From Transistors to Lasers and Light-Emitting Diodes

Nick Holonyak Jr.

Abstract

This article is based on the 2004 Von Hippel Award address by Nick Holonyak Jr. (University of Illinois at Urbana-Champaign). Holonyak received the award for "his many contributions to research and development in the field of semiconductors, not least for the first development of semiconductor lasers in the useful visible portion of the optical spectrum." The talk was presented on Holonyak's behalf by Russell Dupuis on December 1, 2004, at the Materials Research Society Fall Meeting in Boston.

With the discovery of the transistor by Bardeen and Brattain in 1947, and as a consequence of carrier injection and collection, the hole indeed became equal to the electron. The semiconductor took on new importance, as did the study of electron–hole recombination, first in the transistor materials Ge and Si, and then in III–V crystals (e.g., GaAs and GaP). Beyond Si and its indirect-gap and heterojunction limitations, the direct-gap III–V materials, particularly III–V alloys, made possible lasers and light-emitting diodes (LEDs)—and thus optoelectronics.

The direct-gap III–V alloy LED after four decades of development exceeds in performance the incandescent lamp (as well as other forms of lamps) in much of the visible range. Beyond growing display applications, it has put conventional lighting under long-range threat with a semiconductor lamp—an "ultimate lamp" that promises unusual performance and energy savings. In principle, the LED or laser, basically a *p–n* junction, is an ultimate lamp that cannot be exceeded.

Keywords: *lasers, LEDs, light-emitting diodes, semiconductor lamps.*

Introduction

Following John Bardeen (Figure 1) and Walter Brattain's 1947 discovery of the transistor[1,2] and their identification of carrier injection,[3–5] making the hole (positive, *p*) equal in performance to the electron (negative, *n*), the semiconductor took on new importance. With the transistor—a new idea, a new principle, a new device, a new name—a new electronics emerged that could not be based on or matched by the capabilities of the vacuum tube, even if we could or would be willing to cover the earth with tubes.

With the emerging dominance of the semiconductor, the study of electron–hole recombination also attracted intense interest, first in the transistor materials Ge and Si, and then in III–V crystals (e.g., GaAs and GaP). Beyond Si and its indirect-gap and heterojunction limitations, the direct-gap III–V materials, particularly III–V alloys, made possible lasers and light-emitting diodes (LEDs)—and thus optoelectronics.

The first practical visible-spectrum LED, not to mention the first III–V alloy device (a first laser), began in the early 1960s with the direct-gap III–V alloy $GaAs_{1-x}P_x$, which is also the beginning of III–V epitaxy. Of special importance, $GaAs_{1-x}P_x$ established the viability of III–V alloys and, with its energy-gap and wavelength "tunability," set the direction for the construction of heterojunctions.

The progression over four decades—from the direct-gap alloy $GaAs_{1-x}P_x$, the prototype; to later $Al_xGa_{1-x}As$ and, next, $In_{1-x}Ga_xP$; to then the shorter-wavelength, direct-gap alloy $In(Al_xGa_{1-x})P$ and, more recently, $In(Al_xGa_{1-x})N$—has led to heterojunction LEDs that cover all of the visible spectrum. In addition, various III–V alloys, with increasingly sophisticated crystal growth and processing technology, have led to high-performance quantum-well lasers and quantum-well LEDs over a broad range of wavelengths and power.

The direct-gap III–V alloy LED after four decades of development exceeds in performance the incandescent lamp (as well as other forms of lamps) in much of the visible range. Beyond growing display applications, it has put conventional lighting under long-range threat with a semiconductor lamp—an "ultimate lamp" that promises unusual performance and energy savings. By "ultimate lamp," I mean a lamp that cannot be exceeded in efficiency of converting electrical energy to optical energy. In principle, the LED or laser, basically a *p–n* junction (or heterojunction),[6,7] is an ultimate lamp (see sidebar article on p. 515).

Figure 2 is one of Bardeen's seminar transistor diagrams, the original transistor. At low voltage (low impedance) at the so-called emitter, a cloud of holes (circles) is injected into the crystal and is compensated by electrons (–) that shift into proximity with the holes' positive charge to preserve charge neutrality. The holes travel by diffusion, like an expanding cloud of smoke, to the nearby collector and are extracted at a higher voltage (higher impedance) than injected at the emitter. This is the basis for gain, or the operation of an amplifier: current transfer from a low-impedance input to a high-impedance output via a "transfer resistor," or transistor. But as it happens, not all of the holes are successful in getting from the emitter (input) to the collector (output). Some are annihilated by recombining with electrons and might, in the "right" circumstances, deliver light ($h\nu$). Could this, assuming it occurs, be useful in itself, and yield an LED?

The transistor introduced a totally new idea! Who would have believed that it could be something as far-fetched as minority carrier injection (hole injection into an *n*-type crystal, Figure 2) at low impedance at one electrode, which Bardeen called the "emitter" (in a later comment to Holonyak[5]), and carrier collection at high impedance at the other electrode (a reverse-biased remote "collector") thus providing gain (amplification)? At the time (1947), semiconductors such as Ge and Si, which were well known after World War II radar use, were thought to be direct-gap (momentum $\mathbf{k}_e = \mathbf{k}_h$), not indirect-gap, with the capability of possessing long electron–hole lifetimes, thus making possible carrier injection and collection over a significant distance. In Figure 2, we have added to Bardeen's transistor diagram the possibility that

