 30–50 mA. The median lifetimes τ_m for 50°C and 70°C are 1.4×10^5 h and 1.7×10^4 h, respectively. Assuming an Arrhenius-type relationship given by Eq. (14.5.1), the activation energy is estimated to be 0.9 eV.

Hakki et al.[31] have estimated the activation energy of 1.3-μm InGaAsP buried-heterostructure lasers from a measurement of the degradation rate of lasers operating at the same output power (3–5 mW/facet) at two different temperatures. The degradation rates were measured after the initial degradation mode was completed, i.e., beyond the "knee" in Fig. 14.12. The

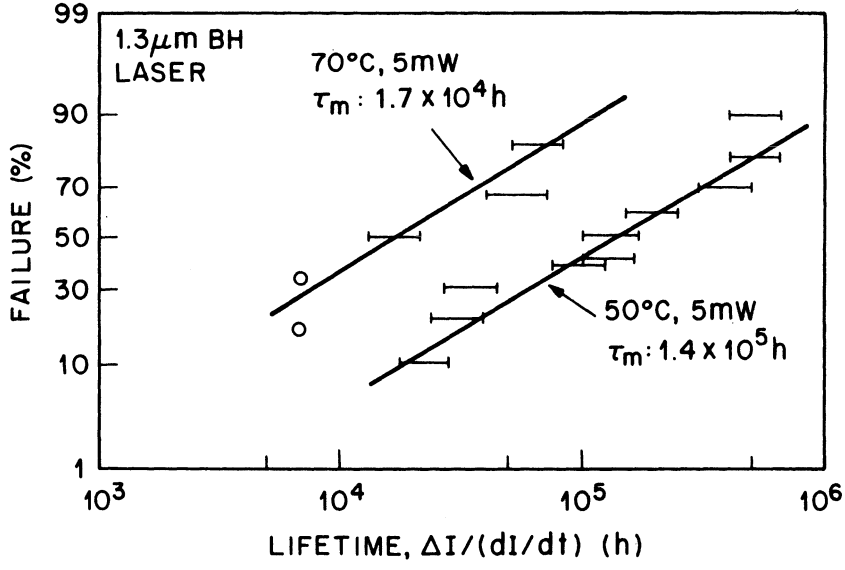


Fig. 14.14 Log-normal lifetime distributions of buried-heterostructure lasers aged at 50°C and 70°C with 5 mW/facet. Horizontal bars show lifetimes estimated from measured increases in the drive current after 8,500 h of operation. Open circles denote actual operating lifetimes. Straight lines were used to estimate the median lifetime τ_m. (After Ref. 34)

activation energy is obtained assuming an Arrhenius-type relationship for the degradation rate R; i.e.,

$$R = R_0 \exp\left(\frac{-E_a}{k_B T}\right) \tag{14.5.2}$$

where R equals the rate of change of the operating current for 5 mW/facet output power. The activation energy is then given by

$$E_a = \frac{k_B T_1 T_2}{T_2 - T_1} \ln\left(\frac{R(T_2)}{R(T_1)}\right). \tag{14.5.3}$$

Figure 14.15 shows the activation energies obtained from 26 sets of measurements. The median value of the activation energy is 1 eV, and the standard deviation σ is 0.13 eV. Figure 14.15 also indicates that 96% of the values of the activation energy fall between 0.76 and 1.28 eV. Table 14.1 gives the means, medians, and standard deviations of the measured values[31] of E_a in different temperature ranges. The mean activation energy is relatively

