Semiconductors: Epitaxial Growth of Laser Diodes

As a process for fabricating semiconductor devices, epitaxial growth has long been important for some types of silicon integrated circuits and for gallium arsenide (GaAs) devices, such as microwave oscillators and light emitting diodes (LED's). But a specific form of epitaxial growth has been essential to the development of a cousin of the LED—the laser diode. These solid state lasers, made from thin films of the semiconductors GaAs and aluminum gallium arsenide [(AlGa)As], may one day be the coherent light sources in communications systems that use light waves as the carrier of information, much as radio waves or microwaves are used in today's telecommunications systems.

In the jargon of the crystal grower, epitaxy refers to the growth of a crystalline substance on a substrate of the same material or one that has a very similar crystal structure (heteroepitaxy). The substrate acts as a seed for the epitaxially growing layer so long as the mismatch between the sizes of the two crystal lattices is no more than a few percent.

Laser diodes made in this way have been commercially available for a few years. These lasers operate in a pulsed mode, emitting up to a few watts of power in short bursts 10^{-7} second in duration, and are used in instrumentation ranging from intrusion alarms to infrared illuminators for night vision devices. But it is a continuous wave (CW) version that emits less power (a few milliwatts) which may find use in optical communications. So far, the CW laser diode is not ready for service because of difficulties in making lasers with the 20-year lifetime needed for reliable systems. A CW laser may, however, be marketed soon for such applications as scanning documents, where it would compete with helium-neon gas lasers.

A number of factors make the GaAs–(AlGa)As system attractive for the laser diode application. The laser light propagates efficiently in the glass fibers planned to carry optical signals over long distances (*Science*, 12 October 1973, p. 151), because the wavelength emitted is in the range 8000 to 9000 angstroms. Moreover, since the laser diode is activated by an applied electrical signal, the modulation needed to impress signals onto the laser carrier can be applied directly to the laser without the need for an external device.

In the future, it may be possible to fabricate the optical equivalent of electrical integrated circuits (*Science*, 14 September 1973, p. 1032), in which laser light sources, waveguides to conduct the light, modulators, and photodetectors may all be constructed in thin film form on a single small substrate chip. Since each of these devices can be made separately from (AlGa)As, integrated optical systems may likewise be possible.

When a GaAs diode is biased, free electrons are drawn into an area near the p-n junction, where they recombine with free holes. Depending on the number and type of crystalline imperfections present (such as impurities, lattice vacancies, or dislocations), the recombination process may be radiative or nonradiative; that is, light may or may not be given off. In the former case, one has an LED. If an optical cavity is created by, for example, cleaving mirror-smooth faces on two ends of the diode, and if the current is increased sufficiently, laser action can begin. The LED then becomes a laser diode.

The first successful attempts to make lasers from GaAs diodes in the early 1960's required enormous currents (50,000 amperes per square centimeter) to activate lasing. Such large currents caused the lasers to overheat rapidly. The problem was that, in contrast to the situation with a glass laser rod or a gas laser tube, there was no physical structure to confine either the laser medium or the laser light. The high threshold current was needed to overcome losses that occurred because free electrons diffused away from the p-n junction without recombining and photons spread into adjacent regions of the diode.

In 1969, scientists at two laboratories announced the operation of heterojunction laser diodes made of GaAs and (AlGa)As. Heterojunction refers to the joining of different semiconductors. Groups headed by M. B. Panish and I. Hayashi of Bell Laboratories, Murray Hill, New Jersey, and Henry Kressel of RCA Laboratories, Princeton, New Jersey, constructed single heterojunction lasers. (Hayashi is now at the Nippon Electric Company in Japan.) A single heterojunction laser has a single p-type (AlGa)As layer covering the p-type side of a GaAs diode.

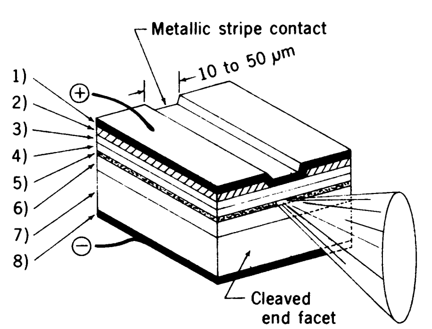

Fig. 1. Schematic diagram of an (AlGa)As double heterojunction laser diode. (1) and (8) Metallic electrical contacts. (2) Silicon dioxide insulating layer which confines the lasing to a 10-μm-wide stripe in order to better dissipate heat generated during operation of the laser, to increase the uniformity of the laser light, and to isolate the active region from free surfaces. (3) GaAs cap, used for ease in making electrical contact. (4) and (6) (AlGa)As confining layers. (5) GaAs layer from which laser light is emitted. (7) GaAs substrate. [Source: H. Kressel, RCA Laboratories]

Shortly thereafter, a third group at the A. F. Ioffe Physical-Technical Institute, Leningrad, under the direction of Zh. I. Alferov, as well as the Bell scientists, produced the double heterojunction laser diode, which has two (AlGa)As layers. The double heterojunction laser has since emerged as the only suitable candidate for the room-temperature CW operation needed for optical communications.

