Introduction to Thermionic Valves (Vacuum Tubes)

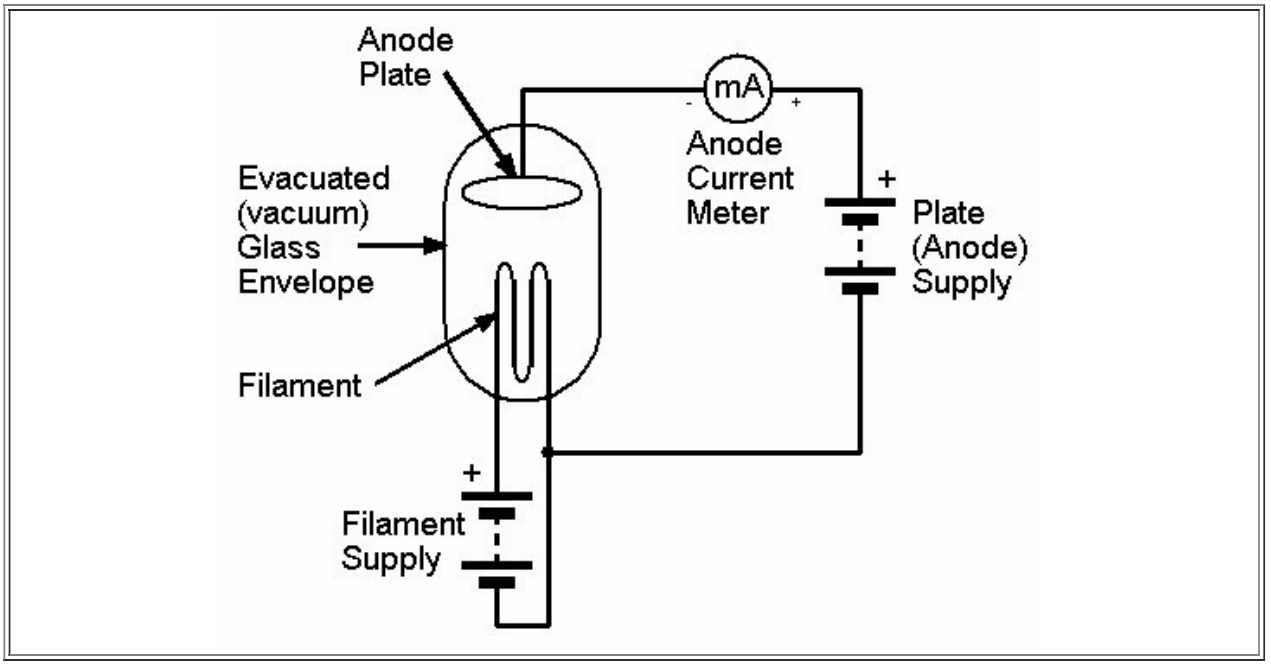
The Diode Valve

The diode valve is basically little more than an electric light bulb (in the form of a wire filament heated by an electric current) with the addition of a metal plate, known as the anode. Negatively charged electrons are given off by the hot filament, and if the anode is connected to a voltage that is positive relative to the filament, the electrons are attracted to the anode creating an electric current. What is special about the diode is that electrons can only flow in one direction, as electrons are not given off by the cold anode so can not complete the circuit when the cathode is positive relative to the anode. This is therefore a device that can rectify alternating current (AC) producing current that may vary with the amplitude of the AC signal, but always flows in one direction, and with suitable filtering, this can be converted into steady direct current (DC), for example allowing power to be distributed as AC then converted to DC for use.

Of course, there must be a hard vacuum (about one millionth of normal air pressure at sea level) inside the bulb (modern light bulbs often contain gas), otherwise gas molecules would block the movement of electrons. To preserve this hard vacuum, after a thermionic valve has had all the air pumped out of it, a "getter" made of reactive metal such as barium is fired by means of an electrical discharge and/or RF induction heating through the glass wall, which creates a silvery coating on the inside of the glass that will absorb any remaining gas molecules, including ones that can be released from the valve materials.

British scientist Sir John Ambrose Fleming (1848-1945) is attributed with the invention of the vacuum diode. In fact, the first person to use a metal plate inside an electric light bulb as an anode was Thomas Alva Edison (1847-1931), recipient of the Congressional Gold Medal in 1928, who in 1883 was trying to stop a blackening of electric lamps by putting a metal plate inside, which he connected to one end of the light filament. He found that a current flowed when the plate was connected to the positive end of the filament, but not when it was connected to the negative end. Two German scientists - Julius Elster and Hans Friedrich Geitel, extended the experiment by adding a battery between the filament and the plate, demonstrating that current flow required a positive voltage on the plate.

Edison patented the discovery in 1883, but did not find a practical application. This became known as the Edison effect. He had stumbled across the basic valve diode rectifier, as shown below:



Edison's work preceded the discovery of electrons by Joseph J. Thomson, who proposed their theoretical existence and in 1897 demonstrated that they could be attracted or repelled by an electric field. The concept of a flow of charged electrons could not therefore have been easily envisaged by Edison in 1883.

But as Sir John Ambrose Fleming once said, "The little things of today may develop into the great things of tomorrow"

Fleming was employed by the Edison Company as a consultant, and in 1885 he was invited to become the first Professor of Electrical Engineering at University College, London (UCL), where he founded England's first Electrical Engineering Department.

In 1899 he became a consultant to the Marconi Company, designing the equipment which Guglielmo Marconi used to make the first transatlantic radio transmission in 1901.

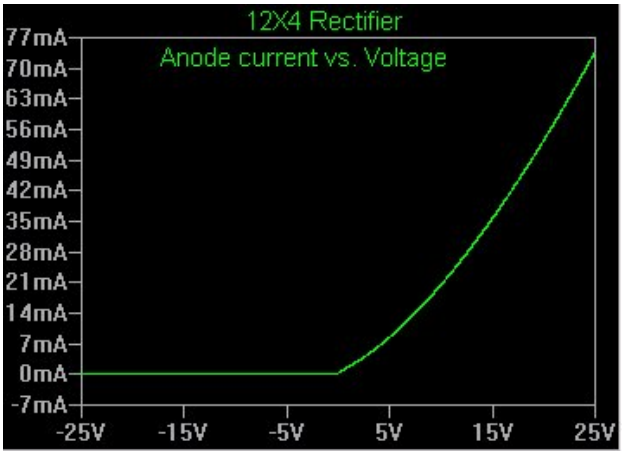
It was because of his involvement with Marconi radio, that... eventually... Fleming realised that the Edison Effect might be capable of providing a better detector of radio waves (by rectifying AC radio signals, the pulses of Morse Code or audio applied to the radio signals by amplitude modulation may be detected).

Using one of the Edison effect experimental lamps from the 1880s, he found that it worked, and applied for a patent for the first practical vacuum-tube rectifier on November 16th, 1904 (although the first solid-state detector was to follow only 6 years later- the cat's whisker crystal device invented by B.F. Miessner in 1910).

Fleming was knighted in 1929 for his contributions to science and industry.

Examples of diode rectifier valves still manufactured today include the 5AR4, 5U4, 5V4, and 5Y3.

The graph below shows the anode current/voltage characteristic of a 12X4 (EZ90) rectifier, showing current flow in forward bias, but no current in reverse bias. The effect is easily understood as the negatively charged electrons emitted by the cathode are the only current carriers, and given that the cold anode does not emit electrons, when the anode is negative relative to the cathode, the electrons are repelled from the anode.



The 12X4 (EZ90) is a miniature glass 7-pin B7G base double diode with a common cathode for full-wave rectification, was introduced in 1945, is 2 5/8 inches tall, 7/8 inch in diameter, has a 1375V peak inverse voltage, and 75mA maximum mean anode current.

The anode current is approximated by the Child-Langmuir law:

$Ia = P * Va^{(3/2)}$

where Va = anode voltage,
P is a constant known as perveance, which is characteristic for a particular valve.
P includes an inverse dependence on the square of the cathode-anode spacing d, and is proportional to cathode area A, i.e:

$P = K * A / d^2$

At higher anode voltages, the current eventually levels out as there is a limit to the number of electrons available from the cathode, and when this limit is reached, the valve is said to be in saturation. At lower currents, electrons tend to collect in a "space charge" around the cathode, a cloud of electrons whose mutual repulsion (electric charges of the same sign repel) suppresses further emission from the cathode unless a positively charged electrode starts to remove them.

The filament cathode, as in the Edison lamp, emits electrons from a directly heated wire surface, as soon as the thermal energy added to the electrons exceeds an energy level characteristic of the material called the work function. A tungsten filament has a work function of 4.52 electron-volts (eV), and with thorium added ("thoriated filament"), the work function is lowered to 2.6eV, reducing the heat required. A good example of a valve with a directly heated filament cathode is the 300B triode, manufactured since the 1930s and still in use today (we will discuss the triode in more detail later).

The following diagram (to the left) shows the structure of a directly heated thoriated filament valve of the 1920s-1930s (similar to a Mullard ORA). To the right is a picture of a "modern" indirectly heated cathode valve for comparison.

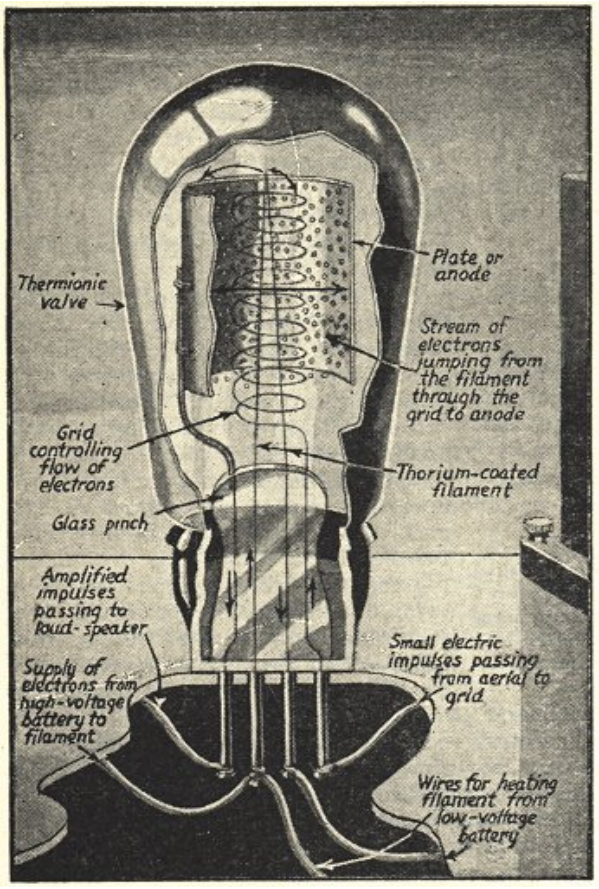


Illustration of the construction of a directly heated thoriated filament valve of the 1920s-1930s. The internals are similar to a Mullard ORA, which is 32 mm x 80 mm, but the envelope shown is similar to that of the MOV/Mazda PX25 in a "balloon envelope" which is 60mm in diameter and 137mm tall excluding pins.



