Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

6.   Third case study –

the solid state advent

29 December 1939 "It has occurred to me that an amplifier using semiconductors rather than a vacuum is in principle possible."
16 December 1947. An entry in Brattain's notebook clearly shows that on this date the point-contact transistor "was born."
Tuesday, 23 December 1947 was the date of the private demonstration for executives that is described in Brattain's famous notebook entry of Christmas Eve, 1947.
Brattain's *24 December* notebook entry continues after the report of the 23 December demonstration to record that an oscillator was constructed on 24 December and did, indeed, oscillate. The birth of the point-contact transistor was a magnificent Christmas present for the group as a whole.

William Shockley
IEEE Trans Electron Dev **ED-23** (No.7) July 1976 p597

"Our work was therefore directed toward a fundamental understanding of the problem, although we were well aware of the technical importance of a semiconductor amplifier – if one could be made."

Walter Brattain
Adventures in Experimental Physics **5** pp1-31 (1976)

"It may be appropriate to speculate at this point [21 October 1950] about the future of transistor electronics. Those who have worked intensively in the field share the author's feeling of great optimism regarding the ultimate potentialities. It appears to most of the workers that an area has been opened up comparable to the entire area of vacuum and gas-discharge electronics. Already several transistor structures have been developed and many others have been explored to the extent of demonstrating their ultimate practicality, and still other ideas have been produced which have yet to be subjected to adequate experimental tests. It seems likely that many inventions unforeseen at present will be made based on the principles of carrier injection, the field effect, the Suhl effect, and the properties of rectifying junctions. It is quite probable that other new physical principles will also be utilized to practical ends as the art develops."

William Shockley
Electrons and Holes in Semiconductors

6.Third case study – the solid state advent

6.1 Introduction

The first two case studies provided an opportunity to contrast the changes in two regimes. In the first, innovation rent was generated by the introduction and organic growth of electric lighting. The desire to remove the constraints imposed by the absence of natural daylight created a demand which grew continuously as economies of scale reduced input costs. Technological superiority provided the motivation for paradigm shifts, from the luminous flame to the arc discharge and then from the arc to the incandescent filament.

The second case study examined the course of communications engineering as it traversed changes of comparable magnitude. It illustrated the path from historical precursors, based purely on human sensory phenomena, to cable systems made possible by the discovery of basic electro-magnetism and then to radio, which was initially predicted theoretically and, subsequently, demonstrated practically and refined into a working system.

In contrast to electric lighting, the growth of which was limited only by socio-economic considerations, cable and wireless communications were subject to external constraints. Intrinsically, their use transcended national territorial boundaries and required common standards for inter-operability. Transmissions could interfere with one another, so wireless was regulated by diplomatic agreement. Efficient communications had a significant military application. They therefore created a demand, the impact of which varied with the incidence of war and peace. Commercial markets were, for many applications, inelastic, reducing the propensity to reduce input costs by way of economies of scale. For example, a ship would need only one wireless system and, hence, the total market was constrained by the number of ships and could expand only with an increase in maritime activity.

The events examined in the two case studies took place concurrently, permitting the effects of differing micro-economic factors to be highlighted. The paradigm changes were also sufficiently remote in time for them to be examined with an historic perspective, as their consequences are, by now, fully played out. A direct corollary of this, however, is that, as the innovations can only be studied using the information that

6.Third case study – the solid state advent

is currently available, the conclusions will clearly lack the insights which can be provided by observation at first hand.

The objective of the third case study is to investigate in greater detail the effect of technological considerations on the dynamics of paradigm shifts. This is achieved by examining a relatively recent, hi-tech innovation, the factors of which are within current experience. During the last fifty years or so there has been an explosive increase in the numbers of learned journals and technical newspapers, so contemporary thinking is well documented. The case study also gains credibility both from knowledge gained by the author's direct participation in the relevant research and development during the early 1960s and from a personal acquaintance with many of the companies and individuals who were involved with decisions which shaped the industry and guided its transition from parochial to global organisation.

The investigation traces the development of the transistor through its entire Schumpeter A-phase. The example chosen is particularly important because, throughout the period in question, the industry was growing at an average rate of thirty per cent per annum. Initially, the cost of market entry was low, but, even in the later stages, the magnitude of this growth meant that the payback time for investment in innovation was short and therefore cost considerations would not greatly inhibit change of direction if technological indicators suggested that a paradigm shift was the right course to take.

Following the pattern of the previous case studies, the development of the solid-state amplifier is traced from its nineteenth century origins, through early abortive attempts, to the Nobel prize-winning invention made just before Christmas 1947. The contributions of theoreticians, experimental physicists, materials scientists and applications engineers are documented, together with the battle against technological dragons which appeared to threaten the innovation in its early stages.

Of particular interest is the way in which the physical properties of materials, especially their metallurgy, were characterised and then harnessed to create device topographies with desired electrical parameters. The technology of fabrication techniques underpinned the entire development, but would other factors, notably inventors, patents and communications, have a significant impact? The legacy of the

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

growth of the telephone created a situation in which the company which fostered the invention of the transistor was constrained by US anti-trust law restrictions and could not exploit the invention directly. This provided a unique opportunity to view the impact of competition law at the outset of an innovation rather than when it has reached maturity, which is the more usual circumstance.

Although the thermionic valve had given good service as an electronic amplifying device since the turn of the century, it had a number of disadvantages. It was based on the physical phenomenon of thermionic emission of electrons which involved heating a cathode to around 1000ºC. This process required a relatively large amount of energy and generated a high level of noise in the signals which were to be amplified. The heat also caused the degradation of the electrical characteristics of other components mounted in physical proximity.

The thermionic valve was a device which had a high failure rate. Constructed of fine wires and welded metal plates which were mounted in a geometrically-critical relationship in a vacuum-sealed glass bulb, it was not robust. Furthermore, the cathode material exhibited a fall-off in efficiency with passage of time, which set a finite limit on its service life.

When the transistor arrived it promised the same performance as the valve whilst consuming less than one thousandth of the power. Because the new device generated a much smaller amount of heat, it was potentially more reliable and could be installed in circuits at a much higher packing density. This was facilitated by the fact that it was very small. Mass production would lead to low cost and large volumes of uniform devices. A solid-state amplifier which exhibited these characteristics would therefore rank highly in Schmookler's tests for desired inventions [Schmookler 1976] and its success could be assured.

6.2 Early discoveries and inventions

Many of the physical properties that characterise semiconductors had been observed empirically by the end of the nineteenth century. [Pearson 1955, p1794] In 1833, Faraday demonstrated that the electrical resistance of silver sulphide decreased with increasing temperature. [Faraday 1839, p122-124] Six years later, Becquerel noticed that a

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

voltage appeared at a junction between a semiconductor and an electrolyte. Photoconductivity was discovered in 1873 by W. Smith whilst working on selenium. [Braun 1982, p15] The next year, Ferdinand Braun, a professor of physics at Marburg, discovered that a contact between metals and galena (lead sulphide) or pyrites (iron sulphide) exhibits a nonlinear current-voltage relationship [Braun 1982, p11] and Schuster observed rectification at a contact between tarnished and untarnished copper wires (probably having a surface film of cuprous oxide or sulphide). [Schuster 1874, p251-252] Adams and Day developed a selenium photocell shortly after [Adams 1876, p113-117] and, in 1883, Fritz constructed the first large-area dry rectifier. [Fritz 1883, p475-472]

Interest in wireless telegraphy at the turn of the century provided the impetus for the next flurry of activity. In September 1901, Bose constructed a wireless coherer detector using galena. [US Pat 755840] Dunwoody used a bulk silicon carbide detector for radio waves in 1906 [US Pat 837616] and Pierce constructed a rectifier from this material in 1907. [US Pat 879061] He subsequently extended his work to titanium dioxide [US Pat 879062] and silver telluride [US Pat 879117] using point contact crystal rectifiers as radio detectors. [Pierce 1907, p31-60] Austin contemporaneously used semiconductors to construct a thermo-electric detector. [Austin 1907, p508-510]

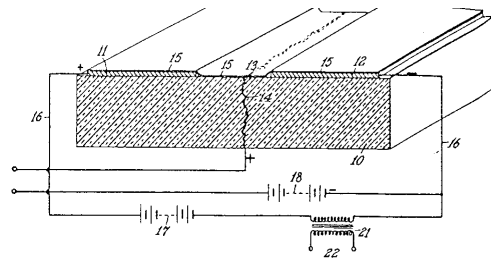## 6.3 The first inventor of the transistor – Julius Lilienfeld

The part played by Julius Lilienfeld in the development of the transistor was the twentieth-century analogue of Heinrich Göbel's role in the invention of the carbon filament lamp. In 1905, he gained a PhD in physics at the University of Berlin, where he was a contemporary of Max Planck and Max von Laue. In 1910 he became a professor of physics at the University of Leipzig where he stayed until 1927. He was granted several patents for inventions concerned with x-ray tubes and also worked with Count Ferdinand von Zeppelin on the design of hydrogen-filled dirigibles. He emigrated to the United States in 1927 to become director of research at Ergon Research Laboratories in Malden, Massachusetts. [Sweet 1988]

In the mid-1920s, he turned his attention to solid state devices and filed a patent application for a dry cell rectifier. [US Pat 1611653] Extending his interest in semiconductors, he proposed a structure for a three-electrode device. [US Pat 1745175] Nowadays, this would

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

be described as a grain-boundary field effect transistor. Its purpose was "for controlling the flow of an electric current between two terminals of an electrically conducting solid by establishing a third potential between said terminals" which would be

