Vacuum Tube Devices in Electronics

Leonel Davila

First Edition, 2012

Table of Contents

Chapter 1

Vacuum Tube

In electronics, a **vacuum tube**, **electron tube** (in North America), or **thermionic valve** (elsewhere, especially in Britain) is a device that relies on the flow of electric current through a vacuum. Vacuum tubes may be used for rectification, amplification, switching, or similar processing or creation of electrical signals. Vacuum tubes rely on thermionic emission of electrons from a hot filament or *cathode*, that then travel through a vacuum toward a positively-charged anode or *plate*. Additional electrodes interposed between the cathode and anode can alter the current flow, making the device an amplifier.

Modern vacuum tubes, mostly miniature style.

Vacuum tubes were critical to the development of electronic technology, which drove the expansion and commercialization of radio communication and broadcasting, television, radar, sound reproduction, large telephone networks, analog and digital computers, and industrial process control. Although some of these applications had counterparts using earlier technologies, such as the spark gap transmitter or mechanical computers, it was the invention of the triode vacuum tube and its capability of electronic amplification that made these technologies widespread and practical.

For the most part vacuum tubes have been replaced by solid-state devices such as transistors and other semiconductor devices. Solid-state devices last much longer, are smaller, more efficient, more reliable, and cheaper than equivalent vacuum tube devices. However, tubes still find particular uses where solid state devices have not been developed or are not practical. Tubes are still produced for such applications and to replace those used in existing equipment such as high-power radio transmitters.

Classification

Vacuum tubes with two active elements ("diodes") are used for rectification. Ones with 3 or more elements ("triodes", "tetrodes", etc.) are used for amplification or functions which rely on amplification such as oscillators. Tubes used in consumer electronic equipment are often classified as "receiving tubes," as opposed to the much larger "transmitting tubes" used to generate high power radio signals for transmission.

On the other hand there are vacuum tubes used in different manners, such as cathode ray tubes which create a beam of electrons for display purposes (such as the television picture tube) in addition to more specialized functions such as electron microscopy and electron beam lithography. X-ray tubes are also vacuum tubes. Phototubes and photomultipliers also rely on electron flow through a vacuum, though in this case the emission of electrons from the cathode depends on energy from photons rather than thermionic emission. Since these sorts of "vacuum tubes" have functions other than electronic amplification and rectification they are described in their own articles.

Modern applications

Specialized applications for amplifying vacuum tubes continue to this day, such as the magnetron which is used to generate microwave energy in the household microwave oven and some radar systems. The klystron is commonly deployed by broadcasters as a high-power UHF television transmitting tube. Hi-fi equipment using tubes is still popular among certain audiophiles for its distinct sound signature and other tube equipment is maintained for its aesthetic appeal.

Gas-filled tubes

There are also varieties of current-conducting tubes filled with one or another gas at a higher or lower pressure; the common fluorescent bulb is a familiar example. However certain types such as the voltage regulator tube and thyristor physically resemble

commercial vacuum tubes and fit in sockets designed for vacuum tubes. Their distinctive orange, red, or purple glow during operation betrays the presence of gas; electrons flowing in a vacuum do not produce light within that region. Although not properly termed vacuum tubes, they may still be referred to as "electron tubes" as they do perform electronic functions, and are briefly discussed below under "Special-purpose tubes."

Description

A vacuum tube consists of two or more electrodes in a vacuum inside an airtight enclosure. Most tubes have glass envelopes, though ceramic and metal envelopes (atop insulating bases) have also been used. The electrodes are attached to leads which pass through the envelope via an airtight seal. On most tubes, the leads, in the form of pins, plug into a tube socket for easy replacement of the tube. (Tubes were by far the most common cause of failure in electronic equipment, and consumers were expected to be able to replace tubes themselves).
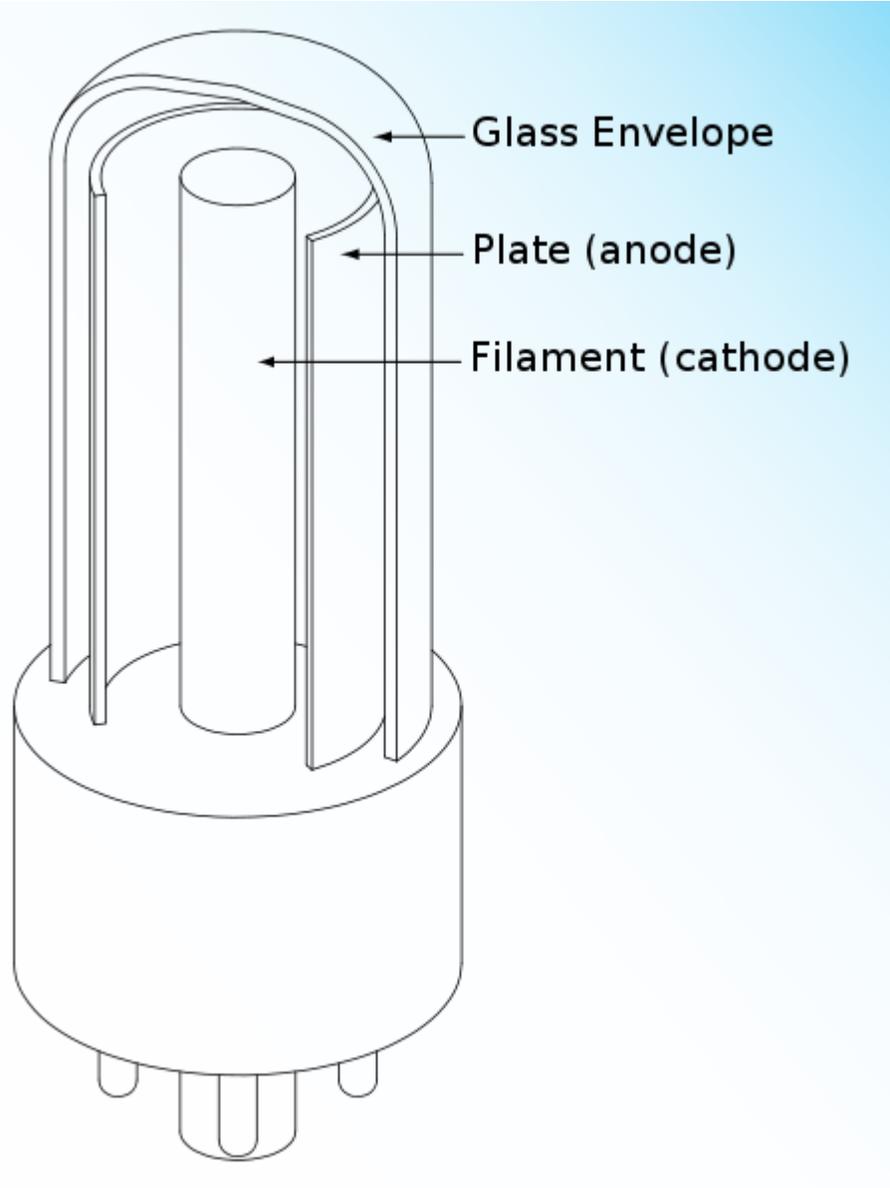
Vacuum tube diode: electrons from the hot cathode flow towards positive anode, but not visa versa.

The earliest vacuum tubes resembled, and in fact evolved from incandescent light bulbs, containing a filament sealed in an evacuated glass envelope. When hot, the filament releases electrons into the vacuum, a process called thermionic emission. These electrons will be drawn to a more positive electrode, the anode or *plate*. The result is a net flow of electrons from filament to plate. However current cannot flow in the reverse direction because the plate is not heated and does not emit electrons. Such a tube with only two electrodes is termed a diode, and is used for rectification. Since current can only pass in one direction, such a diode (or *rectifier*) will convert AC to DC. This is therefore used in a DC power supply, but is also used as a demodulator of amplitude modulated (AM) radio signals, and similar functions.

While early tubes used the directly-heated filament as the cathode, most (but not all) more modern tubes employed indirect heating. A separate element was used for the cathode. Inside the cathode, and insulated from it, was the filament or *heater*. The heater warmed the cathode sufficiently to undergo thermionic emission, but avoided any electrical connection. This allowed the tubes in a radio set to be heated through a common circuit, while allowing each cathode to arrive at a voltage independly of the others, removing an unwelcome constraint on circuit design.

During operation vacuum tubes require constant heating of the filament, so that they require considerable power even when amplifying signals at the microwatt level. In most amplifiers further power is consumed due to the quiescent current between the cathode and the plate (anode), resulting in heating of the plate. In a power amplifier heating of the plate can be quite considerable, and has a potential for self-destruction if the tube is driven beyond its safe limits. Since the tube requires a vacuum to operate, convection cooling of the plate is not generally possible (except in special applications where the anode forms a part of the vacuum envelope; this is avoided in consumer products due to the shock hazard it entails). Thus anode cooling occurs mainly through black-body radiation.
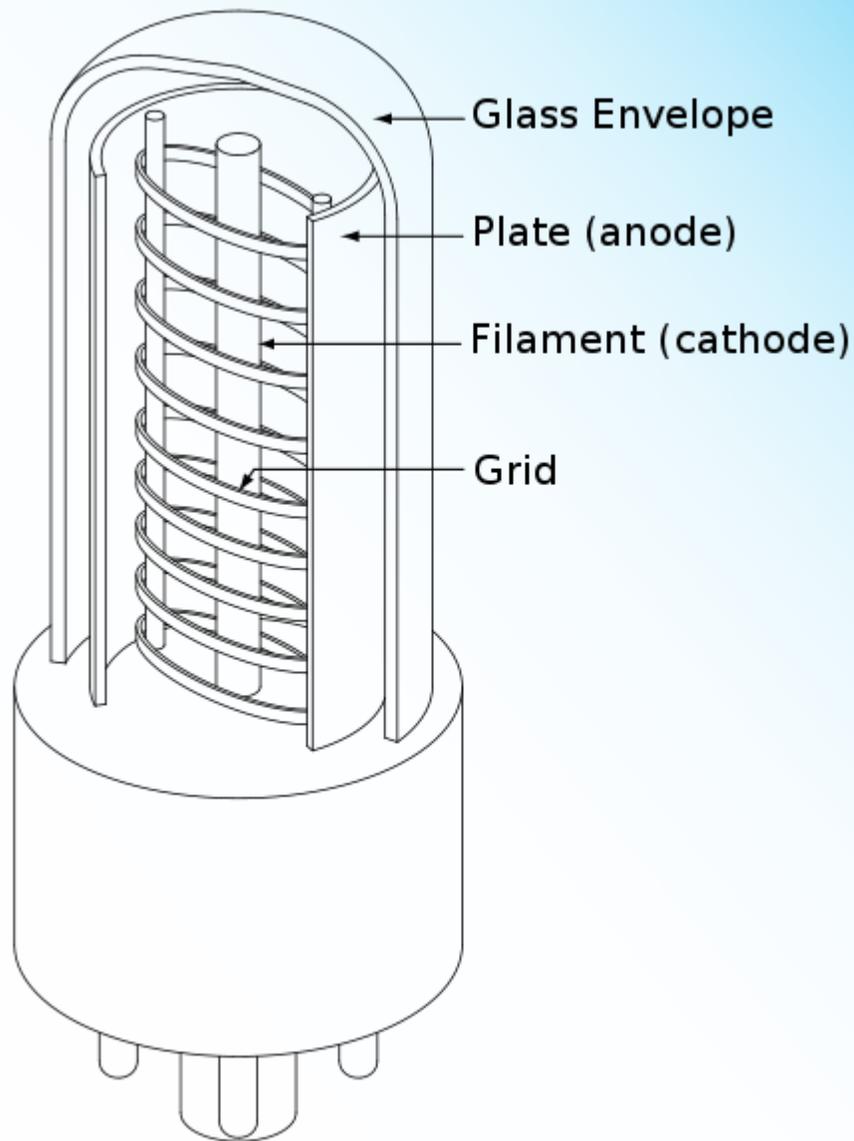
Vacuum tube triode: voltage applied to the grid controls plate current.

With the exception of diode tubes, another electrode, called a control grid, is placed between the cathode and the plate. The vacuum tube is then known as a "triode." With additional grids they are called tetrode, pentode, etc. These intervening electrodes are all called *grids* as they are not solid electrodes but sparse elements through which electrons can pass on their way to the plate. The control grid (and sometimes other grids) turn the diode into a *voltage-controlled device*, that is, the voltage that is applied to the control grid will affect the current flow between the cathode and the plate. A negative electrostatic field from the control grid repels electrons emitted by the cathode, rather than allowing them to continue toward the plate, thus reducing or even completely stopping the current flow. As long as the control grid stays more negative than the cathode, essentially no current flows into it, yet a change of several volts on the control

grid is sufficient to make a large difference in the plate current, possibly changing the output by hundreds of volts (depending on the load of the circuit). The solid-state device most closely resembling the pentode tube is the JFET, although vacuum tubes typically operate at over a hundred volts, unlike most semiconductors in most applications.

History and development

Early RCA triode vacuum tube, type 808

The 19th century saw increasing research with evacuated tubes, such as the Geissler and Crookes tubes. Famous scientists who experimented with such tubes included Thomas Edison, Eugen Goldstein, Nikola Tesla, and Johann Wilhelm Hittorf among many others. With the exception of early light bulbs, such tubes were only used in scientific research

or as novelties. The groundwork laid by these scientists and inventors, however, was critical to the development of subsequent vacuum tube technology.

Although thermionic emission was originally reported in 1873 by Frederick Guthrie, it was Thomas Edison's 1884 investigation that spurred future research, the phenomenon thus becoming known as the "Edison Effect." Edison patented what he found, but he did not understand the underlying physics, nor did he have an inkling of the potential value of the discovery. It wasn't until the early 20th century that the rectifying property of such a device was utilized, most notably by John Ambrose Fleming who used the diode tube to detect (demodulate) radio signals. Lee De Forest's 1906 "audion" was also developed as a radio detector, and soon led to the developement of the triode tube. This was essentially the first electronic amplifier, leading to great improvements in telephony (such as the first coast-to-coast telephone line in the US) and revolutionizing the technology used in radio transmitters and receivers. The electronics revolution of the 20th century arguably began with the invention of the triode vacuum tube.
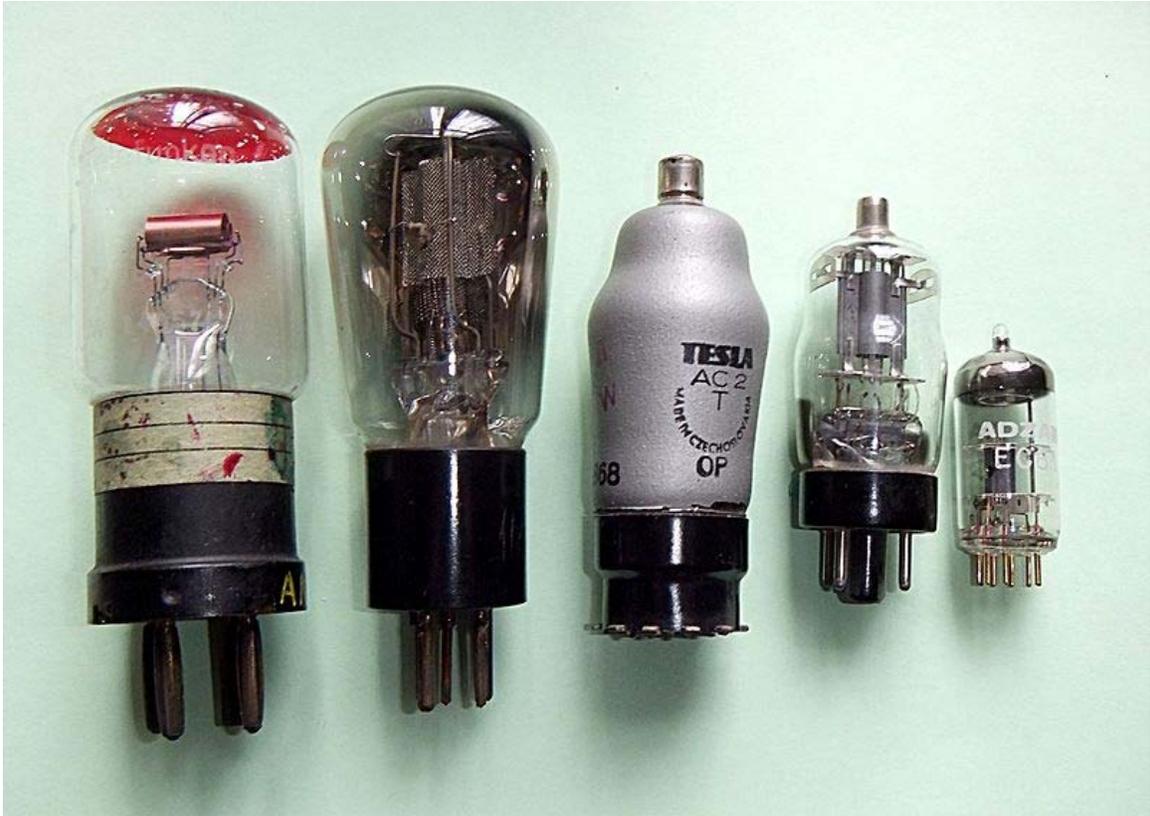
Diodes

The English physicist John Ambrose Fleming worked as an engineering consultant for firms including Edison Telephone and the Marconi Company. In 1904, as a result of experiments conducted on Edison effect bulbs imported from the USA, he developed a device he called an "oscillation valve" (because it passes current in only one direction). The heated filament, or cathode, was capable of thermionic emission of electrons that would flow to the *plate* (or *anode*) when it was at a higher voltage. Electrons, however, could not pass in the reverse direction because the plate was not heated and thus not capable of thermionic emission of electrons.

Later known as the Fleming valve, it could be used as a rectifier of alternating current and as a radio wave detector. This greatly improved the crystal set which rectified the radio signal using an early solid-state diode based on a crystal and a so-called cat's whisker. Unlike modern semiconductors, such a diode required painstaking adjustment of the contact to the crystal in order for it to rectify. The diode tube was a reliable alternative for rectifying radio signals. Higher power diode tubes or *power rectifiers* found their way into power supply applications until they were eventually replaced by silicon rectifiers in the 1960s.
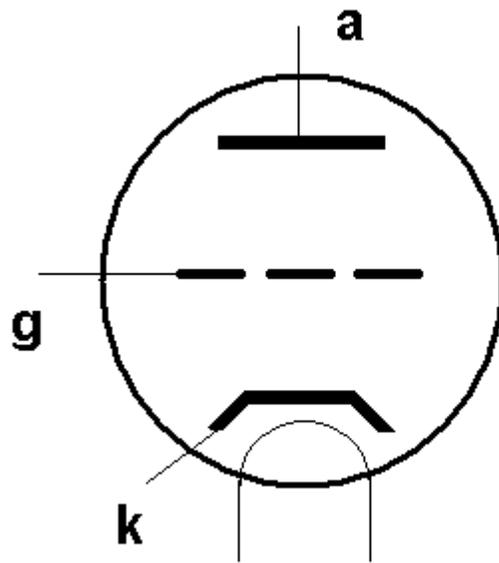
Triodes
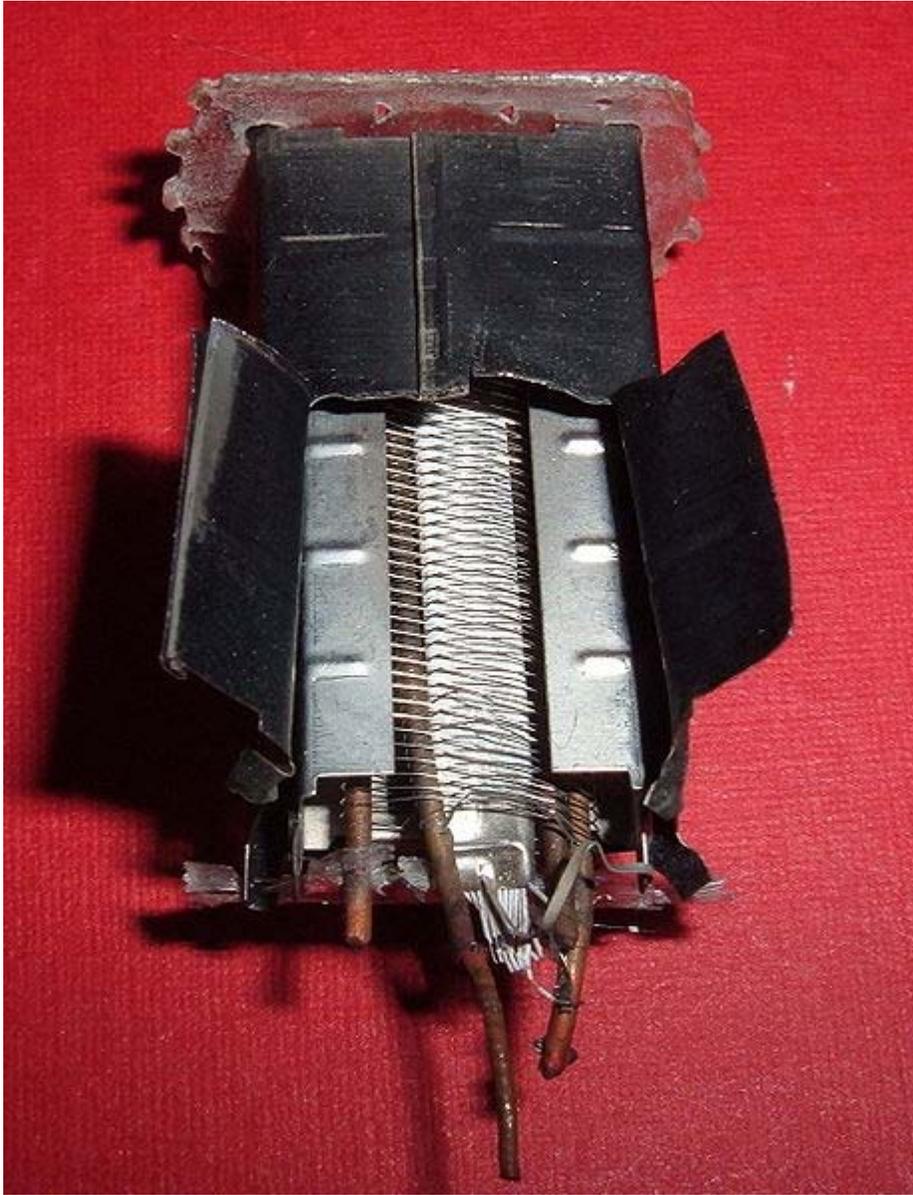
Vacuum tube with plate cut open revealing grid.

Triodes as they evolved over 40 years of tube manufacture, from the RE16 in 1918 to a 1960's era miniature tube.

Triode symbol. From top to bottom: plate (anode), control grid, cathode, heater (filament)

Originally, the only use for tubes in radio circuits was for rectification, not amplification. In 1906 Robert von Lieben filed for a patent for a cathode ray tube which included magnetic deflection. This could be used for amplifying audio signals and was intended for use in telephony equipment. He would later go on to help refine the triode vacuum tube.

However it was Lee De Forest who in 1907 is credited with inventing the triode tube while continuing experiments to improve his original Audion tube, a crude forerunner of the triode. By placing an additional electrode in between the filament (cathode) and plate, he discovered the ability of the resulting device to amplify signals of all frequencies. As the voltage applied to the so-called control grid (or simply "grid") was lowered from the cathode's voltage to somewhat more negative voltages, the amount of current flowing from the filament to the plate would be reduced. The negative electrostatic field created by the grid in the vicinity of the cathode would inhibit thermionic emission and reduce the current to the plate. Thus a few volts difference at the grid would make a large change in the plate current and could lead to a much larger voltage change at the plate, resulting in voltage and power amplification. In 1907, De Forest filed for a patent for such a three-electrode version of his original Audion tube for use as an electronic amplifier in radio communications. This eventually became known as the triode.

De Forest's device was not strictly a vacuum tube, as he erroneously believed that it depended on the presence of residual gas remaining after evacuation. The De Forest company, in its Audion leaflets, even warned against operation which might lead to too high a vacuum! The Finnish inventor Eric Tigerstedt significantly improved on the original triode design in 1914, while working on his sound-on-film process in Berlin, Germany. The first true vacuum triodes in production were the Pliotrons developed by Irving Langmuir at the General Electric research laboratory (Schenectady, New York) in 1915. Langmuir was one of the first scientists to realize that a harder vacuum would improve the amplifying behaviour of the triode. Pliotrons were closely followed by the French 'R' Type which was in widespread use by the allied military by 1916. These two types were the first true *vacuum* tubes; early diodes and triodes performed as such despite a rather high residual gas pressure. Techniques to produce and maintain better vacuums in tubes were then developed. Historically, vacuum levels in production vacuum tubes typically ranged from 10 μPa down to 10 nPa.

The non-linear operating characteristic of the triode caused early tube audio amplifiers to exhibit harmonic distortions at low volumes. This is not to be confused with the so-called overdrive distortion that tube amplifiers exhibit when driven beyond their linear region (known as the tube sound). To remedy the triode's nonlinear characteristics, engineers plotted curves of the applied grid voltage and resulting plate currents, and discovered that there was a range of grid voltages allowing for relatively linear operation. In order to use this range, a negative voltage had to be applied to the grid to place the tube in the "middle" of the linear area with no signal applied. This was called the idle condition, and the plate current at this point the "idle current". Today this current would be called the quiescent or bias current. The controlling voltage was superimposed onto this fixed "bias" voltage, resulting in a linear variation of plate current in response to both positive

and negative variation of the input voltage around that point. This concept is called *grid bias*. Many early radio sets had a third battery called the "C battery" (not to be confused with the modern C cell) whose positive terminal was connected to the cathode of the tubes (or "ground" in most circuits) and whose negative terminal supplied this bias voltage to the grids of the tubes. More modern circuits used cathode biasing in lieu of a separate negative power supply.

When triodes were first used in radio transmitters and receivers, it was found that tuned amplification stages had a tendency to oscillate unless their gain was very limited. This was due to the parasitic capacitance between the plate (the amplifier's output) and the control grid (the amplifier's input), known as the Miller capacitance. Eventually the technique of *neutralization* was developed whereby the RF transformer connected to the plate would include an additional winding in the opposite phase. This winding would be connected back to the grid through a small capacitor, and when properly adjusted would cancel the Miller capacitance. This technique was employed and led to the success of the Neutrodyne radio during the 1920s. However neutralization required careful adjustment and proved unsatisfactory when used over a wide ranges of frequencies.

Tetrodes and pentodes
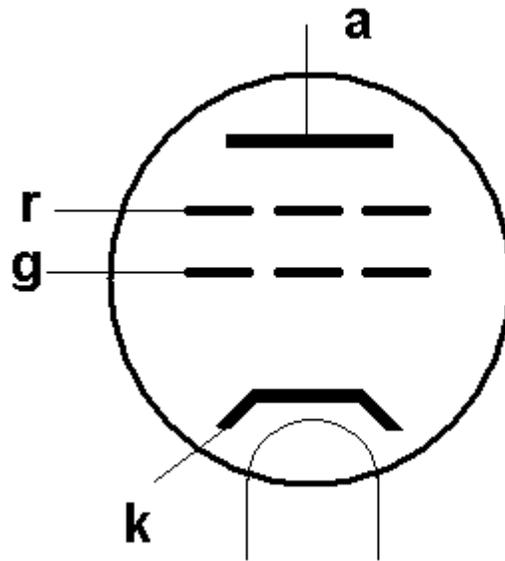
Tetrode symbol. From top to bottom: plate (anode), screen grid, conrol grid, cathode, heater (filament)

In order to combat the stability problems and limited voltage gain due to the Miller effect, the noted physicist Walter H. Schottky invented the tetrode tube in 1919. He showed that the addition of a second grid, located between the control grid and the plate, known as the *screen grid*, could solve these problems. ("Screen" in this case refers to electrical

"screening" or shielding, not physical construction: all "grid" electrodes in between the cathode and plate are "screens" of some sort rather than solid electrodes since they must allow for the passage of electrons directly from the cathode to the plate). A positive voltage slightly lower than the plate voltage was applied to it, and was bypassed (for high frequencies) to ground with a capacitor. This arrangement decoupled the anode and the control grid, essentially eliminating the Miller capacitance and its associated problems. Consequently higher voltage gains from a single tube became possible, reducing the number of tubes required in many circuits. This two-grid tube is called a *tetrode*, meaning four active electrodes, and was common by 1926.

However, the tetrode has one new problem. In any tube, electrons strike the anode with sufficient energy to cause the emission of electrons from its surface. In a triode this so-called secondary emission of electrons is not important since they are simply re-captured by the more positive anode. But in a tetrode they can be captured by the screen grid since it is at also at a high voltage, thus robbing them from the plate current and reducing the amplification of the device. Since secondary electrons can outnumber the primary electrons, in the worst case, particularly as the plate voltage dips below the screen voltage, the plate current can actually go down with increasing plate voltage. This is termed negative resistance and can itself cause instability. This is the so-called "tetrode kink". Another consequence of secondary emission is that in extreme cases the current reaching the screen grid can cause it to overheat to the point of destroying the tube.

Vacuum tubes in an Australian radio of the late 1930s

The solution was to add one more grid in between the screen grid and the plate, called the suppressor grid (since it supressed secondary emission current toward the screen grid). This grid was held at the cathode (or "ground") voltage and its negative voltage (relative to the anode) electrostatically repelled secondary electrons so that they would be collected by the anode after all. This three-grid tube is called a pentode, meaning five

electrodes. The pentode was invented in 1928 by Bernard D. H. Tellegen and became generally favoured over the simple tetrode. A refinement of the tetrode or pentode for power applications is the beam tetrode or "beam power tube", discussed berlow.

Multifunction configurations
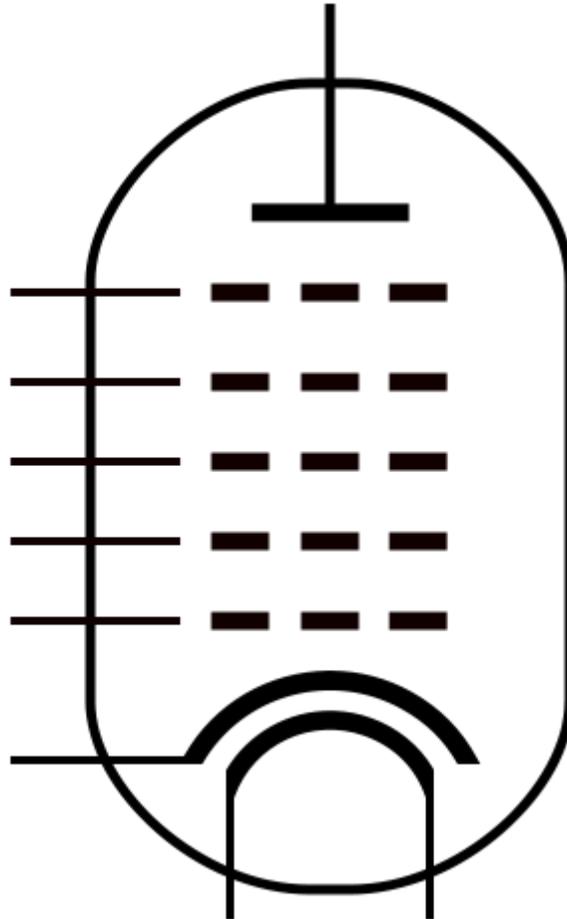
The pentagrid converter contained no less than 5 grids in between the cathode and the plate.