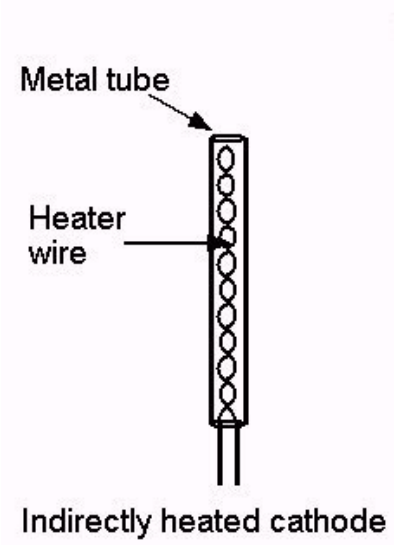
Picture of a PCF802, which includes a triode and a pentode designed for television line oscillator applications, a model introduced no later than 1962, and in common use in the 1960s-1970s. The post-WW2 style B9A ("Noval") base (from 1948, when introduced with the 12AU7, 12AX7 and 12AT7) is made with Kovar alloy pins, matched for thermal expansion to the glass in which they are sealed without the need for a separate holder, and the glass envelope is 20 mm in diameter and 46 mm tall excluding pins.

Although this illustration describes "Amplified impulses passing to loud-speaker", in practice a transformer is required between the valve anode circuit and a loudspeaker, as the valve anode circuit is high impedance (high voltage, low current), whereas a loudspeaker is low impedance, requiring high current at low voltage.

Most modern valves use an indirectly heated cathode, which consists of a metal cylinder containing a wire heater element that is separated from the cathode by a high-temperature insulation material such as aluminium oxide, so that the heater can be independently powered (for example from a low-voltage AC winding on the mains transformer) while the cathode is connected to its own DC bias potential (e.g. via a self-bias resistor to ground).

Their major disadvantage is that they are more sensitive to "cathode poisoning", which is a loss of electron emission caused by absorbing impurity ions present inside the valve. Secondly, it is sometimes possible for the insulation to break down, causing a "heater-cathode short", both well known causes of failure of domestic valve-based radio and television- which usually required the valve to be replaced.

The following diagram illustrates the indirectly heated cathode.



The cathode cylinder (typically of nickel) is coated with a special material that does not need to be heated as much as a carbon or tungsten lamp filament in order to emit electrons. For example, barium and strontium oxide gives a work function of 1.1 eV.

The oxide layer is formed by first coating the nickel tube with a mixture of barium and strontium carbonates, which is then baked to remove binder material and reduce the carbonates to oxides.

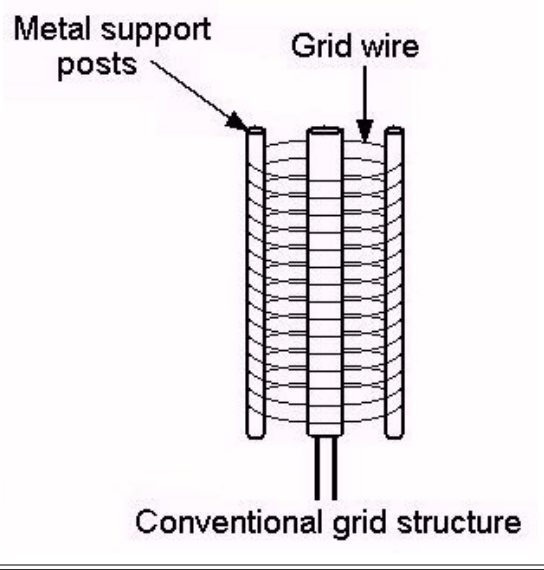
The effectiveness of an oxide cathode for electron emission is related to an excess of the metal atoms in the crystal structure. Barium is an excellent donor of electrons, which results in the cathode coating acting rather like an n-type semiconductor.

Directly heated cathodes can also be coated with oxide, but it is a must for indirectly heated ones because otherwise it would be hard to maintain sufficiently high temperatures on their outside in order to emit electrons.

The Triode Valve

In 1906 the American engineer, Lee De Forest (1873-1961) was looking for improvements in radio signal detectors, and added a third electrode, the control grid, which was a zig-zag of fine wire placed close to the cathode in a modified Fleming diode. This was found to give control of the anode voltage using a much lower voltage on the grid, i.e. the device had a voltage gain. However, its performance was poor as he mistakenly believed that some gas was necessary, instead of a good vacuum, for its operation. De Forest applied for a patent for the new device in 1907, which he called an "Audion", which later become the "De Forest valve", and eventually the triode valve as it has three principal electrodes. He was awarded a special Oscar in 1959 for pioneering inventions which brought sound to the motion picture.

The conventional method of grid construction that is used in most valves up to the present day is illustrated below.

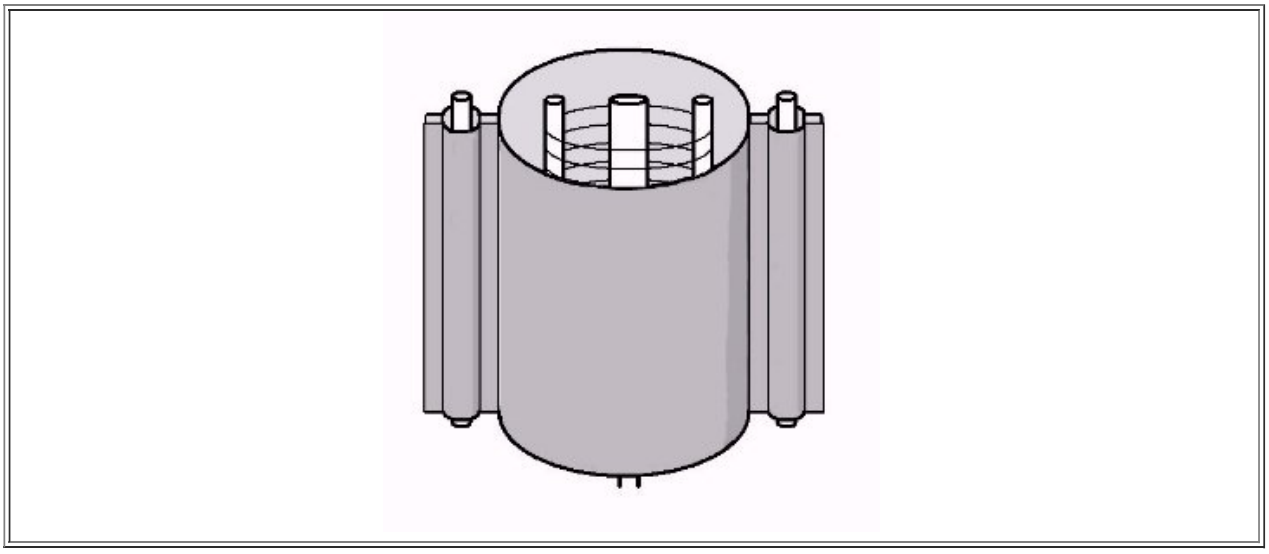


As you can see if you look closely at a typical glass valve from an audio amplifier, the control grid is formed by wrapping a thin wire

around two metal support wires of approximately 0.5-1mm diameter.

The grid wire itself is approximately 1/1000 of an inch in diameter, and may be spaced by only a few 1/1000s of an inch from the cathode. It must be rigid and heat-resistant, and may be plated to minimise secondary electron emission.

The anode commonly takes the form of a cylindrical or oval tube structure, or a box with rounded corners, open at the ends, which surrounds the cathode and grid. It is normally welded or crimped to support wires. The support wires of anode and grid, as well as the cathode cylinder, are normally located in matching holes in insulating washers made of a material such as mica, which is rigid, has good insulating properties, and does not release ions when in a vacuum. The mica washer is normally made a close fit to the valve envelope to act as a support.



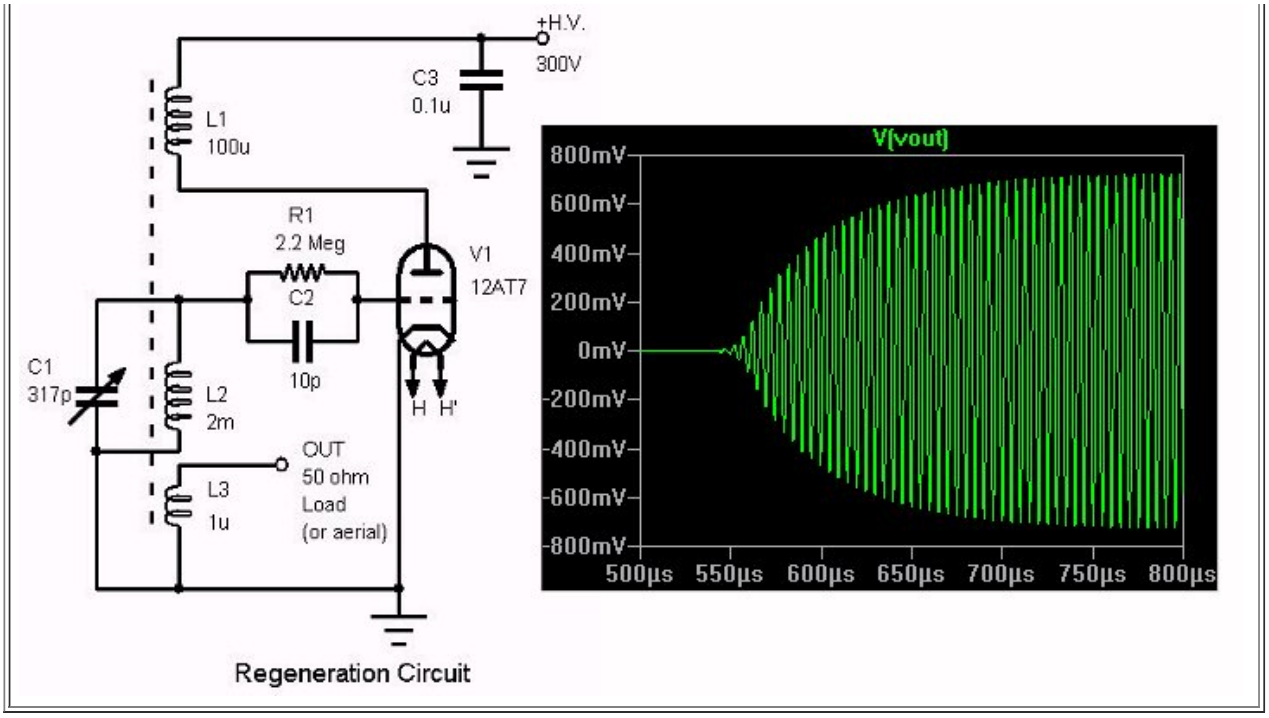
Anodes are often surface treated to make them a dull grey or dark colour to radiate heat more effectively, and may be formed with fins on the outside.