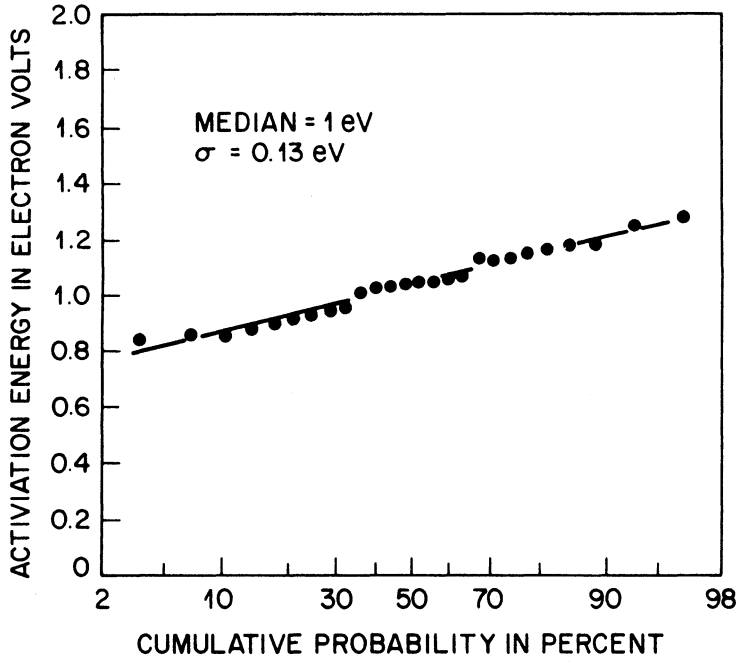


Fig. 14.15 Distribution of activation energies obtained from temperature stress aging of different InGaAsP buried-heterostructure lasers. (After Ref. 31)

Table 14.1 Activation Energy and Temperature-stress Sequence

ACTIVATION ENERGY (eV)	TEMPERATURE STEP (°C)					
	80 TO 70	70 TO 60	60 TO 70	70 TO 50	50 TO 70	40 TO 50
Mean	1.09	1.07	0.92	1.07	1.01	0.88
Median	1.02	1.17	0.90	1.10	1.03	0.82
Standard deviation	0.31	0.33	0.08	0.15	0.31	0.30

Source: Ref. 31.

invariant in the temperature range of 40–80°C. This observation shows that it is reasonable to estimate the actual operating lifetime from the measured high-temperature degradation rate.

The mean time to failure (MTTF) is an important measure of the reliability of semiconductor lasers. From the data in Fig. 14.14, the MTTF extrapolated at 25°C exceeds 10^7 hours. Experimental data[1] on the reliability of

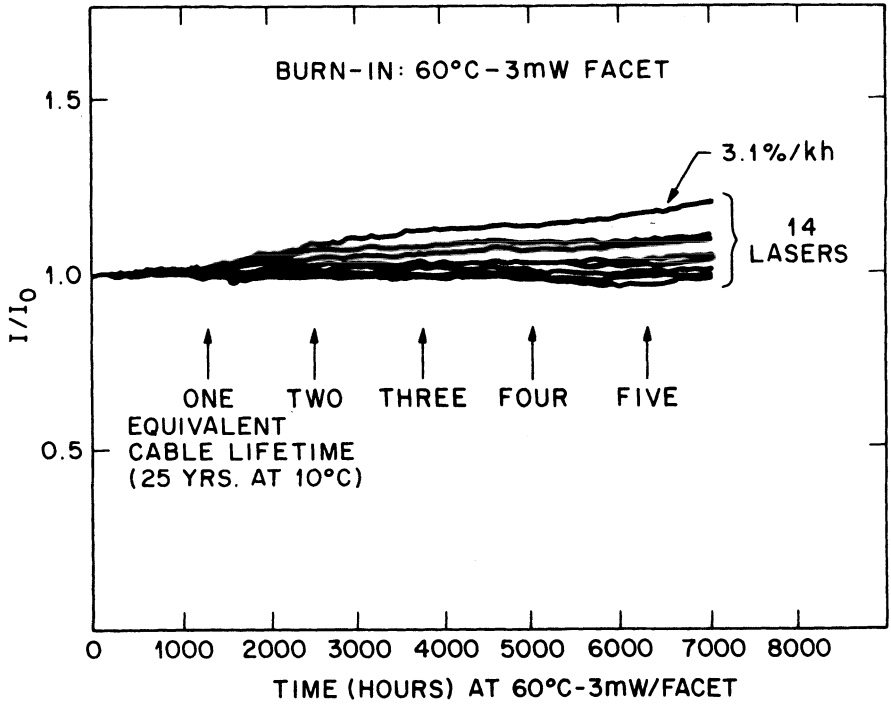


Fig. 14.16 Accelerated aging rate at 60°C with 3 mW/facet of a group of 14 prescreened lasers. Arrows are 1,200 h apart, the time duration equivalent to 25 years at 10°C. (After Ref. 1)