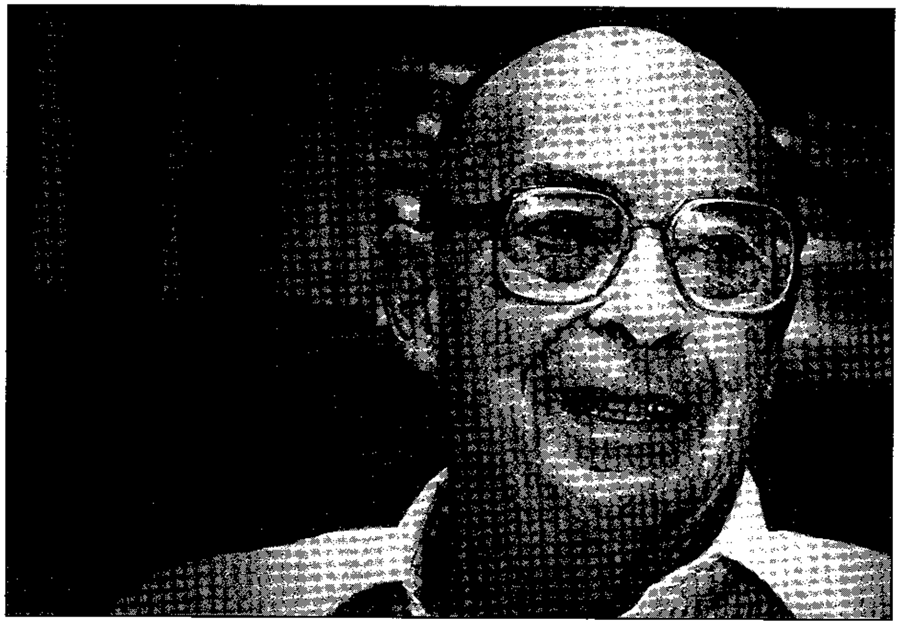


Figure 1. John Bardeen, at the age of 80; Nobel Laureate in physics for the transistor (1956) and for superconductivity (1972).

electron–hole recombination generates a photon, $h\nu$.

Crystal Epitaxy and III–V Alloys

As transistor development proceeded (1950–1960), crystal growth and materials processing grew in scale and sophistication. The transistor engendered—actually drove—a new era of intensive materials study and development and led to new processing methods such as vapor-phase epitaxy (VPE), which of course could be extended beyond just Ge and Si transistors and switches.

The III–V materials, which were interesting to some of us in the late 1950s and early 1960s (1959–1962) for tunnel diodes[8] and parametric diodes, in which capacitance varies with voltage, not to mention light-emitters, were technologically more mysterious than the elemental materials Ge and Si—and more intractable. Despite considerable skepticism in some quarters, some of us nevertheless felt that III–V materials offered important opportunities and should be explored also. For example, a III–V compound such as GaAs made possible higher-voltage tunnel diodes than Ge or Si (Figure 3),[8] but then introduced other problems, such as reliability, and the need for further studies. Fortunately, some agencies—the Air Force and Army, among others—agreed and supported exploratory III–V materials and device work, and by 1960 we could demonstrate III–V alloy and heterojunction work employing VPE that was the basis for an early publication[9] and was sufficiently basic for the issuance of a patent (Figure 4).

Figure 4 shows a sketch of a simple closed-tube halide-transport III–V epitaxial crystal growth process, which, as shown, was used as early as 1960 to grow GaAs on GaAs, GaAs on Ge, GaP on GaP, GaP on GaAs, and $GaAs_{1-y}P_y$ on $GaAs_{1-x}P_x$. The last is of special interest because it shows that co-transport of GaAs and GaP could be used to grow the alloy $GaAs_{1-x}P_x$,[9] either as a uniform crystal or as a heterojunction ($GaAs_{1-x}P_x$ on $GaAs_{1-y}P_y$).

Until the III–V VPE crystal growth of Figure 4, some expert opinions held that a III–V alloy such as GaAsP could be synthesized by simply diffusing phosphorus into gallium arsenide, which, of course, would have required an inordinate amount of time (years, decades, longer) and never was heard of again. Figure 5 is of special interest because it shows the epitaxial layer growth (1961) of a thin layer of GaP on GaAs with a thin layer of $GaAs_{1-x}P_x$ between the two binary layers. It does not take much imagination to see that this will ultimately become a variety of heterostructure, eventually including layers thin enough to exhibit quantum size effects and to serve as quantum wells in heterostructure devices.

III–V Alloys and Light-Emitting Diodes

The transistor and carrier injection, with a current (Figure 2), made it apparent (from the time of Bardeen and Brattain's

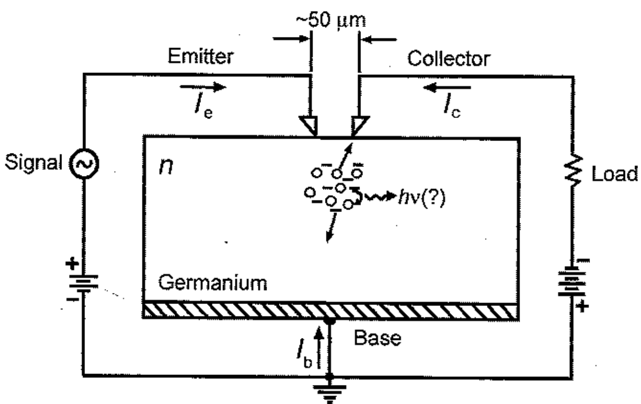


Figure 2. Bardeen seminar diagram of the original 1947 transistor, with an electron–hole recombination event added, suggesting the possibility of radiative recombination ($h\nu$). I_e is emitter current, I_c is collector current, and I_b is base current.

