This article reviews the development of electronic devices used over the last century in wireless communication. It looks at early receiving devices such as the coherer, the magnetic detector and the cat's whisker, progressing to the thermionic valve, the semiconductor, the microchip and digital signal processing.

On the transmission side, the early devices discussed include the spark-gap generator, the voltaic-arc generator and static frequency multipliers. This is followed by a brief description of more modern power devices, including thermionic valves and electron-velocity control tubes.

From the coherer to DSP

M. Lemme and R. Menicucci (Vatican Radio)

This article reviews the development of electronic devices used over the last century in wireless communication. It looks at early receiving devices such as the coherer, the magnetic detector and the cat's whisker, progressing to the thermionic valve, the semiconductor, the microchip and digital signal processing.

On the transmission side, the early devices discussed include the spark-gap generator, the voltaic-arc generator and static frequency multipliers. This is followed by a brief description of more modern power devices, including thermionic valves and electron-velocity control tubes.

1. Introduction

The electromagnetic-wave generator that was first conceived by Hertz, in 1887, consisted of a copper rod 3 m long which had a 30 cm zinc sphere attached to either end. The rod was divided in the middle to include a spark gap of 0.75 cm between two small brass spheres. This assembly – which we now call a *Hertz dipole* – was connected across the secondary winding of an induction coil. When the primary coil was connected to a battery, a spark would appear across the gap, generating an oscillating current. As a consequence, damped electrical waves would propagate across the room.

The wavelength produced by the Hertz resonator depended on the inductance and the capacitance of the circuit (the large zinc spheres could be moved along the rod to vary the wavelength of the oscillations). Hertz was aiming to reproduce phenomena similar to light and the shortest wavelength he managed to generate was about 30 cm, equivalent to an oscillation frequency of 1000 megacycles/second (1000 MHz in modern parlance).

The energy radiated by the dipole was called "Force Diffusion" by Hertz; it was Lord Kelvin who later gave it the name *electromagnetic wave*. Sir William Crookes was very surprised when he noticed that electromagnetic waves could pass through *"the walls and the fog of London"*. In 1892, he gave an explanation of this phenomenon and was even able to foretell wireless telegraphy.

2. Early receiving devices

2.1. Hertz resonator

The first and simplest detector of electromagnetic waves was the *Hertz resonator.* It consisted of an open metal ring with a small metal sphere attached at either end. A spark jumped across the small gap between the two spheres, whenever a spark was generated in the transmitting apparatus.

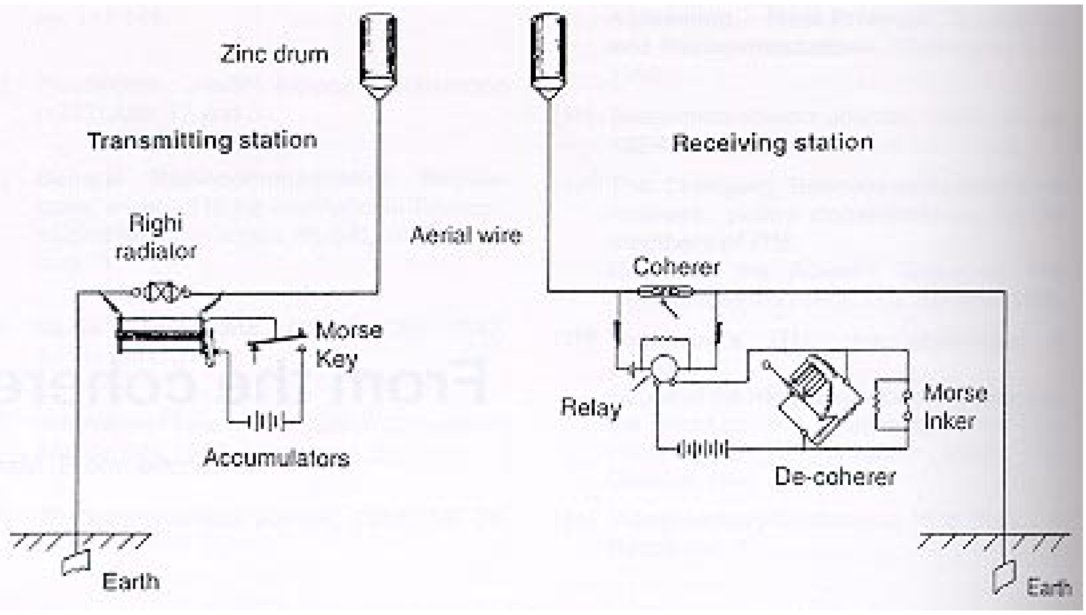


Figure 1
Circuit diagram of the transmit-receive system used during Marconi's Lavernock Point trials in May 1897
(as published by the General Post Office in the Engineer-in-Chief's report of 1903).

2.2. Coherer

An improved detector, the *coherer*, was invented by the Frenchman, Edouard Branly. It was based on the discoveries of the Anglo-American, D.E. Hughes, and the Italian, Calzecchi Onesti – which had been further updated by the Englishman, Sir Oliver Lodge, as well as the Russian, A.S. Popov.

The coherer consisted of a tube filled with fine metal filings (Marconi used 95% nickel and 5% silver). When RF energy was passed through the filings, the particles cohered and the resistance dropped. The coherer could be made responsive to a change of state, i.e. to the presence or absence of a signal. When a signal was present, a battery-powered circuit attracted the moving element of a relay and, consequently, the whole assembly could be used to receive Morse code signals from a remote transmitter. In order to separate the particles inside the bulb, so that the apparatus could be operated again, some form of "de-coherer" such as a solenoid tapper or a mechanical shaker was necessary.

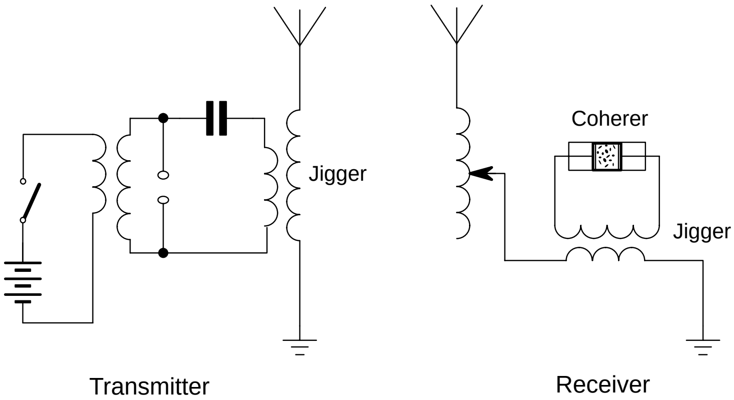


Figure 2
Marconi's system for long-range radio reception, as fitted to the American liner *St. Paul* in 1899 (Patent No. 7777).

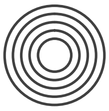
A coherer with an antenna and an earth connection formed the receiving part of the equipment used by Marconi in 1896/7 to carry out his experiments at Pontecchio near Bologna, Italy, and at Lavernock Point near Cardiff, Wales. A circuit diagram of the latter installation is shown in *Fig. 1*.

The coherer was also used in the system Marconi developed to communicate with ships at the turn of the century. In this system, according to his patent No. 7777 *(Fig. 2)*, Marconi connected the transmitting antenna to the secondary winding (jigger) of a "Tesla transformer" (formerly the antenna had been connected to ground via the metal spheres of the exciter). The primary winding of this transformer was connected in series with the exciter and a capacitor, formed by Leyden jars. At the receiving station, the antenna was connected to ground through the primary winding of a Tesla transformer, whose jigger was connected to the coherer. The first apparatus of this type was installed by Marconi in the American liner *St. Paul* at the end of 1899.

Although the coherer was used successfully for several years, it soon showed its poor efficiency and sensitivity. In December 1901, at St. John's in Newfoundland, Canada, Marconi replaced it with a mercury drop detector.

2.3. Magnetic detector

The early detectors were rather unstable and irregular in operation. These problems, along



with some polemics, induced Marconi to aim his research at a new type of detector. Inspired by a paper published in 1896 by Prof. Rutherford, Marconi developed the *magnetic detector* which was a significant improvement over the coherer. The magnetic detector employed clockwork mechanical power to magnify the weak incoming pulses from a spark transmitter so that they were strong enough to drive a pair of headphones.

