The Semiconductor Story

1: The new crystal triode

by K. J. Dean, M.Sc., Ph.D., and G. White†, M.Phil., B.Sc.*

The paper which first announced the discovery of the transistor appeared in the *Physical Review* in July 1948. To commemorate the 25th anniversary of this event, *Wireless World* is publishing a series of four articles presenting a critical survey of the semiconductor industry, past and present, from the U.K. point of view. Part 1 describes the early development of germanium diodes and transistors, while parts 2 and 3 describe respectively the exploitation of the transistor and the integrated circuit to the present day. The final part discusses some of the problems, both technical and commercial, which have faced the industry in recent years. The roles of careful research, happy chance, technical skill and industrial pressure make a facinating story of our times.

The new crystal triode, as the transistor was first called, seemed in 1948 to be poor competition for the Goliath sized valve manufacturing industry. But a veritable David it turned out to be! *Wireless World* reported the discovery in an article in October 1948, entitled "The Amplifying Crystal". How many people reading that report then realized its implications for the future? The transistor was the end result of research which started 140 years ago in 1833 with Michael Faraday. He noted that while most conductors have a positive temperature coefficient of resistance, a substance called silver sulphide had a negative coefficient. Thus a substance later to be classed as a semiconductor was identified. Rectification, photoconductivity and photo-e.m.f. effects were all observed before 1900. Theoretical work on semiconductors after Faraday's original discovery gathered momentum, so that, by the early 1930s, quantum mechanics was applied to the theory of conduction. Energy band diagrams, electrons and holes then started to be discussed. The stage was set for the discovery in America by J. Bardeen and W. H. Brattain of the transistor—a semiconductor triode. This was the first three terminal semiconductor device which could amplify, and that was only 25 years ago. Now the impact of the transistor is universal, it has applications ranging from aviation and broadcasting to washing machines and Xerography.

Cat's whiskers

Semiconductor crystals were used in the early days of radio communications, the crystal rectifier being used as the detector in radio receivers. A typical detector was made by soldering or clamping a minute piece of the crystal in a small brass cup and the point contact made with a flexible wire called the cat's whisker, which was held in light contact with the crystal. The discovery of the thermionic triode by Lee de Forest in 1907, and its subsequent developments, made the crystal rectifer obsolete in radio receivers. However, the point contact crystal could not be replaced for detecting and monitoring u.h.f. power. At the other end of the scale, at low frequencies the copper oxide rectifier and selenium rectifier have been commercially successful but they are however not point contact rectifiers The rectification property of these is obtained by the contact of a thin film of semiconductor with the metal on which it is deposited. They are therefore termed contact rectifiers.

Wartime research

The second World War, like all military ventures, provided the cash to oil the wheels of research, so important at times of national emergency. It saw the development of radar, which gave a great impetus to u.h.f. crystal rectifier design. Research was concentrated on using silicon, germanium and boron. Boron prepared with selected impurities, i.e. "doped", showed sufficient conductivity to be of interest, but its typical characteristic curve was S shaped and symmetrical about the origin, thus the project was then dropped. Silicon showed great promise, being used for most of the commercially available devices. At this time the importance of starting with extremely pure silicon was appreciated. The "red-dot" crystal diode developed by the General Electric Company, for example, was derived from silicon crystals prepared from melts made from highly purified silicon powder, to which was added a fraction of a per cent of aluminium and beryllium. The resulting crystal could dissipate relatively large amounts of power without appreciably impairing its performance as a mixer. These were therefore known as "high-burnout" crystals.

The method of adjusting the cat's whisker at this time is interesting to note. The contact pressure was increased until a predetermined characteristic was obtained, and the cartridge was then tapped with a light mallet. Careful tapping caused the forward resistance to drop and the reverse resistance to rise. The cartridge was then impregnated with wax to provide mechanical stability and to make it impervious to water. Further work in 1943 led to high purity silicon, doped with only 0.001% boron, which produced an extremely good device and made prolonged tapping unnecessary. The small amount of the impurity needed indicates how material technology had to keep pace with the demands of the semiconductor device manufacturer. At this time, work on germanium led to the high-inverse voltage rectifier; so called because it could withstand up to 100V applied in the reverse direction. The doping agent used was tin, although it was found that similar effects could be obtained with some other elements. Germanium, however, could not compete with silicon above 30MHz. These methods of preparing the germanium crystal and polishing its surface were to be used later in the manufacture of the first transistor.

In 1946 H. Q. North showed that the point-contact used in these devices could be welded to the crystal surface by passing a high density current (in the order of 10^7 amps/sq. in) for a short time through the contact point. Although this did not improve their performance little was lost either. This technique too was later to be of value in three-terminal point contact devices.

