TECHNICAL DEVELOPMENT OF RADIO AND THE THERMIONIC VALVE

The three main technical areas of early radio were

(a) generation of high-frequency signals

(b) transmission and reception

(c) detection of signals

Hertz, 1886, demonstrated a system which proved the existence of electro-magnetic waves and action at a distance. He originally used the discharge of a Leyden jar as the source of power; but he understood electrical resonance and was able to tune the receiver and the transmitter to one another by adjusting inductance and capacitance, and this enabled him to get his transmission adequately using an induction coil as source. He measured the wavelength of the waves, their velocity (which he found equal to that of light), their frequency, and demonstrated reflection, interference, standing waves, etc. Using frequencies around 100 MHz he could detect the radiation with his small coil and air-gap over paths of many metres.

It is clear that Hertz laid the foundation for radio, but he did not foresee radio as a means of long-distance communication. It was Marconi who did this, and in 1895 he transmitted signals at Bologna over a distance of 1Km. He used for generation of the signals a sphere spark gap excited by an induction coil. For transmission and reception he used long vertical wire aerials suspended from tall masts. For reception he used the coherer. This he improved over preceeding designs by reducing its dimensions, and by employing nickel and silver filings between two silver terminal blocks in an evacuated sealed glass tube. Marconi patented his system and having failed to arouse interest in Italy, got the co-operation of Sir Wm Preece, the Engineer in Chief of the British Post Office. The P.O. co-operated very fully, and a whole series of successful experiments started in 1896. The tuning arrangements were called at first "syntony", and Sir Oliver Lodge and others collaborated in their development, with experiments first on Salisbury Plain. The sequence was then:

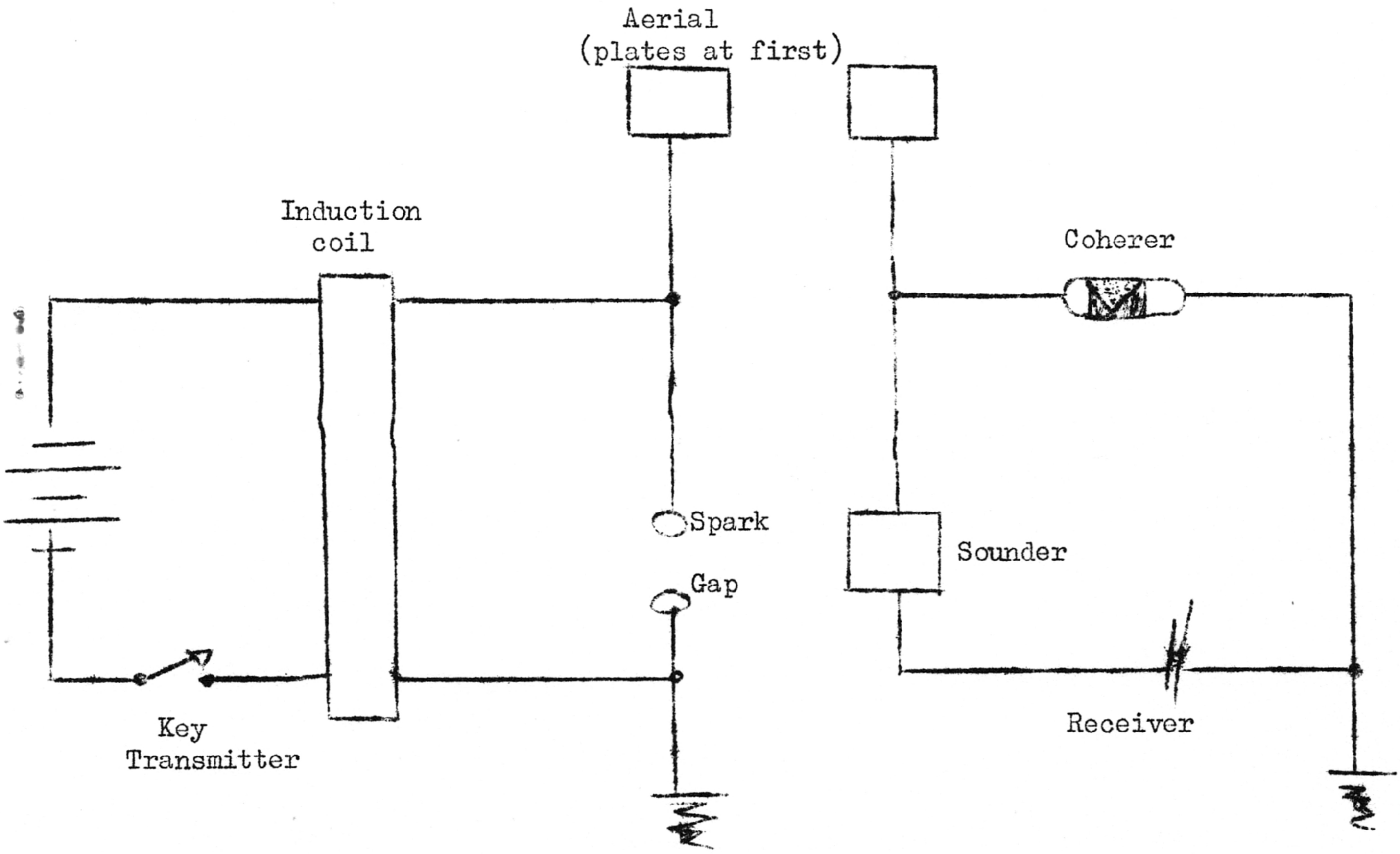
1897: 8-mile link across Bristol Channel