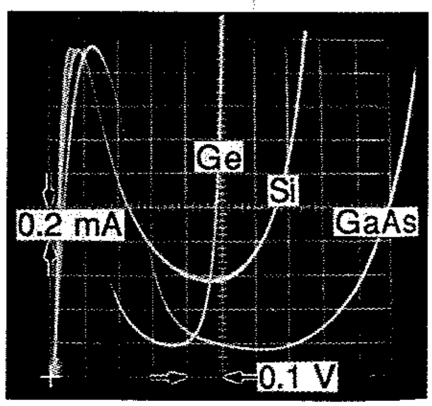


Figure 3. Ge-Si-GaAs tunnel diode current–voltage characteristics, showing the progression to higher voltage with larger energy gaps.

first transistor)[1–3] why a p–n junction could emit light. Moreover, the III–V compounds possessed energy-band properties and energy gaps (wavelengths) that made it possible to seek and make LEDs and even semiconductor lasers. In fact, at least two of us (myself at GE and Rediker at Lincoln Laboratory–MIT) felt that the energy-band property that made possible a laser (1962)—that is, a direct gap ($k_e = k_h$) and conservation of electron and hole momentum in carrier recombination—was what was also required for the best LEDs. Thus, visible-red direct-gap GaAsP was a better LED candidate than, say, shorter-wavelength but indirect-gap GaP, with its offset electron and hole momentum ($k_e \neq k_h$).

In a (not always friendly) debate that started around 1962 and continued for decades, it became clearer and clearer that lasers and LEDs required III–V alloys and not just the usual binary crystals such as GaAs or GaP, which, in fact, have now been relegated mainly to use as substrates. The prototype III–V alloy system proved to be GaAsP, which indeed operated as one of the first semiconductor lasers (1962) and was introduced also as the first practical LED.

The first GaAsP laser is shown in Figure 6.[10] The diode was made from a small halide VPE n-type GaAsP "ingot"[9] (as in Figure 6a) that was first cut into wafers, then polished, etched, Zn-diffused, processed into diodes with Fabry–Perot facets, and assembled as shown. In Figure 6b, the same GaAsP laser diode is shown tilted to reveal one of the polished Fabry–Perot facets of the laser resonator. Figure 7 shows another GaAsP laser similar to that of Figure 6, operating in the visible red with its output beam (from a polished Fabry–Perot facet) photographed (1962) directly on colored film without the benefit of any kind of infrared converter or detector. This is the first semiconductor laser photographed by its own light, in this case, the red light of a direct-gap III–V alloy—a homemade GaAsP crystal.[9,10] The III–V alloys had arrived.

It is worth mentioning that, generically, the GaAsP diode laser is an LED; that is, it belongs to the LED family. It possesses, however, a major difference: a cavity in the form of a Fabry–Perot resonator that aids in the recombination and feedback process and encourages stimulated electron–hole recombination. It solves a problem: It "points" the stimulated recombination