Post war development

After World War II the immediate problems of survival gave place to the interests of commercial enterprise, and researchers were able to return to more general semiconductor problems, although under industrial patronage. Silicon and germanium were chosen for the research effort because they are simpler to understand than most other semiconductors. A lot of expertise on these materials had been accumulated during the war, particularly in America. Fig. 1 shows the structure of silicon or germanium crystals. Each atom has four neighbours, all

---

*South East London Technical College.
†Twickenham College of Technology.

at the same distance from it, and all at equal distance from each other. Each atom and one of its neighbours is attached by an electron pair bond, which consists of sharing two electrons to form a stable bond. Each atom has four electrons available to form bonds (valence electrons), therefore the conditions are exactly right for the diamond structure of Fig. I.

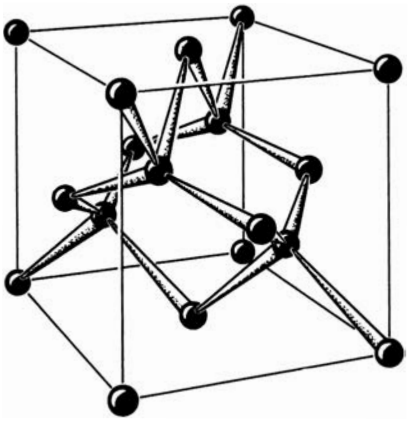


Fig. 1. The crystal structure of germanium and silicon.

The electronic properties are also dependent upon the electrons present in the bonding. By introducing impurities into the crystal the bonding can be modified. Therefore the electronic properties can be tailored as required by the controlled addition of impurities. The unoccupied bonds on the extreme edge of a perfect crystal cannot be used by internal atoms, but they are capable of accepting electrons. These are called acceptor or surface states. Crystal defects and absorbed foreign atoms will have similar effects and also create surface states. It was the thorough investigation of these states that led to the somewhat accidental discovery of the transistor effect. It is strange that surface states are now something to be avoided in transistor manufacture, because they would provide a low impedance path to current flow that is controlled inside the material.

Amplification using semiconductors was first achieved by using the negative resistance characteristic of thermistors. As the current through the thermistor increased, the heat generated caused a reduction in the resistance, and hence a drop in the voltage. The frequency of operation is limited by the temperature which has to follow the current changes. However, by making the physical dimensions small and the thermal conductivities high, oscillations of up to 100kHz have been produced. Bell Telephone Laboratories' aim after the war was to produce a purely electronic, rather than thermal, semiconductor amplifier. The work was initiated by W. Shockley who directed work on investigating the modulation of the conductance of a thin film of semiconductor. The conductance was controlled by an electric field applied by an electrode insulated from the film. It was hoped that the conductance would be modified by changes in the surface states caused by the applied field. The experiment gave disappointing results, since only about 10% of the expected change in conductance occurred. The effect was explained by J. Bardeen who in 1947 proposed a double layer at the surface, formed by the charge in the surface states and the induced space charge. Further research was carried out to measure the characteristics of the surface states.

The transistor discovered

The effect of having the crystal surface immersed in a liquid was studied. The characteristics of a high-inverse voltage germanium rectifier with a field applied by an electrolyte were investigated by J. Bardeen and W. H. Brattain. They proposed that a portion of the current was being carried by holes flowing near the surface. When the electrolyte was replaced with a metal object, transistor action was discovered. The discovery was first published as a short letter to the editor of the *Physical Review* journal in July 1948. This marked the beginning of the transistor era. A more detailed paper was published in the following year.

The transistor is a semiconductor triode The prefix "trans" designates the translational property of the device, while the root "istor" classifies it as a circuit element in the same general family with resistor, varistor, and thermistor. The transistor was commercially made in a similar form to the point contact diode, except for a second cat's whisker mounted very close to the first. The device is shown schematically in Fig. 2. A germanium ingot was prepared in the same manner as that used for the high inverse voltage diodes, and then a slice of this ingot was ground flat on both sides. The slice was copper-plated and tinned on one side, and diced into small squares with a diamond wheel. One of these squares was then sweated onto the brass base plug and the germanium surface treated. The unit was force fitted into a cylindrical cartridge, which had been shaped to accept the contact assembly. The contacts consisted of two 0.005in phosphor bronze wires, which had been bevelled and polished.

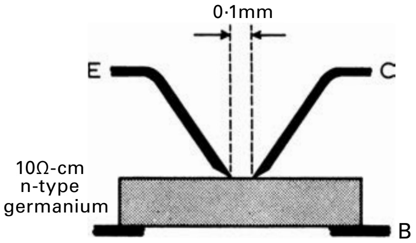


Fig. 2. Schematic of the point contact transistor.

The characteristics of the thermionic diode and the semiconductor diode are fairly similar, and methods of adding a "grid" to control the current in the forward direction as had been achieved with the triode, were looked at. The transistor, however, is not operated in this quadrant, because the output is reverse biased in the high resistance direction. The current is enhanced and controlled by the forward biased emitter contact. This device was designated the type A transistor to distinguish it from possible future varieties. The transistor effect is the injection of holes into the n-type material by the emitter, which are collected as an increment of the collector current. The common terminal called the base electrode is physically the base of the crystal. Devices which operate on different principles, such as the field effect, have since been called transistors. Therefore, transistor electronics is used generally to describe the art of controlling electron movements in a solid, hence is sometimes called solid state electronics. One of the first point contact transistors to be manufactured in the United Kingdom is illustrated. The patent numbers show the advantage of a strong development facility by using experience gained in the construction of point contact diodes to help in the manufacture of transistors. Patent number 591092, which was applied for in 1945, describes a method for holding the contact in place after construction. This is achieved by filling the cartridge with a wax-like substance which will harden on heating. The other patent number, 592659, was applied for in 1941, and deals with the preparation of the crystal and the subsequent treatment of its surface. The germanium had to have a spectroscopic purity of 99.95% for good results.

