Great Experiments in Physics

1. Discovery of Transistor Effect that Changed the Communication World

Amit Roy

Amit Roy is at the Nuclear Science Centre, New Delhi building a superconducting linac booster for the Pelletron accelerator. He spent over two decades at the Tata Institute of Fundamental Research, Mumbai investigating nuclei using accelerators and probing symmetries in physics. His hobbies are books and music.

Transistors have revolutionised the electronics and communications industries. An account of the discovery of the transistor effect is given along with the experiments leading to the discovery.

Today, irrespective of whichever remote corner of the earth you may find yourself in, even on the moon or in outer space, you can be sure of one thing, that is you can establish contact with your people either through radio or phone. The tremendous explosion of information technology owes its genesis to the discovery of the transistor. No other electronic device has influenced modern life more fundamentally than the transistor. It is a ubiquitous element of electronic circuits ranging from the simplest amplifier or radio to complicated electronic computers. Though integrated circuits have lately replaced circuits with discrete transistors in many areas, the former are merely arrays of transistors and other components built from a single chip of semiconductor material.

The transistor effect was discovered by three American physicists, William Shockley, John Bardeen and Walter Brattain, working at Bell Labs in 1947. Were they looking for such a device? It certainly was a goal for them but they arrived at the discovery quite unexpectedly while trying to understand the fundamental behaviour of semiconductor surfaces (see Suggested Reading). They received the Nobel Prize in Physics for this discovery in 1956.

It is interesting to consider the experiments and the studies that preceded this discovery. The rectifying property of a junction between metals and sulphides of lead and iron (galena and pyrites) was discovered by F Brown in 1874 and later in contacts made between copper and copper oxide by A Schuster. By 1904 point contacts on galena, silicon carbide, tellurium, silicon, etc. were found to be well suited for the detection of radio waves which were demonstrated a few years earlier in 1888 by H Hertz. In India, Jagadis Chandra Bose while working on microwaves had obtained a US patent on one such device based on galena in 1904. To understand the action of semiconductors we need to know the quantum theory of solids. Such a theory was developed in 1931 by A H Wilson and it has formed an integral part of semiconductor physics.

To understand the action of semiconductors we need to know the quantum theory of solids.

However, with the advent of vacuum tubes, the interest in such point contact devices diminished. By 1945 semiconductors were being used as diodes, varistors and thermistors. There was a hope of making an amplifying device with a semiconductor, as a triode. Following the analogy of the vacuum tube triode, it was thought that if a grid could be placed in the space charge layer at the contact, it would be possible to control the flow of electrons across the contact. However, a space charge layer is only about one mm in width and introducing a grid within such a narrow space is quite difficult. R Hilsch and R W Pohl succeeded in 1938 in demonstrating the validity of this principle by inbuilding a triode in an alkali halide crystal in which the width of the space charge layer was about 1 cm. This device could amplify frequencies less than 1 Hz and certainly was not a practical device.

In 1946, a Bell laboratory group headed by William Shockley started work on properties of semiconductors, mainly silicon and germanium. By this time they were in a position to make polycrystalline ingots of either n- or p-type silicon or germanium of specified resistance. Shockley had suggested an alternative way for amplification of AC signals. This was to control the conductance of a thin film or a slab of semiconductor by the application of a transverse electric field. The slab would form one electrode of a parallel plate condenser, the other plate being the control electrode as sketched in *Figure 1*. The idea was to apply a voltage across the condenser thus inducing

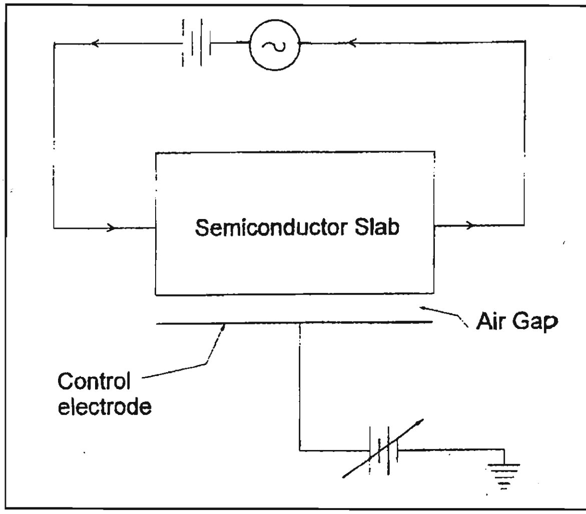


Figure 1. Sketch of the first experimental arrangement designed to look for the field effect predicted by Shockley.