If the grid voltage is the same as the cathode, i.e. electrically neutral, and as the spacing between grid wires is much greater than the wire diameter, when a positive voltage is on the anode, electrons mostly pass straight through the grid, and will arrive at the anode just as if the valve were a diode.

If we apply a negative voltage to the grid, the electrons being negatively charged will be repelled from the grid wires. As the voltage is increased (more negative), eventually the repelling electric field between the grid wires becomes sufficient to cancel out the attractive field from the anode and to cut off the anode current completely.

If we apply a positive voltage to the grid, electrons are attracted to the grid, most passing through to the anode and resulting in increased anode current, but some of them can also now enter the grid wire and flow as a current in the grid circuit.

The ability of the triode valve to amplify signals meant that not only could weak signals be amplified as much as necessary to drive a loudspeaker, but if the output of the valve was fed back to its input, it could be made to generate a signal by means of positive feedback, which was called reaction or regeneration. This allowed generation of pure sinusoidal RF signals instead of the ill-defined noisy frequency bands generated by spark-gap transmitters, and enabled voice and music radio transmission. Radio receivers could use a regenerative valve circuit to boost the received signal at the same time as demodulating the signal to audio frequency to drive headphones or a loudspeaker. The self-oscillation when combined with a morse code transmission allowed its reception as a tone generated by the beat frequency between the signal and the regeneration.



In this example circuit, which can generate a 200kHz signal into a 50 ohm load, or be used as a regenerative receiver (with headphones inserted in series with the anode power supply), the output of V1 is fed back to the grid circuit via a magnetic coupling between the anode inductance L1 (which was sometimes called a "tickler") and the tuned circuit inductance L2.

L2 is resonated at 200kHz as a tuned circuit with C1. The tuned circuit is coupled to the grid via C2, and R1 is present to provide grid bias. The negative bias is created because the positive peaks of the AC signal at the grid drive the grid positive with respect to the cathode. This causes current to flow which drops a negative voltage across R1, and C2 serves as a reservoir capacitance to maintain the negative DC bias throughout the RF cycle (The time constant of R1 and C2 is 22 microseconds, and the duration of a 200 kHz cycle is 5 microseconds).

Such circuits are rarely used any more for radio reception, because of the potential of the regeneration oscillations to radiate from an antenna and cause interference to other users of the radio spectrum.

The electrical characteristics of the triode depend on the geometry of the electrodes and properties of the materials used. Data sheets normally list characteristics in the form of tabulated numbers (e.g. maximum anode voltage, nominal anode voltage), and graphs showing the characteristic relationships between two or more parameters, such as anode voltage and current.

Ratings

With valve ratings, as the rating systems used have changed over time, care must be taken to understand what kind of rating method is being used.

- **Absolute Maximum Ratings** (any year): As for solid state, these must not be exceeded under any circumstances, including margins for the worst-case combination of valve characteristic tolerances, power supply variations, component tolerances, etc. However, it is expected that parameters can go above the limits for brief periods without destroying a valve, whereas in solid state when we are given maximum ratings, they are understood to be absolutes that must certainly not be exceeded at high risk of destroying the device.
- **Design Centre Ratings** (1939-1957): Nominal values for an average valve, to be used as the nominal conditions for the design. They already have a built-in margin so that there is negligible risk of damage for normal tolerances, and power supply voltage variations of up to ±10%.
- **Design Maximum Ratings** (from 1957): Must not be exceeded for an average valve under the worst probable combination of component tolerances, supply voltage fluctuations, and signal levels. Anode voltage and power dissipation are about 10% higher than Design Centre Ratings under this system.

Characteristic Curves

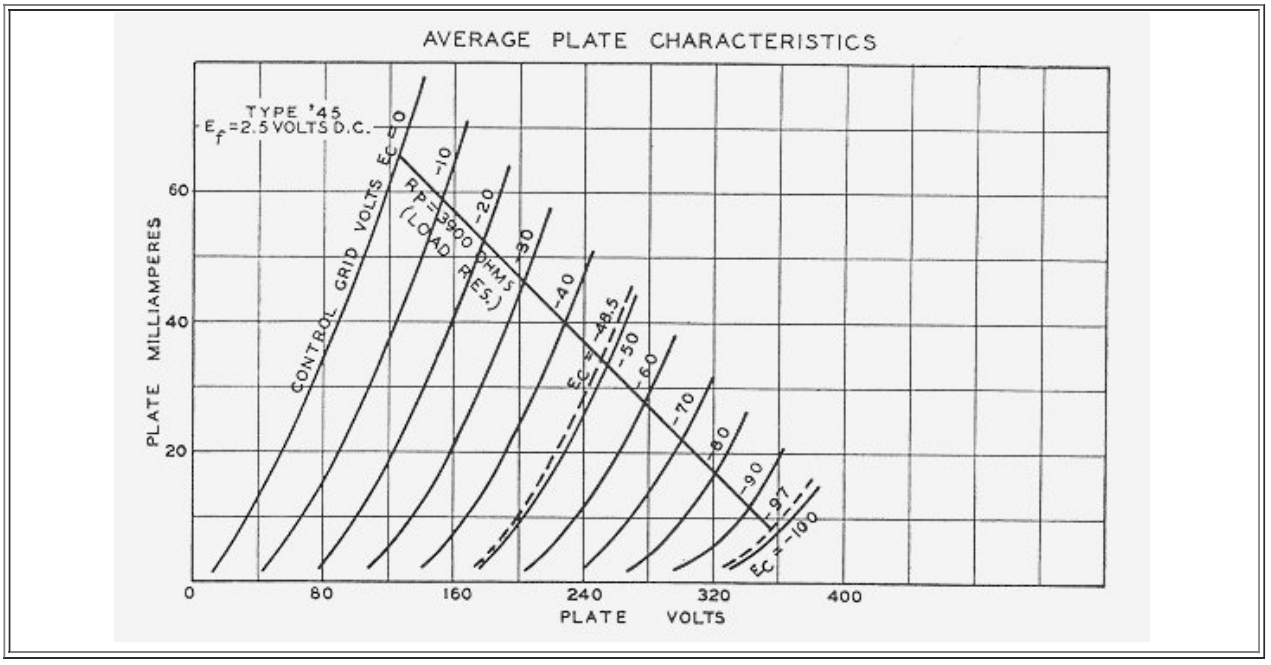
There are a wide variety of possible characteristic plots, but the most commonly useful ones are:

- The anode characteristic, where anode current is plotted on the vertical axis against applied anode voltage on the horizontal axis. A family of curves is normally plotted, each one for a different constant value of grid voltage.

- The mutual characteristic, where anode current is plotted on the vertical axis against applied grid voltage on the horizontal axis. A selection or a family of curves may be plotted, each one for a different constant value of anode voltage.

Nowadays, SPICE modelling is used in preference to older methods of graphical analysis using characteristic curves, but they remain useful for comparing and selecting valves and for selecting the load line and operating point for power driver and output stages. Characteristic curves for individual valves do vary from the nominal ones quoted on data sheets, so allowance should be made in designs (for example, by designing the operating point such that a certain amount of variation is tolerable, or including an adjustable bias, or by using a feedback bias method such as cathode self-bias to provide a degree of self-correction against variations).

The anode current/voltage characteristic curves (each curve is for a different grid bias voltage) are shown below for an RCA-245 triode. The single line crossing them all is a "load line":



The Load Line

The anode characteristic is usually given for an anode connection directly to a test voltage source with no load resistor in series, but in normal operation a load is inserted in the anode circuit. To relate this to the anode characteristic, a "load line" may be drawn on the graph. This is simply a plot of the current that would flow in the load resistance if it is connected between the anode voltage on the horizontal axis and the power supply.

The current in the load resistor is given by:

I(Rl) = (Vsupply - Va)/Rl
I(Rl) = Vsupply/Rl - Va/Rl

Where I(Rl) is the current in the load resistor,
Vsupply is the power supply voltage,
Va is the anode (plate) voltage,
and Rl is the value of load resistance.

If the anode voltage on the horizontal axis were simply connected across the resistor, the current flow would be, according to Ohm's law:

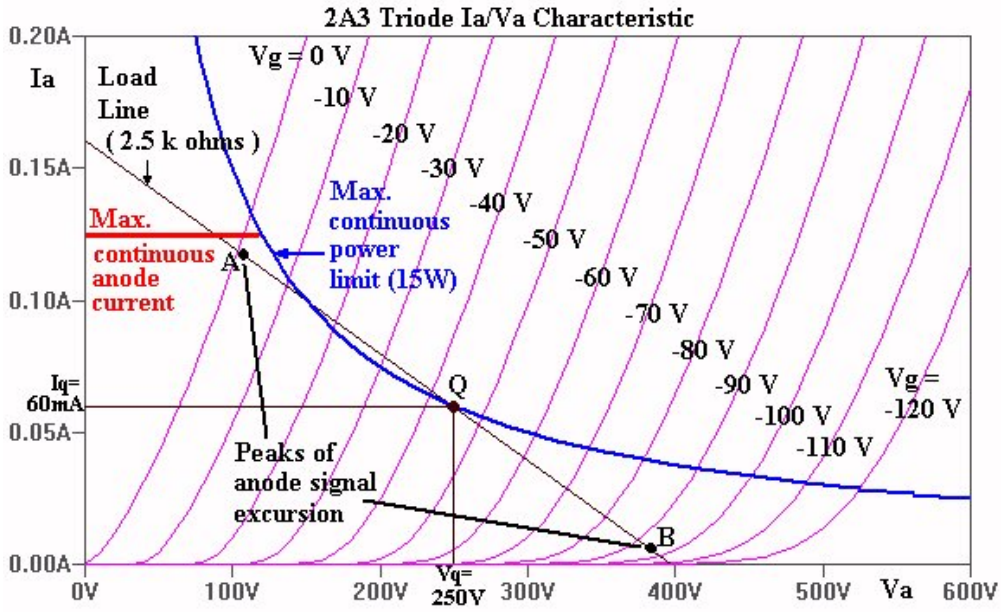
I(Rl) = Va/Rl

Note that the load line has a minus sign in the term "-Va/Rl" which implies a negative slope equivalent to the conductance, -1/Rl, (not resistance, because the valve characteristics are conventionally drawn with the current axis vertical and voltage axis horizontal).