Basically, the magnetic detector operated as follows *(Fig. 3)*. A loop of hard-drawn iron wires *(a)* , wound together in the form of a “rope”, was made to move endlessly through two coils, by means of a clockwork drive. The inner coil *(b)* – a single-layer winding – was connected between the aerial and the earth via a tuning filter. The outer coil *(c)* was made from a large number of turns of fine wire and was connected directly to the headphones *(T)*. Two permanent magnets *(d)*, with their “like-poles” together, produced a quadropole field that caused the moving iron rope to divide into two magnetic domains separated by a “wall” which was stationary relative to the two coils.

Whenever a high-frequency “pulse” was received from the transmitter, the resultant magnetic field acted – via the primary coil – on the wire rope. Incoming half-waves of one polarity would cause the magnetic domain wall to “flick” backwards and produce –via the secondary coil – a “click” on the headphones, while incoming half-waves of the other polarity had little effect on the wall. Thus, the device could reproduce the dots and dashes sent out from a spark transmitter, just like a crystal detector – except that the magnetic detector also offered considerable gain.

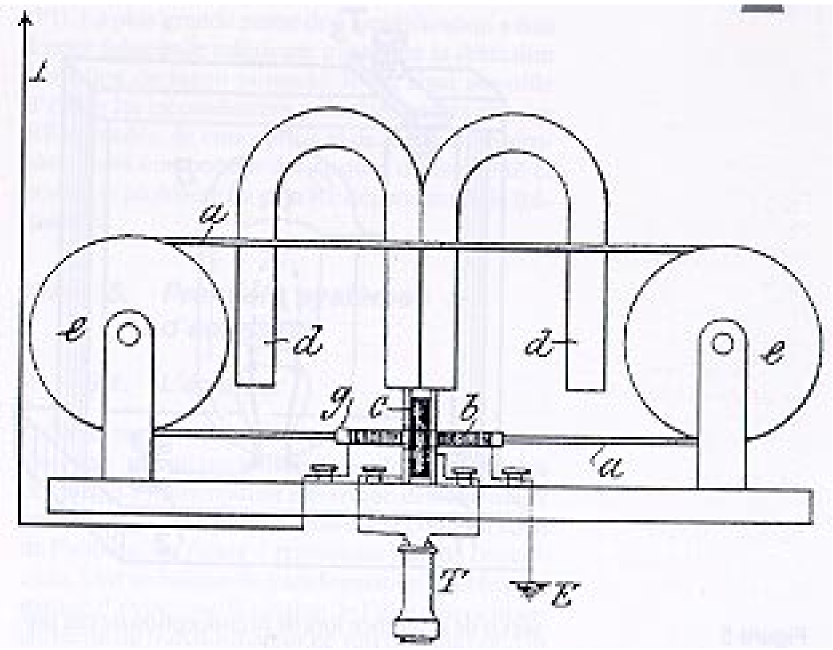


Figure 3
Diagram from Marconi’s Patent No. 10245 of 1902 showing the magnetic detector.

Marconi’s magnetic detector was patented in 1902 and was widely used for about 20 years.

Looking at the pictures and circuit diagrams of the time, an observer with today’s technical knowledge could be excused for thinking that such devices were rather simple and rudimentary. However, reading the journals of those earlier experimenters, one can get a feeling of the problems they had to overcome. Every circuit component caused problems, the solution of which needed months of work and repeated

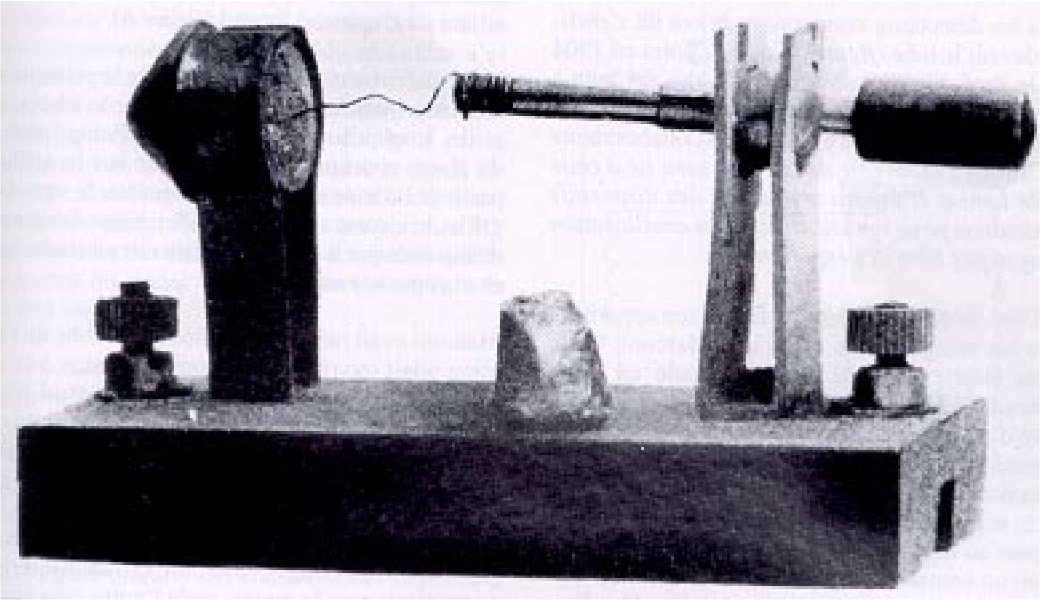


Figure 4
A very early receiver using a galena crystal and cat’s whisker as the detector.

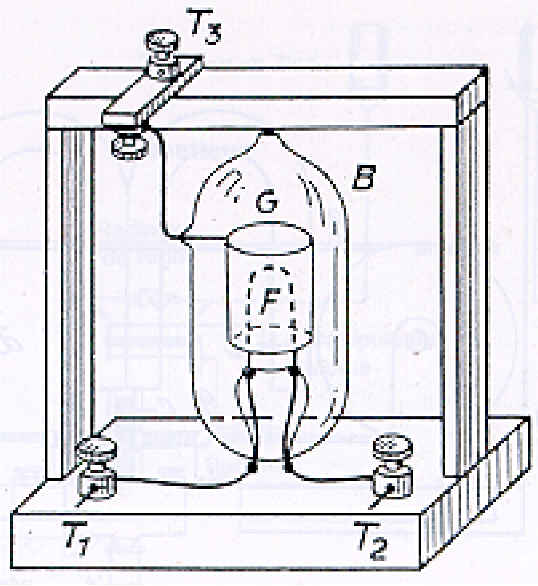


Figure 5
An early thermionic diode, circa 1904.

experiments. Only rough materials were used: for example, the prototype of the magnetic detector was enclosed in a Havana cigar box. There were no suitable instruments to measure RF currents, nor was there any well-grounded theory to aid the design of these early wireless installations.

2.4. Cat's whisker

Another simple detector around this time was the famous *cat's whisker (Fig. 4)*. It consisted of a fine wire electrode whose pointed tip was pressed against the crystal of a detector (very often a crystal of galena). When connected to a suitable tuned circuit, an antenna, an earth point and a headphone, this simple circuit formed a very useful receiver for use with local stations.

3. Thermionic valves

3.1. Diodes

All the above detectors eventually gave way to the thermionic valve *(Fig. 5)* which was conceived by Prof. Fleming in 1904. Marconi soon realized the practical importance of this invention and told one of his collaborators at Clifden: *"the future of the radio will be based upon this small lamp. Its use will require the help of special devices. Many experiments will be needed to improve it"*.

In 1905, Fleming's diode was installed in the receivers of the Marconi company. The diode consisted of a glass bulb containing a carbon filament (the *cathode*) surrounded by a cylinder of thin metal (the *anode* or plate). When the filament was heated until white hot, the emitted electrons reached the anode plate. If the latter was at a positive potential with respect to the filament, it attracted the electrons so that a current flowed in the filament-plate circuit. If the plate became negative, no current flowed through the valve. Thus, by using a thermionic diode, very small HF currents could be rectified into unidirectional currents.

3.2. Triodes

In 1906, Lee de Forest experimented with Fleming's valve and ingeniously conceived a third electrode, the *control grid*, in order to control the flow of electrons. When the grid was negative with respect to the filament, the electrons were repelled so that the anode current was reduced. When the potential of the grid changed to positive with respect to the filament, the anode current increased so as to produce an amplifying effect. At first, the triode valve proved to be a sensitive detector; it was not until 1912 that Lee de Forest appreciated and exploited its possibilities as an amplifier.

