

◆ The Foundation of the Silicon Age

Ian M. Ross

The invention of the transistor almost fifty years ago was one of the most important technical developments of this century. It has had profound impact on the way we live and the way we work. This paper describes the events that led to the invention of the point-contact transistor in December of 1947. It continues with the development of the theory of the junction transistor in early 1948 and the fabrication of the first grown-junction transistor in 1950. The paper next covers the major hurdles that had to be overcome and the major breakthroughs that had to be made to turn an exciting invention into a far-reaching technical innovation. The final part of the paper suggests some of the reasons why such an important technological progress could occur in a relatively short period of time.

Setting the Course

In the summer of 1947, any of a number of technical societies could have held a symposium to commemorate the 50th anniversary of J. J. Thomson's discovery of the electron. That 1897 event could surely qualify as the start of the electronics discipline and the industry that followed. It was the new understanding of the properties of the electron that created the field of electronics, and that, combined with our developing capability in the electrical, magnetic, and mechanical arts, enabled a rich array of new products and services.

The symposium would have been an upbeat event. Vacuum tube technology had fully matured with a wide range of tubes—diodes, pentodes, CRTs, klystrons, and traveling-wave tubes—in high-volume manufacture. Vacuum tubes were the key component in an array of electronic equipment that seemed to meet all conceivable information needs.

Mervin Kelly, the then Director of Research at Bell Labs who later became Bell Labs president, might well have been invited to submit a paper to the symposium. And he would also have been upbeat. Electromechanical relay technology was making possible fully automatic telephone dialing and switching. Microwave radio was providing high-quality telephone transmission across the continent. Again,

available technology appeared capable of meeting conceivable needs.

Yet Kelly would also have raised a word of caution. Although relays and vacuum tubes were apparently making all things possible in telephony, he had predicted for some years that the low speed of relays and the short life and high power consumption of tubes would eventually limit further progress in telephony and other electronic endeavors. Not only had he predicted the problem, he had already taken action to find a solution. In the summer of 1945, Kelly had established a research group at Bell Labs to focus on the understanding of semiconductors. The group also had a long-term goal of creating a solid-state device that might eventually replace the tube and the relay.

Kelly's vision triggered one of the most remarkable technical odysseys in the history of mankind, a journey that has continued through fifty years. The semiconductor odyssey produced a revolution in our society at least as profound as the introduction of steel, of steam engines and the total Industrial Revolution. Today electronics pervades our lives and affects everything we do.

My purpose in this paper is to discuss the events that led to the invention of the transistor plus the hurdles that had to be overcome and the breakthroughs

that were needed to make the semiconductor revolution a reality. In doing this, I have tried to select those events that made *the* difference rather than cover the multitude of contributions that made *a* difference. Admittedly, making this selection has required some judgment on my part.

The Scientific Phase

By January of 1946, Kelly's semiconductor group was in place at Bell Labs under the leadership of Stanley Morgan and William Shockley, both from the Physical Research area. A very capable physicist, Bill Shockley was also an analyst and a man with a fascination for finding practical applications of science. Two other key members of the team were John Bardeen and Walter Brattain. John Bardeen was a remarkably talented theoretical physicist, as evidenced by the fact that he was awarded two Nobel Prizes in physics, each in a field of major significance. Walter Brattain was also an accomplished physicist with a flair for ingenious experiments. Other members included Gerald Pearson, Bert Moore, and Bob Gibney. The team was embedded in the unusually creative environment that existed in Bell Labs, Murray Hill, after World War II. As such it was able to seek the advice of resident experts in almost any relevant discipline.

The group had as well a number of other assets to call on in pursuit of Kelly's goal. There existed a large body of empirical knowledge of semiconductor devices based on experience with diodes for detection of radio signals. These diodes ranged from the "cat's whisker" crystal diodes at the heart of early radio receivers to the microwave diodes used in great quantities during the war for radio and radar detection. In addition, considerable experience had been gained with power rectifiers such as copper oxide diodes. These devices were made from a variety of materials including selenium, lead sulfide (galena), copper oxide, germanium, and silicon. All were semiconductor materials, most were highly impure, and none was single crystal. There was much art, much tinkering, but little engineering understanding and almost no science.

There was already some basis for understanding the physics of semiconductor materials. The concept of band gaps existed. Two types of conduction,

already named n-type and p-type, had been identified in semiconductors and attributed to the presence of certain impurities in very small concentrations. P-n junctions had been found within ingots formed by melting and re-freezing the purest silicon then commercially available. Their electrical and electro-optical characteristics had been explored. Considerable progress had already been made at Purdue University, Bell Labs, and elsewhere in producing semiconductor materials of increasing purity and in understanding their properties.

However, there was also much uncertainty, much still unknown. The highest purity silicon available—99.8%—was characteristic of a soap advertisement and orders of magnitude short of that eventually needed. Semiconductor materials were polycrystalline at best and frequently used in powder form. Single crystals of adequate perfection had yet to be grown. The key properties of these materials relevant for device applications had yet to be fully understood and evaluated.

Finally, there was a long and persistent history of proposals for a solid state amplifier. Most were based on the so called "field-effect" mechanism. The concept was that an electric field applied through the surface of a semiconductor could modify the density of mobile charge in the body of the material and thereby change its conductivity. Typically the field was to be created by applying voltage to a metal plate close to but insulated from the base material. Modulating the voltage on the plate would modulate a current flow through the base material with the possibility of power gain. The first documented invention of this kind was made by J. E. Lilienfeld as early as 1925.¹ All attempts to make such a device had failed, however.