The G.E.C. crystal triode type GET 1, one of the earliest point contact transistors to be made in the U.K. The reverse of the packet, shown here with the transistor, carried a warning "To prevent permanent damage to the triode, it is recommended that whenever possible d.c. limiter resistors be placed in series with both emitter and collector ... Great care should always be taken to connect supplies of the correct polarity to the electrodes."

Transistor amplifiers

The journal Audio Engineering published an article in August 1948 entitled "Experimental Germanium Crystal Amplifier", only one month after Bardeen and Brattain's original letter. This described how to construct a germanium crystal amplifier—such was the rate of progress even in 1948. The article highlights the similarity between point contact diodes and the type A transistor because the construction starts with two diodes. They are dismantled and the crystal used, with the two whiskers carefully adjusted on the surface. Difficulty was experienced in finding active spots, due to the relatively impure crystals being used at that time. Manufacturers were aware of the need for high quality germanium. In 1964 the first extraction plant in the United Kingdom was built at Brimsdown for Johnson Matthey for the bulk production of germanium and other semiconductor materials.

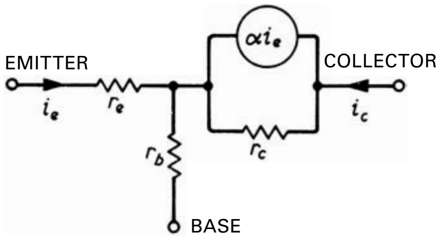


Fig. 3. Equivalent Tee circuit of a transistor.

The type A transistor can be represented by the equivalent circuit shown in Fig. 3, with the following average values for its for parameters:

emitter resistance $r_e = 240\Omega$

base resistance $r_b = 290\Omega$

collector resistance $r_c = 19k\Omega$

amplification factor $a = 1.8$

Unfortunately the active area of the device is very small and hence the collector dissipation is only about 0.2W, although a power gain of 17dB with a power output of 5mW was achieved. The small size of the device, however, gives it a wide frequency response, with an upper limit of approximately 10MHz. It was soon noted that the transistor could be greatly improved by passing large reverse currents through the collector point. This technique, called forming, resulted in amplification factors as high as 5. This process was explained by the formation of a p-n hook at the collector which reduced the height of the potential energy hill at the collector, so allowing a considerable increase in the number of electrons diffusing from the collector into the floating p region.

The movement of holes was thought to be mainly confined to the surface region but in 1949 J. N. Shive proved that the flow of charges could be through the bulk of the material. This was shown by constructing the double surface transistor, which was produced with germanium in the shape of an acutely tapered wedge, the two contacts being opposite each other near the thin edge. This transistor was developed into the coaxial transistor which was much easier to manufacture. Here the germanium was cut into a pill shaped cylindrical wafer with a dimple ground into the centre of both sides, so that the thickness of the centre was only a few thousandths of an inch. The emitter and collector contacts then bear on opposite sides of the semiconductor in the dimples, and are arranged coaxially to fit into a cartridge. This method of constructtion avoided the problem of placing two spring contacts within a few thousandths of an inch of one another. The components used were similar to the parts used in the manufacture of point-contact rectifiers.

Junction transistors

In 1949 W. Shockley proposed that transistor action could be achieved with p-n junctions within a single crystal, thus breaking away completely from the surface effects of point contact devices. The device was therefore called a junction transistor. In principle it consisted of a bar of single crystal n-type germanium, for an n-p-n device. In the centre of the bar was formed a thin layer of p type germanium as part of the single crystal. Ohmic non-rectifying contacts were attached to each of the three regions, the outer two being the collector and emitter and the centre the base. The method of operation is essentially the same as the point contact, hence the electrodes have the same names, although the base is now in the middle. The equivalent circuit chosen for comparison is the same one as used for the point contact i.e. Fig. 3; the new values for the parameters are:

$r_e = 25\Omega$

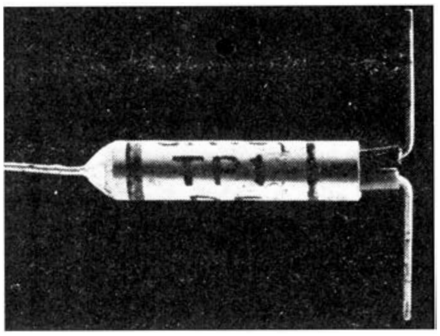
$r_b = 250\Omega$

$r_c = 5M\Omega$

$a = 0.95$

The amplification factor α for junction transistors is less than unity, hence the amplification in common-base operation is due to the differencein impedance levels. A junction transistor was developed with a p-n hook collector, which acted similarly to the point contact transistor as far as the gain was concerned. This was achieved by a four-layer p-n-p-n device, but the transistor had a poor high-frequency response. Little further work was carried out, even though high amplification factors were obtained.