It is a straight line as for Ohm's law, but instead of passing through the origin at Va=0, it passes through the vertical axis at Ia=Vsupply/Rl, and passes through the horizontal axis at the supply voltage, as the anode voltage can only equal the full supply voltage when the voltage drop in the load resistor is zero.

For any anode circuit with a load resistance, the anode current and voltage must lie on both the relevant anode characteristic curve and the load line, i.e. where they intersect. As the grid voltage varies, the intersection point moves along the load line.

A load line overlaid on the anode current/voltage characteristic curves for a 2A3 triode is shown below:



The 2A3 is a power amplifier triode introduced around 1931 (and still manufactured today), for typical use as the output stage of AC-powered radio receivers. It has a medium 4-pin base, a 2.5V, 2.5A directly heated filament cathode, normal anode voltage 250V, normal grid voltage -45V, anode current of 60mA, anode resistance (nominal Class A use) of 800 ohms, amplification factor of 4.2, transconductance of 5.25mS, and output power of 3.5W.

From typical operating data for the 2A3 as a Class A amplifier (drawing anode current throughout the AC cycle and with grid voltage always negative relative to the cathode), the anode voltage is 250V and the zero-signal anode current is 60mA. The grid bias is -45V. This allows us to plot the quiescent operating point Q on the above graph. This is right on the maximum power curve, which is a design centre value. The maximum anode current is 125mA, which I have marked with a horizontal red line. This represents the maximum allowed excursion of the signal, except that the zero voltage grid bias line occurs before the signal gets there, and the input voltage should not exceed this value under normal Class A operating conditions.

The recommended load resistance is given as 2500 ohms, and power output is given as 3.5 W. Given the operating point Q, at zero anode voltage the anode current intercept for the vertical axis should therefore be:

Ia(Va=0) = Iq + 250/2500 = 160mA.

The load line can now be drawn through these two points, and you can see that it intercepts the horizontal axis at Va=400V. This is also the voltage given by adding the voltage drop across the load resistor, i.e.

Va(Ia=0) = Vq + Iq * 2500 = 400V.

Given the output power of 3.5W, the peak voltage across the load must be given by the power formula, and substituting V_{peak} = sqrt(2) * V_{rms}:

$P = V_{rms}^2 / Rl$

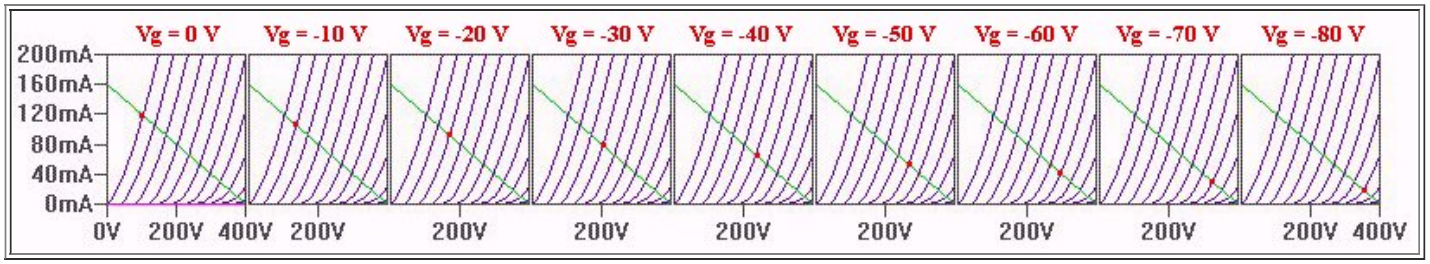
$P = V_{peak}^2 / 2Rl$

which, rearranged, gives:

V_{peak} = sqrt(2 * P * Rl)
V_{peak} = 132 V.

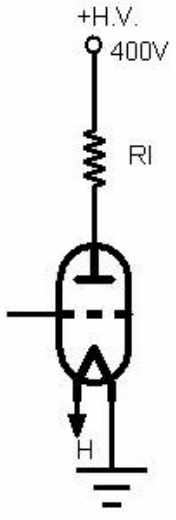
The voltage excursion for the signal at the anode can now be plotted as points A and B on the load line, by subtracting or adding 132 V to the quiescent voltage at point Q. You can see that point A just keeps to within zero grid voltage, and point B is just short of cutting off the anode current.

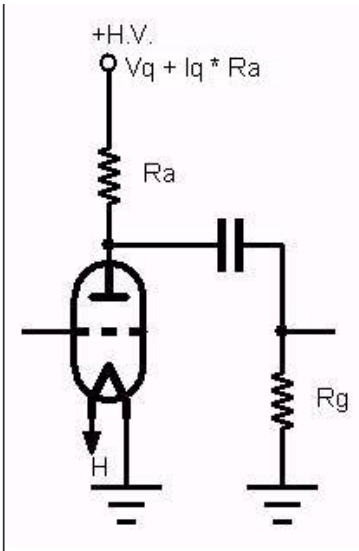
The following sequence of graphs illustrate how the intersection of the load line and the anode characteristic curve (marked by a red dot) changes as the grid voltage changes, giving us the instantaneous anode current and voltage as a function of the grid voltage:



If the AC load is the same as the DC load, the instantaneous operating point lies as already described on the load line.

If the AC load is different from the DC load, for example, if the 2A3 anode were connected through the primary of a loudspeaker transformer to a 250V power supply, the load line can be drawn as a dynamic load line which passes through the nominal DC operating point Q, but will have a slope corresponding to the total A.C. load which will be the combination of any DC load resistance (i.e. an anode load resistor) and an AC coupled resistance (e.g. the input resistance of a following stage coupled through a coupling capacitor, or the reflected resistance of a load appearing in a transformer). This is explained further for various different coupling arrangements as tabulated below:

 Rl is anode load	Load line slope = -1/Rl Power supply voltage = Vq + Iq * Rl All output power is dissipated in Rl
	The total AC load, R, for a capacitively-coupled load (Rg of a following stage) is given by the formula for parallel resistances (expressed as the conductance, because that is what is plotted on the load line): 1/R = (1/Ra) + (1/Rg)

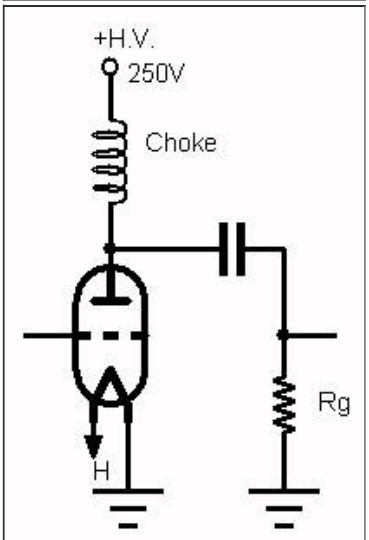


RC coupled stages

The load line is a line with a slope 1/R, drawn through point Q. This will have a higher slope than the one for the anode load resistor on its own, indicating that a given anode voltage swing will result in a higher current swing due to the current contribution from the AC coupled load. It may be necessary to adjust the operating point to avoid exceeding the power ratings or to minimise distortion.

Power supply voltage = Vq + Iq * Ra

Output power is split between Ra and Rg



LC coupled stages

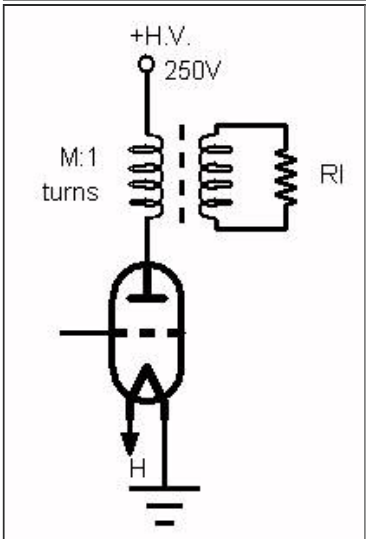
The total AC load, R, for an anode connected to power via a choke inductor (or a parallel-tuned circuit) and to the next stage load via a capacitor is given, assuming the choke to have a very high impedance at the operating frequency, by:

1/R = (1/Rg)

This type of coupling is typically used for driving power output stages, or in RF amplifiers, to avoid too much loss of power in the anode load resistor. The power supply voltage is approximately equal to the quiescent point anode voltage.

Power supply voltage = Vq

Almost all output power is dissipated in Rg



Transformer coupled stages

The total AC load, R, for a transformer-coupled load (Rl connected to the secondary of a transformer with M times as many turns in the anode/primary circuit as in the load/secondary) is given by the formula for the resistance reflected in the transformer primary:

$1/R = 1/(Rl * M^2)$

The load line is again a line with a slope 1/R, drawn through point Q. Of course, with the above example we would contrive the reflected resistance to be 2500 ohms, i.e. an 8 ohm load through a transformer with M=18, so the graph would look the same, but the power supply voltage instead of being 400V is now 250V (ignoring any small voltage loss in the transformer winding).

Power supply voltage = Vq

Almost all output power is dissipated in Rl

Small-signal AC Parameters

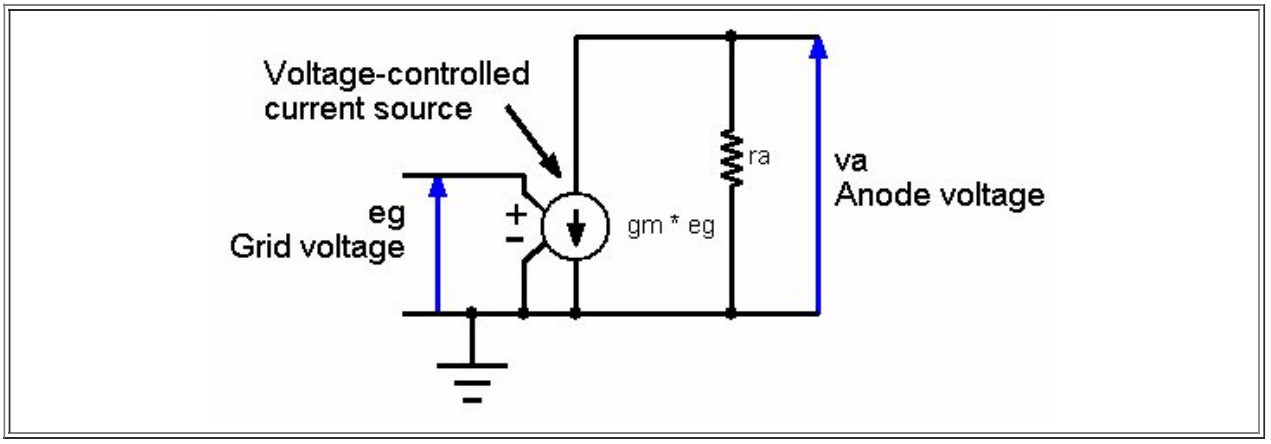
There are three important dynamic (AC) parameters usually quoted for triodes and other types of valve:

- μ, the dimensionless amplification factor, which is the ratio of anode voltage to applied grid voltage, idealised by holding the anode current constant.
- gm, the mutual conductance (units: Siemens), which is the ratio of anode current to grid voltage, idealised by holding the anode voltage constant.