charges in the slab. If the induced charges are mobile carriers then the conductance would vary with change of voltage on the control electrode. Experiments were tried in earnest to observe this field effect predicted by Shockley, but the effect observed was several orders of magnitude less than predicted, and several other predictions of the theory did not agree with the experimental results. The reason for this failure was attributed to the electron localization at the surface, trapping a large fraction of the charge carriers. The field effect suggested by Shockley had to wait for improvements in semiconductor technology and it was possible later to make electronic amplifiers with high gain using the field effect principle. While the evidence on surface states was fairly convincing, it was all of an indirect nature. Shockley, Brattain and Bardeen carried out a number of experiments that could yield more concrete evidence about the surface barrier. Shockley's prediction that a difference in contact potential would be found between *n*-and *p*-type specimens with large impurity concentration was observed. Brattain argued that illuminating the surface region would excite electrons and holes which would be separated by the field due to surface charges and thus result in a change of the surface charge and contact potential. Brattain used the apparatus shown schematically in *Figure 2* to measure the change in contact

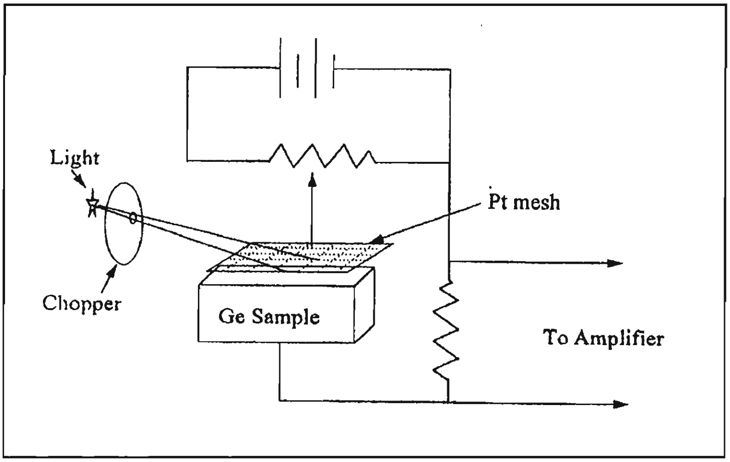


Figure 2. Brattain's experiment to measure changes in contact potential due to light.

potential. The reference electrode was made of platinum in the form of a screen so that light could pass through it. Chopped light falling on the surface with the electrode held fixed, generated an alternating voltage across the condenser. Brattain tried several ambient atmospheres and different temperatures and observed large effects when a liquid dielectric filled the space between the electrode and the semiconductor surface. He along with a colleague, Gibney introduced different electrolytes and observed the effects attributed to large changes in the surface barrier with voltage applied across the electrolyte. Evidently ions piling up at the surface created a very large field which penetrated through the surface states.

The use of electrolytes to change the surface barrier led to another approach to observe the field effect predicted by Shockley. Bardeen and Brattain took a slab of p-type Si which was treated to give an n-type surface and made a point contact by pushing a metal point onto this surface. They covered the metal with a thin layer of wax and surrounded it with a drop of water, which acted as an electrolyte. An electrode dipped in the electrolyte was used to apply a strong electric field in the vicinity of the contact. The experimental arrangement is shown in *Figure 3*. If the point is biased positive with respect to the block, it is the reverse bias or high resistance direction and a part of the reverse current consists of electrons flowing through the

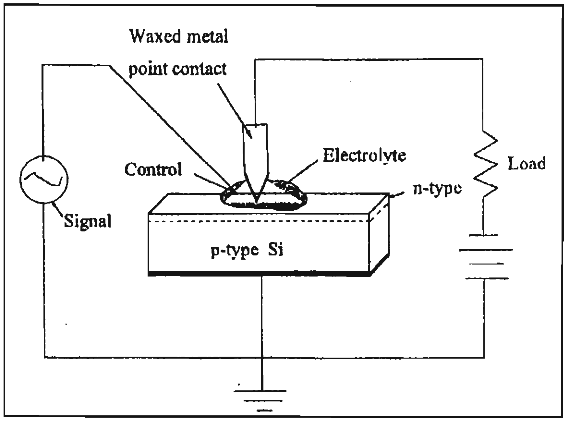


Figure 3. Schematic view of the apparatus used by Brattain and Bardeen to observe the field effect.

n-type inversion layer to the contact. It was found that applying a voltage on the electrolyte probe resulted in changing the conductance of the inversion layer, and could also change the magnitude of the current. Since under static condition only small current flows through the electrolyte, the set up could be used as an amplifier. In the initial tests current and power amplification was achieved but no voltage amplification was observed. A negative voltage on the probe was found to decrease the current flowing in the reverse direction to the contact.