In the first triode receivers, the radiofrequency signal was taken direct from the antenna to the valve input via a Tesla transformer with a tuning capacitor, thus providing a tuned-resonance circuit. Rectification was achieved by means of either a single crystal or diode, if only half-wave rectification was required, or a pair of crystals or diodes if full-wave rectification was required.

Around this time, it had also been "discovered" that the triode valve could be used to generate continuous radiofrequency oscillations and, moreover, these oscillations could be modulated at audio frequencies. The stage was now set for rapid developments in the field of *radiotelephony* and Marconi was just one of many to exploit this new communications medium. In 1914, he used valve oscillators to carry out regular radio-telephony transmissions for the Italian Navy. Marconi was able to achieve distances of around 110 km by using devices based upon the oscillating circuit of Captain Round *(Fig. 6)*.

A triode oscillator works on the principle of feed-back from the anode circuit to the grid circuit. When an oscillating voltage of suitable amplitude and phase is fed back from the anode plate to the grid so as to increase the grid signal, the circuit starts to oscillate with a frequency determined by the circuit constants (inductance and capacitance) of the anode and the grid.

Marconi was right when he foretold that Fleming's valve would open up new horizons for radiotelephony. However, there were still many problems to be solved, first of all concerning the vacuum in the valve bulbs. A small amount of gas in the glass bulb of de Forest's *Audion* seemed at first to be advantageous for it to function. However, the operation of the de Forest and Fleming valves was irregular due to the gas within them. On the one hand, the amount of gas increased with time; on the other hand, a variation of the bias between the anode plate and the filament under such conditions was sufficient to cause the breakdown of the valve, due to the increased bombardment of positive ions onto the filament.

Langmuir, an American, discovered that vacuum reduction in the bulb was caused by gas atoms contained in the molecules of the filament and in the metal and glass parts of the valve. He constructed new pumps to create a hard vacuum in the bulb. He also studied ways of expelling the gas atoms from the metal and glass parts of the valve by the so-called "clean-up effect", and he devised a method to seal the filaments inside the glass bulb at a very high temperature. Lee de Forest was the first to introduce metal filaments in place of the carbon filaments used previously. Further experiments led to the metal filaments being coated with alkaline oxides which improved the efficiency of the thermionic triode even more.

3.3. Tetrodes and pentodes

The plate-grid capacity – the cause of many problems for technicians – was reduced to almost one hundredth that of a triode by providing a second grid (the so-called *screen grid*) to act as a screen and prevent the control grid from being influenced by the anode voltage changes. The screen grid reduced the anode-grid feedback and, hence, it prevented instability.

Later, the pentode (containing a third grid called the *suppression grid*) and other types of valves were developed, leading to further progress in the design of receivers and transmitters. In addition to the amplification and detection abilities of these devices, they could also be used to generate controlled oscillations within the receiver. Thus, for example, these more sophisticated valves allowed Prof. Fessenden to construct the first *heterodyne* receiver, which achieved signal detection by beating the incoming RF signal against one produced by a local oscillator in the receiver. In this way, an unmodulated RF carrier was made audible, as the beat note was at an audible frequency.

4. Superheterodyne receivers

Thereafter, the *superheterodyne* (superhet) receiver was developed. It had several important advantages over earlier types of receivers. In the superhet circuit, the incoming RF signal was changed to a lower frequency known as the intermediate frequency (IF). The major part of the amplification then took place at this lower frequency before detection occurred in the normal manner. In this way, it was possible to avoid the drawbacks of tuned radiofrequency (TRF) circuits, which were simple to design and construct but which presented the risk of instability and had a problem of the RF gain being dependent upon frequency.

5. Early transmitting devices

5.1. The spark gap

Let us turn now to the first transmitters. Following on from the early experiments of Marconi and others, a lot of improvements were achieved in both the circuits and the power supplies of transmitters, due to the cooperation of Fleming. One such circuit is shown in *Fig. 7*. In order to avoid too high a voltage at the secondary of the transformer, a dual transformation technique was used. Here, a capacitor was first charged with a voltage of 20 kV, then its oscillatory discharge was

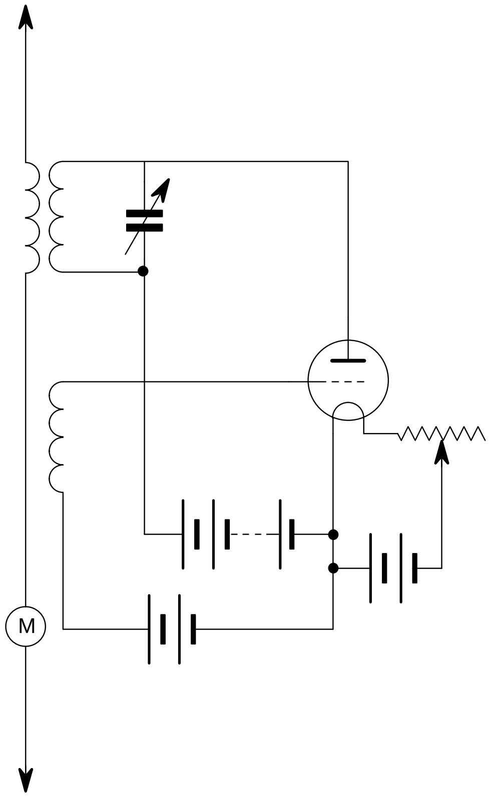


Figure 6
The oscillating circuit developed by Captain H.J. Round.

supplied through a Tesla transformer to a second oscillating circuit provided with a spark gap. This circuit was connected to the antenna via the jigger (secondary winding) of a second Tesla transformer.

The spark gap consisted of two aligned steel disks, almost tangential to each other with a narrow spark gap between their edges. The lower portion of each disk was dipped in its own separate vessel filled with mercury, to avoid the overheating of the disks during electrical discharges. It was an installation such as this which succeeded in repeatedly transmitting the famous Morse code signal for the letter "S" (dot-dot-dot) from Poldhu in Cornwall, England, to St. John's in Newfoundland, Canada, in December 1901.

This momentous event gave new impetus to the researchers. The simple spark gap needed to be improved. After arcing, it did not lose its conductivity immediately, because of ionization effects, and this prevented the capacitor from recharging quickly, resulting in a greater energy consumption and the generation of fewer wave-trains. The simple spark gap was duly replaced by the quenched spark gap and, afterwards, by the rotary discharger (Marconi, 1905) in order to cause the electrode to quench and the sparks to be interrupted more quickly.

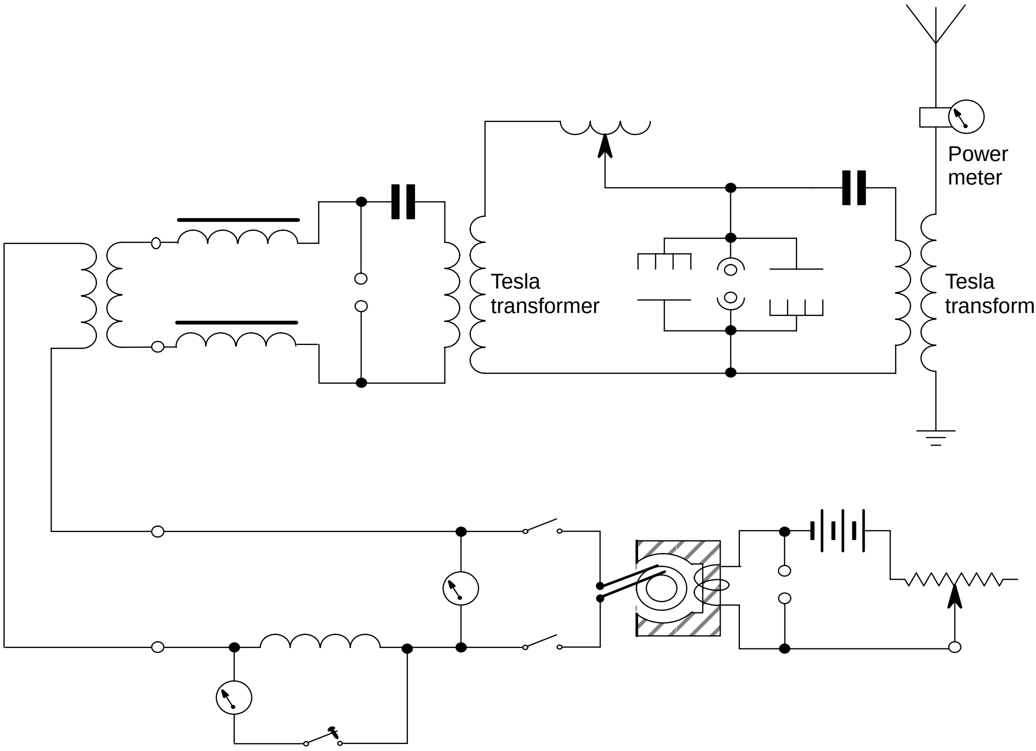
However, there was still the problem of damped waves and low efficiency, which allowed only the transmission of telegraphy signals; intermittent operation was not adaptable to telephony modulation.