The first junction transistors were a great improvement over the point contact devices. Power gains of 40dB, with class A operation of 49% efficiency were achieved against 23dB gain and an efficiency of 30% for point contact transistors. The higher power gain is due to the increase in the output impedance, and the almost ideal characteristics show that the junction transistor can operate close to the 50% maximum for a class A amplifier. Junction transistors will operate with extremely low input power of around $0.6\mu W$. This is about one ten-thousandth of the power required to operate the point contact transistor, or one millionth of the power to heat the cathode of a typical thermionic valve. Unfortunately the frequency of operation at that time was limited to about 1MHz. This was due to the time taken for the charge carriers to diffuse across the base. The equivalent effect in thermionic valves is the transit time, that is, the time taken by the electrons to travel from the cathode to the anode. The type of case used by S.T.C. for an early junction transistor is shown in the photograph. Although the TP1 device shown was a point contact transistor, it was made at the same time, and externally looks identical to the TS1 junction transistor.



The S.T.C. point contact transistor TP1 appeared about the same time as the G.E.C. GET 1. It was soon withdrawn and replaced by the TS 1, a junction transistor.

Several methods have been used to improve the high-frequency response of junction transistors. The most obvious answer is to reduce the base width; this is limited, however, by the problem of punch through. A second contact added to the base by Wallace et al in 1952 effectively reduced the base area and the base resistance. This increased the cut-off frequency to about 50MHz. Further improvements were realised by advances in material technology, in particular by the diffusion process which started in 1952, and by the production of extremely pure silicon. The purification was achieved by zone refining. This process is based upon the relatively high rate of diffusion of impurities in the molten zone of a crystal, compared with the much slower rate in the solidification zone. The raw single crystal is passed slowly through a localized radio-frequency heating coil. The crystal within the coil is in the molten state, and on passing through the coil re-solidifies into a single crystal again. The impurities tend to remain in the molten region and therefore are swept to the end of the crystal. The process is repeated several times. The end with the impurities is discarded and the concentration of impurities in the main section can be reduced to about 10^{17} atoms/cu.m.

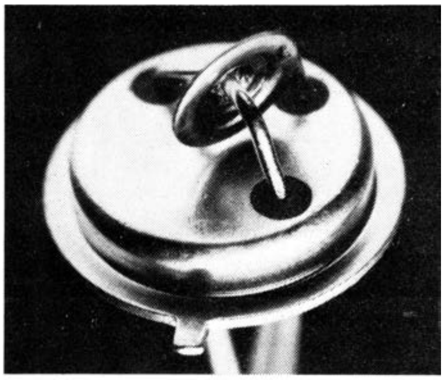
Field effect

The field effect transistor experiments that failed were the beginning for the point contact and junction transistors. In 1952 W. Shockley proposed a unipolar field effect transistor which overcame the earlier problems of surface states. The point contact and junction transistors are called bipolar because charge carriers of both signs are involved. In the field effect the controlled conductance between input and output terminals results from changes in the number of carriers of one type, hence the name unipolar. The field effect transistor has several advantages, the most important being the high input impedance. The input is a reverse biased p-n junction, and the depletion layers created control the conductance through the channel. The difference in operation is reflected in the names for the electrodes, the emitter and collector being called the source and drain respectively. The controlling electrode is now called the gate instead of the base. It was not until fairly recently that the technology needed to be able to mass produce these devices has been developed. In the meantime the junction transistor has built up a commanding lead.

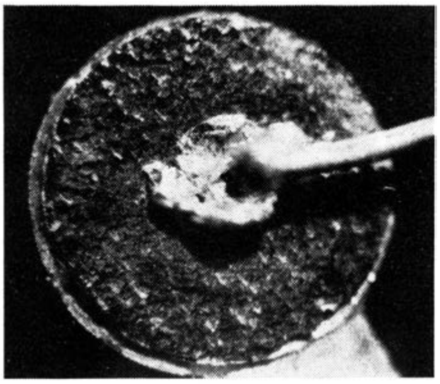
Circuit design

Early work on transistor circuit design tended to start with a well tried thermionic valve circuit, and then modify it for use with transistors, even though the parameters are radically different. The grounded cathode triode is a voltage amplifying device with a high input impedance and a relatively low output impedance. Conversely the grounded base transistor is a current amplifier with a low input impedance and a relatively high output impedance. The early papers on transistor circuit design referred to the transistor's characteristics as peculiar, because they were different to those of a valve. On looking further at the parameters, it was noted that, if the roles of current and voltage were changed over, the devices were similar enough for quantitative designs starting from the valve circuits. This background led to the circuit performance of transistors being less than they might have been, until designers began to take account of the transistor's peculiarities and use them to advantage. One of the major advantages which would be unheard of with valves is the use of complementary circuitry, allowed by having n-p-n and p-n-p transistors.