The same experiment was repeated with a block of n-type germanium. Since water would evaporate quickly, they changed the electrolyte to glycol borate. A positive change in the voltage on the probe decreased the reverse current and considerable voltage and current amplification was observed and the experiment showed that a large part of the reverse current consisted of holes flowing in an inversion layer near the surface. Since this electrolyte has long time constants, amplification occurred for only low frequencies, less than 8Hz. A metal control electrode insulated from the surface by a thin oxide layer then replaced the electrolyte. The surface was first anodised and a thin layer of gold was then evaporated on several spots on the anodised surface. A point contact was placed very close to one of the spots and biased in the reverse direction. A small effect on the reverse current was observed when the spot was biased positively, but was opposite to that observed with an electrolyte. An increase in positive bias increased rather than decreased the reverse current to the point contact. The effect was large enough to give some voltage, but no power amplification. This experiment suggested that holes 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, where the current is being carried predominantly by minority carriers. Power amplification was observed by placing metal contacts at a distance of 0.005 cm from each other by evaporating gold on a wedge of polystyrene and then separating the gold at the point of the wedge with a razor blade to make two closely placed contacts (See *Figure 4*). The close separation was later achieved by using two appropriately shaped point contacts placed close together. This was the birth of the point contact transistor. Using this transistor Brattain had set up a circuit to amplify speech and observed a distinct gain in speech level which could both be heard as well as seen on the oscilloscope, with no noticeable change in quality. The device could also be made to act as an oscillator the next day. J R Pierce, another famous colleague of Brattain at the Bell laboratories, coined the name transistor. The discovery of the transistor was publicly announced in three contiguous letters to the *Physical Review*

J R Pierce, another famous colleague of Brattain at the Bell laboratories, coined the name transistor.

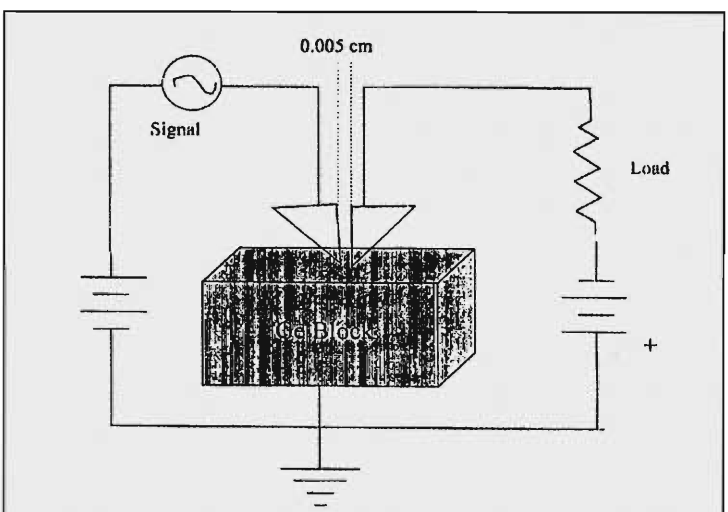
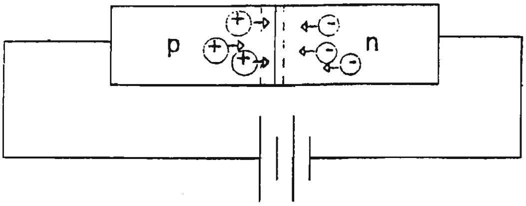
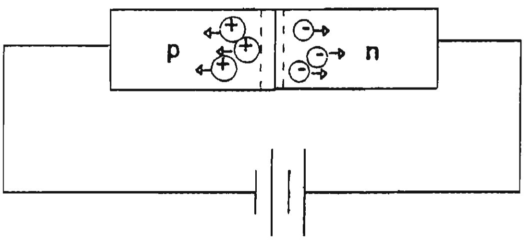


Figure 4. Sketch of the arrangement in which the full transistor effect was first observed .