Both before and after the war, Shockley had studied and analyzed possible field-effect structures and had concluded that the effect should lead to amplification in achievable structures. Shockley's existence proof that amplification was theoretically possible in practical semiconductor materials provided major encouragement that the challenge undertaken by the Bell Labs group could indeed be accomplished.

By January 1946, two critical decisions had been made. The first was to focus the group's attention on

crystals of silicon and germanium and ignore other more complex materials frequently used in prior investigations. It was recognized that silicon and germanium were stable elements that readily assumed the crystalline state and therefore showed the best promise of being made into high-purity, high-perfection single crystals. Such materials would permit the investigation to move forward on a sound scientific base. The second decision was to pursue the field-effect principle as the one having the most assurance of leading to a useful device.

Numerous attempts to demonstrate the field effect in semiconductors had been made over the years, and all had failed. Before the war, Shockley had participated in one such failure using a structure with a grid of metal filaments buried in the body of a semiconductor. Given the renewed focus, a number of new experiments were carried out by J. R. Haynes, H. J. McSkimin, W. A. Yager, and R. S. Ohl in attempts to observe the field effect. All gave negative results. Bardeen proposed that these experiments failed because the electric field was not penetrating the body of the semiconductor material but was terminated by immobile charges trapped in states at the semiconductor surface.² He calculated that a quite small number of such surface states, low compared to the density of surface atoms, would be adequate to shield the body from any measurable field effect.

Bardeen and Brattain attempted to confirm this theory by experimenting with metal probes on the surface of germanium. The theory seemed to be correct. Thus for the first time there was some understanding of the persistent failure to observe the field effect and an opportunity to intervene. In the course of their work, they tried to modify the surface states with electrolytes surrounding the metal contacts to the germanium surface. Following a suggestion by Gibney,³ they found that applying voltage to the electrolyte created major changes in the current flow through a reverse biased contact. Brattain later replaced the electrolyte with an evaporated gold spot adjacent to the point contact. Finally, he replaced both contacts by an ingenious arrangement of two strips of gold foil separated by just a few mils and pressed onto the germanium surface. With one gold contact for-

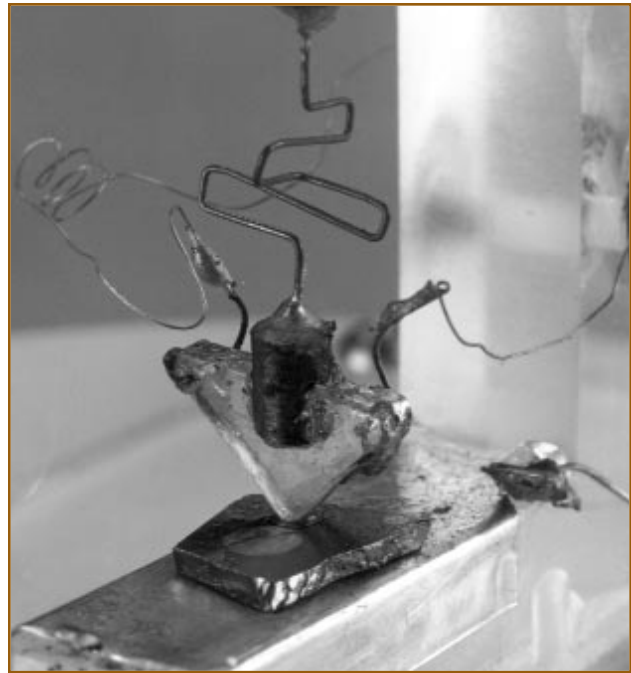


Figure 1.
The original point-contact transistor.

ward biased and the other reverse biased he observed power gain (see **Figure 1**). The transistor effect had been discovered.⁴ This was on December 16, 1947, a mere two and a half years after the formation of the Shockley group.

On Christmas Eve of 1947, the transistor action was demonstrated by Brattain and Moore for the top management of Bell Labs. This time the device was operated as an oscillator, an acid test of the existence of power gain. The announcement of the transistor discovery was delayed, however, until June 1948. This six-month period was used to gain more understanding of the device and its possible applications and to obtain an adequate patent position. Shockley, Bardeen, and Brattain were awarded the 1956 Nobel Prize in physics for the invention of the transistor (see **Figure 2**).

The above is an abbreviated account of the events that led to the invention of the transistor. I believe it to be essentially correct. It is consistent with a memorandum written in December 1949 by W. S. Gorton, an assistant to the Director of Research of Bell Labs.⁵ Gorton had been asked by his management, while the memories were reasonably clear, "to write an account

