TRANSISTOR *Tales*

by Gregg Grant

PART 2

Sussing out Silicon

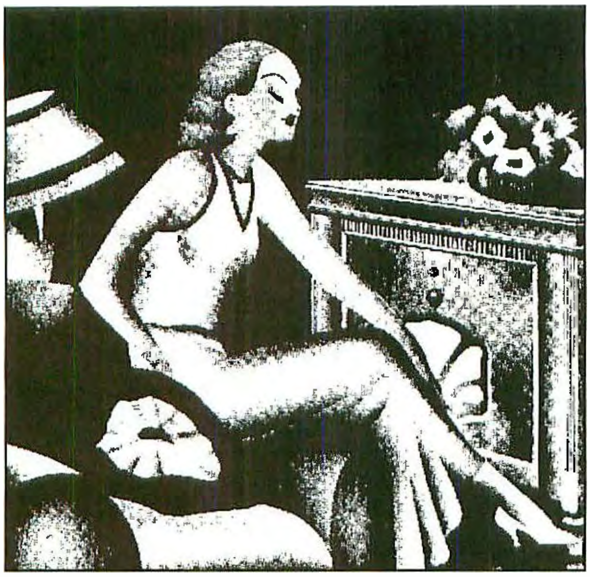


Figure 1.

The 1920s and 1930s are remembered - if at all - for the Great Depression and the myriad problems it brought. One such problem was that there was virtually no capital available for investment in technical developments and, even if there had been, it's questionable if it would have been allocated for projects aimed at reducing the size of things.

At this time the valve was king, new types of the component being brought out on a fairly regular basis. The Pentode for example - developed by Tellegen and Hoist at Philips in Eindhoven - was a child of this period. More to the point the valve, for its purpose, was considered satisfactory by both the radio manufacturing industry and the general public.

For those fortunate enough to be in work during this period, the cosy relationship between manufacturer and customer worked to the advantage of both. One of the most potent symbols of the importance of a job was the radio set in the living room. As Figure 1 illustrates. this was a large unit, frequently amounting to a piece of furniture.

Figure 2 illustrates one of the advertising slants of the period, much of which revolved around the 'more is better' theme, where the number of valves in the equipment signified but one thing: quality. Therefore developing smaller components using less power - whether a valve, capacitor or anything else - sent entirely the wrong indicator to the marketplace, in fact the sort of signal the crystal set radiated: poor quality.

Nevertheless, the period was not entirely barren where scientific and technical advances were concerned. It was in this period for example that the theory behind what would later be called semiconductors was developed and the first transistor patents were filed.

The Theory Men

Between 1926 and 1932, a number of able physicists published a series of papers on the quantum electron theory of metals. Among the authors were the Nobel Prizewinners Felix Bloch, Werner Heisenberg and Wolfgang Pauli as well as Rudolf Peierls and Arnold Sommerfeld.

As a result, there developed a basic quantum theory of solids, which became the bedrock on which future advances were built. The next generation of theorists began to examine the properties of what would shortly be termed semiconductors.

In the US, the principal centres of this research were evolving by the early 1930s, one at the Massachusetts Institute of Technology, the MIT, the other at Princeton.

In the UK, there were also two centres where this sort of research was carried out. One location was Bristol University, where Neville Mott and Harry Jones were the principal investigators. The other was at Cambridge University, where Alan Wilson was also examining the physics of semiconductor materials.

Both Mott and Wilson began by attempting to understand the action of the metal/semiconductor rectifier, of the type invented by Grondahl and Geiger .

Wilson thought that the action was a quantum mechanical one, in which the charge carriers 'tunnelled' through the narrow 'response curve' - a mere 10^{7} cm wide - of Figure 3. This concept seemed to indicate that the direction of low resistance to current flow - what would come to be termed the Forward direction - was from metal to semiconductor.

In fact this prediction by Wilson was exactly the opposite to that observed and later experiments indicated that the barrier width was around 10^{5} cm. which suggested that tunnelling was not the explanation for the rectifying action. Mott - and independently the German research engineer Walther Schottky - put forward a different explanation, that the majority carriers were thermally agitated over a wider barrier, one in the region of 10^{4} to

Figure 2.

R·G 41

If they will be honest about it, and admit it, most keen Radio listeners will tell you that their aim and ambition is a good Radiogram : for there is something so intensely satisfying about owning a Radiogram. Radiograms however are big instruments, hold a lot of mechanism, and usually cost a lot of money. Here is a splendid All-wave Radiogram built with all the care characteristic of Bush Radio, at a price which will do much to break down the money barrier. As a point of technical interest, an L.S. valve is used in front of the output valve, which is essential if really good quality is to be expected or even hoped for either on radio or records.