- ra, the anode (internal) resistance (units: Ohms), which is the ratio of anode voltage to anode current, idealised by holding the grid voltage constant.

These parameters may usually be obtained from the valve data sheet, and are small-signal approximations valid only at a particular set of bias conditions, as the device is of course non-linear, but if the signal levels are small compared with the standing voltages and currents, the small signal model is all you need for a basic analysis.

An equivalent circuit is shown below, which is in fact suitable for triodes, tetrodes and pentodes as they can all be described by the appropriate values of these parameters:



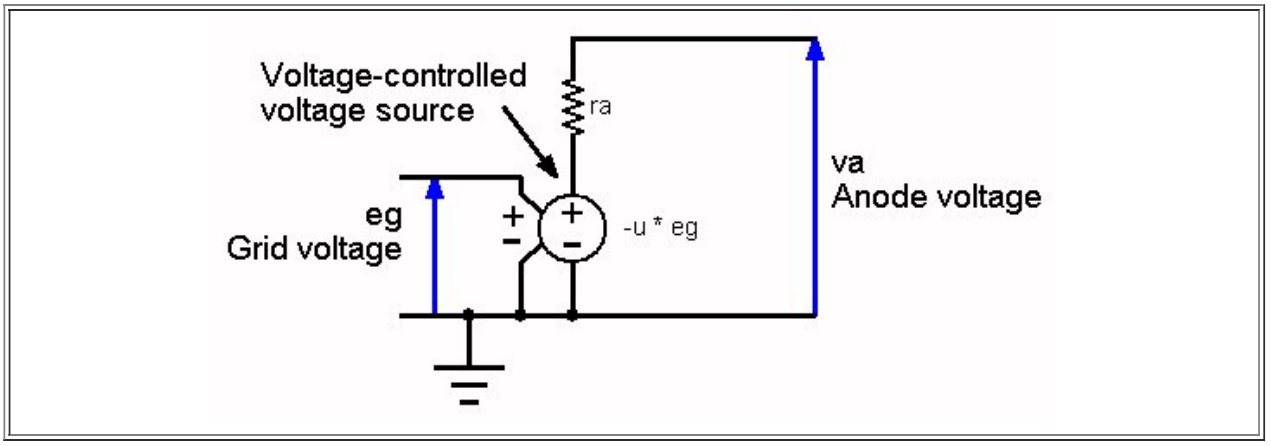
In an ideal circuit environment, the triode could theoretically obtain a maximum realisable voltage gain of μ, but there would have to be no external anode resistor loading effects (and this would require an infinite resistance or a constant current source, i.e. anode current is held constant).

The physical action of the grid voltage is however understood as controlling the flow of electrons from cathode to anode, therefore the device can be understood as a voltage-controlled current source (VCIS), which is the idealised element shown in the above figure. The anode current is equal to mutual conductance (g_m) multiplied by grid voltage e_g, where mutual conductance (g_m) is the ratio of anode current to grid voltage (while anode voltage is constant). The anode characteristic resistance is equivalent to the ratio of AC anode voltage to AC anode current (while grid voltage is constant), i.e. it depends on the slope of the anode voltage/current characteristics at the operating point. These parameters are related by:

$\mu = -g_m * r_a$

There is a minus sign, because the voltage at the anode decreases as grid voltage and anode current increases.

The following equivalent circuit using an ideal voltage-controlled voltage-source (VCVS) describes exactly the same behaviour:



A practical circuit may have a grid resistor that is high enough not to affect the grid signal voltage significantly, but the anode load resistor is usually significant. To AC signals the load resistor R_a appears in parallel with the characteristic anode resistance r_a, according to the VCIS model, or as a part of a potential divider with r_a according to the VCVS model.

According to the VCIS model:

Gain A = $-g_m * r_a \parallel R_a$

Where "||" denotes the parallel resistor formula, i.e. R = R1 || R2 = (R1*R2)/(R1+R2)
Therefore:

Gain A = $-g_m * r_a * R_a / (r_a + R_a)$

According to the VCVS model:

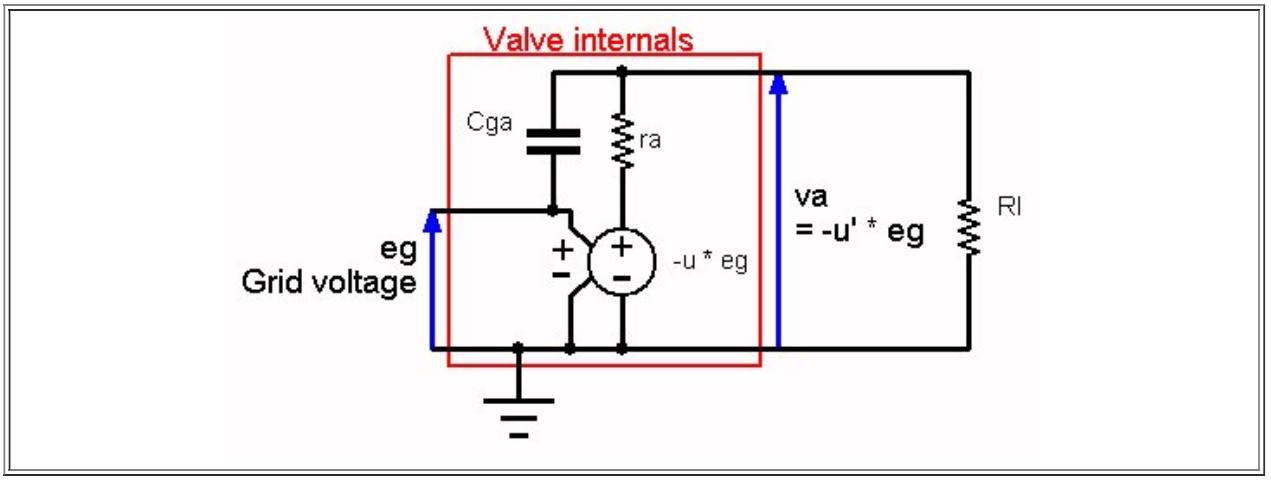
Gain A = $\mu * R_a / (r_a + R_a)$

As $\mu = -g_m * r_a$, the two formulae are equivalent, as expected. If $R_a = 4 * r_a$, then the gain is $-\mu * 4/5$, i.e. 80% of ideal.

The Miller Effect

In 1920, John Miller described the feedback effect in triodes that limited their high-frequency response, especially at radio frequencies [1]. It is observed that the triode appears to behave as if it has a shunt input capacitance that is much larger than any of the inter-electrode capacitances of the triode, and this acts as a low-pass filter to attenuate high frequencies.

The Miller Effect is applicable, not just to triodes, but to any amplifying device that has a stray capacitance between its input and output. Consider the valve as a VCVS model, including the grid-anode capacitance Cga, as shown below:



To take a short cut, I am considering only the Cga capacitance, although there are other stray capacitances, as this is known to have the dominant influence, and it is the one generating the Miller effect. The ideal gain is μ, but the gain to the load Rl will be less, and to take another short cut I have assumed this to be a voltage gain -A, which is somewhat lower than μ. The input impedance is the input voltage divided by the input current. Note that the grid is modelled as an ideal infinite-impedance input, therefore all the input current (other than a tiny amount in an external grid-leak resistor) must flow in Cga. To find the current in Cga, we need to know the voltage across it, which is equal to eg-va:

eg-va = eg - (-A * eg) = (A+1) eg

The impedance of the capacitor is Xc, and the current through it is:

ICga = (A+1) eg / Xc

But the current in Cga if it was connected directly between the grid voltage and ground like an ordinary component (or if the anode circuit were decoupled at AC) would be:

ICga_grounded = eg / Xc

Therefore the effective impedance at the grid, causing a current (A+1) eg / Xc, is

Xc' = Xc / (A+1)

Since Xc = 1 / (2 π * f * C) where f is the frequency and C is capacitance, the effective capacitance at the grid is

Cg = (A+1)Cga

Thus we have arrived at the essential Miller effect: The effective input capacitance of an amplifier with a gain of A is equal to the capacitance between input and output multiplied by (A+1). If A is large, and the other stray capacitances such as grid to cathode (assuming the cathode is AC decoupled to ground) are similar, then clearly the Miller capacitance will be the biggest and dominant one. It is also important to note that there is some advantage in terms of Miller effect to using a triode stage with a low voltage gain.

It is possible to improve the frequency response of triodes by using neutralising circuits, which aim to cancel the effect of the Miller capacitance. Unfortunately, such circuits were often fiddly to set up, had to be adjusted for each valve, and if they went out of adjustment they could make the valve regenerate (oscillate) instead of amplify.

There was consequently a good incentive to look for an improved type of valve, which leads us on to the Tetrode.

The Tetrode Valve

The Tetrode has an extra grid, known as a screen grid, between the control grid (as in the triode) and the anode. This came about as a result of work by Walter Schottky (1886-1976) at Siemens & Halske GMBH in Germany between 1915 and 1919 (who is also famous for explaining the metal-semiconductor Schottky barrier diode effect), and by Dr. Albert W. Hull (1880-1966) and N. H. Williams of the General Electric Company from 1919 to 1928.