The small size and ruggedness of transistors opened new fields and their small power requirements meant that the components used with them could be miniaturised also. The type A transistor of 1949 occupied one-fiftieth of a cubic inch, with a collector voltage of 30V. In 1952 the junction transistor could be fitted into one five hundredth of a cubic inch with a collector voltage of 2V. Bell Telephone Laboratories studied the problem of manufacturing complete circuit packages under an American Signal Corps contract in 1952. At that time the package of a laboratory circuit model required about one-tenth the space and power of an equivalent package built with thermionic valves. The importance of designing sub-sections of a system, which would be used in quantity, and manufacturing them as packages was realized from the beginning of the transistor's development, and has been a goal ever since.



A modern germanium alloy junction transistor still in production at Newmarket Transistors. The emitter lead is in the foreground and the base lead at the right connects to a metal disc in which the semiconductor pellet is held.



A photomicrograph of an early medium power germanium alloy junction transistor. The pellet of impurity and the emitter lead connected to it are clearly shown in the centre of the picture.

The general manufacture of transistors began in 1952, after Bell Telephone Labs. held a symposium, where they offered know how to all who wanted it for the price of an admission ticket ($25,000). The era of the practical transistor had now begun. Photographs show the construction of an early alloy junction and the progress achieved since then by comparison with a modern alloy junction transistor. The successive developments to improve the parameters and to find transistor structures, which lend themselves to easier manufacture are related in part 2 "The search for the best transistor". The originators of transistor electronics, J. Bardeen, W. H. Brattain and W. Shockley were awarded the Nobel prize for physics in 1956 in recognition of their work in the theory of semiconductors, when it was beginning to be recognized that they had not just invented the transistor, but had laid the foundations of the worldwide multi-million pound microelectronics industry.

(To be continued)

Bibliography

1. "The amplifying crystal", *Wireless World*, October 1948.
2. Bardeen, J., "Surface states and rectification at a metal semiconductor contact", *Phys. Rev.*, 71, pp. 717-727, 15 May 1947.
3. Bardeen, J., and Brattain, W. H., "The transistor, a semiconductor triode", *Phys. Rev.*, 74, pp. 230-231, 15 July 1948.
4. Shockley, W., "The theory of p-n junctions in semiconductors and p-n junction transistors", *Bell. Sys. Tech. Journ.*, 28, pp. 435-489, July 1949.
5. Shockley, W., "Electrons and holes in semiconductors", Van Nostrand, 1950.
6. Morton, J. A., "Present status of transistor development", *Bell Sys. Tech. Journ.*, 31, pp. 411-422, May 1952.
7. North, H. Q., "Properties of welded germanium rectifiers", *Journal of Applied Physics*, pp. 912-923, Nov. 1946.

The Semiconductor Story

2: Search for the best transistor: continuing a four part series of articles commemorating the 25th anniversary of the transistor

by K. J. Dean, M.Sc., Ph.D., and G. White†, M.Phil., B.Sc.*

At the start of the 1950s the transistor was a novelty. Industry needed to be convinced of its advantages over valves and electro-mechanical devices such as relays and magnetic amplifiers. Besides, there were a number of types being developed—which was the best? Even the textbooks of the period hedged their bets, taking as much space over point contacts as over junction transistors. But the electronics industry, at least, was just beginning to take notice. In 1952 the Post Office Research Station at Dollis Hill had demonstrated the first line amplifier to be made in the U.K. which used junction transistors, while a year later in America, Texas Instruments produced their first pocket transistor radio.

1953 was an important year for the U.K. semiconductor industry. One might almost say that was its birth, for in that year a number of companies set up manufacturing plants, among them G.E.C., Mullard, Ferranti and Pye, who were not then in the Philips group. One of the problems at that time was that the available germanium transistors did not have worthwhile gain at radio frequencies. Naturally, therefore one of the first commercial applications that they chose to exploit was that of transistor amplifiers for hearing aids. The Post Office was the authority for National Health hearing aids and under its guidance Mullard developed the OC56 and OC57 junction transistors specifically for this market. At the same time, Pye at Cambridge had interested Acousticon Ltd, manufacturers of valve-operated hearing aids, in transistors and the first 300 were delivered at the end of 1955. Some of these early devices were packaged in glass cases which were filled with silicone grease and were then painted to prevent the photoelectric effect (amplified by the transistor) making the other current changes due to transistor action. Many an engineer carefully scratched the paint away to use them as sensitive photocells until the manufacurers foiled this dodge by using metal cans. Some of the first metal cases were sealed with solder, leading to examples of flux contamination. The Post Office was not satisfied with these types of encapsulation and insisted on hermetic sealing.

So difficult was the technology of junction devices to master that one manufacturer in those early days recorded that the yield in the first week of production was one device and another calculated that his first working transistor represented an investment of £1 million.

