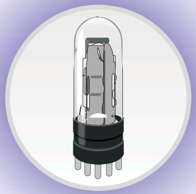

## Tube Renaissance

A possible expalnation of why tubes may sound better than transistors.

From the late 1960's, tubes were largely, though not entirely, superceded by semiconductors in audio frequency amplifier designs. This was an inevitable consequence of a continuing quest for new techniques.

Semiconductors (Transistors and Integrated Circuits) have certain and obvious advantages:their small size, absence of heaters, low voltage operation and consequent opportunity to dispense with output transformers may appear to make tubes obsolete. However, from about 1975 onward, there has been a resurgence of interest in tubes; and it seems worthwhile to consider why. It is said by 'hi-fi' enthusiasts that tube amplifiers sound better, that their distortion is either lower or less noticable. Carefully conducted listening tests seem to bear this out, although their results are difficult to interpret. If there really are subjective differences to a listener between tubes and semiconductor amplifiers, can they be explained technically?

One thing should be clearly understood: it is possible to design either a tube or a semiconductor amplifier so that over a certain range of output power its distortion will be so small as to be imperceptable to the ear. Therefore, if two similarly rated well-designed high fidelity amplifiers, one using tubes and the other using semiconductors, are compared in the same listening conditions, correctly operated, their performance should be indistinguishable - and subjectively perfect.

Now, on the basis of measured performance, many modern high fidelity semiconductor amplifiers are actually superior to the older tube amplifiers, which were already good enough for their distortion to be imperceptable; so how can here be subjective differences? It seems that there cannot be any, while the amplifiers are correctly operated: and this may be the key to the mystery, for there are two major problems: one is that it is extremely difficult to avoid occasional over driving of an amplifer, because of the very large dynamic range of the audio signal; and the other is that the loading is not always resistive. It is under these (usually unintentional) wrong conditions that differences may show up.

Let us consider the over driving first. Owing to continual improvements in recording and playback technique, the possible dynamic range of music signals- from either CD or record - is greater now than it used to be. As a tentative estimate, it appears that the loadest passage of a modern disc recording maybe 40dB above the average sound level.

Now it may be said that amplifiers in a high fidelity system ought theoretically to be able to reproduce the loudest of loud bursts without distortion. However, to allow for 40dB above 50mW - not a very high listening level - a power capability of 500W would be required; and further developments may make the figure even greater. One seems to hear a cry of "where is it all going to end?" Anyway, when setting up an amplifier system one adjusts the gain to give the prefered average sound level. One has no way in advance of knowing in advance whether there is an exceptionally loud passage coming that will over drive the amplifier. Bursts in excess of 30dB above the average are quite rare.

If we accept, then, that occasional over driving is virtually inevitable, how will the amplifier behave? We now come to the first possible reasons why tubes and semiconductors may "sound different".

Presented with an over large signal, tubes merely clip the peaks, delivering a flat-topped waveform while the over driving is taking place. The limiting may occur at the grid as the circuit resistances are commonly such as to prevent it from being driven more than slightly positive, or it may be the results of coalescence of characteristic curves at lower voltages. The ear is surprisingly tolerant of such clipping when it occurs only on these occasional load bursts.

The semiconductors used in audio amplifiers are virtually always bipolar transistors, either discrete or integrated. They require base current to be applied in order to make collector current flow. Now transistor amplifiers normally incorporate a large amount of negative feedback, and, when such an amplifier limits, some of its stages are driven very hard, so that extra large base currents are drawn. If any capacitors are affected by such current pulses, the result may well be that a brief over driving is followed by a comparatively long recovery signal, which would be much more noticeable than mere clipping of peaks. Even with dc coupling, there may still be capacitors thatcan cause such extra signals.

There is a further effect that takes place in the transistor itself, because of the phenomenon of charge storage. A transistor that has been conducting does not switch off immediately when the forward base bias is removed, but continues to take collector current until all the relevant charge carriers that are in transit have been swept out. The effect is most pronounced in a transistor that has been turned on hard: in fact the larger the base current the longer will be the turn-off time. In audio transistors that have been over driven this time may be of the order of hundreds of microseconds, so this effect can also give rise to spurious signals.

When it is realised that even the most critical listener cannot detect peak clipping of occasional short loud bursts by as much as 6dB, we can understand why it is sometimes said that a 50W tube amplifier can sound equal to some 200W transistor amplifiers. A tube amplifier can be quite grossly over driven with little or no subjective effect on sound quality, whereas most transistor amplifiers probably cannot.

The other kind of unintentional wrong operation we have to consider is incorrect loading. The impedance of a loudspeaker system is by no means constant: a so-called 8 ohm system may well present anything from 2 to 16 ohms over the audio frequency range, and be highly reactive at some frequencies. It is under reactive load conditions with large signals that another major difference appears between tubes and transistors. The combination of simultaneous high voltage drop and high current occurring for brief periods at certain parts of the elliptical load line does not normally affect tubes, may cause a catastrophic second breakdown effect, in which a permanent short circuit develops - not to be confused with ordinary avalanche break down, which is a reversible phenomenon.

The risk of second break down may be avoided by using transistors with sufficiently high ratings to be well clear of the effect, if available; but the alternative commonly employed is to incorporate protective circuitry that cuts the signal whenever the output transistors are subject to a dangerous combination of voltage and current, and this obviously has a very unpleasant effect on the sound. The purpose of these remarks is not to denigrate transistor amplifiers, but to present a case for tubes, and to show that there may be technical reasons for the supposedly subjective effects that have been attributed to transistors.

Ways may be found of eliminating or obviating these effects in a transistor amplifier design; but there is a simplicity about tube circuitry which may appeal to many audio engineers, both professional and amateur.

A further point can be made in favour of tubes, concerning cooling. Output transistors have to be conduction cooled, preferably by some method that does not heat up other semiconductors in the circuit.

This requires some rather bulky metalwork thermally isolated from the rest of the chassis. Glass envelope tubes, on the other hand, looses most of their heat by a mixture of convection and radiation.

A brief reference may be in order here about what is usually considered to be the main disadvantage of a tube amplifier, the output transformer.
It is indeed a heavy and costly item, to be set against the relative simplicity of circuit and various other advantages that have here been attributed to tube equipment. However it can enable the amplifier to work into more than one load impedance, whilst a transformerless semiconductor amplifier designed to drive an 8 ohm load would usually deliver only half its normal power into a 16 ohms, and might be damaged if operated with 4 ohms.

Also, with an output transformer provided that it is correctly loaded, the amplifier input sensitivity without feedback is the same whatever the value of load impedance; and by taking the negative feedback connection from a fixed point on the secondary winding the sensitivity with feedback can be made similarly independent of load impedance: in other words, the number of decibels of feedback and therefore the reduction of distortion, damping factor and so on, are the same whatever the load.

So there is something to be said for having an output transformer.

Perhaps enough has been said to suggest at least that the advantages are not entirely on the side of semiconductors, and that points can be made in favour of tubes, concerning both performance and convenience in use.

Semiconductors may produce un-welcome effects on over driving, so difficult to avoid in practise; and not only the output stage, but also low level stages are involved in these.

Tubes have a distinct advantage in operation with reactive loads, and are easier to cool. Even the need for an output transformer is not quite such an unmitigated drawback as it may sometimes seem.

These may be some of the reasons why a substantial part of the audio amplifier market has stayed with tubes during the "transistor era", and why there has recently been such a remarkable "Tube Renaissance".