US Pat 1745175

Fig. 6.1 Lilienfeld's grain boundary field-effect transistor

"particularly adaptable to the amplification of oscillating currents such as prevail, for example, in radio communication." The object was "to dispense entirely with devices relying upon the transmission of electrons through an evacuated space and especially to devices of this character wherein the electrons are given off from an incandescent filament" providing "a simple, substantial and inexpensive relay or amplifier not involving the use of excessive
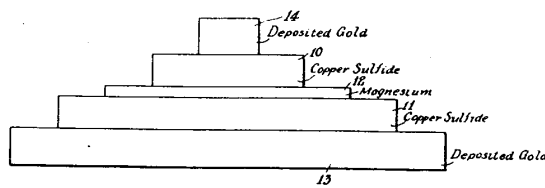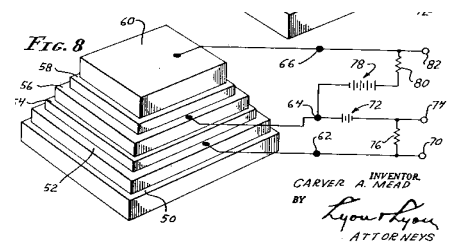

Fig. 6.2 Lilienfeld's MSMSM transistor

US Pat 3056073

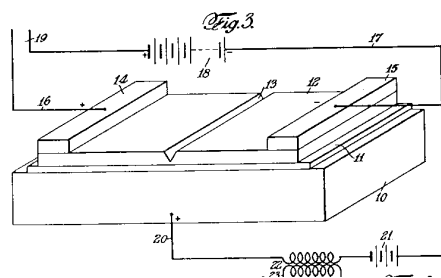Fig. 6.3 Mead's MIMIM transistor (1960)

Fig.6.4 Lilienfeld's MOSFET

Fig.6.5 Weimer's MOSFET

voltages … More particularly, the suitable incoming-oscillations, a current characteristics that said current will be changes."

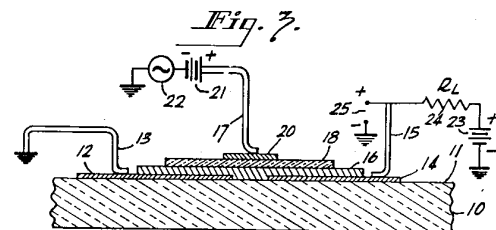invention consists in affecting, as by in an electrically conducting solid of such affected by and respond to electrostatic

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

Lilienfeld went on to patent two further solid state amplifier devices, the precursors of the metal-insulator-metal-insulator-metal (MIMIM) transistor and the metal-oxide-semiconductor field-effect transistor (MOSFET). The MIMIM device was re-invented by Carver Mead in 1960 [US Pat 3056073] whilst Lilienfeld's MOSFET greatly resembles the thin film devices developed in the 1960s by P.K. Weimer at the Radio Corporation of America. [US Pat 3258663]

John Bardeen, co-inventor of the point contact transistor, summarised Lilienfeld's contribution in a letter which was published when the American Physical Society established the Julius Edgar Lilienfeld prize in 1988. [Sweet 1988]

> "Lilienfeld's patents predated the work of Brattain, Shockley and me by nearly 20 years. A considerable effort was made by J.B. Johnson and others to reproduce his results. For patent reasons it was necessary to distinguish our work from the prior art. Nothing was found that would suggest that the bipolar principle had been anticipated in the prior art of Lilienfeld and others, as pointed out in the article by Johnson…
>
> Lilienfeld deserves great credit for his pioneering efforts to make a semiconductor amplifier. This was not long after copper oxide and copper sulfide rectifiers were discovered. He had the basic concept of controlling the flow of current in a semiconductor to make an amplifying device. It took many years of development of theory and materials technology to make his dream a reality."

Lilienfeld also made very fundamental proposals for a field-emission cathode based on a sharply-pointed electrode coated with a low-work-function metal [US Pat 1578045] and may have influenced later pioneering work by Walter Schottky, with whom he had correspondence before he emigrated to the USA.

Because suitable materials were not available, Lilienfeld was unable to construct viable amplifiers. He abandoned interest in them and went on to work on electrolytic capacitors. In the mid-1930s he moved to the Virgin Islands with his American-born wife and disappeared into obscurity. As with Göbel and the incandescent lamp, his proposals were used by defendants in litigation concerning the transistor patents of the Bell System, but the courts decided that they did not anticipate the principles on which the transistor was based.

6.4 The second inventors of the transistor – Hilsch and Pohl

Contemporaneously with Lilienfeld's proposals, Grondahl and Geiger demonstrated rectification at a copper-cuprous oxide barrier [Grondahl 1927] and, shortly

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

afterwards, Lange showed that a selenium-metal structure exhibited rectification and a photo-electric effect. **Lange 1936**

In 1931, Wilson proposed the electron theory of semiconductors **Wilson 1931** from which Mott **Mott 1939** and Schottky **Schottky 1939** developed a theory of rectification at a metal-semiconductor barrier. Pohl, at the University of Göttingen, performed experiments on colour centres in alkali halide crystals and, with Hilsch, **Hilsch 1938** constructed a solid-state, three-electrode device, based on the analogy of a thermionic triode. In this, a metallic wire coated with potassium served as a cathode. A control grid of platinum wire 0.2mm in diameter was embedded in the crystal approximately 2mm from the cathode. The anode was a platinum layer in contact with the surface of the crystal. This device, which operated at an anode voltage of around 100-150V and a positive grid voltage of around 10V, had an amplification factor of about 20. It was, however, only of academic interest, since it had such a low cut-off frequency that it would not amplify any useful signal.
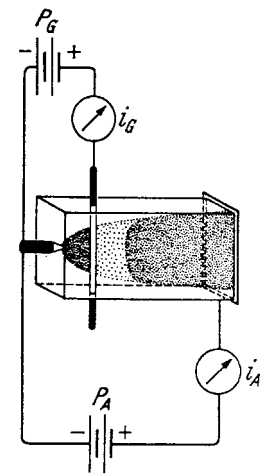
Hilsch 1938
Fig. 6.6
Control of current flow in ionic crystals

6.5 "Third time lucky!" – the invention of the transistor at Bell Labs.

At the time of the invention of the transistor, there was a feeling in United States' industry that research of a fundamental character would yield practical benefits. **Shockley 1956, p100** This view was held particularly strongly at Bell Telephone Laboratories where it stemmed from attitudes of four men who had successively been its research director and president. Each of these men, H.D. Arnold, F.E. Jewett, O.E. Buckley and M.J. Kelly, had a background in governmental or civic affairs and each was thoroughly conversant with the culture of research, having undertaken a doctorate in physics. A team of experts was gathered together and set to work investigating the basic properties of semiconductor materials.

6.5.1 The advent calendar – key stages in the development

The successful development of the first viable solid-state amplifier device is summarised in Table 6.1, which shows the key steps, as seen by William Shockley, who

6.Third case study – the solid state advent

led the team. [Shockley 1976] JB, WHB, and WS refer to John Bardeen, Walter H. Brattain, and William Shockley in whose laboratory notebook the entry (nbe) appears. Dates considered to be important by Shockley appear in bold face type.

Table 6.1

29 Dec 39	WS (nbe)	**disclosure of what would now be called "Schottky-gate field-effect transistor."**
29 Feb 40	WS (nbe)	improved version of 29 Dec 39.
16 Apr 45	WS (nbe)	field-effect idea using a p-n junction.
23 Jun 45	WS (nbe)	**quantitative estimate of the degree of failure of thin-layer field-effect experiments.**
19 Mar 46	JB (nbe)	**proposal of surface states to explain failure of 23 Jun 45.**
12 Mar 47	WS (nbe)	very preliminary theory of p-n junction resistance.
4 Apr 47	WS (nbe)	lightning arrestor using alternating p- and n- type layers.
24 Apr 47	WS (nbe)	**diffusion theory of minority current to a reverse-biased p-n junction.**
19 May 47	WS (nbe)	theory of reverse currents in p-n junctions including Zener effect.
16 Sep 47	WS (nbe)	n-p-n structure for a high-frequency thermistor involving electron flow over the potential maximum of the p-layer.
17 Nov 47	WHB (nbe)	**overcoming surface states using Gibney's suggestion of changing bias of the electrolyte.**
4 Dec 47	WHB (nbe)	**success of several device ideas including WS proposal to modulate p-n junction resistance with voltage applied to a drop of electrolyte over the junction.**
8 Dec 47	WS (nbe)	**concepts of a junction, field-effect transistor and of voltage gain by using reverse bias on a p-n junction.**
8 Dec 47	WHB (nbe)	**achievement of voltage and power gain by using reverse voltages on "high-back voltage" germanium.**
16 Dec 47	WHB (nbe)	**first record of voltage and power gain in a point-contact transistor.**
31 Dec 47	WS (nbe)	**a p-n-p structure almost – but not quite – involving minority carrier injection into a base layer.**
23 Jan 48	WS (nbe)	**the conception of the junction transistor.**
18 Feb 48		J.N. Shive reported at a conference his observations on the double-surface transistor.
7 Jun 48		J.R. Haynes (nbe) first version of Haynes-Shockley experiment showing that a positive point injects holes into n-type germanium that drift in an electric field.
26 Jun 48	WS:	**patent application filed for junction transistor.**
1 Dec 48		E.J. Ryder and WS published as letter (received 1 Dec 48) by the Physical Review reporting hole injection lowers bulk resistance of n-type germanium. (The experiment was done several months earlier.)
7 Apr 49		"Existence proof" for the junction transistor. Made by dropping molten p-type germanium onto hot n-type and sawing the resulting junction to make two strips of p-type lying on an n-type plane. Preparation by R. M. Mikulyak under the direction of M. Sparks.
'49 and 50		Good p-n junction prepared by M. Sparks and G. K. Teal by changing the doping of the melt in a germanium crystal grower by "pill dropping" and subsequent double doping to make n-p-n structures. The development and publication of junction transistor theory added momentum to the program.
Early 51		Realization and demonstration of the first micro-watt, grown-junction transistors.