Superheterodyne receivers require a local oscillator and mixer, which required two tubes. With the development of the pentagrid converter, these functions were combined inside a single tube which applied the RF signal to the control grid, but also implemented the local oscillator using additional grids. Various alternatives such as using a combination of a triode with a hexode and even an octode have been used for this purpose. The additional grids include both control grids (at a low potential) and screen grids (at a high voltage). Many designs used such a screen grid as a second 'leaky' plate to provide feedback for the oscillator oscillator function, whose current was added to that of the incoming radio frequency signal. Due to the large oscillating signal nonlinearity of the tube response caused frequency mixing, seen on the plate current (output) of such a "converter" circuit. The difference frequency between that of the incoming signal and that

of the oscillator was selected by a tuned transformer, becoming the input to the receivers's intermediate frequency (IF) amplifier.

The pentagrid converter such as the 12BE6 thus became widely used in AM receivers including the minature tube version of the "All American Five". Octodes such as the 7A8 were rarely used in the US, but much more common in Europe particularly in battery operated radios where the lower power consumption was an advantage.

To further reduce the cost and complexity of radio equipment, two separate vacuum tubes could be combined in the bulb of a single tube, a so-called *multisection tube*. An early example was the Loewe 3NF. This 1920s device had 3 triodes in a single glass envelope together with all the fixed capacitors and resistors required to make a complete radio receiver. As the Loewe set had only one tube socket, it was able to substantially undercut the competition since, in Germany, state tax was levied by the number of sockets. However, reliability was compromised, and production costs for the tube were much greater. In a sense, these were akin to integrated circuits. In the US, Cleartron briefly produced the "Multivalve" triple triode for use in the Emerson Baby Grand receiver. This Emerson set also had a single tube socket, but because it used a four-pin base, the additional element connections were made on a "mezzanine" platform at the top of the tube base.
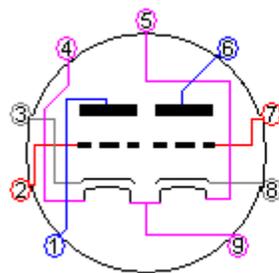
1. Anode Triode Number 2
2. Grid Triode Number 2
3. Cathode Triode Number 2
4. Heater (Triode 2)
5. Heater (Triode 1)
6. Anode Triode Number 1
7. Grid Triode Number 1
8. Cathode Triode Number 1
9. Heater Center tap

Popular 12AX7 dual triode

Compactron tube: 12AE10, dual pentode

By 1940 multisection tubes had become commonplace. There were constraints, however, due to patents and other licencing considerations. Constraints due to the number of external pins (leads) often forced the functions to share some of those external connections such as their cathode connections (in addition to the heater connection). The RCA Type 55 was a double diode triode used as a detector, automatic gain control rectifier and audio preamp in early AC powered radios. These sets often included the 53 Dual Triode Audio Output. Another early type of multi-section tube, the 6SN7, is a "dual triode" which, for most purposes, can perform the functions of two triode tubes, while taking up half as much space and costing less. The 12AX7 is a dual high voltage gain (or *high mu*) triode in a minature enclosure, and became widely used in audio signal amplifiers, instruments, and.guitar amplifiers.

The introduction of the miniature tube base (see below) which could have 9 pins, also allowed many other multi-section tubes, such as the 6GH8 triode + pentode. Along with a host of similar tubes, the 6GH8 was quite popular in television receivers. Some color TV sets used exotic types like the 6JH8 which had two plates and beam deflection electrodes (it was known as the 'sheet beam' tube). Vacuum tubes used like this were designed for demodulation of synchronous signals, an example of which is color demodulation for television receivers. The desire to include additional functions in one envelope resulted in the General Electric Compactron which had 12 pins (minature tubes had only 7 or 9 pins). A typical example, the 6AG11, contained two triodes and two diodes.

Beam power tubes
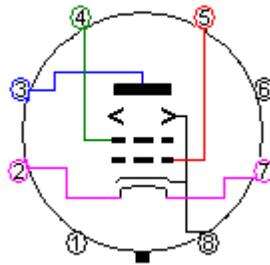
Beam power tube symbol and pinout for 6L6

The beam power tube is usually a tetrode with the addition of beam-forming electrodes, which take the place of the suppressor grid. These angled plates focus the electron stream onto certain spots on the anode which can withstand the heat generated by the impact of massive numbers of electrons, while also providing pentode behavior. The positioning of the elements in a beam power tube uses a design called "critical-distance geometry", which minimizes the "tetrode kink", plate-grid capacitance, screen-grid current, and secondary emission effects from the anode, thus increasing power conversion efficiency. The control grid and screen grid are also wound with the same pitch, or number of wires per inch.

6L6 tubes in glass envelopes

Aligning the grid wires also helps to reduce screen current, which represents wasted energy. This design helps to overcome some of the practical barriers to designing high-power, high-efficiency power tubes. 6L6 was the first popular beam power tube, introduced by RCA in 1936. Corresponding tubes in Europe were the KT66, KT77 and KT88 by GEC (the KT standing for "Kinkless Tetrode").

Variations of the 6L6 design are still widely used in tube guitar amplifiers, making it one of the longest lived electronic device families in history. Similar design strategies are used in the construction of large ceramic power tetrodes used in radio transmitters.

Miniature tubes

Miniature tube, alongside euro coin
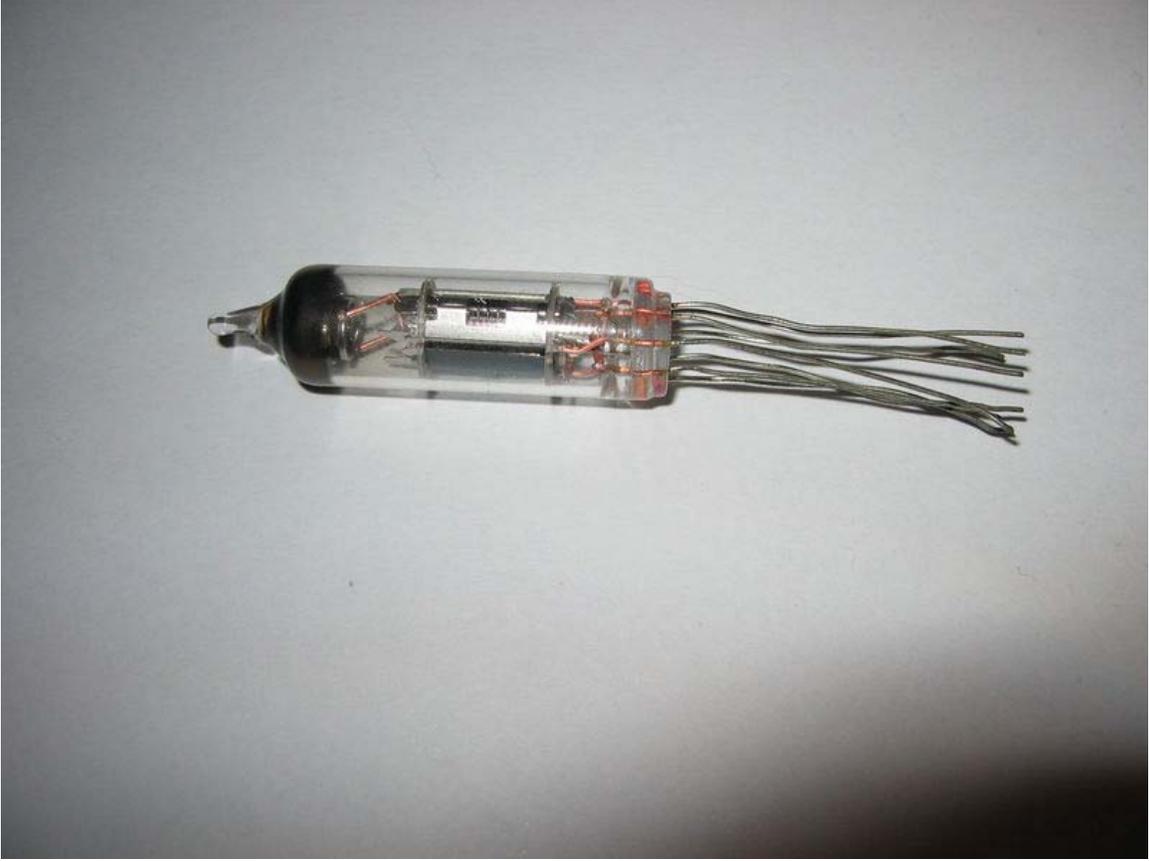
Subminiature CV4501 tube, 35 mm long x 10 mm diameter (excluding leads).

RCA 6DS4 "Nuvistor" triode, ca. 20 mm high by 11 mm diameter.

Early tubes used a metal or glass envelope atop an insulating bakelite base. In 1938 a technique was developed to instead use an all glass construction with the pins fused in the glass base of the envelope. This was used in the design of a much smaller tube outline, known as the miniature tube, having 7 or 9 pins. Making tubes smaller reduced the voltage that they could work at, and also the power of the filament. Miniature tubes became predominant in consumer applications such as radio receivers and hi-fi amplifiers. However the larger older styles continued to be used especially as higher power rectifiers, in higher power audio output stages and as transmitting tubes.

Subminiature tubes with a size roughly that of half a cigarette were used in hearing-aid amplifiers. These tubes did not have pins plugging into a socket but were soldered in

place. The "acorn" valve (named due to its shape) was another such example. Another very small tube style was called the nuvistor. About the size of a thimble, these metal cased tubes were made small not mainly for compactness, but for use at very high frequencies, notably in UHF television tuners.

Improvements in construction and performance

The very earliest vacuum tubes strongly resembled incandescent light bulbs and were made by lamp manufacturers, who had the equipment for manufacture of glass envelopes and the powerful vacuum pumps required to evacuate the enclosures. After World War I, specialized manufacturers using more economical construction methods were set up to fill the growing demand for broadcast receivers. Bare tungsten filaments operated at a temperature of around 2200 °C. The development of oxide-coated filaments in the mid 1920s reduced filament operating temperature to a dull red heat (around 700 °C), which in turn reduced thermal distortion of the tube structure and allowed closer spacing of tube elements. This in turn improved tube gain, since the gain of a triode is inversely proportional to the spacing between grid and cathode.

Indirectly heated cathodes

The desire to power electronic equipment using AC mains power faced a difficulty with respect to the powering of the tubes' filaments, as these were also the cathode of each tube. Powering the filaments directly from a power transformer would introduce 50 or 60 Hz hum into audio stages using tubes whose filaments were powered in such a manner. The invention of the "equi-potential cathode" reduced this problem, with the filaments being powered by a balanced AC power transformer winding having a grounded center tap.

A superior solution, and one which allowed each cathode to "float" at a different voltage, was that of the indirectly-heated cathode. Now, a filament inside a cylinder of oxide-coated nickel, provided for a cathode electrically isolated from the filament which could then just as well be powered by AC. In such tubes, the filament is frequently referred to as a *heater* to distinguish it as an inactive element. In the 1930s indirectly heated cathode tubes became widespread in equipment using AC power. However directly heated filament tubes continued to prevail in battery operated equipment, as the power requirements for these filaments were substantially lower than required by the heaters used to heat cathodes indirectly.

World War II

Near the end of World War II, to make radios more rugged, some aircraft and army radios began to integrate the tube envelopes into the radio's cast aluminium or zinc chassis. The radio became just a printed circuit with non-tube components, soldered to the chassis that contained all the tubes. During WWII in 1942, rugged metal vacuum tubes were mounted in anti-aircraft shells. These proximity fuzes made anti-aircraft shells 6 times more effective. In the fall of 1944, artillery shells with proximity fuses were used.

The tiny tubes were later known as "subminiature" types. They were widely used in 1950s military and aviation electronics.

Use in early electronic computers

The 1946 ENIAC computer used 17,468 vacuum tubes and consumed 150kW of power.

While the development of vacuum tubes made electronic computing possible for the first time, the cost and reliability of early tubes made such developments rather impractical. It was only the pressure of World War II that led to the development of Colossus and early electronic computers using tubes. The general practicality of electronic computers was only realized with the development of transistors over a decade later.

Colossus

Colossus (and its successor Colossus Mk2) was built by the British during World War II to substantially speed up the task of breaking the German high level Lorenz encryption. Based on 1500 vacuum tubes, Colossus replaced an earlier machine based on relay and

switch logic (the Heath Robinson). Colossus was able to break in a matter of hours messages that had previously taken several weeks. Colossus Mk2 used a total of around 2000 vacuum tubes. Colossus was the first ever use of vacuum tubes on such a large scale for a single machine. The largest project previously had used just 150 tubes and had proven to be extremely unreliable. The main design problem at Colossus's inception was how to make vacuum tube based equipment reliable when the tubes were used in large numbers.

The Colossus computer's designer, Dr. Tommy Flowers, had a theory that most of the unreliability was caused during power down and (mainly) power up. Once Colossus was built and installed, it was switched on and left switched on running from dual redundant diesel generators (the wartime mains supply being considered too unreliable). The only time it was switched off was for conversion to the Colossus Mk2 and the addition of another 500 or so tubes. Another 9 Colossus Mk2s were built, and all 10 machines ran with a surprising degree of reliability. The 10 Colossi consumed 15 kilowatts of power each, 24 hours a day, 365 days a year—nearly all of it for the tube heaters.

Whirlwind

To meet the reliability requirements of the early digital computer Whirlwind, it was necessary to build special "computer vacuum tubes" with extended cathode life. The problem of short lifetime was traced to evaporation of silicon, used in the tungsten alloy to make the heater wire easier to draw. Elimination of the silicon from the heater wire alloy (and paying extra for more frequent replacement of the wire drawing dies) allowed production of tubes that were reliable enough for the Whirlwind project. The tubes developed for Whirlwind later found their way into the giant SAGE air-defense computer system. High-purity nickel tubing and cathode coatings free of materials that can poison emission (such as silicates and aluminium) also contribute to long cathode life. The first such "computer tube" was Sylvania's 7AK7 of 1948. By the late 1950s it was routine for special-quality small-signal tubes to last for hundreds of thousands of hours, if operated conservatively. This increased reliability also made mid-cable amplifiers in submarine cables possible.

Heat generation and transfer

The anode of this transmitting triode has been designed to dissipate up to 500W of heat

A considerable amount of heat is produced when tubes operate, both from the filament (heater) but also from the stream of electrons bombarding the plate. The requirements for heat removal can significantly change the appearance of high-power vacuum tubes. Although the miniature tube style became predominant in consumer equipment, high power audio amplifiers and rectifiers would still require the larger "octal" style of enclosure. Transmitting tubes could be much larger still.

Most tubes produce heat from two sources during operation. The first source is the filament or heater. Some tubes contain a *directly heated cathode*. This is a filament

similar to an incandescent electric lamp; some types glow brightly like a lamp, but most glow a dim orange-red. The "bright emitter" types possess a tungsten filament alloyed with 1-3 % thorium which reduces the work function of the metal, giving it the ability to emit sufficient electrons at about 2000 degrees Celsius. The "dull emitter" types also possess a tungsten filament but it is coated in a mixture of calcium, strontium and barium oxides, which emit electrons easily at much lower temperatures due to a monolayer of mixed alkali earth metals coating the tungsten; these only reach 800-1000 degrees Celsius.

The second form of cathode is the *indirectly heated* form which usually consists of a nickel cylinder, coated on the outside with the same strontium, calcium, barium oxide mix used in the "dull emitter" directly heated types. Inside the cylinder is a tungsten filament to heat it. This filament is usually uncoiled and coated in a layer of alumina (aluminium oxide) in order to insulate it electrically from the actual cathode. This form of construction allows for a much greater electron emitting area and allows the cathode to be held at a potential difference, typically 150 volts more positive than the heater or 50 volts more negative than the heater. For small-signal tubes such as used in radio receivers, heaters consume between 50 mW and 5 watts, (directly heated), or between 500 mW and 8 watts for indirectly heated types. Thus even a small signal amplifier might consume a watt of power just to warm its heater, compared to the milliwatts (or less) that a modern semiconductor amplifier would require for the same function. Even in power amplifiers the filament power may be responsible for an appreciable reduction in efficiency.

The second source of heat is generated at the plate, as electrons accelerated by the anode voltage strike the plate, depositing their kinetic energy on it and raising its temperature. In tubes used in power amplifiers or transmitter output stages, this source of heat will far exceed the power due to the cathode heater. The plates of improperly operated or overloaded beam power tubes can sometimes become visibly red hot; this should never occur under normal operation of consumer electronics and is a precursor to tube failure.

Heat escapes the device by black body radiation from the anode/plate as infrared radiation. Convection is not possible in most tubes since the anode is surrounded by vacuum. Considerations of heat removal can affect the overall appearance of some tubes. The anode or plate is often treated to make its surface less shiny and darker in the infrared. The screen grid may also generate considerable heat, which is radiated toward the plate which must reradiate that additional heat along with the heat it generates itself. Limits to screen grid dissipation, in addition to plate dissipation, are listed for power devices. If these are exceeded then tube failure is likely.

Tubes used as power amplifier stages for radio transmitters may have additional heat exchangers, cooling fans, radiator fins, or other measures to improve heat transfer at the anode. High power transmitting tubes may have the surface of their anodes external to the tube, allowing for water cooling or evaporative cooling. Such a water cooling system must be electrically isolated to withstand the high voltage present on the anode.

Tubes which generate rather little heat, such as the 1.4 volt filament directly heated tubes designed for use in battery powered equipment, often have shiny metal anodes. 1T4, 1R5 and 1A7 are examples. Gas filled tubes such as thyratrons may also use a shiny metal anode since the gas present inside the tube allows for convection of heat from the anode to the glass enclosure.

The outer electrode in most tubes is the anode. Some small signal types, such as sharp and remote cut-off R.F. and A.F. pentodes and some pentagrid converters have a shield fitted around all the electrodes enclosing the anode. This shield is sometimes a solid metal sheet, treated to make it dull and gray so that it can itself reradiate heat generated from within. Sometimes it is fabricated from expanded metal mesh, acting as a Faraday cage but allowing sufficient infrared radiation from the anode to escape. Types 6BX6/EF80 and 6BK8/EF86 are typical examples of this shielded type using expanded mesh. Types 6AU6/EF94 and 6BE6/EK90 are examples which use a gray sheet metal cylindrical shield.

Tube packaging

Metal cased tubes with "octal" bases
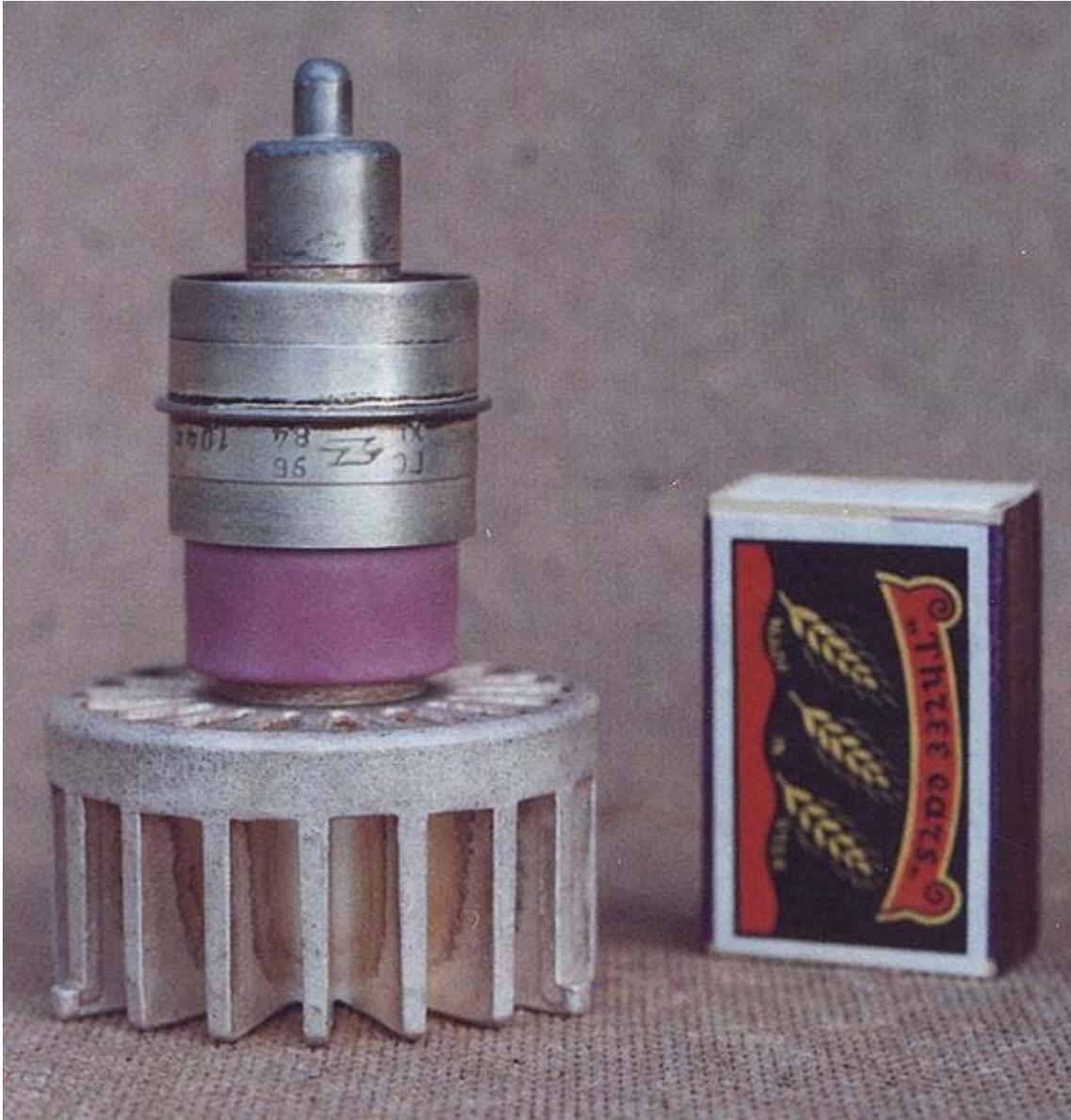
High power GS-9B triode transmitting tube with heat sink at bottom.

Most modern tubes have glass envelopes, but metal, fused quartz (silica), and ceramic have also been used. The first version of the 6L6 used a metal envelope sealed with glass beads, while a glass disk fused to the metal was used in later versions. Metal and ceramic are used almost exclusively for power tubes above 2 kW dissipation. The nuvistor was a modern receiving tube using a very small metal and ceramic package.

Tubes have always had their internal elements connected to external circuitry using pins at their base which plug into a socket. After all, tubes needed to be replaced rather frequently unlike modern semiconductor devices which are mostly soldered in place. Subminiature tubes were produced using wire leads rather than sockets, however these were restricted to rather specialized applications. In addition to the connections at the

base of the tube, many early triodes connected the grid using a metal cap at the top of the tube; this was done in order to reduce stray capacitance between the grid and the plate leads. Tube caps were also used for the plate (anode) connection, particularly in transmitting tubes and tubes using a very high plate voltage.

High power tubes such as transmitting tubes have packages designed more to enhance heat transfer. In some tubes, the metal envelope is also the anode. The 4CX1000A is an external anode tube of this sort. Air is blown through an array of fins attached to the anode, thus cooling it. Power tubes using this cooling scheme are available up to 150 kW dissipation. Above that level, water or water-vapor cooling are used. The highest-power tube currently available is the Eimac 4CM2500KG, a forced water-cooled power tetrode capable of dissipating 2.5 megawatts. (By comparison, the largest power transistor can only dissipate about 1 kilowatt.)

Special-purpose tubes

Voltage regulator tube in operation. Low pressure gas within tube glows due to current flow.

Some special-purpose tubes are constructed with particular gases in the envelope. For instance, voltage regulator tubes contain various inert gases such as argon, helium or neon, and take advantage of the fact that these gases will ionize at predictable voltages. The thyratron is a special-purpose tube filled with low-pressure gas or mercury vapor. Like vacuum tubes, it contains a hot cathode and an anode, but also a control electrode which behaves somewhat like the grid of a triode. When the control electrode starts conduction, the gas ionizes after which the control electrode no longer can stop the current; the tube "latches" into conduction. Removing plate (anode) voltage lets the gas

de-ionize, restoring its non-conductive state. Some thyratrons can carry large currents for their physical size. One example is the miniature type 2D21, often seen in 1950s jukeboxes as control switches for relays. A cold-cathode version of the thyratron, which uses a pool of mercury for its cathode, is called an Ignitron. Some ignitrons can switch thousands of amperes. Thyratrons containing hydrogen have a very consistent time delay between their turn-on pulse and full conduction, and have long been used in radar transmitters. Thyratrons behave much like modern silicon-controlled rectifiers, also known as thyristors in reference to their tube predecessor.

An extremely specialized tube is the Krytron, which is used for extremely precise, rapid high-voltage switching. Due to their intended purpose, the initiation of the precise sequence of detonations used to set off a nuclear weapon, they are heavily controlled at an international level.

X-ray tubes are used in medical imaging among other uses. X-ray tubes used for continuous duty operation in fluoroscopy and CT imaging equipment may use a focused cathode and a rotating anode to dissipate the large amounts of heat thereby generated. They are housed in an aluminum housing which is filled with oil for cooling. Nuclear medicine imaging equipment and liquid scintillation counters require photomultiplier tube arrays to detect scintillation due to ionizing radiation; the photomultiplier is a rare example of a vacuum tube which doesn't employ thermionic emission.

Powering the tube

Batteries

Batteries provided the voltages required by tubes in early radio sets. Three different voltages were generally required, using three different batteries designated as the **A'**, *B*, *and* **C** battery. The "A" batteries or LT (low-tension) battery provided the filament voltage. Tube heaters were designed for single, double or triple-cell lead-acid batteries, giving nominal heater voltages of 2 V, 4 V or 6 V. In portable radios, dry batteries were sometimes used with 1.5 or 1 V heaters. Reducing filament consumption improved the life span of batteries. By 1955 towards the end of the tube era, tubes using only 50 mA down to as little as 10 mA for the heaters had been developed.

The plate voltage was provided by the "B" batter or the HT (high-tension) supply or battery. These were generally of dry cell construction and typically came in 22.5, 45, 67.5, 90 or 135 volt versions.
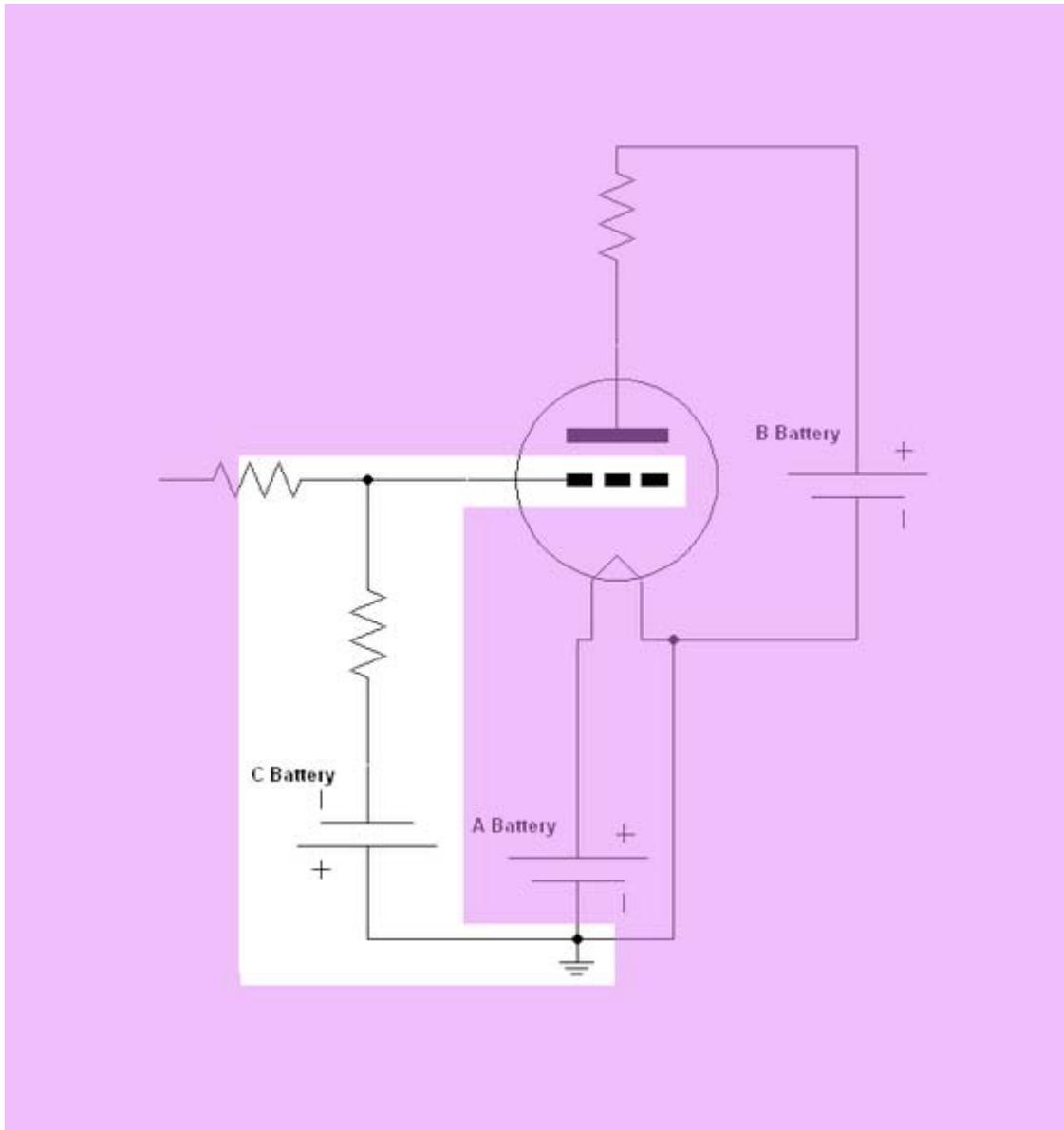
Batteries for a vacuum tube circuit. The C battery is highlighted.

Early sets used a grid bias battery or "C" batteries which was connected to provide a *negative* voltage. Since virtually no current flows through a tube's grid connection, these batteries had very low drain and lasted the longest. Even after AC power supplies became commonplace, some radio sets continued to be built with C batteries, as they would almost never need replacing. However more modern circuits were designed using cathode biasing, eliminating the need for a third power supply voltage; this became practical with tubes using indirect heating of the cathode.

Note that the "C battery" is a designation having no relation to the 1.5 volt "C cell" (nor for the A and B batteries, discussed above).

AC power

Replacement of batteries was a major cost of operation for early radio receiver users. The development of the battery eliminator, and, in 1925, batteryless receivers operated by household power, reduced operating costs and contributed to the growing popularity of radio. A power supply using a transformer with several windings, one or more rectifiers (which may themselves be vacuum tubes), and large filter capacitors provided the required direct current voltages from the alternating current source.

As a cost reduction measure, especially in high-volume consumer receivers, all the tube heaters could be connected in series across the AC supply, and the plate voltage derived from a half-wave rectifier directly connected to the AC input, eliminating the need for a heavy power transformer. As an additional feature, these radios could be operated on AC or DC mains. This arrangement resulted in a limited plate voltage, however with advances in tube technology tubes could run reasonably effectively with only 150 volts on their plates. A filament tap on the rectifier tube provided the 6 volt, low current supply needed for a dial light.