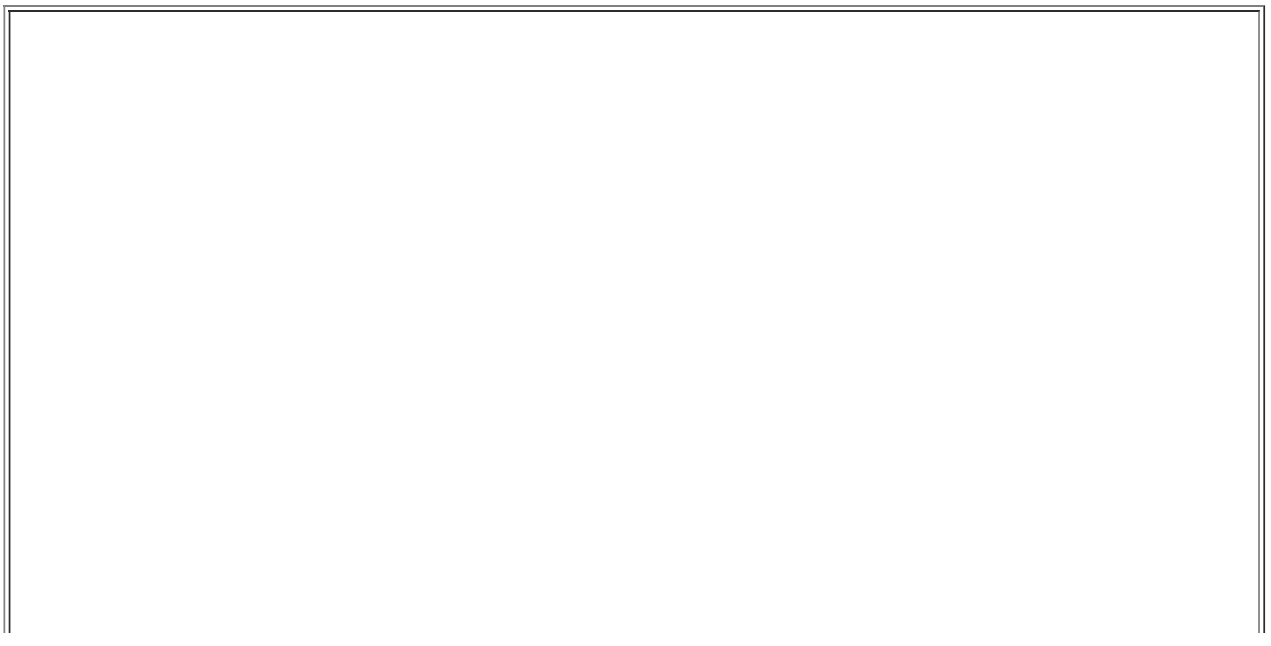
Schottky's grid was biased with a positive supply that was less than the anode bias (typically it is about 2/3 of the anode bias), and it accelerated electrons from the cathode just as the anode does in a triode, except that they mostly pass through the screen grid and are collected by the anode. However, when the anode voltage varies, the screen grid voltage does not, so unlike the triode, the AC voltage on the anode does not tend to counteract the effect of the grid voltage on the anode current, thus eliminating internal negative feedback and increasing the gain of the valve. The anode resistance is increased up to about 0.5 Megohm.

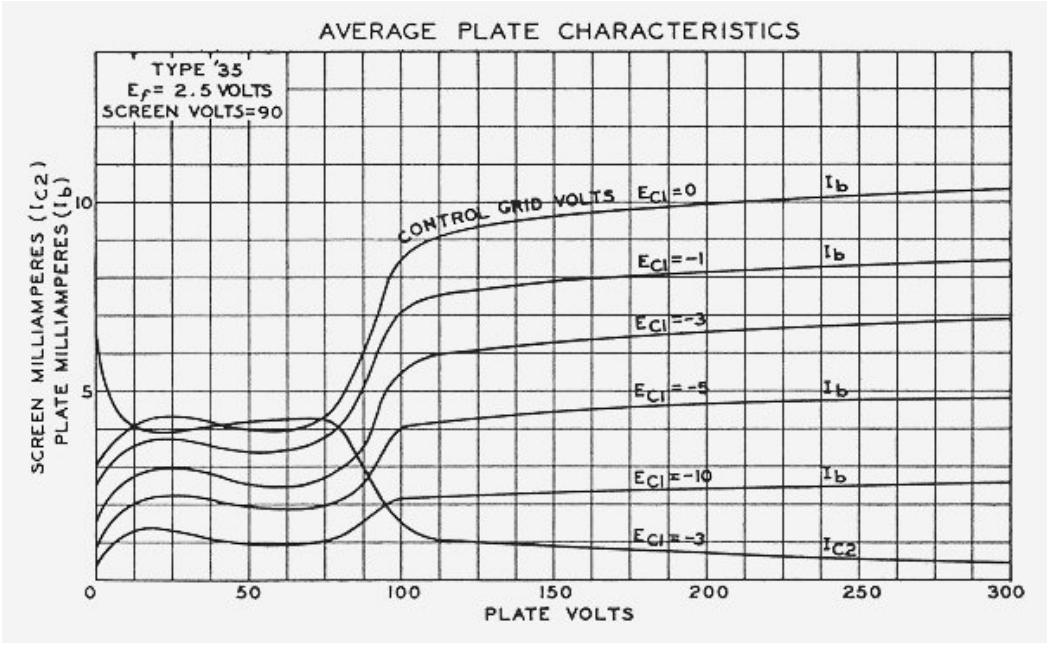
To reduce Cga by a sufficient amount, better screening was required, and Hull's contribution was to ensure that all possible electrostatic paths between anode and grid (for example, between the wires connecting the electrodes to the external valve pins) were shielded as much as possible.

A well designed tetrode valve has a Cga which is reduced by a factor of several hundred to 1000 compared with a triode, and the gain is increased from about 15 to about 40.

The screen grid is usually fed from a resistor voltage dropper and decoupled to ground with a capacitor, so that it remains an effective electrostatic shield at AC, so this capacitor must be chosen to have a low impedance at the lowest frequency to be passed by the amplifier.

The anode current/voltage characteristic curves (each curve is for a different grid bias voltage) are shown below for an RCA-235 tetrode:





The Tetrode "Kink"

Unfortunately, it seemed that two steps forward provided by the screen grid was accompanied by a step back in the form of a "kink" in the tetrode's anode current versus anode voltage characteristic, as can be seen in the RCA-235 curves.

Consider that one electron-volt (eV) is the potential energy gained by an electron when it moves through an electrical potential difference of one volt. The problem is that, given typical anode voltages, an electron moving from cathode to anode gains many times more electron-volts of energy than the energy of a few eV that will liberate one electron from the anode metal, and several electrons may therefore be emitted for each primary electron that arrives. This is known as secondary emission. When this happens in a tetrode, and the anode voltage is lower than that of the screen grid, the screen can attract the ejected electrons, causing a parasitic current to flow between the anode and screen grid.

Over a particular section of the anode voltage / current curve, where the anode voltage is below the screen grid voltage, an increase of anode voltage gives a decrease in anode current. To an AC signal, this appears as a negative resistance; remembering that ordinary positive resistance causes power loss, negative resistance can actually cancel positive resistance and supply power, such that a resonant LC circuit connected between a tetrode anode and its power supply, with no other feedback or coupling circuits, will oscillate.

However, the negative resistance flaw of the tetrode was put to good use; the cancellation of positive resistance and generation of oscillations in a resonant LC circuit anode load is the basis for an extraordinarily simple and versatile oscillator circuit known as the Dynatron.

M. G. Scroggie [2] described the Dynatron as follows:

> "Almost any screen-grid tetrode with the anode fed at a substantially lower voltage than the screen will work as a dynatron, but the Mazda AC/S2 excels. For detailed information there is a lengthy article on it ("Applications of the Dynatron", Wireless Engineer, Oct., 1933, pp. 527-540), together with 35 references to other literature on the subject. The dynatron is worth studying, because once the mode of operation has been grasped it is one of the most useful tools for experimental work. No feed-back is necessary for oscillation; the typical curves of Fig. 19 show that when the anode voltage Va is between the limits of about 10 to 90 per cent. of the screen voltage Vs the anode current (Ia) curve slopes the opposite way from the usual, which means that the anode A.C. resistance is negative and that any oscillatory circuit connected in series will oscillate if its dynamic resistance is numerically greater. As the curves

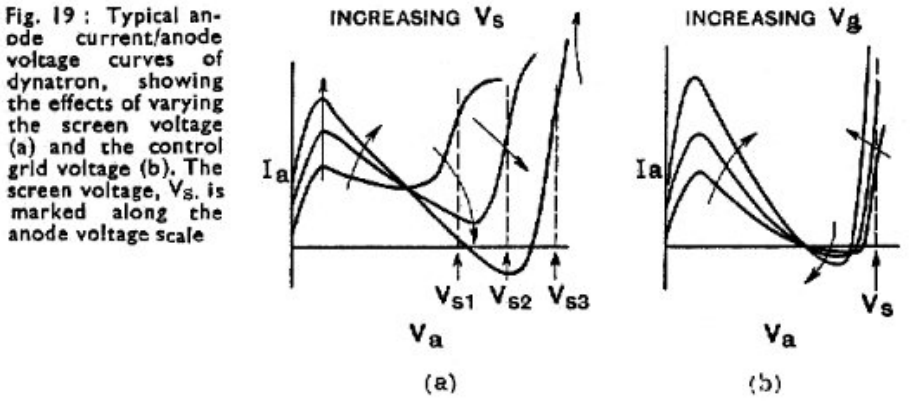
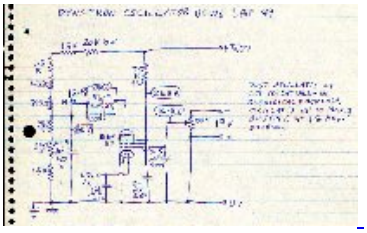


Fig. 19 : Typical anode current/anode voltage curves of dynatron, showing the effects of varying the screen voltage (a) and the control grid voltage (b). The screen voltage, V_S, is marked along the anode voltage scale

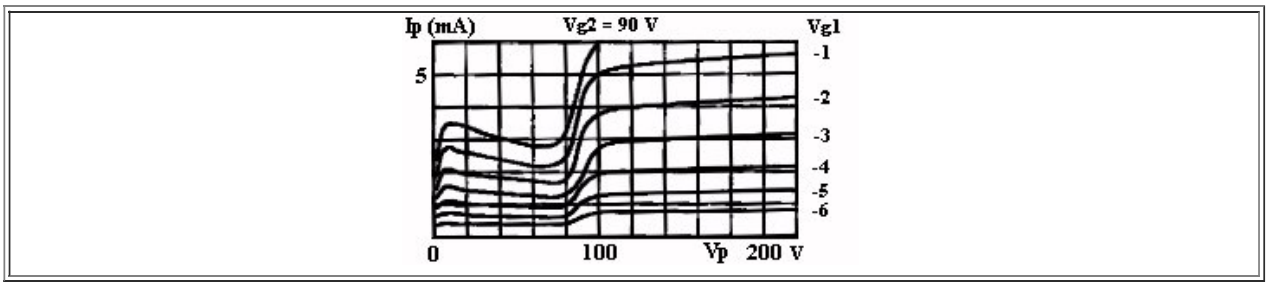
show, the negative resistance can be varied either by Vs or by grid voltage Vg. Generally it is convenient to fix Vs at the lowest that will give the required negative resistance and sufficient amplitude of oscillation, and to use Vg for increasing the resistance until oscillation is only just maintained. Under these conditions it sweeps over the practically straight downward slope and the purity of waveform is exceedingly good. Convenient voltages are Vs 100, Va about 20 for small amplitudes and 50 for maximum amplitude, and Vg variable to about -8. With zero bias the anode resistance of the AC/S2 goes down to about -6000 ohms, which is capable of setting even heavily damped circuits into oscillation; but when run like this there is a risk of the dynatron properties deteriorating fairly quickly. The dynatron is therefore not advised for ultra-high frequency work, in which the impedance of the oscillatory circuit is inevitably low. An important practical point is to make any potential divider used to tap off the anode voltage considerably lower in resistance than the negative resistance of the dynatron. The advantages of the dynatron are its ability to set up oscillation in a simple two-terminal circuit, which may even be screened and inaccessible; the ease and precision of control; the straightness of its working characteristic; and the frequency-stability of its oscillations *provided that they do not sweep beyond this working slope*."