5.2. Voltaic-arc generators

There was a need to go back to basic theory. In 1850, Sir W. Thompson had defined the relationship between the resistance, inductance and capacitance that was needed to provide an oscillatory discharge from a circuit. Based on those ideas, generators of undamped waves were eventually developed, along with voltaic-arc generators which used carbon electrodes.

The operating principle of the voltaic-arc generator *(Fig. 8)* is based upon the negative resistance of the arc. A series capacitor and inductor, connected in parallel with an arc, are charged by a DC power supply. This causes the voltage across the arc and the LC circuit to increase. However, the capacitor is charged to a greater voltage than that of the arc, due to the effect of the inductor. When the capacitor starts to discharge, the current across the arc increases, thus causing the voltage to decrease. When the voltage of the capacitor equals that of the arc, the inductor maintains the discharge current, causing the voltage across the capacitor to become lower than that of the arc. At the end of the discharge cycle, the capacitor is charged again and the process is repeated.

Figure 7
Circuit diagram of a Marconi transmitter around the time of the Poldhu - St. John's trials (1901).



Poulsen then experimented with arcs dipped into a large tank of hydrogen. This type of generator produced quasi-continuous waves which could be amplitude modulated for telephony applications. The system offered the great advantage of simplicity and economy, although it generated harmonics and consequently, RF interference. In spite of its drawbacks, many researchers deemed that the Poulsen arc was the successor of the spark-gap system and devoted themselves to improve it even further. Prof. Majorana – who experimented with multiple-spark generators for radiotelephony applications in 1903-1904 – used the Poulsen arc to transmit telephony signals over a range of about 400 km (the maximum distance that could be reached by radiotelephony at the time).

5.3. *Static frequency multipliers*

Special alternators were developed to generate continuous RF waves, but they had many mechanical problems due to the very high frequencies that had to be reached. One method of overcoming such problems was to use a "static frequency multiplier" – a magnetic-core device, similar to a peaking transformer, which provided harmonics by distorting a sinewave. The multiplier was supplied directly by the alternator and operated as an oscillation exciter in a circuit which was tuned to the desired frequency and coupled to the antenna.

This type of system was limited to longwave applications. The cost of the installation was prohibitively high because, in addition to the alternators, several 200 m high towers were required to support the horizontal longwave radiators. This prompted the researchers to develop a simpler, more economic, generator of RF oscillations.

6. *Thermionic valve generators*

It was around this time that the first thermionic-valve generators appeared. A lot of time was devoted to their development, before the advantages of their application could be enjoyed. At that point, there were many clashes between the proponents of valve generators and the supporters of systems which used RF alternators.

In 1915, considerable progress was obtained by the stations at Arlington and Honolulu. An oscillatory valve was coupled to the grid circuit of a valve of greater power. The plate current of the latter drove in turn the grid circuits of a further 10 valves in parallel. An array of a further 300 valves with their anodes in parallel was driven by the above circuits and supplied its output energy to the high antenna of the Arlington station.

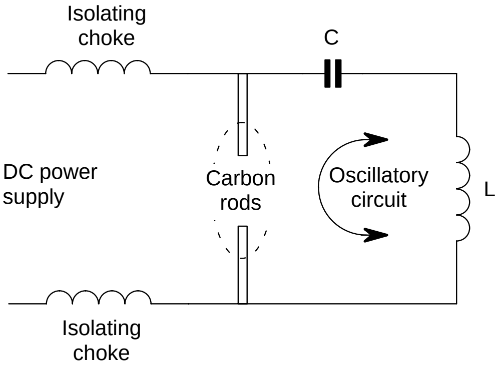


Figure 8
Circuit diagram of a voltaic-arc generator.

Marconi had already experimented with valve transmitters for the navy in 1914. With his own experience and with that of Arlington and Honolulu, he became convinced that a transmitter which used alternators, a voltaic-arc generator or rotary discharger could not provide a commercial service more convenient than the wired service. Therefore, he decided to develop systems that would require new valve designs which were capable of delivering power levels of greater than 1500 W.

In order to achieve the required power levels, Marconi connected 48 valves in parallel and increased the anode voltage to 10 kV. The resultant current in the aerial was then 330 A. A new problem to be overcome was that of keeping the frequency constant.

In 1921, Franklin devised the "Master Oscillator" consisting of an oscillating valve circuit which generated signals of constant frequency, independent of temperature changes. It was connected to a number of amplifier stages in order to reach the desired power for supplying the aerial. Thereafter, as it was easier to stabilize an oscillator that operated at a low frequency, the "drive oscillator" was arranged to generate a frequency which was a submultiple of the frequency of the transmitter. The drive oscillator was followed by several frequency doublers and triplers which also operated as amplifiers. Very stable oscillators with "quartz crystals" (based upon piezo-electrical phenomena) were later developed: a suitably-cut quartz plate is a resonator of excellent quality, especially if it is kept at a constant temperature.

In addition to the problem of producing high-power radiofrequency energy, solved by the use of valves, a new problem arose: which frequency was the most suitable for a given transmission?

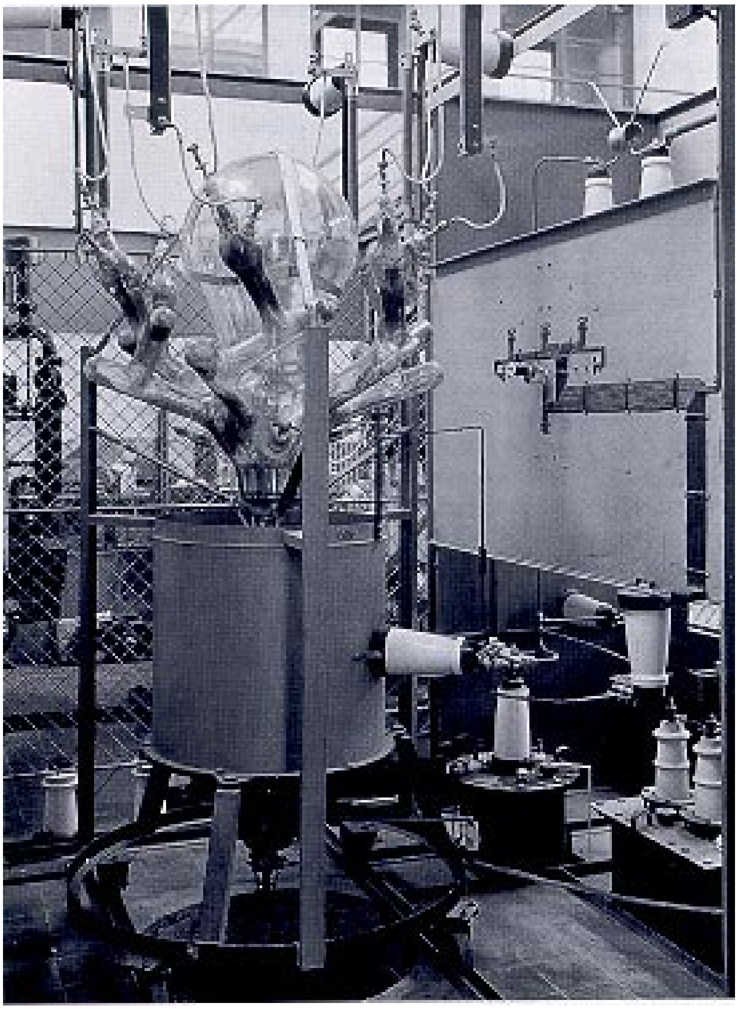


Figure 9
A high-voltage mercury-arc rectifier from 1933.