Rectifying the AC mains directly did have one safety issue: the chassis of the receiver was connected to one side of the mains, presenting a shock hazard. This hazard was reduced by enclosing the chassis in an insulated case and running the AC power through a so-called interlock connection at the removable back side of the receiver. This would come disconnected whenever the radio was opened (for instance, to test and replace the tubes) preventing such a shock hazard. (Technicians and tinkerers routinely bypassed this by using a separate cord, known colloquially as a "cheater cord" or "widowmaker.") Many consumer AM radio manufacturers of the era used a virtually identical circuit with the tube complement of 12BA6, 12BE6, 12AV6, 35W4, and 50C5, giving these radios the nickname All American Five or simply "Five Tube Radio." Although millions of such receivers were produced, they have now become collector's items.

Reliability

Tube tester manufactured in 1930

One reliability problem of tubes with oxide cathodes is the possibility that the cathode may slowly become "poisoned" by gas molecules from other elements in the tube, which reduce its ability to emit electrons. Trapped gases or slow gas leaks can also damage the cathode or cause plate-current run away due to ionization of free gas molecules. Vacuum hardness and proper selection of construction materials are the major influences on tube lifetime. Depending on the material, temperature and construction, the surface material of the cathode may also diffuse onto other elements. The resistive heaters that heat the cathodes may break in a manner similar to incandescent lamp filaments, but rarely do, since they operate at much lower temperatures than lamps.

The heater's failure mode is typically a stress-related fracture of the tungsten wire or at a weld point and generally occurs after accruing many thermal (power on-off) cycles. Tungsten wire has a very low resistance when at room temperature. A negative temperature coefficient device, such as a thermistor, may be incorporated in the equipment's heater supply or a ramp-up circuit may be employed to allow the heater or filaments to reach operating temperature more gradually than if powered-up in a step-function. Low-cost radios had tubes with heaters connected in series, with a total voltage equal to that of the line (mains). Following World War II, tubes intended to be used in series heater strings were redesigned to all have the same ("controlled") warm-up time. Earlier designs had quite-different thermal time constants. The audio output stage, for instance, had a larger cathode, and warmed up more slowly than lower-powered tubes. The result was that heaters that warmed up faster also temporarily had higher resistance, because of their positive temperature coefficient. This disproportionate resistance caused them to temporarily operate with heater voltages well above their ratings, and shortened their life.

Another important reliability problem is caused by air leakage into the tube. Usually oxygen in the air reacts chemically with the hot filament or cathode, quickly ruining it. Designers developed tube designs that sealed reliably. This was why most tubes were constructed of glass. Metal alloys (such as Cunife and Fernico) and glasses had been developed for light bulbs that expanded and contracted in similar amounts, as temperature changed. These made it easy to construct an insulating envelope of glass, while passing connection wires through the glass to the electrodes.

When a vacuum tube is overloaded or operated past its design dissipation, its anode (plate) may glow red. In consumer equipment, a glowing plate is universally a sign of an overloaded tube. However, some large transmitting tubes are designed to operate with their anodes at red, orange, or in rare cases, white heat.

Vacuum
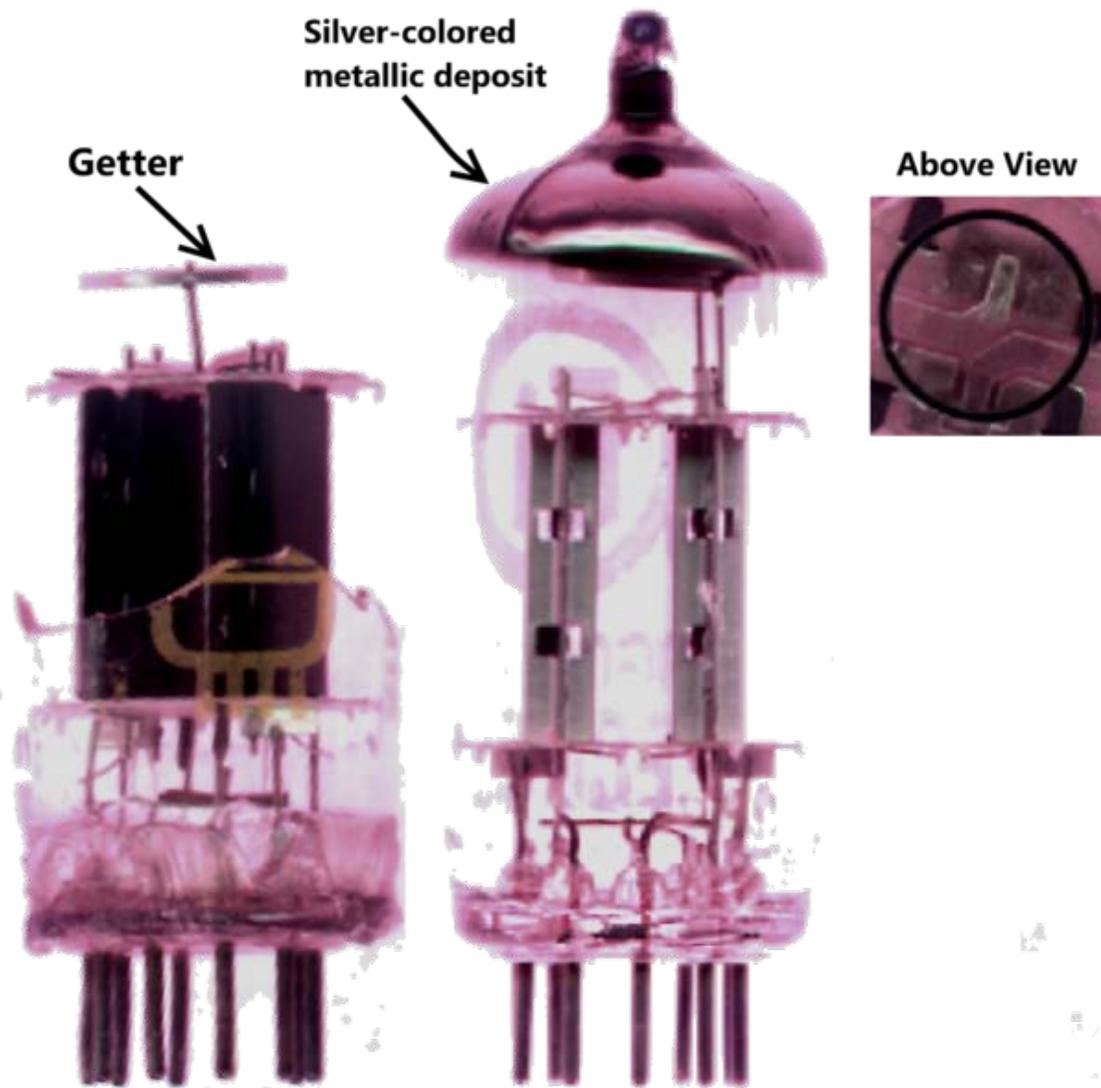
Getter in opened tube; silvery deposit from getter

The highest possible vacuum is desired in a tube. Remaining gas atoms will ionize and conduct electricity between the elements in an undesired manner. In a defective tube residual air pressure will lead to ionization, becoming visible as a pink-purple glow discharge between the tube elements.

To prevent gases from compromising the tube's vacuum, modern tubes are constructed with "getters", which are usually small, circular troughs filled with metals that oxidize quickly, barium being the most common. While the tube envelope is being evacuated, the internal parts except the getter are heated by RF induction heating to help free any remaining gases from the metal parts. The tube is then sealed and the getter is heated to a high temperature, again by radio frequency induction heating. This causes some material from the getter to evaporate, reacting with any residual gases and usually leaving a silver-

colored metallic deposit on the inside of the envelope of the tube. The getter continues to absorb small amounts of gas that may leak into the tube during its working life. If a tube develops a serious leak in the envelope, this deposit turns a white color as it reacts with atmospheric oxygen. Large transmitting and specialized tubes often use more exotic getter materials, such as zirconium. Early gettered tubes used phosphorus based getters and these tubes are easily identifiable, as the phosphorus leaves a characteristic orange or rainbow deposit on the glass. The use of phosphorus was short-lived and was quickly replaced by the superior barium getters. Unlike the barium getters, the phosphorus did not absorb any further gases once it had fired.

Transmitting tubes

Large transmitting tubes have carbonized tungsten filaments containing a small trace (1% to 2%) of thorium. An extremely thin (molecular) layer of thorium atoms forms on the outside of the wire's carbonized layer and, when heated, serve as an efficient source of electrons. The thorium slowly evaporates from the wire surface, while new thorium atoms diffuse to the surface to replace them. Such thoriated tungsten cathodes usually deliver lifetimes in the tens of thousands of hours. The end-of-life scenario for a thoriated-tungsten filament is when the carbonized layer has mostly been converted back into another form of tungsten carbide and emission begins to drop off rapidly; a complete loss of Thorium has never been found to be a factor in the end-of-life in a tube with this type of emitter. The highest reported tube life is held by an Eimac power tetrode used in a Los Angeles radio station's transmitter, which was removed from service after 80,000 hours (~9 years) of operation. It has been said that transmitters with vacuum tubes are better able to survive lightning strikes than transistor transmitters do. While it was commonly believed that at rf power levels above approx. 20 kilowatts, vacuum tubes were more efficient than solid state circuits, this is no longer the case especially in medium wave (AM broadcast) service where solid state transmitters at nearly all power levels have measurably higher efficiency. FM broadcast transmitters with solid state power amplifiers up to approx. 15 kW also show better overall mains-power efficiency than tube-based power amplifiers.

Receiving tubes

Cathodes in small "receiving" tubes are coated with a mixture of barium oxide and strontium oxide, sometimes with addition of calcium oxide or aluminium oxide. An electric heater is inserted into the cathode sleeve, and insulated from it electrically by a coating of aluminium oxide. This complex construction causes barium and strontium atoms to diffuse to the surface of the cathode when heated to about 780 degrees Celsius, thus emitting electrons.

Failure modes

Catastrophic failures

A catastrophic failure is one which suddenly makes the vacuum tube unusable. A crack in the glass envelope will allow air into the tube and destroy it. Cracks may result from stress in the glass, bent pins or impacts; tube sockets must allow for thermal expansion, to prevent stress in the glass at the pins. Stress may accumulate if a metal shield or other object presses on the tube envelope and causes differential heating of the glass. Glass may also be damaged by high-voltage arcing.

Tube heaters may also fail without warning, especially if exposed to over voltage or as a result of manufacturing defects. Tube heaters do not normally fail by evaporation like lamp filaments, since they operate at much lower temperature. The surge of inrush current when the heater is first energized causes stress in the heater, and can be avoided by slowly warming the heaters, gradually increasing current. Some tubes intended for series string operation of the heaters across the supply will have a definite controlled warm-up time to avoid excess voltage on some heaters as others warm up. Directly-heated filament-type cathodes as used in battery-operated tubes or some rectifiers may fail if the filament sags, causing internal arcing. Excess heater-to-cathode voltage in indirectly heated cathodes can break down the insulation between elements and destroy the heater.

Arcing between tube elements can destroy the tube. An arc can be caused by applying plate potential before the cathode has come up to operating temperature, or by drawing excess current through a rectifier which damages the emission coating. Arcs can also be initiated by any loose material inside the tube, or by excess screen voltage. An arc inside the tube allows gas to evolve from the tube materials, and may deposit conductive material on internal insulating spacers.

Degenerative failures

Degenerative failures cause the performance of the tube to slowly deteriorate with time.

Overheating of internal parts, such as control grids or mica spacer insulators, can result in trapped gas escaping into the tube; this can reduce performance. A getter is used to absorb gases evolved during tube operation, but has only a limited ability to combine with gas. Control of the envelope temperature prevents some types of gassing. A tube with very bad internal gas may have a visible blue glow when plate voltage is applied.

Gas and ions within the tube contribute to grid current which can disturb operation of a vacuum tube circuit. Another effect of overheating is the slow deposit of metallic vapors on internal spacers, resulting in inter-element leakage.

Tubes on standby for long periods, with heater voltage applied, may develop high cathode interface resistance and display poor emission characteristics. This effect

occurred especially in pulse and digital circuits, where tubes had no plate current flowing for extended times.

Cathode depletion describes the loss of emission after thousands of hours of normal use. Sometimes emission can be restored for a time by raising heater voltage either for a short time or a permanent increase of a few percent. Cathode depletion was uncommon in signal tubes but was a frequent cause of failures of monochrome television cathode-ray tubes.

Other failures

Vacuum tubes may have or develop defects in operation that makes an individual tube useless in one device, but which may not prevent its satisfactory operation in another system. *Microphonics* refers to internal vibration of tube elements, which modulates the signal from the tube in an undesirable way; sound or vibration pick-up may affect the signals, or even cause uncontrolled howling if a feedback path develops between a microphonic tube and, for example, a loudspeaker. Leakage current between AC heaters and the cathode may couple into the circuit, or electrons emitted directly from the ends of the heater may also inject hum into the signal. Leakage current due to internal contamination may also inject noise.

Cooling

Like any electronic device, vacuum tubes produce heat while operating. This waste heat is one of the principal factors that affect tube life. In power amplifiers, the majority of this waste heat originates in the anode though screen grids may also require cooling. For example, the screen grid in an EL34 is cooled by two small radiators or "wings" near the top of the tube. A tube's heater (filament) also contributes to the total waste heat. A tube's data sheet will normally identify the maximum amount of heat that each element may safely dissipate.

The method of anode cooling is dependent on the construction of the tube itself. Tubes used in consumer equipment have internal anodes, so cooling occurs through black body radiation from the anode to the glass envelope. Natural convection (air circulation) then removes the heat from the envelope. Tube shields that aided heat dispersal could be retrofitted on certain select types of tubes. These shields act by improving heat conduction from the surface of the tube to the shield itself by means of tens of copper tongues in contact with the glass tube, and have an opaque, black outside finish for improved heat radiation. The ability to remove heat may be further increased by implementing forced air cooling, and adding an external heat sink attached to the anode through the tube's enclosure. These measures are both implemented in the 4-1000A transmitting tube whose anode was designed to safely operate at red hot temperatures, dissipating up to one kilowatt.

The amount of heat that may be removed from a tube with an internal anode is limited. Tubes with external anodes may be cooled using forced air, water, vapor, and multiphase.

The 3CX10,000A7 is an example of a tube with an external anode cooled by forced air. The water, vapor, and multiphase cooling techniques all depend on the high specific heat and latent heat of water. The 8974 is an example of a water cooled tube and is among the largest commercial tubes available today.

In a water cooled tube, the anode voltage appears directly on the cooling water surface, thus requiring the water to be an electrical insulator. Otherwise the high voltage can be conducted through the cooling water to the radiator system; hence the need for deionized water. Such systems usually have a built-in water-conductance monitor which will shut down the high-tension supply (often tens of kilovolts) if the conductance becomes too high.

Other vacuum tube devices

Many devices were built during the 1920–1960 period using vacuum-tube techniques. Most such tubes were rendered obsolete by semiconductors; some techniques for integrating multiple devices in a single module, sharing the same glass envelope have been discussed above, such as the Loewe 3NF. Vacuum-tube electronic devices still in common use include the magnetron, klystron, photomultiplier, x-ray tube, traveling-wave tube and cathode ray tube. The magnetron is the type of tube used in all microwave ovens. In spite of the advancing state of the art in power semiconductor technology, the vacuum tube still has reliability and cost advantages for high-frequency RF power generation. Photomultipliers are still the most sensitive detectors of light.

The cathode ray tube (CRT) is a vacuum tube used particularly for display purposes. Many televisions, oscilloscopes and computer monitors still use cathode ray tubes, though flat panel displays are becoming more predominant as prices drop. At one time many radios used "magic eye" tubes, a specialized sort of CRT used in place of a meter movement to indicate signal strength, or input level in a tape recorder. A modern indicator device, the vacuum fluorescent display (VFD) is also a sort of cathode ray tube.

Secondary emission is the term for what happens when electrons in a vacuum strike certain materials, and the impacts cause electrons to be emitted. For some materials, more electrons are emitted than originally hit the surface. Such devices, called electron multipliers, amplify the current represented by the incoming electrons. Several stages (as many as 15 or so) can be cascaded for high gain, and are essential parts of very sensitive phototubes, usually called photomultipliers or multiplier photoubes. The image orthicon TV studio camera tubes also used multistage photomultipliers.

For decades, electron-tube designers tried to use secondary emission to obtain more amplification in vacuum tubes with hot cathodes, but they suffered from short life because the material used for the secondary-emission electrode (called a dynode) "poisoned" the tube's hot cathode. (For instance, the interesting RCA 1630 secondary-emission tube was marketed, but did not last.) However, eventually, Philips of The Netherlands developed the EFP60 tube that had a satisfactory lifetime, and was used in at

least one product, a laboratory pulse generator. However, transistors were rapidly improving, and eclipsed tubes in general.

A variant, called a channel electron multiplier, is a curved tube, such as a helix, coated on the inside with material with good secondary emission. One type had a little funnel to capture incoming electrons. The tube was resistive, and its ends were connected to enough voltage to create repeated cascades of electrons.

Tektronix made a high-performance wideband oscilloscope CRT with a channel electron multiplier plate behind the phosphor layer. This plate was a bundled array of a huge number of short individual c.e.m. tubes that accepted a low-current beam and intensified it to provide a display of practical brightness. (The electron optics of the wideband electron gun could not provide enough current to directly excite the phosphor.)

Some tubes, such as magnetrons, traveling-wave tubes, carcinotrons, and klystrons, combine magnetic and electrostatic effects. These are efficient (usually narrow-band) RF producers and still find use in radar, microwave ovens and industrial heating. Traveling-wave tubes (TWTs) are very good amplifiers; they are used in some communications satellites. High-powered klystron amplifier tubes can provide hundreds of kilowatts in the UHF range.

Gyrotrons or vacuum masers, used to generate high-power millimetre band waves, are magnetic vacuum tubes in which a small relativistic effect, due to the high voltage, is used for bunching the electrons. Gyrotrons can generate very high powers (hundreds of kilowatts). Free electron lasers, used to generate high-power coherent light and perhaps even X rays, are highly relativistic vacuum tubes driven by high-energy particle accelerators.

Particle accelerators can be considered vacuum tubes that work backward, the electric fields driving the electrons, or other charged particles. In this respect, a cathode ray tube is a particle accelerator.

A tube in which electrons move through a vacuum (or gaseous medium) within a gas-tight envelope is generically called an *electron tube*.

Some condenser microphone designs use built-in vacuum tube preamplifiers.

Vacuum tubes in the 21st century

Niche applications

Although vacuum tubes have been largely replaced by solid-state devices in most amplifying applications, there are certain exceptions. In addition to the special functions noted above, tubes have some niche applications even in the current age.

Vacuum tubes are much less susceptible than corresponding solid-state components to the electromagnetic pulse effect of nuclear explosions. This property kept them in use for certain military applications long after transistors had replaced them elsewhere. Vacuum tubes are still used for very high-powered applications such as industrial radio-frequency heating, generating large amounts of RF energy for particle accelerators, and power amplification for broadcasting. Several sorts of tubes are used in microwave applications, such as the household microwave oven in which the magnetron is used to efficiently generate microwave powers of several hundred watts.

70 watt tube audio amplifier selling for $2,680 in 2011. Certain audiophiles happily pay ten times more for the perceived "tube sound."

Many audiophiles, professional audio engineers, and musicians prefer the so-called tube sound compared with amplifiers using transistors. There are companies which still make high end audio amplifiers using tube technology to serve this market. One can also purchase microphone preamplifiers using tubes; these are preferred by professional music recording studios.

Many electric guitar amplifiers sold are still based on tubes. The sound produced by a tube power amplifier when overdriven has defined the texture of some genres of music such as classic rock and blues. Rather than the hard clipping characteristic of solid state power amplifiers, a tube amplifier in conjunction with its output transformer (solid state

amplifiers have direct coupled outputs) produces a more gradual distortion and distinctive sound. Guitarists often cite the sound of tube amplifiers for the warmth of their tone and the natural compression they apply to an input signal in this regime. Although the reliability of solid state amplifiers has greatly improved, tube amplifiers have the advantage that the output tubes can be replaced by the owner, whereas "blown" output transistors require attention by a qualified technician.

Vacuum fluorescent display
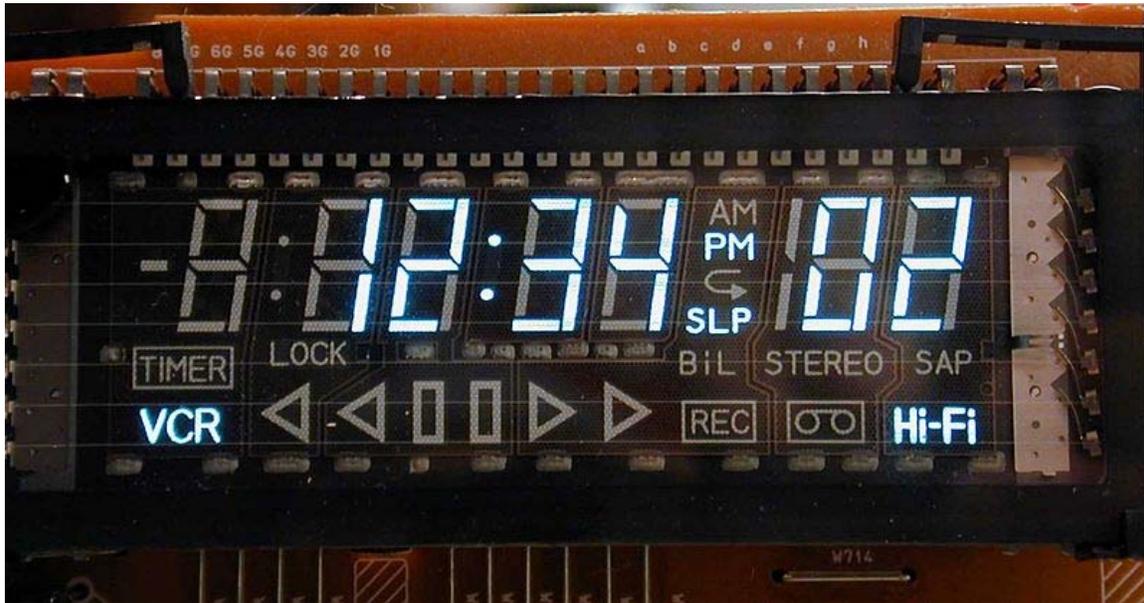
Typical VFD used in a videocassette recorder

A modern display technology using a sort of cathode ray tube is freqeuntly used in videocassette recorders, microwave oven control panels, and automotive dashboards. Rather than raster scanning these vacuum fluorescent displays (VFD) switch control grids and plate voltages on and off to display discrete characters, for instance. This term should not be confused with fluorescent light technology which uses fluorescence from ultraviolet radiation produced using a discharge tube. Rather the VFD uses phosphor-coated anodes as in other display cathode ray tube. Because the filaments are in view, they must be operated at temperatures where the filament does not glow visibly. This is possible using more recent cathode technology, and these tubes also operate with quite low anode voltages (often less than 50 volts) contrary to classical cathode ray tubes. Often found in automotive applications, their high brightness allows reading the display in daylight. VFD tubes are flat and rectangular, as well as relatively thin. Typical VFD phosphors emit a broad-spectrum greenish-white light, permitting use of color filters. This type of phosphor provides a bright glow despite the low energy of the incident electrons.

Vacuum tubes using field electron emitters

In the early years of the 21st century there has been renewed interest in vacuum tubes, this time with the electron emitter formed on a flat silicon substrate, as in integrated circuit technology. This subject is now called vacuum nanoelectronics. The most common design uses a cold cathode in the form of a large-area field electron source (for example a field emitter array). With these devices, electrons are field-emitted from a large number of closely spaced individual emission sites.

Their claimed advantages include greatly enhanced robustness combined with the ability to provide high power outputs at low power consumptions. Operating on the same principles as traditional tubes, prototype device cathodes have been fabricated in several different ways. Although a common approach is to use a field emitter array, one interesting idea is to etch electrodes to form hinged flaps – similar to the technology used to create the microscopic mirrors used in Digital Light Processing) that are stood upright by an electrostatic charge.

Such integrated microtubes may find application in microwave devices including mobile phones, for Bluetooth and Wi-Fi transmission, in radar and for satellite communication. Presently they are being studied for possible applications in field emission display technology, but significant production problems seem to exist.

Modern manufacturers

Vacuum tubes are still being manufactured in the following countries:

China

Manufacturer	Area of expertise
Shuguang Electron Group Co.	Tubes primarily for audio applications.
Tianjin Quanerzhen Electron Tube Technology Co.	Direct heated tubes for audio applications.
Nanjing Sanle Electronic Information Industry Group Co.	Transmitting and industrial tubes. (Including Chinese 3-500C & 4-400C)
JiangXi Jingguang Electronics Co.	Transmitting and industrial tubes.
Huaguang Electric Power & Electronics Co.	Transmitting and industrial tubes
Chengdu Xuguang Electronics Co.	Transmitting and industrial tubes

Russia

Manufacturer	Area of expertise
Ekspopul JSC	Audio tube factory of New Sensor Inc. Known formerly as tube factory of Reflektor JSC.
"Ryazan" Vacuum Components LLC	Direct heated tubes. SV811 and SV572 series for audio, 811A and 572B for RF applications.
"SED-SPb" Svetlana Electron Devices - St. Petersburg. Svetlana JSC	Primarily transmitting and industrial tubes. Manufactures also few models for audio applications.
Voskhod KRLZ JSC	Tubes for small signal RF and audio applications
NEVZ-Soyuz HC JSC	Transmitting and industrial tubes, known formerly as Novosibirsk electron tube plant.

United States

Manufacturer	Area of expertise
Communications & Power Industries Inc.	Transmitting and industrial tubes, formerly known as Eitel-McCullough Inc.
Burle Industries Inc.	Transmitting and industrial tubes, formerly factory of RCA
MPD Components Inc.	Planar triodes and magnetrons, formerly Ken-Rad and later GE tube factory
MU Incorporated	Contract manufacturer.
Western Electric Export Corporation	300B Tubes Former Factory of AT&T's Western Electric moved equipment from Missouri to Tennessee in 2002

United Kingdom

Manufacturer	Area of expertise
e2v Technologies Ltd.	Transmitting and industrial tubes, formerly known as English Electric Valve Co. Ltd.
TMD Technologies Ltd.	Transmitting and industrial tubes. Formerly THORN Microwave Devices Ltd.

France

Manufacturer	Area of expertise
	Transmitting and industrial tubes. Products marketed by Richardson Electronics.
Covimag SA	Formerly Philips transmitting tube factory.

Thales Electron Devices SA	Transmitting and industrial tubes. Formerly known as Thomson-CSF.

Czech Republic

Manufacturer	Area of expertise
Emission Labs	Direct heated tubes for audio applications
KR Audio Electronics s.r.o.	Direct heated tubes for audio applications
Tesla Electrontubes s.r.o	Transmitting and industrial tubes

Slovakia

Manufacturer	Area of expertise
JJ-Electronic	Tubes primarily for audio applications, factory was formerly part of Tesla Electrontubes
Euro Audio Team	Tubes for high-end audio.

Chapter 2

Valve Amplifier

The glow from four power tubes lights up the inside of a guitar amplifier

A **valve amplifier** or **tube amplifier** is a type of electronic amplifier that makes use of vacuum tubes to increase the power and/or amplitude of a signal. Low to medium power valve amplifiers for frequencies below the microwaves were largely replaced by solid state amplifiers during the 1960s and 1970s. Valve amplifiers are used for applications such as guitar amplifiers, satellite transponders such as DirecTV and GPS, audiophile

stereo amplifiers, military applications (such as radar) and very high power radio and UHF television transmitters.

History

Origins

Until the invention of the transistor in 1947, all practical amplifiers were made using thermionic valves. The simplest valve was invented by John Ambrose Fleming while working for the Marconi Company in London in 1904 and named the diode, as it had two electrodes. The diode conducted electricity in one direction only and was used as a radio detector and a rectifier.

In 1906 Lee De Forest added a third electrode and invented the first electronic amplifying device, the triode, which he named the *Audion*. This additional *control grid* modulates the current that flows between cathode and anode. The relationship between current flow and plate and grid voltage is often represented as a series of "characteristic curves" on a diagram. Depending on the other components in the circuit this modulated current flow can be used to provide current or voltage gain.

The first application of valve amplification was in the regeneration of long distance telephony signals. Later, valve amplification was applied to the 'wireless' market that began in the early thirties. In due course amplifiers for music and later television were also built using valves.
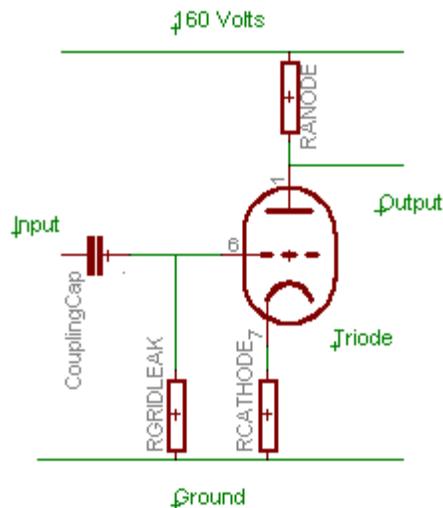
Circuit diagram of a single-ended triode

The overwhelmingly dominant circuit topology during this period was the single-ended triode gain stage, operating in class A, which gave very good sound (and reasonable measured distortion performance) despite extremely simple circuitry with very few components: important at a time when components were hand made and extremely expensive. Before World War II, almost all valve amplifiers were of low gain and with linearity dependent entirely on the inherent linearity of the valve itself, typically 5% distortion at full power.

Negative feedback (NFB) was invented by Harold Stephen Black in 1927, but initially little used since at that time gain was at a premium. This technique allows amplifiers to trade gain for reduced distortion levels (and also gave other benefits such as reduced output impedance). The introduction of the Williamson amplifier in 1947, which was extremely advanced in many respects including very successful use of NFB, was a turning point in audio power amplifier design, operating a push-pull output circuit in class AB1 to give performance surpassing its contemporaries.

Post-war developments

World War II stimulated dramatic technical progress and industrial scale production economies. Increasing affluence after the war brought about for the first time a substantial and expanding consumer market. This enabled more advanced valve designs to be marketed at affordable prices, with the result that the 1960s saw the increasing spread of electronic gramophone players, and ultimately the beginnings of "high fidelity". Hifi was able to drive full frequency range loudspeakers (for the first time often with multiple drivers for different frequency bands) to significant volume levels. This combined with the spread of TV, produced a 'golden age' in valve development and also in the development of the design of valve amplifier circuits.

A range of topologies with only minor variations (notably different phase splitter arrangements and the "Ultra-Linear" transformer connection for tetrodes) rapidly became widespread. This family of designs remains the dominant high power amplifier topology to this day for music application. This period also saw continued growth in civilian radio, with valves being used for both transmitters and receivers.

Decline

From the 1970s the silicon transistor became increasingly pervasive. Valve production was sharply decreased, with the notable exception of cathode ray tubes (CRTs), and a reduced range of valves for amplifier applications. Popular low power tubes were dual triodes (ECCnn, 12Ax7 series) plus the EF86 pentode, and power valves were mostly being beam tetrode and pentodes (EL84, EL34, KT88 / 6550, 6L6), in both cases with indirect heating. This reduced set of types remains the core of valve production today.

The Soviets retained valves to a much greater extent than the West during the Cold War, for the majority of their communications and military amplification requirements, in part

due to valves' ability to withstand instantaneous overloads (notably due to a nuclear detonation) that would destroy a transistor .

