Chapter 1

Minority Carriers and the First Two Transistors

by Michael Riordan & Lillian Hoddeson[1]

The first transistor—the "point-contact transistor," invented at Bell Labs in December 1947 by John Bardeen and Walter Brattain—had almost no commercial impact in the decade following its invention. Only one firm seriously attempted to produce such a transistor in quantity—the Western Electric Company, the manufacturing division of American Telephone and Telegraph, which began using the device in such telephone applications as direct dialing as early as 1952.[2]

It was the second transistor—the "p-n junction transistor," which William Shockley conceived in January 1948—that became "the first technologically important device of the solid state era," as he put it.[3] The successful manufacture of the junction transistor and its offspring stimulated the exponential growth of the semiconductor industry. By contributing crucially to the rise of such industry giants as Texas Instruments, SONY and Fairchild Semiconductor during the explosive fifties, the junction transistor made Silicon Valley possible.

Given this impact, it is surprising how little scholarship has been devoted to the history of the junction transistor, and how many important historical issues about the device have not even been explored.[4] At the 1992 Indianapolis meeting of the American Physical Society, Nick Holonyak pursued one of them—the scientific understanding of "minority carrier injection," a physical process that was critical to the invention of the junction transistor.[5] Whereas Shockley claimed sole authorship of this idea,[6] Holonyak argued on the basis of several pieces of indirect evidence that Bardeen deserves the credit.[7] In this paper we reexamine the controversy about the origins of this concept, using more immediate documentary evidence from laboratory notebooks, patent applications, scientific articles written in 1947–48, and interviews conducted with participants.[8]

The Invention of the Point-Contact Transistor

The roots of the junction transistor—in which n- type and p-type layers of semiconductors are sandwiched together alternately—cannot be separated from those of the point-contact transistor—in which the crucial interfaces are where two sharp metal points contact a slab of semiconductor.[9] Both inventions emerged from the Bell Labs program of basic research on solid state physics that Mervin Kelly, then executive vice president, initiated in 1945. Within this new program, Shockley proposed in April 1945 the device now called the "field-effect transistor." Here an externally applied electric field is arranged so that it can increase or decrease the number of charge carriers in a thin film of semiconductor, thus altering its conductivity and regulating the current

flowing through it. By applying suitable voltages to two circuits passing through the semiconductor, Shockley predicted that an input signal applied to one circuit could yield an amplified signal in the other.

Unfortunately, all the attempts to fabricate Shockley's field-effect device failed[10] as did Shockley's theoretical attempt to explain why. On the basis of the available semiconductor rectification theory by Nevill Mott and Walter Schottky, his conceptual field-effect device did not work as predicted.[11]

In October 1945 John Bardeen joined the new semiconductor group that was being formed under Shockley within Bell's larger solid state department. Shockley asked Bardeen to check calculations that he had made to examine why attempts to demonstrate his design for a field-effect amplifier had failed. By March 1946 Bardeen had an answer. He explained the lack of significant modulation of the conductivity using a creative heuristic model, the idea of "surface states."[12] In this model, electrons drawn to the semiconductor surface by the applied field become trapped in the postulated states and are thus unable to act as charge carriers. As Shockley later put it, the surface states "blocked the external field at the surface and . . . shielded the interior of the semiconductor from the influence of the positively charged control plate."[13]

But were these postulated states real? And if so, how did they generally behave? These questions became of primary interest to the semiconductor group, which in the following months responded to Bardeen's surface state theory with an intensive research program to explore these states. Bardeen worked closely on the problem with the group's experimental physicists, Walter Brattain and Gerald Pearson.[14]

Brattain made a discovery on 17 November 1947. Drawing on a suggestion by Robert Gibney, a physical chemist in the semiconductor group, he found he could neutralize the field-blocking effect of the surface states by immersing a semiconductor in an electrolyte.[15] "This new finding was electrifying," observed Shockley. "At long last, Brattain and Gibney had overcome the blocking effect of the surface states."[16] It set in motion the events that would culminate one month later in the first transistor.

Four days after this discovery, Bardeen and Brattain tried to use the results to build a field-effect amplifier. Their approach was based on Bardeen's suggestion to use a point-contact electrode pressed into a specially prepared silicon surface. Rather than use the thin films employed in the 1945 experiments by Shockley and his collaborators, Bardeen proposed the use of an *n*-type "inversion layer" a few microns thick that had been chemically produced on the originally uniform surface of *p*-type silicon. (The existence of such inversion layers—of *n*-type semiconductor on a *p*-type substrate, or vice versa—had been demonstrated by earlier experimental studies at Bell Labs by Russell Ohl and Jack Scaff). Because charge carriers—in this case, electrons—would have a much higher mobility in such an inversion layer than they had in the vapor-deposited films employed in the 1945 experiments, Bardeen believed this approach would work better in a field-effect amplifier.[17] In particular, this layer would act as a narrow channel in which the population of charge carriers could be easily modulated by an applied external field.[18] The device tested on November 21 used a drop of electrolyte on the surface as one contact and a metal point as the other; Bardeen and Brattain obtained significant power amplification, but the frequency response was poor.[19]

The next crucial step occurred on December 8. At Bardeen's suggestion, Brattain replaced the silicon with an available slab of *n*-type, "high-back-voltage" germanium, a material developed during the wartime radar program by a research group at Purdue directed by Karl Lark-Horovitz.[20] They obtained a power gain of 330—but with a negative potential applied to the droplet instead of positive, as they had expected. Although the slab had not been specially prepared, Bardeen proposed that an inversion layer was being induced electrically, by the strong fields under the droplet. "Bardeen suggests that the surface field is so strong that one is actually getting P type conduction near the surface," wrote Brattain that day, "and the negative potential on the grid is increasing the P type or hole conduction."[21]