A typical double heterojunction laser (Fig. 1) is composed of five layers: an n-type GaAs substrate, an n-type (AlGa)As layer with a thickness of 3 μm, a GaAs layer 0.2 to 0.5 μm thick from which light is emitted, a p-type (AlGa)As layer 1.5 μm thick, and a final layer of p-type GaAs 1.5 μm thick. More complicated seven-layer structures have also been made. Photolithographic, cleaving, and sawing operations result in diodes that are about 300 to 500 μm long and 200 to 500 μm wide (of which the active region where lasing occurs is about 10 μm wide). Thus the appellation "pinhead laser" is quite appropriate.

The band gap of a semiconductor corresponds to the minimum energy required to create a free electron-hole pair. The difference in the band gaps at the heterojunction between the two p-type layers of GaAs and (AlGa)As acts as a potential barrier to prevent electrons from diffusing out from the p-n junction. Thus, the electrons are confined by the epitaxial layers of high band gap semiconductor [the (AlGa)As] surrounding the active laser material (the GaAs).

Light is also confined by the multilayer epitaxial structure, because a waveguide effect is provided by the combination of a medium with a low index of refraction [the high band gap layers of (AlGa)As] cladding a core with a high index of refraction (the low band gap GaAs layer).

The band gap of (AlGa)As can be altered by changing its composition, since the ternary alloy (AlGa)As is a continuous solid solution of the two compounds GaAs and aluminum arsenide (AlAs), with the general formula $Al_xGa_{1-x}As$, where x is the fraction of AlAs in the solution. Since the band gap of (AlGa)As increases monotonically from that of GaAs as the fraction x increases, the required epitaxial structure can be made by controlling that fraction in each layer of the laser diode.

Values of x are typically 0.2 to 0.4. Furthermore, (AlGa)As is unique in that the size of the basic unit in the (AlGa)As crystal lattice is nearly independent of the amount of aluminum, a factor that allows great flexibility in the process of fabrication.

A form of solution growth called liquid phase epitaxy (LPE) has turned out to provide a way to fabricate (AlGa)As heterojunction laser diodes. A molten gallium solution containing the requisite amounts of aluminum, arsenic, and appropriate doping impurities is slowly cooled, and excess (AlGa)As precipitates onto the substrate in a sufficiently orderly manner to build up crystalline layers.

By 1970, scientists had reduced the threshold current for lasing to the neighborhood of 1000 amperes per square centimeter. Since then, interest in making long-lived room-temperature CW laser diodes has become worldwide. The principal subject of this interest has been how to achieve long lifetimes reproducibly. The earliest lasers often failed because of catastrophic destruction of the cleaved mirror faces by the high photon densities in double heterojunction laser diodes (10 milliwatts of laser light corresponds to a power density of up to 1 megawatt per square centimeter). The catastrophic mode of failure has not been a big problem with CW lasers, however, because much less optical power is produced. The problem with CW lasers is that the intensity of the laser light tends to gradually decrease with time. When the light intensity is too low to sustain lasing, the laser shuts off.

Speaking of Science

Weisskopf on the Frontiers and Limits of Science

Victor Weisskopf is a physicist's physicist and something of a Renaissance man. Both of these talents were displayed in a recent talk in which he celebrated the scientific world view and yet raised a humanistic note of caution as to the limits of its applicability. Weisskopf predicted, in effect, that science may eventually explain all observable phenomena and yet remain incomplete, and he criticized exclusively rational views of human experience by comparing them to medieval religious excesses.

Weisskopf's career has spanned two continents, including graduate study in Europe under Wolfgang Pauli and wartime work on the U.S. atomic bomb project. He has been director of the European Center for Nuclear Research (CERN) near Geneva, Switzerland, one of two Americans to have held the post, and was on the faculty of the Massachusetts Institute of Technology from 1945 until his retirement last fall. By avocation he is a pianist of near-professional competence, and he has a broad interest in many aspects of literature, philosophy, and international relations. In retirement, he has continued to write and lecture. The address he gave to the American Academy of Arts and Sciences* reflects his growing concern that science has lost its human anchor, that both its critics and many of its practitioners have lost sight of the relationship in which science properly stands to human affairs.