The dramatic reduction in size, power consumption, reduced distortion levels and above all cost of electronics products based on transistors has made valves obsolete for mainstream products since the 1970s. Valves remained in certain applications such as high power RF transmitters and the microwave oven.

Resurgence

A single-ended class 'A' guitar amplifier chassis, with additional GZ34 valve rectifier installed.

In audio applications, valves continue to be highly desired by some users, both in the higher-end home audio market and in the guitar amplifier market. Amongst stereo enthusiasts, there is a subgroup of audio buffs who advocate the use of tube amplifiers for home listening; they argue that tube amplifiers produce a "warmer" or more "natural" valve sound. Companies in Russia, China and Eastern Europe continue to produce valves to cater to this market.

In the guitar amplifier market, most performers continue to use tube amps today, including in folk, blues, roots rock, and in harder genres such as metal, where tube amps are used to create heavy distortion. Audio engineers suggest that the subjectively pleasing aspects of tube amplification may be due to the non-linear overdrive that is produced with tubes.

Tube amplifiers respond differently from transistor amplifiers when signal levels approach and reach the point of clipping. In a tube amplifier, the transition from linear

amplification to limiting is less abrupt than in a solid state unit, resulting in a less grating form of distortion at the onset of clipping. For this reason, some guitarists prefer the sound of an all-tube amplifier; the aesthetic properties of tube versus solid state amps, though, are a topic of debate in the guitarist community.

Characteristics

Valves are high voltage and low current devices in contrast to transistors, which typically operate at lower voltages and higher currents for the same power level. The high working voltage makes valves well suited for radio transmitters, for example, and valves remain in use today for very high power radio transmitters, where there is still no other technology available. However, for most applications requiring an appreciable output current, a matching transformer is required. The transformer is a critical component and heavily influences the performance (and cost) of the amplifier.

Many power valves have good linearity but modest gain or transconductance. Signal amplifiers using tubes are capable of very high frequency response ranges – up to radio frequency. Indeed, many of the directly heated single-ended triode (DH-SET) audio amplifiers are in fact radio transmitting tubes designed to operate in the megahertz range. In practice, however, tube amplifier designs typically "couple" stages either capacitively, limiting bandwidth at the low end, or inductively with transformers, limiting the bandwidth at high end.

Advantages

- Very linear (especially triodes) making it viable to use them in low distortion linear circuits with little or no negative feedback.
- Inherently suitable for high voltage circuits.
- Can be constructed on a scale that can dissipate large amounts of heat (some extreme devices even being water cooled). For this reason valves remained the only viable technology for very high power applications such as radio and TV transmitters long into the age when transistors had displaced valves in most other applications.
- Very low "drift" (of specifications) over a wide range of operating conditions, specifically high heat and high power. Semiconductors are very heat sensitive by comparison, forcing compromises in solid state amplifier designs.
- Electrically very robust, they can tolerate overloads for minutes which would destroy bipolar transistor systems in milliseconds.
- Easily removable for testing or replacement – it is much harder to replace a faulty transistor.
- Softer clipping when overloading the circuit, which many audiophiles and musicians think gives a more pleasant sound.

Disadvantages

- Heater supplies are usually required for the cathodes, and high voltages are usually required for the anodes.
- They are significantly larger than equivalent solid-state devices
- High impedance and low current output is unsuitable for direct drive of many real world loads, notably various forms of electric motors.
- Valves may have a shorter working life than solid state parts due to various failure mechanisms (such as heat, cathode poisoning, breakage, or internal short-circuits).
- Available in a single polarity only whereas transistors are available in complementary polarities (e.g., NPN/PNP), making possible many circuit configurations that cannot be realized directly with valves.
- Valve circuits must avoid introduction of noise from ac heater supplies.
- Microphonics – valves may sometimes be sensitive to sound or vibration, inadvertently acting like a microphone.
- Power consumption due to the heater requirements.

Operation

All amplifier circuits are classified by "class of operation" as A, B, AB and C etc. Some significantly different circuit topologies exist compared to transistor designs.

- The grid (where the input signal is presented) needs to be biased substantially negative with respect to the cathode This makes it extremely difficult to direct couple the output of one valve to the input of a following valve as is normally done in transistor designs.

- Valve stages are coupled with components rated to withstand several hundred volts, typically a capacitor, occasionally a coupling transformer. The phase shifts introduced by coupling networks can become problematic in circuits that have feedback.

- There is no valve analog of the complementary devices widely used in "totem pole" output stages of silicon circuits. Push-pull valve topologies therefore typically require a transformer.

- The very high output impedance of valves (compared with transistors) usually requires matching transformers to drive low impedance loads such as loudspeakers or cutting lathe heads. The transformer is used as the load, in place of the resistor usually used in small-signal and driver stages. The impedance of the transformer primary at the frequencies in use is much higher than the DC resistance of the windings, often kilohms. High performance transformers are however severe engineering compromises, are expensive, and in operation are far from ideal. Transformers dramatically increase the cost of a valve amplifier circuit compared to a direct-coupled transistor alternative.

- The open loop linearity of valves, especially triodes, makes it possible to use little or no negative feedback in circuits whilst retaining acceptable or even excellent distortion performance (especially for small-signal circuits).

Topologies

- Linear small signal circuits almost invariably use a triode in the single ended gain stage topology (in class A), including the output stage.
- Broadband valve amplifiers typically use class A1 or AB1.
- Modern high power output stages are usually push pull, often necessitating some form of phase splitter to derive a differential/balanced drive signal from a single ended input, typically followed by a further gain stage (the "driver") prior to the output tubes. For example SRPP)
- single ended" power stages using very large valves exist and dominate in radio transmitter applications. A sidebar is the observation that the niche "DH-SET" topology favored by some audiophiles are extremely simple and typically constructed using valve types originally designed for use in radio transmitters
- more complex topologies (notably the use of active loads) can improve linearity and frequency response (by removing Miller capacitance effects).

Output impedance

The high output impedance of tube plate circuits is not well matched to low-impedance loads such as loudspeakers or antennas. A matching network is required for efficient power transfer; this may be a transformer at audio frequencies, or various tuned networks at radio frequencies.

Applications

Audio Frequency (AF) and Broadband amplifiers

Valves remain in widespread use in guitar and high-end audio amplifiers due to the sound quality they produce, which is subjectively preferred by some users. They are largely obsolete for most other applications, mainly due to the cost effectiveness advantages of the transistor, and due to the lower weight and heat production of transistor amps.

Telephony

Telephony was the original, and for many years was a driving application for audio amplification. A specific issue for the telecommunication industry was the technique of multiplexing many (up to a thousand) voice lines onto a single cable, at different frequencies.

The advantage of this is that a single valve "repeater" amplifier can amplify many calls at once, this being very cost effective. The problem is that the amplifiers need to be extremely linear, otherwise "intermodulation distortion" (IMD) will result in "crosstalk"

between the multiplexed channels. This stimulated development emphasis towards low distortion far beyond the nominal needs of a single voice channel.

Audio

Today the main application for valves is audio amplifiers for high-end hi-fi and musical performance use with electric guitars, electric basses, and Hammond organs, although these applications have different requirements regarding distortion which result in different design compromises, although the same basic design techniques are generic and widely applicable to all broadband amplification applications, not only audio.

Post World War II, the majority of valve power amplifiers are of the Class AB-1 "push pull" ultralinear topology, but niche products using the DH-SET and even OTL topologies still exist in small numbers.

A pre-amplifier design using all power tubes instead of small signal tubes

A power amplifier design with 70Wrms/ch output

Instrumentation amplifiers

The basic moving coil voltmeter and ammeter itself takes a small current and thus loads the circuit to which it is attached. This can significantly alter the operating conditions in the circuit being measured. The vacuum tube voltmeter (VTVM) uses the high input impedance of a valve to buffer the circuit being measured from the load of the ammeter.

Valve oscilloscopes share this very high input impedance and thus can be used to measure voltages even in very high impedance circuits. There may typically be 3 or 4 stages of amplification per display channel. In later oscilloscopes, a type of amplifier using a series of tubes connected at equal distances along transmission lines, known as a distributed amplifier was employed to amplify very high frequency vertical signals before application to the display tube. Valve oscilloscopes are now obsolete.

In the closing years of the valve era, valves were even used to make "operational amplifiers" – the building blocks of much modern linear electronics. An op-amp typically has a differential input stage and a totem pole output, the circuit usually having a minimum of five active devices. A number of "packages" were produced that integrated such circuits (typically using two or more glass envelopes) into a single module that could be plugged into a larger circuit (such as an analog computer). Such valve op-amps

were very far from ideal and quickly became obsolete, being replaced with solid-state types.

Narrow band and radio frequency tuned amplifiers

Historically (pre-WWII) "transmitting tubes" were among the most powerful tubes available. These usually had directly heated thoriated filaments cathodes that glowed like light bulbs. Some tubes were capable of being driven so hard that the anode would itself glow cherry red; the anodes were machined from solid material (rather than fabricated from thin sheet) to withstand heat without distorting. Notable tubes of this type are the 845 and 211. Later tetrodes and pentodes such as 817 and (direct heated) 813 were also used in large numbers in (especially military) radio transmitters

RF circuits are significantly different from broadband amplifier circuits. The antenna or following circuit stage typically contains one or more adjustable capacitive or inductive component allowing the resonance of the stage to be accurately matched with carrier frequency in use, to optimize power transfer from and loading on the valve, a so called "tuned circuit".

Broadband circuits require flat response over a wide range of frequencies. RF circuits by contrast are typically required to operate at high frequencies but often over a very narrow frequency range. For example, an RF device might be required to operate over the range 144 to 146 MHz (just 1.4%)

Today, radio transmitters are overwhelmingly silicon based, even at microwave frequencies. However an ever decreasing minority of high power radio frequency amplifiers continue to have valve construction.

Chapter 3

Gas-Filled Tube

A **gas-filled tube**, also known as a **discharge tube**, is an arrangement of electrodes in a gas within an insulating, temperature-resistant envelope. Although the envelope is typically glass, power tubes often use ceramics, and military tubes often use glass-lined metal. Gas tubes are tubes whose electrical characteristics are substantially influenced by the pressure and composition of gas contained inside.

Gas-filled tubes exploit phenomena related to electric discharge in gases, operating by ionizing the gas with applied voltage to start electrical conduction. Both hot cathode and cold cathode type devices are encountered. Depending on application, either the glow from the gas or the electric arc or electric glow discharge may be the desired function.

Gases in use

Hydrogen

Hydrogen is used in tubes used for very fast switching, e.g. some thyratrons, dekatrons, and krytrons, where very steep edges are required. The build-up and recovery times of hydrogen are much shorter than in other gases. Hydrogen thyratrons are usually hot-cathode. Hydrogen (and deuterium) can be stored in the tube in the form of a metal hydride, heated with an auxiliary filament; releasing hydrogen by heating such storage element can be used to replenish cleaned-up gas, and even to adjust the pressure as needed for a thyratron operation at a given voltage.

Deuterium

Deuterium is used in ultraviolet lamps for ultraviolet spectroscopy, in neutron generator tubes, and in special tubes (e.g. crossatron). It has higher breakdown voltage than hydrogen. In fast switching tubes it is used instead of hydrogen where high voltage operation is required. For a comparison, the hydrogen-filled CX1140 thyratron has anode

voltage rating of 25 kV, while the deuterium-filled and otherwise identical CX1159 has 33 kV. Also, at the same voltage the pressure of deuterium can be higher than of hydrogen, allowing higher rise rates of rise of current before it causes excessive anode dissipation. Significantly higher peak powers are achievable. Its recovery time is however about 40% slower than for hydrogen.

Noble gases
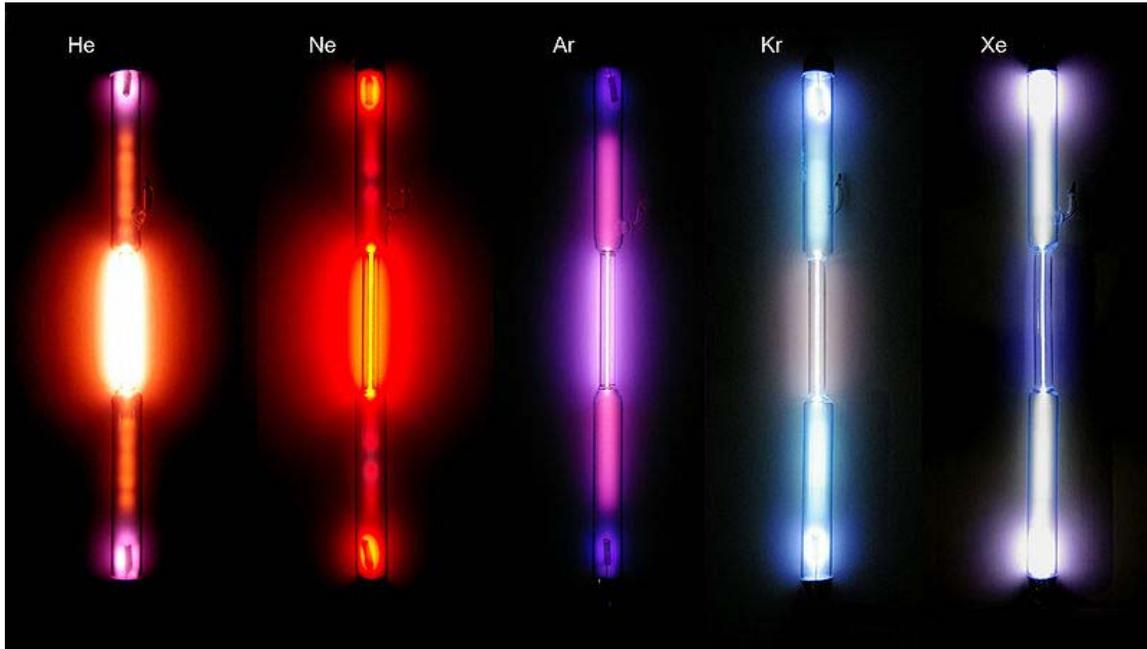
Noble gas discharge tubes; from left to right: helium, neon, argon, krypton, xenon

Noble gases are frequently used in tubes for many purposes, from lighting to switching. Pure noble gases are employed in switching tubes. Noble gas filled thyratrons have better electrical parameters than mercury-based ones. The electrodes undergo damage by high-velocity ions. The neutral atoms of the gas slow the ions down by collisions, and reduce the energy transferred to the electrodes by the ion impact. Gases with high molecular weight, e.g. xenon, protect the electrodes better than lighter ones, e.g. neon.

- Helium is used in helium-neon lasers and in some thyratrons rated for high currents and high voltages. Helium provides about as short deionization time as hydrogen, but can withstand lower voltage, so it is used much less often.
- Neon has low ignition voltage and is frequently used in low-voltage tubes. Discharge in neon emits relatively bright red light; neon-filled switching tubes therefore also act as indicators, shining red when switched on. This is exploited in the decatron tubes, which act as both counters and displays. Its red light is exploited in neon signage. Used in fluorescent tubes with high power and short length, e.g. industrial lighting tubes. Has higher voltage drop in comparison with argon and krypton. Its low atomic mass provides only a little protection to the electrodes against accelerated ions; additional screening wires or plates can be

used for prolonging the anode lifetime. In fluorescent tubes it is used in combination with mercury.

- Argon was the first gas used in fluorescent tubes and is still frequently used due to its low cost, high efficiency, and very low striking voltage. In fluorescent tubes it is used in combination with mercury. It was also used in early rectifier tubes; first thyratrons were derived from such argon-filled tubes.
- Krypton can be used in fluorescent lamps instead of argon; in that application it reduces the total energy losses on electrodes from about 15% to 7%. The voltage drop per lamp length is however lower than with argon, which can be compensated by smaller tube diameter. Krypton-filled lamps also require higher starting voltage; this can be alleviated by using e.g. 25%-75% argon-krypton mixture. In fluorescent tubes it is used in combination with mercury.
- Xenon in pure state has high breakdown voltage, making it useful in higher-voltage switching tubes. Xenon is also used as a component of gas mixtures when production of ultraviolet radiation is required, e.g. in plasma displays, usually to excite a phosphor. The wavelength produced is longer than with argon and krypton and penetrates the phosphors better. To lower the ionization voltage, neon-xenon or helium-xenon are used; above 350 torr, helium has lower breakdown voltage than neon and vice versa. At concentrations of 1% and less of xenon, the Penning effect becomes significant in such mixtures, as most of xenon ionization occurs by collision with excited atoms of the other noble gas; at more than few percents of xenon, the discharge ionizes xenon directly due to most energy of the electrons being spent on direct ionization of xenon.
- Penning mixtures are used where lower ionization voltage is required, e.g., in the neon lamps, Geiger-Muller tubes and other gas-filled particle detectors. A classical combination is about 98–99.5% of neon with 0.5–2% of argon, used in, e.g., neon bulbs and in monochrome plasma displays.

Elemental vapors (metals and nonmetals)

- Mercury vapors are used for applications with high current, e.g. lights, mercury arc valves, ignitrons. Mercury is used because of its high vapor pressure and low ionization potential. Mercury mixed with an inert gas is used where the energy losses in the tube have to be low and the tube lifetime should be long. In mercury-inert gas mixtures, the discharge is initially carried primarily by the inert gas; the released heat then serves to evaporate enough mercury to reach the desired vapor pressure. Low-voltage (hundreds volts) rectifiers use saturated mercury vapor in combination with a small amount of inert gas, allowing cold start of the tubes. High-voltage (kilovolts and more) rectifiers use pure mercury vapor at low pressure, requiring maintenance of maximum temperature of the tube. The liquid mercury serves as a reservoir of mercury, replenishing the vapors that are used up during the discharge. Unsaturated mercury vapor can be used, but as it can not be replenished, the lifetime of such tubes is lower. The strong dependence of vapor pressure on mercury temperature limits the environments the mercury-based tubes can operate in. In low-pressure mercury lamps, there is an optimum mercury pressure for the highest efficiency. Photons emitted by ionized mercury atoms can

be absorbed by nearby nonionized atoms and either reradiated or the atom is deexcited nonradiatively, too high mercury pressure therefore causes losses of light. Too low mercury pressure leads to too few atoms present to get ionized and radiate photons. The optimum temperature for low-pressure mercury lamps is at about 42 °C, when the saturated vapor pressure of mercury (present as a drop of about 1 mg of liquid mercury in the tube, as a reservoir compensating for losses by clean-up) reaches this optimum. In lamps intended for operation at higher ambient temperatures, and at a wider temperature range, mercury is present in the form of an amalgam with e.g. bismuth and indium; the vapor pressure above amalgam is lower than above liquid mercury. Mercury is used in fluorescent tubes as a source of visible and ultraviolet light for exciting the phosphor; in that application it is usually used together with argon, or in some cases with krypton or neon. Mercury ions deionize slowly, limiting the switching speed of mercury-filled thyratrons. Ion bombardment with mercury ions of even relatively low energies also gradually destroys oxide-coated cathodes.

- Sodium vapors are used in sodium-vapor lamps.
- Sulfur vapors are used in sulfur lamps.
- Vapors of many metals, alone or together with a noble gas, are used in many lasers.

Other gases
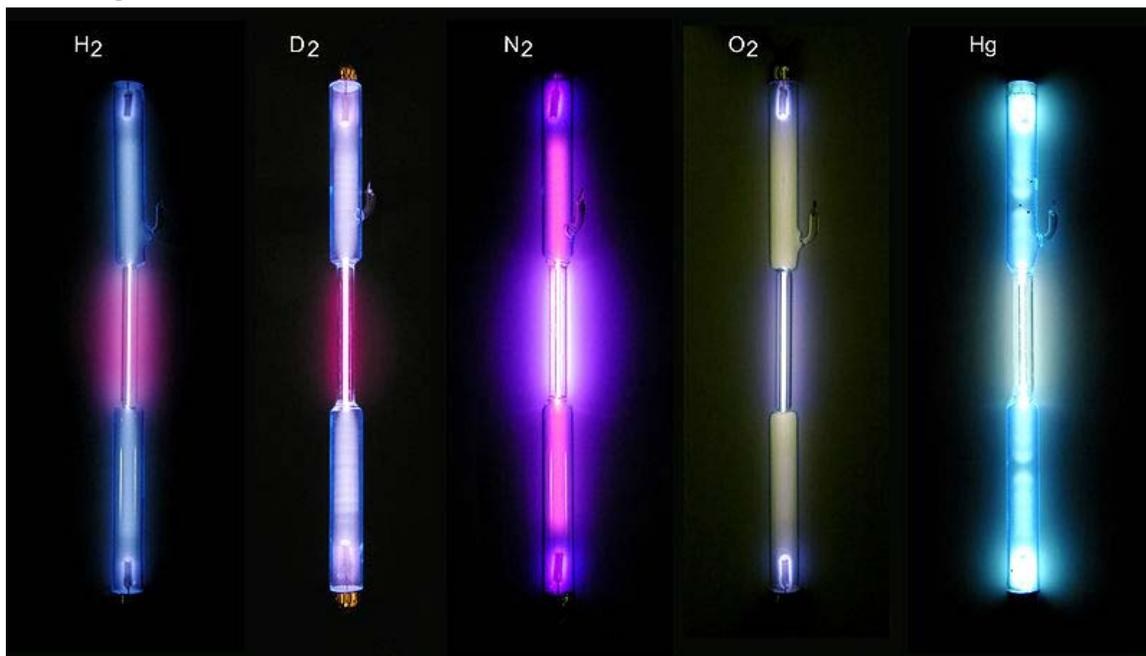
Other gases in discharge tubes; from left to right: hydrogen, deuterium, nitrogen, oxygen, mercury

- Air can be used in some low-demanding applications.
- Nitrogen at relatively high pressure tends to be used in surge arresters, due to its short build-up time, giving the tubes fast response time to voltage surges.

- Halogens and alcohol vapors absorb ultraviolet radiation and have high electron affinity. When added to inert gases, they quench the discharge; this is exploited in e.g. Geiger-Muller tubes.

Insulating gases

In special cases (e.g., high voltage switches), gases with good dielectric properties and very high breakdown voltages are needed. Highly electronegative elements, e.g., halogens, are favored as they rapidly recombine with the ions present in the discharge channel. One of the most popular choices is sulfur hexafluoride, used in special high-voltage applications. Other common options are dry pressurized nitrogen and halocarbons.

Gas-tube physics and technology

The gas pressure may range between 0.001 and 1000 torr; most commonly, pressures between 1-10 torr are used. The gas pressure influences the following factors:

- breakdown voltage (also called ignition voltage)
- current density
- operating voltage
- backfire voltage
- tube lifetime (lower pressure tubes tend to have shorter lifetimes due to using up of the gas)
- cathode sputtering, reduced at higher pressures

Above a certain value, the higher the gas pressure, the higher the ignition voltage. High-pressure lighting tubes can require a few kilovolts impulse for ignition when cold, when the gas pressure is low. After warming up, when the volatile compound used for light emission is vaporized and the pressure increases, reignition of the discharge requires either significantly higher voltage or reducing the internal pressure by cooling down the lamp. For example, many sodium vapor lamps cannot be re-lit immediately after being shut off; they must cool down before they can be lit up again.

The gas tends to be used up during the tube operation, by several phenomena collectively called **clean-up**. The gas atoms or molecules are adsorbed on the surfaces of the electrodes. In high voltage tubes, the accelerated ions can penetrate into the electrode materials. New surfaces, formed by sputtering of the electrodes and deposited on e.g. the inner surfaces of the tube, also readily adsorb gases. Non-inert gases can also chemically react with the tube components. Hydrogen may diffuse through some metals.

For removal of gas in vacuum tubes, getters are used. For resupplying gas for gas-filled tubes, **replenishers** are employed. Most commonly, replenishers are used with hydrogen; a filament made from a hydrogen-absorbing metal (e.g. zirconium or titanium) is present in the tube, and by controlling its temperature the ratio of absorbed and desorbed hydrogen is adjusted, resulting in controlling of the hydrogen pressure in the tube. The

metal filament acts as a hydrogen storage. This approach is used in e.g. hydrogen thyratrons or neutron tubes. Usage of saturated mercury vapor allows using a pool of liquid mercury as a large storage of material; the atoms lost by clean-up are automatically replenished by evaporation of more mercury. The pressure in the tube is however strongly dependent on the mercury temperature, which has to be controlled carefully.

The gas used dramatically influences the parameters of the tube. The breakdown voltage depends on the gas composition and electrode distance; the dependencies are described by Paschen's law.

Pure inert gases are used where the difference between the ignition voltage and the burning voltage has to be high, e.g. in switching tubes. Tubes for indication and stabilization, where the difference has to be lower, tend to be filled with Penning mixtures; the lower difference between ignition and burning voltages allows using lower power supply voltages and smaller series resistances.

Large rectifiers use saturated mercury vapor with a small amount of an inert gas. The inert gas supports the discharge when the tube is cold.

The mercury arc valve current-voltage characteristics are highly dependent on the temperature of the liquid mercury. The voltage drop in forward bias decreases from about 60 volts at 0 °C to somewhat above 10 volts at 50 °C and then stays constant; the reverse bias breakdown ("arc-back") voltage drops dramatically with temperature, from 36 kV at 60 °C to 12 kV at 80 °C to even less at higher temperatures. The operating range is therefore usually between 18–65 °C.

The gas in the tube has to be kept pure to maintain the desired properties; even small amount of impurities can dramatically change the tube values; presence of non-inert gases generally increases the breakdown and burning voltages. The presence of impurities can be observed by changes in the glow color of the gas. Air leaking into the tube makes the discharge look pale, milky, or reddish. Traces of mercury vapors glow bluish, obscuring the original gas color. Magnesium vapor colors the discharge green. To prevent outgassing of the tube components during operation, a bake-out is required before filling with gas and sealing. Thorough degassing is required for high-quality tubes; even as little as 10^{-8} torr of oxygen is sufficient for covering the electrodes with monomolecular oxide layer in few hours. Non-inert gases can be removed by suitable getters. for mercury-containing tubes, getters that do not form amalgams with mercury (e.g. zirconium, but not barium) have to be used. Cathode sputtering may be used intentionally for gettering non-inert gases; some reference tubes use molybdenum cathodes for this purpose.

Lighting and display gas-filled tubes

Fluorescent lighting, CFL lamps, mercury and sodium discharge lamps and HID lamps are all gas-filled tubes used for lighting.

Neon lamps and neon signage (most of which is not neon based these days) are also low-pressure gas-filled tubes.

Specialized historic low-pressure gas-filled tube devices include the Nixie tube (used to display numerals) and the Decatron (used to count or divide pulses, with display as a secondary function).

Xenon flash lamps are gas-filled tubes used in cameras and strobe lights to produce bright flashes of light.

The recently developed sulfur lamps are also gas-filled tubes when hot.

### Gas-filled tubes in electronics

Some important examples include the thyratron, krytron, and ignitron tubes, which are used to switch high-voltage currents. A specialized type of gas-filled tube called a Gas Discharge Tube (GDT) is fabricated for use in surge protectors, to limit voltage surges in electrical and electronic circuits.

Chapter 4

Cavity Magnetron

The **cavity magnetron** is a high-powered vacuum tube that generates microwaves using the interaction of a stream of electrons with a magnetic field. The 'resonant' cavity magnetron variant of the earlier magnetron tube was invented by Randall and Boot in 1940. The high power of pulses from the cavity magnetron made centimetre-band radar practical. Shorter wavelength radars allowed detection of smaller objects. The compact cavity magnetron tube drastically reduced the size of radar sets so that they could be installed in anti-submarine aircraft and escort ships. At present, cavity magnetrons are commonly used in microwave ovens and in various radar applications.

Construction and operation
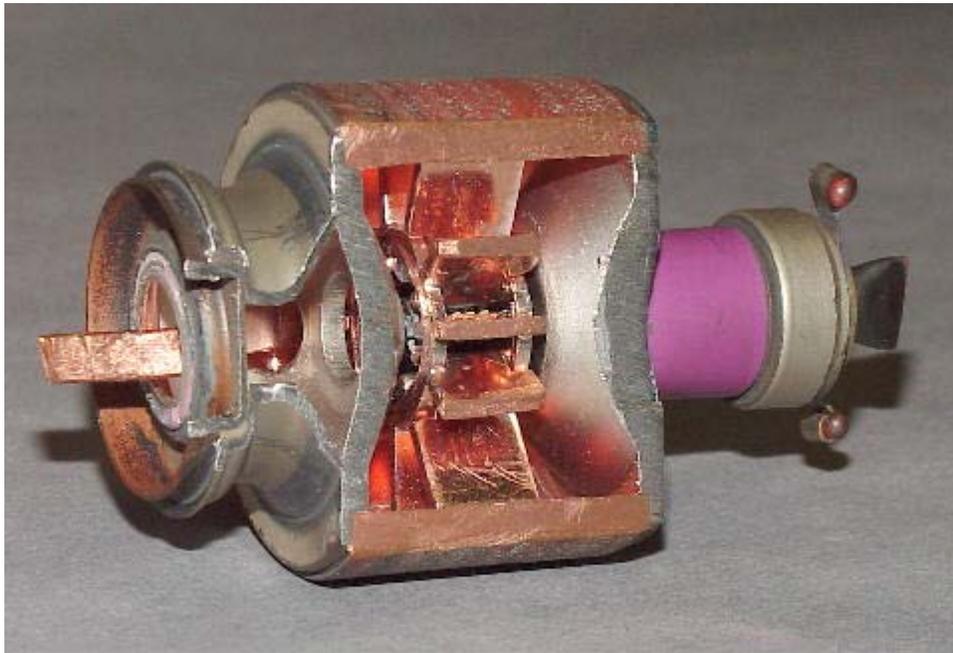
Magnetron with section removed (magnet is not shown)
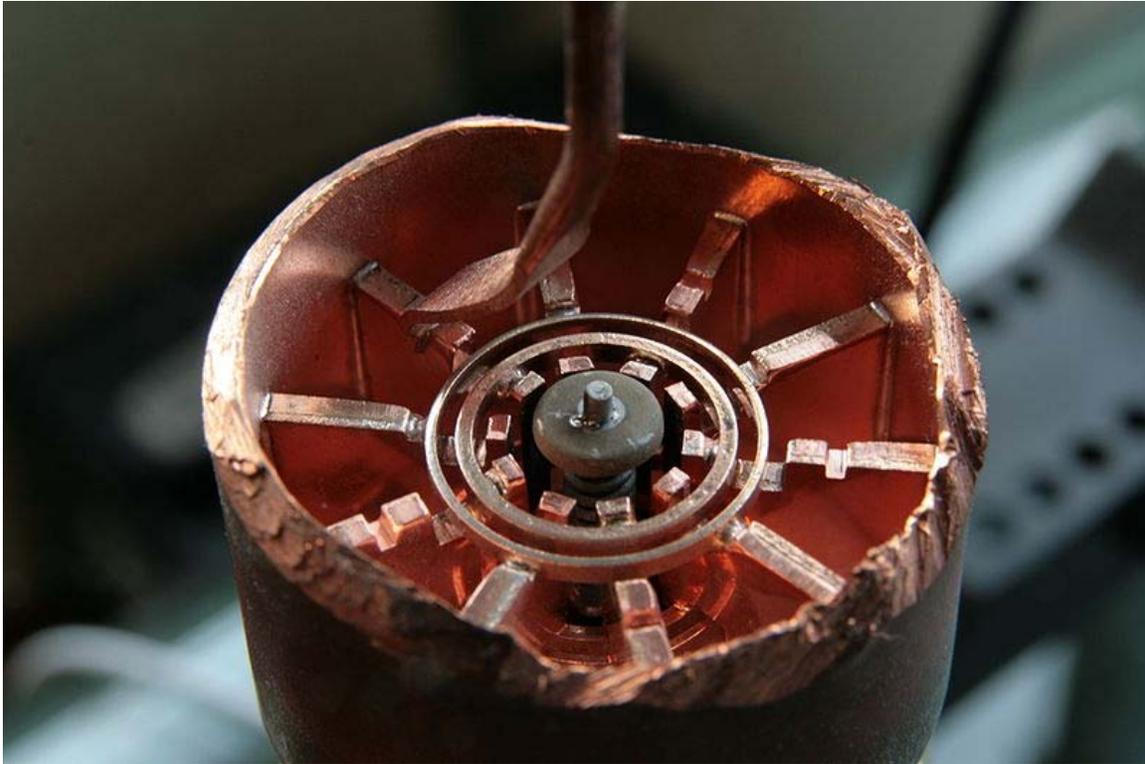
A similar magnetron with a different section removed (magnet is not shown).