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

6.5.2 The point contact transistor

The discovery of the transistor effect occurred during work which commenced at the Bell Telephone Laboratories in early 1946. **Bardeen 1956, p77**   The study of semiconductors was just one   aspect of a programme of solid state research led by S.O. Morgan and William Shockley. In Shockley's initial group were Walter Brattain who worked mainly with surface properties and rectification, G.L. Pearson, concerned with bulk properties of materials and John Bardeen, a theoretician.   They were subsequently joined by a physical chemist, R.B. Gibney and a circuit expert, H.R. Moore.   The general aim of the programme was to obtain as complete an understanding as possible of semiconductor phenomena, not in empirical terms, but on the basis of atomic theory.

Initial attempts to make an amplifier device were based on the field effect – electrodes were applied to the surface of a semiconductor wafer to create an electric field which would influence a current flowing in the semiconductor. Difficulty was experienced because the effect measured was several orders of magnitude below what was expected from theoretical calculations.

John Bardeen eventually postulated an explanation which was based on the existence of "surface states" – fields created by hanging bonds resulting from the crystalline nature of the semiconductor. In an attempt to characterise these surface states, he and Walter Brattain applied a potential to a germanium wafer and made measurements using two pointed metallic probes which were pressed against the surface. Serendipitously, they found that a small current flowing between one electrode
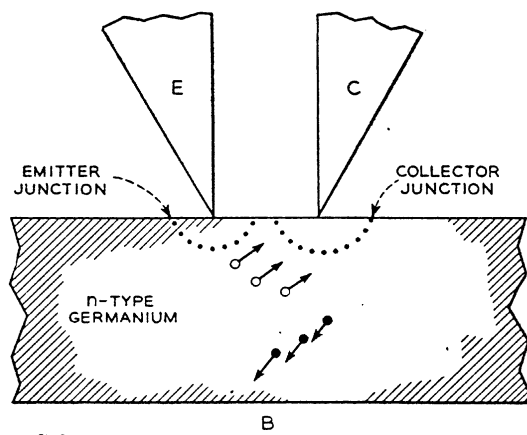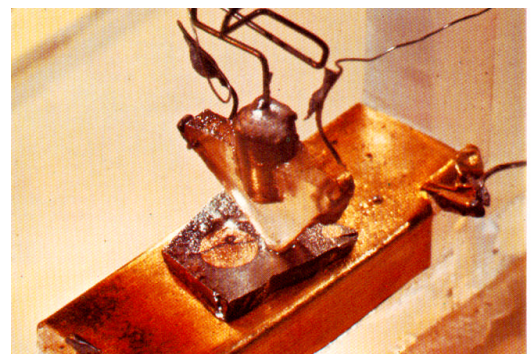

Morton 1952

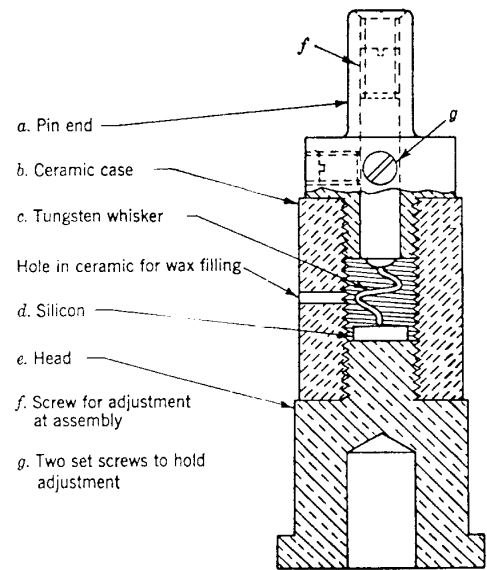Fig. 6.7 Principle of the point contact transistor

Weber 1981

Fig. 6.8 Bardeen and Brattain's point contact transistor

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

and the wafer could control a much larger current flowing between the two electrodes. They connected this arrangement into a circuit with feedback and obtained oscillation. This demonstrated that the device would serve as an amplifier.
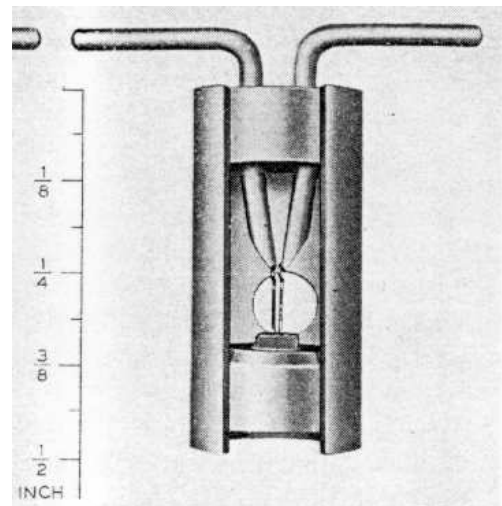
The Type A, the first production version of the point contact transistor, used techniques which had been developed for radar detector diodes during the second world war. **Torrey 1948, p16** Two phosphor bronze wires 12.5μm in diameter were pointed, bent and welded to nickel mounting wires which were previously moulded into a plastic plug. The plug was pressed into a metal cylinder. A germanium wafer 500μm thick and 1.25mm square was soldered on to a brass mounting-plug which was also pressed into the cylinder. The phosphor bronze whiskers were placed in contact with the germanium wafer and "formed" by passing an electric current which welded them to the crystal. The unit was then vacuum-impregnated with wax through a hole in the cylinder. **Shockley 1950**

a. Pin end
b. Ceramic case
c. Tungsten whisker
Hole in ceramic for wax filling
d. Silicon
e. Head
f. Screw for adjustment at assembly
g. Two set screws to hold adjustment

Torrey 1948

Fig. 6.9 Bell System point contact diode

Morton 1952

Fig. 6.10 The Type A, the first production version of the point contact transistor

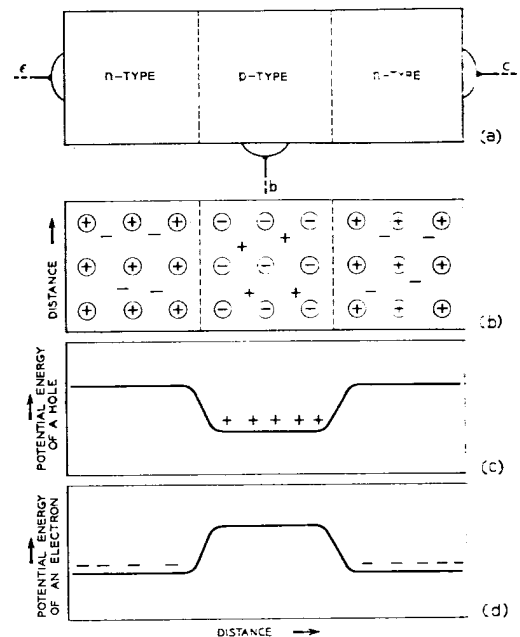### 6.5.3 The junction transistor

Whilst Bardeen and Brattain were carrying out their experimental investigation, Shockley was looking at semiconductors from a theoretical viewpoint. **Shockley 1950** Starting from Wilson's electronic theory, **Wilson 1931** he considered the tetrahedral structure of the germanium crystal and proposed a device fabricated from a single crystal in which a narrow region having one extrinsic type of conductivity (due to donor

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

or acceptor impurities) was sandwiched between two regions of the other extrinsic conductivity type. He calculated that, for a defined width of the central region, a minority carrier current flowing between the two outer regions will bear a fixed relationship to the current flowing to an electrode connected to the central region. This could be the basis for amplification. Although an application for a patent for this idea was filed on 28 June 1948, it was not possible at that date to make a working device.
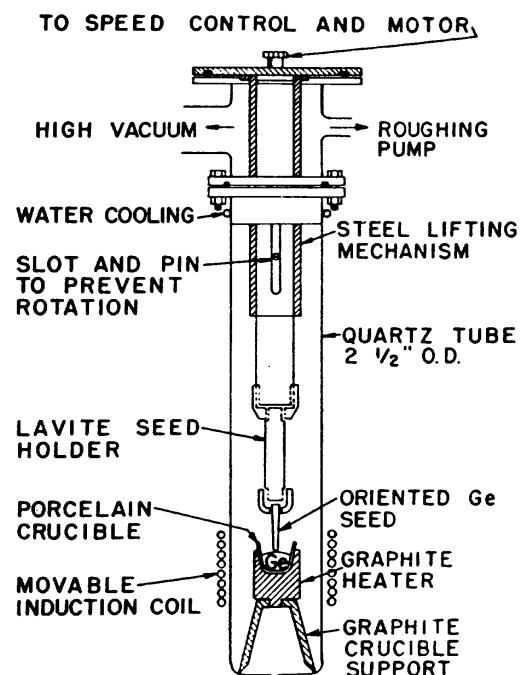

Shocley 1952b

Fig. 6.11 Shockley's model of a junction transistor

6.5.4 Growth and purification of single crystal material

6.5.4.1 Czochralski growth

Semiconductor research received a great impetus during the second world war, when point contact diodes, based on silicon and germanium ingots were used as detectors in radar receivers. Lark-Horowitz and his co-workers at Purdue University investigated the properties of different dopant materials. They found that the addition of small quantities of other elements such as tin modified the electrical characteristics of germanium and increased the inverse voltage which diodes made from it could withstand.