18-mile link Poole - Isle of Wight

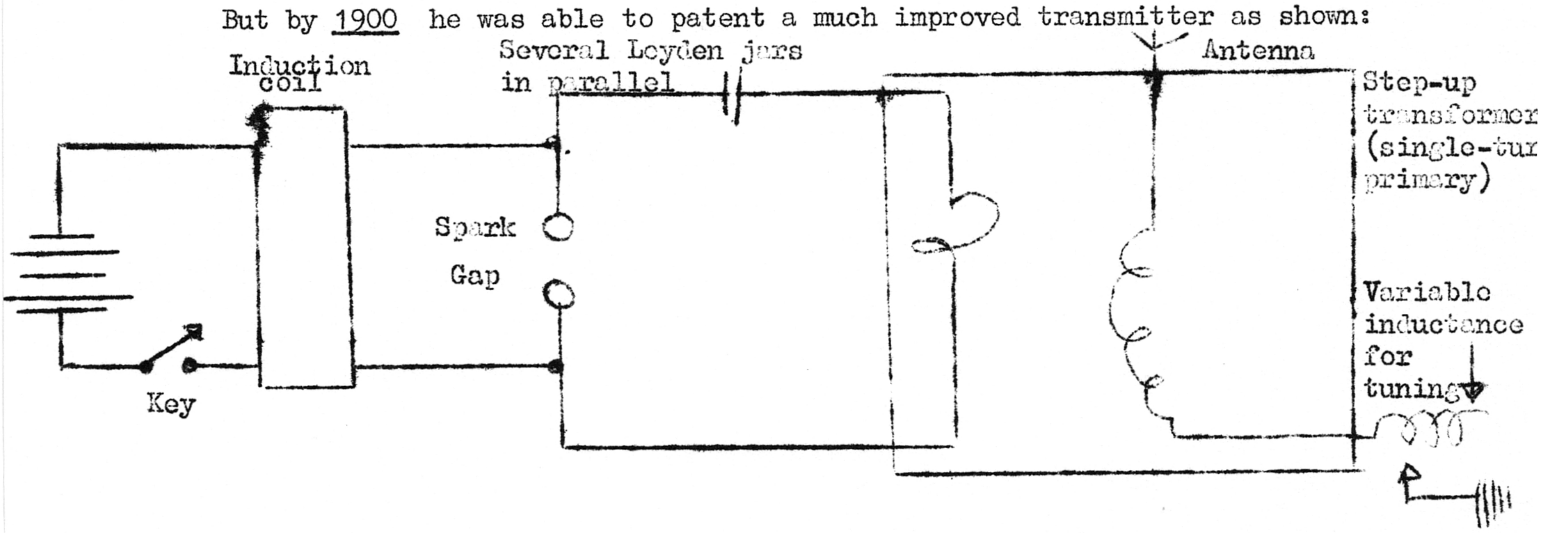
30-mile link Poole - St. Catherine's Point, using two simultaneous transmissions separated by tuning