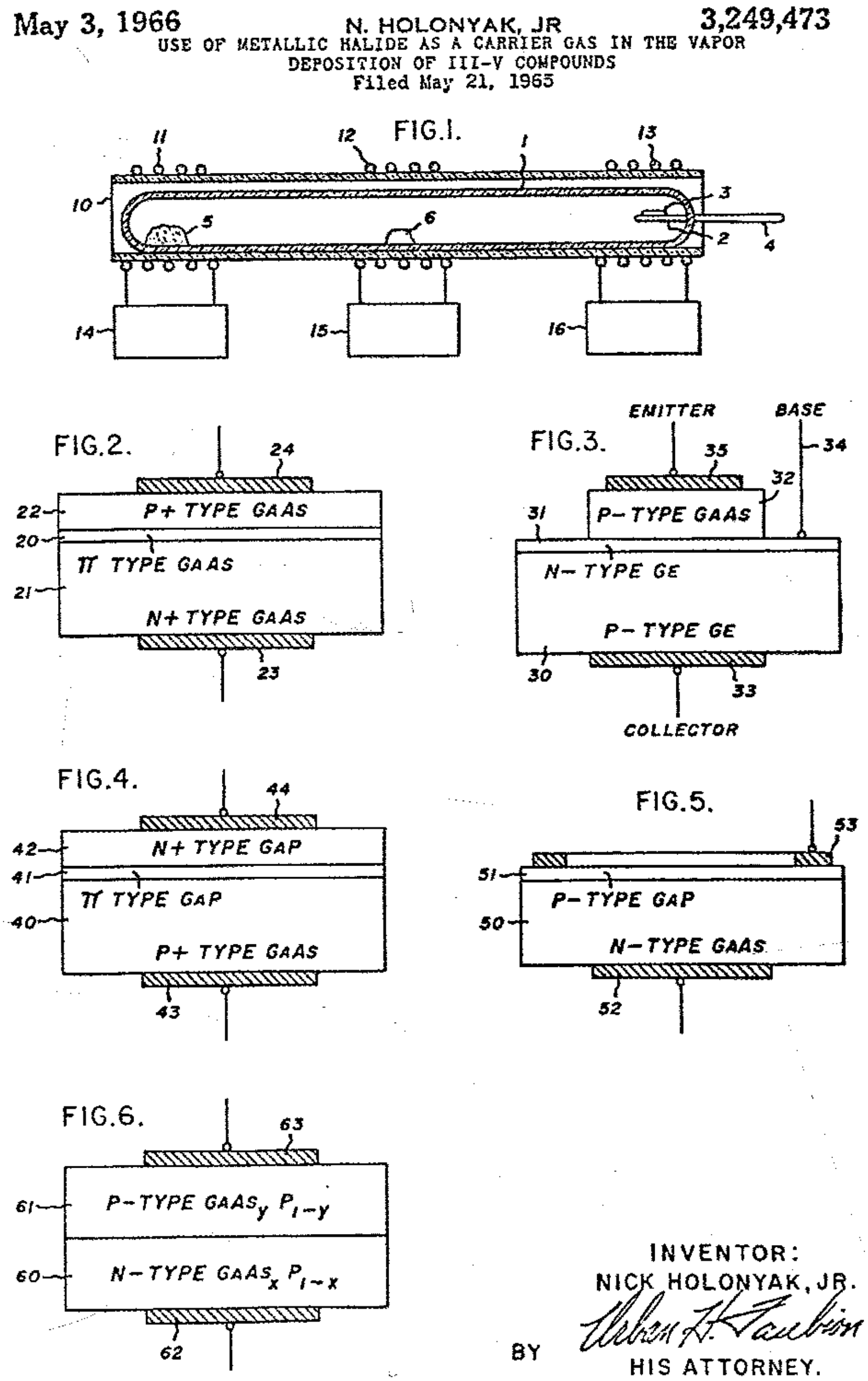


Figure 4. Page from author's patent for closed-tube halide VPE crystal growth (1960–1962) of III–V crystals, including GaAsP and various heterojunctions.

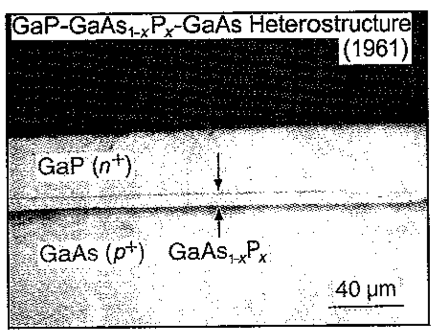


Figure 5. Micrograph of a closed-tube halide VPE GaP-GaAsP-GaAs heterostructure (1961).

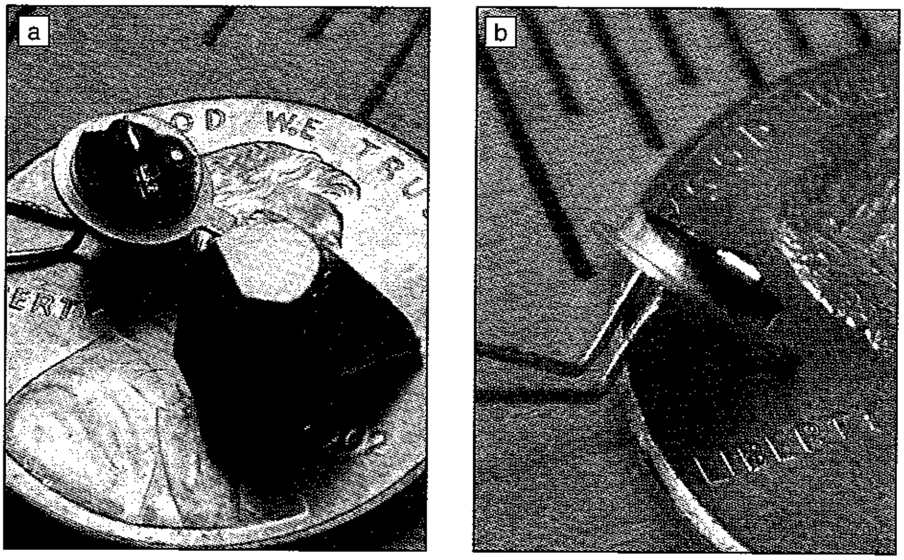


Figure 6. (a) Typical 1960–1962 halide VPE GaAsP crystal shown with the first GaAsP visible-spectrum laser (1962). (b) Same laser as in (a), tilted to show one of the polished Fabry–Perot facets.

radiation out of a Fabry–Perot facet of the crystal (the diode laser) and yields a large signal instead of the weak signal of spontaneous recombination that partly escapes but mainly reflects around in the crystal and is absorbed, unless special device geometries are devised.

After the 1962 IEEE Solid-State Device Research Conference in New Hampshire, some of us understood the need for a cavity, and a number of others mistakenly believed that for a laser, even a diode laser, discrete levels were required, which of course led to some bizarre notions. My thought was to take a visible-spectrum (red) GaAsP diode ($\mathbf{k}_e = \mathbf{k}_h$) and operate it in an external cavity. I knew the importance of the cavity, and I wanted to see what I was doing. In addition, I knew how to make visible-spectrum *p–n* junctions. Why not work in the visible (GaAsP), not the infrared (GaAs)?

When I discussed my ideas for a visible-spectrum diode laser in August 1962 with my Schenectady colleague R.N. Hall, with whom I shared an Air Force contract, he informed me that he was already at work making a GaAs diode laser in which he would use the crystal itself as the infrared photon resonator, and he intended "merely" to polish the crystal sides into the form of a Fabry–Perot resonator. Hall's thinking was clever. I immediately suggested that he cleave the crystal to form Fabry–Perot facets, but he preferred to polish it. As it happened, I tried unsuccessfully to cleave my large-grain polycrystalline GaAsP, but quickly switched to polishing when Hall's "boss," L. Apker, called me (Schenectady to Syracuse) one early fall day in 1962 to tell me Hall's GaAs diodes were operating as lasers. Hall could see the laser diffraction pattern with a snooper scope—basically, a night-vision device that sees infrared photons—the pattern I expected to see directly in the visible red (GaAsP).

I quickly devised the cavity polishing scheme sketched in Figure 8. Hall preferred his polishing method, and I my own, while others[11] apparently struggled to realize a cavity. After diffusing Zn (an acceptor) 10–20 μm deep into an *n*-type GaAsP wafer, placing black wax half-round masking threads parallel across the wafer surface, and etching deep slots to form long mesa junctions, we could saw-cut through the wafers at right angles to the long mesas and mount the resulting *p–n* junction strips as shown in Figure 8, and then proceed to polish Fabry–Perot facets (Figure 6b). We then immediately assembled successful red GaAsP diode lasers and proved at once the value of III–V alloys.