One seldom stops to think why the U.K. semiconductor industry developed as it did. Where did the money come from? Who made the decisions that got it all started? Many companies owed their place in transistor research to the encouragement of C.V.D. (Commercial Valve Development!) This government committee, on which the services, the Post Office and our national research establishments were represented, placed contracts for the development of transistors. It is always popular to blame governmet for wrong decisions or for no decisions at all, but without C.V.D. help few U.K. companies would have got started. One exception was Mullard, owned by the Dutch Philips Group, whose research was funded from the profits of selling valves. In fact their early transistors used valve nomenclature: A for diodes, B for double diodes and C for triodes. The first symbol of the type number was reserved for the heater voltage, zero for transistors of course. So the OC70 was clearly a triode with no heater.

Difficulties with germanium

The first transistors were germanium devices but for a long time the material which would eventually be best was in doubt. Supplies of germanium were limited as there were only three known ores. Two sources were in Zaire (then Belgian Congo) not a particularly stable part of the world; a third ore, germanite, came originally from South Africa, but the mines were exhausted there so that its chief source was from ores imported into Germany before World War I. In addition certain coals contain germanium and at that time the principal supplier in the U.K. was Johnson Matthey who indicated that their main source was from flue dust. Hence, the price of pure germanium was high—about £100 per lb. Meanwhile in Japan the Tokyo Gas Company was extracting germanium from waste coal-gas liquid—one of the first signs of competition from the Far East. It was estimated that one ton of germanium would make 200 million transistors and that in a few years 40 tons per annum would be needed for the world market, against the current production of 3 tons per annum, including the germanium needed for other purposes. Something had to be done.

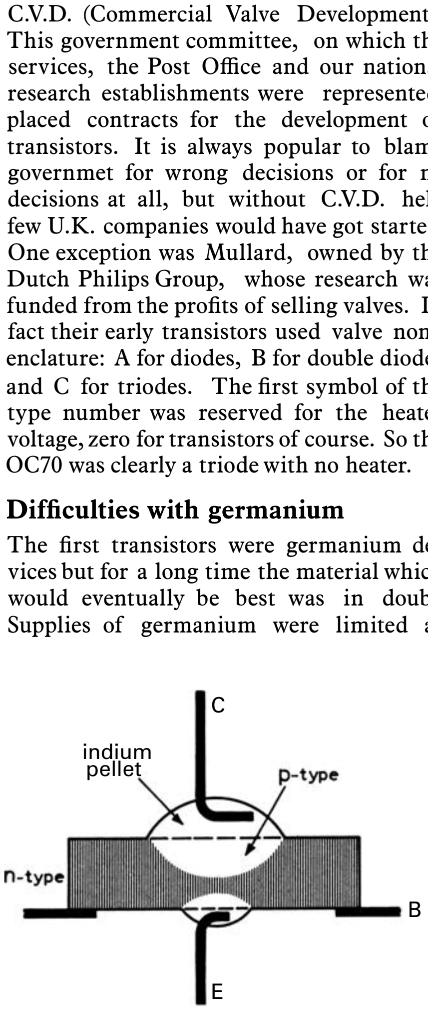


Fig. 1. Slab of n-type germanium with two indium-doped pellets alloyed to it so that it will be modified to p-type immediately below them after heating. The resulting alloy junction transistor was illustrated by a photomicrograph in Part 1 of this series.

Silicon was the obvious contender. Like germanium it is a group IV element; also, after oxygen it is the most common element in the earth's crust, but its melting point is 1420°C compared with 937°C for germanium. The purification of germanium requires a heating and cooling cycle of seven hours, one hour of which was at 1050°C in an atmosphere of pure dried hydrogen. The temperatures for silicon are correspondingly higher. Large quantities of expensive argon are used, which had to be reclaimed, and there were difficulties with phosphorus and boron impurities. Also the quartz (that is, silica) of the crucibles used tended to dissolve in the silicon. As late as 1955, S.T.C. (Standard Telephones and Cables) reported that their own attempts to purify silicon to the extremely high standard of purity required had not been successful. "No further work was done," the report adds, "due to the loss of the man doing it." Nowadays a large proportion of manufacturers are content to buy-in purified semiconductor material in slices for them to process.

Successes with silicon

Texas Instruments were first in the field with silicon transistors in 1952 and had a virtual monopoly for three years. At first the current gain was low and the frequency response was poor due to the lower mobility of charges compared with germanium.

---

*South East London Technical College.
†Twickenham College of Technology.

The Semiconductor Story

3: Solid circuits — a new concept

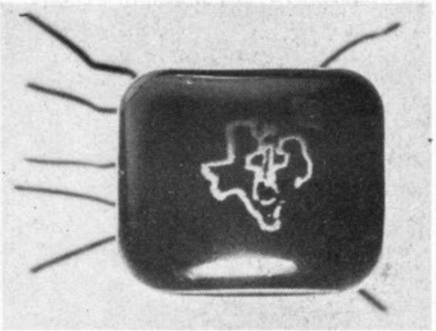
by K. J. Dean, M.Sc., Ph.D., and G. White†, M.Phil., B.Sc.*