1899: South Foreland Lighthouse - Wimereux, across English Channel

Marconi's earliest circuits can be represented thus:



But by 1900 he was able to patent a much improved transmitter as shown:



He had a similarly-tuned receiver. With these arrangements he could attain ranges of several hundred Kms and could use several simultaneous channels. This 1900 British patent was numbered 7777 and became famous through the legal battles fought over it.

On the basis of these improvements, Marconi was ready to try bridging the Atlantic in 1900. The first high power radio station was built at Poldhu in Cornwall, using (inplace of the induction coil) an alternater, driven by a 25 h.p. oil engine, with transformers stepping up the voltage to 20,000. The aerial system comprised 50 wires supported from 200 ft masts. Towards the end of 1901 Marconi travelled to St. Johns, Newfoundland where he used aerials supported by kites and balloons. The wavelength used was just under 1 Km. Three-dot signals were continuously being sent out from Poldhu, and on 13 December 1901 Marconi was able to receive them.

Although there was continuous development on the generation of signals, there were only two significant developments in their detection before the thermionic valve came into use. So we shall dispose of these now. The coherer had always been rather uncertain and unreliable and had required "decohering" after each train of waves by mechanical means. So the electromagnetic detector was attractive. This had its origin in the work of Ernest Rutherford in 1896. He used a needle which became magnetized when electromagnetic waves acted on it. Marconi improved it and patented his design in 1902. The principle seems very obscure and the descriptions are not very satisfactory: Dunsheath says "A loop of fine wires was made to travel slowly down the centre of two concentric coils which thus formed a transformer. One of these coils was connected in the receiving aerial circuit and the other to a telephone receiver. Near to the moving wires were disposed two horseshoe permanent magnets which thus magnetized the wire as it travelled along. When the high-frequency current from the receiving aerial passed through the primary coil it interfered with the magnetism in the tape and produced a "click" in the telephone. As the wave trains arrived at a frequency of some hundreds a second, the result was a continuous musical note in the telephone as long as the waves were arriving. By subdividing these on the Morse code system through the transmitting Key, telegraph messages were received by the operator wearing headphones.

The magnetic detector soon ousted the coherer.

Although the thermionic diode was invented by Fleming in 1904 (see later), it proved for a long time to be unreliable and erratic and did not come into general use. It was thus possible for another, and very cheap type of detector to come into use in 1906: this was the crystal detector, which was discovered by H.H.C. Dunwoody in 1906, using the rectifying properties of carborundum (silicon carbide). This was so very cheap that it led to the growth of amateur radio.

In the development of generation of signals, it was soon recognized that bursts of damped vibrations were not very satisfactory. A continuous wave system was required. W. Duddell in 1900 proposed the electric arc coupled to a resonant circuit. This only became really successful when Poulsen (a Dane) in 1903 enclosed the arc in hydrogen and placed it in a strong magnetic field. The frequency produced was about 1MHz.

Another C.W. generator was the high-frequency alternator. This originated with R. A. Fessenden, who got C. P. Steinmetz of General Electric (USA) to build a 10KHz generator in 1903. In 1904 Fessenden ordered a 100KHz generator and assumed it also would be designed by Steinmetz, who used a wooden core. Instead E. F. W. Alexanderson (a lower-ranking engineer) was given the project and used a laminated iron armature. Fessenden objected and the two argued stubbornly. Fessenden placed another order with another company in the end; GE completed the machine Fessenden had ordered according to his own design and it was a failure; the other machine was only partially successful, operating satisfactorily only up to 50KHz. It was the machine used for Fessenden's first successful radiotelephone at Xmas 1906.