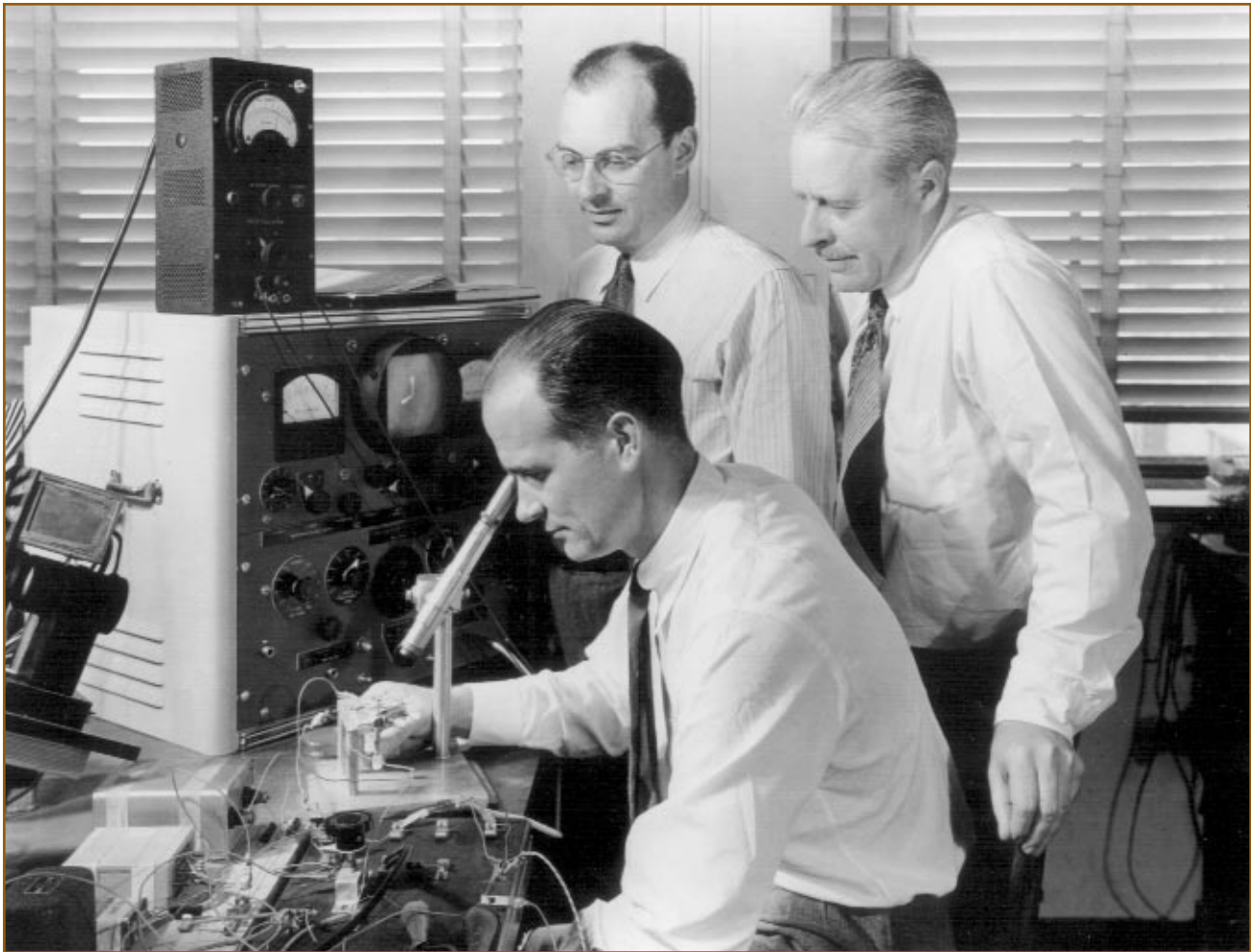


Figure 2.
The three inventors of the transistor: (left to right) William Shockley, John Bardeen, and Walter Brattain, who were awarded the 1956 Nobel Prize in physics.

of the thinking, work, and events which resulted in the transistor.”⁶ Gorton’s memorandum is probably the most authentic summary in existence. In preparing his account, Gorton addressed the question of giving full credit to all who had contributed. Gorton’s memorandum includes the names of twelve people who had taken a substantial part in the work. Those names all appear in the foregoing account.

With the invention of the point-contact transistor—the gold foil having been replaced by two closely spaced point contacts—and with the demonstration of transistor action, the door had been opened to a whole new era of electronics. But the understanding of the transistor still had a long way to go. Transistor action had been observed, but no one understood just what the mechanism was. Was it a surface effect or was the

action occurring in the semiconductor body? Ironically, the mechanism certainly was not the field effect that had helped guide the whole effort.

Bardeen and Brattain leaned in the direction of a surface effect and continued experiments on that basis. Shockley, however, had recognized the role of minority carriers and, by late January of 1948, he had completed a thorough formulation of p-n junction theory and the role played by the injection of minority carriers in forward bias and their collection in reverse bias. His analysis concluded with the invention of a junction transistor, a sandwich of lightly doped n-type material between two regions of p-type—or the other way around. With one p-n junction forward biased and the other reverse biased, minority carriers would be injected from the forward-biased junction into the

n-type material. They could then diffuse across the n-type region and, if it were thin enough, a large fraction would be collected at the reverse junction. Thus current generated in a low-impedance circuit, the emitter, would create a similar current flow in a high-impedance circuit, the collector, and power gain would result.⁷ But this so far was just theory.

One month later, in February of 1948, John Shive carried out a critical experiment.⁸ He applied two phosphor-bronze contacts to the opposite sides of a 0.01-cm-thick slice of germanium. With this arrangement he observed transistor action from one contact to the other with substantial power gain. The length of the surface path around the semiconductor slice effectively ruled out a surface effect. The action had to take place through the semiconductor body. The behavior he observed was nicely explained by Shockley's recently developed theory of the junction transistor. Thus, while the point-contact transistor may have exhibited some surface effects, bulk propagation was also surely taking place and was probably the dominant effect.

The next major advance was made in 1948. G. K. Teal and J. B. Little succeeded in growing a single crystal of germanium by slowly pulling a seed crystal from a melt of high-purity germanium.⁹ Using such material it was at last possible to detect and characterize minority carriers injected by metal contacts into filaments of germanium. Various elegant experiments by Haynes, Pearson, Suhl, and Shockley confirmed the behavior of both types of minority carriers and yielded measurements on injection efficiency, mobility, diffusion coefficients, and lifetime.¹⁰ These results showed that useful devices could be made according to Shockley's junction transistor theory. All that remained was to make one.