The Dynatron might seem an odd way to make an oscillator, but I have built one myself and it works just like described by Scroggie, who used the principle extensively for laboratory measurements.



Dynatron Oscillator - Circa 1976
Note: Uses a pentode with suppressor grid connected to screen grid to simulate a tetrode
Click for full size image

The anode voltage / current characteristics below are for the RCA Radiotron 36 tetrode, and as for the Type '35 chart shown earlier, clearly show significant negative resistance kinks. The straightness of the negative resistance portions differs between types though, which may be important when designing a Dynatron oscillator. When designing a true tetrode amplifier using a load line, you would have to keep the load line well to the right (i.e. ensuring the minimum anode voltage swing is above the screen grid voltage). This results in a loss of efficiency.



Note that typical simulation models for tetrodes do not cover the full tetrode characteristic, which to simulate the negative resistance, must calculate the fractional split of the cathode current between the anode and screen grid accurately in this region. Instead, models are often simplified by having the parameter fixed rather than varying properly according to the full range of the anode/screen grid voltage relationship [3], [4]; therefore SPICE models will normally only be good for the region above the kink (Most valve SPICE models are written for triodes, beam tetrodes, and pentodes).

Of course, with a negative resistance lurking in the characteristic, the tetrode amplifier could suffer from unintentional oscillation, and

any voltage swing on the anode below the screen grid bias could result in a nasty distortion. These disadvantages naturally led to the development of the pentode.

The Pentode

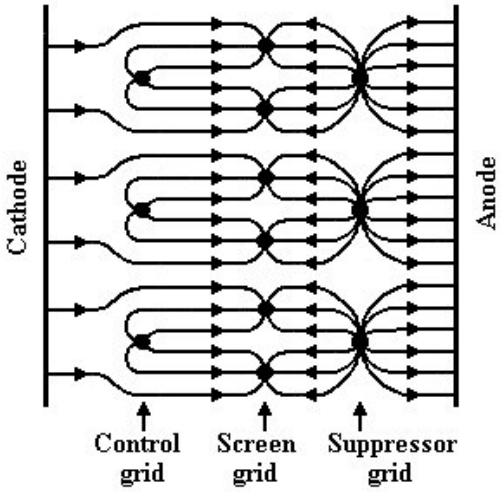
After adding a screen grid, what could be more obvious than to try to solve the problems of the tetrode "kink" by adding yet another grid? Nevertheless, it worked.

The suppressor grid was added between the screen grid and anode, and with the largest spacing between grid wires of all the grids.

This innovation is attributed to the work of Bernard D.H. Tellegen (1900-1990) of the Philips Physics Laboratory in Eindhoven, Holland in 1926 (The Philips B443 AC Output Pentode was the first commercially available pentode, released in 1927).

The German engineer G. Jobst [5], [6] has also been cited in association with pentode development in 1926.

The suppressor grid is usually connected to the cathode, and establishes an electric field at the anode that is always directed toward the anode, which suppresses the flow of secondary electrons to the screen grid, as most do not have sufficient kinetic energy to overcome the field. The electrons that reach the suppressor grid from the cathode have been accelerated by the electric field of the control and screen grids, and arrive with high kinetic energy, enabling them to easily pass through the suppressor grid and reach the anode.



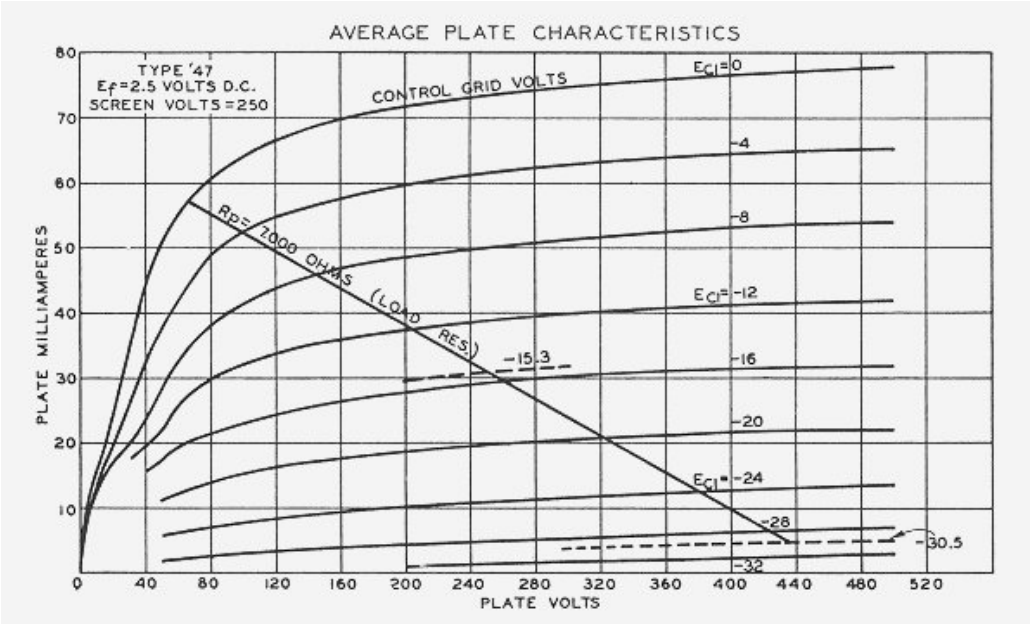
Electric fields within the Pentode

Examples of pentodes are the 6SJ7, EL34 (6CA7), EL84 (6BQ5), SV83 and EF86.

The shape of the pentode anode current / voltage characteristic curves is like a resistor at low voltage, where anode current is proportional to anode voltage, but as anode voltage increases to a "knee" threshold the curve flattens with anode current changing little with anode voltage, more like a current source. This most closely resembles the characteristics of the junction FET. However, junction FETs are usually limited to low-voltage, low-power applications. There is also a similarity to the MOSFET, except that the MOSFET curves tend to have a sharper "knee" at a lower voltage, and a very flat constant-current region.

The suppressor grid also adds (with the screen grid) to the electrostatic shielding between anode and grid, making the pentode suitable for radio-frequency (RF) and intermediate-frequency (IF) stages.

The characteristic curves for the RCA type '47 power output pentode are shown below.



The Beam Tetrode

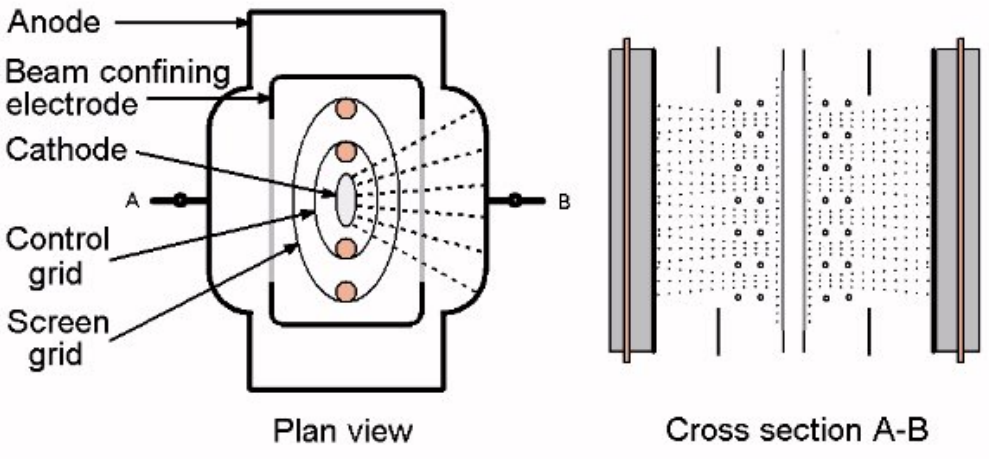
The beam tetrode, for example the 6V6, 6L6, 6550, 807, and 813, was developed to solve the negative resistance "kink" of the ordinary tetrode. Although the pentode had already been developed as a solution to this problem, the beam tetrode was developed by engineers of the UK company HMV in the early 1930s, because the UK's Marconi-Osram Valve (MOV) company, who owned part of HMV, wanted a valve that did not infringe the existing Mullard patent for the pentode.

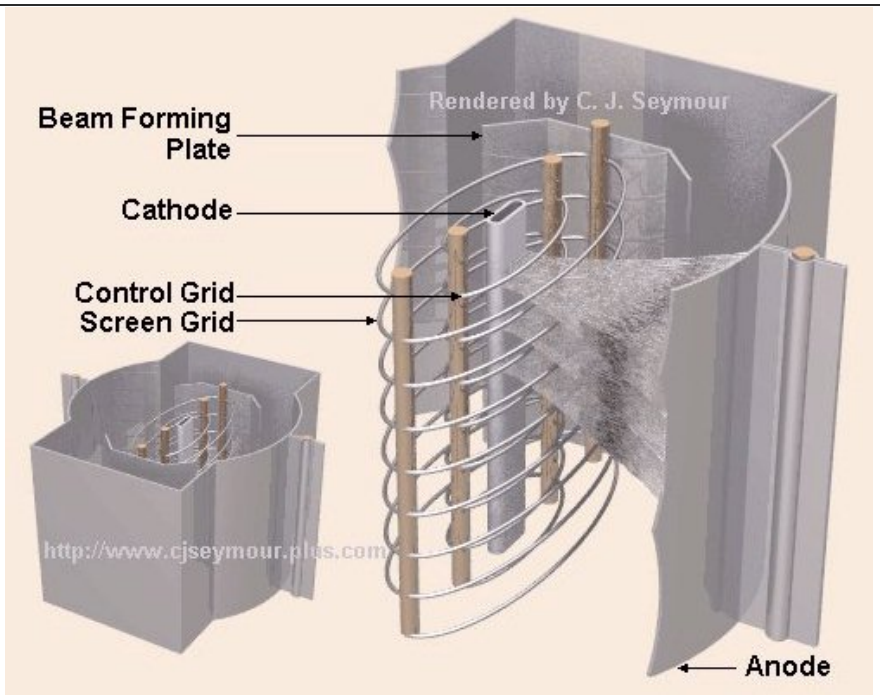
MOV had an exchange agreement with RCA of the USA, who developed the 6L6 for production starting in 1936. MOV developed the KT66 and the KT88, which are still in production and in use today.