ALL WAVE RADIOGRAM

5 valve (inc. rect.) 7 stage superhet with 6 tuned circuits, for A.C. Mains Waveranges: Short 16.5-51 m. Medium 189-550 m. Long 800-2,000 m. Full A.V.C. Attractive scale station calibrated. Each band individually illuminated. Continuously variable tone control. Large, sensitive M.C. speaker. Provision for external speaker. Very lovely walnut cabinet. Size 3' 2½" high, 23" wide, 15½" deep.

FOR 20 GUINEAS

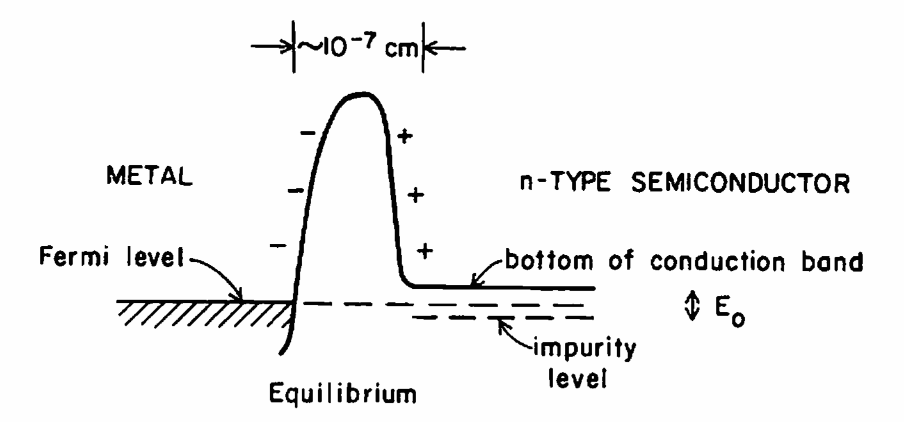


Figure 3.

10^{4} cm. Their concept is illustrated in Figure 4. This is essentially a ramp in which - in equilibrium - the electrons on both sides of the divide see the same barrier height. When, however, a negative voltage is applied on the metallic side as Figure 4b, it raises the electron energy levels in the metal relative to those in the semiconductor.

The result is an increase in the barrier space charge and the semiconductor electrons 'see' an increase in barrier height. Therefore electron flow from the semiconductor to the metal decreases.

Conversely, the application of a positive voltage to the metal as in Figure 4c, decreases the metal's electrons energy levels in comparison to those of the semiconductor material, so that its conduction electrons 'see' a barrier that has been lowered. The result is a considerable increase in current flow in the forward direction, from semiconductor to metal.

Although both the Wilson and Mott-Schottky theories had their faults - one of which was that both hypotheses took no account of the role of minority carriers - they were among the earliest attempts to explain the metal/semiconductor forward-reverse conduction mechanisms. This then paved the way for the later, American work on the development of the transistor.

In fact by 1931, Alan Wilson was using '. . the so-called band theory to study the processes by which conduction occurs' [1] in semiconductors and this groundbreaking work remains the basis of our understanding of semiconductor operation. His theory gave us the concept of doping semiconductor material with precisely measured amounts of impurities, which produced what he termed defect and excess types of semiconductor.

Ten years after Wilson's research Jack Scaff, a Bell Laboratories (Bell Labs) metallurgist, renamed these substances p-type and n-type, the names by which they are known today.

By early 1935, Bell Labs had decided to carry out research into electronic conduction in solids, the Bell believing that '. . the knowledge of materials and processes acquired will ultimately be of value in the solution of engineering problems.' [2] This research triggered a new interest in crystal detectors, Bell wishing to develop '. . a crystal detector for microwaves because vacuum tubes would not work at those frequencies.' [3]

Consequently one of Bell's radio engineers, George Southworth, asked his metallurgist and chemist colleagues if they could produce silicon whose properties would be more controllable and predictable than that used in the 'cat's whisker' radio detectors. Whilst Bell Labs concentrated on silicon, a research group at Purdue University, led by Professor Karl Lark-Horovitz, investigated germanium, with a view to using the material as a radar detector. Yet this period was not one of theory entirely. Two German researchers actually applied for patents for solid state devices at this time, one in the US, the other in the UK.

The Patent Men

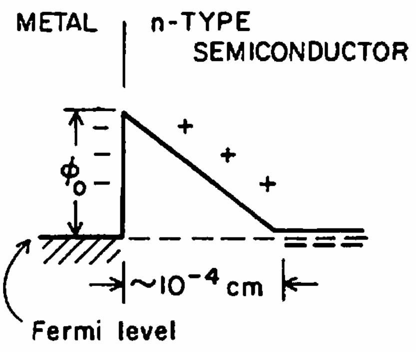
In March 1928, Dr. Julius Lilienfeld filed a patent application with the US Patent Office for what he termed "a device for controlling electric current."

