Physics Nobel Prize Honors Roots of Information Age

> Jack Kilby, coinventor of the integrated circuit, and Zhores Alferov and Herbert Kroemer, early pioneers of semiconductor heterostructures, made today's electronic and communications technology possible.

The Royal Swedish Academy of Sciences has awarded the 2000 Nobel Prize in Physics to Zhores Alferov, director of the Ioffe Physico-Technical Institute in Saint Petersburg, Russia; Herbert Kroemer of the University of California, Santa Barbara; and Jack Kilby, now retired from Texas Instruments Inc (TI) in Dallas, Texas, "for basic work on information and communication technology." Alferov and Kroemer will share half of the SKr 9 million (about $900 000) prize "for developing semiconductor heterostructures used in high-speed and opto-electronics." Kilby is to receive the other half "for his part in the invention of the integrated circuit."

The twin pillars of today's electronic society are computers and communication systems. Both have grown very rapidly over the past decade, as attested by the increased role of the Internet and the World Wide Web in both business and private life. In acknowledging the impact of these technologies, the Nobel committee singled out two inventions that have been driving this development: the integrated circuit (IC) and semiconductor heterostructures.

The transistor is often singled out as the basis for today's electronic society, but it was the IC that gave the transistor its modern form and allowed its power to be harnessed in a reliable, low-cost way. Heterostructures form the basis of semiconducting lasers and light-emitting diodes (LEDs) as well as high-speed transistors and photocells.

Putting the pieces together

The invention of the transistor in 1947, for which John Bardeen, Walter Brattain, and William Shockley received the Nobel Prize in Physics in 1956, was a watershed event in the history of electronics (see PHYSICS TODAY, December 1997, page 34). By the late 1950s, the switch from large, power-hungry vacuum tubes to small and efficient transistors was well under way, and transistorized electronic products—including hearing aids and radios—were commercially available.

The strong push for the miniaturization of electronics motivated Kilby to join TI in the summer of 1958. His first few months there are now the stuff of legend. As a new employee, Kilby could not take vacation that summer, and so he looked for a project to work on. "I had looked at the alternate miniaturization schemes that were around at that time," recalls Kilby, "and I concluded that the best thing TI could make was semiconductors." His idea: Use semiconductors not just for transistors and diodes, but also for resistors and capacitors, putting everything on the same substrate. The bulk resistivity of the semiconductor and its diffusion-doped layers could be exploited for fabricating resistors; p–n junctions could provide capacitance. By September, Kilby had made working devices—a phase-shift oscillator and a flip-flop—from available blocks of germanium, and within a few months had made ICs from silicon.[1]

In the discrete electronic circuits at that time, the need to manually solder all the components together was labor intensive and error prone. In Kilby's first circuit, some of the components were interconnected by virtue of their arrangement on the germanium, with the necessary additional interconnections made with gold wire bonds. This approach worked as a proof of principle, but for larger ICs, better solutions were needed.

In early 1959, at the upstart Fairchild Semiconductor Corp, Robert Noyce had a different idea for fabricating ICs, drawing on the planar process his colleague Jean Hoerni had developed. In the late 1950s, mesa transistors—which had protruding plateaus—were the state of the art. Although they had fairly good performance, these transistors were exposed, prey to contamination and surface leakage, and

ZHORES ALFEROV

JACK KILBY

TONY J. MASTRES/UCSB

HERBERT KROEMER

© 2000 American Institute of Physics, S-0031-9228-0012-320-8

difficult to connect.

Hoerni's approach yielded transistors that were buried beneath a layer of silicon dioxide. From the work of John Moll and Carl Frosch at Bell Labs, the merits of SiO_2 were already beginning to be appreciated as a protective coating, as a diffusion mask for controlling the boundaries of the doping process, and as a good natural insulator that has few inherent charge traps and interface states. Largely because of the desirable properties of its oxide, Si was starting to replace Ge as the semiconductor of choice in electronics (see PHYSICS TODAY, October 1997, page 50).