buried-heterostructure lasers are shown in Fig. 14.16. Normalized operating currents are shown for 14 lasers screened for ~7,000 h at 60°C and 3 mW/facet. The maximum degradation rate observed is 3.1% per 1,000 h. The measured degradation rate can be used to obtain an expected MTTF using a 50% change in I/I_0 as a failure criterion. The MTTF at an operating temperature of 10°C (the ocean-bottom temperature) is then obtained using Eq. (14.5.1) and an activation energy of 0.9 eV. The distance between arrows in Fig. 14.16 (1,200 h of burn-in time at 60°C and 3 mW/facet) represents 25 years of equivalent operating lifetime at 10°C, which is the expected cable lifetime and temperature of submarine lightwave systems.[35] It is evident from Fig. 14.16 that state-of-the-art 1.3-μm InGaAsP lasers can meet the system lifetime requirement (25 years at 10°C) if high-stress aging tests are used to eliminate the poor devices.

14.6 DFB LASER RELIABILITY

A DFB laser should emit light in a single longitudinal mode during its entire lifetime. Reliability of DFB lasers is governed not only by the MTTF but also by the spectral stability. A parameter that determines the performance of a DFB laser is the mode suppression ratio (MSR), i.e., the ratio of the intensity of the dominant lasing mode to that of the next intense side mode. DFB lasers have been fabricated that have aging rates comparable to that shown in Fig. 14.16. An example of the spectrum before and after aging is shown in Fig. 14.17. Similar measurements have been done as a function of current. These measurements show that the MSR of good DFB lasers with low light-emitting aging rates does not change significantly after aging, confirming spectral stability of the emission.

For some applications such as coherent lightwave systems, the absolute wavelength stability of the laser is important. Measured changes in the emission wavelength at 100 mA before and after aging of several devices are shown in Fig. 14.18. Most of the devices do not exhibit any change in wavelength, and the standard deviation of the change is less than 0.2 nm. The absolute wavelength stability of these devices is adequate for coherent transmission applications.

Another parameter of interest in certifying the spectral stability of a DFB laser is the change in the dynamic line width (or frequency chirp) with aging. Since the chirp depends on data rate and the bias level, a certification of the chirp stability is, in general, tied to the requirements of a specific transmission system. A measure of the effect of chirp on the performance of a lightwave system is the dispersion penalty. Measurements of dispersion penalty at 600 Mb/s for a total dispersion of 1,700 ps/nm have been done to study the effect of aging on DFB lasers. The median change in dispersion penalty was

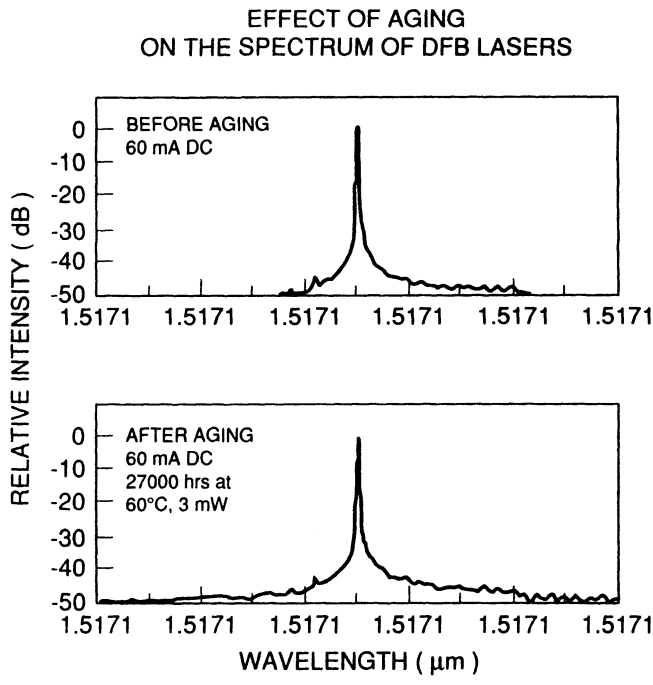


Fig. 14.17 CW spectrum of a DFB laser before and after aging.