The first works of Marconi were based upon the use of medium and long waves. Apparatus of high power and large antennas supported by high towers were needed, which required considerable investment and resulted in a very expensive radio service. The theory of Loewenstein (concerning the composition of the atmosphere and the influence of its layers upon the propagation of the electrical waves), along with experiences made in the short-wave field, prompted Marconi to confide to one collaborator: *“Only the short waves will be able to save the radio from the blind alley of the medium and long waves”*.

In 1923, Marconi carried out important experiments with Franklin between the ship “Elettra” at anchor by St. Vincent in the Cape Verde Islands, (off the coast of West Africa) and the station at Poldhu in Cornwall, England, where a short-wave transmitter (25 – 90 metres) had also been installed. At a distance of 4130 km the received signals were strong and clear, even if the transmission power was reduced to only 1 kW. Under such conditions, the signals were stronger than those transmitted from Caernarvon (Wales) and Nauen (Germany) on a wavelength several hundred times greater, and with a power of 200 kW.

Thereafter, studies were made to improve the modulation percentage (i.e. modulation index) and to reduce the distortion. One such method was *grid modulation* which was easier to produce but which had low fidelity. Heising introduced *anode modulation* where the signal, after suitable amplification, was superimposed on the anode voltage of the final radiofrequency valve. In contrast to the other methods, a modulated wave with a high modulation percentage and a low distortion was obtained.

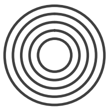
Finally they had to solve the problem of how to cool the valves which operated at power levels of above 2500 W. Around 1930, the anode was cooled by oil circulation. Later, after the solution of anode-insulation problems, distilled water circulation was used to keep the anode cool. Thus, the high power required in the final stages of the transmitter, especially for long waves, could now be obtained by a lesser number of valves operating in parallel. After a lot of improvements, the tetrodes of today are capable of delivering several hundreds of kilowatts.

7. Modern transmitting devices

Increases in the power needed for AM transmitters, accompanied by increases in the cost of energy supply, has prompted the development of high-efficiency RF power amplifiers and modulation systems (Pulse Duration Modulation, Pulse Step Modulation, etc). More recently, completely-digital AM transmitters with very high efficiency have been developed.

Over a few decades, the increased application of radio-telecommunication systems, and the growing need for more radiofrequency spectrum to accommodate these systems, has led to the use of ever higher frequencies: hundreds of MHz, tens of GHz ... up to the frequencies of light. New devices, based on the control of the electron beam current, have taken the place of the old electronic tubes.

The maximum operating frequency of triodes, tetrodes and pentodes is limited by the reactances and resistances associated with the electrodes. These effects, for instance, reduce the input and



output impedances, and the efficiency as well. Additionally, when the electron transit time is comparable with a quarter of the period of the microwave, the total efficiency of the valve, when operating either as an oscillator or an amplifier, is reduced to less than half. As a result, it is difficult to produce triode or tetrode amplifiers for use above 1 GHz which have an output power of several kilowatts and an acceptable efficiency for continuous operation.

7.1. Electron-velocity control tubes

Tubes which operate by modulating the velocity of an electron beam have been developed for microwave use. Such tubes are mainly of two types:

- *klystrons and travelling wave tubes,* where the electron beam flows linearly;
- *magnetrons,* where the electron beam follows a curved path under the action of orthogonal electric and magnetic fields.

Klystrons and travelling wave tubes find a wide application in broadcasting and microwave telecommunication systems. Both types of device use an axial magnetic field to "collimate" the electron beam, i.e. to prevent the beam from spreading out as a result of small initial divergent angles, and the repulsion effect of the negative charges which make up the beam. The speed of each moving electron is controlled by an electric field and is calculated from the following equation for kinetic energy:

$$Ve = \frac{1}{2}mv^2$$

where: V = the accelerating voltage acting on the electron
e = the charge on the electron (-1.6×10^{-19} C)
v = the velocity of the electron
m = the mass of the electron (9.1×10^{-31} kg).

From this:

$$v = \sqrt{\frac{2Ve}{m}10^7} \quad \text{(cm/sec)}$$

And hence:

$$v = 5.93 \times 10^7 \sqrt{V} \quad \text{(cm/sec)}.$$

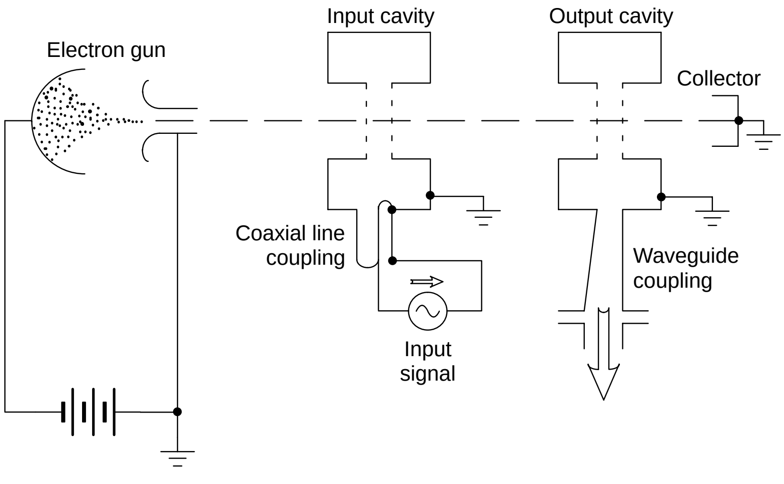
7.1.1. Klystron

The Klystron was invented in 1937 and basically operates as follows *(Fig. 10)*. Electrons are fired from an electron gun which contains a cathode heated by a filament, and a voltage source which produces acceleration of the beam. The microwave input signal at the first cavity creates voltages across the gap traversed by the beam, causing the electrons in the beam to accelerate or decelerate. As it travels towards the collector, the electron beam is thus "velocity modulated" by the input signal; groups of electrons bunch together to form "electron packets". When these electron packets arrive at the output cavity, they generate a microwave resonance at an increased power level relative to the input signal, due to the energy supplied by the electron beam.

Additional "intermediate" cavities are provided on some klystrons to improve the bunching process, resulting in a higher gain and improved efficiency. By varying the size of a cavity, it is possible to adjust its resonant frequency. Thus in a multi-cavity device, if all the cavities are tuned to the same frequency, the gain is maximized but the bandwidth is reduced. On the other hand, if the cavities are tuned to slightly different frequencies, the gain is reduced but the bandwidth is increased.

Klystrons with two or more cavities are used as narrow-band linear power amplifiers. They can generate about 10 kW under continuous operating conditions, at a frequency of several GHz. The maximum gain achieved by a klystron is some tens of decibels and its efficiency ranges from about 30 % to 50 %.

Figure 10
Schematic diagram of a two-cavity klystron amplifier.



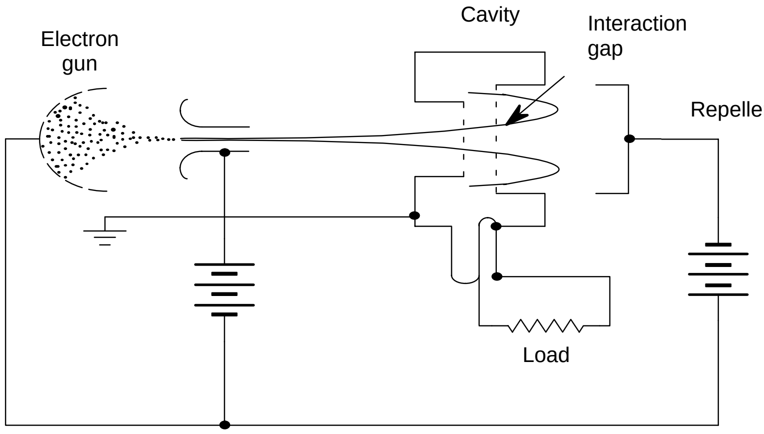


Figure 11
Schematic diagram of a reflex klystron amplifier.