The development of the transistor, described last month in part 2 of this series, had been a strange mixture of chance and directed scientific research, of skill with difficult processes and of commercial brinksmanship in which some went too near the abyss and never recovered or withdrew from competing. However, there were occasions when someone intimately involved in the struggle was able to look beyond the immediate technical difficulties and point to an idea not then matched by technological skill, but for which the technology would one day be available. Remarkably enough there are two instances of this happening in the same year, 1952, only four years after the discovery of the transistor effect by Bardeen and Brattain. In both cases the prophesies, for that is what they were, came true in the years to come. W. Shockley, writing in the Proceedings of the American Institute of Radio Engineers (now the I.E.E.E.) laid down the theory of the field effect transistor, fourteen years before it was to become a commercial proposition. G. W. A. Dummer of the Royal Radar Establishment (now at Malvern) speaking at a transistor conference in Washington pointed out that semiconductors could be used to make resistors, capacitors, diodes and transistors so that the possibility of putting a number of all these elements on a single piece of semiconductor existed — in fact that it was possible to make an integrated circuit. It was however to be seven years or so before this idea reached any sort of fruition and about sixteen years before these two, the integrated circuit and the field effect device, came together as a complex commercial product.

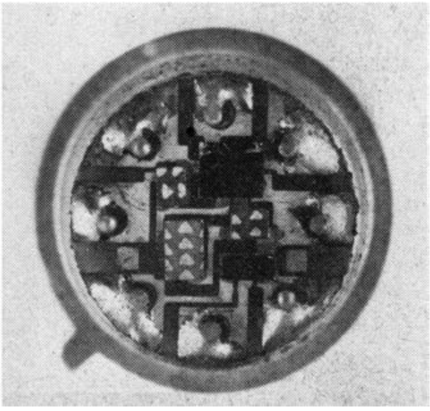
Of course the germanium technology of 1952 was quite inadequate to put Dummer's idea into practice and it was five years before the Plessey Company, who were by then more interested in precise photo-chemical processes, were given a contract in association with the R.R.E. to investigate the possibility of a solid circuit. In 1957 an international symposium on electronic components was held in Malvern at which, reported *Wireless World* in November, the solid circuit was little more than an idea to be discussed in the same breath as ferrite blocks and resin-potted circuits. But there was one point which was significant — the solid circuits being proposed in 1957 were silicon, not germanium.

Technology available

By this time a number of other companies both in the U.S.A. and in Europe were interested in solid circuits, amongst them Texas Instruments (in Bedford as well as in the U.S.A.) and Fairchild. Not only were silicon transistors available but the mesa process had also been recently developed, largely by Texas Instruments. Now this process has the important advantage of requiring diffusion from only one face of the silicon slice. Hence it was thought possible to place various active and passive components side by side on a single slice and then inter-connect them. In 1958 this is what Texas were able to show they could do. As was the case with transistors where increasing skill with technology and governmental patronage produced a variety of transistor types, changes in solid state techniques had a vital impact on the development of integrated circuits. The key technology was the development by Fairchild of the planar process, so that even by 1960 it was clear that planar devices would most easily lend themselves to interconnection as solid circuits. In fact it can be argued that two of the major efforts since that time have been to minimize the profile contours of silicon chips and reduce the size of transistors within the chip. These have been brought about using modifications of the planar process.



Early Texas mesa integrated circuit showing the vitreous enamel package, from the underside of which the connecting leads protrude.



Multi-chip integrated circuit by S.T.C. mounted in a T05 header on a printed circuit board on which the interconnection pattern has been etched. The circuit, sold in 1964, is that of a d.t.l. gate.

The patronage which proved decisive and turned, alas once more, a British idea into a foreign product, came from the U.S. Government. The Minuteman project was at the end of the 1950s the American contribution to the U.S./U.S.S.R. arms race and represented the ultimate then possible in electronic sophistication. It was funds from this project, principally. to Fairchild but also to Texas which provided the immediate incentive to devise high component-density circuits of great reliability for use in the limited space and very difficult environment of a missile. Thus the early integrated circuits were born. Although by this time a technology to make a form of integrated circuit was available on a laboratory basis, it had a number of limitations, both of cost and as a production method. Failure to produce a reliable isolation technique meant that multi-chip circuits were the best that many companies could do. One chip might carry a single transistor another might carry a resistive network and a third might consist of diodes. The chips were at first mounted on a suitable sub-divided printed

---

*South East London Technical College.
†Twickenham College of Technology.