All cavity magnetrons consist of a hot cathode with a high (continuous or pulsed) negative potential by a high-voltage, direct-current power supply. The cathode is built into the center of an evacuated, lobed, circular chamber. A magnetic field parallel to the filament is imposed by a permanent magnet. The magnetic field causes the electrons, attracted to the (relatively) positive outer part of the chamber, to spiral outward in a circular path rather than moving directly to this anode. Spaced around the rim of the chamber are cylindrical cavities. The cavities are open along their length and connect the common cavity space. As electrons sweep past these openings, they induce a resonant, high-frequency radio field in the cavity, which in turn causes the electrons to bunch into groups. A portion of this field is extracted with a short antenna that is connected to a waveguide (a metal tube usually of rectangular cross section). The waveguide directs the extracted RF energy to the load, which may be a cooking chamber in a microwave oven or a high-gain antenna in the case of radar.

Resonant cavity magnetron high-power
high-frequency oscillator
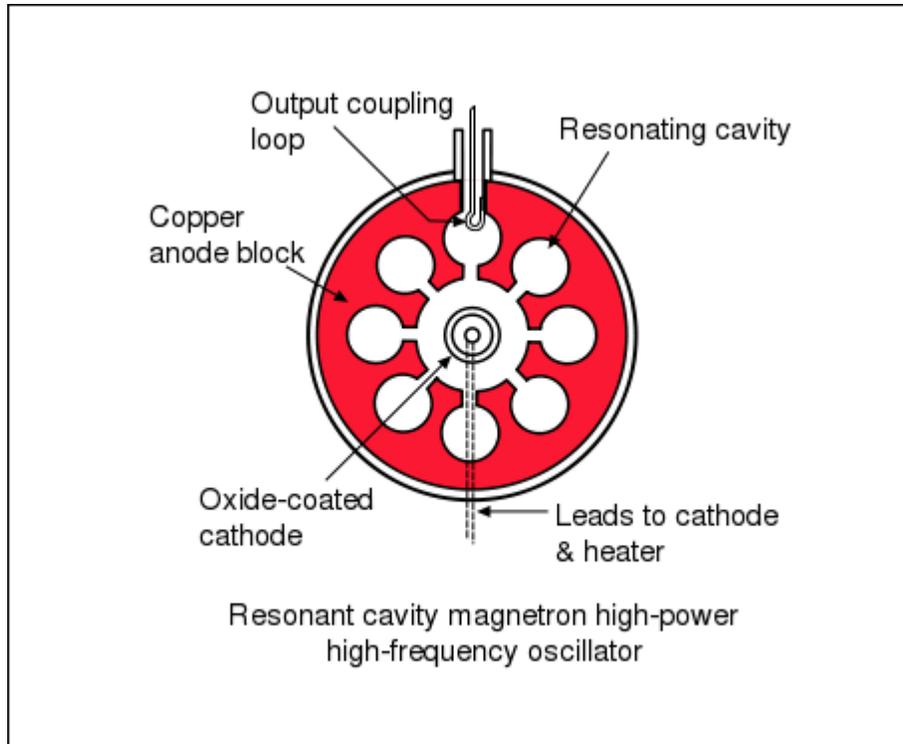
A cross-sectional diagram of a resonant cavity magnetron. Magnetic lines of force are parallel to the geometric axis of this structure.

The sizes of the cavities determine the resonant frequency, and thereby the frequency of emitted microwaves. However, the frequency is not precisely controllable. The operating frequency varies with changes in load impedance, with changes in the supply current, and with the temperature of the tube. This is not a problem in uses such as heating, or in some forms of radar where the receiver can be synchronized with an imprecise magnetron frequency. Where precise frequencies are needed, other devices such as the klystron are used.

The magnetron is a self-oscillating device requiring no external elements other than a power supply. A well-defined threshold anode voltage must be applied before oscillation will build up; this voltage is a function of the dimensions of the resonant cavity, and the applied magnetic field. In pulsed applications there is a delay of several cycles before the oscillator achieves full peak power, and the build-up of anode voltage must be coordinated with the build-up of oscillator output.

The magnetron is a fairly efficient device. In a microwave oven, for instance, a 1.1 kilowatt input will generally create about 700 watt of microwave power, an efficiency of around 65%. (The high-voltage and the properties of the cathode determine the power of a magnetron.) Large S-band magnetrons can produce up to 2.5 megawatts peak power with an average power of 3.75 kW. Large magnetrons can be water cooled. The magnetron remains in widespread use in roles which require high power, but where precise frequency control is unimportant.

Applications
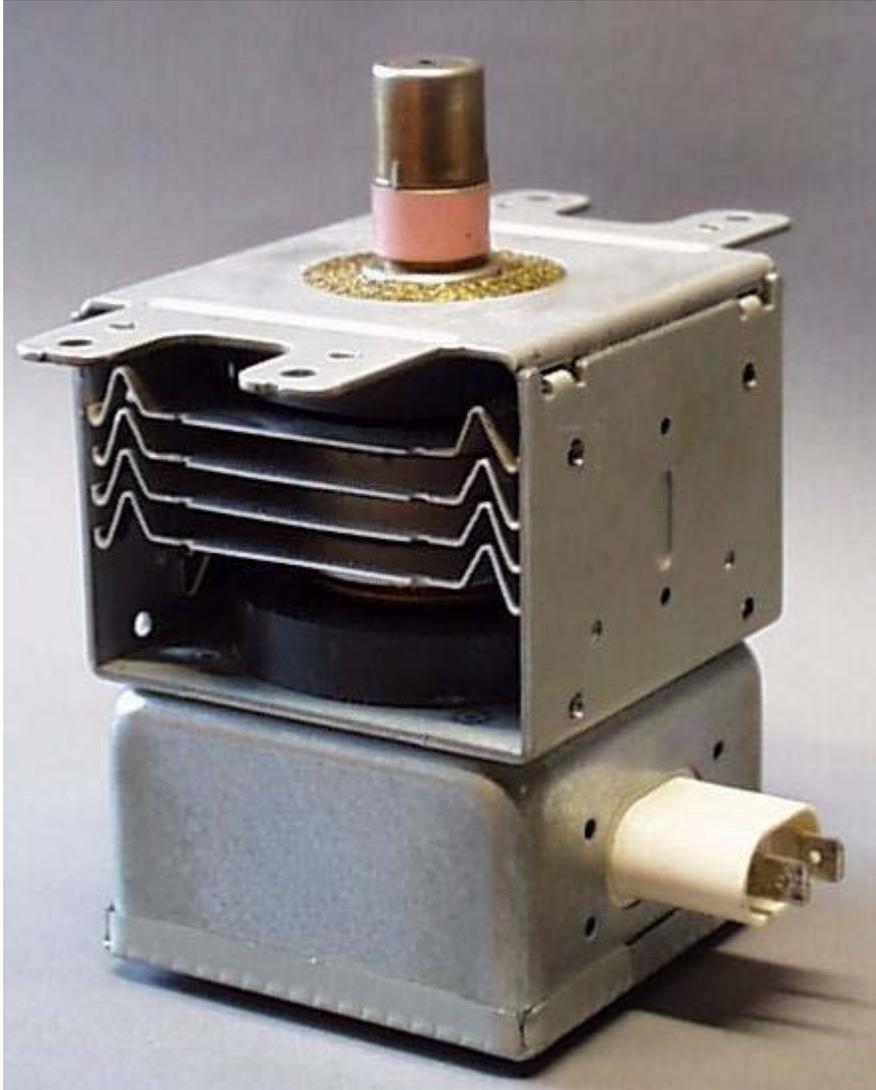
Magnetron from a microwave oven with magnet in its mounting box. The horizontal plates form a heat sink, cooled by airflow from a fan

Radar

In radar devices the waveguide is connected to an antenna. The magnetron is operated with very short pulses of applied voltage, resulting in a short pulse of high power microwave energy being radiated. As in all radar systems, the radiation reflected off a target is analyzed to produce a radar map on a screen.

Several characteristics of the magnetron's power output conspire to make radar use of the device somewhat problematic. The first of these factors is the magnetron's inherent instability in its transmitter frequency. This instability is noted not only as a frequency shift from one pulse to the next, but also a frequency shift within an individual transmitter

pulse. The second factor is that the energy of the transmitted pulse is spread over a wide frequency spectrum, which makes necessary its receiver to have a corresponding wide selectivity. This wide selectivity permits ambient electrical noise to be accepted into the receiver, thus obscuring somewhat the received radar echoes, thereby reducing overall radar performance. The third factor, depending on application, is the radiation hazard caused by the use of high power electromagnetic radiation. In some applications, for example a marine radar mounted on a recreational vessel, a radar with a magnetron output of 2 to 4 kilowatts is often found mounted very near an area occupied by crew or passengers. In practical use, these factors have been overcome, or merely accepted, and there are today thousands of magnetron aviation and marine radar units in service. Recent advances in aviation weather avoidance radar and in marine radar have successfully implemented semiconductor transmitters that eliminate the magnetron entirely.

Heating

In microwave ovens the waveguide leads to a radio frequency-transparent port into the cooking chamber.

Lighting

In microwave-excited lighting systems, such as a sulfur lamp, a magnetron provides the microwave field that is passed through a waveguide to the lighting cavity containing the light-emitting substance (e.g., sulfur, metal halides, etc.)

History

The first simple, two-pole magnetron was developed in 1920 by Albert Hull at General Electric's Research Laboratories (Schenectady, New York), as an outgrowth of his work on the magnetic control of vacuum tubes in an attempt to work around the patents held by Lee De Forest on electrostatic control.

Hull's magnetron was not originally intended to generate VHF (very-high-frequency) electromagnetic waves. However, in 1924, Czech physicist August Žáček (1886–1961) and German physicist Erich Habann (1892–1968) independently discovered that the magnetron could generate waves of 100 megahertz to 1 gigahertz. Žáček, a professor at Prague's Charles University, published first; however, he published in a journal with a small circulation and thus attracted little attention. Habann, a student at the University of Jena, investigated the magnetron for his doctoral dissertation of 1924. Throughout the 1920s, Hull and other researchers around the world worked to develop the magnetron. Most of these early magnetrons were glass vacuum tubes with multiple anodes. However, the two-pole magnetron, also known as a split-anode magnetron, had relatively low efficiency. The cavity version (properly referred to as a *resonant-cavity magnetron*) proved to be far more useful.

While radar was being developed during World War II, there arose an urgent need for a high-power microwave generator that worked at shorter wavelengths (around 10 cm

(3 GHz)) rather than the 150 cm (200 MHz) that was available from tube-based generators of the time. It was known that a multi-cavity resonant magnetron had been developed and patented in 1935 by Hans Hollmann in Berlin. However, the German military considered its frequency drift to be undesirable and based their radar systems on the klystron instead. But klystrons could not achieve the high power output that magnetrons eventually reached. This was one reason that German night fighter radars were not a match for their British counterparts.

In 1940, at the University of Birmingham in the United Kingdom, John Randall and Harry Boot produced a working prototype similar to Hollman's cavity magnetron, but added liquid cooling and a stronger cavity. Randall and Boot soon managed to increase its power output 100 fold. Instead of abandoning the magnetron due to its frequency instability, they sampled the output signal and synchronized their receiver to whatever frequency was actually being generated. In 1941, the problem of frequency instability was solved by coupling alternate cavities within the magnetron.

Because France had just fallen to the Nazis and Britain had no money to develop the magnetron on a massive scale, Churchill agreed that Sir Henry Tizard should offer the magnetron to the Americans in exchange for their financial and industrial help (the Tizard Mission). An early 6 kW version, built in England by the General Electric Company Research Laboratories, Wembley, London (not to be confused with the similarly named American company General Electric), was given to the US government in September 1940. At the time the most powerful equivalent microwave producer available in the US (a klystron) had a power of only ten watts. The cavity magnetron was widely used during World War II in microwave radar equipment and is often credited with giving Allied radar a considerable performance advantage over German and Japanese radars, thus directly influencing the outcome of the war. It was later described as "the most valuable cargo ever brought to our shores".

The Bell Telephone Laboratories made a producible version from the magnetron delivered to America by the Tizard Mission, and before the end of 1940, the Radiation Laboratory had been set up on the campus of the Massachusetts Institute of Technology to develop various types of radar using the magnetron. By early 1941, portable centimetric airborne radars were being tested in American and British planes. In late 1941, the Telecommunications Research Establishment in Great Britain used the magnetron to develop a revolutionary airborne, ground-mapping radar codenamed H2S. The H2S radar was in part developed by Alan Blumlein and Bernard Lovell.

Centimetric radar, made possible by the cavity magnetron, allowed for the detection of much smaller objects and the use of much smaller antennas. The combination of small-cavity magnetrons, small antennas, and high resolution allowed small, high quality radars to be installed in aircraft. They could be used by maritime patrol aircraft to detect objects as small as a submarine periscope, which allowed aircraft to attack and destroy submerged submarines which had previously been undetectable from the air. Centimetric contour mapping radars like H2S improved the accuracy of Allied bombers used in the strategic bombing campaign. Centimetric gun-laying radars were likewise far more

accurate than the older technology. They made the big-gunned Allied battleships more deadly and, along with the newly developed proximity fuze, made anti-aircraft guns much more dangerous to attacking aircraft. The two coupled together and used by anti-aircraft batteries, placed along the flight path of German V-1 flying bombs on their way to London, are credited with destroying many of the flying bombs before they reached their target.

Since then, many millions of cavity magnetrons have been manufactured; while some have been for radar the vast majority have been for microwave ovens. The use in radar itself has dwindled to some extent, as more accurate signals have generally been needed and developers have moved to klystron and traveling-wave tube systems for these needs.

Health hazards

Caution: radiowaves hazard

Among more speculative hazards, at least one in particular is well known and documented. As the lens of the eye has no cooling blood flow, it is particularly prone to overheating when exposed to microwave radiation. This heating can in turn lead to a

higher incidence of cataracts in later life. A microwave oven with a warped door or poor microwave sealing can be hazardous.

There is also a considerable electrical hazard around magnetrons, as they require a high voltage power supply.

Some magnetrons have beryllium oxide (beryllia) ceramic insulators, which are dangerous if crushed and inhaled, or otherwise ingested. Single or chronic exposure can lead to berylliosis, an incurable lung condition. In addition, beryllia is listed as a confirmed human carcinogen by the IARC; therefore, broken ceramic insulators or magnetrons should not be directly handled.

Chapter 5

Klystron

High-power klystron used at the Canberra Deep Space Communications Complex.
(Klystrons used for generating heterodyne reference frequencies in radar receivers are
about the size of a whiteboard pen.)

A **klystron** is a specialized linear-beam vacuum tube (evacuated electron tube). Klystrons are used as amplifiers at microwave and radio frequencies to produce both low-power reference signals for superheterodyne radar receivers and to produce high-power carrier waves for communications and the driving force for modern particle accelerators.

Klystron amplifiers have the advantage (over the magnetron) of coherently amplifying a reference signal so its output may be precisely controlled in amplitude, frequency and phase. Many klystrons have a waveguide for coupling microwave energy into and out of the device, although it is also quite common for lower power and lower frequency klystrons to use coaxial couplings instead. In some cases a coupling probe is used to couple the microwave energy from a klystron into a separate external waveguide.

All modern klystrons are amplifiers, since reflex klystrons, which were used as oscillators in the past, have been surpassed by alternative technologies.

The name *klystron* comes from the stem form κλυσ- (*klys*) of a Greek verb referring to the action of waves breaking against a shore, and the end of the word *electron*.

History

The brothers Russell and Sigurd Varian of Stanford University are the inventors of the klystron. Their prototype was completed in August 1937. Upon publication in 1939, news of the klystron immediately influenced the work of US and UK researchers working on radar equipment. The Varians went on to found Varian Associates to commercialize the technology (for example to make small linear accelerators to generate photons for external beam radiation therapy). In their 1939 paper, they acknowledged the contribution of A. Arsenjewa-Heil and Oskar Heil (wife and husband) for their velocity modulation theory in 1935.

The work of physicist W.W. Hansen was instrumental in the development of the klystron and was cited by the Varian brothers in their 1939 paper. His resonator analysis, which dealt with the problem of accelerating electrons toward a target, could be used just as well to decelerate electrons (i.e., transfer their kinetic energy to RF energy in a resonator). During the second World War, Hansen lectured at the MIT Radiation labs two days a week, commuting to Boston from Sperry gyroscope company on Long Island. His resonator, called a "hohlraum" by nuclear physicists and coined "rhumbatron" by the Varian brothers, is used in 2009 in the National Ignition Facility investigating nuclear fusion. Hansen died in 1949 as a result of exposure to beryllium oxide (BeO).

During the second World War, the Axis powers relied mostly on (then low-powered) klystron technology for their radar system microwave generation, while the Allies used the far more powerful but frequency-drifting technology of the cavity magnetron for microwave generation. Klystron tube technologies for very high-power applications, such as synchrotrons and radar systems, have since been developed.

Explanation

Klystrons amplify RF signals by converting the kinetic energy in a DC electron beam into radio frequency power. A beam of electrons is produced by a thermionic cathode (a heated pellet of low work function material), and accelerated by high-voltage electrodes (typically in the tens of kilovolts). This beam is then passed through an input cavity. RF energy is fed into the input cavity at, or near, its natural frequency to produce a voltage which acts on the electron beam. The electric field causes the electrons to bunch: electrons that pass through during an opposing electric field are accelerated and later electrons are slowed, causing the previously continuous electron beam to form bunches at the input frequency. To reinforce the bunching, a klystron may contain additional "buncher" cavities. The RF current carried by the beam will produce an RF magnetic field, and this will in turn excite a voltage across the gap of subsequent resonant cavities. In the output cavity, the developed RF energy is coupled out. The spent electron beam, with reduced energy, is captured in a collector.

Two-cavity klystron amplifier
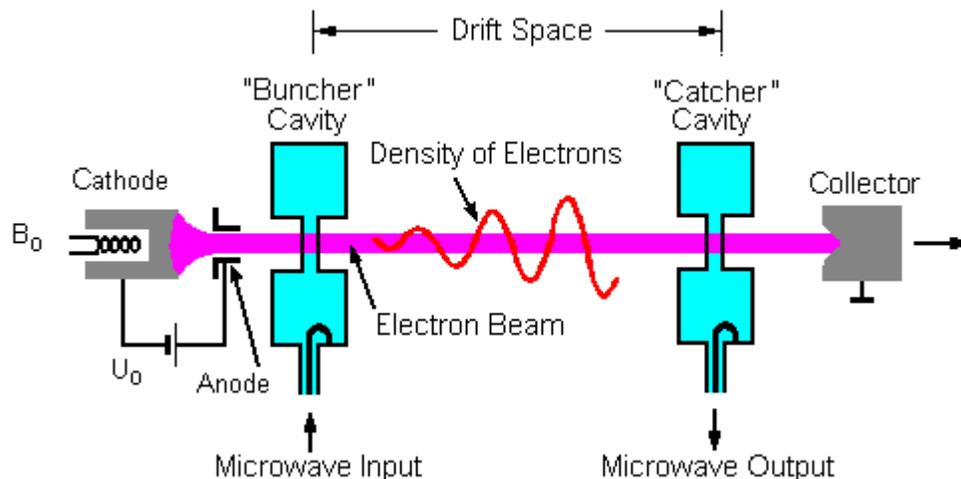
In the two-chamber klystron, the electron beam is injected into a resonant cavity. The electron beam, accelerated by a positive potential, is constrained to travel through a cylindrical *drift tube* in a straight path by an axial magnetic field. While passing through the first cavity, the electron beam is velocity modulated by the weak RF signal. In the moving frame of the electron beam, the velocity modulation is equivalent to a plasma oscillation. **Plasma oscillations** are rapid oscillations of the electron density in conducting media such as plasmas or metals.(The frequency only depends weakly on the wavelength). So in a quarter of one period of the plasma frequency, the velocity modulation is converted to density modulation, i.e. bunches of electrons. As the bunched electrons enter the second chamber they induce standing waves at the same frequency as the input signal. The signal induced in the second chamber is much stronger than that in the first.

Two-cavity klystron oscillator

The two-cavity amplifier klystron is readily turned into an oscillator klystron by providing a feedback loop between the input and output cavities. Two-cavity oscillator klystrons have the advantage of being among the lowest-noise microwave sources available, and for that reason have often been used in the illuminator systems of missile targeting radars. The two-cavity oscillator klystron normally generates more power than the reflex klystron—typically watts of output rather than milliwatts. Since there is no reflector, only one high-voltage supply is necessary to cause the tube to oscillate, the voltage must be adjusted to a particular value. This is because the electron beam must produce the bunched electrons in the second cavity in order to generate output power. Voltage must be adjusted to vary the velocity of the electron beam (and thus the frequency) to a suitable level due to the fixed physical separation between the two cavities. Often several "modes" of oscillation can be observed in a given klystron.

Reflex klystron
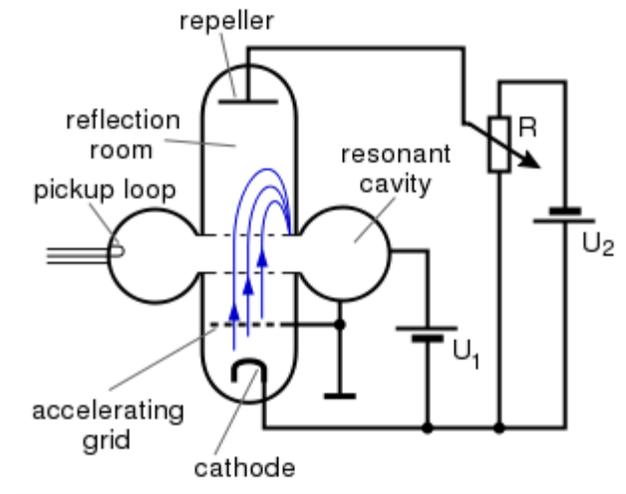
In the reflex klystron (also known as a 'Sutton' klystron after its inventor), the electron beam passes through a single resonant cavity. The electrons are fired into one end of the tube by an electron gun. After passing through the resonant cavity they are reflected by a negatively charged reflector electrode for another pass through the cavity, where they are then collected. The electron beam is velocity modulated when it first passes through the cavity. The formation of electron bunches takes place in the drift space between the reflector and the cavity. The voltage on the reflector must be adjusted so that the bunching is at a maximum as the electron beam re-enters the resonant cavity, thus ensuring a maximum of energy is transferred from the electron beam to the RF oscillations in the cavity. The voltage should always be switched on before providing the input to the reflex klystron as the whole function of the reflex klystron would be destroyed if the supply is provided after the input. The reflector voltage may be varied slightly from the optimum value, which results in some loss of output power, but also in a variation in frequency. This effect is used to good advantage for automatic frequency control in receivers, and in frequency modulation for transmitters. The level of

modulation applied for transmission is small enough that the power output essentially remains constant. At regions far from the optimum voltage, no oscillations are obtained at all. This tube is called a reflex klystron because it repels the input supply or performs the opposite function of a klystron.

There are often several regions of reflector voltage where the reflex klystron will oscillate; these are referred to as modes. The electronic tuning range of the reflex klystron is usually referred to as the variation in frequency between half power points—the points in the oscillating mode where the power output is half the maximum output in the mode. The frequency of oscillation is dependent on the reflector voltage, and varying this provides a crude method of frequency modulating the oscillation frequency, albeit with accompanying amplitude modulation as well.

Modern semiconductor technology has effectively replaced the reflex klystron in most applications.

Multicavity klystron

Large klystrons as used in the storage ring of the Australian Synchrotron to maintain the energy of the electron beam

In all modern klystrons, the number of cavities exceeds two. A larger number of cavities may be used to increase the gain of the klystron, or to increase the bandwidth.

Tuning a klystron

Some klystrons have cavities that are tunable. Tuning a klystron is delicate work which, if not done properly, can cause damage to equipment or injury to the technician. By adjusting the frequency of individual cavities, the technician can change the operating

frequency, gain, output power, or bandwidth of the amplifier. The technician must be careful not to exceed the limits of the graduations, or damage to the klystron can result.

Manufacturers generally send a card with the unique calibrations for a klystron's performance characteristics, that lists the graduations to be set to attain any of a set of listed frequencies. No two klystrons are exactly identical (even when comparing like part/model number klystrons), and so every card is specific to the individual unit. Klystrons have serial numbers on each of them to uniquely identify each unit, and for which manufacturers may (hopefully) have the performance characteristics in a database. If not, loss of the calibration card may be an economically insoluble problem, making the klystron unusable or perform marginally un-tuned.

Other precautions taken when tuning a klystron include using nonferrous tools. Some klystrons employ permanent magnets. If a technician uses ferrous tools, (which are ferromagnetic), and comes too close to the intense magnetic fields that contain the electron beam, such a tool can be pulled into the unit by the intense magnetic force, smashing fingers, injuring the technician, or damaging the unit. Special lightweight nonmagnetic (aka diamagnetic) tools made of beryllium alloy have been used for tuning U.S. Air Force klystrons.

Precautions are routinely taken when transporting klystron devices in aircraft, as the intense magnetic field can interfere with magnetic navigation equipment. Special overpacks are designed to help limit this field "in the field," and thus allow such devices to be transported safely.

Optical klystron

In an optical klystron the cavities are replaced with undulators. Very high voltages are needed. The electron gun, the drift tube and the collector are still used.

Floating drift tube klystron

The floating drift tube klystron has a single cylindrical chamber containing an electrically isolated central tube. Electrically, this is similar to the two cavity oscillator klystron with a lot of feedback between the two cavities. Electrons exiting the source cavity are velocity modulated by the electric field as they travel through the drift tube and emerge at the destination chamber in bunches, delivering power to the oscillation in the cavity. This type of oscillator klystron has an advantage over the two-cavity klystron on which it is based. It only needs one tuning element to effect changes in frequency. The drift tube is electrically insulated from the cavity walls, and DC bias is applied separately. The DC bias on the drift tube may be adjusted to alter the transit time through it, thus allowing some electronic tuning of the oscillating frequency. The amount of tuning in this manner is not large and is normally used for frequency modulation when transmitting.

Collector

After the RF energy has been extracted from the electron beam, the beam is destroyed in a collector. Some klystrons include depressed collectors, which recover energy from the beam before collecting the electrons, increasing efficiency. Multistage depressed collectors enhance the energy recovery by "sorting" the electrons in energy bins.

Applications

Klystrons produce microwave power far in excess of that developed by solid state. In modern systems, they are used from UHF (hundreds of MHz) up through hundreds of gigahertz (as in the Extended Interaction Klystrons in the CloudSat satellite). Klystrons can be found at work in radar, satellite and wideband high-power communication (very common in television broadcasting and EHF satellite terminals), medicine (radiation oncology), and high-energy physics (particle accelerators and experimental reactors). At SLAC, for example, klystrons are routinely employed which have outputs in the range of 50 megawatts (pulse) and 50 kilowatts (time-averaged) at frequencies nearing 3 GHz.

Popular Science's "Best of What's New 2007" described a company, Global Resource Corporation, using a klystron to convert the hydrocarbons in everyday materials, automotive waste, coal, oil shale, and oil sands into natural gas and diesel fuel.

Chapter 6

Pentagrid Converter

The **pentagrid converter** is a radio receiving valve (vacuum tube) with five grids used as the frequency mixer stage of a superheterodyne radio receiver.

Grids of a 12SA7GT Pentagrid converter showing all 5 grids.

The pentagrid was part of a line of development of valves that were able to take an incoming RF signal and change its frequency to a fixed intermediate frequency which was then amplified, and detected in the remainder of the receiver circuitry. The device was generically referred to as a *frequency changer* or just *mixer*.

Origins

The first devices designed to change frequency in the manner described above seem to have been developed by the French who simply put two grids into what would otherwise have been an ordinary triode valve (the bi-grille). Although technically a four electrode device, neither the term tetrode nor the tetrode valve as we know it today had yet appeared. Each grid was able to accept one of the incoming signals and the non-linearity of the device produced the sum and difference frequencies. The valve would have been very inefficient but, most importantly, the capacitive coupling between the two grids would have been very large. It would therefore have been quite impossible to prevent the signal from one grid coupling out of the other. At least one reference claims that the bi-grille was self oscillating, but this has not been confirmed.

When Edwin Armstrong invented the superheterodyne receiver in 1918, although the tetrode had been invented a couple of years earlier, he nevertheless employed only triodes in his design. Armstrong employed a pair of triodes as his mixer stage. One operated in a conventional oscillator circuit, but he was able to employ the other as a mixer by coupling the oscillator signal into the mixer's cathode, and the received signal to the grid. The sum and difference frequencies were then available in the mixer's anode (or plate) circuit. Once again, the problem of coupling between the circuits would be ever present.

Shortly after Armstrong invented the superhet, a triode mixer stage design was developed that not only mixed the incoming signal with the local oscillator, but the same valve doubled as the oscillator. This was known as the autodyne mixer. Early examples had difficulty oscillating across the frequency range because the oscillator feedback was via the first intermediate frequency transformer primary tuning capacitor which was too small to give good feedback. Also keeping the oscillator signal out of the antenna circuit was difficult.

The invention of the tetrode demonstrated the idea of screening electrodes from each other by using additional earthed (grounded) grids (at least as far as the signal was concerned). In 1926, Philips invented a technique of adding yet another grid to combat the secondary emission that the tetrode suffered from. All the ingredients for the pentagrid were now in place.

The Pentagrid
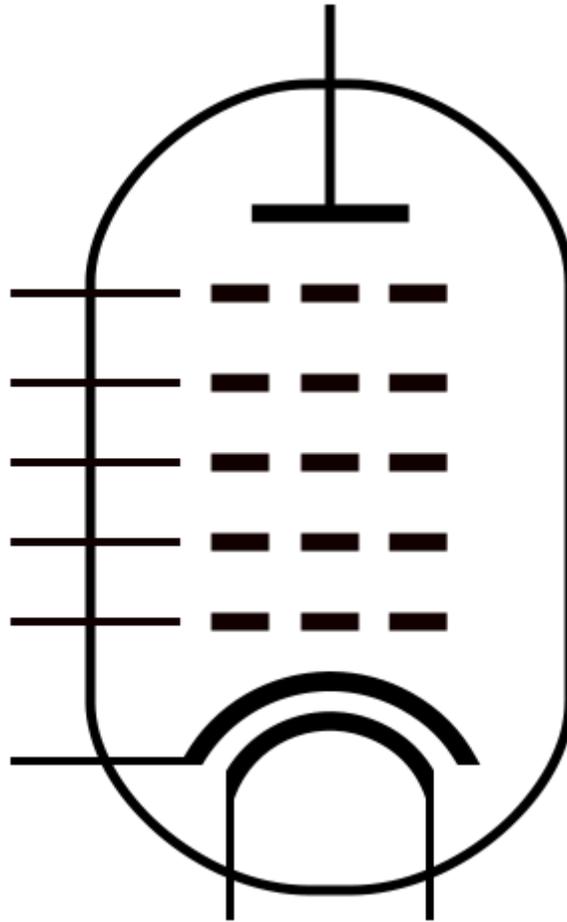
Circuit symbol of a heptode

The development of the pentagrid or heptode valve was a novel development in the mixer story. The idea was to produce a single valve that not only mixed the oscillator signal and the received signal and produced its own oscillator signal at the same time but, importantly, did the mixing and the oscillating in different parts of the same valve.