That required further refinement of the techniques of crystal growth and particularly of the controlled doping of the crystals during growth. In April 1950, a team of Shockley, Sparks, and Teal succeeded in growing a crystal containing a thin region of p-type embedded in n-type material. The crystal was cut into n-p-n rods and contacts were applied (see **Figure 3**). The electrical properties of the resulting devices were largely consistent with the Shockley theory.¹¹

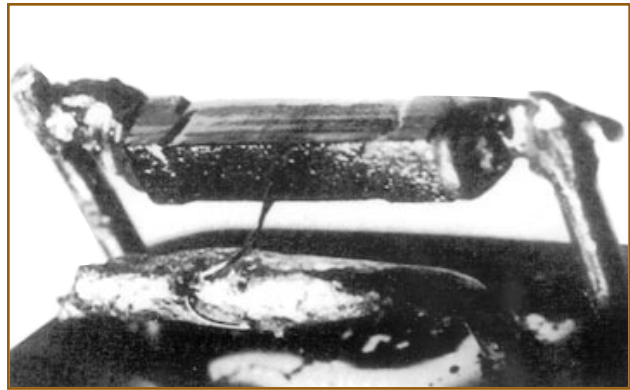


Figure 3.
The first grown-junction transistor.

Transistor electronics now had a solid foundation.

One other event completed this phase of the transistor saga. That was the publication in 1950 of Shockley's book *Electrons and Holes in Semiconductors*.¹² This was an exquisite account of the current understanding of semiconductors and transistors. It makes enlightening reading even today, after almost fifty years. In the '50s it provided an excellent means, and almost the only means, for scientists and engineers to get up to speed on a rapidly developing technology. It was required reading for those entering the business in its early days, particularly if you found yourself reporting to its author, as I did in March of 1952.

So in a period of only five years from the establishment of the semiconductor group at Bell Labs, the invention of the transistor was essentially complete, understood and documented. The scientific phase was coming to an end. The next phase would focus on solving development and engineering issues so that a brilliant invention could be converted into an revolutionary innovation.

The Development and Engineering Phase

Following the invention of the transistor, the challenge was then to find ways to design a product that could be manufactured and that could sustain a market. This phase took the industry approximately eight years, during which many challenging problems were addressed and solved. Whereas the scientific phase had been dominated by Bell Labs, other companies were now in the business, and they also made major innovations.

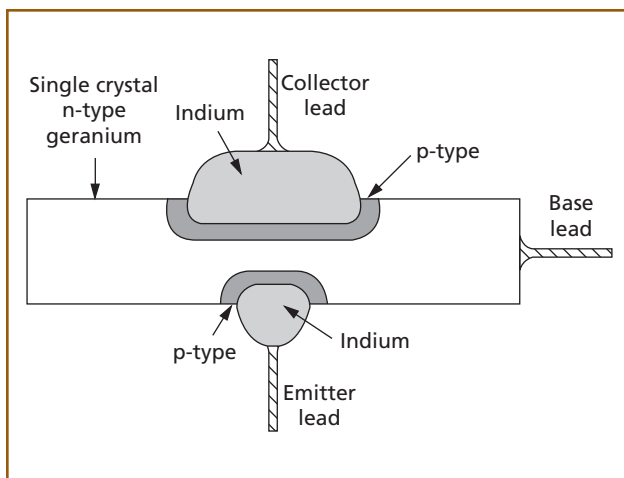


Figure 4.
Schematic of p-n-p alloyed germanium junction transistor.

What follows is an attempt to select and describe some of the major hurdles that had to be overcome and the major breakthroughs that were made. As I have previously stated, many events made a difference, but my focus here is on those that made *the* difference.

The Early Manufacturing Problems

In early 1951, two transistor structures had been proven to work, but neither of them was suitable for large-scale manufacture. The point-contact transistor had all the frailties of its cat's whisker heritage. It was difficult to make and its electrical characteristics, far from ideal, were very variable, hard to control, and inherently unstable. Point-contact transistors were, nevertheless, manufactured for ten years, but they were never popular with the manufacturing engineer or the circuit designer.

The junction transistor, on the other hand, had predictable and more desirable electrical characteristics. It was, however, prodigal in its use of precious semiconductor material and it required tricky contacting techniques not conducive to automation.

The grown-junction transistor was manufactured starting in 1952. In the same year, J. E. Saby at General Electric announced the development of the alloy junction transistor.¹³ The original version was made by alloying dots of indium, an acceptor material, on opposite sides of thin slices of n-type germanium (see **Figure 4**). The starting point was the growth of

uniformly doped crystals that were relatively easy to produce. Slices were cut from the crystal, most of which could be used. Arrays of indium dots could be positioned in jigs on either side of the slices and, after alloying, each slice could be diced to yield a great many individual transistors. Contacts were easy to apply. The alloy transistor had well-behaved performance characteristics, made efficient use of semiconductor material, and could be manufactured with some degree of batch processing and automation. The alloy device was the first transistor to be readily manufactured and, for some years, was the mainstay of the industry. One drawback, however, was that precise control of dimensions and alloying temperatures were required to create base layers thin enough for high-frequency performance.

The Quest for Silicon

It was understood from the beginning that silicon would be a better transistor material than germanium for most applications. This mainly resulted from the higher energy gap of silicon—1.1 eV compared to 0.67 eV for germanium. In germanium at room temperature the thermal generation of minority carriers led to substantial reverse currents in p-n junctions. The reverse current in silicon was orders of magnitude smaller and made a far superior rectifier.

The most serious problem with silicon was that critical chemical and metallurgical processes all took place at substantially higher temperatures. For example, the melting point of silicon was 1415°C compared to 937°C for germanium. Silicon was also more chemically reactive than germanium. For example, silicon would react with the quartz crucibles that were used to contain germanium during crystal growth and purification by zone refining.