Materials technology was the key to the post-war advance at Bell Labs. A clutch of metallurgists, chemists and physicists

Roth 1952

Fig. 6.12 Growing a single crystal by the Czochralski method

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.
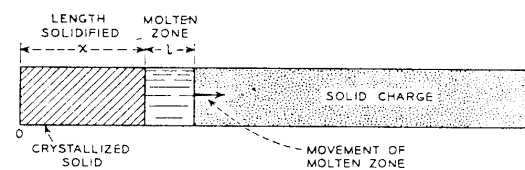
6.Third case study – the solid state advent

worked on the refining and characterisation of germanium. In order to simplify the programme, a decision was made to experiment with single crystals as well as the polycrystalline ingots which were generally available. Large single crystals were prepared by dipping a small seed crystal into a crucible of molten germanium, a development of a technique which had been developed many years earlier. **Czochralski 1917** Germanium expands on solidification, so crystals float on top of the molten liquid. By cooling the melt from the top it is possible to prepare large crystals. In this process, germanium solidifies on a seed, extending its crystal structure. Small dice were prepared from the drawn crystals by sawing, with a diamond saw, along the crystallographic planes. Work damage which was incurred during mechanical sawing and lapping processes was removed by chemical etching with etchants based on hydrofluoric acid, which left a highly polished surface on the germanium wafers. These etchant materials (in particular, the so-called CP4 and CP8) played a major part in the later successful development of junction transistors, but they received little publicity or acknowledgement. (As a general comment, it may be remarked that the contribution of the materials scientists, which was the major factor in the success of the development of the transistor, has been greatly underplayed in comparison with the contribution of the device physicists, in particular, Bardeen, Brattain and Shockley, who received a Nobel Prize for their work.)

6.5.4.2 Zone refining and zone levelling

The solubility of impurities is not usually the same in the liquid and solid phase of a host material. This property of matter was the key to development in Bell Laboratories of methods of purification of germanium and silicon which led to the practical realisation of Shockley's junction transistor.

W.G. Pfann, a metallurgist working in Shockley's team, used a modification of a method developed by P.W. Bridgeman **Bridgeman 1925** in the 1920s to produce single crystals of germanium. In this method, a charge of a molten alloy is slowly frozen
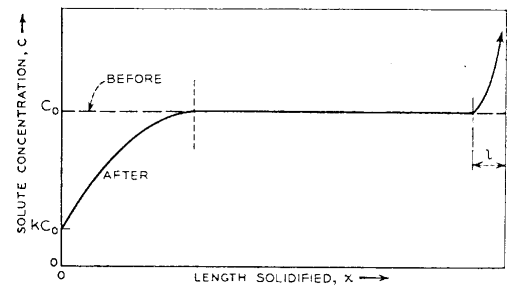
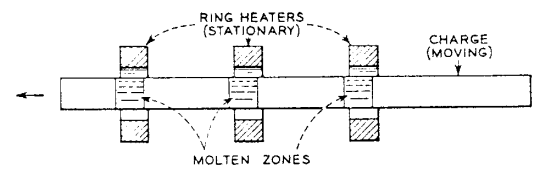

Pfann 1952a

Fig. 6.13
Solidification by zone melting

6. Third case study – the solid state advent

from one end. Pfann adapted the technique by starting with a solid charge. He created a small molten zone within this charge using radio-frequency induction heating. By moving the charge in relation to the RF field, he caused the molten zone to traverse the crystal. Due to preferential segregation of impurities in the liquid phase, he created a fairly uniform concentration of impurities in the central region, with a depleted zone at the starting end and a zone of excess impurity at the finishing end. This was due to the solidification of the last molten zone which contained a large proportion of the impurity which was more soluble in the liquid phase.

Pfann 1952a

Fig. 6.14
Concentrations of impurity before and after single-pass zone melting

Pfann 1952a

Fig. 6.15
Multi-pass zone refining

A reverse pass of the molten zone increased the uniformity of distribution of impurity. **Pfann 1952a, p749** Repeated passes of the molten zone caused most of the impurity to be transferred to one end of the crystal which could thus be progressively purified by this process.

Germanium was purified by means of this zone refining technique by placing it in a graphite boat and moving it horizontally in relation to the RF field. Silicon, on the other hand, was much more reactive and could not, for example, have been placed in a graphite crucible. The silicon crystal was therefore mounted vertically and the molten zone kept in position by surface tension. This technique was known as float zone refining.

6.5.4.3 The taming of deathnium

The successful operation of bipolar transistors depends on the preservation of minority charge carriers as they diffuse across the base region. Minority carrier lifetime was measured by the creation of hole-electron pairs by shining light on to a semiconductor specimen. It was found in the early days to be a parameter which varied greatly from crystal to crystal. Various mechanisms were proposed for the creation of

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

recombination centres – faults in the crystal lattice and surface imperfections were amongst the agencies considered. As the materials technology improved, so the effect of these mechanisms was reduced, but it was still found that certain crystals had an inexplicably high recombination rate. This was attributed to a phenomenon known as "deathnium". **Shockley 1952b, p1297** Eventually, the presence of deathnium was found to correlate with the presence of deep traps within the energy-level diagram of the semiconductor. These traps were created by small levels of certain impurities such as gold or iron which were inadvertently introduced during some of the early stages of processing that the crystals underwent. A particular source of the problem was iron introduced from materials processing furnaces.

Once the cause of the phenomenon was identified, its effect was minimised by techniques such as confining the use of particular furnaces to specific processes. It was then turned to advantage by the development of a process for deliberately introducing gold into silicon planar transistors to enhance their switching speed. **US Pat 3184347** This was achieved by depositing gold on the wafer from which the transistors were to be formed and heating the silicon to a high temperature to allow the impurity to diffuse into it.

6.5.4.4 Vapour phase deposition

At Bell Labs in the early 1950s, Gordon Teal and Howard Christensen investigated the pyrolitic deposition of volatile semiconductor halides. The desired introduction of small amounts of donor and acceptor impurities was effected by the inclusion of small amounts of volatile compounds of the dopant elements into a stream of gases which was passed over slices of semiconductor crystals which were heated to high temperatures in a furnace known as a reactor. With appropriate temperatures and gas flows, it was found that it was possible to deposit thin layers of semiconductor which extended the crystalline lattice of the heated substrate **US Pat 2692839** By changing the relative proportions of the compounds in the gas stream, it was possible to control the electrical characteristics of devices which were constructed subsequently. This process of epitaxial growth was to exert a major influence on the evolution of semiconductor device manufacture. It first began to be used as a production technique

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

in the late 1950s, but did not play a significant part in the economic differentiation of the players in the semiconductor industry because, due to a Consent Decree, which the Bell System was forced to concede as a result of anti-trust proceedings, the process was made universally available to all US companies.

6.5.5 The grown junction transistor

The first junction transistors were made by a modification of the Czochralski method. A seed crystal was drawn from a melt of lightly-doped *n*-type germanium. After the first part of the crystal had formed on this seed, a pellet of indium (an acceptor impurity) was added to the melt, swamping the influence of the first dopant and changing the conductivity of the next stage of the crystal to *p*-type. Shortly after, pellets of antimony (a donor impurity) were added to the melt, changing the conductivity of the drawn crystal to *n*-type once more.   Thus the crystal which finally resulted was predominantly *n*-type, but had a thin region of *p*-type material in the middle.   The location of this sandwich was determined by a combination of chemical etching and photo-electric probing. The crystal was then sawn into small rods, each containing an *npn* structure.   Contacts were applied to the individual regions and the device encapsulated.

The addition of dopant to the semiconductor melt meant that only one sandwich could be produced for each crystal pulled because, if the process were to be repeated, the doping levels would be too great to make satisfactory transistors.   Most of the material in the crystal was not useful. Although it could be re-purified and thus recycled, this was expensive and time-consuming.
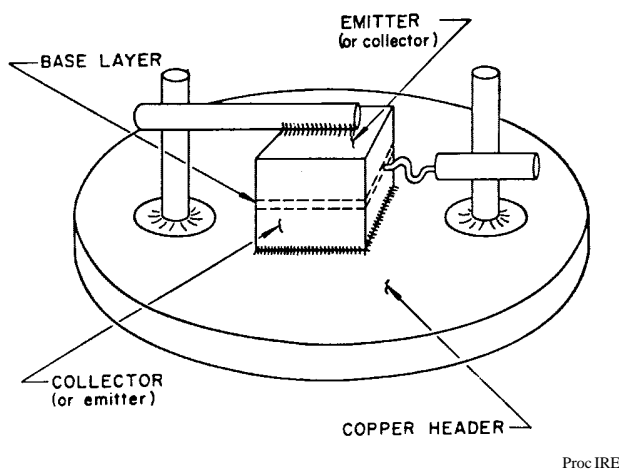

Fig 6.16 The construction of a grown junction transistor

In a later development, it was found that *npn* sandwiches could be formed by starting with a melt which contained both donor and acceptor dopants and varying the

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.
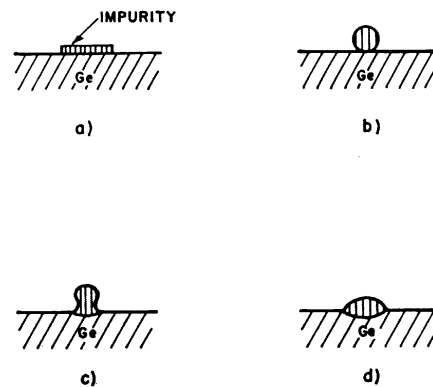
6.Third case study – the solid state advent

rate at which the crystal was pulled. This increased the yield of devices per crystal, but it was still difficult to produce them to a predetermined electrical specification. Grown junction transistors also suffered from the fact that the region of high-resistivity semiconductor adjacent the collector-base junction created a parasitic resistance which had a deleterious effect in electronic circuits.