Patent No. 1,900,018 was subsequently granted on the 7th of March 1933 for the device illustrated in Figure 5. This has had the letters G, S and D added, mimicking the current field-effect transistor terminology of Gate, Source and Drain.

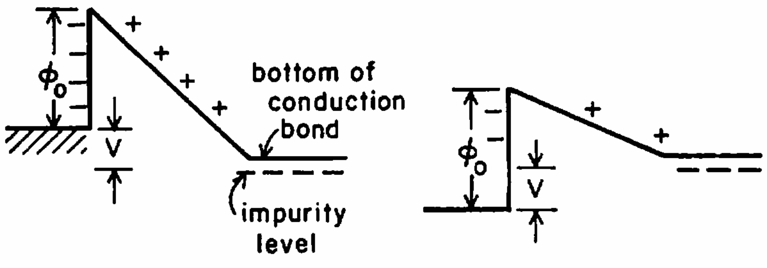
The materials used were unusual, to say the least. The Gate - 10 in the illustration - is made from aluminium and insulated by a film of aluminium oxide, 11. The p-type semiconductor - 12 - is made from cuprous sulphide and is of molecular thickness at the point shown as 13. The Source and Drain contacts are shown as 14 and 15 and are electrical conductors. Both they and the semiconductor were vacuum evaporated. How well did this early quasi-transistor work?

In 1964 J. B. Johnson, the discoverer of 'Johnson Noise,' who was - at that time - a physicist with the Instrument Division of Thomas A. Edison Industries attempted to find out whether Lilienfeld's device would indeed work.

Johnson had his doubts since - despite his having carefully followed Lilienfeld's



a. Equilibrium



b. Voltage in back direction

c. Voltage in forward direction

Figure 4.

specifications - he noted that no amplification or even modulation took place. He thought that one reason for this '. . was probably the very low mobility of holes in the Cu$_2$S (~ 1cm^{2}V-sec), and the effect of surface states on the free surface of the film.'[7]

In 1991, another experimental effort was made to create the device outlined in the Lilienfeld patent. In his Master of Science thesis for the University of Vermont, entitled Th invention of the transistor, Bret Crawford set out to reproduce the original using the same materials as had been available in the late 1920s.

Faithfully adhering to the patent, Crawford did produce working models, although their performance was less that spectacular, they being unstable. Cuprous Sulphide as a semiconductor was simply no match for the present-day, purified and carefully doped silicon. Nevertheless, Crawford suggested that Lilienfeld actually did manufacture the device he described in the patent, as well as theorising about it because his - Crawford's - own observations were close to those of Lilienfeld, as described in the patent application.

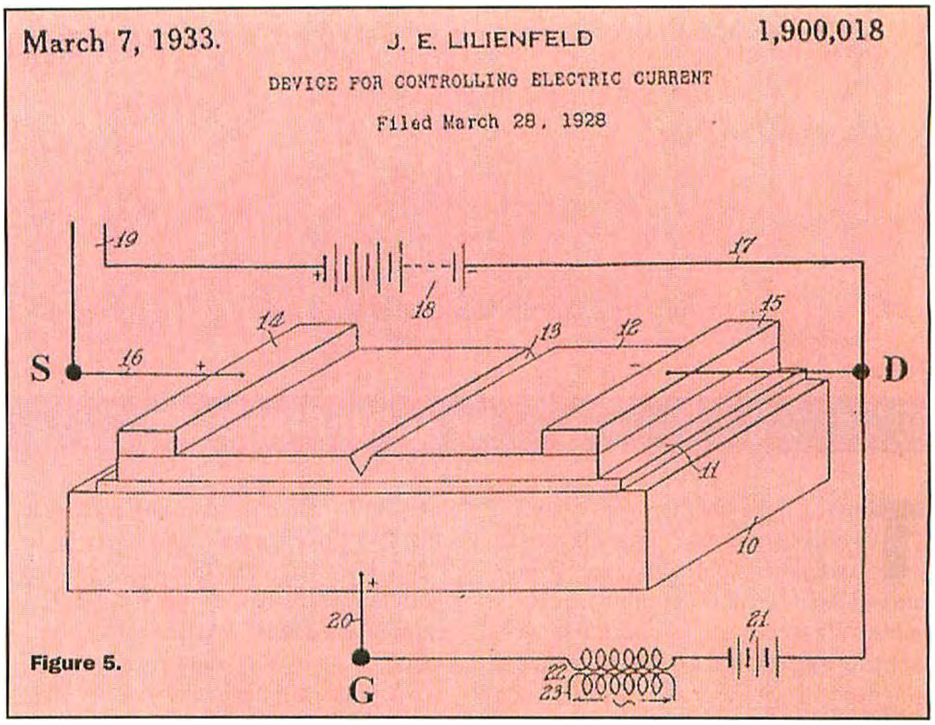


Figure 5.