Later that week Brattain evaporated a gold control plate onto a specially prepared germanium slab (which therefore *had* an inversion layer, a priori); he was trying to improve the frequency response by eliminating the droplet. He thought the gold would be insulated from the germanium by a thin oxide layer, but through a failure in his procedure it was instead directly in contact with the semiconductor surface. This oversight proved to be a critical step toward the point-contact transistor.[22]

The following Monday, December 15, Bardeen and Brattain were surprised to find that they could still modulate the output voltage and current at a point contact positioned close to the gold plate, however only when the plate was biased *positively*—the opposite of what they had expected![23] "An increase in positive bias *increased* rather than decreased the reverse current to the point contact," wrote Bardeen almost ten years later. This finding, as he explained in his 1956 Nobel lecture, suggested "that holes

5

were flowing into the germanium surface from the gold spot and that the holes introduced in this way flowed into the point contact to enhance the reverse current. This was the first indication of the transistor effect."[24] Although Brattain and Bardeen failed to observe any power amplification with this configuration, Bardeen suggested that it would occur if two narrow contacts could be spaced only a few mils apart.[25] Brattain achieved the exacting specifications by wrapping a piece of gold foil around one edge of a triangular polystyrene wedge and slitting the foil carefully along that edge. He then pressed the wedge—and the two closely spaced gold contacts—down into the germanium surface using a makeshift spring (See Figures 1 and 2). In their first tests, on December 16, the device worked just as expected. It achieved both voltage and power gains at frequencies up to 1000 Hz. The transistor had finally been born. A week after that, on 23 December 1947, the device was officially demonstrated to Bell Labs executives in a circuit that allowed them to hear amplified speech through a pair of headphones.[26] (See Figure 3 for Brattain's record of this event in his notebook).

The Flow of Charge Carriers

The crucial issue for us here is, how did Bardeen and Brattain conceptualize the flow of charge carriers while they were developing the first transistor? Memory is imperfect, and later accounts are often subject to "retrospective realism," a term coined by historian and sociologist of science Andrew Pickering to describe the process whereby conjectures are later imbued with an aura of certainty, or embellished with details that became known only at a later

time.[27] Fortunately, we have available several telling entries Bardeen, Brattain and Shockley made in their laboratory notebooks during those pivotal weeks before and after Christmas 1947.[28]
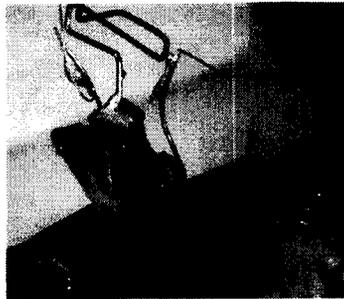

Figure 1: Photograph of the point-contact transistor invented by Bardeen and Brattain in December 1947. A strip of gold foil slit along one edge is pressed down into the surface of a germanium slab by ˜ polystyrene wedge. (Reprinted with the permission of AT&T Archives)
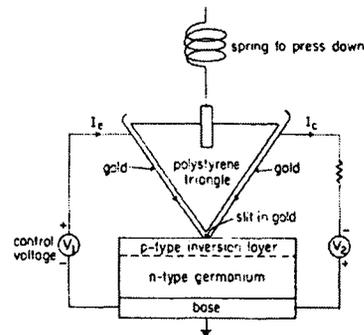
Figure 2: Schematic diagram of the first transistor (Figure 1). The signal current I_e flows through the input circuit, generating holes in a p-type inversion layer that modulate the flow of current I_c in an output circuit. (Reprinted from L. Hoddeson, "The Discovery of the Point-Contact Transistor")

On December 19, three days after the first successful test of their device, Brattain wrote: "It would appear then that the modulation obtained when the grid point is bias + is
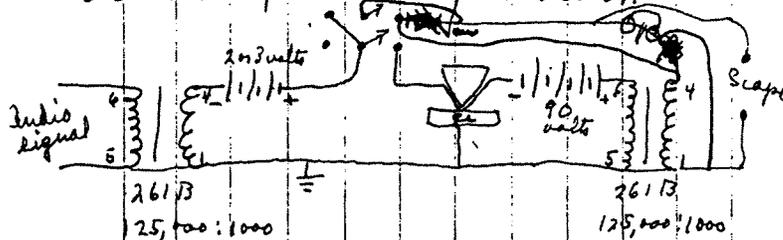
Figure 3: Entry in Brattain's lab notebook describing the circuit used in the 23 December 1947 demonstration of the point-contact transistor to Bell Labs executives. (Reprinted with the permission of AT&T Archives)

due to the grid furnishing holes to the plate point."[29] By grid point and plate point, he was referring to what we now call the emitter and collector—he was obviously using a familiar vacuum tube analogy. Although we cannot determine from this passage how he conceived the details of their flow, we can be certain he understood that positively charged quantum-mechanical entities were the charge carriers responsible for modulation.

Bardeen gave a more detailed explanation in a notebook entry dated 24 December 1947, the day after the team made its official demonstration. After describing the setup, which used a slab of n-type germanium specially prepared to produce a very thin inversion layer of p-type conductivity near its surface, he portrayed the phenomenon as follows (see Figure 4):

> When A is positive, holes are emitted into the semi-conductor. These spread out into the thin P-type layer. Those which come in the vicinity of B are attracted and enter the electrode. Thus A acts as a cathode and B as a plate in the analogous vacuum tube circuit.[30]

Again it was clear that Bardeen also attributed the transistor action to the holes, basing his model on the same vacuum-tube analogy Brattain used. But he went a step farther and stated that the flow of these holes occurred within the inversion layer.