The critical breakthrough came in 1953 with H. C. Theuerer's development of the floating zone method.¹⁴ Theuerer was able, in a vertical rod of silicon, to create a zone of molten material contained only by surface tension (see **Figure 5**). Thus the zone refining technique could be used for silicon, and it resulted in crystals of purity comparable to the best obtained in germanium.

In 1954, Teal, who had moved to Texas Instruments, made the first silicon transistor using the grown junc-

tion method.¹⁵ All the pieces were then in place for silicon devices to assume a major role.

The Bob Wallace Revelation

Once the hurdle of being able to make transistors with some degree of reproducibility was overcome, replacing the vacuum tube in as many applications as possible became the goal. Accomplishing this was not as simple as it first appeared. Transistors were easiest to make in small sizes, which inherently led to limited power-handling capability. High-frequency response called for smaller, not larger, devices. In seeking higher power at higher frequencies, we seemed to be bucking nature.

One day I was in a small meeting at Bell Labs with a colleague named Bob Wallace. As was frequently the case, we were discussing our problems in emulating the vacuum tube when Bob suddenly commanded our attention with a comment that went something like this:

Gentlemen, you've got it all wrong! The advantage of the transistor is that it is inherently a small-size and low-power device. This means that you can pack a large number of them in a small space without excessive heat generation and achieve low propagation delays. And that's what we need for logic applications. The significance of the transistor is not that it can replace the tube but that it can do things that the vacuum tube could never do!

And this was a revelation to us all. We realized that in chasing the vacuum tube we had the wrong emphasis.

I am sure that the same idea occurred independently to other people in other organizations at about that time. The net result was that the semiconductor community began to relax about replacing the tube and focused on developing the transistor in its own right.

There is a lesson in this story. Having the clear goal of an application for an invention is a powerful stimulus for innovation. But frequently it turns out that the original application is not the most important application.

The Speed Problem—Controlling the Depth Dimension

The fundamental determinant of the frequency response of a junction transistor was the transit

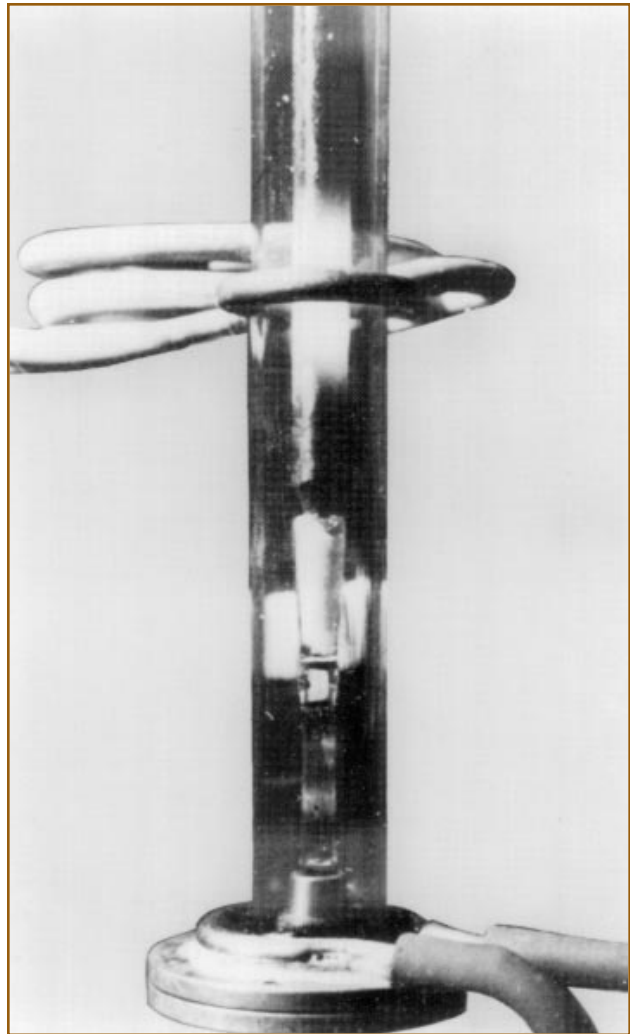


Figure 5.
Floating zone apparatus used to purify silicon crystals.

time of minority carriers across the base region, and therefore the thickness of the base layer. In practice, alloy transistors were manufactured with bases as thin as 10μ , yielding a frequency response approaching 10 MHz. Although this was quite a feat of manufacturing engineering, performance up to a few gigahertz was needed to support a full range of electronic applications.

The base width problem was solved by using the process of diffusing donors and acceptors into the semiconductor surface, a process which eventually yielded precise control of the depths of diffused layers in the range from 20μ to a fraction of a micron.

In 1954, C. A. Lee made the first germanium diffused transistor.¹⁶ He diffused a base layer of arsenic to

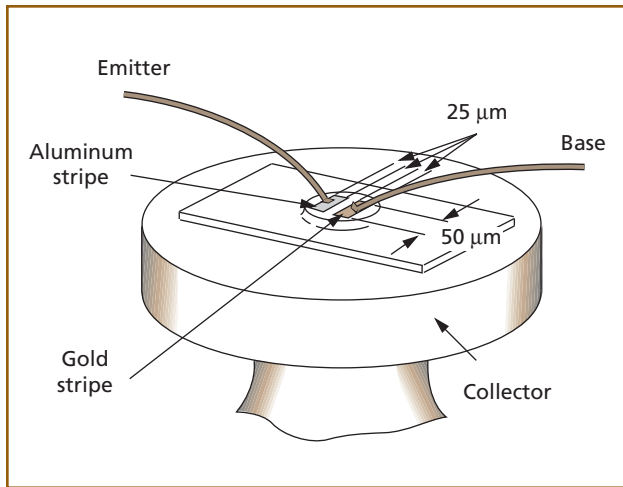


Figure 6.
Schematic of the first diffused-base transistor.

a depth of $1.5 \mu\text{m}$ and created an emitter region by alloying aluminum to a depth of $0.5 \mu\text{m}$, producing base thickness of about $1.0 \mu\text{m}$. This first diffused p-n-p transistor had a cut-off frequency of 500 MHz. A year later the first diffused silicon transistor was made. It had a frequency cutoff at 120 MHz.¹⁷

The speed problem was almost solved—but not quite. The frequency limitation had moved from the base region to the collector region. The collector had the highest resistivity of the three regions—an inevitable result of the additive nature of the diffusion process. This led to significant series resistance in the collector and that, combined with the capacitance of the collector junction, limited the frequency response.