The invention of the device at first sight doesn't seem to be obscure, but it would appear that it was developed in both America and the United Kingdom, more or less at the same time. However, the UK device is different to its American counterpart.

It is known that Donald G. Haines of RCA applied for a patent for the pentagrid on 28 March 1933 (subsequently granted on 29 March 1939) under US patent number 2,148,266. The pentagrid also featured in a UK patent (GB426802) granted on 10 April 1935. However, the Ferranti company of Great Britain entered the valve business with the first known UK produced pentagrid, the **VHT4** late in 1933 (though it must have been in development, and would certainly have existed as a prototype well before this time).

The pentagrid proved to be a much better mixer. Since the oscillator circuit was more or less self contained, good feedback for reliable oscillation across the frequency range was easy to obtain. Some manufacturers that had adopted the autodyne mixer converted some, if not all, of their designs to pentagrid mixers.

What was the goal to develop a reliable self oscillating mixer? The reasons were to differ from the UK to America. The UK radio manufacturers had to pay a royalty of £1 per valve holder to the British Valve Association to cover use of their members' patent rights. Further, they dictated than not more than one electrode structure could be contained in a single envelope (which would have evaded the royalty - at least in part). The Americans appeared to be driven by the desire to produce a low cost 'every expense spared' design which was to lead to the All American Five. By making the mixer self oscillate, the necessity of providing a separate oscillator valve is avoided. The All American Five was to use a pentagrid converter from when it first appeared in 1934, right up until valves became obsolete when transistors took over.

In the UK, the five grids operated thus. Grid 1 acted as the oscillator grid in conjunction with grid 2 which acted as its anode. Grid 4 accepted the incoming signal with the remaining two grids, 3 and 5 connected together (usually internally) which acted as screen grids to screen the anode, grid 4 and grid 2 from each other. Because grid 2 was a 'leaky' anode in that it allowed part of the modulated electron stream through, the oscillator was coupled into the mixing section of the valve. In fact, in some designs, grid 2 consisted of just the support rods, the actual grid wire itself being omitted.

In America, the configuration was different. Grid 1 acted as the oscillator grid as before, but in this case, grids 2 and 4 were connected together (again usually internally). Grid 2 functioned as both a screen and the oscillator anode, in this case the grid wire had to be present to provide the screening. Grid 3 accepted the incoming signal. Grid 4 screened this from the anode, and grid 5 was a suppressor grid to suppress secondary emission. This configuration limited the oscillator design to one where the oscillator 'anode' was operated from the HT+ (B+) rail. This was often accomplished by using a Hartley Oscillator circuit and taking the cathode to the tap on the coil.

It will be noted that the UK version would have had significant secondary emission and would also have had a tetrode kink. This was exploited in providing the non linearity necessary to produce good sum and difference signals. The American devices although having no secondary emission due to the suppressor grid, nevertheless were able to get the required non linearity by biasing the oscillator such that the valve was overdriven. The American version was also a little more sensitive because the grid that accepted the signal was closer to the cathode increasing the amplification factor.

The pentagrid converter in either guise operated extremely well, but it suffered from the limitation that a strong signal was able to 'pull' the oscillator frequency away from a weaker signal. This was not considered a major problem in broadcast receivers where the signals were likely to be strong, but it became a problem when trying to receive weak signals that were close to strong signals. Some short wave radios managed quite

satisfactorily with these devices. Special high frequency versions appeared after World War II for the 100 MHz FM bands. Examples are the **6SB7Y** (1946) and the **6BA7** (1948). The pulling effect had a beneficial side effect in that it gave a degree of automatic tuning.

Another disadvantage was that in spite of the presence of the screen grids, the electron beam, modulated by the oscillator electrodes, still had to pass through the signal grid, and coupling of the oscillator into the signal circuit was inevitable. The American Federal Communication Commission (FCC) started requiring radio manufacturers to certify that their products avoided this interference under Part 15 of their rules. In the UK the Postmaster General (who was responsible for radio licensing), laid down a set of stringent rules concerning radio interference.

The Hexode
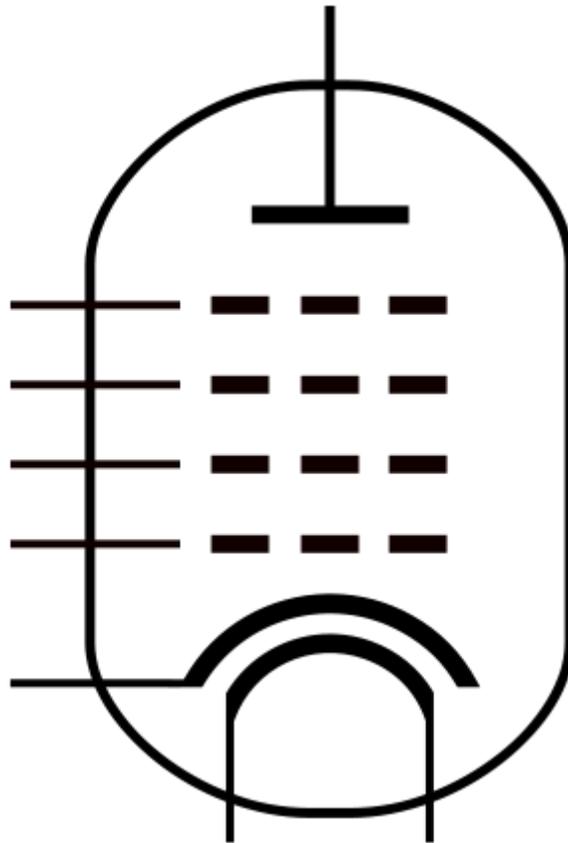
Circuit symbol of a hexode

It may come as a surprise that the Hexode was actually developed after the heptode or pentagrid. It was developed in Germany as a mixer but was designed from the start to be used with a separate triode oscillator. Thus the grid configuration was grid 1, signal input; grids 2 and 4 screen grids (connected together - again, usually internally) and grid 3 was

the oscillator input. The device had no suppressor grid. A major advantage was that by using grid 1 as the signal input grid, the device was more sensitive to weak signals.

It was not long before they put the triode and hexode structures in the same glass envelope - by no means a new idea. The triode grid was usually internally connected to the hexode grid 3, but this practice was dropped in later designs when the mixer section operated as a straight IF amplifier in AM/FM sets when operating on FM, the mixing being carried out in a dedicated FM frequency changing section.

The UK manufacturers were initially unable to use this type of mixer because of the BVA prohibition on multiple structures (and indeed separate valves because of the levy). Indeed one UK company, MOV, successfully enforced the cartel rules against the German Lissen company in 1934 when they attempted to market a radio in the UK which had the triode-hexode mixer.

Following pressure from the UK manufacturers, the BVA were compelled to relax the rules and the UK started to adopt triode-hexode mixers. The Mullard **ECH35** was a popular choice.

One company, Osram made an ingenious move. One of their popular pentagrid converter designs was the **MX40**, initially marketed in 1934. They put on sale in 1936, the **X41** triode-hexode frequency changer. The clever bit was that the **X41** was a direct plug in pin compatible replacement for the MX40. Thus a pentagrid radio could be easily converted to a triode-hexode without any other circuit modifications.

It is interesting to note, that America never really adopted the triode-hexode and it was seldom used, even though the **6K8** triode-hexode was available to manufacturers in 1938.

In some designs, a suppressor grid was added to produce yet another heptode design. Mullard's **ECH81** became popular with the move to miniature 9 pin valves.

The Octode

Although not strictly a pentagrid (in that it does not have 5 grids), this device nevertheless operates on the pentagrid principle. It resulted simply from the addition of a suppressor grid to the UK version of the pentagrid heptode. This was done mainly to reduce the power consumption for use in radio sets operated by dry batteries that were becoming increasingly popular.

In North America, the only octode manufactured was the **7A8**. Introduced by Sylvania in 1939 (and used mostly by Philco), this valve was the product of adding a suppressor grid to type **7B8**, which was the loctal version of type **6A7**. Adding the suppressor allowed Sylvania to lower the current of the 6.3 volt heater from 300 milliamperes to 150 milliamperes while maintaining the same conversion transconductance (550 microsiemens). This allowed Philco to use this valve in every line of radio throughout the 1940s.

One octode design worthy of mention was the Philips **EK3** Octode. This was designated as a 'beam octode'. The novel part about the design was that grids 2 and 3 were constructed as beam forming plates. This was done in such a way that Philips claimed that the oscillator electron beam and the mixer electron beams were separated as much as possible and thus the pulling effect was minimised. No information is available as to the degree of success.

The Pentode

The use of a pentode would seem an unlikely choice for a frequency converter. However, during the Great Depression, many American radio manufacturers used pentode types **6C6**, **6D6**, **77** and **78** in their lowest priced AC/DC receivers because they were cheaper than pentagrid type **6A7**. In these circuits, the suppressor (grid 3) acted as the oscillator grid.

One UK company, Mazda, produced a triode-pentode frequency changer, the **AC/TP**. Designed for low cost AC radios, the device was deliberately designed to allow strong signals to pull the oscillator without the risk of radiating the oscillator signal from the aerial. The cathode was common to both sections of the valve. The cathode was connected to a secondary coil on the oscillator coil and thus coupled the oscillator into the pentode mixer section, the signal being applied to grid 1 in the conventional manner. The AC/TP was one of the AC/ range of valves designed for low cost radios. They were considered durable for their time (even the AC/TP frequency changer, which was normally problematic). Any AC/ valves encountered today are likely to be brand new as service shops stocked up on spares which were seldom required.

Nomenclature

In order to distinguish between the two versions of the heptode or pentagrid, manufacturers data often describes them as 'heptode of the Hexode type' for a heptode without a screen grid, and a 'heptode of the octode type', where a screen grid is present.

Significant types

North American types

- For battery-powered home and portable radios
    - **1A6** - Dual-tetrode pentagrid used in 1930s battery-operated radios. Has 2.0 volt filament powered by 2 volt storage battery or, with ballast valve, 3 volt "air cell." Due to low filament current (60 ma), this valve is not used in circuits that tune shortwave frequencies above 10 meghertz. Octal version: **1D7-G**
    - **1C6** - Like type **1A6**, but with a higher filament current (120 ma) that allows it to tune frequencies up to 20 megahertz. Octal version: **1C7-G**

- o **1A7-GT** - Re-engineered version of types **1A6** and **1D7-G**, designed for use in portable AC/DC/Dry-cell battery radios introduced in 1938. Has 1.4 volt, 50 milliampere filament.
  - o **1B7-GT** - Re-engineered version of types **1C6** and **1C7-G**, designed for use in dry-cell battery radios with shortwave bands. Has 1.4 volt, 100 milliampere filament which makes it incompatible with 50 milliampere AC/DC filament strings.
  - o **1LA6** (loctal) and later **1L6** (7-pin miniature) – battery pentagrid for Zenith Trans-Oceanic short wave radio.
  - o **1LC6** - Similar to type **1LA6**, but with higher conversion transconductance.
  - o **1U6** - Nearly identical to type **1L6**, but with a 1.4 volt, 25 milliampere filament that makes it incompatible with 50 milliampere AC/DC filament strings.
  - o **1E8** - Subminiature "triode-pentode" pentagrid type
  - o **1V6** - Subminiature triode-pentode non-pentagrid converter
  - o **1R5** - "Triode-pentode" pentagrid used in many portable radios from 1941 to the beginning of the transistor radio era.

- For AC or AC/DC powered radios
  - o **2A7** and **6A7** – The first of the RCA pentagrids 1933, "dual-tetrode" design.
  - o **6A8** and **12A8** - Octal versions of types **2A7** and **6A7**. Type **12A8** was used in the first 150 milliampere AC/DC models.
  - o **7B8** and **14B8** - Loctal versions of types **6A8** and **12A8**.
  - o **7Q7** and **14Q7** - Loctal versions of types **6SA7** and **12SA7**.
  - o **6BE6** and **12BE6** - Seven-pin miniature versions of types **6SA7** and **12SA7**.
  - o **6J8-G** - American triode-heptode, 1938.
  - o **6K8** and **12K8** – American Triode-Hexode, 1938.
  - o **6SB7Y** (octal), **6BA7** and **12BA7** (9-pin miniatures) – Pentagrids for VHF use 1946
  - o **7A8** – the only octode produced in America by Sylvania, 1939. Used mostly in Philco radios.
  - o **18FX6** - Low-current (100 ma heater) version of type **12BE6**.

- For equipment powered by 12-volt car batteries (all have 12.6 volt heater, screen and plate voltages)
  - o **12FA6** - Car radio version of **12BE6**.
  - o **12GA6** - Similar to type **12FA6**, but with lower conversion transconductance.
  - o **12FX8** - Triode-heptode converter for car radios. Screen and both plate voltages rated at 12.6 volts.

European types

- **VHT1** – Ferranti pentagrid 1933.
- **MX40** – Osram pentagrid 1934.
- **X41** – Osram Triode Hexode 1936 – Plug-in replacement for **MX40**
- **ECH35** – Mullard Triode-Hexode
- **EK3** – Beam octode produced by Philips.
- **ECH81** – Mullard Triode-Heptode (of the hexode type) as well the **6И1П** Soviet version.
- **15A2**, **15D1** and **15D2** - British (Brimar) dual-triode pentagrids similar to American type 6A7. Home receiver type 15A2 has a 4.0 volt, 650 ma heater. Car radio types 15D1 and 15D2 have 13.0 volt heaters with 200 ma and 150 ma current ratings, respectively.

This list is by no means exhaustive.

The All American Five used a number of valve types in its history and the reader is referred to that article for a listing of the different types used.

Other mixer valves

- **6L7** - Superheterodyne mixer used in high-end console receivers of the late 1930s and early 1940s. (Most commonly used by Zenith.) Grid 1 is control grid (remote cutoff). Grid 3 is local oscillator signal injector. Local oscillator is always a separate valve, often a **6C5**. Due to low relatively conversion transconductance (350 μmhos), this valve is always fed by at least one tuned RF amplifier.
- **1612** - Audio mixer version of type **6L7**. Both control grids (1 and 3) are sharp-cutoff. Engineered to be less likely to develop microphonics.
- **1LB6** - Superheterodyne mixer for battery-operated radios. Like type **6L7**, works with a separate oscillator and tuned RF amplifier. Can be used with AC/DC/Dry-cell battery power supplies.
- **FM-1000** - Combined oscillator and quadrature FM detector for early post-war Philco radios. Manufactured only by Sylvania with either their markings or Philco markings.
- **6CS6** - Heptode with two sharp-cutoff control grids (1 and 3). Same basing as type **6BE6**. Used as sync separator in television sets. Also: **3CS6**, **4CS6**, **12CS6**.
- **6BY6** - Similar to type **6CS6**, but with higher transconductance. Also: **3BY6**.
- **12EG6** - Dual sharp-cutoff control grid mixer (grids 1 and 3). Has 12.6 volt plate and screen voltage. Designed for use with audio equipment powered by a car battery.

Photomultiplier

Photomultiplier

Dynodes inside a photomultiplier tube

Photomultiplier tubes (**photomultipliers** or **PMT**s for short), members of the class of vacuum tubes, and more specifically phototubes, are extremely sensitive detectors of light in the ultraviolet, visible, and near-infrared ranges of the electromagnetic spectrum. These detectors multiply the current produced by incident light by as much as 100 million times (i.e., 160 dB), in multiple dynode stages, enabling (for example) individual photons to be detected when the incident flux of light is very low.

The combination of high gain, low noise, high frequency response, and large area of collection has earned photomultipliers an essential place in nuclear and particle physics, astronomy, medical diagnostics including blood tests, medical imaging, motion picture film scanning (telecine), and high-end image scanners known as drum scanners. Semiconductor devices, particularly avalanche photodiodes, are alternatives to photomultipliers; however, photomultipliers are uniquely well-suited for applications requiring low-noise, high-sensitivity detection of light that is imperfectly collimated. While photomultipliers are extraordinarily sensitive and moderately efficient, research is

still underway to create a photon-counting light detection device that is much more than 99% efficient. Such a detector is of interest for applications related to quantum information and quantum cryptography. Elements of photomultiplier technology, when integrated differently, are the basis of night vision devices.

History

Combining two scientific discoveries

The invention of the photomultiplier is predicated upon two prior achievements, firstly discovering the photoelectric effect and secondly discovering secondary emission (i.e., the ability of electrons in a vacuum tube to cause the emission of additional electrons by striking an electrode).

Photoelectric effect

The first demonstration of the photoelectric effect was carried out in 1887 by Heinrich Hertz who demonstrated it using ultraviolet light. Significant for practical applications, Elster and Geitel two years later demonstrated the same effect using *visible* light striking alkali metals (potassium and sodium). The addition of caesium, another alkali metal, has permitted the range of sensitive wavelengths to be extended towards longer wavelengths in the red portion of the visible spectrum.

Historically, the photoelectric effect is associated with Albert Einstein, who relied upon the phenomenon to establish the fundamental principle of quantum mechanics, in 1905, an accomplishment for which Einstein received the 1921 Nobel Prize. It is worthwhile to note that Heinrich Hertz, working 18 years earlier, had not recognized that the kinetic energy of the emitted electrons is proportional to the frequency but independent of the optical intensity. This fact implied a discrete nature of light, i.e. the existence of *quanta*, for the first time.

Secondary emission

The phenomenon of secondary emission was first limited to purely electronic inventions (i.e., those lacking photosensitivity). In 1902, Austin and Starke reported that the metal surfaces impacted by electron beams emitted a larger number of electrons than were incident. The application of the newly discovered secondary emission to the amplification of signals was only proposed after World War I by Westinghouse scientist Joseph Slepian in a 1919 patent.

The first photomultiplier

The race towards a practical electronic television camera

The ingredients for inventing the photomultiplier were coming together during the 1920s as the pace of vacuum tube technologies accelerated. The primary goal for many, if not

most, workers was the need for a practical television camera technology. Television had been pursued with primitive prototypes for decades prior to the 1934 introduction of the first practical camera (the iconoscope). Early prototype television cameras lacked sensitivity. Photomultiplier technology was pursued to enable television camera tubes, such as the iconoscope and (later) the orthicon, to be sensitive enough to be practical. So the stage was set to combine the dual phenomena of photoemission (i.e., the photoelectric effect) with secondary emission, both of which had already been studied and adequately understood, to create a practical photomultiplier.

First photomultiplier, single-stage (early 1934)

The first documented photomultiplier demonstration dates to the early 1934 accomplishments of an RCA group based in Harrison, NJ. Harley Iams and Bernard Salzberg were the first to integrate a photoelectric-effect cathode and single secondary emission amplification stage in a single vacuum envelope and the first to characterize its performance as a photomultiplier with electron amplification gain. These accomplishments were finalized *prior* to June 1934 as detailed in the manuscript submitted to Proceedings of the Institute of Radio Engineers (Proc. IRE). The device consisted of a semi-cylindrical photocathode, a secondary emitter mounted on the axis, and a collector grid surrounding the secondary emitter. The tube had a gain of about eight and operated at frequencies well above 10 kHz.

Magnetic photomultipliers (mid 1934–1937)

Higher gains were sought than those available from the early single-stage photomultipliers. However, it is an empirical fact that the yield of secondary electrons is limited in any given secondary emission process, regardless of acceleration voltage. Thus, any single-stage photomultiplier is limited in gain. At the time the maximum first-stage gain that could be achieved was approximately 10 (very significant developments in the 1960s permitted gains above 25 to be reached using negative electron affinity dynodes). For this reason, multiple-stage photomultipliers, in which the photoelectron yield could be multiplied successively in several stages, were an important goal. The challenge was to cause the photoelectrons to impinge on successively higher-voltage electrodes rather than to travel directly to the highest voltage electrode. Initially this challenge was overcome by using strong magnetic fields to bend the electrons' trajectories. Such a scheme had earlier been conceived by inventor J. Slepian by 1919.

Accordingly, leading international research organizations turned their attention towards improving photomultipliers to achieve higher gain with multiple stages. This work proceeded against a background of economic boom and bust, tyrannical dictatorship, and cataclysmic war clouds collecting on the horizon.

In the USSR, RCA-manufactured radio equipment was introduced on a large scale by Joseph Stalin to construct broadcast networks, and the newly formed All-Union Scientific Research Institute for Television was gearing up a research program in vacuum tubes that was advanced for its time and place. Numerous visits were made by RCA scientific

personnel to the USSR in the 1930s, prior to the Cold War, to instruct the Soviet customers on the capabilities of RCA equipment and to investigate customer needs. During one of these visits, in September 1934, RCA's Vladimir Zworykin was shown the first multiple-dynode photomultiplier, or *photoelectron multiplier*. This pioneering device of 28-year-old Leonid A. Kubetsky achieved gains of 1000x or more when demonstrated in June 1934. The work was submitted for print publication only two years later, in July 1936 as emphasized in a recent 2006 publication of the Russian Academy of Sciences (RAS). which terms it "Kubetsky's Tube." The Soviet device used a magnetic field to confine the secondary electrons and relied on the Ag-O-Cs photocathode which had been demonstrated by General Electric in the 1920s.

By October 1935, Vladimir Zworykin, George Ashmun Morton, and Louis Malter of RCA in Camden, NJ submitted their manuscript describing the first comprehensive experimental and theoretical analysis of a multiple dynode tube — the device later called a *photomultiplier* — to Proc. IRE. The RCA prototype photomultipliers also used a Ag-O-Cs (silver oxide-caesium) photocathode. They exhibited a peak quantum efficiency of 0.4% at 800 nm.

Electrostatic photomultipliers (1937–present)

Whereas these early photomultipliers used the magnetic field principle, electrostatic photomultipliers (with no magnetic field) were demonstrated by Jan Rajchman of RCA Laboratories in Princeton, NJ in the late 1930s and became the standard for all future commercial photomultipliers. The first mass-produced photomultiplier, the Type 931, was of this design and is still commercially produced today.

Improved photocathodes

Also in 1936, a much improved photocathode, Cs_3Sb (caesium-antimony), was reported by P. Gorlich. The caesium-antimony photocathode had a dramatically improved quantum efficiency of 12% at 400 nm, and was used in the first commercially successful photomultipliers manufactured by RCA (i.e., the 931-type) both as a photocathode and as a secondary-emitting material for the dynodes. Different photocathodes provided differing spectral responses.

Spectral response of photocathodes

In the early 1940s the JEDEC (Joint Electron Devices Engineering Council), an industry committee on standardization, developed a system of designating spectral responses. The philosophy included the idea that the product's user need only be concerned about the response of the device rather than how the device may be fabricated. Various combinations of photocathode and window materials were assigned "S-numbers" (spectral numbers) ranging from S-1 through S-40, which are still in use today. For example, S-11 uses the caesium-antimony photocathode with a lime glass window, S-13 uses the same photocathode with a fused silica window, and S-25 uses a so-called "multialkali" photocathode (Na-K-Sb-Cs, or sodium-potassium-antimony-caesium) that

provides extended response in the red portion of the visible light spectrum. No suitable photoemissive surfaces have yet been reported to detect wavelengths longer than approximately 1700 nanometers, which can be approached by a special (InP/InGaAs(Cs)) photocathode.

Role of RCA

For decades, RCA was responsible for performing the most important work in developing and refining photomultipliers. RCA was also largely responsible for the commercialization of photomultiplers. The company compiled and published an authoritative and very-widely used *Photomultiplier Handbook*. RCA made printed copies available for free upon request. The handbook, which continues to be made available online at no cost by the successors to RCA, is considered to be an essential reference.

Following a corporate break-up in the late 1980s involving the acquisition of RCA by General Electric and disposition of the divisions of RCA to numerous third-parties, RCA's photomultiplier business became an independent company.

Lancaster, Pennsylvania facility

The Lancaster, Pennsylvania facility was opened by the U.S. Navy in 1942 and operated by RCA for the manufacture of radio and microwave tubes. Following the Allied victory in World War II, the naval facility was acquired by RCA. *RCA Lancaster*, as it became known, was the base for development and production of commercial television products. In subsequent years other products were added, such as cathode ray tubes, photomultiplier tubes, motion-sensing light control switches, and closed-circuit television systems.

Burle Industries

Burle Industries, as a successor to the RCA Corporation, carried the RCA photomultiplier business forward after 1986, based in the Lancaster, Pennsylvania facility. The 1986 acquisition of RCA by General Electric resulted in the divestiture of the RCA Lancaster New Products Division. Hence, 45 years after being founded by the U.S. Navy, its management team, led by Erich Burlefinger, purchased the division and in 1987 founded Burle Industries.

In 2005, after eighteen years as an independent enterprise, Burle Industries and a key subsidiary were acquired by Photonis, a European holding company Photonis Group. Following the acquisition, Photonis was composed of Photonis Netherlands, Photonis France, Photonis USA, and Burle Industries. Photonis USA operates the former Galileo Corporation Scientific Detector Products Group (Sturbridge, Massachusetts), which had been purchased by Burle Industries in 1999. The Group is known for microchannel plate detector (MCP) electron multipliers—an integrated micro-vacuum tube version of photomultipliers. MCPs are used for imaging and scientific applications, including night vision devices.

On 9 March 2009 Photonis announced that it would cease all production of photomultipliers at both the Lancaster, Pennsylvania and the Brive, France plants.

Other companies

The Japan-based company Hamamatsu Photonics (also known as Hamamatsu) has emerged since the 1950s as a leader in the photomultiplier industry. Hamamatsu, in the tradition of RCA, has published its own handbook, which is available without cost on the company's website. Hamamatsu uses different designations for particular photocathode formulations and introduces modifications to these designations based on Hamamatsu's proprietary research and development.

Structure and operating principles
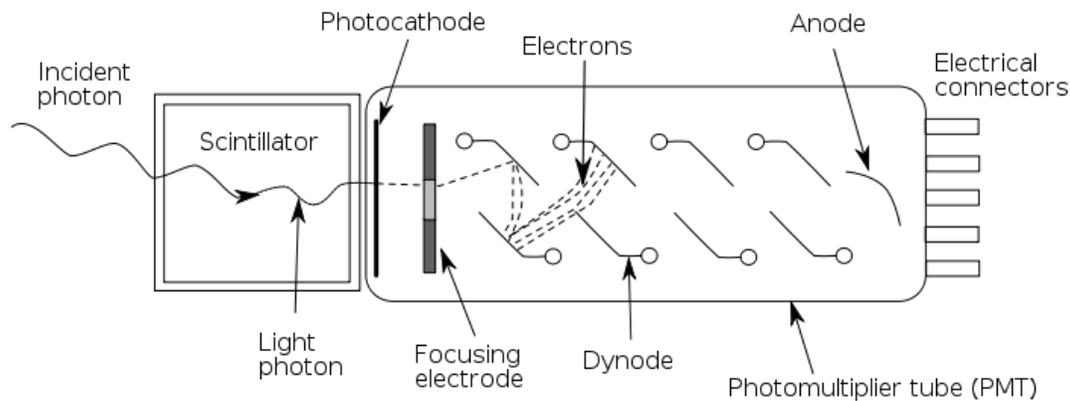
Schematic of a photomultiplier tube coupled to a scintillator

Photomultipliers are constructed from a glass envelope with a high vacuum inside, which houses a photocathode, several dynodes, and an anode. Incident photons strike the photocathode material, which is present as a thin deposit on the entry window of the device, with electrons being produced as a consequence of the photoelectric effect. These electrons are directed by the focusing electrode toward the electron multiplier, where electrons are multiplied by the process of secondary emission.

The electron multiplier consists of a number of electrodes called *dynodes*. Each dynode is held at a more positive voltage than the previous one. The electrons leave the photocathode, having the energy of the incoming photon (minus the work function of the photocathode). As the electrons move toward the first dynode, they are accelerated by the electric field and arrive with much greater energy. Upon striking the first dynode, more low energy electrons are emitted, and these electrons in turn are accelerated toward the second dynode. The geometry of the dynode chain is such that a cascade occurs with an ever-increasing number of electrons being produced at each stage. Finally, the electrons reach the anode, where the accumulation of charge results in a sharp current pulse indicating the arrival of a photon at the photocathode.

There are two common photomultiplier orientations, the *head-on* or *end-on* (transmission mode) design, as shown above, where light enters the flat, circular top of the tube and passes the photocathode, and the *side-on* design (reflection mode), where light enters at a particular spot on the side of the tube, and impacts on an opaque photocathode. Besides the different photocathode materials, performance is also affected by the transmission of the window material that the light passes through, and by the arrangement of the dynodes. A large number of photomultiplier models are available having various combinations of these, and other, design variables. Either of the manuals mentioned will provide the information needed to choose an appropriate design for a particular application.

Photocathode materials

The photocathodes can be made of a variety of materials, with different properties. Typically the materials have low work function and are therefore prone to thermionic emission, causing noise and dark current, especially the materials sensitive in infrared; cooling the photocathode lowers this thermal noise. The most common photocathode materials are:

- **Ag-O-Cs**: also called **S1**. Transission-mode, sensitive from 300–1200 nm. High dark current; used mainly in near-infrared, with the photocathode cooled.
- **GaAs:Cs**: caesium-activated gallium arsenide. Flat response from 300 to 850 nm, fading towards ultraviolet and to 930 nm.
- **InGaAs:Cs**: caesium-activated indium gallium arsenide. Higher infrared sensitivity than GaAs:Cs. Between 900–1000 nm much higher signal-to-noise ratio than Ag-O-Cs.
- **Sb-Cs**: caesium-activated antimony. Used for reflective mode photocathodes. Response range from ultraviolet to visible. Widely used.
- **Bialkali** (**Sb-K-Cs**, **Sb-Rb-Cs**): caesium-activated antimony-rubidium or antimony-potassium alloy. Similar to Sb:Cs, with higher sensitivity and lower noise. Can be used for transmission-mode; favorable response to a NaI:Tl scintillator flashes makes them widely used in gamma spectroscopy and radiation detection.
  - **High-temperature bialkali** (**Na-K-Sb**): can operate up to 175 °C, used in well logging. Low dark current at room temperature.
- **Multialkali** (**Na-K-Sb-Cs**): wide spectral response from ultraviolet to near-infrared; special cathode processing can extend range to 930 nm. Used in broadband spectrophotometers.
- **Solar-blind** (**Cs-Te**, **Cs-I**): sensitive to vacuum-UV and ultraviolet. Insensitive to visible light and infrared (CsTe has cutoff at 320 nm, CsI at 200 nm).

Window materials

The windows of the photomultipliers act as wavelength filters; this may be irrelevant if the cutoff wavelengths are outside of the application range or outside of the photocathode sensitivity range, but special care has to be taken for uncommon wavelengths.

