Class A/AB mosfet power amplifier

A discussion of the effects on performance of capacitors and transistors, and a practical design to illustrate some solutions

J.L. LINSLEY HOOD

The design of audio amplifiers, like that of any other equipment for use in the sound reproduction chain, suffers from the difficulty that, since its purpose is to produce a response from a human sensory organ, the quality of the final result cannot be determined, with absolute confidence, from engineering measurements alone, nor can anyone be certain that the stage has been reached at which no further worthwhile improvements could be made.

Many attempts have been made to relate engineering specifications to perceived sound quality, but these have been complicated by the fact that the ear, like any other sensory organ, varies from person to person, and from time to time. It is also a very poor instrument for assessing sound quality and its memory of sound characteristics is even worse. Nevertheless, in spite of its apparent insensitivity to some quite major defects in the reproduction of the audio chain – such as significant amounts of second harmonic distortion – it can be exceedingly perceptive of some others, especially if trained to listen for them.

THE EMERGENCE OF THE 'SUBJECTIVISTS'

It is a matter of historical observation, and some considerable regret, that circuit design engineers have, in their enthusiasm to exploit new technology, allowed new and unsuspected forms of signal distortion to occur because of their reliance on test procedures such as measurements of total harmonic distortion at full output power, which had not shown anything amiss.

This discrepancy between relatively poor observed sound quality and high claimed performance specification was noted by the lay users of the equipment and tended to undermine their confidence in the validity of engineering specifications as a whole, rather than causing them to demand that fuller, and more searching, test measurements should be made.

It also led to the growth of the opinion that specifications, on their own, were meaningless as a measure of performance, and to the emergence of a minor host of self-appointed pundits, together with a number of magazines dedicated to their views, who claimed particular skills in assessing the quality of equipment, by listening to its performance on a suitable range of sound recordings.

This abandonment of instrumental tests in favour of 'subjective' judgments has led to the proliferation of claims, some of which are exceedingly unlikely on any engineering basis, about the benefits of a host of add-on bits and pieces, and has now led also to the evolution of design procedures based on ideas which are supposed to be good for sound quality, without reference to any instrumental test results.

Since whether or not these design techniques do indeed lead to better sound quality is often judged by the same people who proposed these ideas, this approach tends to be self reinforcing and self sustaining and renders their proponents impervious to any arguments based on physics or engineering principles.

A recent article by Self[1] provided a salutary reminder that it is impossible to make progress in any form of technical development without performance standards which are both measurable and verifiable, against which the effect of design changes can be seen, and against which the validity of design theories or calculations can be tested.

In general, I agree entirely with Self's views, though I entertain a few reservations which I made in a subsequent letter.[2] These arise because I am well aware of the mistakes which have been made in the past, when circuit designers have offered designs which were clearly less good than they should have been – in respect of residual 'crossover distortion' artefacts; or because of proneness to 'slew-rate limiting; or because of inadequate loop-stability margins when used with akward LS loads; or because of poor transient response under reactive load conditions; or because of output device protection systems which caused premature 'clipping' on LS systems which had a low impedance at some part of their frequency response; and so on and on – and I lack adequate confidence that contemporary test procedures will reveal all of the faults which may remain.

In particular, I feel that while a great deal of work has been done in reducing the magnitude of steady-state non-linearities, not enough attention has been paid to circuit behaviour under discontinuous or transient signals, where prominent intermodulation effects may arise. Measurable malfunctions may therefore still lurk in this area.

This concentration on steady-state harmonic distortion figures is probably due, for commercial reasons, to the excessive importance which the layman attaches to the number of zeros behind the decimal point in the quoted THD figure as a criterion of quality.

Steady-state measurements may also tend to minimize the result of sudden changes in signal level upon components which are sensitive to thermal or voltage-dependent effects, such as capacitors and semiconductor devices, and I do not think that we are adequately knowledgeable to be confident that no audibly untoward effects whatever will occur as a consequence of these known shortcomings – particularly when these phenomena can be quite clearly seen with other physical test procedures.

CAPACITORS

Capacitors are the most complex of all the 'passive' components, in respect of their underlying physical behaviour, and differ considerably from the notional 'pure' capacitance which one might depict with the symbol shown in **Fig. 1(a)**. A broad distinction can be drawn between 'polar' (i.e., 'electrolytic'), and 'non-polar' (i.e., film, mica or ceramic dielectric) types, in terms of the effective equivalent circuit introduced by the component but, in general, this will be more nearly that of Fig. 1(b).

In this, C is the effective capacitance of the unit, which will be somewhat dependent on frequency, temperature, and operating voltage. In series with this element of capacitance is a resistance, R_k, representing the dielectric-loss factor, which is strongly dependent on temperature and operating frequency, and in parallel with C is the leakage resistance R_l – also very temperature dependent.