This emerging theory of the transistor based on the flow of holes at or near the surface of the germanium developed further during the following six months, the period in which Bell Labs kept the discovery of the transistor "laboratory secret," while patent applications were being drawn up. A drawing found in Bardeen and Brattain's

Figure 4: Entry in Bardeen's lab notebook dated 24 December 1947, giving his conception of how the point-contact transistor functions. (Reprinted with the permission of AT&T Archives)

patent application of 17 June 1948 (revised from a version submitted on February 26) suggests that although the flow of charge carriers was thought to occur largely within the *p*-type inversion layer, they were by this time allowing that some holes might diffuse into the body of the *n*-type germanium (see Figure 5). The text of their application states:

> . . . potential probe measurements on the surface of the block, made with the collector disconnected, indicate that *the major part* of the emitter current travels on or close to the surface of the block, substantially laterally in all directions away from the emitter 5 before crossing the barrier 4.[31]

Oct. 3, 1950   J. BARDEEN ET AL   2,524,035
THREE-ELECTRODE CIRCUIT ELEMENT UTILIZING
SEMICONDUCTIVE MATERIALS
Filed June 17, 1948   3 Sheets-Sheet 1
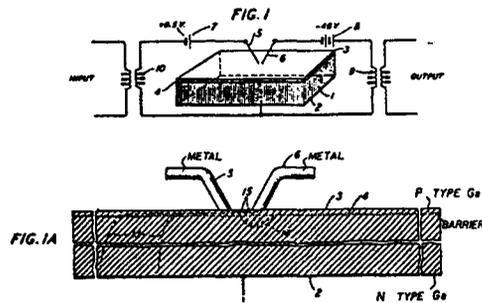
Figure 5: Figures 1 and 1A from Bardeen and Brattain's U.S. patent (No. 2,524,035) on the point-contact transistor. The dominant flow between emitter and collector occurs through a *p*-type inversion layer on a slab of *n*-type germanium (U.S. Patent Office).

    In their famous letter submitted to the *Physical Review* on June 25 1948, they wrote that as a result of the existence of the thin *p*-type inversion layer next to the germanium surface, "the current in the forward direction with respect to the block is composed in large part of holes,

i.e., of carriers of sign opposite to those normally in excess in the body of the block."[32] In a subtle shift from their earlier conception, they envisioned that holes flow predominantly in the *p*-type inversion layer, but with a portion that can also flow through the *n*-type layer beneath it.

It is not clear from these entries just how and why this shift occurred. But both the revised patent application and the *Physical Review* letter are dated well after Shockley's disclosure of the junction transistor in late January and a crucial mid-February experiment (discussed below) by John Shive.

Research at Purdue

Bell Labs was not the only institution at which researchers were interested in the behavior of germanium semiconductors. During the wartime efforts to develop crystal rectifiers for radar receivers, Lark-Horovitz's research group at Purdue had greatly advanced under-standing of the element germanium. As part of this study, the group explored a phenomenon known as the "spreading resistance" of germanium—the resistance to the current flow that diverges from a metal point-contact pressed into the surface of the semiconductor. Lark-Horovitz asked graduate student Ralph Bray to study this spreading resistance, in part to determine whether the local resistivity near a point contact matched the bulk resistivity of the entire sample.[33]

The Purdue group continued this work after the war. In early 1947 Bray noticed that a spreading resistance much *lower* than predicted by theory occurs when high field pulses of positive voltage were applied through a metal wire to certain samples.[34] At the time, he did not realize that the

effect was being caused by holes injected into the n-type germanium at the point where the wire touched the germanium. "The spreading resistance was sort of a mystery and nobody understood it," he admitted much later. But after reading Bardeen and Brattain's transistor paper, he could finally explain everything they had seen by the production of holes. "I guess we were on the verge of it in thinking about such [a possibility]," Bray said, "but it hadn't really formulated itself in our minds."[35]

Even so, Brattain and Bardeen were concerned the Purdue group might beat them to the discovery of the transistor. They were especially nervous during the period of laboratory secrecy.[36] Jane Bardeen recalls that "John gave her hell" for mentioning that John was working on semiconductors in a letter to a friend who was acquainted with people in the Purdue group. He felt her friend could possibly "let the cat out of the bag."[37]

Similarly, Brattain found himself in an awkward position when he heard Seymour Benzer, another member of the Purdue group, mention the spreading resistance in late January 1948 at the New York American Physical Society meeting. He understood by then why the resistivity was decreasing near the point—that it was the result of hole emission at the point. Brattain recalled listening quietly to Benzer in the corridor, until Benzer remarked, "I think if somebody put another point contact down on the surface, close to this point, and measured the distribution of potential around the point, then we might be able to understand what this [effect] is about." To that Brattain replied, "Yes, I think maybe that would be a very good experiment" and walked away.[38]

The Conception of the Junction Transistor

In the weeks that followed the invention of the point-contact transistor, Shockley was torn by conflicting emotions. Although Bardeen and Brattain's invention had been a "magnificent Christmas present" to Bell Labs, he was chagrined that he had not had a direct role in this crucial breakthrough. "My elation with the group's success was tempered by not being one of the inventors," he recalled. "I experienced frustration that my personal efforts, started more than eight years before, had not resulted in a significant inventive contribution of my own."[39]