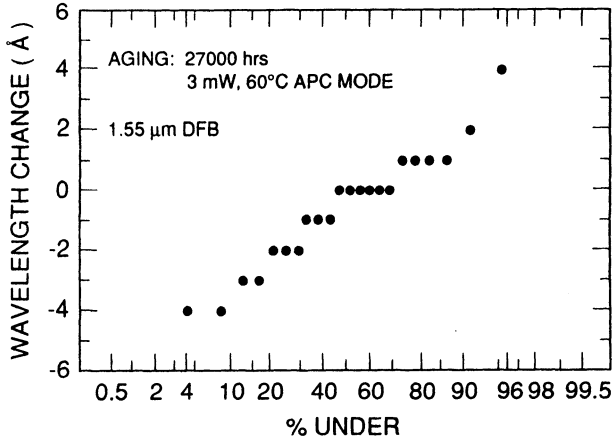


Fig. 14.18 Distribution of the wavelength change of DFB lasers before and after aging measured at 100 mA under CW operation. The aging time is the same as in Fig. 14.17.

less than 0.1 dB, and not a single device showed a change larger than 0.3 dB. This suggests that the dynamic line width or the chirp under modulation is stable.

DFB lasers are currently being used in many commercial systems. The lasers have proven to be adequately reliable even for the stringent requirements of undersea transmission. The lifetime for single-wavelength operation is assured by measuring the optical spectrum of the device at fixed intervals during the certification process. For applications where replacement costs are very high, the development of a certification process for the assurance of single-wavelength operation over the lifetime of the system is very valuable. The DFB lasers now available are robust enough for submarine cable applications. A transatlantic lightwave system (TAT-9), employing DFB lasers modulated at 590 Mb/s, became operational in 1992. Since their initial demonstration in 1962, semiconductor lasers have come a long way in terms of their reliability. A whole book devoted to laser reliability was published in 1991.[36]

PROBLEMS

14.1 Why do semiconductor lasers degrade with aging? How does degradation manifest itself in practice?

14.2 Describe various kinds of defects that can form inside the active region of a semiconductor laser. What experimental techniques can be used to observe these defects?

14.3 What is meant by electroluminescence, photoluminescence, and cathodoluminescence? Use diagrams to show how these phenomena can be observed experimentally.

14.4 Define the concept of activation energy. How can it be used to estimate the lifetime of semiconductor lasers?

14.5 What is the basic idea behind the concept of stress aging. How can it be used to assure the lifetime of semiconductor lasers?

REFERENCES

1. Nash, F. R., W. J. Sundburg, R. L. Hartman, J. R. Pawlik, D. A. Ackerman, N. K. Dutta, and R. W. Dixon. *AT&T Tech. J.* **64,** 809 (1985).
2. The reliability requirements of a submarine lightwave transmission system are discussed in a special issue of *AT&T Tech. J.* **64,** 3 (1985).
3. DeLoach, B. C., Jr., B. W. Hakki, R. L. Hartman, and L. A. D'Asaro. *Proc. IEEE* **61,** 1042 (1973).
4. Petroff, P. M., and R. L. Hartman. *Appl. Phys. Lett.* **23,** 469 (1973).
5. Johnston, W. D., and B. I. Miller. *Appl. Phys. Lett.* **23,** 1972 (1973).
6. Petroff, P. M., W. D. Johnston, Jr., and R. L. Hartman. *Appl. Phys. Lett.* **25,** 226 (1974).