6.5.6 The alloyed junction transistor

A major influence on the early development of the transistor was the army of metallurgists working in the various companies. One of the problems they had to solve was the difficulty of making contact to the active regions of the germanium crystal. Many of the metals which might have been considered formed brittle alloys with germanium. These alloys possessed different expansivities and thermal cycling in the

Jenny 1953

Fig. 6.17 Stages in making an alloyed contact to germanium

operation of the device caused cracks, that is, if failure had not already occurred in making the initial contact.

Among those metals which were tried successfully, was indium. As well as being ductile, it was an acceptor dopant. An alloyed contact to *n*-type germanium would thus create a *pn* junction which would exhibit rectifying properties. Hall and Saby at General Electric **US Pat 2999195, Saby 1951** and Pankove at Radio Corporation of America **US Pat 3005132** proposed the fabrication of a transistor by alloying pellets of indium to opposite sides of a wafer of *n*-type germanium. This proved to be a

Proc IRE

Fig. 6.18 The structure of an alloyed-junction transistor

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

reproducible way of making transistors and was far more efficient than the grown-junction method since a large proportion of each Czochralski crystal could be divided by a combination of diamond sawing, lapping, ultrasonic slurry cutting and chemical etching into a multiplicity of starting wafers.
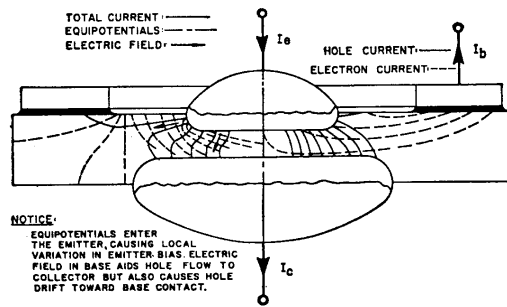

Fig. 6.19 Pankove's asymmetric alloyed-junction transistor structure

Initially, alloyed-junction transistors were made using identical pellets of indium, but Pankove at RCA calculated that improved performance would be achieved by using a larger pellet for the collector electrode. The resultant structure would be less susceptible to minority carrier recombination in the base region and would thus exhibit a high current-gain characteristic.

High emitter efficiency was achieved by Philips by constructing the emitter from a material (gallium) which had a high segregation coefficient [US Pats 3078397, 3078195] As pure gallium has a very low melting point, it was not suitable in its pure state to serve as an emitter electrode, It was therefore alloyed with indium which served as a suitable carrier.

A similar principle was adopted for the fabrication of alloyed junction *npn* transistors. The main donor impurities (arsenic and antimony) were not ductile, so they were alloyed with an inert carrier such as lead. Lead-antimony alloy was also used to make the ohmic contacts to the base region of *pnp* devices.

6.5.7 The drift transistor

The current gain of a bipolar transistor depends inversely on the distance between the collector-base and emitter-base junctions. If there is no electric field in the base region, charge carriers move between the emitter and collector electrodes by diffusion. The transit time, which controls the high-frequency performance, depends on the minority carrier mobility and the distance to be traversed. The first parameter is a characteristic of the semiconductor material from which the device is made, but the base width is dependent on the device design and method of fabrication. For better high

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

frequency performance, a narrow base width is desirable, but this is difficult to achieve at high yield with the alloyed-junction technique, since alloying depth varies with furnace temperature and the crystalline characteristics (surface etch-pit density and lattice perfection) of the semiconductor wafer. The latter, in particular, in the early and mid-1950s, when alloying was the predominant technique, were extremely variable.

Kroemer, at RCA, suggested [Kroemer 1953, 1954] that improved high frequency performance might be achieved by constructing the base region with a varying impurity distribution. This would have the effect of creating a background field – the drift field – which would speed the passage of minority carriers through the base. He made wafers suitable for such devices by diffusion of donor impurities. The drift transistor improved on the conventional alloyed-junction transistor by a factor of about 5:1, but the trade-off was that it suffered from low emitter-base breakdown voltages due to the high surface impurity concentration resulting from the powder diffusion process used to create the drift-field impurity gradient.
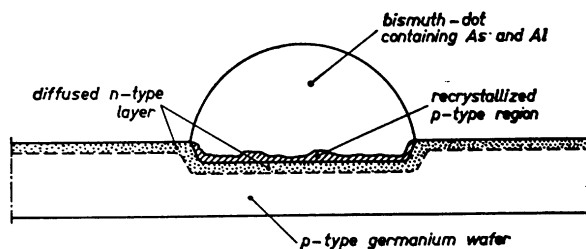
6.5.8 The post-alloy diffused transistor

Fig. 6.20 Formation of a *pn*-junction by a process of post-alloy diffusion [Proc IRE]

There were strong technical links between Philips and RCA and this led to cross-fertilisation of ideas. J.R.A. Beale, [Beale 1957] working for Philips' British subsidiary, Mullard, combined the principles of the drift transistor and the high-efficiency alloyed-emitter transistor. Commonly-used donor materials have relatively high diffusivities and low segregation coefficients. The complementary acceptor materials, on the other hand, have lower rates of diffusion, but high segregation coefficients. By combining an acceptor impurity and a donor impurity with a ductile inert carrier, and holding the carrier in contact with a germanium wafer for a protracted time at an elevated temperature, the donor impurity is induced to diffuse into the crystal, forming an *n*-type region remote from the interface. On cooling, the acceptor impurity re-crystallises, forming a *p*-type region adjacent the interface.

6.Third case study – the solid state advent

This technique was adapted to make the post-alloy diffused transistor (PADT). **US Pat 3160797, US Pat 3160799** A thin layer of *n*-type semiconductor was formed on a slice of *p*-type material by powder diffusion. Two small pellets of

Proc IRE

Fig. 6.21 The post-alloy diffused transistor

lead or bismuth, one containing small amounts of aluminium and arsenic, the other containing just the *n*-type impurity, were placed adjacent one another and the slice was heated to allow the impurities to diffuse from the pellets. The slices were then removed from the furnace and allowed to cool. The alloyed pellets were then protected by a small blob of wax and the exposed surface of the germanium slice etched, leaving an *n*-type semiconductor island under the pellets. The slice was then soldered to a nickel tab and leads attached to the pellets, which formed the emitter and base electrodes. The entire module was then hermetically sealed in a metal encapsulation.
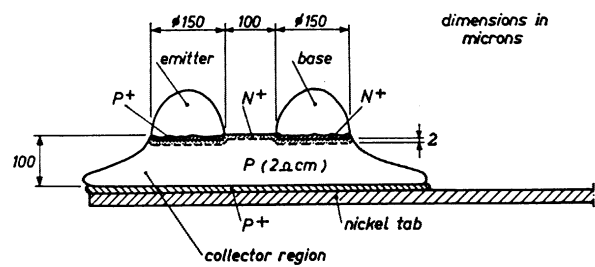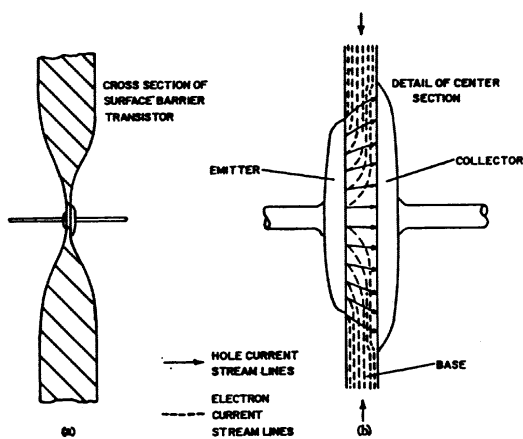
6.5.9 The surface barrier transistor – a technological *cul-de-sac*

Towards the end of the nineteenth century, Schuster and Braun independently discovered **Schuster 1874, Braun 1982** that rectification occurred at a metal-semiconductor contact. This was the basis for the selenium and cuprous oxide rectifiers which were

Bradley 1953

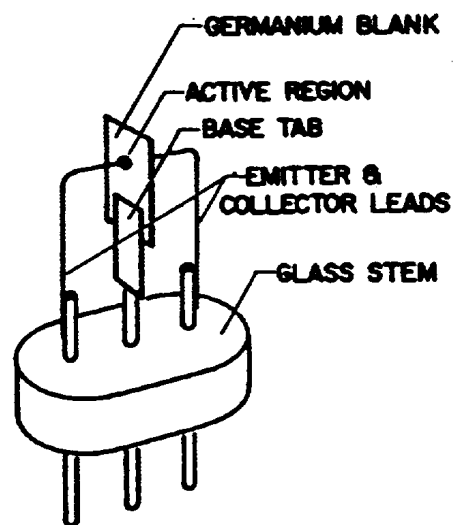Fig. 6. 22 The structure of the surface barrier transistor

Tiley 1953

Fig. 6.23 Mounted surface barrier transistor wafer
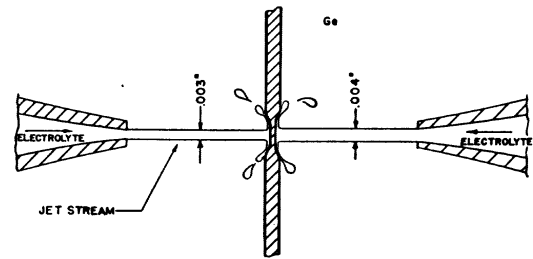
Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

developed commercially during the 1930s. **Grondahl 1927, Lange 1936** Schottky **Schottky 1933 Schottky 1939** and Mott **Mott 1939** developed the theory of rectification at a metal-semiconductor contact.