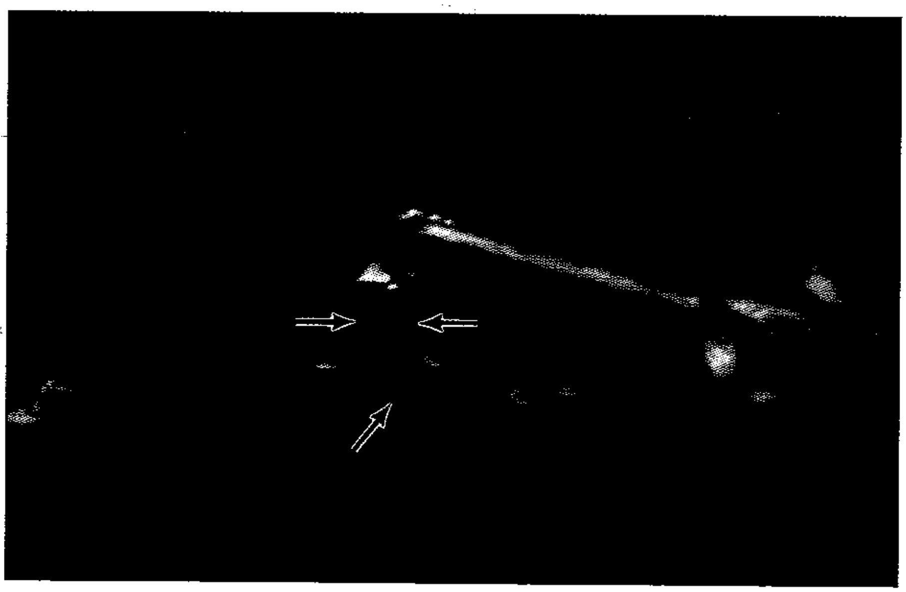


Figure 7. First visible-spectrum diode laser (1962, GaAsP III–V alloy), photographed by its own light. (See Figure 6 for scale; the same diode platform was used in both figures.) Horizontal arrows delineate the diode; the other arrow points out a diffraction spot.

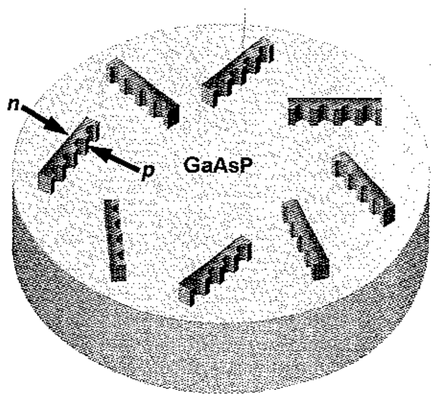


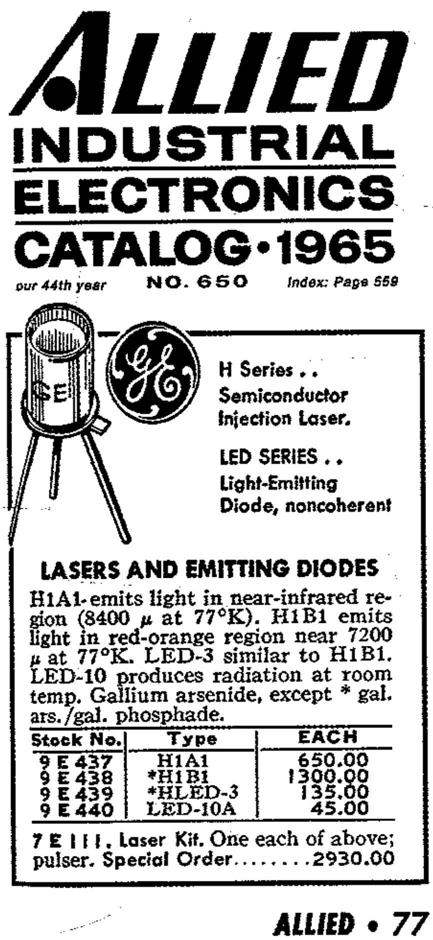
Figure 8. Mounted p–n *junction diode strips ready for polishing. The Fabry–Perot cavity polishing method was used at GE (Syracuse) for the first GaAsP lasers (1962).*

The same GaAsP *p–n* junction material that was used for the lasers of Figures 6a, 6b, and 7, except shifted in composition to a shorter wavelength (more P), could be packaged in small pigtail diode packages (the transparent glass type used for small Si diodes) and used as LEDs. At the first-ever semiconductor laser conference in Schenectady in November 1962, which was held by GE for invited Department of Defense guests, glass-packaged GaAsP LEDs were given to some of the participants from the DOD. The first practical LEDs and the first ones offered for sale (by GE, 1962) were made from the III–V alloy $GaAs_{1-x}P_x$ and sold by GE (and a little later at a 50% reduced price) through the Allied radio catalog (Figure 9).

In fact, in an interview with Harlan Manchester of *Reader's Digest* published in February 1963, we even claimed that visible-red GaAsP LEDs would exceed lasers in importance, and that eventually a white LED would be possible. As it happens, to this day, LED sales exceed semiconductor laser sales and, as the LED now becomes a full-fledged lamp, LED use and sales promise to become enormous—as do the energy savings in comparison with less efficient forms of conventional lighting (for more on addressing grand energy challenges through advanced materials, see the article by Dresselhaus in this issue).