The eventual solution was to add a totally different process, that of the epitaxial growth of a lightly doped layer of single crystal semiconductor on a substrate of a heavily doped single crystal—a process called epitaxial growth. A transistor base and emitter layer was then diffused into the epitaxial layer (see **Figure 6**). The results were published in June 1960 by Theuerer, Kleimack, Loar, and Christensen.¹⁸

Oxide Masking and Photolithography—Controlling the Surface Dimensions

In 1955, C. J. Frosch and L. Derick made a very important observation. They had been studying the pitting of the surface of silicon wafers during the diffusion process, and its dependence on the presence of oxygen. They discovered that a few-thousand-

angstrom layer of silicon dioxide grown on the surface prior to diffusion could mask the diffusion of certain donor and acceptor atoms into the silicon. They also demonstrated that diffusion would occur unimpeded through windows etched in the oxide layer.¹⁹ Shortly afterward, J. Andrus and W. L. Bond showed that certain photoresists deposited on the oxide surface would prevent etching of the oxide.²⁰ Hence optical exposure of the resist by projection or contact masks could be used to create precise window patterns in the oxide and in turn provide precise control of areas in which diffusion would occur.

Thus, in the course of a few weeks, four people had invented the complete process of oxide masking of diffusion and the application of photolithography to the precise control of the geometry of diffused regions. This was a natural batch process that has since been developed to the point that junction areas can be controlled to a fraction of a micron. This complemented the precision of the depth control of junctions diffused into the silicon surface providing the means to control the fabrication of silicon devices in three dimensions to the precision of a fraction of a micron.

It also ended the role of germanium as a major player. No material that would provide diffusion masking for germanium was found, and germanium became the niche material for specialty devices that rely critically on some special property.

The Reliability Problem

It was found in the early days that the transistor was very sensitive to its environment and particularly to humidity. This lack of reliability was a huge setback and embarrassment to the semiconductor community. The transistor had been lauded as a device with no failure mechanisms, with nothing to wear out. Instead it had a severe reliability problem, one that took almost twenty years to solve completely.

The immediate remedy was to hermetically seal the devices in packages using the metal to glass seals from vacuum tube technology. This was a further blow to the pride of the semiconductor engineer. The packaging art evolved using a variety of empirical procedures including vacuum baking, dry gas baking, and gettering. It is remarkable that with these unscientific approaches, germanium transistors were eventually

manufactured with failure rates less than 10 per billion operating hours.

Systematic studies to try and understand the problem and find a more fundamental solution were ongoing. At Bell Labs, Brattain continued his experimental work on surface states, as did Shive. M. M. Atalla and his development group studied the surface properties of silicon in the presence of a silicon dioxide layer. They speculated that growing an oxide layer under very clean and controlled circumstances on the surface of carefully cleaned silicon could lead to a reduced density of states at the silicon surface and might serve to protect the surface against further change. In 1959 they did confirm that the presence of an oxide layer could reduce the density of surface states to such a level that the field effect could be observed. However, they had difficulties gaining enough control of the process to get reproducible results. Nevertheless, the concept that an oxide layer might provide a solution to the reliability problem was a major step forward.²¹

The final breakthrough in the solution of the reliability problem came with an invention by J. A. Hoerni at Fairchild in late 1957 or early 1958. His idea was later reduced to practice and published in 1960.²² Hoerni proposed that, in the course of fabricating diffused silicon transistors, the silicon dioxide layer that was used as a diffusion mask be left in place. The junctions thus intersected the silicon surface under the oxide layer, and Hoerni speculated that that the oxide could protect the junction areas from contamination. He indeed found that such junctions had acceptable characteristics without further treatment. This was a startling result, particularly for those who believed that a passivating oxide would need to be grown under meticulously clean conditions.

This was not the end of the story, but the Hoerni result was a big step in the right direction. It was later found that not all “diffusion” oxides gave adequate initial performance and that all were subject to degradation with time. It was not until about 1966 that techniques were developed to produce satisfactory oxide layers and to “overcoat” them to retain their properties. Silicon devices then only needed to be further encapsulated in plastic for pro-

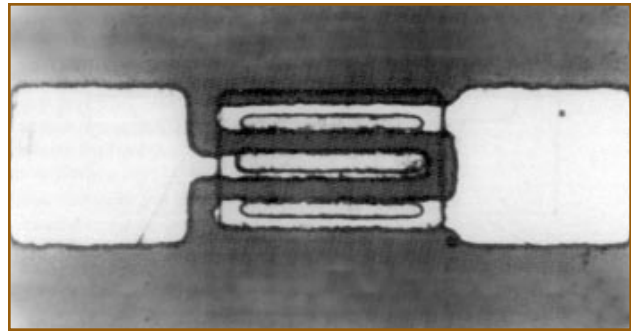


Figure 7.
A view of the surface of the planar transistor.

tection against gross environmental effects. Transistors, after all of twenty years, no longer looked like small vacuum tubes.