The Semiconductor Story

4: Large scale intentions. Conclusion of a series of articles commemorating the 25th anniversary of the transistor

by K. J. Dean, M.Sc., Ph.D., and G. White†, M.Phil., B.Sc.*

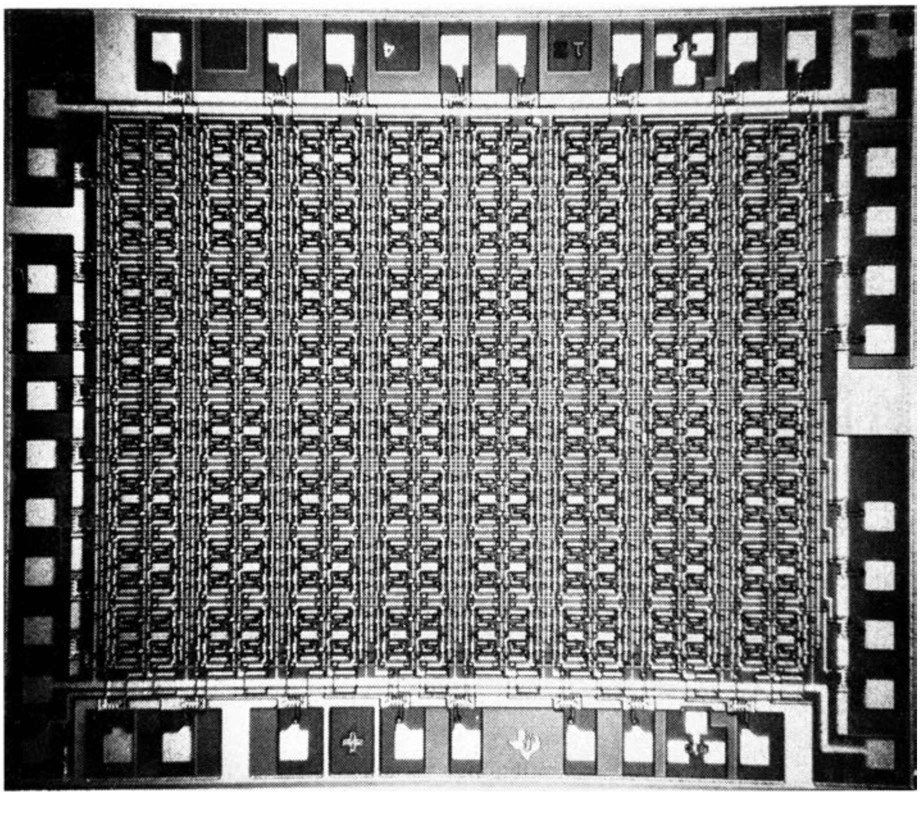
Since 1945 the industrial society in which we live has been one where technological change has been the normal state of affairs. It is not easy to plan such changes; indeed there has been very little worthwhile market and technological forecasting. Our national research establishments have been involved in bringing changes about, but it does not seem to have been a part of their role or that of industry to formulate clear research and development goals based on market assessment. To a surprising extent the semiconductor industry has been a victim of circumstances rather than their master. Its fortunes were founded on the arms race and further encouraged by the U.S. space programme. Again we have seen that military confrontation seems necessary to bring about major scientific developments. (There must surely be some other way.) Defence contracts helped establish large production plants when yield efficiencies were small, so that increasing skill and consequent falling production costs brought overproduction and "dumping". Fierce competition resulted in casualties despite the larger market which became available. A situation was arising which, though so clearly visible in retrospect. no one appeared to notice then.

Larger chips

As the move to put more electronics on a single chip got under way even greater attention was paid to the problems of increasing yield. There are, perhaps, three golden rules if high yields are to be achieved but, like all such rules they are easier to state than to implement.

First, the processing should be simple The main difficulty here is with gold doping which is a particularly critical process necessary because charge-storage takes place in the lowest concentration area of doping, which is usually in the collector region of the transistor. Gold doping decreases life-time and so reduces charge-storage. However, this effect can also be mitigated by a diode between collector and base, so that the overdrive current goes through this anti-bottoming diode rather than the collector region. Unfortunately, if a silicon junction diode is used it has the same forward characteristics as the silicon junction transistor which it is trying to speed up. This difficulty was overcome by using the Shottky barrier diode formed by aluminium on the silicon, which has a knee voltage of 0.3V instead of 0.5V for a silicon junction diode. The use of Shottky diodes to clamp a transistor was originally developed by Texas in 1964. In some devices the storage is in the base region. In this case, a second emitter is provided for the transistor on the chip and connected to a Shottky diode to remove the charge. Devices where speed is obtained from Shottky diodes are compatible on the same chip with linear circuits whereas gold doped circuits are not. They are also compatible in the same system, but not the same chip, with similar designs for gates, but which use gold doping. An example of this is the Texas 74 series which has been second-sourced by a number of other **manufacturers.**

Secondly, the number of stages in the process must be kept small. But as the circuits which industry require become more sophisticated, such as gates with good speed-power ratio and high fan-in and fan-out and capability for wired-OR connections, so the number of stages tends to rise. There are for example typically three masks needed for a single transistor, eight for a t.t.l. gate and ten for some linear amplifiers. Only five masks are needed for m.o.s. gates but here speed problems exist, particularly where m.o.s. gates are interfaced to external connection. The yield of a single diffusion is inversely proportional to the area of the chip. That of a transistor or integrated circuit with *n* diffusions is proportional to the yield for a single diffusion raised to the power of *n*. Thus the yield for planar transistors with three diffusions must be extremely high



Dual 64-bit shift register first available commercially in the U.K. in 1967 and typical of the state-of-the-art at that time. (photo: Texas Instruments)

---

*South East London Technical College.
†Twickenham College of Technology.