Man, Weisskopf notes, has been curious about the world since the beginning of culture, but it was only about 500 years ago that a more focused and limited version of that curiosity began to result in what we now call science. According to Weisskopf, it was the renunciation of immediate contact with absolute truth and the investigation of particular phenomena that created a framework for understanding the natural world. The growing discovery that nature is indeed comprehensible led to a world view—what Weisskopf calls "the scientific myth of the universe"—that in its twentieth century version depends heavily on a description of matter derived from quantum mechanics. Thus matter is composed of atoms, and atomic phenomena—encompassing chemistry, biology, geology, and most naturally occurring phenomena on earth—are essentially electric in character. These phenomena exemplify for Weisskopf the internal frontier of science, in which the basic principles (the coulomb force and the quantum mechanical selection rules) are believed known but understanding is limited by the complexity of phenomena. The external frontier, on the other hand, includes subnuclear physics and astronomy, where the fundamental principles are still unknown.

Weisskopf is hugely optimistic that the frontiers of scientific knowledge will continue to recede. He believes that "it is reasonable to predict that man will eventually understand all of nature scientifically"—all observable phenomena. But he qualifies this sweeping claim by asserting that scientific insights will not cover every aspect of human experience. For example, "one can understand a sunset or the stars in the night sky in a scientific way, but there is something about experiencing these phenomena that lies outside science." Quoting Wittgenstein and the Swiss philosopher Fierz, Weisskopf goes on to develop the point that science does not always illuminate the most important aspects of human experience, that there are limits to the scientific world view. Indeed, he believes that overemphasis on the scientific way of thinking can be dangerous in that it leads to the neglect of other modes of experience, and he draws an analogy to religion in the Middle Ages. "The religious and the scientific emphases have each released creative forces, but as one-sided approaches, both have also produced serious abuses. In the Middle Ages, one can point to the Crusades and to the complete neglect of corporal suffering; in our time there has been overrationality with respect to definitions of the quality of life and political decisions and an excessive concern with the production of material goods."

It is to this overemphasis of the scientific approach and the corresponding neglect of modes of experience captured in art, music, and literature that Weisskopf assigns the blame for much of the current prejudice against science and technology and for the rise of such pseudosciences as astrology and ESP—"perverse" forms that are, he claims, the result of natural urges suppressed because "the scientific approach is considered the only 'serious' way of dealing with human experience."

It is an uncompromisingly humanist manifesto, and one that carries all the more force coming from a scientist of Weisskopf's stature. In an interview, Weisskopf said that he does not wish to be considered a spokesman for Carlos Castaneda, Theodore Roszak, or other antirationalist critics of science. But he did acknowledge that a connection exists between the narrowness and specialization of graduate training and research practice in the sciences today and the problems they, and he, have addressed. "Specialization has made of the rational method a profession and not an avocation," was how he put it. Many "professional" scientists may wish to disagree with Weisskopf's diagnosis, but they might do well not to dismiss it out of hand.—ALLEN L. HAMMOND

*V. F. Weisskopf, *Bull. Am. Acad. Arts Sci.* **27**, 15 (March 1975).

Depending on the care taken during fabrication of the laser, devices lasting more than 1½ years can be made. Moreover, R. L. Hartman and R. W. Dixon of Bell Laboratories have estimated from lifetime studies at higher than room temperature (where the rate of degradation is increased) that 100,000 hours of laser operation should be obtainable at room temperature. Their conclusion was based on an apparent relationship between the observed lifetime and the temperature of the test.

Several effects may cause steady degradation of lasers with time. Many of them seem to be related to the initial presence and subsequent motion of crystal imperfections of various types. A class of defects known as dark line defects, because they show up as dark areas with a decreased light output on the face of the laser, can shut down the operation of the laser diode in only a few hours either by absorbing the laser light or by introducing centers for nonradiative recombination. Dark line defects are thought to be created by the migration of crystal lattice vacancies to and subsequent interaction with dislocations. There is some evidence that the mobility of such vacancies may be vastly greater during the time when electrons and holes are recombining than when the diode is not operating. These imperfections could then easily migrate under the influence of various forces, such as residual strains from the preparation of the diode.

The preparation of long-lived lasers in part involves avoiding the conditions that give rise to dark line defects. Hartman at Bell Laboratories notes that anything that can be done to prevent the incorporation of strain or imperfections into the diode during its fabrication minimizes the effect of dark line defects. Kressel and Ivan Ladany at RCA observe that the migrating imperfections tend to come from such regions of the diodes as free surfaces. Thus, they report that if the lasing active layer of the diode can be isolated from such surfaces, no degradation is observed over periods of several thousand hours. Nonetheless, there appear to be other, more subtle degradation mechanisms whose nature is not yet clearly known.