Since the failure of his field-effect idea more than two years earlier, Shockley had paid only passing attention to semiconductor research. During the several months before the invention, he had mainly been working cn the theory of dislocations in solids, a problem he had become interested in after attending a conference on the topic the previous August. Shockley had, however, devoted some thought to the physics of *p-n* junctions and their use in such practical devices as lightning arresters and high-speed thermistors.[40]

Brattain and Gibney's discovery in November 1947 galvanized him into action. "The breakthrough observation of 17 November that surface-states could be overcome," he later wrote, "stimulated *the will to think*—and act—in minds conditioned to search for semiconductor amplifiers."[41] A few days later he suggested fabricating an amplifier using a drop of electrolyte deposited across a *p-n* junction in silicon or germanium; this approach worked when Brattain and Pearson tried it.[42] On 8 December 1947, more than a week before the point-contact transistor was

invented, Shockley described an idea in his laboratory notebook for an *n-p-n* sandwich that had current flowing laterally in the *p*-layer and with the *n*-layers acting as control electrodes.[43]

The 16 December invention of the point-contact transistor and Bardeen's interpretation of its action in terms of the flow of holes also stimulated Shockley's thinking. Bardeen's above-quoted analogy with the operation of a vacuum tube—in which the current carriers were holes instead of electrons—was in fact due to Shockley,[44] who applied it in his first attempt at a junction transistor, written in a room in Chicago's Hotel Bismarck on New Year's Eve of 1947. A letter he wrote to his mother a few days earlier aboard the New York Central's 20th Century Limited captures some of the flavor of his frenetic activity:

> I shall attend the Physical Society Meeting Tues and Wed and then dig in at a Hotel and try to write some articles until about Monday or Tuesday when I shall come out to give a lecture at the Institute for the Study of Metals at the University of Chicago.[45]

Holed up in his hotel room, Shockley recorded his ideas on a pad of paper, pages of which he mailed to Bell Labs, where they were witnessed by Bardeen and group co-leader Stanley Morgan before being pasted into his lab notebook. In this first stab at a junction transistor, one can see a clear analogy with a vacuum tube; its "control" electrode acts as a grid to control the flow of holes from a "source" to a "plate" (see Figure 6).[46] About this disclosure of a *p-n-p* device, Shockley admitted that he "failed to recognize the possibility of minority carrier injection into a base layer. . . . What is conspicuously lacking [in these pages] is any suggestion of the possibility that holes might
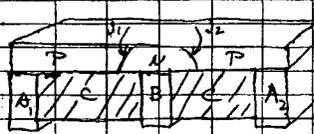
Figure 6: Page in Shockley's lab notebook showing one of his first attempts at a junction transistor. This entry was written by Shockley in Chicago on a piece of paper that he later (after Bardeen and Morgan had witnessed it) pasted into his notebook. (Reprinted with the permission of AT&T Archives)

be injected into the *n*-type material of the strip itself, thereby becoming minority carriers in the presence of electrons."[47]

A little more than three weeks later, this time working at his home in Madison, New Jersey, Shockley conceived another design in which *n*-type and *p*-type layers were reversed and electrons rather than holes were the current carriers (see Figure 7). Applying a positive potential to the *p*-layer should then cause holes to flow into it and thereby lower its potential for electrons; this he realized would "increase the flow of electrons over the barrier exponentially."[48] As Shockley noted in an account written almost thirty years later, this *n-p-n* device finally contained the key concept of "exponentially increasing minority carrier injection across the emitter junction."[49] Minority carriers, in this case the electrons, had to flow in the presence of the dominant majority carriers, in this case the holes of the *p*-type layer. Yet at the time he did not seem to have been terribly excited with the significance of this invention, given that it was written on Friday, 23 January 1948, and witnessed four days later by his assistant John R. Haynes. "If I had *really* appreciated the impact that the junction transistor would have," he wrote over a quarter century later, "I would have driven the few miles necessary to obtain a witnessing signature the *same* day."[50]