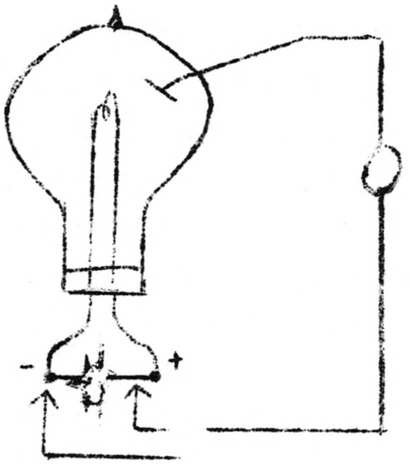
Alexanderson built one of his own generators for the American Telephone and Telegraph Co. in 1907, working at 100KHz. He also added a triode amplifier, and patented a multiple-tuned aerial. Marconi was very impressed by Alexanderson's system and saw it as the basis for his world-wide radio system. Unfortunately his plans for this were eventually thwarted by World War I. Nevertheless high-power, high-frequency alternators were widely used for long-distance radio until the high-power thermionic valve displaced them.

Efforts to transmit speech by radio began early. Fessenden was the main inventor in this field. He thought out the principle of amplitude-modulation (or better nowadays "envelope modulation"). As early as 1901 he had demonstrated voice transmission over a short distance. His idea was to place a telephone microphone in the aerial circuit and so modulate the otherwise constant output from an h.f. alternator. He probably understood the basis of modulation by non-linear elements as he had patented a magnetic modulator. He was finally completely successful in 1906, when on Xmas Eve he broadcast a programme of music and speech. Other people followed him, but the development of voice radio was surprisingly slow. Radio-telephony was still experimental after the World War I in 1919.

THE THERMIONIC VALVE

Modern radio has developed through the exploitation of the thermionic valve, and we cannot take the history of radio any further without looking first at the development of the valve. The very name "valve" gives away the fact that the earliest concepts were only of rectification, but in Britain the name stuck. The Americans called the device a "vacuum tube" or just "tube".

It all started with the "Edison Effect". In 1883, Edison had placed a metal plate supported on a wire inside the glass bulb of an electric incandescent lamp. The filament was of carbon. He connected a meter in the plate-to-filament cicuit, thus:-



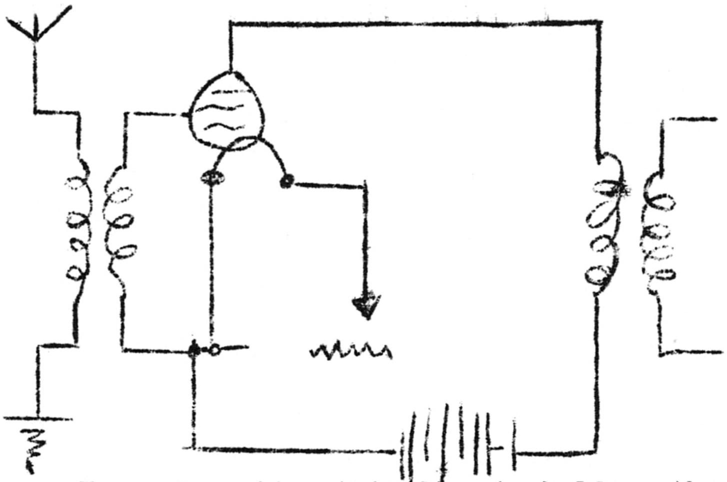
When the meter wire was connected to the positive terminal of the filament, a current flowed through the meter. When connected to the negative terminal no current flowed.

For a very long time this effect was not understood. Physicists working on X-rays and cathode-rays had, however, been gradually developing the idea of the "electron" as a tiny negatively-charged particle, and in 1903, O. W. Richardson recognized the Edison effect as comprising an emission of electrons from the cathode, and showed that the number of electrons leaving the cathode would depend on the cathode temperature. A positive voltage on the plate would attract them, a negative voltage would repel them.