As Noyce realized, the flat transistors resulting from Hoerni's planar process could be interconnected by depositing metal over etched holes in the oxide. This fabrication process is still used today, "although it's been improved almost beyond recognition," comments Kilby. Noyce, who died in 1990, and Kilby are generally regarded as coinventors of the IC.

Gordon Moore, who left Fairchild with Noyce in 1968 to found Intel, noted in 1965 that the complexity of ICs was doubling roughly every year. With a slight revision in time constant, to about 18 months, this exponential growth of semiconductor devices—known as Moore's law—is still going strong, driven by increased control over processing, the quality of materials, and especially the refinements in lithography that allow reduced feature size and higher packing densities. "The modern IC is a result of the efforts of tens or hundreds of thousands of the best engineers," notes Kilby, "with almost a continuous series of advances."

Mixing and matching

Doping a semiconductor like Si or GaAs provides control over the sign and density of the charge carriers. By combining different semiconductors in heterostructures, one gains control over much more, including the band gap energy, refractive index, carrier mass and mobility, and other fundamental parameters.

Not long after the development of the transistor, ideas for controlling the holes and electrons in them started to emerge—in particular, what's come to be known as bandgap engineering. In his 1948 transistor patent application, Shockley mentioned the idea of using a wider-gap semiconductor for the emitter than for the rest of the transistor, a configuration analyzed by Aleksandr Gubanov at Ioffe in 1951.

Lights out at LEP

As we were going to press, the CERN directorate announced that the 2 November shutdown of LEP was indeed the e^+e^- collider's final termination. The 27-km LEP tunnel will now be made ready for the Large Hadron Collider, which should be providing 14-TeV proton–proton collisions by 2005. LEP had already enjoyed a one-month reprieve in response to recent tantalizing hints of a possible Higgs boson with a mass of about 115 GeV/c^2. But, after heated discussion, it was finally decided that the extra October run did not uncover sufficient confirmatory evidence to justify a costly further extension of LEP's 11-year life.

BERTRAM SCHWARZSCHILD

In 1957, Kroemer, then at RCA in Princeton, New Jersey, returned to an idea he had had a few years earlier: varying the energy gap in semiconductors.[2] He realized that a spatially varying energy gap, and the resulting variations in the semiconductor's conduction and valence bands, creates quasi-electric fields that act differently on electrons in the conduction band and holes in the valence band. The resulting forces on electrons and holes could even act in the same direction.

Kroemer realized that this effect should make it possible to create devices that are fundamentally impossible to achieve in homostructures. "While the examples I could come up with—transistors with a graded bandgap in the base, or with a wide-gap emitter—represented improvements," he told us, "they did not fully exploit the new principle." This would change abruptly six years later with the idea of the double-heterostructure (DHS) laser.

Starting in the early 1960s, much attention in the semiconductor community was aimed at developing a laser. The indirect bandgap of Si and Ge made those elements unlikely candidates, so the focus was on III–V semiconductors such as GaAs. The goal was attained in 1962 by several groups: Robert Hall and coworkers at General Electric Co in Schenectady, New York, Marshall Nathan and colleagues at IBM Corp, and Robert Rediker and coworkers at MIT's Lincoln Laboratory demonstrated lasing in GaAs, and Nick Holonyak Jr and Sam Bevacqua at GE in Syracuse created a visible laser in the alloy GaAsP.

These early devices used forward-biased p–n homojunctions made of a single semiconductor with different dopants. Electrons from the n side and holes from the p side recombined in the junction region, giving off light. The lasers were not very efficient because of high optical and electrical losses. Thus the threshold currents were high and low temperatures were necessary.