- **Borosilicate glass** is commonly used for near-infrared to about 300 nm. Glass with very low content of potassium can be used with bialkali photocathodes to lower the background radiation from the potassium-40 isotope.
- **Ultraviolet glass** transmits visible and ultraviolet down to 185 nm. Used in spectroscopy.
- **Synthetic silica** transmits down to 160 nm, absorbs less UV than fused silica. Different thermal expansion than kovar (and than borosilicate glass that's expansion-matched to kovar), a graded seal needed between the window and the rest of the tube. The seal is vulnerable to mechanical shocks.
- **Magnesium fluoride** transmits ultraviolet down to 115 nm. Hygroscopic, though less than other alkali halides usable for UV windows.

Usage considerations

Photomultiplier tubes typically utilize 1000 to 2000 volts to accelerate electrons within the chain of dynodes. The most negative voltage is connected to the cathode, and the most positive voltage is connected to the anode. Negative high-voltage supplies (with the positive terminal grounded) are preferred, because this configuration enables the photocurrent to be measured at the low voltage side of the circuit for amplification by subsequent electronic circuits operating at low voltage. Voltages are distributed to the dynodes by a resistive voltage divider, although variations such as active designs (with transistors or diodes) are possible. The divider design, which influences frequency response or rise time, can be selected to suit varying applications. Some instruments that use photomultipliers have provisions to vary the anode voltage to control the gain of the system.

While powered (energized), photomultipliers must be shielded from ambient light to prevent their destruction through overexcitation. If used in a location with strong magnetic fields, which can curve electron paths, steer the electrons away from the dynodes and cause loss of gain, photomultipliers are usually shielded by a layer of mu-metal. This magnetic shield is often maintained at cathode potential. When this is the case, the external shield must also be electrically insulated because of the high voltage on it. Photomultipliers with large distances between the photocathode and the first dynode are especially sensitive to magnetic fields.

Typical applications

- Photomultipliers were the first electric eye devices, being used to measure interruptions in beams of light.
- Photomultipliers are used in conjunction with scintillators to detect nuclear and particle radiation in physics experiments.
- Photomultipliers are used in research laboratories to measure the intensity and spectrum of light-emitting materials such as compound semiconductors and quantum dots.
- Photomultipliers are used in numerous medical equipment designs. For example, blood analysis devices used by clinical medical laboratories utilize

photomultipliers to determine the relative concentration of various components in vials of blood drawn in doctors' offices, in combination with optical filters and incandescent lamps.

High sensitivity applications

After fifty years, during which solid-state electronic components have largely displaced the vacuum tube, the photomultiplier remains a unique and important optoelectronic component. Perhaps its most useful quality is that it acts, electronically, as a nearly perfect current source owing to the high voltage utilized in extracting the tiny currents associated with weak light signals. There is no Johnson noise associated with photomultiplier signal currents even though they are greatly amplified, e.g., by 100 thousand times (i.e., 100 dB) or more. The photocurrent still contains shot noise.

Photomultiplier-amplified photocurrents can be electronically amplified by a high-input-impedance electronic amplifier (in the signal path, subsequent to the photomultiplier), thus producing appreciable voltages even for nearly infinitesimally small photon fluxes. Photomultipliers offer the best possible opportunity to exceed the Johnson noise for many configurations. The aforementioned refers to measurement of light fluxes that, while small, nonetheless amount to a continuous stream of multiple photons.

For smaller photon fluxes, the photomultiplier can be operated in photon counting or Geiger mode. In Geiger mode the photomultiplier gain is set so high (using high voltage) that a single photo-electron resulting from a single photon incident on the primary surface generates a very large current at the output circuit. However, owing to the avalanche of current, a reset of the photomultiplier is required. In either case, the photomultiplier can detect individual photons. The drawback, however, is that not every photon incident on the primary surface is counted either because of less-than-perfect efficiency of the photomultiplier, or because a second photon can arrive at the photomultiplier during the "dead time" associated with a first photon and never be noticed.

A photomultiplier will produce a small current even without incident photons; this is called the *dark current*. Photon counting applications generally demand photomultipliers designed for low dark current.

Nonetheless, the ability to detect single photons striking the primary photosensitive surface itself reveals the quantization principle that Einstein put forth. Photon-counting (as it is called) reveals that light, not only being a wave, consists of discrete particles (i.e., photons).

Chapter 8

Battery (Vacuum Tube) & Cold Cathode

Battery (Vacuum Tube)
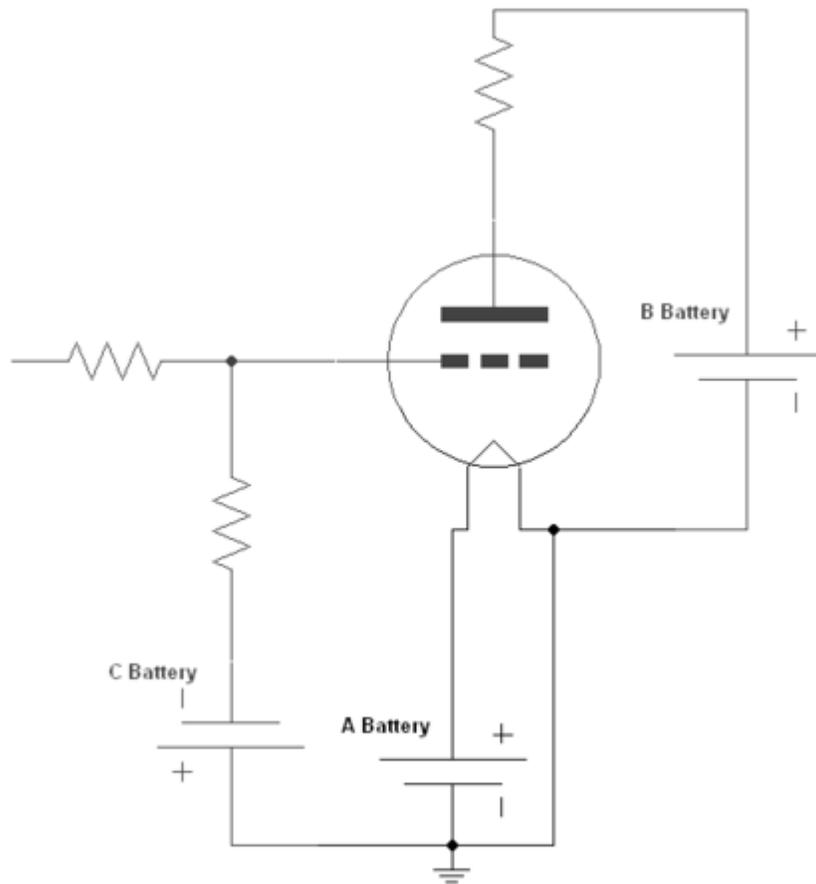
A generic triode circuit showing "A", "B" and "C" batteries

In the early days of electronics, vacuum tube devices (such as radios) were powered by batteries. Each battery had a different designation depending on which **vacuum tube** element it was associated with.

Initially, the only such device was a diode with only a filament (cathode) and a plate (anode). Following the direction of electron flow, these electrodes were identified as "A" and "B", respectively and thus the associated batteries were referred to as the "A" battery and "B" battery, respectively. Later, when the control grid element was added to create the triode tube, it was logically assigned the letter "C" and supplied from a "C" battery. Subsequent addition of further internal elements to improve the performance of the triode did not require an extension to this series of batteries - these elements were either resistively-biased from the existing batteries, connected to ground or to the cathode.

This nomenclature was used primarily within North America. Different battery names were used elsewhere in the English speaking world.

A battery

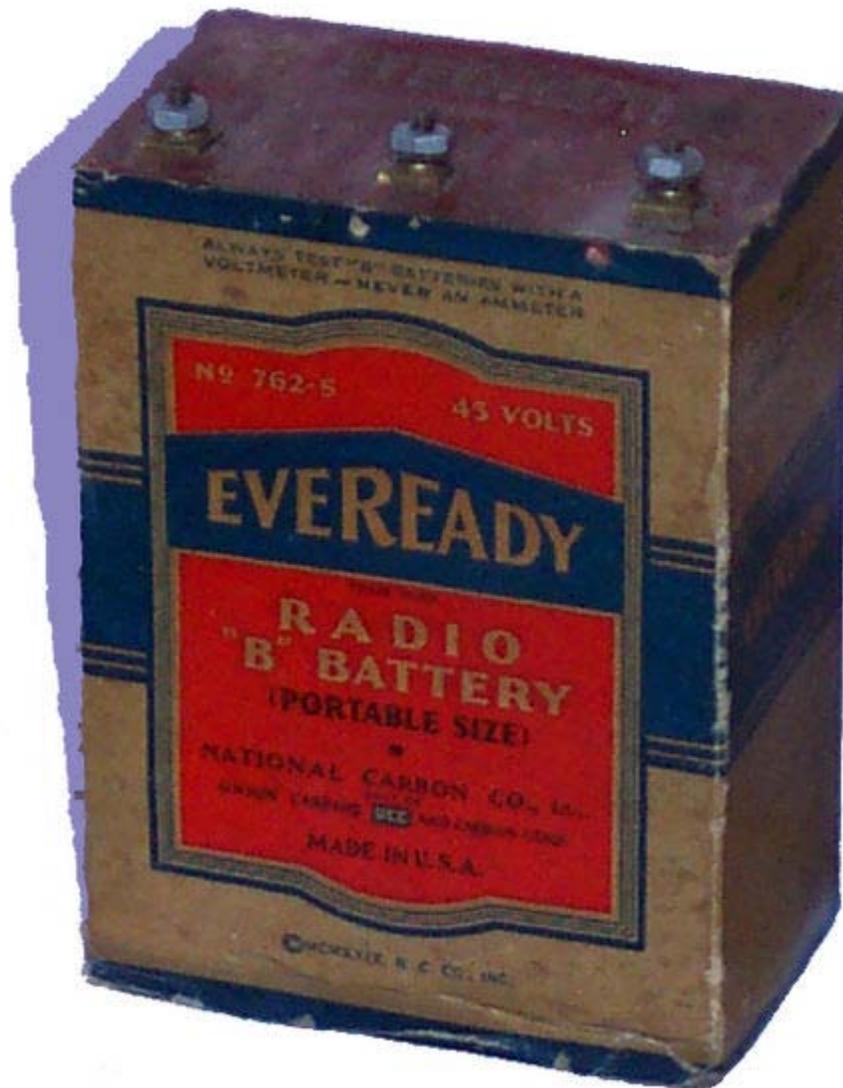

Eveready #742 1½ volts
Portable "A" battery

An **A battery** is any battery used to provide power to the filament of a vacuum tube. It is sometimes colloquially referred to as a "wet battery" (although there's no reason why a "dry" battery of suitable voltage couldn't be utilised for the purpose). The term comes from the days of valve (tube) radios when it was common practice to use a dry battery for the plate (anode) voltage and a rechargeable lead/acid "wet" battery for the filament voltage. (The filaments in vacuum tubes consumed much more current than the anodes, and so the "A" battery would drain much more rapidly than the "B" battery; therefore, using a rechargeable "A" battery in this role reduced the need for battery replacement. In contrast, a non-rechargeable "B" battery would need to be replaced relatively infrequently.)

A batteries were initially 2 volts, being lead acid accumulators, but with the introduction of all dry battery radios, 1.4 volts became more common. Other voltages can be encountered. For example 7.5 volt batteries were sometimes used to power a series connected set of 1.4 volt valves.

In Britain and in some other countries, the "A" battery was known as the "LT" (low tension) battery.

B battery

Eveready #762-S 45 volt radio "B" battery

A **B battery** is any battery used to provide the plate voltage of a vacuum tube. It is sometimes colloquially referred to as a "dry battery" (although there's no reason why a "wet" battery of suitable voltage couldn't be utilised for the purpose).

The filament is primarily a heat source and therefore the A battery supplies significant current and rapidly discharges. The B battery experiences comparatively little current draw and retains its stored capacity far longer than an A battery. Early B batteries used with bright emitter tubes were 120 volts, but these quickly became obsolete as they were

replaced with examples having voltages of typically 45 volts, 67½ volts, or 90 volts as more efficient tubes became available. Some examples had taps every 22½ volts.

Even when the plate voltage rail is fed by a power supply rather than a battery, it is generally referred to as the "B+" line in American schematics.

Because plate voltages can be as high as 300V DC, multiple B batteries may be connected together in series to additively provide the required operating voltages. The much higher available voltage of B batteries means that they must be handled more carefully than other battery types due to their ability to shock and/or burn the person handling them.

In Britain and in some other countries, the "B" battery was known as the "HT" (high tension) battery.

C battery

Eveready #761 4½ volt "C" battery

In electronics, a **C battery** is any battery used to provide bias to the control grid of a vacuum tube. Until the early 1930s this was common practice in valve (tube) radio sets but was largely superseded by grid leak resistors or voltage divider biasing.

Grid bias batteries are still manufactured today, but not for radio use. They are popular in schools and colleges as a convenient variable voltage source in science classes. The most popular battery is the 9 volt type with taps every 1½ volts that accept a banana plug.

A rare form of "C" battery was the *bias cell*, a button-size miniature battery designed to deliver a constant voltage with no current drain. These were briefly popular between 1936 and 1945 as the bias cell was less costly than a resistor/capacitor bias network.

In Britain and in some other countries, the "C" battery was known as the "GB" (grid bias) battery.

Cold Cathode

Cold cathode fluorescent lamp

A **cold cathode** is an element used within nixie tubes, gas discharge lamps, discharge tubes, and some types of vacuum tube. The term *cold cathode* refers to the fact that the cathode is not independently heated, but may still operate at elevated temperature.

Terminology

Types of cold cathode lamps include the following:

- Cold Cathode Fluorescent Lamps (**CCFLs**)

- Neon lamps are a very common example of a cold cathode lamp.

Cold cathode lamps remain popular for LCD backlighting and enthusiast computer case modders.

In the lighting industry, "cold cathode" historically refers to luminous tubing which is larger than 15mm in diameter and operates on a minimum current of 60 milliamps. This larger diameter tubing is often used for interior cove and accent lighting. The term "neon lamp" refers to tubing that is smaller than 15 mm diameter and typically operates at approximately 40 milliamps. These lamps are commonly used for neon signs.

Details

The cathode is the negative electrode. Any gas discharge lamp has a pair of electrodes, acting as cathode and anode (the positive electrode). Both electrodes alternate between acting as an anode and a cathode when these devices run with alternating current.

A standard computer case fitted with blue and green cold cathode tubes.

A *cold cathode* is distinguished from a hot cathode that is heated to induce thermionic emission of electrons. Discharge tubes with hot cathodes have an envelope filled with low pressure gas and containing a pair of cathodes. Examples are most common fluorescent lamps, high pressure discharge lamps and completely-evacuated electron tubes and vacuum fluorescent displays.

The interior surface of cold cathodes are capable of producing secondary electrons at a ratio greater than unity (amplification) upon electron and ion impact. For acceleration of the ions to a sufficient velocity for creating free electrons from the cathode material, cold cathode discharge lamps need higher voltages than hot cathode ones, causing a strong electric field near the cathodes.

Another mechanism for generating free electrons from a cold metallic surface is field electron emission. It is used in some x-ray tubes, the field electron microscope (FEM), and field emission displays (FEDs).

Cold cathodes sometimes have a rare earth coating on them to enhance electron emission. Some types contain a source of beta radiation to start ionization of the gas that fills the tube. In such a tube, glow discharge around the cathode is usually minimized, in favor of a so called positive column, filling the tube. The best example is the humble neon lamp. Another good example is nixie tubes. Nixie tubes too are cold cathode neon displays that also happen to be in-line, but not in-plane display devices.

A common cold cathode application is in neon signage and other locations where the ambient temperature is likely to drop well below freezing, The Clock Tower, Palace of Westminster (Big Ben) uses cold cathode lighting behind the clock faces where continual striking and failure to strike in cold weather would be less than ideal. Other examples include the thyratron, krytron, sprytron, and ignitron tubes. Large-scale Cold Cathode Fluorescent Lamps (CCFLs) have been produced in the past, and are still used today when shaped, long life linear light sources are required. Nowadays, miniature CCFLs are extensively used as backlights for computer liquid crystal displays, as well as LCD televisions. CCFLs' lifespans vary in LCD televisions depending on transient voltage surges, humidity and temperature levels in usage environments. Additionally, CCFLs are used by computer modders to light the insides of their customized computer cases.

Due to its efficiency, Cold Cathode Fluorescent Lamp (CCFL) technology has expanded into room lighting. Costs are competitive to those of fluorescent bulbs, but with several advantages. The light emitted is easier on the eyes, bulbs turn on instantly to full output and are also dimmable.

In systems using alternating current but without separate anode structures, the cathodes alternate as anodes and the impinging electrons can cause substantial localized heating, often to red heat. The cathode may take advantage of this heating to facilitate the thermionic emission of electrons when it is acting as a cathode. (*Instant start* fluorescent lamps definitely do employ this aspect; they start as cold-cathode devices but soon localized heating of the fine tungsten wire cathodes causes them to operate as ordinary hot cathode lamps.)

This aspect is problematic in the case of backlights used for LCD TV displays. New energy efficiency regulations being proposed in many countries will make variable backlighting highly desirable for LCD TV sets, as well as improving the perceived contrast range. However, CCFLs are strictly limited in the degree to which they can be

dimmed, both because a lower plasma current will lower the temperature of the cathode, causing erratic operation, and because running the cathode at too low temperature drastically shortens the life of the lamps. Much research is being directed to this problem, but high-end manufacturers are now turning to high-efficiency white LEDs as a better solution.

Cold cathode devices typically use a complex high-voltage power supply with some mechanism for limiting current. Although creating the initial space charge and the first arc of current through the tube may require a very high voltage, once the tube begins to heat up, the electrical resistance drops, thus increasing the electrical current through the lamp. To offset this effect and maintain normal operation, the supply voltage is gradually lowered. In the case of tubes with an ionizing gas, the gas can become a very hot plasma and electrical resistance is greatly reduced. If operated from a simple power supply without current limiting, this reduction in operational resistance would lead to damage to the power supply or the tube electrodes from overheating.

Hot Cathode & Fleming Valve

Hot Cathode
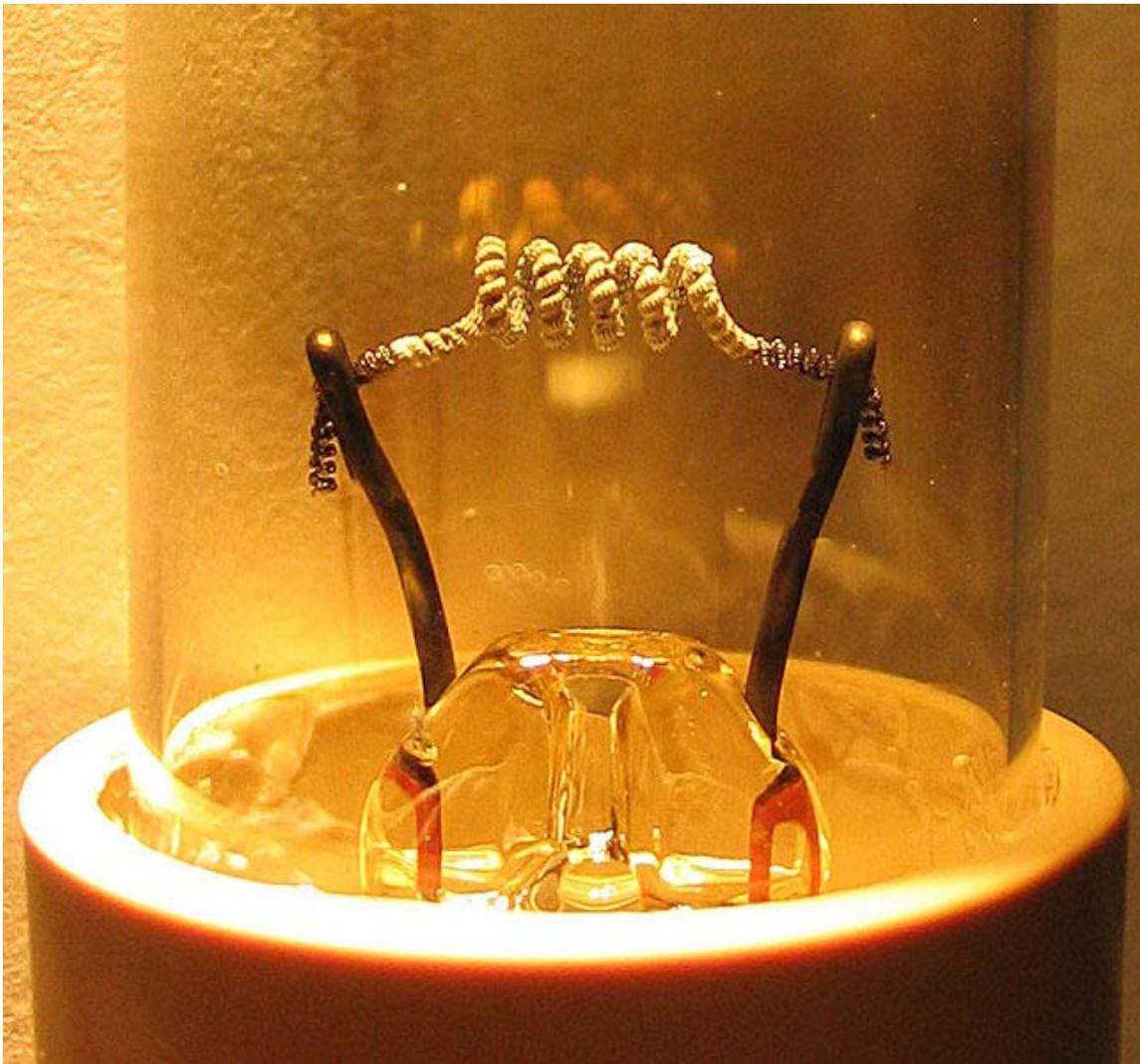
Closeup of the filament on a low pressure mercury gas discharge lamp showing white thermionic emission mix coating on the central portion of the coil. Typically made of a

mixture of barium, strontium and calcium oxides, the coating is sputtered away through normal use, often eventually resulting in lamp failure.

In vacuum tubes, a **hot cathode** is a cathode electrode which emits electrons due to thermionic emission. In the accelerator community, these are referred to as thermionic cathodes. (*Cf.* cold cathodes, where field electron emission is used and which do not require heating.) The heating element is usually an electrical filament. Hot cathodes typically achieve much higher power density than cold cathodes, emitting significantly more electrons from the same surface area.

Hot cathodes are the main source of electrons in electron guns in cathode ray tubes, electron microscopes, vacuum tubes, and fluorescent lamps.

Principles

Hot cathodes may be either **directly heated**, where the filament itself is the source of electrons, or **indirectly heated**, where the filament is electrically insulated from the cathode; this configuration minimizes the introduction of hum when the filament is energized with alternating current. The filament is most often made of tungsten. With indirectly heated cathodes, the filament is usually called the *heater* instead. The cathode for indirect heating is usually realized as a nickel tube which surrounds the heater.

The first cathodes consisted simply of a tungsten filament heated to white incandescence (known as bright emitters). Later cathodes are typically covered with an **emissive layer**, made of a material with lower work function, which emits electrons more easily than bare tungsten metal, reducing the necessary temperature and lowering the emission of metal ions. Cathodes can be made of pure sintered tungsten as well; tungsten cathodes in the shape of a parabolic mirror are used in electron beam furnaces. Thorium can be added to tungsten to increase its emissivity, due to its lower work function. Some cathodes are made of tantalum.

Variants

Oxide-coated cathodes

A common type is an **oxide-coated cathode**. The earliest material used was barium oxide; it forms a monoatomic layer of barium with an extremely low work function. More modern formulations utilize a mixture of barium oxide, strontium oxide and calcium oxide. Another standard formulation is barium oxide, calcium oxide, and aluminium oxide in a 5:3:2 ratio. Thorium oxide is used as well. Oxide-coated cathodes operate at about 800-1000 °C, orange-hot. They are used in most small glass vacuum tubes, but are rarely used in high-power tubes since they are vulnerable to high voltages and oxygen ions, and undergo rapid degradation under such conditions.

For manufacturing convenience, the oxide-coated cathodes are usually coated with carbonates, which are then converted to oxides by heating, and then the metal monolayer

is formed in a process called **electrode activation**. The activation may be achieved by microwave heating, direct electric current heating, or electron bombardment while the tube is on the exhausting machine, until the production of gases ceases. The purity of cathode materials is crucial for tube lifetime.

Boride cathodes

Lanthanum hexaboride hot cathode

Lanthanum hexaboride hot cathodes

Lanthanum hexaboride (LaB_6) and cerium hexaboride (CeB_6) are used as the coating of some high-current cathodes. Hexaborides show low work function, around 2.5 eV. They are also resistant to poisoning. Cerium boride cathodes show lower evaporation rate at 1700 K than lanthanum boride, but it becomes equal at 1850 K and higher. Cerium boride cathodes have one and a half times the lifetime of lanthanum boride, due to its higher resistance to carbon contamination. Boride cathodes are about ten times as "bright" as the tungsten ones and have 10-15 times longer lifetime. They are used e.g. in electron microscopes, microwave tubes, electron lithography, electron beam welding, X-Ray tubes, and free electron lasers. However these materials tend to be expensive.

Other hexaborides can be employed as well; examples are calcium hexaboride, strontium hexaboride, barium hexaboride, yttrium hexaboride, gadolinium hexaboride, samarium hexaboride, and thorium hexaboride.

Thoriated filaments

Thoriated filaments are another option, discovered in 1914 and made practical by Irving Langmuir in 1923. A small amount of thorium is added to the tungsten of the filament. The filament is heated white-hot, at about 2400 °C, and thorium atoms migrate to the surface of the filament and form the emissive layer. Heating the filament in a hydrocarbon atmosphere carburizes the surface and stabilizes the emissive layer. Thoriated filaments can have very long lifetimes and are resistant to high voltages. They are used in nearly all big high-power vacuum tubes for radio transmitters, and in some tubes for hi-fi amplifiers. Their lifetimes tend to be longer than those of oxide cathodes.

Thorium alternatives

Due to concerns about thorium radioactivity and toxicity, efforts have been made to find alternatives. One of them is **zirconiated tungsten**, where zirconium dioxide is used instead of thorium dioxide. Other replacement materials are lanthanum(III) oxide, yttrium(III) oxide, cerium(IV) oxide, and their mixtures.

Other materials

In addition to the listed oxides and borides, other materials can be used as well. Some examples are carbides and borides of transition metals, e.g. zirconium carbide, hafnium carbide, tantalum carbide, hafnium diboride, and their mixtures. Metals from groups IIIB (scandium, yttrium, and some lanthanides, often gadolinium and samarium) and IVB (hafnium, zirconium, titanium) are usually chosen.

In addition to tungsten, other refractory metals and alloys can be used, e.g. tantalum, molybdenum and rhenium and their alloys.

A barrier layer of other material can be placed between the base metal and the emission layer, to inhibit chemical reaction between these. The material has to be resistant to high temperatures, have high melting point and very low vapor pressure, and be electrically conductive. Materials used can be e.g. tantalum diboride, titanium diboride, zirconium diboride, niobium diboride, tantalum carbide, zirconium carbide, tantalum nitride, and zirconium nitride.

Cathode heater

A *cathode heater* is a heated wire filament used to heat the cathode in a vacuum tube or cathode ray tube. The cathode element had to achieve the required temperature in order for these tubes to function properly. This is why older electronics often needed some time to "warm up" after being powered on; this phenomenon can still be observed in the

cathode ray tubes of some modern televisions and computer monitors. The cathode heats to a temperature that causes electrons to be 'boiled out' of its surface into the evacuated space in the tube, a process called thermionic emission. The temperature required for modern oxide-coated cathodes is around 800–1,000 °C (1,472–1,832 °F)

The cathode is usually in the form of a long narrow sheet metal cylinder at the center of the tube. The heater consists of a fine wire or ribbon, made of a high resistance metal alloy like nichrome, similar to the heating element in a toaster but finer. It runs through the center of the cathode, often being coiled on tiny insulating supports or bent into hairpin-like shapes to give enough surface area to produce the required heat. The ends of the wire are electrically connected to two pins protruding from the end of the tube. When current passes through the wire it becomes red hot, and the radiated heat strikes the inside surface of the cathode, heating it. The red or orange glow seen coming from operating vacuum tubes is produced by the heater.

There is not much room in the cathode, and the cathode is often built with the heater wire touching it. The inside of the cathode is insulated by a coating of alumina (aluminum oxide). This is not a very good insulator at high temperatures, therefore tubes have a rating for maximum voltage between cathode and heater, usually only 200 - 300 V.

Heaters require a low voltage, high current source of power. Miniature receiving tubes for line-operated equipment used on the order of 0.5 to 4 watts for heater power; high power tubes such as rectifiers or output tubes would have used on the order of 10 to 20 watts, and broadcast transmitter tubes might need a kilowatt or more to heat the cathode. The voltage required was usually 5 or 6 volts AC. This was supplied by a separate 'heater winding' on the device's power supply transformer that also supplied the higher voltages required by the tubes' plates and other electrodes. A more common approach used in transformerless line-operated radio and television receivers such as the All American Five was to connect all the tube heaters in series across the supply line. Since all the heaters were rated at the same current, they would share voltage according to their heater ratings. Battery-operated radio sets used direct-current power for the heaters, and tubes intended for battery sets were designed to use as little heater power as necessary, to economize on battery replacement. Radio receivers were built with tubes using as little as 50 mA for the heaters, but these types were developed at about the same time as transistors which replaced them. Where leakage or stray fields from the heater circuit could potentially be coupled to the cathode, direct current was sometimes used for heater power. This would eliminate a source of noise in sensitive audio or instrumentation circuits.

Failure modes

The emissive layers degrade slowly with time, and much more quickly when the cathode is overloaded with too high current. The result is weakened emission and diminished power of the tubes, or brightness of the CRTs.

The activated electrodes can be destroyed by contact with oxygen or other chemicals (e.g. aluminium, or silicates), either present as residual gases, entering the tube via leaks, or released by outgassing or migration from the construction elements. This results in diminished emissivity. This process is known as **cathode poisoning**. High-reliability tubes had to be developed for the early Whirlwind computer, with filaments free of traces of silicon.

Slow degradation of the emissive layer and sudden burning and interruption of the filament are two main failure modes of vacuum tubes.

Transmitting tube hot cathode characteristics

Material	Operating temperature	Emission efficacy	Specific emission
Tungsten	2500 K	5 mA/W	500 mA/cm^2
Thoriated tungsten	2000 K	100 mA/W	5 A/cm^2
Oxide coated	1100 K	500 mA/W	10 A/cm^2
Barium aluminate	1300 K	400 mA/W	4 A/cm^2
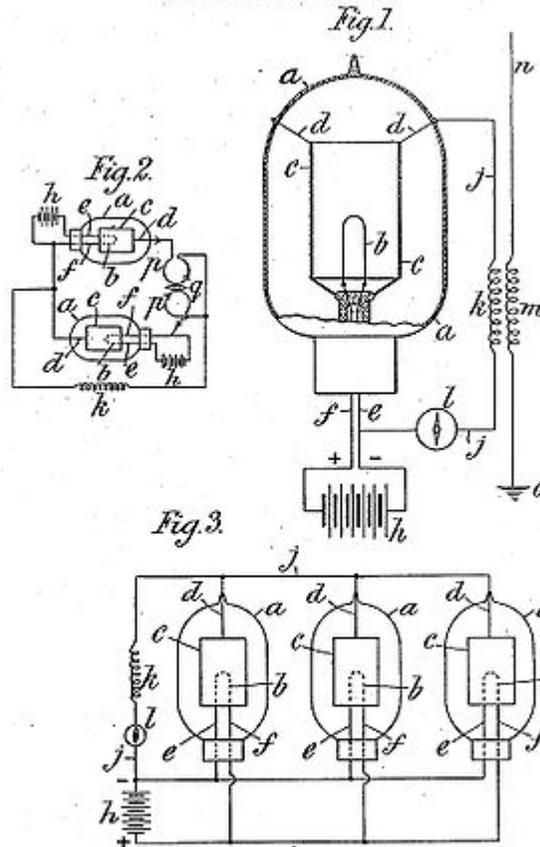
Fleming Valve

Fleming valve schematic from US Patent 803,684.