Another version of the device, known as a *reflex klystron*, is used as an RF oscillator *(Fig. 11)*. In this design, the output cavity has been eliminated and an electrode (repeller), at a suitable negative potential, is used to repel electron packets back to the two grids of the sole cavity.

7.1.2. Travelling wave tubes

In a travelling wave tube (TWT) *(Fig. 12)*, the electrons generated by the heated cathode travel along the axis of the tube, constrained by focussing coils, until they reach the collector. Spaced closely around the electron beam is a helix which is capable of propagating a slow-moving electromagnetic wave.

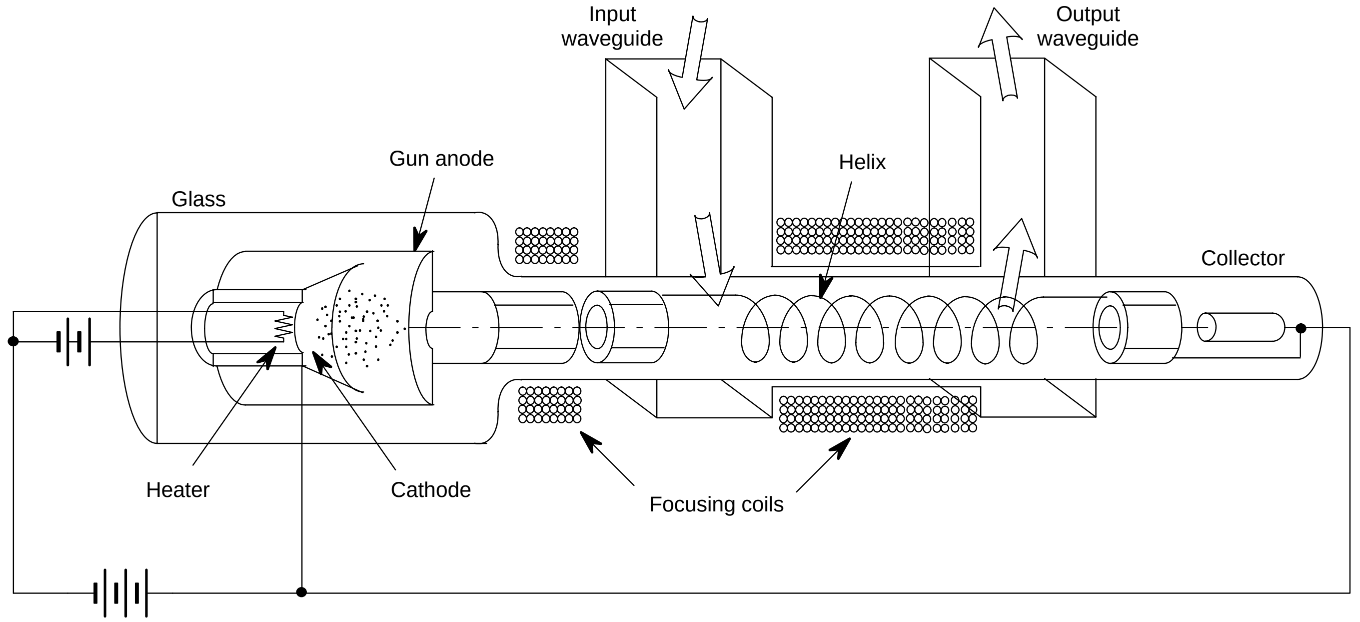
The RF signal enters the tube via the input waveguide and travels along the helix *wire* at velocities close to the speed of light. However, along the *axis* of the helix, the velocity of propagation (the so-called *phase velocity*) is much lower, typically 0.1 to 0.3 times the speed of light; it is equal to the product of the velocity of light and the ratio between the pitch and circumference of the helix.

The electrons generated by the TWT are accelerated by a high voltage such that, by the time they reach the helix, their velocity is similar to that of the axial phase velocity of the RF input signal along the helix. When the RF signal enters the helix, the longitudinal part of its field interacts with the electron beam, causing some electrons to accelerate and others to decelerate such that electron packets are formed in a similar manner to that of the klystron. The result is a progressive rearrangement in phase of the electrons in relation to the RF wave. In turn, the modulated electron beam induces additional waves on the helix and this process of mutual interaction continues along the length of the tube. The outcome is that the DC energy of the electron beam is transformed into RF energy in the helix, and the RF wave is amplified.

In order to avoid self-oscillations, the amplified RF wave should not be reflected back to the input of the helix; any reflected portion of the signal must be attenuated along the helix. To achieve this, the wire of the helix is made of, or coated with, a magnetic material. As a result, the potential gain of a TWT is reduced but a good compromise between gain and stability can usually be reached.

Unlike the klystron, a travelling wave tube has no specific resonating elements. It can be fabricated

Figure 12
Schematic diagram of a travelling wave tube.



with a pass-band of 600 – 800 MHz and can yield a gain of some tens of decibels. The maximum power levels produced by a TWT are of the order of several hundred watts at 10 GHz, and its efficiency is around 30 %.

8. Transistors and other semiconductors

Although we can refer to the cat's whisker crystal as the earliest ancestor of modern semiconductor devices, reference should also be made to the silicon point-contact diode, which was developed during the Second World War for use in microwave radar receivers, and the germanium junction diode which was developed shortly after.

However, the new electronic age began in 1948 when W. Shockley, J. Berdeen and W.H. Brattain of Bell Telephone Laboratories announced the first transistor (a name abridged from "transfer resistor"). This important electronic device had a strong impact on the field of electronics, in both engineering and economic terms. Transistors and, in general, semiconductors are based upon the controlled presence of imperfections in otherwise nearly-perfect crystals. The materials generally used are germanium and silicon.

8.1. Semiconductor physics

The germanium or silicon atom may be considered to consist of four valence electrons surrounding a nucleus which has a positive charge of four electron units. The atom as a whole is thus neutral. This symmetry is broken if we introduce some chemical impurities, the so-called *donors* or *acceptors*, into silicon and germanium crystals. The presence of these classes of impurities causes an excess of electrons or holes, thus making the crystal a conductor or rather a semiconductor. Typical donor elements for silicon and germanium are antimony and arsenic which each have five valence electrons. Consequently, one electron for every atom of impurity becomes an excess electron and wanders throughout the crystal (in the n-type semiconductor).

Aluminum and gallium are typical acceptors for silicon and germanium. These elements have three valence electrons rather than four. As a result, they cannot complete the paired electron structure of the four bonds surrounding them, when they replace an atom in the crystal lattice. If an electron from a bond somewhere else is used to fill in this hole, the acceptor atom will acquire a localized negative charge and a new hole is created in the atom which had given up its spare electron. The crystal becomes a p-type semiconductor because a positive particle wanders throughout the crystal lattice.

It is evident that in a single crystal there may be both p-type and n-type regions. The boundary between such regions is called a *p-n junction*. Such p-n junctions have interesting electrical and optical properties. Due to the different charge on the particles which wander throughout the two regions (electrons and holes), an electric field is established across the junction and no current will flow across the junction under conditions of thermal equilibrium.

The application of a negative voltage to the p-region and a positive voltage to the n-region (reverse bias) will increase the electrostatic potential difference between the two regions, so that no current due to the majority carriers flows into the device. If the reverse voltage increases to the breakdown value (Zener voltage), the current flowing into the device will rapidly increase to the point of reverse saturation. This is due to the minority carriers which are present in the crystal as a result of breaking the covalent bonds. When forward bias is applied, the current increases exponentially to the point of forward saturation.

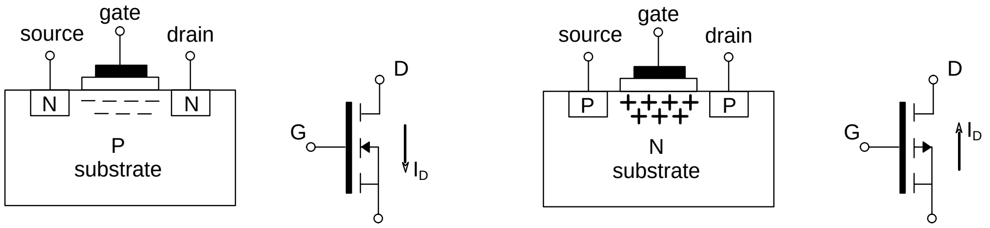
8.2. Junction transistor

The junction transistor consists of a single crystal of germanium or silicon divided into three regions: these regions are alternately n-type, p-type and n-type and are called the *emitter*, *base* and *collector* respectively. When an n-p-n transistor is used as an amplifier, the positive bias is applied to the collector region in order to bias the p-n junction between the collector and the base regions in the reverse direction. If a positive signal is applied between the emitter and the base, a large number of electrons will be able to diffuse into the base.