A Crucial Experiment

It is a curious fact that almost another month went by before Shockley revealed his breakthrough idea to anyone in his group other than Haynes. It is especially curious given that Bardeen and Morgan witnessed Shockley's earlier junction transistor design. Why did Shockley keep the
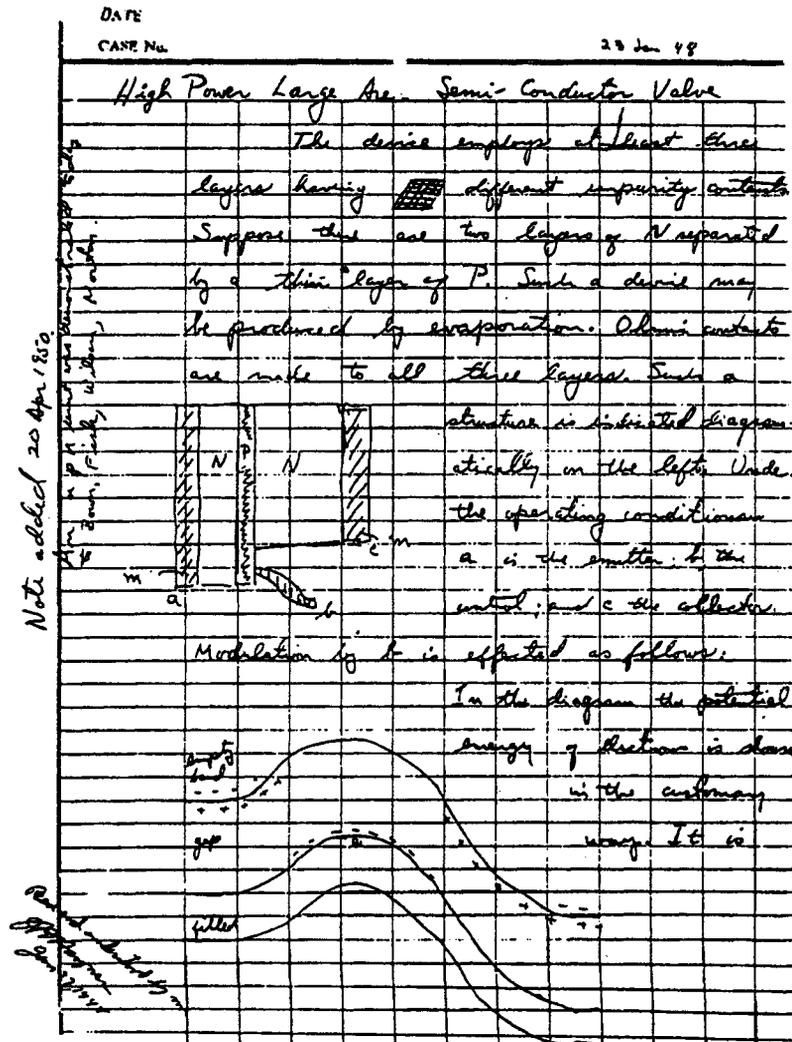
Figure 7: Entry in Shockley's lab notebook dated 23 January 1948 giving his disclosure of the junction transistor. He wrote this page at home on a piece of paper and later (after J. R. Haynes had witnessed it) pasted it into his notebook. (Reprinted with the permission of AT&T Archives)

information to himself? Did he recognize that he had made a major conceptual breakthrough but decide to keep it quiet to give himself the additional time to follow up its theoretical and practical ramifications? Was he afraid Bardeen and Brattain were so close to making the discovery themselves that knowledge of his idea would push them to publish before him?[51] Or was Shockley simply so unsure of the work that he avoided showing it to colleagues until he could think about it further? Unfortunately, documentary evidence provides no firm answer to these questions.

It *is* true that in order to function, Shockley's *n-p-n* device required something more than the point-contact transistor did. To work, Shockley's device needed new physics—namely, it was crucial to understand that minority carriers are in fact able to diffuse *through* the base layer in the presence of majority carriers. In late January such behavior was not obvious from the work of Bardeen and Brattain, who at that time were preoccupied with preparing patent documents dealing with the point-contact device. (The very next week, in fact, was the New York APS meeting at which Brattain had his discussion with Benzer.) As noted above, Bardeen and Brattain still believed then that essentially all the hole flow occurred in a micron-deep *p*-type layer at the semiconductor surface.

Evidence for the required diffusion of the minority carriers into the bulk material was not long in coming. In a closed meeting at Bell Labs on February 18, John Shive revealed that he had just tested a successful point-contact transistor using a very thin wedge of *n*-type germanium, but with the emitter and the collector placed on two opposite faces of the wedge (see Figure 8).[52] At the place where the two contacts touched it, the wedge was only 0.01 cm thick,

while the distance between them along the germanium surface was much larger. Shockley immediately recognized what this revelation meant. In this geometry, the holes were flowing by diffusion in the presence of the majority carriers, the electrons in the n-type germanium, through the bulk of the semiconductor; they were not confined to an inversion layer on the surface, as Bardeen and Brattain had been suggesting occurred in their device. Shockley recounted that "as soon as I had heard Shive's report, I presented the ideas of my junction transistor disclosure and used them to interpret Shive's observation."[53]
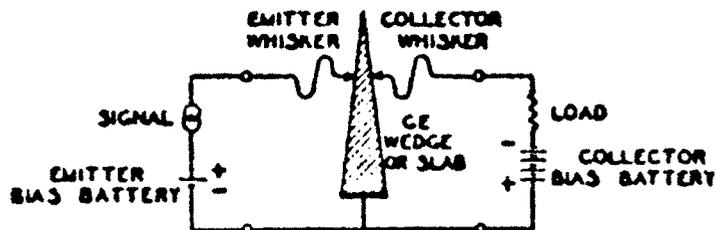
Figure 8: Schematic diagram of the "double-surface" transistor first demonstrated by John Shive on 18 February 1948. Holes must flow by diffusion through the bulk of the germanium—and not along the surface—in order for this device to work. (Reprinted from Shive, "The Double-Surface Transistor.")

The most likely (but still not certain) interpretation of this sequence of events is that Shockley felt he needed to establish that minority carriers could indeed diffuse through the bulk of the semiconductor, something that was by no means obvious in late January 1948. Shive's serendipitous

experiment, done without any knowledge of Shockley's January 23 disclosure, gave him the evidence he required. The experiment also may have injected a heady dose of urgency into the solid state physics group. On February 26, Bell Labs applied for a total of four patents on semiconductor amplifiers, including Bardeen and Brattain's original patent application on the point-contact transistor. Their two landmark papers, titled "The Transistor, a Semi-Conductor Triode" and "Nature of the Forward Current in Germanium Point Contacts," went off to the *Physical Review* four months later, on 25 June 1948. One day after that, Bell applied for Shockley's patent on the junction transistor. And then on June 30 it announced the invention of the transistor in a famous press conference.