When the junction transistor was invented, a group of engineers working at Philco considered the possibility that a metal-semiconductor barrier might replace the *pn* junctions of a junction transistor. **Bradley 1953, Schwartz 1953, Tiley 1953** Their solution also addressed a drawback of the junction transistor, the difficulty of making a consistently narrow base region to achieve good high-frequency performance and high gain.

Philco developed a sophisticated automated plant for the manufacture of surface barrier transistors. A germanium slice was mounted on a jig and two streams of electrolyte directed at opposing surfaces. A potential difference was applied between the electrolyte and the semiconductor and opposing depressions progressively etched in its surface. As germanium is translucent to infra-red radiation, this property was utilised for process control. Infra-red radiation was directed at the slice, using the light guiding effect of the electrolyte jet. On the other side of the slice, a sensor detected the level of radiation, which
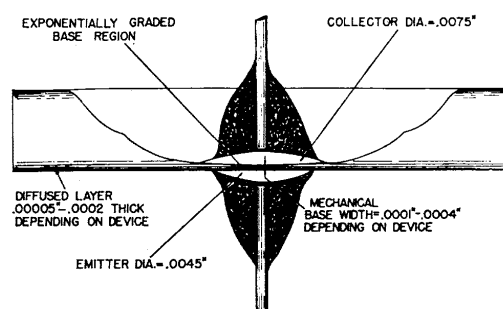

Tiley 1953

Fig, 6.24 Electrolytic etching a germanium wafer
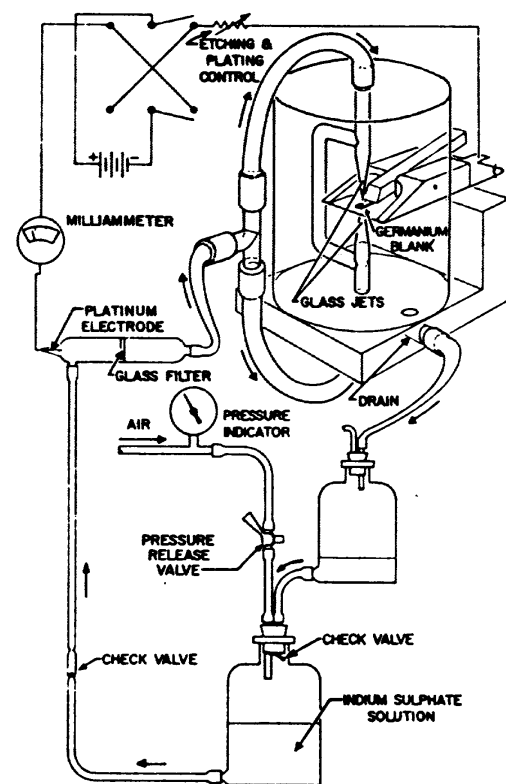
Tiley 1953

Fig. 6.25 Apparatus for making surface barrier transistors

Proc IRE

Fig. 6.26 Surface barrier analogue of the drift transistor

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

increased as the distance between the two etched depressions decreased. At a predetermined endpoint, corresponding to a specified thickness of the slice, the etching was halted. The electrolyte was changed and the surface of the germanium was electroplated with indium, forming a metal-semiconductor barrier.

Transistors formed in this way had a high-frequency performance which was extremely good compared with that of their competitors. They were, however, prone to catastrophic failure from a phenomenon known as "punch-through", which was a consequence of the structure, which combined a narrow base region with a relatively high resistivity semiconductor. Under an inverse bias, a charge-carrier depletion layer associated with the collector-base junction increased in size, eventually reaching the emitter-base junction when the voltage exceeded a critical value. At this stage, a large current flowed, destroying the transistor. Another major drawback was that the complex manufacturing plant required constant maintenance by skilled engineers. The very thin wafers were also very fragile.

With the surface barrier transistor there were several developments which were analogous to those which were taking place in the mainstream of technological progress in the industry. Following the introduction of the drift transistor, slices were diffused with impurity to provide a graded base region. These slices were etched from one surface only and the emitter electrode was applied to the diffused layer. In another development, aimed at reducing electrode capacitance, micro-alloy contacts were applied to the etched germanium region.

An interesting anecdote, is that the production rejects from the factory of the British licensee (Plessey) provided the launch-pad for the career of Clive Sinclair. He purchased the transistors by the barrow-load, tested them and selected those which were good enough for the purpose to assemble into kits for electronic equipment which he sold by mail order to eager home constructors.
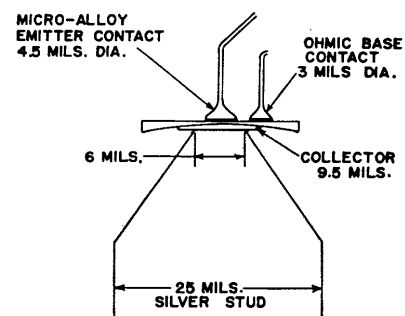

Proc IRE

Fig. 6.27 Micro-alloy transistor

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

6.5.10 The double-diffused mesa transistor

All the initial work on transistors was carried out with germanium. Device physicists were aware that silicon, an element from the same group of the periodic table, was a possible alternative, but it was much more reactive chemically and this was thought to be an intractable problem in the early days. Gordon Teal, one of the original team which
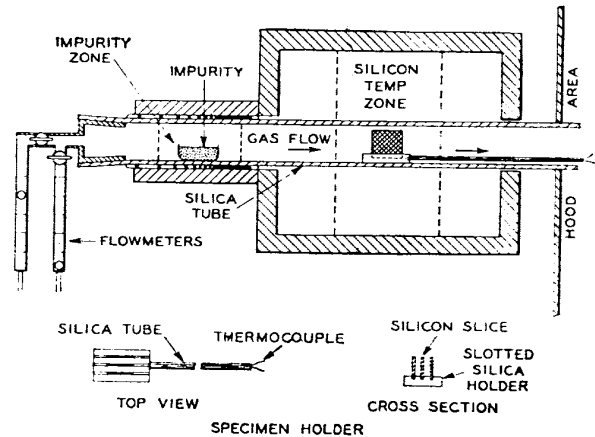

Proc IRE

Fig. 3.28 Silicon slice diffusion furnace

developed the transistor at Bell, moved to Dallas to work for Texas Instruments, a small company with no previous track record in electronic component manufacture. He caused consternation in the industry at an IRE National Conference in Dayton, Ohio, on May 10,1954, when he announced the successful manufacture of the first silicon transistors. [Appendix 8]

Although silicon possessed a wider band gap than germanium and silicon devices might therefore be expected to be less prone to breakdown, charge carrier mobility was lower and, for similar device geometries, high frequency performance was inferior. The metallurgy of silicon alloys was also less felicitous. Indium and lead, which were ductile metals used for contacts to germanium, did not alloy successfully with silicon, whilst aluminium, which did, formed a brittle alloy with a markedly different expansivity from that of the semiconductor crystal. The consequence of this was that thermal cycling during operation caused electrical degradation and, on occasion, mechanical failure of alloyed junction devices produced by techniques which had been developed for germanium. There was, as a result, a powerful incentive to develop fabrication techniques which would be better suited to the properties of silicon.

The most pressing need was to increase the high frequency performance to compensate for the lower minority carrier mobility of silicon. An obvious way to do this was by making the base region narrower.
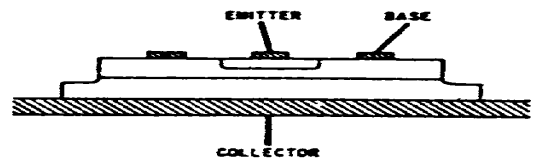
6.Third case study – the solid state advent

Silicon differs from germanium as it has a stable oxide which can be formed as a layer on the surface of a slice by heating it to a high temperature in an oxidising atmosphere. Frosch and Derick, at Bell Labs [US Pat 2804405] found that the

Petritz 1962

Fig. 6.29 Double-diffused mesa transistor

glassy layer of silicon dioxide acted as a barrier which would inhibit the diffusion of impurities into the silicon slice. By selectively removing the oxide from areas of the slice surface using a combination of photolithography and chemical etching with hydrofluoric acid-based etchants, they could selectively diffuse impurities into predetermined regions of the silicon slice.

Slices were prepared by diffusing boron (an acceptor) into the surface of high-resistivity *n*-type silicon. During this processing step, a layer of silicon dioxide about one micrometre thick was grown on the surface. Using photolithographic methods which had been developed by the printing industry, windows were etched in this oxide layer. The slices were then placed in a second furnace and phosphorus (a donor) allowed to diffuse into the silicon to form localised *n*-type regions which would serve as the transistor emitters. A small region round each n-type diffused region was masked photolithographically and the remainder of the p-type layer etched away. The resulting flat-topped structures, which were separated into individual dice by a process of scribing the slice with a diamond stylus and breaking by applying stress to the crystal lattice, were known as mesa transistors due to their similarity to the flat-topped mountains found in Mexico.

Contact to the emitter and base regions was made by means of a layer of aluminium evaporated on to the surface of the silicon slice *in vacuo*. At first, this evaporation was confined to the individual regions by use of a mask, but, as photolithography improved, the delineation of the contacts was performed by chemical etching.

Devices which were produced in this way had a base width of about 0.5μm compared with about 25μm for an alloyed junction germanium transistor. This greatly reduced base width permitted operation at frequencies of more than 100MHz.