The application of this to radio came through the work of John Ambrose Flem Prof. of Elec. Eng. at Univ. Coll. London. He had done some preliminary experiments in 1889 and 1890 on the Edison effect, and so had Sir Wm Preece of the Post Office, but the latter's conclusions that the current was produced by particles of carbon was incorrect. Fleming was nearer in referring to the discharge from the filament as "negative electricity". Fleming took an interest in radio, but found his increasing deafness a handicap, since the magnetic detector produced only an audible signal. He wanted a visual signal, and thought Kelvin's mirror galvanometer (used on long submarine cables for indicating feeble telegraph signals) would be suitable if he could find a way of rectifying high-frequency currents. So his thoughts turned back to the Edison effect, and in 1904 he connected his diode valve (patented!) in the aerial circuit of a Marconi receiver. It worked. But it was years before it became a reliable device.

The history of communication by electricity is full of patent litigation, and the invention of the thermionic valve led to plenty. It so happened that another clever man, this time an American, named Lee De Forest, also worked on the problem of detection and invented a thermionic valve. He approached it through noticing the effect of electromagnetic waves on a gas flame, deciding a naked flame was unsuitable for use on a ship, and thinking he could get the same sort of effect in an incandescent electric filament lamp. So he had made a carbon filament lamp with a platinum plate, and this, in 1905 or 1906, he used as a detector - but not as a rectifier, for he followed the analogy of the magnetic detector, having a battery circuit and telephone as in the magnetic detector (or even in the coherer circuit) with the d.c. influenced by the incoming wave. De Forest always claimed that he had not been aware of Fleming's work, and this may be true as his approach was quite different. What made De Forest's work important (at this stage it was well behind Fleming's) was that he went on to add a third electrode - at first a piece of foil round the outside of the tube, then a second plate the other side of the filament, and finally a wire grid between the first plate and the filament.

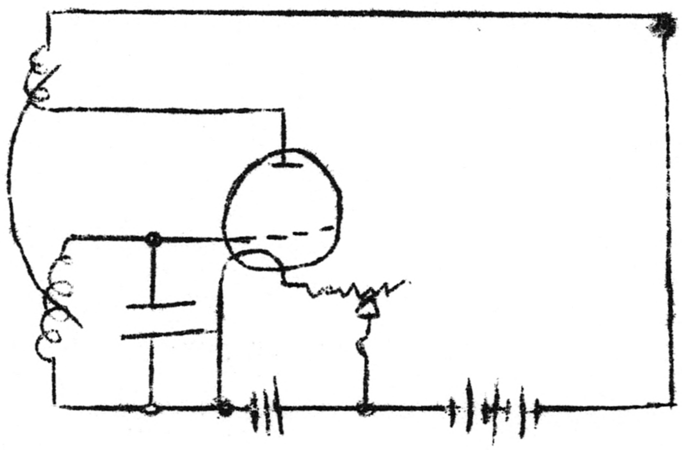
He connected the aerial to this third electrode, and had in effect the triode valve, which he patented in October 1906 as an amplifier and in January 1907 as part of a wireless apparatus. This led to this kind of circuit:



De Forest does not seem to have understood his invention at any stage, but he did recognize its importance, Fleming naturally claimed it was only a technical improvement of his diode, and a bitter patent litigation ensued, lasting for nearly 36 years. For a long time this dispute held-up the use of the valve, as no-one knew who held the patent rights. Also it was a long time before the wide versatility of the valve was recognized. World War I helped a lot in bringing it out.

The understanding of triode operation in terms of characteristic curves was developed by E. H. Armstrong, in 1914, and he thus was able to see how to use the triode as combined detector and amplifier by biassing the grid back to cut-off, so that only positive half-cycles of signal caused anode current to flow.