In July, Shockley proved that hole "injection" (as he dubbed the flow of minority carriers in transistor action) was indeed occurring in n-type germanium. Working with Haynes, he showed that the charge carriers traveling from emitter to collector were in fact "positive particles with a mobility of about 1.2×10^3 cm^2/volt-sec."[54] Their paper was published in early 1949 immediately after Shive's paper on the two-sided transistor.

Shockley had another blind spot to overcome in his thinking about the minority carriers before it finally became possible to fabricate functioning junction transistors using his ideas. One of the problems behind the failure of his field-effect transistor had been the slow rate at which charge carriers diffused through the polycrystalline silicon and germanium films used in the experiments. Gordon Teal, one of Shockley's colleagues at Bell Labs, realized the merits of using single crystals of germanium and eventually silicon. Teal knew by this time that in polycrystalline films these

carriers do not survive long enough to make it from the emitter to the collector in sufficient numbers and that minority carriers had lifetimes 20 to 100 times longer in single crystals. He tried to convince Shockley of this critical advantage of single crystals, but Shockley ignored the suggestion and was satisfied to use polycrystalline films of germanium.[55]

Fortunately Jack Morton, who headed Bell Lab's efforts to develop the point-contact transistor into a commercially viable product, took Teal's suggestion seriously and late in 1949 gave him a modicum of support to pursue this avenue of research. Working with Morgan Sparks, Teal modified the crystal-making machine that he and a colleague John Little had recently developed for pulling single crystals out of molten germanium. The modification allowed doping the germanium in a controlled manner and made it possible for Sparks and Teal to fabricate the first practical *n-p-n* junction transistor in April 1950.[56] On the date of its demonstration to Bell Labs executives, 20 April 1950, Shockley penned a note in the margin of his 23 January 1948 disclosure (see Figure 7), stating, "An *n-p-n* unit was demonstrated today to Bown, Fisk, Wilson, Morton."[57]

Conclusion

From this examination of the different ways in which the inventors of the first two transistors interpreted the flow of charge carriers in these devices, it seems clear that Bardeen made the fundamental breakthrough in establishing that minority carriers—in this germanium case, holes—are responsible for transistor action. One might quibble that

until Shive's experiment he had an imperfect understanding of the detailed geometry of the flow. And one might argue that holes trickling through a *p*-type inversion layer, as Bardeen at first conceived their flow, are not really "minority carriers" *in that layer,* where the dominant charge carriers are in fact holes.[58] Still, Bardeen's realization of the crucial role the holes were playing was the "great leap forward" that made semiconductor amplifiers a reality. It was no small leap of the imagination in those postwar years to realize that holes—"carriers of sign opposite to those normally in excess in the body of the block"[59]—could serve this function. Had the Purdue physicists made the same interpretation of the spreading resistance they had observed almost a year earlier, they might well have beaten Bell Labs to the transistor.

In 1980 Bardeen reflected in an interview on the difference between Shockley's view and his in the interpretation of transistor action:

> The difference between ourselves [Bardeen and Brattain] and Shockley came in the picture of how the holes flow from the emitter to the collector. They could flow predominantly through the inversion layer at the surface, which does contain holes. And the collector would be draining out the holes from the inversion layer. They could also flow through the bulk of the semiconductor, with their charge compensated by the increased number of electrons in the bulk . . . .[60]

After acknowledging Shive's "definitive experiment," which "showed definitely that the holes were flowing through the bulk of the germanium,"[61] he noted:

> . . . what you were modulating was the conductance of minority carriers in the inversion layer. And we probably had

> that too much on our minds, when we were interpreting the
> first results of the point-contact transistor, in that we knew
> that there was an inversion layer there which would provide a
> channel for holes going from the emitter to the collector. The
> question is, how important that aspect was, compared with the
> flow of the holes through the bulk, which Shockley called
> "injection" to differentiate it from the word "emitter," which
> we'd already introduced.[62]

Shive's experiment—and that of Haynes and Shockley later that year—resolved this question. Note Bardeen's emphasis above on the semantic difference between hole injection and hole emission. According to Bardeen, Shockley introduced the word "injection" because he wanted a different word to describe the flow of minority carriers in the bulk of the semiconductor.

This is not to diminish Shockley's important work in discerning a new physical process necessary to design semiconductor amplifiers that could be easily mass-produced with relatively stable and predictable properties. And he came up with the junction transistor idea almost a month *before* Shive's pivotal experiment demonstrated that minority carriers did indeed diffuse through the bulk material. Bardeen and Brattain made the fundamental scientific breakthrough by recognizing that holes, not electrons, were the crucial charge carriers involved in their transistor. Shockley took their work a major step further, providing a more detailed understanding of the flow of minority carriers—which could be *either* holes or electrons. Their combined breakthroughs in semiconductor physics are what opened the road to Silicon Valley.

1   An early version of this paper was presented by one of us (Riordan) at the 1994 Pittsburgh meeting of the American Physical Society. It is based on research done in preparation for writing our book *Crystal Fire: The Birth of the Information Age* (New York: W.W. Norton & Co., 1997).

2   W.J. Pietenpol, "Bell System and Commercial Applications of Transistors," Manuscript, 17 June 1958, AT&T Archives.

3   W. Shockley, "The Path to the Conception of the Junction Transistor," *IEEE Transactions on Electron Devices* ED-23, no. 7 (July 1976): 597-620, hereafter cited as "Path." Quote on p. 597.