Because of the thickness of the base region, we can consider that all the electron current entering the base region from the emitter arrives at the collector. Only a negligible fraction of the emitter current flows in the base circuit. All changes in the emitter current appear in the collector current and, thus, the junction transistor behaves much like an ideal triode but requires no cathode power and offers the hope of practically unlimited life.

8.3. Other transistor types

In addition to the junction or *bipolar* transistor (so called because the conduction involves electrical charges of both signs), *unipolar* tran-



Typical N-MOS

Typical P-MOS

Figure 13
The basic structure of a MOSFET.

sistors have also been developed. In these devices, conduction is due to electrical charges of only one sign; the *junction field effect transistor* (JFET) and *metal oxide semiconductor field effect transistor* (MOSFET) are perhaps the best-known examples.

JFETs are active devices with three terminals (*gate*, *source* and *drain*) which come in two versions: p-channel and n-channel. Both versions can operate in either a linear or a non-linear mode. Between the channel and the gate there is a p-n junction, the polarization of which controls the current flowing between the source and the drain. In order to operate in the linear mode, the gate-channel junction should be reverse-biased. The output and transfer characteristics of a JFET are very similar to those of a triode.

The MOSFET transistor *(Fig. 13)* consists of a p-type or an n-type substrate. The gate is insulated from the substrate by silicon oxide, while the source and drain are reverse-doped regions with respect to the substrate. The source and substrate are innerly connected so that they are at the same potential. The channel is formed by the free charges in the source region which are attracted to the gate by its polarization voltage.

FET transistors are widely deployed in the integrated amplifiers used for microwave applications up to frequencies of around 14 GHz. The so-called *monolithic microwave integrated circuit* (MMIC) is fabricated by using gallium arsenide technology.

For higher frequency applications, *parametric amplifiers* with negative resistance are used. A typical negative-resistance element is the gallium arsenide varactor.

Further important semiconductor devices are the *luminescent diodes* (LED: light emitting diode) and the *opto-electronic components* (photodiode, phototransistor, etc.) which have been developed especially for use with optical fibres.

9. Microchips

By the end of the 1940s, as electronic systems grew in importance and size, more power was consumed, their weight and costs increased while their reliability decreased. The cold war and the space race both demanded cheaper, smaller and more reliable electronic systems. Following on from the invention of the transistor in 1948, J.S. Kilby completed the first integrated circuit in 1958; it consisted of a phase oscillator formed entirely of semiconductors (transistors, resistors and capacitors) in a single block called a *microchip*.

The use of large computers to design ICs, together with rapid developments in the chemical and photo-optical technologies applied to semiconductor electronics, has led to:

- greater purity of the silicon wafer used as the substrate;
- more accurate control of the doping of the different substrate regions which provide the basic components of the electronic circuit (transistors, diodes, resistors, capacitors, etc.);
- the highly-accurate fabrication of microscopic material layers (junctions, connecting tracks, connection soldering);
- continuous improvements in the performance of chips due the use of new materials, such as gallium arsenide and new manufacturing technologies (e.g. planar process, MOS technology);
- a higher integration capability to reduce the size of the circuits so that a small chip can include a greater number of functions and components.

Nowadays, integrated circuits are found in every field of electronics; they have entirely replaced circuits built from discrete components.

10. Digital signal processing

The processing of an electrical signal essentially involves:

- signal processing both in the time and frequency domains;

Michele Lemme *graduated in Electrotechnical Engineering from the University of Naples in 1952 and initially stayed on as a Research Assistant at the University.*

In 1956, he joined Vatican Radio at a time when the St. Maria di Galeria transmitting centre was in the process of being set up.

Since then, until his recent retirement, Mr Lemme has devoted his whole working life to St. Maria di Galeria, where he was the Director and the prime motor behind all its developments.

Rolando Menicucci *graduated in Radiocommunications from the G. Galilei Technical Industrial Institute of Rome in 1958.*

After completing a special course in Nuclear Physics at the National Research Council (CNR), he joined the Selenia company – where he was granted a scholarship – and worked on the development of radar systems in the project laboratory.

Mr Menicucci also worked in a company which produced car radio receivers before joining Vatican Radio in 1964. He now heads the telecommunications section of Vatican Radio.

- deleting unwanted information (e.g. filtering out noise);
- extracting specific information from a signal;
- deleting redundant and/or irrelevant information;
- adding extra information (e.g. encryption);
- storing, manipulating and transmitting signals with the best fidelity and reliability.

The processing of analogue signals (amplification, modulation, etc.) is very difficult and is restricted by real limiting factors that can sometimes be improved upon, but never removed entirely:

- the circuits need to be very linear with low noise;
- a special design is needed for almost every application;
- the circuits depend on sensitive components (resistors, capacitors) which need adjustment;
- circuits tend to drift in both the short-term and the long-term;
- the signals are difficult to record and store;
- the signals are not easy to compress.

Digital signals, on the other hand, use devices (such as amplifiers) which have only two operating states: saturation (on) and cutoff (off). It is thus possible to obtain better use, reliability and efficiency from each device in the chain. Futhermore, the on-off signals can be regenerated many times throughout the chain, with no losses whatsoever.

In 1938, Alec H. Reeves patented an invention concerning the transmission of voice signals by using coded pulses of constant amplitude, similar to those used in telegraphy. This invention was called *pulse code modulation* (PCM). However, it was not until 1962 that such an invention found an application in the field of telephony, with the production of the first 24-channel PCM system by Bell Systems of the USA.

In a PCM system, the conversion of the analogue signal into a digital signal is carried out through three different steps: *sampling*, *quantizing* and *coding*.

In the sampling step, the analogue input signal is transformed into a timed discrete signal, by pulse amplitude modulation (PAM) of a carrier formed of equally time-spaced pulses. The modulation is carried out by changing the amplitude of the transmitted pulses in sympathy with changes in the amplitude of the input analogue signal. The resultant PAM signal is then “quantized” by assigning to each sample a discrete level of amplitude, selected according to a definite level scale (e.g. 256 quantized levels). The amplitudes of the PAM samples are then “coded” according to a succession of elementary bits (8-bit words in the case of 256 quantized levels). The sampling may also be carried out by either modulating the *duration* or the *position* of the sampling pulse, instead of its *amplitude*.

Michele Lemme *graduated in Electrotechnical Engineering from the University of Naples in 1952 and initially stayed on as a Research Assistant at the University.*

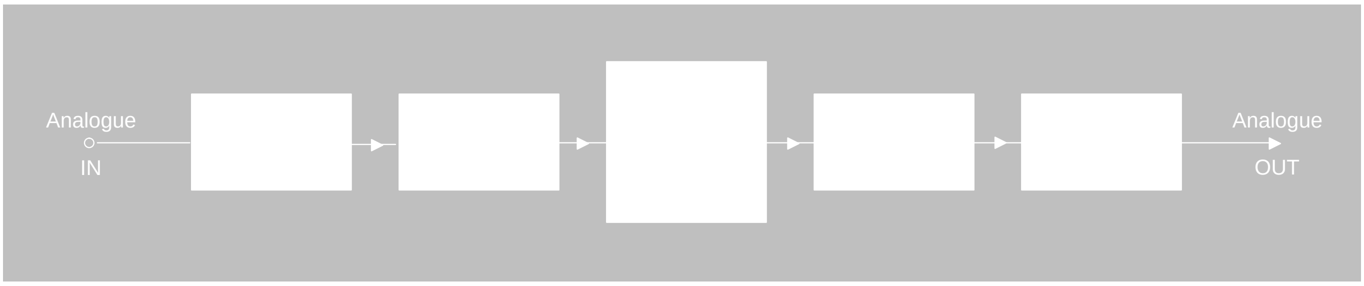
In 1956, he joined Vatican Radio at a time when the St. Maria di Galeria transmitting centre was in the process of being set up.