Box 1



Forward Biased Junction



Reverse Biased Junction

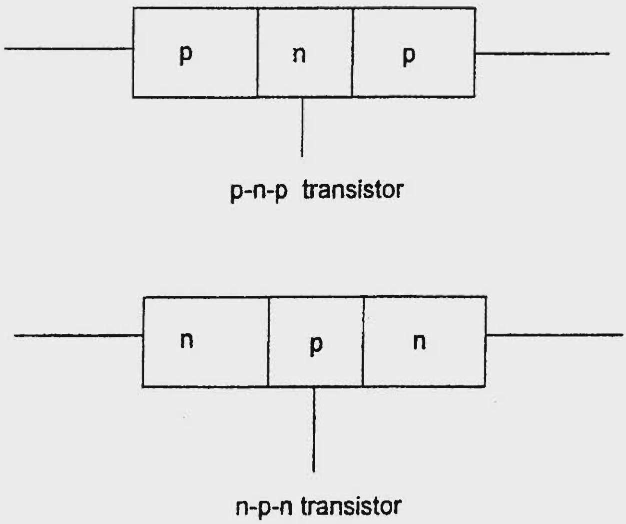
Schematic representation of a p-n junction diode in forward and reverse bias.

Semiconductors have a resistance in between that of a good conductor (metals) and an insulator. The conductivity of semiconductors (normally group IV elements like Si, Ge) can be altered by introducing as impurities either group III elements (B, Al, Ga, In) giving rise to *p*-type semiconductors or group V elements (P, As etc.) giving rise to *n*-type semiconductors. In the *p*-type semiconductor the bulk of the current is carried by holes (majority carriers are positively charged) and the electrons are the minority carriers, whereas in *n*-type semiconductors the situation is the reverse. When a *p*-type and an *n*-type semiconductor are brought together a space charge layer is formed at the junction due to diffusion of holes from the *p*-side to the *n*-side and that of electrons from the *n*- to *p*-side. If a voltage is applied across a *p-n* junction, the current flow across the junction depends on the polarity of the voltage applied. If positive voltage is applied to the *p*-type side, the junction is forward biased and the majority carriers contribute the current. A negative voltage on the *p*-type side results in a much lower current carried by the minority carriers and the junction is reverse biased. Such a *p-n* junction thus acts as a diode, which can be considered as a two-terminal device.

several months later ([3],[4]). Subsequently Shockley went on to work out the theory of the *p-n* junction (see *Box 1*) and predicted the *p-n-p* and *n-p-n* junction transistors (see *Box 2*), which were made by M Sparks at Bell laboratories.

Further developments in the understanding of semiconductors and improvements in manufacturing techniques led to the development of a new family of solid state devices. This in turn has brought about a remarkable technology called microelectronics which has given us the tiny integrated circuits which then shrank into the large scale integrated circuits (LSI) and then further into the very large scale integrated (VLSI) circuits.

Box 2



Schematic representation of a p-n-p and an n-p-n transistor.

A point contact diode evolved out of the devices where phosphor bronze whiskers were pressed onto crystals of galena, pyrites, quartz etc. to demodulate the amplitude-modulated radio waves. The junction transistor is an extension of point contact diodes in the sense that two such point contacts are made on a semi-conductor surface. The transistor is a three terminal device consisting of two *p-n* junctions, each of which has the characteristics of a diode. The three terminals are connected to the three sections of the transistor designated for historical reasons – emitter, base and collector. There exist two types of transistors, one in which the emitter and collector are *n*-type and the base is *p*-type and the second in which the emitter and collector are *p-type* whereas the base is *n*-type. These are called *n-p-n* and *p-n-p* transistors, respectively. For a good account of the transistor and its properties, see [1]. Of the three terminals, the current between any two terminals can be controlled by means of small voltages on the third terminal and power amplification results.

Suggested Reading

[1] **D N Bose. Transistors, from Point to Single Electron. *Resonance*, Vol.2, No. 12, 1997.**

[2] **J Bardeen and W H Brattain. The Transistor, a Semi-Conductor Triode. *The Physical Review*. 74, 230, 1948.**

[3] **W H Brattain. Nature of the Forward Current in Germanium Point Contacts. *The Physical Review*. 74, 232, 1948.**

[4] **W H Brattain. Genesis of the Transistor. *Physics Teacher*. 6,112, March 1968.**

Address for Correspondence
Amit Roy
Nuclear Science Centre
Aruna Asaf Ali Marg
P.O. Box 10502
New Delhi 110 067
Email: roy@nsc.ernet.in

As long as men are free to ask what they must, free to say what they think, free to think what they will, freedom can never be lost and science can never regress.

J Robert Oppenheimer