7. Matsui, J., R. Ishida, and Y. Nannichi. *Jpn. J. Appl. Phys.* **14,** 1555 (1975).
8. Casey, H. C., and M. B. Panish, Chap. 8 in *Heterostructure Lasers.* New York: Academic Press, 1978.
9. Kressel, H., and J. K. Butler. *Semiconductor Lasers and Heterojunction LEDs.* New York: Academic Press, 1977.
10. Petroff, P. M., and D. V. Lang. *Appl. Phys. Lett.* **31,** 60 (1977).
11. Ueda, O., I. Umebu, S. Yamakoshi, and T. Kotani. *J. Appl. Phys.* **53,** 2991 (1982).
12. Ueda, O., S. Yamakoshi, S. Komiya, K. Akita, and T. Yamaoka. *Appl. Phys. Lett.* **36,** 300 (1980).
13. Yamakoshi, S., M. Abe, O. Wada, S. Komiya, and T. Sakurai. *IEEE J. Quantum Electron.* **QE-17,** 167 (1981).
14. Temkin, H., C. L. Zipfel, and V. G. Keramidas. *J. Appl. Phys.* **52,** 5377 (1981).
15. Johnston, W. D., Jr., G. Y. Epps, R. E. Nahory, and M. A. Pollack. *Appl. Phys. Lett.* **33,** 992 (1978); Johnston, W. D., Chap. 7 in *GaInAsP Alloy Semiconductors,* ed. T. P. Pearsall. New York: John Wiley & Sons, 1982.
16. Mahajan, S., W. D. Johnston, Jr., M. A. Pollack, and R. E. Nahory. *Appl. Phys. Lett.* **34,** 717 (1979).
17. Ueda, O., S. Yamakoshi, S. Komiya, and T. Kotani, *Proc. Defects and Radiation Effects in Semiconductors,* Ser. no. 59, Institute of Physics, Bristol, 1981.
18. A. K. Chin, in *Scanning Electron Microscopy,* Vol. III, p. 1069. Chicago: SEM Inc., 1982.
19. Mahajan, S., A. K. Chin, C. L. Zipfel, D. Brasen, B. H. Chin, R. T. Tung, and S. Nakahara. *Mater. Lett.* **2,** 184 (1984).
20. Chin, A. K., C. L.Zipfel, M. Geva, I. Camlibel, P. Skeath, and B. H. Chin. *Appl. Phys. Lett.* **45,** 37 (1984).
21. Fukuda, M., K. Wakita, and G. Iwane. *J. Appl. Phys.* **54,** 1246 (1983).
22. Yamakoshi, S., M. Abe, S. Komiya, and Y. Toyamer. Proc. Int. Electron. Dev. meeting, p. 122 (1979).
23. Temkin, H., A. Mahajan, M. A. DiGiuseppe, and A. G. Dentai. *Appl. Phys. Lett.* **40,** 562 (1982).
24. Ueda, O., H. Imai, A. Yamaguchi, S. Komiya, I. Umebu, and T. Kotani. *J. Appl. Phys.* **55,** 665 (1984).
25. Shaw, D. A., and P. R. Thornton. *Solid-State Electron.* **13,** 919 (1970).
26. Kressel, H., and H. Mierop. *J. Appl. Phys.* **38,** 5419 (1967).
27. E. J. Flynn (personal communication).
28. Mizuishi, K., M. Sawai, S. Todoroki, S. Tsuji, M. Hirao, and M. Nakamura. *IEEE J. Quantum Electron.* **QE-19,** 1294 (1983).
29. Gordon, E. I., F. R. Nash, and R. L. Hartman. *IEEE Electron Device Lett.* **ELD-4,** 465 (1983).
30. Ikagami, T., K. Takahei, M. Fukuda, and K. Kuroiwa. *Electron. Lett.* **19,** 282 (1983).
31. Hakki, B. W., P. E. Fraley, and T. F. Eltringham. *AT&T Tech. J.* **64,** 771 (1985).
32. Hartman, R. L., and R. W. Dixon. *Appl. Phys. Lett.* **26,** 239 (1975).
33. Joyce, W. B., K. Y. Liou, F. R. Nash, P. R. Bossard, and R. L. Hartman. *AT&T Tech. J.* **64,** 717 (1985).
34. Mizuishi, K., M. Hirao, S. Tsuji, H. Sato, and M. Nakamura. *Jpn. J. Appl. Phys. Part 1* **21,** 359 (1982).
35. Runge, P. K., and P. R. Trischitta. *J. Lightwave Technol.* **LT-2,** 744 (1984).
36. Fukuda, M. *Reliability and degradation of semiconductor lasers and LEDs.* Norwood, MA: Artech House, 1991.

INDEX