On learning of the large losses, Alferov in the Soviet Union and Kroemer in the US—neither of whom had any prior background in lasers—independently and simultaneously proposed the solution: use a DHS. DHS lasers, with a low-bandgap material sandwiched between two wide-bandgap semiconductors, have two significant inherent advantages over homostructures. Because of the lower conduction band energy and higher valence band energy in the middle layer, the electrons and holes are trapped there. But lower bandgap materials also have higher indices of refraction. Thus, in addition to carrier confinement, the DHS also provides optical confinement of the emitted light to the lower-bandgap region. Alferov proposed DHS lasers in a 1963 Soviet patent application with Rudolf Kazarinov.[3] Kroemer published a similar suggestion in the US, and filed for a US patent just one week after Alferov.[4]

Alferov organized an effort at Ioffe to explore heterostructure applications, and many US labs continued work on semiconductor devices. However, Kroemer, working in industry at Varian Associates in Silicon Valley at that time, was refused support for working on DHS. "I was told that the new device could not possibly have any practical applications." He turned his attention to the Gunn effect, but in the early 1970s returned to heterostructures, focusing on heterojunction bipolar transistors (HBTs).

Although the ideas for DHS were there, the technology wasn't. The primary question was which materials to use, and the dominant consideration was matching the lattice constants of the semiconductors. GaAs–GaAlAs eventually emerged as the pairing of choice, once researchers correctly recognized the close values of their lattice constants and realized that, whereas AlAs will oxidize in air, the alloy $Ga_xAl_{1-x}As$ is chemically stable for low to moderate Al concentrations.

Heterostructure research picked up steam in 1967, when Alferov in the Soviet Union and Hans Rupprecht and Jerry Woodall at IBM independently reported the first GaAs–GaAlAs heterojunctions, grown by liquid-phase epitaxy. Whereas the US groups were working on single heterostructures, Alferov's group was committed to DHS.

Over the next few years, his group achieved efficient carrier injection, the predicted electrical and optical confinement, and low-threshold pulsed room-temperature lasing in such devices. And in 1970, continuous-wave (CW) room-temperature semiconductor lasers were at last reported, in GaAs–GaAlAs DHS devices, by Alferov in the Soviet Union and, one month later, by Izuo Hayashi and Morton Panish at Bell Labs in the US.[5]

'Heterostructures for everything'

GaAs–GaAlAs heterostructures were by no means the only devices being pursued. Ternary compounds, such as GaAsP and InGaAs, and quaternary compounds like InGaAsP also received much attention. Armed with new deposition techniques—molecular beam epitaxy, developed by Alfred Cho, John Arthur, and their Bell Labs coworkers, and metalorganic chemical vapor deposition of semiconductors, pioneered by Harold Manasevit, Russell Dupuis, and Paul Dapkus at Rockwell International Co—researchers could fabricate almost any structure that could be drawn, ultimately with single-atom thicknesses.

The controlled layer deposition facilitated the fabrication and study of heterostructure superlattices by Leo Esaki, Raphael Tsu, and Leroy Chang at IBM, which have in turn led to resonant tunnel diodes and, recently, to quantum cascade lasers developed by Federico Capasso at Bell Labs Lucent Technologies. The improved control has also allowed the reduction of the middle layer of the DHS sandwich. In 1974, Raymond Dingle and company at Bell Labs reported observing quantized size effects. Exploiting this reduced dimensionality, Dupuis, Dapkus, Holonyak, and coworkers reported a CW room-temperature quantum well laser in 1978. From heterostructures with two-dimensional quantum wells, researchers have gone to even fewer dimensions—one-dimensional quantum wires and zero-dimensional quantum dots. And vertical device geometries, with high-quality optical cavities fabricated into the structure, have emerged with vertical-cavity surface-emitting lasers.

Near-infrared semiconductor lasers and LEDs are now commonplace. The telecommunications industry has blossomed with the development of long-wavelength InGaAsP heterostructure lasers, lattice-matched to InP substrates, for use with optical fibers (see PHYSICS TODAY, September 2000, page 30). And the recent development of AlGaInN devices (see PHYSICS TODAY, October 2000, page 31) has produced blue and violet heterostructure lasers and LEDs.