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

External electrodes were connected by means of thin aluminium or gold wires using the technique of thermocompression bonding. $^{\text{US Pat 3006067}}$ The silicon die was mounted on a metallic header and heated to a temperature slightly below the melting point of the contact wire which was positioned using a glass capillary tube. Pressure was applied with the end of this tube, lowering the melting point and creating a firm bond when the pressure was removed.

6.5.11 The planar transistor

The electrical characteristics of transistors were degraded by spurious surface effects in the vicinity of the emitter and, particularly, the collector junction. Many early transistors failed as a result of surface contamination. An indication of the significance of this phenomenon is that many patents relating to point contact and germanium junction transistors were concerned with surface preparation and hermetic encapsulation. $^{\text{Appendix 7}}$

Petritz 1962

Fig. 6.30 Double-diffused planar transistor
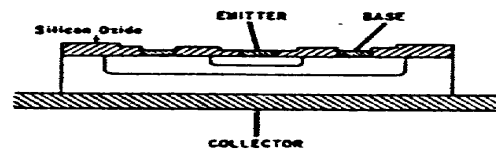
Jean Hoerni at Fairchild Semiconductor discovered that if the layer of silicon dioxide (silica) which formed during the diffusion of impurities into the wafer were left in place, the resulting transistors were much more stable. In transistors made in this way, the base regions were formed by localised diffusions, like the emitters, rather than by etching mesas. The devices had much flatter geometry and were thus known as planar transistors.

As the quality of materials improved, the lifetime of minority carriers in the base region increased. Many transistors were used for switching applications and this persistence of minority carriers gave a large switch-off time (t_{off}). Hoerni discovered $^{\text{US Pat 3184347}}$ that the diffusion of gold into the transistor reduced the minority carrier lifetime and improved switching performance. Thus he harnessed the influence of deep traps which had given rise to "deathnium" and had been responsible for failure of early germanium devices. What had been an undesirable effect was now a positive virtue.

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6. Third case study – the solid state advent

6.5.12 The epitaxial transistor

The electrical performance of a transistor is degraded by parasitic effects. For instance, an inversely-biased junction, such as the collector-base junction acts as a capacitor, whilst the semiconductor substrate between the collector junction and the header on which the chip is mounted has a resistance which is dependent on the conductivity of the semiconductor material. The first double-diffused and planar transistors were fabricated in silicon which had a resistivity of the order of 1-2 ohm-cm. (It was necessary to use this material to prevent failure of the junction at low reverse voltages.) As a consequence, the parasitic resistance was also relatively high.

Teal and Christensen at Bell Labs had shown [US Pat 2556711, US Pat 2692839] that it was possible to deposit a layer of germanium from the vapour phase by pyrolytic decomposition of its halides. Control of the resistivity could be achieved by adding a predetermined amount of the halide or other volatile compound of a significant impurity. Thus *n*-type germanium could be deposited by passing a mixture of germanium trichloride and arsenic trichloride over a heated slice of germanium. With the correct combination of temperature and gas flow, this deposited layer would extend the crystal lattice of the substrate. The process was known as epitaxial growth.

Since the doping level of an epitaxial layer could be controlled, it became possible to grow a high-resistivity layer on a low-resistivity substrate. Transistors formed in the high-resistivity layer would have high breakdown voltages, but, as they also had a low resistivity substrate, they would not suffer to the same extent from parasitic resistance in series with the collector.

6.5.13 The integrated circuit

As early as 1952, G.W.A. Dummer had suggested that electronic components might be formed in a single solid block. [Morris 1990, p45] Jack Kilby, at Texas Instruments, constructed the first functional monolithic integrated circuit in 1958. [US Pat 3138743] He used mesa etching to separate
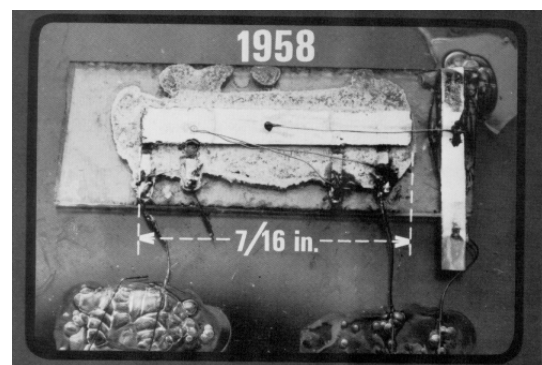

Fig. 6.31 Kilby's first integrated circuit

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

the components and made the ancillary connections with thin gold wires, which were attached by thermocompression bonding. Contemporaneously, Bob Noyce, at Fairchild Semiconductor, formed individual circuit elements in a single block of silicon by the planar process. These components were interconnected by a miniature printed circuit etched in a layer of aluminium which had been evaporated on to the silicon dioxide surface layer.

The first integrated circuits were relatively simple – a single transistor with a load resistor and a group of diodes connected to provide separate inputs. The combination functioned as a NOR gate. As skills in the techniques of diffusion, epitaxy and photolithography improved, so the complexity of integrated circuits and their manufacturing yields increased.
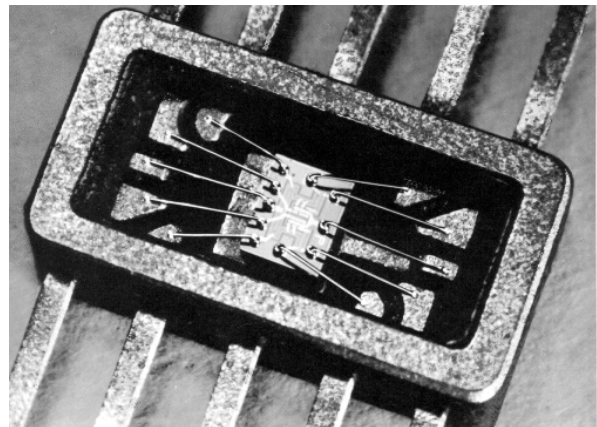
Weber 1981

Fig.6.32 Early RTL integrated circuit

The invention of these processes marked the transition from the Schumpeter A-stage to the B-stage of the evolution of the semiconductor industry.

6.6 Influences on the development of the semiconductor industry

A three-electrode semiconductor amplifying device was first invented in the 1920s by Julius Lilienfeld. He proposed device structures which were brought to practical realisation some 30-40 years later. His inventions were patented, but not developed commercially because the materials available to him were not capable of sustaining the minority carrier charge flow necessary for viable operation. Although his ideas were soundly based, Lilienfeld did not have the resources or perseverance to push them through to fruition.

Hilsch and Pohl revisited the concept of a solid-state amplifier in the 1940s. They did not succeed because they tried to simulate a thermionic triode structure in an ionic crystal. Although they achieved a measurable current gain, the properties of the

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

materials inhibited operation at frequencies which would have been commercially useful.

Wartime activity, in particular at Purdue University, was responsible for the development of the materials which provided the springboard for the discovery of the transistor effect. Another essential component was the existence of a culture in the Bell Telephone System which permitted the expenditure of substantial funds on a programme of fundamental research. Although such research could conceivably lead to the holy grail of a solid state replacement for the unreliable thermionic valve, there was no requirement for it to be so targeted.

The Bell Laboratories were a centre of excellence, and the team which was assembled was of the highest quality. It was also multi-skilled, including theoretical and applied physicists, chemists, metallurgists and electronic engineers. As befits a major US corporation, it was well-briefed legally and performed a thorough job of protecting the intellectual property generated by the research.

Constraints which had arisen from anti-trust actions prevented the Bell System from exploiting the invention of the transistor in the most direct way – by making and selling the devices. An imaginative licensing programme was therefore introduced by seminars which provided a taster of the significance of the invention. This was followed up by well-organised transfer of know-how to the licensee organisations. Success was further assured by inviting only companies of substance to participate.
[1965] RPC 335

An artificial situation existed initially as an anti-trust consent decree effectively forced Bell to pass on its semiconductor know-how to others. The transfer of know-how was soon reinforced by movement of personnel as Shockley and Teal, who were members of the original team, sought pastures new. Another early influence was Shockley's irascible personality. Scientists and engineers who had joined his Palo Alto start-up, found that he was impossible to work with, and left to create their own companies. This was the beginning of the "Silicon Valley Effect" – technology transfer through the founding of new companies.

There was a great *cameradie* amongst the Californian semiconductor community, which shared technical knowledge, oblivious to the proprietary nature of the know-how.

6.Third case study – the solid state advent

The semiconductor manufacturers possessed a vast core of technical skills. As soon as a new idea was made public, many rivals were in a position to exploit it, a factor which is confirmed by the short lead time between seminal patents and daughter inventions. Patents were either cross-licensed or ignored **Appendix 1** so the lead-time advantage of innovation was negligible.
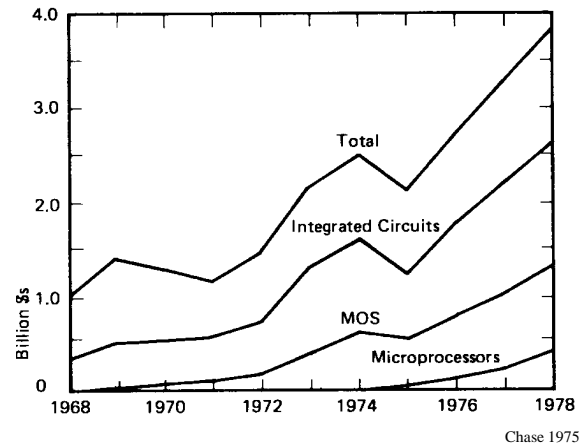
Chase 1975

Fig. 6.33 Annual fluctuations in semiconductor production in USA

Although the first wave of manufacturers was drawn mainly from the ranks of those who made thermionic valves, the predecessor product, by the time the industry was ten years old, many of these had fallen by the wayside and newcomers, such as Texas Instruments and Fairchild, had taken their places. The Not-Invented-Here syndrome was probably playing a significant part.