The Planar Transistor

In his 1960 paper, Hoerni also described the planar transistor. In this concept, both the base and emitter regions were diffused through windows in silicon dioxide masks so that both collector and emitter junctions terminated at the surface (see **Figure 7**). The masking oxides were left in place and provided protection and eventually passivation of the silicon surface. Ohmic contact was made to both emitter and base regions through windows in the oxide layer. It was noted that connection to the collector region could also be made on the top surface if that were desirable. The metal used for all contacts was aluminum, which G. E. Moore and R. N. Noyce had previously shown would make good contact to either n or p-type silicon.²³ Moore had also shown that the aluminum could be extended over the oxide to form larger pads to ease connections to the chip. Somewhat later the epitaxial process was added to the planar transistor to minimize collector resistance.

This structure brought it all together. All the key development and engineering problems were either solved or on course for an elegant solution. There was a sound foundation for the long-term manufacture of semiconductor devices. Silicon, the semiconductor of choice, could be produced with a crystalline perfection and purity more than adequate to the task. Critical dimensions in all three directions could if necessary be controlled to a fraction of a micron. Electrical contacts could be made with a single metal and without the need for microscopic precision. The resulting devices

would eventually be solidly reliable. And all this could be done with batch processing with the promise of high yield and low unit cost.

Now—some thirteen years after its discovery—the transistor had a sound engineering foundation. This provided the base for the next giant step. The integrated circuit was invented in 1958 by J. S. Kilby at Texas Instruments²⁴ with a major added contribution from Noyce at Fairchild.²⁵

What Made It Possible?

The pace of accomplishment in the early years of the development of transistor technology is remarkable. It took a mere two and a half years after the formation of the Shockley group to the invention of the point-contact transistor. In a period of only five years from the establishment of the group, the invention of the transistor was essentially complete, understood, and documented. After fifteen years, all the technology was in place to support the development of the integrated circuit with its spectacular rates of progress that continue to this day—fifteen years to provide the foundation for the silicon age.

It is instructive to speculate on what fundamental characteristics made this possible. I see four key elements.

The Search for Understanding

From its beginning, the exploration of the transistor was accompanied by a search for sound scientific understanding. Kelly set this direction by establishing a research group, albeit a group with a mission that contemplated important practical applications. This concept was reflected in the composition of the group he formed and particularly the three principal members. They strongly believed in the importance of basic understanding and avoiding the empirical approach.

This attribute remained with the industry. Even during the many years when empirical solutions were applied to the reliability problem, the search for a basic solution continued and eventually won out.

A Willingness to Share Information

The semiconductor industry operated then, as it does now, with an unusual willingness to share information. This of course derived from the special nature of AT&T as the manager of the Bell System.

The spirit of communication probably was further sustained by the institutional climate of Silicon Valley. Movement of key people among companies was so easy and occurred so often that open communication was inevitable.

Leadership

The leaders of the semiconductor industry largely came up through the technology side of the business. That is not to deny that some of the successful leaders had no technical experience and fared well. But their colleagues and competitors mostly had “silicon under their finger nails.”

I believe that the prevalence of technical knowledge in the leadership contributed to the remarkable success of the industry. It took a deep understanding of a complex technology to appreciate what present limitations were and to anticipate what improvements were likely to occur next. It also took deep understanding and confidence to make the large commitments to the creation of the next generation of the technology.

An Element of Luck

A number of events that made *the* difference in the transistor story must be recognized as being very fortunate. It was surely fortunate that both of the elemental semiconductors, germanium and silicon, were relatively easy to purify and produce as single crystals and had properties suitable for transistor action. The energy gaps, dielectric strength, minority carrier lifetimes, carrier mobility were all in favorable ranges. Had any one of these factors been one order of magnitude less favorable, the hurdles may well have been insurmountable.

The capabilities of silicon dioxide are also most fortunate. Silicon dioxide is an excellent insulator, makes a fine dielectric for a capacitor, masks diffusion and, as grown during the diffusion process, provides environmental protection to yield highly reliable devices.

We surely were incredibly lucky to find one material and its oxide that we could use to fabricate and encapsulate high-performance transistors and integrated circuits. But it also took more than luck. As Lee Trevino says about his golf, “The harder I practice, the luckier I get.” It took a lot of hard practice on the part of the scientists and engineers who created this tech-

nology to be smart enough to recognize and build on the luck that nature bestowed.

Acknowledgments

I have relied heavily on accounts in *A History of Engineering and Science in the Bell System, Electronics Technology (1925–1975)* published by Bell Labs, and particularly on the section entitled “The Transistor,” written by John Hornbeck and edited by Friedolf Smits. Friedolf was also helpful in clarifying some of the events covered in the text. I am also grateful to Gordon Moore for inputs on his contributions and the contributions of his colleagues. Bill Troutman of Lucent’s Microelectronic Group provided valuable reference material and even more valuable encouragement. I also wish to thank AT&T and Lucent for their support and particularly for providing material from their archives.

I recognize that in choosing the events that made the difference I have exercised judgment that is far from perfect. There were a multitude of contributions that I class as making a difference. Many of these were at least as meritorious and deserving of recognition as the ones I have described.