The GaAsP laser and LED proved at once, by their device performance, that III–V alloys, although in a sense stochastic, were sufficiently "smooth," uniform, and defect-free so as to be viable, useful systems.

Now, the III–V alloy, being an alloy, could of course be used in heterojunctions.[9] Ultimately, heterojunctions would displace homojunctions and prove to be indispensable also for quantum-well heterostructures and superlattices. Without III–V alloys, the optoelectronics known today could not exist. It is the alloy GaAsP that got this all started,[9,10] that pointed to the so-called alloy road, a term used later by Egon Loebner (Hewlett Packard → Agilent → LumiLeds). It is interesting also that the simple form of GaAsP red LED that we introduced in 1962 is still being manufactured more than 40 years later.


ALLIED INDUSTRIAL ELECTRONICS CATALOG • 1965

our 44th year NO. 650 Index: Page 559

H Series .. Semiconductor Injection Laser.

LED SERIES .. Light-Emitting Diode, noncoherent

LASERS AND EMITTING DIODES

H1A1 emits light in near-infrared region (8400 μ at 77°K). H1B1 emits light in red-orange region near 7200 μ at 77°K. LED-3 similar to H1B1. LED-10 produces radiation at room temp. Gallium arsenide, except * gal. ars./gal. phosphade.

Stock No.	Type	EACH
9 E 437	H1A1	650.00
9 E 438	*H1B1	1300.00
9 E 439	*HLED-3	135.00
9 E 440	LED-10A	45.00

7 E 111. Laser Kit. One each of above; pulser. Special Order........2930.00

ALLIED • 77

Figure 9. Allied Industrial Electronics Catalog listing (1965) of GE's 1962 lasers and light-emitting diodes, with the price reduced by 50% from 1962.

III–V Alloys and High-Brightness LEDs

The III–V semiconductor laser and LED research of the early 1960s, which occurred primarily in industrial laboratories (just as did early transistor research), could equally well have been done in university laboratories. Massive laboratories were not yet needed for semiconductor research, and III–V materials and device research could move into university laboratories, as indeed happened with III–V alloy semiconductor research. Based on an invitation from John Bardeen, we moved our GaAsP research, as well as the study of other III–V alloys, from Syracuse (GE) to Urbana (University of Illinois) in 1963. In the photograph in Figure 10, which dates from the early 1970s, Bardeen and I are looking at red–orange–yellow–green (ROYG) LEDs in the higher-energy InGaP alloy system. In 1970, we were able to show laser operation in this III–V alloy and then began a long period of studying and extending this system to shorter wavelengths (red → green).

With the continued development of VPE crystal growth processes, and then the demonstration of AlGaAs-GaAs heterojunction lasers grown by metalorganic chemical vapor deposition (MOCVD, a VPE process developed by Dupuis et al. in 1977),[12] it was clear that Al-Ga substitution in the $In(Al_xGa_{1-x})P$ system, which is an extension of InGaP, would yield high-performance ROYG heterojunction LEDs. It is worth mentioning that the red–orange lasers in digital video disc recording machines employ the $In(Al_xGa_{1-x})P$ alloy and are mainly grown by Dupuis-style MOCVD.[12]

Figure 10. Nick Holonyak Jr. (left) and John Bardeen in the Electrical Engineering Research Laboratory at the University of Illinois at Urbana-Champaign, examining the prototype red–orange–yellow–green InGaP light-emitting diodes that later, with Al-Ga substitution, became the In(AlGa)P LEDs now approaching the performance of an "ultimate lamp." Photograph dates from around 1972.

For the best LEDs, it is not sufficient merely to have a direct-gap visible-spectrum III–V alloy. It is important also to make the LED in the form of a double heterojunction (DH)—that is, a wider-bandgap hole (p) and electron (n) emitter on either side of a narrower-gap active region, a concept championed very early by Alferov for lasers.[13] This is advantageous for improved carrier injection and in allowing the escape of photons from the active region (less absorption), and of course is part of the design of high-brightness ROYG In(AlGa)P LEDs. This does not take care of all of the problems of absorption, because photons, even in a very efficient quaternary DH, tend to be contained in the dense medium of the crystal (the problem of crystal and free-space mismatch). Hence, the absorbing GaAs substrate on which In(AlGa)P DHs are grown by VPE must be removed and replaced with, for example, a transparent "platform" such as GaP.

Kish et al.,[14] then of Craford's group at Hewlett-Packard (both alumni of Urbana's III–V semiconductor research) accomplished this task and introduced a new family of high-brightness In(AlGa)P LEDs.[14] The improvement in LED performance, depending upon the VPE crystal and wavelength, can be 200% or more. In addition, if the LED crystal is properly shaped (Krames et al.)[15] to reduce multiple photon reflections in the crystal, the LED performance can be further enhanced, for example, reaching an external quantum efficiency as high as 50–60%. The LED, based on this level of performance, is indeed at the stage of becoming an "ultimate lamp."[6,7]

When an already successful product is improved on this scale, the effect is not just evolutionary, it is revolutionary. Figure 11 (after Craford and co-workers) shows where III–V alloy LEDs fit in performance in comparison with other well-known light sources. Besides ROYG In(AlGa)P LEDs, Figure 11 shows also where blue and green In(AlGa)N LEDs (Nichia), which also are part of the III–V alloy family (and are grown mainly by MOCVD),[12] fit in the same comparison. Figure 11 shows distinctly that visible-spectrum III–V alloy LEDs now exceed the standard incandescent lamp in performance (lumens per watt), which is itself a sufficient achievement. With further improvement, III–V alloy LEDs threaten to outperform a number of other light sources.