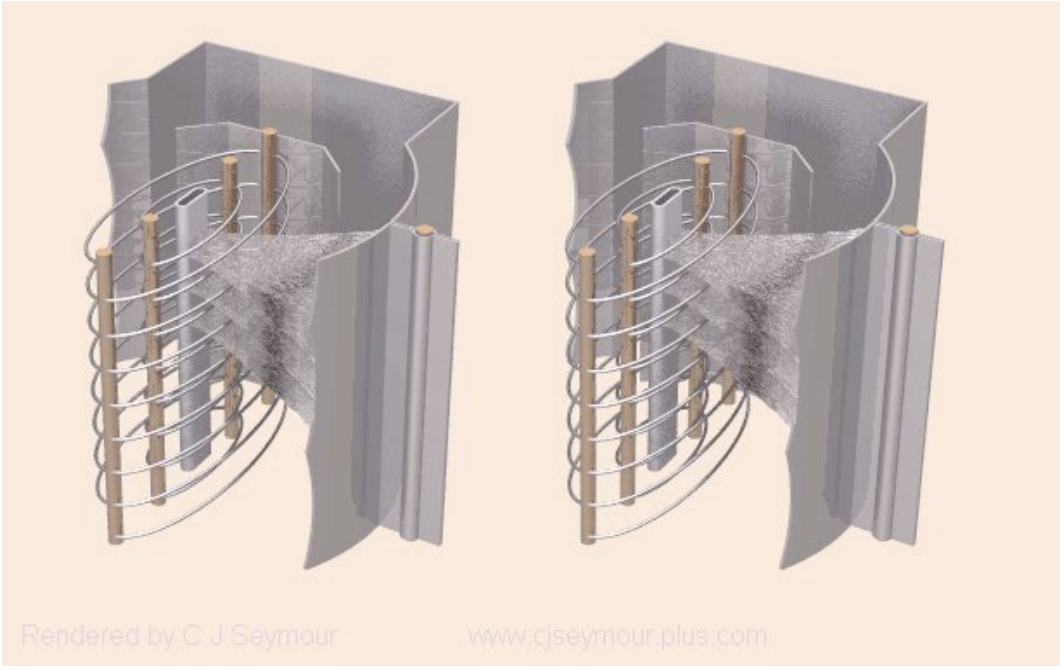
Like the ordinary tetrode, the beam tetrode has a heated cathode, control grid, screen grid, and anode. The difference is in design features that aim to focus the electron flow into tight beams.

The first important feature of the beam tetrode are that the grid wires are carefully aligned so that electrons passing between the control grid wires are not obstructed by the screen grid; this also reduces screen grid current, which is reduced to about 5-10 percent of the anode current compared with about 20 percent for a pentode.

The second important feature is that there are beam-forming electrodes that prevent electrons reaching the anode through the obstructed region of the grid support wires, focus the electron beams through the grids in between the grid support wires, and repel secondary electrons that come from the anode. Where the electron beams arrive at the anode, any secondary electrons ejected back towards the cathode are repelled by the negative charge carried in the high electron density in the electron beams. The resulting suppression of secondary electrons was sufficient to create a "kinkless tetrode", providing a characteristic very similar to that of a pentode. The diagram below shows how the electrodes are arranged in a beam tetrode. The beam-forming electrode may be fabricated as a pair of electrodes, or as a box with rectangular cut-outs where the electron beams are to be focused.



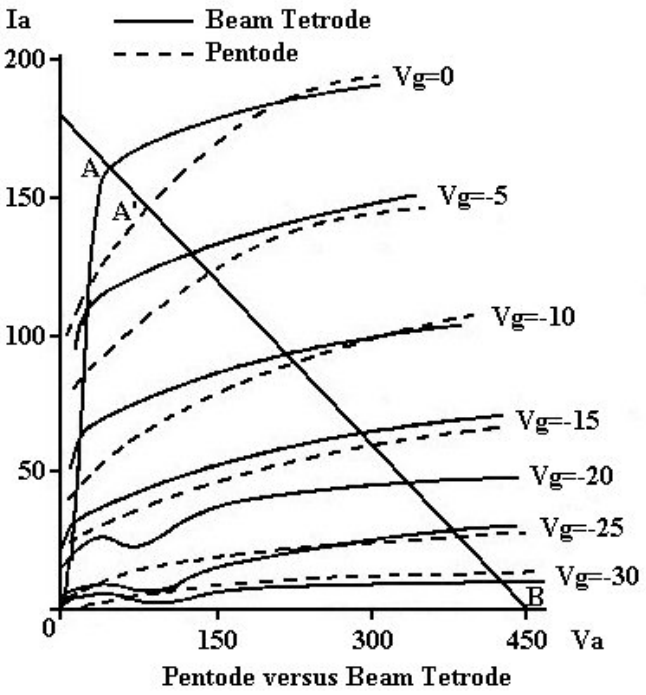




This version is a cross-eye stereoscopic pair. To view in 3D, look at the mid-point between the two images, then cross your eyes to look at a point closer than the screen, (Putting your finger near the screen and moving it towards you while looking at your finger tip may help). Allow the two images to merge and line up, and allow your eyes to focus on the merged image in the middle.

Another advantage of the beam tetrode is that it produces less third harmonic distortion than the pentode. The ear is more tolerant of second harmonic distortion, and if a pair of beam tetrodes are used in a push-pull configuration, this tends to cancel out the second harmonic (but not the third).

The characteristics of the pentode and beam tetrode are compared in the following diagram, where the two types are approximately matched in maximum power dissipation ratings.



Pentode versus Beam Tetrode

Note that the extent of the linear range for comparable distortion levels of the load-line excursion in the direction of increasing anode current is at A for the beam tetrode, but at A' (which is less) for the pentode. This also produces an increase in efficiency for the beam tetrode compared to the pentode.

The beam tetrode does however have some disadvantages compared with the pentode:

To obtain full power output may require driving the grid positive, which will cause current to flow in the grid circuit. This is known as Class A2 (Class A1, which is normally what is meant by "Class A", means that grid current is never drawn). It requires the grid circuit and its driver stage to be designed to handle the current flow without distortion.

The beam tetrode generates higher intermodulation distortion than the pentode. This is a process by which, when two signal frequencies are present in a non-linear circuit, distortion products are generated at frequencies other than the original frequencies or their harmonics, e.g. if input signals are applied at 1000 Hz and 1100 Hz, distortion products may appear at 900 Hz and 1200 Hz, and other frequencies. This may, if excessive, result in an unpleasant sound to the ear.

Info

References

[1] John M. Miller, "Dependence of the input impedance of a three-electrode vacuum tube upon the load in the plate circuit", Scientific Papers of the Bureau of Standards, 15(351):367-385, 1920.

[2] M. G. Scroggie, "Radio Laboratory Handbook", Second Edition (Reprint), The Wireless World, Iliffe & sons Ltd. 1943

[3] S. Zheng and John Keane, "Modeling and Simulation of the Power Amplifier for the RHIC 28 MHz Accelerating Cavity", Collider-Accelerator Department, Brookhaven National Laboratory, Upton, NY 11973

[4] W. Marshall Leach , Jr, SPICE Models for Vacuum-Tube Amplifiers, Georgia Institute of Technology, School of Electrical and Computer Engineering, Atlanta, GA., J. Audio Eng. Soc., Vol. 43, No. 3, March 1995

[5] "Zusammenfassung über Radio-Röhren" (Summary and overview of radio valve development)

[6] The Development of Tubes, Rundfunkmuseum der Stadt Fürth (Radio Museum of the City of Fürth)

[7] Eric Barbour, "The Cool Sound of Tubes", IEEE Spectrum, The Institute of Electrical and Electronics Engineers, Inc.

[8] NJ7P Tube Database Search

[9] The National Valve Museum

[10] Duncan's Amp Pages (includes Tube Data Sheet Locator)

[11] Stefano Perugini, "Vacuum diode Models & PSpice Simulations", originally in Glass Audio, Vol.10,Nr. 4, 1998

[12] John's Radio Web [Dutch] - click "Lampen" to look up valve data

Data

Zip file containing Circuit Scribe project files:
Vequiv.cct - VCCS and VCVS based equivalent circuits for valves
vbasics.cct - Illustration of Edison Effect diode test circuit
trifab.cct - Triode fabrication (cathode and grid)
regencct.cct - Regeneration circuit using 12AT7
Miller.cct - Equivalent circuit for Miller effect
coup_LL.cct - Coupling circuits to illustrate load lines
beamtetf.cct - Plan view and cross section of beam tetrode
anodefab.cct - Diagram showing anode surrounding grid and cathode
ValveCCT.zip [38KB]

Zip file containing SwitcherCAD (LTSPICE) simulation files:
curveLL.asc - Plots triode characteristics with load line
curves.asc - Plots triode characteristics with maximum power line
Tetcurve.asc - Plots beam tetrode characteristics
regencct.asc - Regeneration circuit using 12AT7 triode model
triodeosc.asc - Tuned anode oscillator using 12AT7 triode model
VDIODES.lib - Valve diode models constructed according to [10]
VDiodecurve.asc - Plots diode characteristic for 12X4 or GZ30 using VDIODES.lib
Vdiode.asy - Valve diode symbol for VDiodecurve.asc
Valvespi.zip [12KB]

For Circuit Scribe software, see: Software

Linear Technology Corporation's LTSpice (formerly SwitcherCAD 3) software:
Linear Technology - Home Page

Copyright © All Rights Reserved. No portion of this site may be reproduced in any way without express permission of the publisher.

Page created May 21st, 2005
Last updated September 1st, 2005