A most important development in the use of triodes was the discovery of feedback by C. S. Franklin in 1913. He, and others almost simultaneously, were then able to use the triode for generating h.f. ocillations. Franklin's feed-back circuit of 1913 was thus:-



The feedback principle also led to great improvements in receivers. The feedback was positive, and it is interesting that nowadays we always use negative feedback for its many valuable properties, e.g. stability, widening of bandwidth, linearity, etc. But these valuable properties are obtained at the expense of gain. Nowadays, gain costs next-to-nothing to obtain; but in the early days of valves, gain was hard to obtain, and expensive. Thus it was positive feedback which was valuable, for it led to greatly increased gain, and the disadvantages of instability, narrow bandwidth, non-linearity, etc., were only a small price to pay. Thus the regenerative receiver came into use; its invention was claimed by both Armstrong and De Forest in 1913, but Armstrong had the stronger case, and won at first in a series of court battles. Then in 1934 the U.S. Supreme Court reversed the decision and awarded priority to De Forest. This decision was, however, deplored by both the I.R.E. and the Franklin Institute, who stated firmly their conviction that Armstrong was the true inventor! But the money and time spent on the case shows how valuable the invention was. Armstrong later invented (this time unchallenged) the Super-regenerative receiver, in which the circuit was kept just on the verge of self-oscillation; receipt of even very weak radio signals would cause oscillation and thus very sensitive detection was possible.

Armstrong was also concerned in several other important developments, notably the superheterodyne and frequency-modulation.

The heterodyne principle itself was due to Fessenden, who in 1902 proposed it, and had it tested by the U.S.Navy in 1910, in some radio tests in which a range of 3000 miles was attained. He used a h.f. alternator as the local source, near in frequency to the received signal with which it was mixed. Audio beat-notes could be obtained in telephones. It was in 1917 that Armstrong developed the idea to help the Americans keep up with the British and French, who were far ahead in radio technology with valves which could operate up to 1 MHz. Since the Americans had only low-frequency tubes, Armstrong's idea was to change the received frequency to a lower one which could be amplified by the American valves available. It was also found easier to tune this "superheterodyne" receiver as the selectivity was at a fixed intermediate frequency. By 1924 the superhet was the basis of broadcast receivers.

The principle of frequency-modulation was developed by Armstrong as a means of combating "static" or atmospheric noise. The work was financed by the profits of super-regeneration.

Returning to the valve itself, many improvements took place. Langmuir and his colleagues (in U.S.A.) began to understand how valves worked and so could improve design. Following World War I it was discovered that the highest possible vacuum was desirable, that tungsten filaments with an oxide coating were effective, but that if thorium were mixed with the tungsten the emission was so improved that the filament temperature could be reduced from incandescence at 2000° C to a dull heat - so the change from "bright emitters" to "dull emitters" came about. The Germans found that barium and strontium would do as well as thorium. Then the indirectly-heated cathode came in, and ousted the directly-heated cathode in the 1930's.