Conclusions and Future Outlook

At this point, we can look back over a considerable period (more than 40 years, Figure 11) and see what has happened to the LED, which, of course, is itself a consequence of transistor-era developments. The $GaAs_{1-x}P_x$ laser gave an unambiguous start to the first practical LED, which in Figure 11 is the point the far lower left (arrow, 1962). This figure, which was supplied by my former student and colleague of many years, George Craford (LumiLeds), shows how LEDs have evolved in performance over the years. It is not necessary to describe all the steps and plateaus in LED performance in this figure except to state that the figure begins on the lower left with the direct-gap III–V alloy $GaAs_{1-x}P_x$, the prototype red-spectrum alloy; then at the far right (~2000), it climbs to well above the performance of conventional incandescent lamps, because of the direct-gap alloy $In_{0.6}(Al_xGa_{1-x})_{0.5}P$, grown lattice-matched to GaAs. Direct-gap III–V alloys, in fact, have prevailed as LEDs. They have eliminated everything else and still have room for considerable improvement (see Figure 11).

It is clear what this means: On a logarithmic scale of decades (10, 20, 50, 100 years), the semiconductor will become the "ultimate lamp." This is possible in theory—it is what p–n junction theory allows.[6] We know this from ideas I began to assemble over 40 years ago, to make sure that we were not making a mistake, and then taught for many years to graduate students before much later publication.[6] It is possible now, based on what semiconductor technology will permit, to speak of an "ultimate lamp."[7] Every form of display will eventually be possible, as well as every form of lamp. In short, more than the transistor has come from the transistor and from the uniqueness of the semiconductor, which is a universal substance employing and equating in importance the electron, hole, and photon.

Wherever my teacher John Bardeen is, I am sure he is pleased with what has happened. Perhaps we should take another look at Figure 11 and realize how much time and effort are needed to research, to learn, to build, and to accomplish. Obviously, not everything should be done and measured with a short-term perspective and the expectation of immediate gain or yield. That is not how we got here, to the high-brightness LED lamp and the performance and energy savings it promises.

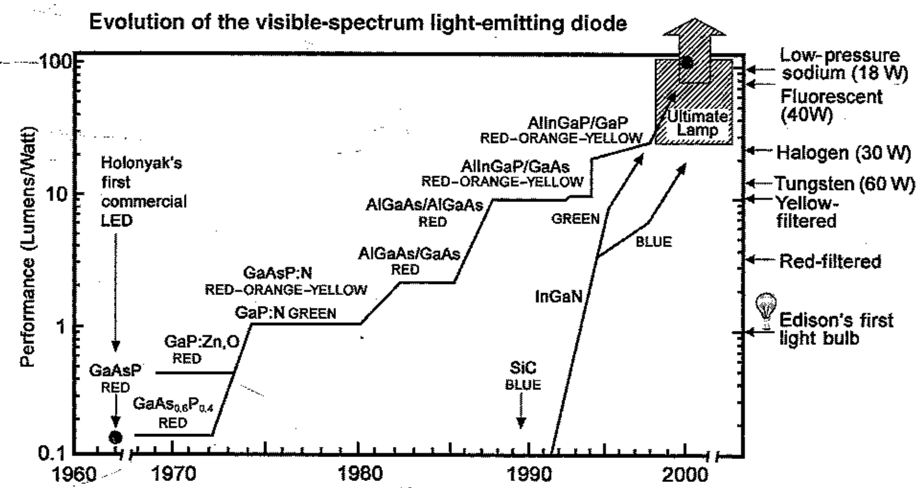


Figure 11. Diagram showing the evolution over 40 years in the performance of LEDs, from the time of the first GaAsP laser and red light-emitting diode in 1962, to the III–V alloys now prevailing in the approach to an "ultimate lamp."

Acknowledgments

In closing, I want to mention that our early work, the first successful III–V alloy laser and LED work, was supported by the Air Force, and more recently, in the era of quantum wells, by the Army Research Office, the National Science Foundation, and DARPA. I want to thank my students and colleagues, and other colleagues in many other places, for all of their contributions, and for all of their efforts in making the LED into the high-brightness LED lamp. For help with the figures and manuscript, I am grateful to G. Walter, B.L. Payne, and R. Chan. I don't think any of us can repay our debt to John Bardeen. I am especially grateful to John Bardeen and to Frederick Seitz for what they brought to Urbana and gave us, and to Frederick Seitz for nominating me for the 2004 MRS Von Hippel Award. I wish to thank all of the MRS people for their generosity.