Market growth of the order of 30% per annum attracted many newcomers to the industry. Cyclic profitability and zero average net profit caused many of them to quit. Market leaders came and went with the introduction of new manufacturing methods.

During the fifteen years after the initial discovery, the dominant manufacturing technology changed every three to four years. The point contact was superseded by the grown junction, which was more robust. The alloyed junction, which followed, reduced manufacturing waste and parasitic collector resistance. Silicon replaced germanium to make devices less sensitive to the effects of high temperatures, whilst the associated double-diffusion and mesa etching processes gave much better high frequency performance. Planar surface passivation techniques introduced long-term stability and permitted the use of plastic encapsulation. Finally epitaxy gave rise to flexibility in device topography and presaged the fabrication of complex integrated circuits.

Advances in device technology were governed by properties of materials and the materials scientists and metallurgists were the unsung heroes who made the industry viable. **Petritz 1962** Silicon took over from germanium partly because it had a wider band gap, but mainly because it had a stable glassy oxide which provided surface passivation.

6.Third case study – the solid state advent

Alloyed junction transistors were predominantly *pnp* devices, despite the fact that electron mobility was of the order of three times that of hole mobility, because indium, an acceptor dopant, was ductile and alloyed readily with germanium. Complementary *n*-type dopants, antimony and arsenic, were brittle and volatile at alloying temperatures. They had to be used in conjunction with an inert metal carrier such as lead, which did not wet the surface of germanium as well as indium did. The post-alloy-diffused transistor was made possible because donor impurities had higher diffusivities than acceptors, whereas the segregation coefficients, which were important for alloying, were higher with the *p*-type dopants. Gallium arsenide, which had a wider band gap and higher minority carrier mobility than silicon, did not supersede it because it was difficult to fabricate. Eutectic bonding processes, such as bird-beak and nail-head or ball bonding, used for external lead attachment were a serendipitous consequence of the metallurgy of gold and aluminium. (A counter influence, which caused many catastrophic failures until its mechanism was well understood, was "purple plague", a non-conducting intermetallic compound of gold and aluminium which formed at high temperatures.)

The commercial potential of the transistor was apparent, even before a practical implementation was developed. It was a salesman's dream – a product capable of satisfying a huge latent need. In the early days, when the volume of sales was low, the cost of market entry was commensurate. Potential participants either possessed the skills and equipment necessary, or could acquire them with very little outlay. Later, as markets increased in size, specialist equipment suppliers emerged and device manufacturers relinquished in-house equipment construction. The industry was a technological meritocracy. Companies with the most efficient processes succeeded; those which did not adapt to change, went to the wall.

The transistor and the *pn*-junction diode were the mainstream components of the semiconductor paradigm, as the thermionic triode and diode had been during the first age of electronics. Other thermionic and gas discharge devices also had their semiconductor analogues. Thus the thyristor was the equivalent of the thyratron, whilst the unijunction transistor and the four-layer diode played the role of the gas discharge tubes. The vertical junction field-effect transistor was able to mimic the electrical

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6. Third case study – the solid state advent

characteristics of the pentode valve and the zener diode provided a voltage reference as did the neon tube in valve circuits. The existence of so many comparable devices is indicative of the methodology of circuit design during the genesis of the semiconductor industry. The general approach was to attempt to translate the equivalent valve circuit by substitution of an appropriate semiconductor component. It was not until the invention of the integrated circuit and the move from analogue to digital electronics that engineers truly began to think in new terms.

6.9 Conclusion

As with the electric light and cable and wireless communications, the technological origins of the transistor may be traced to discoveries made some fifty to a hundred years earlier. Again, like the light and cable and wireless, it was invented at least twice before its time. The transistor finally got off the ground, mainly as a result of wartime research which produced new and suitable materials, but also due to the assembly of a polytechnic team of scientists of the highest calibre – analogous to Edison's laboratory at Menlo Park – by a company which allocated sufficient resources to a management which, in turn, had the creative vision to see the project through to a successful conclusion.

The transistor was developed as a result of a well-directed, broadly based investigation, which was carried out from first principles by a high-calibre, multi-disciplinary team. The discovery of the point contact transistor effect was a chance consequence of rigorous, comprehensive and thorough experimental procedure, but it provided a huge fillip to the project. The team leader, William Shockley, had developed the theoretical basis for the transistor over long period of time and was annoyed when Bardeen and Brattain stole his thunder with the fortunate discovery, but the cognitive dissonance which this engendered, resulted in the conception of a viable alternative by Shockley.

Production versions of the point-contact transistor were based on existing, well-established diode manufacturing technology, but they were soon replaced by the first junction transistors which were fabricated by a brute-force method and, like Edison's tar-putty incandescent lamp, were little more than a proof of the concept. This was,

6.Third case study – the solid state advent

however, sufficient to stimulate enthusiastic work on circuit applications. As with the metallic incandescent-filament lamp, many major paradigm shifts came from without and were based on combinations of materials with felicitous physical, metallurgical and electrical properties

The use of single crystal material rather than polycrystalline germanium which was more readily available for the development work, was an early, informed decision based on theoretical principles. Materials were chosen and their properties orchestrated to construct device topographies yielding desired electrical characteristics. Paradigm shifts were the mechanism by which this procedure was optimised.

The key to the successful initial development of the transistor was a polytechnic research team. Contributions were often made by scientists from different disciplines bringing their particular heritage as a contribution to the feast. Materials scientists chemists and metallurgists made major breakthroughs, but received scant praise for their efforts. Many technological barriers were surmounted *ad hoc* as they were encountered. This did not require unduly creative thinking, merely meticulous attention to detail and mastery of the underlying scientific principles.

The combination of many companies, employing strong R&D teams and enjoying free exchange of information meant that new advances could potentially come from any quarter. If successful, they would be rapidly adopted universally. Some new paradigms offered the prospect of commercial advantage, but, after a brief period of exploitation by the single company that introduced them, they were abandoned and that company reverted to the technological path being followed by the remainder of the industry. Not all technologically elegant solutions proved to be commercially viable. However, due to the underlying economic growth pattern of the industry, it was possible to try, then abandon, them if they were unsuccessful.

Electronic circuit applications for the new transistor expanded rapidly as new fabrication techniques yielded devices with ever improved characteristics. The result of the impact of the classic cash-flow J-curve associated with technological paradigm shifts was cyclic profitability and zero or negative average net profit. This caused many participants to quit – market growth of the order of thirty percent per annum, easy access to know-how and low cost of market entry in the early days attracted many newcomers

6.Third case study – the solid state advent

to the industry. Market leaders came and went with the introduction of new manufacturing methods. Research and development were also stimulated because the military potential of the new development attracted a large provision of funds by the US Government.

The most obvious choice of materials for device manufacture, from the point of view of desirable electrical characteristics, was often *not* adopted because alternatives were either easier to fabricate or produced devices which exhibited greater reliability. Silicon, which became the universal material of choice for the mainstream semiconductor industry was a compromise based on these alternative considerations.

As experience increased, threats turned into opportunities. For instance, "deathnium" which destroyed many early transistors, was found to be caused by deep-trap impurities, and the property was harnessed to make faster switching transistors for computers. Planar integrated circuits arose from treatment of a threat (formation of a surface oxide layer to counteract pollution) as an opportunity (use of the oxide layer as support for interconnections). This methodology was applied twice in one company – Fairchild – which, as a consequence, became the dominant company in the industry for a while.

The subject chosen for this case study provided a particularly apt example of change for technological reasons in the Schumpeter A-phase because (a) there was always a strong demand for the end product (b) an extremely high rate of growth year-on-year which meant that the payback from paradigm shifts was quick, and (c) mistakes could be abandoned without significant penalty. Change was also facilitated by formal and informal exchange of know-how within the industry.

This case study also provided a unique example of the influence of competition law at the start of a new development rather than when it is mature (which is the usual situation.) Because a Consent Decree was already in force when the transistor was invented, Bell Laboratories was forced to license all comers. To generate income, they provided know-how to permit the licensees to become established in the new industry. Potentially, this could have resulted in a completely different evolutionary characteristic because the sanction was applied from the outset. However, in the long run, the market leaders which emerged were those companies which were responsible for significant

Technological roulette – a multi-disciplinary study of the dynamics of innovation in electrical, electronic and communications engineering.

6.Third case study – the solid state advent

paradigm shifts. In the absence of this competition law influence, these actual market leaders may well have been different, but the overall structure of the industry would, in all probability, have been the same, because it was determined by innovations which offered a technological advance. New innovations came from without, possibly as a result of the "not-invented-here" syndrome. Absence of response-time monopolies, which would have permitted the advantages of technological paradigm shifts to be pressed home, and freedom-to-use paranoia, which negated the effect of intersecting patent monopolies, allowed competitors to play catch-up in characteristic ologopolistic fashion.

Ultimately, the universal adoption of planar diffusion and epitaxial deposition manufacturing techniques, and the transition from discrete devices to integrated circuits, marked the end of the gestation of the semiconductor industry. By this time the thermionic valve precursor was effectively dead, surviving in only a few niche applications and in price-sensitive markets where it still offered a financial advantage because it was a mature product and all costs were fully amortised. The next springboard was very large scale integration (VLSI), the so-called silicon chip, which is the subject of the next case study.