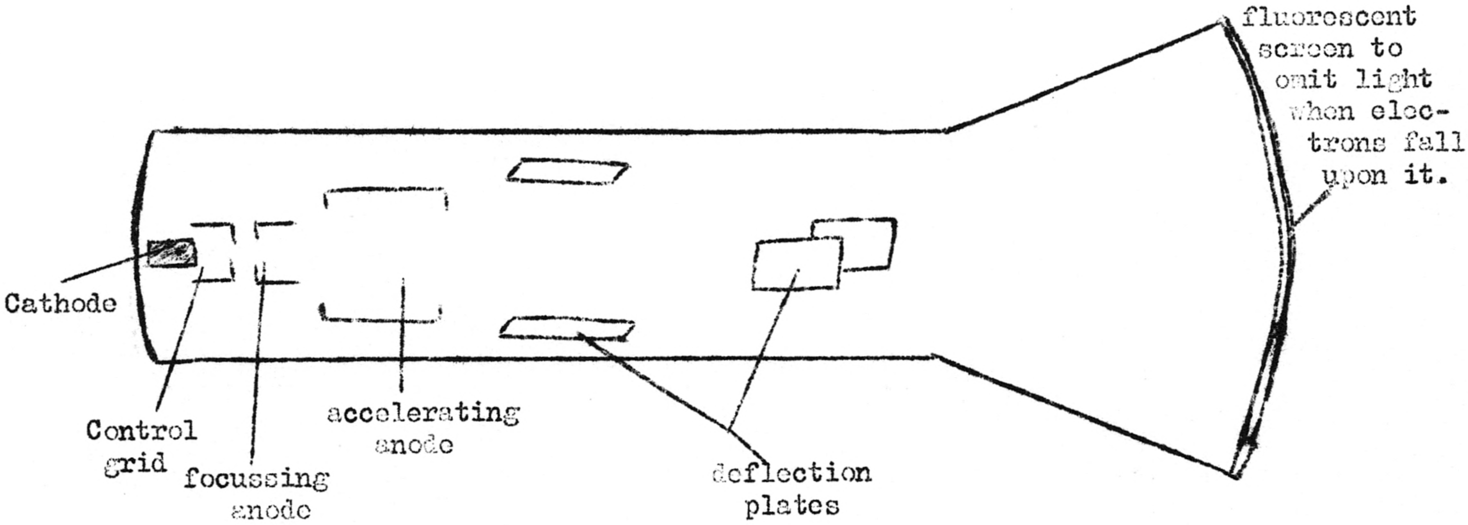
The triode valve suffered from a defect that limited its performance in a number of applications; this was the capacitance between anode and grid which caused a very large effective capacitance to arise in the circuit between grid and cathode - the Miller effect. To reduce the anode-grid capacitance, Scholtkey as early as 1916 tried putting a high-potential screen grid between anode and grid, and this was made practical by the British worker H. J. Round in 1926 in what was called later the Tetrode - at the time called the Screen-grid valve. Apart from its normal uses as an amplifier, this valve had other special properties, e.g. it could easily produce a two-terminal negative resistance which led to the "dynatron oscillator".

Complications with secondary emission of electrons from the anode led to the addition of yet another grid - the suppressor grid, normally at earth potential, between the anode and the screen-grid. This captured the secondary emission and diverted it from the other electrodes. This became the pentode, the most used of all thermionic valves, generally credited to the Dutchman Tellegen and Holst in 1926.

Numerous other kinds of valve have been produced, with built-in diodes and triodes as well as pentodes in the same glass envelope. These extravagantly-electroded valves were mostly British in origin and arose from the need to pay royalties at a fixed rate per valve; thus the number of separate valves had to be minimized.

Other types of valve, working on quite different principles, e.g. the Klyston and the Magnetron, have been introduced for working at very high frequencies such as 1000 to 10,000 MHz, where the transit time of electrons limits the use of conventional valve techniques.

The cathode-ray tube is another type of thermionic valve which has assumed enormous importance since the early 1930's. But it has its origin in the 19th century. By 1897 Braun had made a CRT for experimental purposes. In 1905 Wehnelt achieved the necessary emission for practical purposes by coating a platinum filament with lime. The essential design of the modern CRT was established within a few years, thus:-



Television as a concept preceded the CRT by a very long time. The idea of a mozaic of tiny light-sensitive cells was proposed in 1875 by G. R. Carey (American) and taken up by Ayrton and Perry (British) in 1877 and by Senlecq (French). Each cell was to be connected to a wire, with battery and lamp to reproduce a picture. But of course this was quite impracticable. The concept of scanning was essential to practical television, and the use of revolving discs, vibrating mirrors, etc. was developed, with sequential samples of the picture transmitted electrically over a single circuit. It was evident in 1928 that the principles of television were well understood, and awaited only suitable electronic systems to enable them to be exploited. Picture telegraphy was in commercial use by the early

1930's, and the development of the Emitron television camera in England and the Iconoscope in America (by Zworykin of RCA), together with the now very efficient CRT, enabled television to be introduced as a public service in the late 1930's.

References.

1. P. Dunsheath, "A History of Electrical Engineering", Faber, London, 1962.

2. H. I. Sharlin, "The Making of the Electrical Age", Abeland-Schumann, London and New York, 1963.