References

1. J. Bardeen and W.H. Brattain, *Phys. Rev.* **74** (1948) p. 230.
2. J. Bardeen and W.H. Brattain, U.S. Patent 2,524,035 (October 3, 1950; filed June 17, 1948).
3. J. Bardeen, *Optoelectron. Dev. Technol.* **2** (1987) p. 124.
4. J. Bardeen, "The Early Days of the Transistor," Abstracts, New Materials Conference, Osaka, 1990. Reference supplied by M. Kikuchi, November 5, 1999.
5. N. Holonyak Jr., *Phys. Today* **45** (4) (1992) p. 36.
6. N. Holonyak Jr., *Am. J. Phys.* **68** (2000) p. 864.
7. M.G. Craford, N. Holonyak Jr., and F.A. Kish Jr., *Sci. Am.* **284** (2) (2001) p. 62.
8. N. Holonyak Jr. and I.A. Lesk, *Proc. IRE* **48** (1960) p. 1405.
9. N. Holonyak Jr., D.C. Jillson, and S.F. Bevacqua, *Metallurgy of Semiconductor Materials*, Vol. 15 (Interscience, New York, 1962).
10. N. Holonyak Jr. and S.F. Bevacqua, *Appl. Phys. Lett.* **1** (1962) p. 82.
11. H. Zeiger, *Optics and Photonics News* **15** (2004) p. 14.
12. R.D. Dupuis, *IEEE J. Selected Topics in Quantum Electronics* **6** (2000) p. 1040.
13. Zh.I. Alferov, V.M. Andreev, D.Z. Garbuzov, Y.U. Zhilyaev, E.P. Morozov, E.L. Portnoy, and V.G. Trofim, *Fiz. Tekh. Poluprovodn.* **4** (1970) p. 1826.
14. F.A. Kish, F.M. Steranka, D.C. DeFevere, D.A. Vanderwater, K.G. Park, C.P. Kuo, T.D. Osentowski, M.J. Peanasky, J.G. Yu, R.M. Fletcher, D.A. Steigerwald, M.G. Craford, and V.M. Robbins, *Appl. Phys. Lett.* **64** (1994) p. 2839.
15. M.R. Krames, M. Ochiai-Holcomb, G.E. Hofler, C. Carter-Coman, E.I. Chen, I.-H. Tan, P. Grillot, N.F. Gardner, H.C. Chui, J.-W. Huang, S.A. Stockman, F.A. Kish, and M.G. Craford, *Appl. Phys. Lett.* **75** (1999) p. 2365.

The Semiconductor *p–n* Junction "Ultimate Lamp"

Nick Holonyak Jr.

Abstract

Simple diagrams are used to show the transformation of a thin sample of intrinsic, direct-gap semiconductor from an ideal "flat-band" photopumped recombination-radiation light source into a current-driven, *p–n* junction "ultimate lamp," a light-emitting diode.

Keywords: *p–n junctions, LEDs, light-emitting diodes.*

Introduction

The surprising arrival of the transistor (December 16, 1947)[1,2]—a revolutionary new idea and device—at once made the hole, the unique valence-band "void" (positive, *p*) of the semiconductor, equal to the electron (negative, *n*) in behavior and importance. At first a theoretical notion, the hole suddenly became real. The transistor, to be a "transfer resistor" exhibiting gain, needed hole (+) current as well as electron (–) current. It was the base current, of opposite charge polarity, that separated the low-impedance input from the high-impedance output. The transistor was naturally bipolar and of necessity led to the study of electron–hole nonequilibrium and recombination, not to mention the study of crystal growth and doping (*n*-type or *p*-type) as well as crystal perfection. A new electronics was launched, employing electron, hole, *and* photon.

The semiconductor, with electron and hole coupled across the energy gap (E_g) by a photon, could inherently absorb and thus detect light or, of more concern here, generate light. But how should light-generation be effected, in what kind of crystal, and by what excitation method? What form should the semiconductor take? We show via a fundamental argument, beginning with "light in" to get "light out," that if excitation with a current is specified (and desired), the initially undoped (intrinsic) semiconductor slab has to be modified and takes the form of a *p–n* junction. The *p–n* junction is in fact an "ultimate lamp,"[3] by which I mean a lamp that cannot be exceeded in efficiency of converting electrical energy to optical energy. Also, as proven in practice, a direct-gap semiconductor, with its conduction-band electrons and valence-band holes aligned in momentum **k** ($\mathbf{k}_e = \mathbf{k}_h$, not $\mathbf{k}_e \neq \mathbf{k}_h$), is preferred because of its stronger band-to-band matrix element (coupling).

Light In and Light Out

We begin with an experiment. Figure 1 shows an early 1960s-era GaAsP laser operating on a filament from one polished Fabry–Perot resonator facet to the other facet, with the output from one end at 6708 Å impinging directly, by direct contact, onto one side of a 1–2-μm-thick CdSe platelet. The photoexcitation is essentially fully absorbed and causes the CdSe to operate as a surface-emitting laser 50 meV lower in energy at 6893 Å.[4] This is possible in a dense atomic system with a high loss-absorption coefficient, $\alpha_l \sim 10^4$ cm^{-1}, which can, by photopumping, be inverted into gain ($\alpha_g \sim \alpha_l$) at energy $\hbar\omega \sim E_g$, where $\hbar$ is Planck's constant divided by 2π and ω is the frequency. There is only a 3% reduction in photon energy (Figure 1). We throw away half of the GaAsP laser light (right side) and convert almost all of the remainder (left side) into CdSe laser light, so for the light in (the excitation), we get almost all of the light out (the output). The problem is that we wish to do this with current excitation, not light excitation, and eventually with a III–V semiconductor and not the technologically more intractable II–VI semiconductor CdSe of Figure 1, which is employed here merely for the sake of convenient illustration.

The semiconductor energy bands for the experiment we have described[4] are shown in Figure 2. The expressions in Figure 2 for the electron density *n* and hole density *p* are the usual well-known approximations[5] and depend on the semiconductor intrinsic concentration n_i, the electron quasi-Fermi level (chemical potential) E_{Fn}, the hole quasi-Fermi level E_{Fp}, and the intrinsic level E_i (the reference also for the potential $\psi \equiv E_i/q$, where q is the electron charge).[5] For every electron generated, there is a compensating hole, charge neutrality is preserved, and the energy bands (Figure 2) remain in the flat-band configuration. In the region of excitation, E_{Fn} approaches or is slightly above the conduction-band edge E_c, and E_{Fp} shifts slightly below the valence-band edge E_v ($[E_c - E_v] = E_g$). This is consistent with the expressions for n and p in Figures 2–4. A photovoltage $(E_{Fn} - E_{Fp})/q$ exists in the region of excitation that can be used as a battery, provided it can be contacted.