Heterostructures are also prevalent in transistors. In addition to HBTs, field-effect devices with very high mobilities have been developed. Called HEMTs, for high electron-mobility transistors, these heterostructures, which use dopants in high-bandgap materials to generate a high-density two-dimensional electron gas (2DEG) with low defects in a quantum well, can operate at very high frequencies and with low noise. Such devices, primarily of GaAs–GaAlAs and GaAs–InGaP but increasingly of Si–Ge alloys, can be found in cell phones and satellite receivers.

In addition to being key components in the billion-dollar optoelectronics and consumer electronics industries, heterostructures have also become fundamental structures for basic physics research. The fractional quantum Hall effect was observed by Horst Stormer and Daniel Tsui in 2DEG heterostructures (see PHYSICS TODAY, December 1998, page 17). Ballistic transport and quantum-dot "artificial atoms" are among the many other research topics made possible with heterostructures.

Heterostructures have emerged as the basic building block of semiconductor devices. "Over 99% of the semiconductor research today involves heterostructures," says Alferov. The title of a talk Kroemer gave in 1980 has come true: "Heterostructures for Everything."

Biographies

Born in Vitebsk, Byelorussia, USSR, in 1930, Alferov has been at Ioffe in Saint Petersburg since he graduated in 1952 from the Ulyanov Electrotechnical Institute in the same city. He has been director of the institute since 1987, and is also a member of the Russian Duma.

Kilby was born in Jefferson City, Missouri, in 1923 and received his MS in 1950 from the University of Wisconsin in electrical engineering. He worked for Centralab in Milwaukee until 1958, when he joined TI. A co-inventor of the pocket calculator and the thermal printer and holder of more than 60 patents, Kilby took a leave of absence from TI in 1970, was a professor at Texas A&M University from 1978 to 1984, and now is an independent consultant.

A native of Weimar, Germany, Kroemer got his start in semiconductor physics at age 24 with a PhD dissertation on the then-new transistor from the University of Göttingen in Germany in 1952. He held positions at various laboratories in Germany and the US until 1968, when he joined the faculty of the University of Colorado. In 1976 he moved to the University of California, Santa Barbara, where he is a professor in the department of electrical and computer engineering and the materials department.

RICHARD FITZGERALD

References

1. See J. S. Kilby, IEEE Trans. Electron. Devices **23**, 648 (1976).
2. H. Kroemer, Proc. IRE **45**, 1535 (1957).
3. See Zh. I. Alferov, Phys. Scr. **T68**, 32 (1996).
4. H. Kroemer, Proc. IEEE **51**, 1782 (1963). See also H. Kroemer Phys. Scr. **T68**, 10 (1996).
5. Zh. I. Alferov *et al.*, Sov. Phys. Semicond. **4**, 1573 (1971). I. Hayashi, M. Panish, P. W. Foy, S. Sumski, Appl. Phys. Lett. **17**, 109 (1970).

Nobel Prize in Chemistry Salutes the Discovery of Conducting Polymers

> Conducting polymers have found applications ranging from antistatic coatings to all-polymer integrated circuits.

In 1976, a serendipitous chain of events brought together three individuals from different academic and geographical cultures to study a curious polymer: polyacetylene. The trio soon discovered that doping this polymer can change its behavior from insulating to metallic. For that work, the three—Alan Heeger, a physicist now at the University of California, Santa Barbara; Alan MacDiarmid, a chemist from the University of Pennsylvania then specializing in inorganic chemistry; and Hideki Shirakawa, a polymer chemist who has recently retired from Japan's Tsukuba University—have now earned the 2000 Nobel Prize in Chemistry "for the discovery and development of conductive polymers."

Heralded at the time, the discovery of conducting polymers has become even more significant in hindsight as this class of materials has proven to be