The **Fleming valve**, also called the **Fleming oscillation valve**, was a vacuum tube diode invented by John Ambrose Fleming and used in the earliest days of radio communication. As the first vacuum tube, the IEEE has described it as "one of the most important developments in the history of electronics", and it is on the List of IEEE Milestones for electrical engineering.

Description

Applications

The Fleming valve was the first practical application of the "Edison effect" (thermionic emission) discovered in 1883 by Thomas Edison shortly after his invention of the incandescent light bulb, that is, the emission of electrons by a lamp's heated filament to a nearby metal plate. Edison was granted a patent for this device as part of an electrical indicator in 1884, but did not hit upon any practical use. Professor Fleming of University College London consulted for the Edison Electric Light Company from 1881-1891, and subsequently for the Marconi Wireless Telegraph Company.

In 1901 Fleming designed a transmitter for Guglielmo Marconi to attempt transmission of radio waves across the Atlantic from Poldhu, England, to Nova Scotia, Canada. The distance between the two points was about 3,500 kilometres (2,100 miles). Although widely heralded as a great scientific advance, there was also some skepticism about this claim, in part because the signals had only been heard faintly and sporadically. In addition, there was no independent confirmation of the reported reception, and the transmission, which merely consisted of the three dots of the Morse code letter *S* sent repeatedly, came from a transmitter whose signals were difficult to differentiate from the noise made by atmospheric static discharges. (A detailed technical review of the early transatlantic work appears in John S. Belrose's work of 1995.) The receptions of signals were difficult to detect with a galvanometer. Fleming researched on a way that he could get the received signal to flow in only one direction, its oscillations could be detected with less trouble.

In 1904 Fleming tried an Edison effect bulb for this purpose, and found that it worked well to rectify high frequency oscillations and thus allow detection of the rectified signals by a galvanometer. On November 16, 1904, he applied for a US patent for what he termed an oscillation valve. This patent was subsequently issued as number 803,684 and found immediate utility in the detection of messages sent by Morse code.

Oscillation valves

The Fleming valve proved to be the start of a technological revolution. After reading Fleming's 1905 paper on his oscillation valve, American engineer Lee DeForest in 1906 created a three-element vacuum tube (the Audion) by adding a modulation grid. It could act as an amplifier and oscillator as well as detector. De Forest quickly refined his device into the triode, which was then central to the creation of long-distance telephone and radio communications, radars, and early digital computers. Similarities and differences between the Fleming valve and DeForest's triode caused decades of expensive and disruptive litigation, which were not settled until 1943 when the United States Supreme Court ruled Fleming's patent invalid.

Pentode & SY4307A

Pentode
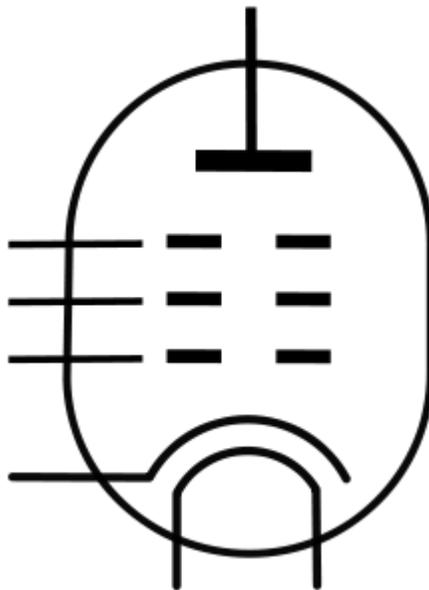
Schematic diagram of a pentode.

A **pentode** is an electronic device having five active electrodes. The term most commonly applies to a three-grid vacuum tube (thermionic valve), which was invented by the Dutchman Bernhard D.H. Tellegen in 1926. Pentodes (termed "triple-grid amplifiers" in some early literature) are closely related to beam tetrodes, and an improvement over conventional tetrodes, which were themselves a development of triodes.

Types of pentodes

- **Variable transconductance** ("**vari-mu**", "**remote-cutoff**" or "**super-control**") tubes in general are those with a non-uniform grid wire spacing to allow them to

handle a wide range of input signal levels without excessive cross-modulation distortion, and so useful in Radio frequency stages where Automatic gain control is applied to the pentode. Examples include: 1T4, 6K7, 6BA6, and the EF83 (while perhaps the EF85/6BY7, and certainly the 6JH6, could be described as "**semiremote-cutoff**" pentodes).

- **Sharp-cutoff** ("high slope" or ordinary) pentodes have the more ordinary uniform spacing of grid wires, and so mutual conductance decreases in an essentially uniform manner with increasing negative bias, and has a more abrupt cutoff. These pentodes are more suitable for audio amplifiers. Examples include: EF37A, EF86/6267, 1N5GT, 6AU6A, 6J7GT. Often in the European valve naming scheme for pentodes an even number indicated a sharp-cutoff device while odd indicated remote-cutoff; the EF37 was an exception to this general trend, perhaps due to its history as an update to the EF36 ("The Mullard EF36, EF37 and EF37A" at the National Valve Museum).
- Power output pentodes, also referred to as "Beam Pentodes" , include the EL34/6CA7, 6K6GT and the EL84/6BQ5. Sometimes the word "pentode" was avoided, in names such as "beam power amplifier" or "beam power tube", at a time when the pentode patent was an important issue.

Note: "triode-pentodes" are not a different type of pentode, but an envelope containing both a triode and a pentode, such as an ECF80 or ECL86.

Advantages over the tetrode

A tetrode could supply sufficient power to a speaker or transmitter, and offered a larger amplification factor than the earlier triode. However, the positively charged screen grid can collect the secondary electrons emitted from the anode, which can cause increased current toward the screen grid, and cause the anode current to *decrease* with increasing anode voltage over part of the Ia/Va characteristic.

A pentode, as introduced by Tellegen, has an additional electrode, or third grid, called the suppressor grid that solves the problem of secondary emission. The suppressor grid does this by being held at a low potential, usually either grounded or connected to the cathode. Secondary emission still occurs, but the electrons that come from it can no longer reach the screen grid as they have less energy than the primary electrons, and, hence, cannot pass the grounded suppressor grid. The electrons from the secondary emission are re-collected by the anode.

Pentodes, therefore, can have higher current outputs and a wider output voltage swing; the anode/plate can even be at a lower voltage than the screen grid yet still amplify well .

Comparisons with the triode
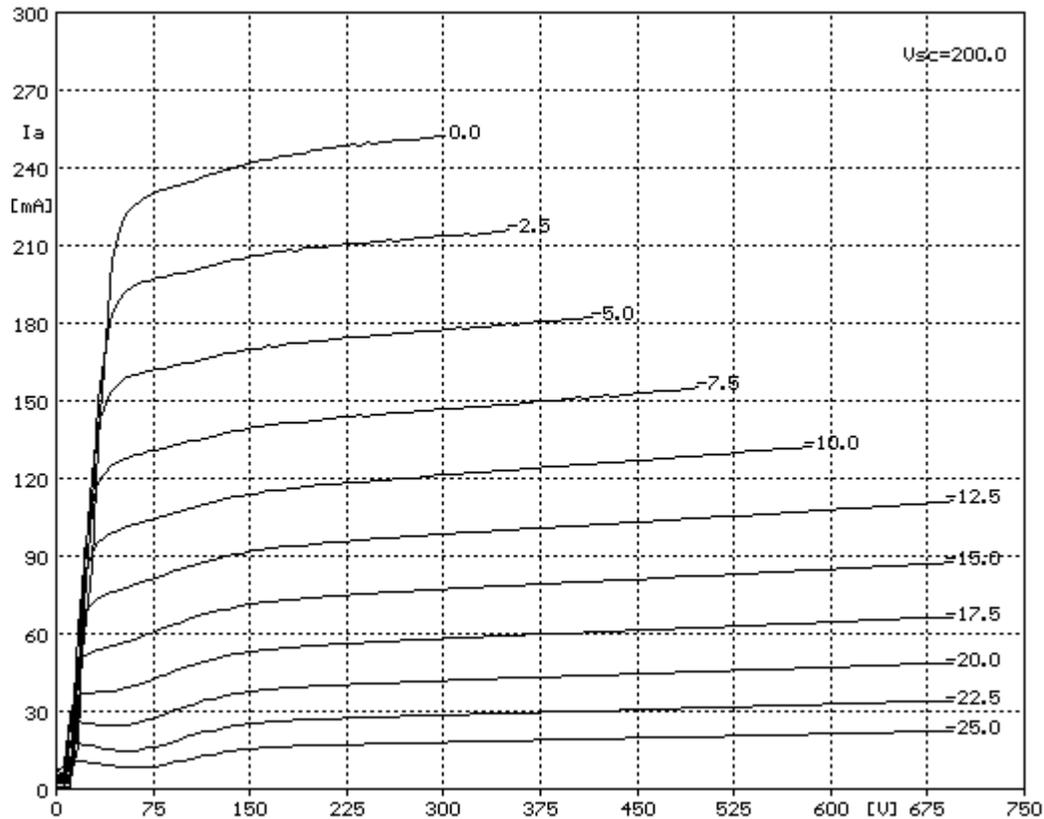
Example pentode current/voltage characteristics for differing control grid voltages.

- Pentodes (and tetrodes) tend to have a much lower feedback capacitance, due to the screening effect of the second grid.
- Pentodes tend to have a higher noise,
- Triodes have a lower internal anode resistance, and hence higher damping factor when used in audio output circuits, compared with pentodes, when negative feedback is absent. That also reduces the potential voltage amplification obtainable from a triode compared with a pentode of the same transconductance, and usually means a more efficient output stage can be made using pentodes, with a lower power drive signal.
- Pentodes are almost unaffected by changes in supply voltage, and can thus operate with more poorly stabilised supplies than triodes. An examination of the characteristics shown here will show that the plate current hardly changes as plate voltage varies.

Usage

A General Electric 12AE10 double pentode

Pentode valves were first used in consumer-type radio receivers. A well-known pentode type, the EF50, was designed before the start of the World War II, and was extensively used in radar sets and other military electronic equipment. The pentode contributed to the electronic preponderance of the Allies. After World War II, pentodes were widely used in TV receivers, particularly the successor to the EF50, the EF80. Vacuum tubes were replaced by transistors during the 1960s. However, they continue to be used in certain applications, including high-power radio transmitters and (because of their well-known valve sound) in high-end and professional audio applications, microphone preamplifiers and electric guitar amplifiers. Large stockpiles in countries of the former Soviet Union

have provided a continuing supply of such devices, some designed for other purposes but adapted to audio use, such as the GU-50 transmitter tube.

Triode-strapped pentode circuits

A pentode (or, less commonly, a tetrode) can have its screen grid (grid 2) connected to the anode (plate) and the resulting "triode-strapped" (or "triode-connected") device has characteristics very similar to a triode (lower anode resistance, lower noise, more drive voltage required). This is sometimes provided as an option in audiophile pentode amplifier circuits, to give the sought-after "sonic qualities" of a triode power amplifier. There are situations where this arrangement is unsafe, for example when doing so (without a series resistor) could exceed the screen grid's power or voltage rating, but remains a valuable option due to the difficulty in obtaining good modern power triodes.

SY4307A

SY4307A

Standard Telephones and Cables brand
Australian Pentode, type SY4307A, c1940

Classification:	Power Output Pentode
Service:	Class C amplifier,

(Radio Frequency)

Cathode	
Cathode type:	Directly heated
Filament voltage:	6.3v
Filament current:	1.16A
Anode	
Socket connections	

Anerican 5 Pin, (UX5)

with side locating pin between pis 1 and 2.

Pin 1, Heater/Cathode

Pin 2, Screen Grid, g2

Pin-3, Control Grid, g1

Pin-4, Supressor Grid

Pin-5, Heater/Cathode

Top Cap, Anode/Plate

The type **SY4307A** is a power output pentode possessing a similar power ratings, but significantly different characteristics to the far more common type 807 thermionic valve/vacuum tube.

Uses

Although the 807 is used for both audio amplification, (usually class AB, and radio frequency amplification in the 1-30 MHz range, (usually class C), this type was usually used for the latter.
This tube/valve is a pentode with a 'directly heated cathode', meaning that the cathode consists of an emissive oxide coated filament. These features suggest the SY4307A could predate and is contemporary with, the 807 which is a beam power tetrode, a more advanced design with a more linear transfer characteristic and a more complex 'indirectly heated cathode'.

Similarity to 807

If the pinout of the SY4307A is compared to that of the 807 the similarity is striking. All pins are assigned equivalent electrodes with the exception of pin 4, which is the suppressor grid, g3, in this type but in the 807 pin 4 is the connected to equivalent of a suppressor grid, the 'beam plates' and also the separate indirectly heated cathode. It is possible to interchange the two with the relatively minor alteration involving the linking of pins 4 and 5 and ensuring the heater supply is "floating". If the heater is supplied from an A.C. source, linking pin 4 to the centre of that source and the cathode bootstrap minimizes hum pickup.

Historical significance

The SY4307A was used extensively during World War II by the Australian armed forces and the example depicted is marked on the Bakelite collar of its base with the "D broad arrow D" commonwealth department of defense mark.

Two SY4307As were used, wired in parallel, as the output stage, or "final", in a transmitter constructed clandestinely by the Australian solders of Sparrow Force in Japanese occupied Portuguese Timor in 1942. This transmitter became known affectionately as "Winnie the War Winner", named after the British wartime Prime Minister, Winston Churchill.

Side view of SY4307A

View of directly heated cathode of the SY4307A

View of UX5 Base of the SY4307A

Krytron & Space Charge

Krytron

KN2 "Krytron" switch tube, made by EG&G (ca. 25 mm tall)

The **krytron** is a cold-cathode gas filled tube intended for use as a very high-speed switch, and was one of the earliest developments of the EG&G Corporation. It is somewhat similar to thyratron. The vacuum version is called **vacuum krytron**, or **sprytron**.

Description

Unlike most other gas switch tubes, the krytron uses arc discharge to handle very high voltages and currents (several kV and several kA peak), rather than the usual low-current glow discharge. The krytron is a development of the triggered spark gaps and thyratrons originally developed for radar transmitters during World War II.

The gas used in krytrons is hydrogen; noble gases (usually krypton, or a Penning mixture) can also be used.

Operation
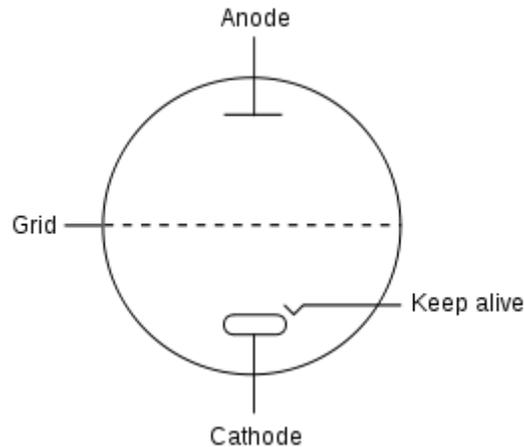
There are four electrodes in a krytron. Two are conventional anode and cathode. One is a keep-alive electrode, arranged to be close to the cathode. The keep-alive has a low positive voltage applied, which causes a small area of gas to ionize near the cathode. High voltage is applied to the anode, but primary conduction does not occur until a positive pulse is applied to the trigger electrode ("Grid" in the image above). Once started, arc conduction carries a considerable current.

The control grid is usually wrapped around the anode, except a small opening on its top.

In place of or in addition to the keep-alive electrode some krytrons may contain a very tiny amount of radioactive material (usually less than 5 microcurie of nickel-63) which emits beta particles (high-speed electrons) to make ionization easier. The amount of radiation in a krytron is very small and not harmful. The radiation source serves to increase the reliability of ignition and formation of the keepalive electrode discharge. The keepalive discharge creates a preionized plasma between the cathode and anode; the time for forming of the arc is considerably shortened as the arc does not have to first create an

ionized channel, which achieves switching times 3–4 orders of magnitude lower than with other cold-cathode tubes.

The gas filling provides ions for neutralizing the space charge and allowing high currents at lower voltage. The keepalive discharge populates the gas with ions, forming a preionized plasma; this can shorten the arc formation time by 3–4 orders of magnitude in comparison with non-preionized tubes, as time does not have to be spent on ionizing the medium during formation of the arc path.

The electric arc is self-sustaining; once the tube is triggered, it conducts until the arc is interrupted by the current falling too low for too long (under 10 milliamperes for more than 100 microseconds for the KN22 krytrons).

Krytrons and sprytrons are triggered by a high voltage from a capacitor discharge via a trigger transformer, in a similar way flashtubes for e.g. photoflash applications are triggered. Devices integrating a krytron with a trigger transformer are available.

Sprytron

Sprytron, also known as **vacuum krytron**, is a vacuum-"filled" version. It is designed for use in environments where high levels of ionizing radiation are present (because the radiation might cause the gas-filled krytron to trigger inadvertently.)

Sprytrons lack the keepalive electrode and the preionization radioactive source. The trigger pulse for the grid must be stronger than for a krytron. Sprytrons are however able to handle higher currents; krytrons tend to be used for triggering a secondary switch, e.g. a triggered spark gap, sprytrons are usually connected directly to the load.

The trigger pulse has to be much more intense, as there is no preionized gas path for the electric current, and a vacuum arc must form between the cathode and anode. An arc first forms between the cathode and the grid, then a breakdown occurs between the cathode-grid conductive region and the anode.

Sprytrons are evacuated to hard vacuum. As kovar and other metals are somewhat permeable for hydrogen, especially during the 600 °C bakeout before evacuation and sealing, all external metal surfaces have to be plated with thick (25 micrometers or more) of soft gold. The same metallization is used for other switch tubes as well.

Sprytrons can be also triggered by a laser pulse. In 1999, the laser pulse energy needed to trigger a sprytron was reduced to 10 microjoules.

Sprytrons are usually designed as rugged metal/ceramic parts. Sprytrons typically have low inductance (10 nanohenry) and low electrical resistance when switched on (10–30 milliohms). After triggering, just before the sprytron switches on in avalanche mode, it becomes slightly conductive (100–200 amperes) briefly. High-power MOSFET

transistors operating in avalanche mode show similar behavior. SPICE models for sprytrons are available.

Performance

This design, dating from the late 1940s, is still capable of pulse-power performance which even the most advanced semiconductors (even IGBTs) cannot match easily. The krytrons and sprytrons are capable of handling high current high voltage pulses, with very fast switching times, constant low time delay between application of the trigger pulse and switching on, and a low jitter of this delay.

Krytrons can switch currents up to about 3000 amperes and voltages up to about 5000 volts. The commutation time of less than 1 nanosecond can be achieved, with the delay between the application of the trigger pulse and switching achievable as low as about 30 nanoseconds. The achievable jitter may be below 5 nanoseconds. The required trigger pulse voltage is about 200–2000 volts. Higher-voltage pulse can to a degree decrease the switching delay. The switching delay is strongly independent on the environment. The commutation time can be somewhat shortened by a faster rise time of the trigger pulse. A given krytron tube, with identical trigger pulse, will show very similar performance (low jitter) between the pulses. The keepalive current ranges in tens to hundreds microamperes. The pulse repetition rate can range from one per minute to tens of thousands per minute.

The switching performance is largely independent on the environment (temperature, acceleration, vibrations, etc.). The formation of the keepalive glow discharge is however more sensitive, which necessitates the use of the radioactive source to aid its ignition.

Krytrons have a limited lifetime, counted in the number of pulses they can switch. The number ranges by type between tens of thousands and tens of millions, but can be as low as few hundreds.

Sprytrons have somewhat faster switching times than krytrons.

Hydrogen-filled thyratrons may be used as a replacement in some applications.

Applications

Krytrons and their variations are still manufactured by Perkin-Elmer Components, and used in a variety of industrial and military devices. They are best known for their use in igniting the exploding-bridgewire detonators and slapper detonators in nuclear weapons, their original application, either directly (sprytrons are usually used in such manner) or by triggering the higher-power spark gap switches. They are also used to trigger thyratrons, large flashlamps in photocopiers, lasers and scientific apparatus, as well as firing ignitors for industrial explosives.

Export restrictions

Because of the potential for use as nuclear triggers, the export of krytrons is tightly regulated. A number of cases involving the smuggling or attempted smuggling of krytrons have been reported, as countries seeking to develop nuclear weapons have attempted to procure supplies of krytrons for igniting their weapons. One prominent case was Richard Kelly Smyth.

Only some types of krytrons and sprytrons are suitable for nuclear weapon triggers. Typical values considered relevant for nuclear weapons are voltages 2500 V and above, currents 100 A and above, and switching delays of under 10 microseconds.

Space Charge

Space charge is a concept in which excess electric charge is treated as a continuum of charge distributed over a region of space (either a volume or an area) rather than distinct point-like charges. This model typically applies when charge carriers have been emitted from some region of a solid—the cloud of emitted carriers can form a space charge region if they are sufficiently spread out, or the charged atoms or molecules left behind in the solid can form a space charge region. Space charge usually only occurs in dielectric media (including vacuum) because in a conductive medium the charge tends to be rapidly neutralized or screened. The sign of the space charge can be either negative or positive. This situation is perhaps most familiar in the area near a metal object when it is heated to incandescence in a vacuum. This effect was first observed by Thomas Edison in light bulb filaments, where it is sometimes called the Edison effect, but space charge is a significant phenomenon in many vacuum and solid-state electronic devices.

Cause

Physical explanation

When a metal object is placed in a vacuum and is heated to incandescence, the energy is sufficient to cause electrons to "boil" away from the surface atoms and surround the metal object in a cloud of free electrons. This is called thermionic emission. The resulting cloud is negatively charged, and can be attracted to any nearby positively charged object, thus producing an electrical current which passes through the vacuum.

Space charge can result from a range of phenomena, but the most important are:

1. Combination of the current density and spatially inhomogeneous resistivity
2. Ionization of species within the dielectric to form heterocharge
3. Charge injection from electrodes and from a stress enhancement
4. Polarization in structures such as water trees

It has been suggested that in AC most carriers injected at electrodes during a half of cycle are ejected during the next half cycle, so the net balance of charge on a cycle is practically zero. However, a small fraction of the carriers can be trapped at levels deep enough to retain them when the field is inverted. The amount of charge in AC should increase slower than in DC and become observable after longer periods of time.

Hetero and Homo Charge

Hetero charge means that the polarity of the space charge is opposite to that of neighboring electrode, and homo charge is the reverse situation. Under high voltage application, a hetero charge near the electrode is expected to reduce the breakdown voltage, whereas a homo charge will increase it. After polarity reversal under ac conditions, the homo charge is converted to hetero space charge.

Mathematical explanation

If the "vacuum" has a pressure of 10^{-6} mmHg or less, the main vehicle of conduction is electrons. The emission current density (J) from the cathode, as a function of its thermodynamic temperature T, in the absence of space-charge, is given by:

$$J = (1 - \tilde{r})A_0 T^2 \exp\left(\frac{-\phi}{kT}\right)$$

where

$A_0 = \dfrac{4\pi emk^2}{h^3} \approx 1.2 \times 10^6$ A m^{-2} K^{-2}
e = elementary positive charge (i.e., magnitude of electron charge),
m = electron mass,
k = Boltzmann's constant = 1.38 x 10^{-23} J/K,
h = Planck's constant = 6.55 x 10^{-34} J s,
φ = work function of the cathode,
$\check{r}$ = mean electron reflection coefficient.

The reflection coefficient can be as low as 0.105 but is usually near 0.5. For Tungsten, $(1 - \check{r})A_0 = 0.6$ to 1.0×10^6 A m^{-2} K^{-2}, and $\varphi = 4.52$ eV. At 2500 °C, the emission is 3000 A/m^2.

The emission current as given above is many times greater than that normally collected by the electrodes, except in some pulsed valves such as the cavity magnetron. Most of the electrons emitted by the cathode are driven back to it by the repulsion of the cloud of electrons in its neighborhood. This is called the *space charge effect*. In the limit of large current densities, J is given by the Child-Langmuir equation below, rather than by the thermionic emission equation above.

Occurrence

Space charge is an inherent property of all vacuum tubes. This has at times made life harder or easier for electrical engineers who used tubes in their designs. For example, space charge significantly limited the practical application of triode amplifiers which lead to further innovations such as the vacuum tube tetrode.

On the other hand, space charge was useful in some tube applications because it generates a negative EMF within the tube's envelope, which could be used to create a negative bias on the tube's grid. Grid bias could also be achieved by using an applied grid voltage in addition to the control voltage. This could improve the engineer's control and fidelity of amplification.

Space charges can also occur within dielectrics. For example, when gas near a high voltage electrode begins to undergo dielectric breakdown, electrical charges are injected into the region near the electrode, forming space charge regions in the surrounding gas. Space charges can also occur within solid or liquid dielectrics that are stressed by high electric fields. Trapped space charges within solid dielectrics are often a contributing factor leading to dielectric failure within high voltage power cables and capacitors.

Child's Law
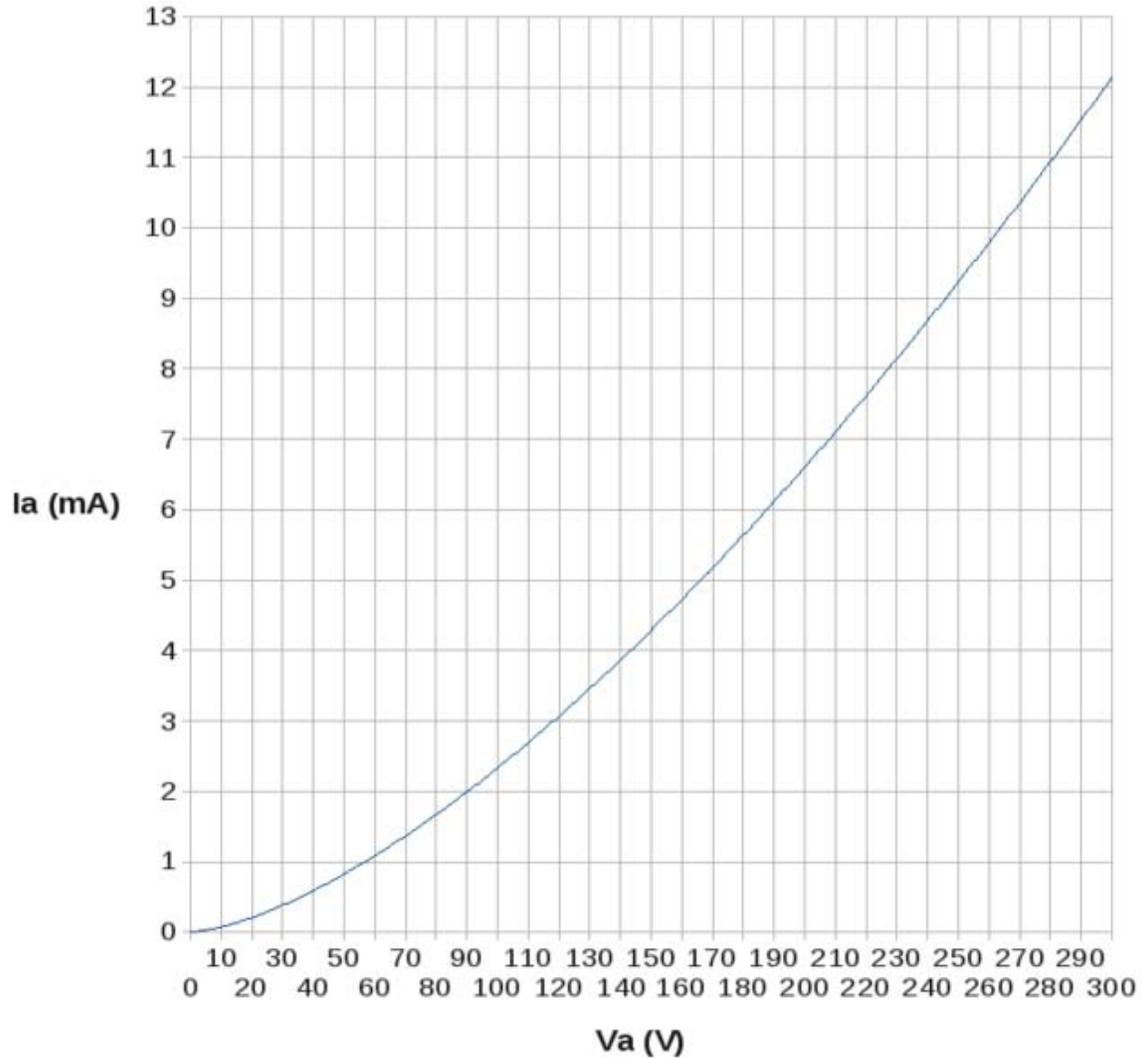
Graph showing Child-Langmuir Law. S and d are constant and equal to 1.

Also known as the Child-Langmuir Law or the Three-Halves Power Law, Child's Law states that the space-charge limited current in a plane-parallel diode varies directly as the three-halves power of the anode voltage V_a and inversely as the square of the distance d separating the cathode and the anode. That is,

$$I_a = JS = \frac{4\epsilon_0}{9}\sqrt{2e/m_e}\frac{SV_a^{3/2}}{d^2}.$$

Where I_a is the anode current, J the current density, and S the area. This assumes the following:

1. The electrodes are planar, parallel, equipotential surfaces of infinite dimensions.

2. The electrons have zero velocity at the cathode surface.
3. In the interelectrode region, only electrons are present.
4. The current is space-charge limited.
5. The anode voltage remains constant for a sufficiently long time so that the anode current is steady.

Mott-Gurney Law

In the low-field regime, velocity of injected carriers can be represented by

$$v = \mu \mathcal{E}$$

Where $\mathcal{E}$ is the applied electric field, μ is the carrier mobility, and v is the carrier velocity. If the current is limited by the drift component of inject carriers, the space-charge-limited conduction current density J can be written as

$$J = \frac{9 \epsilon \mu V_a^2}{8 L^3}$$

where V_a is the applied voltage, and L is the length of the plane-parallel sample. This expression is known as the Mott-Gurney law.

In the velocity-saturation regime, this equation takes the following form

$$J = \frac{2 \epsilon v V_a}{L^2}$$

Note the different dependence of J on V_a in each of the two cases. Interestingly, in the ballistic case (assuming no collisions,) the Mott-Gurney equation takes the form of the more familiar Child-Langmuir law.

It should be noted that the above derivations make the following assumptions:

1. There is only one type of charge carrier present.
2. The material has no intrinsic conductivity, but charges are injected into it from one electrode and captured by the other.
3. The carrier mobility μ and the dielectric permittivity ε are constant throughout the sample.
4. The electric field at the charge-injecting cathode is zero.

As an application example, the steady-state space-charge-limited current across a piece of silicon with a charge carrier mobility of 1500 cm^2/V-s, a dielectric constant of 11.9, an area of 10^{-8}cm^2 and a thickness of 10^{-4}cm can be calculated by an on line calculator as 126.4mA at voltage 3V.

Shot noise

Space charge tends to reduce shot noise. Electrons (and positive charge carriers) come with their own built-in negative feedback.