The CW double heterojunction laser will likely first be used, either in optical communications or elsewhere, for applications that require only a discrete laser that is small, inexpensive, and easily modulated. However, for a truly integrated optical device, the laser should be part of the thin film structure. For this to occur, some way must be found to avoid the cleaved end mirrors that are needed to permit the intensity of the coherent laser light to build up. A device that seems to solve this problem is the distributed feedback laser.

A common procedure for making a distributed feedback laser is to make a corrugated pattern on the top surface of the active layer of the double heterojunction structure. The resulting periodic variation in the thickness of the GaAs performs the same function as the end mirrors. The periodicity of this grating-like structure also determines the wavelength of the laser light. For use in thin film optics, a laser could be made continuous with an optical waveguide or another element of an optical circuit simply by adding a grating to the section designated as the laser.

Recently several groups have reported obtaining distributed feedback operation in GaAs–(AlGa)As heterojunction laser diodes, though only with pulsed mode lasers. D. R. Scifres, R. D. Burnham, and W. Streifer at the Xerox Palo Alto Research Center, Palo Alto, California, have produced the grating needed for making a distributed feedback laser by using standard holographic techniques to make a mask on the surface of a partially completed single heterojunction laser diode. Grooves are etched into the surface of the masked layer by bombarding it with ions, a process known as ion milling. After the grating is complete, LPE growth of the remaining layers of the diode is continued. In accordance with theory, experiments showed that the lasing wavelength could be controlled by varying the spacing of the grooves, a feature not obtainable in the normal laser structure where the wavelength can be difficult to control. The initial experiments were all carried out at 77°K.

M. Nakamura and his associates at Hitachi, Ltd., Tokyo, in cooperation with Amnon Yariv and his co-workers at the California Institute of Technology, Pasadena, have fabricated distributed feedback double heterojunction lasers that operated in the temperature range 80° to 100°K. Room temperature operation was hampered in part by surface damage incurred during the ion milling process, which is thought to introduce nonradiative recombination centers, whose activity tends to increase with temperature.

This problem may have been solved, however, as H. C. Casey, S. Somekh, and M. Ilegems of Bell Laboratories have recently obtained room-temperature operation of a heterojunction laser diode with distributed feedback.

Also at Bell Laboratories, F. K. Reinhart, R. A. Logan, and their colleagues have grown integrated double heterojunction lasers and passive waveguides by LPE growth. One attempt at such a structure involved gradually eliminating the GaAs active layer in the double heterojunction structure. Such tapered structures had previously been demonstrated by other researchers to efficiently couple light from one waveguide into another, in this case from the GaAs into an adjoining layer of (AlGa)As. A grating etched into the passive waveguide portion of this integrated structure was shown by Reinhart, Logan, and C. V. Shank to support laser action at room temperature.

Although the LPE process results in high-quality laser structures, some scientists feel that a relatively new growth process known as molecular beam epitaxy (MBE) may turn out to be superior. The MBE process is a highly sophisticated vacuum evaporation procedure carried out at 10^{-9} torr or less. The advantage of the MBE technique is that very fine control of the thickness and smoothness of the epitaxial layers can be obtained. It may also be possible to exert control over lateral dimensions through the use of masks, as in integrated circuit technology.

A. Y. Cho and Casey at Bell Laboratories have used MBE to make double heterojunction laser diodes. According to Casey, the diodes produced so far have exhibited somewhat poorer performance than the devices made by the LPE process, perhaps because of the introduction of nonradiative recombination centers during the growth process.

The (AlGa)As alloy system in combination with the LPE process has already been successfully exploited to make a number of optoelectronic devices, including laser diodes, LED's, solar cells, and photocathodes. However, the wavelengths that can be produced or detected with this material are limited by the band gaps of GaAs and AlAs. For optical communications, a solid state laser diode emitting 1.06-μm radiation would be even better than (AlGa)As emitting at 0.85 μm.

C. J. Neuse and G. H. Olsen of RCA Laboratories have grown indium gallium arsenide–idium gallium phosphide heterojunction laser diodes that lase at about 1 μm by another epitaxial process—vapor phase epitaxy. Although the performance of these lasers does not equal that of the highly developed (AlGa)As devices, the conclusion seems clear. The various epitaxial techniques, as applied to ternary or even quaternary semiconductors, hold the promise of permitting the tailoring of a wide variety of optical and electronic components simply by adjusting the composition of the various epitaxial layers.

—ARTHUR L. ROBINSON

Additional Reading

1. H. Kressel and H. Nelson, in *Physics of Thin Films*, G. Hass, M. Francombe, R. W. Hoffman, Eds. (Academic Press, New York, 1973), vol. 7, pp. 115–256.
2. M. B. Panish, *IEEE (Inst. Electr. Electron. Eng.) Trans. Microwave Theory Tech.* **MTT-23**, 20 (1975).