4   Other than a few papers written by Shockley himself (which hardly qualify as objective scholarship), there is nothing to compare with the extensive works on the origins of the point-contact transistor. W. Shockley, "Path," 597-620; "The Invention of the Transistor: An Example of Creative-Failure Methodology," *Proceedings of the Second European Solid State Device Research Conference* (London: Institute of Physics, 1973): 55-75; and "The Invention of the Transistor—An Example of Creative-Failure Methodology," *Proceedings of the Conference on the Public Need and the Role of the Inventor,* Monterey, California, 11–14 June 1973 (Washington, National Bureau of Standards Special Publication No. 388, May 1974), 47-89. Andrew Goldstein describes the role of Gordon Teal and his techniques for growing single germanium crystals in fabricating the first grown-

junction transistor in "Finding the Right Material: Gordon Teal as Inventor and Manager," in *Sparks of Genius: Portraits of Electrical Engineering Excellence,* ed. Frederick Nebeker (New York: IEEE Press, 1993), 93-126. See also Ernest Braun, "Selected Topics from the History of Semiconductor Physics and its Applications," in *Out of the Crystal Maze,* eds. L. Hoddeson, et al. (New York: Oxford University Press, 1992), 443-488.

5   The meeting was devoted to the work of the late John Bardeen. N. Holonyak, "John Bardeen and the Point-Contact Transistor," *Physics Today* 45, no. 4 (April 1992): 36-43. The detailed microscopic understanding of the flow of minority carriers in semiconductors became an important factor allowing the subsequent miniaturization of transistors as components of integrated circuits.

6   W. Shockley, "Path," 615.

7   Holonyak's argument is based largely on (1) a statement Bardeen made about his role in 1990, in a filmed interview for Japanese television; (2) Holonyak's vivid memory and informed interpretation of a comment Bardeen made in a graduate seminar Holonyak attended in the spring of 1952 (that if Walter Schottky had asked what the holes were doing in his rectification studies, the transistor would have been invented then); (3) a credit Shockley gave to Bardeen and Brattain in the preface of his widely used textbook, *Electrons and Holes in Semiconductors, with Applications to Transistor Electronics* (Princeton: D. Van Nostrand,

1950), and a handwritten inscription by Shockley in Bardeen's copy of the book, acknowledging Bardeen's important role in the developments treated; and (4) Holonyak's personal experience of Bardeen's generous nature and modesty in presenting his contributions.

8   All interviews cited here are (or will shortly be) on deposit at the American Institute of Physics Center for History of Physics, College Park, Maryland (hereafter "AIP").

9   Historical accounts of the point-contact transistor include: L. Hoddeson, "The Discovery of the Point-Contact Transistor," *Historical Studies in the Physical Sciences* 12, no. 1 (1981): 41-76, hereafter cited as "Point-Contact"; C. Weiner, "How the Transistor Emerged," *IEEE Spectrum* 10 (1973): 24-33; E. Braun and S. MacDonald, *Revolution in Miniature: The History and Impact of Semiconductor Electronics* (Cambridge: Cambridge University Press, 1978); John Bardeen, "Semiconductor Research Leading to the Point-Contact Transistor," *The Nobel Lectures, 1942–1962* (Amsterdam: 1964), 313-386; and Walter Brattain, "Genesis of the Transistor," *The Physics Teacher* 6 (1968): 108-114. A much fuller treatment can be found in Michael Riordan and Lillian Hoddeson, *Crystal Fire: The Birth of the Information Age* (New York: W.W. Norton & Co., 1997).

10   The tests were conducted in May–June 1945 by J. Hayes, H. McSkimin, W. Yager, and Russell Ohl.

11   For more detail, see Hoddeson, "Point-Contact," 62-63.

12 J. Bardeen, Bell Labs Notebook (BLNB) No. 20780, 18 March–23 April 1946, 38-57, AT&T Archives; "Surface States and Rectification at a Metal to Semiconductor Contact," *Physical Review* 71 (1947): 717-727.

13 W. Shockley, "Path," 605.

14 The full program is described in the laboratory notebooks of the members of the group and summarized in "Point-Contact," 64-66. The sequence of steps to the point-contact transistor detailed in this article follow, for the most part, the account in "Point-Contact." It was to become a lifelong pattern for Bardeen to work closely with experimentalists and only on rare occasions—such as in his work on the theory of superconductivity—with other theorists.

15 W. Brattain, Bell Labs Notebook (BLNB) No. 18194, 17–29 November 1947, 142-151, AT&T Archives; "Genesis of the Transistor," *Physical Review* 72 (1948): 345; interview with Brattain by A.N. Holden and W.J. King (AIP).

16 W. Shockley, "Path," 608.

17 Bardeen, "Semiconductor Research," 318-341.

18 In addition, using such a layer circumvented the considerable practical difficulty of depositing a good thin layer of semiconductor material. Bardeen interview by L. Hoddeson, 13 February 1980 (AIP).

19 Bardeen, BLNB No. 20780, entries of 22 November 1947, 61-67 and 23 November, 68-70.

20 The best treatment of this Purdue program is in Paul
    Henriksen, "Solid State Physics Research at Purdue,"
    Osiris 3, Ser. 2 (1987): 237-260. For a vivid account
    of how Lark-Horovitz worked with students in this
    group see Helmut Fritzsche interview by L. Hoddeson,
    31 March 1987 (AIP).

21 Brattain, BLNB No. 18194, 175-76. This is the first
    recorded instance we can find in which Bardeen and
    Brattain recognized the possibility that holes were acting
    as charge carriers. Note that Bardeen still proposed that
    the flow occurred within a narrow inversion layer at the
    semiconductor surface.