Since then, until his recent retirement, Mr Lemme has devoted his whole working life to St. Maria di Galeria, where he was the Director and the prime motor behind all its developments.

Rolando Menicucci *graduated in Radiocommunications from the G. Galilei Technical Industrial Institute of Rome in 1958.*

After completing a special course in Nuclear Physics at the National Research Council (CNR), he joined the Selenia company – where he was granted a scholarship – and worked on the development of radar systems in the project laboratory.

Mr Menicucci also worked in a company which produced car radio receivers before joining Vatican Radio in 1964. He now heads the telecommunications section of Vatican Radio.


Analogue
IN
Analogue
OUT

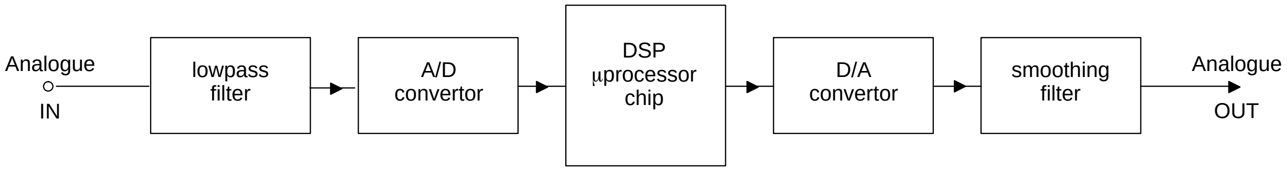


Figure 14
Schematic diagram of a DSP channel.

The PCM system was the forerunner to the modern technique of digitally processing an analogue signal in real time. Digital signal processing (DSP) techniques use the same main procedures as the PCM technique (i.e. sampling, quantizing and coding).

Coding is usually carried out by *data compression*, i.e. by reducing the redundancy which is inherent in the source analogue signal. Such compression is often based upon so-called *adaptive differential coding*, in which the difference between the input signal and its most recent estimate is coded, not the absolute value of the input signal.

The digital values of the coded signals are processed by software techniques (i.e. operations such as filtering, modulation and delay, which were typically executed by hardware, are now executed by software). Thus, once the signal is converted into digital form, it is possible to process it in both the time and frequency domains, by applying the Fourier transform and other mathematical algorithms.

A typical DSP channel is illustrated in *Fig. 14*. The input analogue signal enters the channel via a low-pass filter and is next passed through an A/D converter to ensure that the DSP unit receives a digital, band-limited version of the original input signal.

The DSP chip processes the digital data with the algorithm stored in its memory and then passes the modified data onto a D/A converter. As the output of the D/A converter includes discrete values rather than a pure analogue value, the signal must be passed through a smoothing filter to remove unwanted high-frequency components.

DSP techniques offer the following advantages over analogue processing methods:

– computer simulation to check the analogue system performance;

– increased compactness of the equipment;

– lower sensitivity to the environment and to component ageing;

– no need to tune the components;

– availability of sophisticated functions not covered by conventional analogue techniques;

– mathematical equations can be implemented directly as signal processing algorithms;

– less sensitivity to noise.

11. Summary

The first communication systems were based on the presence or absence of a transmitted signal (e.g. Morse code, telegraphy) and thus were truly *digital systems*. The work of Marconi and other pioneers, however, was devoted to finding better ways to transmit speech, music and other sounds as *analogue signals* and, consequently, a wide range of linear electronic devices were developed over the years to increase the range and quality of these analogue broadcasts.

Now, as the first century of wireless communication draws to a close, development efforts are once again being aimed vigorously in the direction of digital technologies.

Bibliography

[1] Solari, L.: *Storia della Radio*
Treves, 1939.

[2] Riley, J.P.: *Who started ripples in the ether?*
Electronics World + Wireless World,
September 1994.

[3] Pickworth, G.: *The spark that gave Radio to the world*
Electronics World + Wireless World,
July 1994.

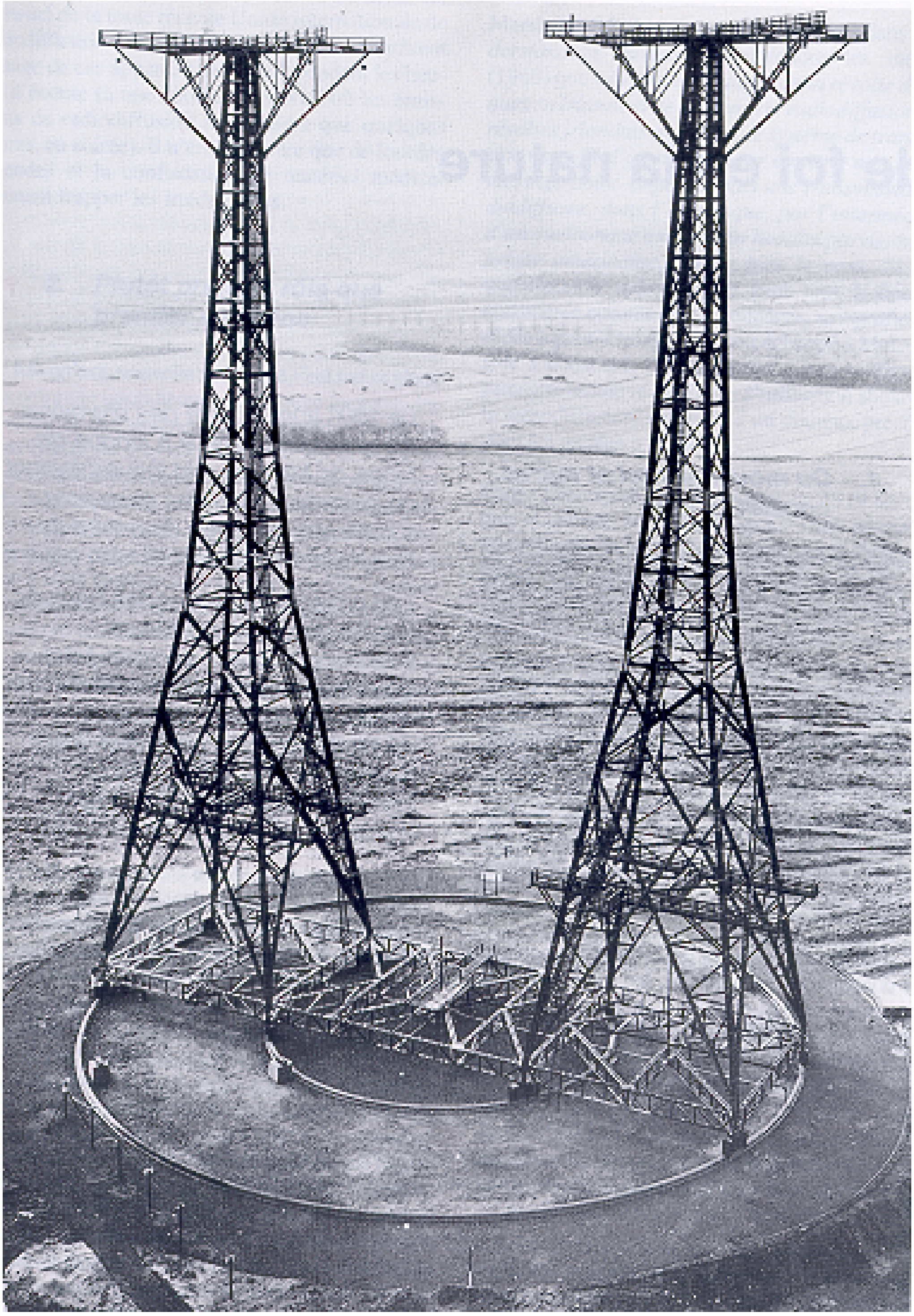
[4] O'Dell, D.H.: *Marconi's magnetic domain that stretches into the ether*
Electronics World + Wireless World,
August 1993.

[5] Shockley, W.: *Transistor Electronics; Imperfections, Unipolar and Analog Transistors*
Proceedings of the I.R.E., Vol. 40, No. 11,
November 1952.

[6] Fuseli, D.: *Elettronica verso l’Integrazione* Zanichelli, 1990.

[7] Setti, G.: *Manuale di Reti di Telecomunicazioni e Trasmissione Dat*i, Calderini, 1991.

[8] Texas Instruments Seminar on DSP, 1993.



Twin HF antenna towers on a rotating platform in the Netherlands (1938).