References

1. J. E. Lilienfeld, “Method and Apparatus for Controlling Electric Currents,” U.S. Patent 1,745,175, filed Oct. 8, 1926, issued Jan. 28, 1930; “Amplifier for Electric Currents,” U.S. Patent 1,877,140, filed Dec. 8, 1928, issued Sept. 13, 1932; “Device for Controlling Electric Current,” U.S. Patent 1,900,018, filed Mar. 28, 1928, issued Mar. 7, 1933.
2. J. Bardeen, “Surface States and Rectification at a Metal Semi-Conductor Contact,” *Physical Review*, Vol. 71, No. 10, May 15, 1947, pp. 717–727.
3. *A History of Engineering & Science in the Bell System, Physical Sciences (1925–1980)*, edited by S. Millman, AT&T Bell Laboratories, 1983, p. 98.
4. J. Bardeen and W. H. Brattain, “The Transistor, A Semi-Conductor Triode,” *Physical Review*, Vol. 74, No. 2, July 15, 1947, pp. 230–231.
5. *A History of Engineering & Science in the Bell System, Physical Sciences (1925–1980)*, p. 97.
6. *A History of Engineering and Science in the Bell System, Physical Sciences (1925–1980)*, p. 101.
7. W. Shockley, “The Theory of p-n Junctions in Semiconductors and p-n Junction Transistors,” *Bell System Technical Journal*, Vol. 28, No. 4, July 1949, pp. 435–489.
8. J. N. Shive, “Double-Surface Transistor,” *Physical Review*, Vol. 75, No. 4, Feb. 15, 1948, pp. 689–690.
9. G. K. Teal and J. B. Little, “Growth of Germanium Single Crystals,” *Physical Review*, Vol. 78, No. 5, June 1, 1950, p. 647.
10. W. Shockley, G. L. Pearson, and J. R. Haynes, “Hole Injection in Germanium—Quantitative Studies and Filamentary Transistors,” *Bell System Technical Journal*, Vol. 28, No. 4, July 1949, pp. 344–366.
11. W. Shockley, M. Sparks, and G. K. Teal, “p-n Junction Transistors,” *Physical Review*, Vol. 83, No. 1, July 1, 1951, pp. 151–162.
12. W. Shockley, *Electrons and Holes in Semiconductors*, D. Van Nostrand, Princeton, New Jersey, 1950.
13. J. E. Saby, “Fused Impurity P-N-P Transistors,” *Proceedings of the IRE*, Vol. 40, November 1952, pp. 1358–1360.
14. *A History of Engineering and Science in the Bell System, Physical Sciences (1925–1980)*, p. 582.
15. G. K. Teal, “Single Crystals of Germanium and Silicon—Basic to the Transistor and Integrated Circuit,” *IEEE Transactions on Electron Devices*, Vol. ED-23, No. 7, July 1976, pp. 621–639.
16. C. A. Lee, “A High-Frequency Diffused Base Germanium Transistor,” *Bell System Technical Journal*, Vol. 35, No. 1, January 1956, pp. 23–34.
17. M. Tanenbaum and D. E. Thomas, “Diffused Emitter and Base Silicon Transistors,” *Bell System Technical Journal*, Vol. 35, No. 1, January 1956, pp. 1–22.
18. H. C. Theuerer, J. J. Kleimack, H. H. Loar, and H. Christensen, “Epitaxial Diffused Transistors,” *Proceedings of the IRE*, Vol. 48, September 1960, pp. 1642–1634.
19. C. J. Frosch and I. Derick, “Surface Protection and Selective Masking During Diffusion in Silicon,” *Journal of the Electrochemical Society*, Vol. 104, September 1957, pp. 547–552.
20. J. Andrus and W. L. Bond, “Photoengraving in Transistor Fabrication,” in *Transistor Technology*, Vol. III, edited by F. J. Biondi, D. Van Nostrand, Princeton, New Jersey, 1958, pp. 151–162.
21. M. M. Atalla, E. Tennenbaum, and E. J. Scheiber, “Stabilization of Silicon Surfaces by Thermally Grown Oxides,” *Bell System Technical Journal*, Vol. 28, No. 4, May 1959, pp. 749–783.
22. J. A. Hoerni, “Planar Silicon Diodes and Transistors,” *IRE Transactions on Electron Devices*, Vol. 8, Mar. 1961, p. 178.
23. G. E. Moore and R. N. Noyce, “Method for Fabricating Transistors,” U. S. Patent 3,108,359, filed June 30, 1959, issued Oct. 29, 1963.
24. J. S. Kilby, “Invention of the Integrated Circuit,” *IEEE Transactions on Electron Devices*, Vol. ED-23, No. 7, July 1976, pp. 648–654.

25. R. N. Noyce, "Semiconductor Device-and-Lead Structure," U. S. Patent 2,981,877, filed July 30, 1959, issued April 25, 1961.

(Manuscript approved September 1997)



IAN M. ROSS was president of Bell Labs from 1979 to July 1991, when he was named President Emeritus of Bell Labs. He holds bachelor's, master's and Ph.D. degrees in electrical engineering from Cambridge University in England. At Bell Labs, Dr. Ross spent twelve years in the research and development of semiconductor devices. During that time, he served as director of the Semiconductor Laboratory and also as director of the Semiconductor Device and Electron Tube Laboratory. He then served as managing director of Bellcomm, the Bell System company that provided systems engineering support for the Apollo manned space-flight program, and later became president of Bellcomm. He subsequently returned to Bell Labs and held several key positions before becoming president of Bell Labs. Dr. Ross is a member of several academies, including the IEEE (Fellow), the National Academy of Engineering, the National Academy of Sciences, and the Royal Academy of Engineering in the United Kingdom. A recipient of numerous awards, such as the IEEE Founder's Medal, the American Electronics Association's Medal of Achievement, and the Semiconductor Industry Association's Robert Noyce Award, he chaired, from 1988 to 1992, the National Advisory Commission on Semiconductors established by Congress and the President. Dr. Ross's current board memberships include the Science & Technology Advisory Board - Taiwan, R. O. C. (Chair); National Science Board; and Board of Directors, NACCO Industries, Inc., B. F. Goodrich, and Thomas and Betts. ♦