22 L. Hoddeson, "Point Contact," 71-72.

23 L. Hoddeson, "Point-Contact," 72, which indicates that
    this event occurred on Thursday, 11 December. A
    closer examination of Brattain, BLNB No. 18194,
    183-92 indicates that there was a period of confusion
    followed by the actual breakthrough on 15 December.
    See Riordan and Hoddeson, *Crystal Fire,* Chapter 7,
    for a more complete discussion of this sequence of
    events.

24 J. Bardeen, "Semiconductor Research Leading to the
    Point Contact Transistor," 338.

25 W. Brattain, BLNB No. 18194, 16 December 1947,
    192-193; Brattain interview by Holden and King (AIP).

26 L. Hoddeson, "Point-Contact," 74.

27 Andrew Pickering, *Constructing Quarks: A
    Sociological History of Particle Physics* (Edinburgh:

University Press, 1984), 7.

[28] It is interesting that some of the entries in Brattain's notebook during those critical weeks in December 1947 are in Bardeen's hand. The two obviously were working side-by-side in the laboratory.

[29] W. Brattain, BLNB No. 21780, 19 December 1947, 3.

[30] J. Bardeen, BLNB No. 20780, 24 December 1947, 72.

[31] J. Bardeen and W. Brattain, "Three-Electrode Circuit Element Utilizing Semiconductive Materials," U.S. Patent No. 2,524,035, Application 17 June 1948, Serial No. 33,466, Sheet 1, Figure 1. See column 10, lines 64-70. (Emphasis added.)

[32] J. Bardeen and W.H. Brattain, "The Transistor, A Semi-Conductor Triode," *Physical Review* 74 (1948): 230-31, on 230.

[33] Paul W. Henriksen, "Solid State Physics Research at Purdue," 253-4.

[34] It depended inversely on both the magnitude of the field and the duration of the pulse. R. Bray, interview by P. Henriksen, 14 May 1982 (AIP).

[35] Bray interview by Paul Henriksen, quoted in Henriksen, "The Emergence of Solid State Physics Research at Purdue University During World War II," Manuscript, University of Illinois, 1983.

[36] J. Bardeen interviews by L. Hoddeson, 1 Dec. 1977, 14 April 1978, and 13 Feb. 1980 (AIP).

[37] Holonyak interview by M. Riordan and L. Hoddeson,

30 July 1993 (AIP).

[38] This exchange recounted in W. Brattain interview by C. Weiner, 28 May 1974 (AIP), 28.

[39] W. Shockley, "Path," 612.

[40] W. Shockley, Bell Labs Notebook (BLNB) No. 20455, 71, 76-78, 80, 88-89, AT&T Archives.

[41] W. Shockley, "Path," 608. (Emphasis in the original.)

[42] W. Brattain, BLNB No. 18194, 169-70; G. Pearson, Bell Labs Notebook (BLNB) No. 20912, 75, AT&T Archives.

[43] W. Shockley, BLNB No. 20455, 91.

[44] W. Shockley, BLNB No. 20455, 106-109; Bardeen credits him with this suggestion in his BLNB No. 20780, 72.

[45] Letter from W. Shockley to M.B. Shockley postmarked 30 December 1947, Shockley Archives, Box 7, Folder 3, Stanford University Special Collections.

[46] W. Shockley, BLNB No. 20455, 110-113.

[47] W. Shockley, "Path," 613-14.

[48] W. Shockley, BLNB No. 20455, 128-132.

[49] W. Shockley, "Path," 614.

[50] W. Shockley, "Path," 615. (Emphasis in the original.)

[51] Holonyak gives a reasonable argument that once the point-contact transistor and the notion of minority carrier flow were on hand, the *p-n* junction transistor was "bound to follow." He recalls a dinner in the mid-

'80s in Urbana at which Bardeen stated that he and Brattain had planned to move on to that device as soon as they completed their taxing work preparing patents for the original transistor, only to find that Shockley had already tied up this area of work with the Bell Laboratories patent attorney (who, in John's words, "was in Shockley's pocket.") Holonyak interview by Hoddeson on 10 January 1992.

52 J.N. Shive, "The Double-Surface Transistor," *Physical Review* 75 (1949): 689-90. See also Shive, Bell Labs Notebook (BLNB) No. 21869, 30-35, AT&T Archives.

53 W. Shockley, "Path," 618.

54 J.R. Haynes and W. Shockley, "Investigation of Hole Injection in Transistor Action," *Physical Review* 75 (1949): 691.

55 Goldstein, "Finding the Right Material," 104. See also Gordon K. Teal and John Little, "Growth of Germanium Single-Crystals," *Physical Review* 78 (1950): 647 (reported at a meeting held 16–18 March 1950).

56 Goldstein, "Finding the Right Material," 104-5.

57 W. Shockley, Notebook No. 20455, 128. We tell the full story of these developments in our book *Crystal Fire*.

58 In a paper about Bardeen delivered in 1992 at Indianapolis, Conyers Herring however stated that the current in the point-contact transistor was "primarily

carried by holes, that is, *minority carriers,* moving into the semiconductor rather than by electrons moving out (emphasis added)." See C. Herring, "Recollections from the Early Years of Solid State Physics," *Physics Today* 45, no. 5 (April, 1992): 26-33, on 31.

[59] J. Bardeen and W. Brattain, "The Transistor," 230.

[60] Interview of Bardeen by L. Hoddeson, 13 February 1980, 2.

[61] Interview of Bardeen, 4.

[62] Interview of Bardeen, 12.