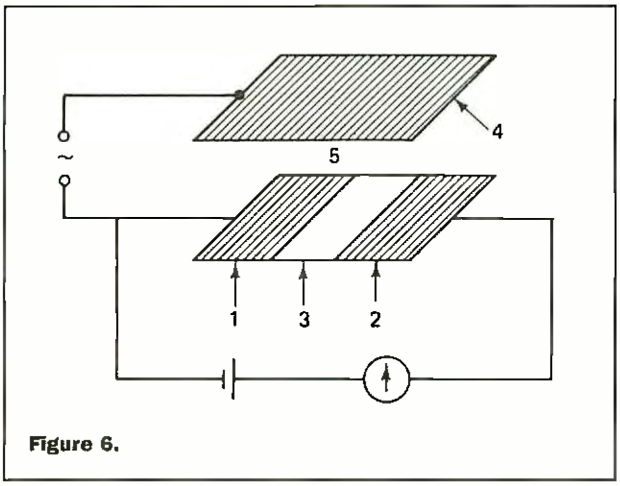


Figure 6.

Four years later another American physicist, Joel Ross, looked into the Lilienfeld device. At the Spring meeting of the American Physical Society, Ross presented his paper Reconstruction of a Lilienfeld transistor, in which he - like Crawford - followed the prescriptions of the original patent. The devices Ross produced, whilst they remained stable over a considerable period and did indeed demonstrate the field effect, had poor conductance due - he thought - to surface states.

In 1935 the University of Berlin's Oskar Heil filed a patent application in the UK for what he termed Improvements in or relating to Electrical Amplifiers and other control arrangements and devices. His concept is illustrated in Figure 6.

The section annotated 3 in the drawing is a very slim area of semiconductor material such as cuprous oxide, iodine, vanadium pentoxide or tellurium, whilst the even smaller areas designated 1 and 2 are ohmic contacts.

The area above the semiconductor material annotated 5 is a metallic control electrode which is insulated - 4 - from the semiconductor. In his explanation, Heil outlined how a signal applied to the control electrode modulated the semiconductor's resistance, producing an amplified signal, which is registered by the current meter.

Today, this device would be termed a Unipolar field-effect transistor having an insulated gate or, to give it its common abbreviation, an IGFET.

In 1939, William Shockley at Bell Labs wrote in his workbook that it '. . occurred to me that an amplifier using semiconductors rather than vacuum is in principle possible.'[5] Yet another decade would pass before the transistor was invented and it would be the early 1960s before a commercial field effect transistor became available to the electronics industry. Why?

The most obvious problem was the material itself. The first thing to note about silicon is that it is NOT a metal. It comes from silicon dioxide, in its familiar form of quartz, and is produced in electric arc furnaces. The silicon produced, via the reducing agents wood chippings and coke, is over 95% pure. The ubiquitous silicon chip is created by crushing the product of the furnace '. . to a powder and then [leaching it] in hydrochloric acid, forming the compound SiHCl$_3$. The compound is then treated with hydrogen; extremely pure silicon is formed.'[6]

The above of course is the modern method of producing the material on which the electronics industry has depended for the last half century or so. In 1939 however Russell Ohl, an electro-chemist turned radio engineer working at Bell Labs, found that contact between the silicon material and a 'cat's whisker' rectified in one direction and then, suddenly, in the opposite direction.

Ohl decided that the way ahead lay in purifying silicon by melting it, and he became the first engineer to discover that special furnaces would be needed to do this.

In the autumn of 1939, Ohl turned to two metallurgist colleagues, Henry Theuerer and Jack Scaff, and asked them to purify the silicon. In the course of doing this, Scaff and Theuerer confirmed what Ohl had already discovered, namely that the direction of rectification varied from one silicon ingot to the next. Obviously, more research was needed on this, at times strange, material.

References

1: Atherton, W.A. (1984): From Compass to Computer. San Francisco Press Inc., San Francisco. Ch 10, Page 240.
2: Hoddeson, Lillian (1981): The Discovery of the Point Contact Transistor. University of California. Page 48.
3: Weiner, Charles (1973): How the Transistor Emerged. IEEE Spectrum Article, January 1973. Page 28.
4: Johnson, J.B. (1964): More on the Solid-State Amplifier and Dr. Lilienfeld. 'Physics Today' article, May 1964, Page 61.
5: Ibid [2], Page 60.
6: Street, Arthur & Alexander, William (1998): Metals in the Service of Man. 11th Edition. Ch 23, Page 214.