In all capacitors, there will be a series element of resistance, R_s, and a series inductance, L_s, simply due to the mechanical

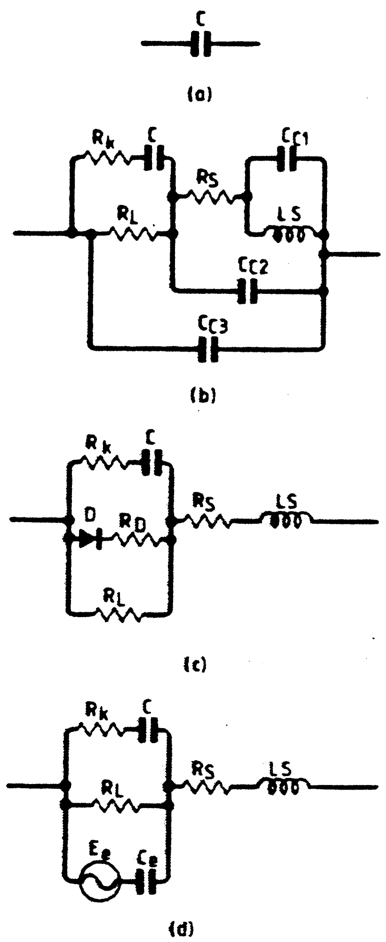


Fig.1. At (a) is a "pure" capacitance, which is more nearly represented by the equivalent circuit at (b). The diode in (c) represents the unidirectional conductive path in an electrolytic capacitor, while (d) shows a generator and resistor to indicate the stored charge and dielectric hysteresis exhibited by film dielectrics.

construction of the component, with a small amount of inherent distributed parasitic capacitance, C_c, which can probably be ignored except at radio frequencies.

Electrolytic types. In these there will also be a unidirectional conductive path, D, in series with a further non-linear resistance R_d as shown in Fig 1(c), which comes into effect if the polarity is reversed, but can also have an effect under zero polarizing voltage conditions when these have persisted for some time, due to the gradual deterioration of the electrolytically formed dielectric layer.

The action of the polarizing voltage has a complex electrochemical/ionic effect and, if reversed polarity conditions are allowed to arise, modifications to the nature of the dielectric layer can permanently affect the other characteristics of the component.

As regards the common electrolytic capacitor types, the tantalum-bead types are more compact for a given capacitance value, have a lower series inductance and a higher reverse breakdown voltage (2-3V vs. about 0.5–1V for aluminium types) and a dielectric layer which is more resistant to deterioration during zero polarizing voltage conditions. On the other hand, the equivalent series resistance (ESR) is significantly greater and even more non-linear than that of equivalent aluminium types. Tantalum bead capacitors are only available in relatively low working voltage forms.

Non-polar film dielectric capacitors. Although these avoid some of the undesirable characteristics of the electrolytic types, they can suffer to a much greater extent from dielectric hysteresis and other stored charge effects of the 'electret' type, represented in Fig. 1(d) by the generator E_e, and the series capacitor C_e.

The possibility of building into the dielectric layer a semi-permanent polarization, usually by heating the material above its first-order transition temperature and then allowing it to cool while exposed to an electric field, has been known and exploited in 'electret' microphone diaphragms for some years, but it can also arise in normal use with suitable materials. In general, the proneness of a dielectric material to this effect is dependent on its molecular structure and upon its crystallinity, physical hardness and rigidity.

Of the commonly used film dielectrics, those such as poystyrene, polycarbonate or polysulphone, from which thin films are made by band casting from a solution, are both limp and amorphous and are therefore less likely to retain molecular-scale electromechanical distortions than the more rigid and highly crystalline types of film such as those based on polypropylene or polyesters which are manufactured by biaxially stretching a thicker extruded sheet.

However, the molecular (polar) asymmetry of the solution-cast materials is typically greater, with the exception of polystyrene, than that of polypropylene, say, which makes a clear preference difficult.

A desirable quality in these components is that they should be compact, and offer a high capacitance/volume ratio. Unfortunately, since both the dielectric constant of the material and the dielectric loss factor are dependent on the asymmetry of the polar groups within the molecule, it is implicit that the desirable qualities of low dielectric loss and high capacitance values cannot be obtained in physically small components.

Stacked film/foil capacitors, where the conductor/dielectric combination is assembled like a pack of cards, offer a lower series inductance (L_s) than spiral wound forms. In all of these types, film/foil components offer both a lower series resistance, (R_s), and a higher leakage resistance, (R_l), than the metallized-film types, but are physically more bulky.

Ceramic dielectric capacitors. Certain piezo-electric ceramic materials, such as titanium dioxide, barium titanate, and barium titanate zirconate, offer dielectric constants in the range 80-50,000, which permits the construction of very small, high-capacitance and low-ESR components. However, the frequency and temperature dependence of capacitance and dielectric-loss values of these capacitors can be very high, which limits their use to RF applications where the overriding consideration is for a low ESR.

Other types. Both mica and air dielectric components are free of most of the problems noted above, but are only available in small capacitance values. Waxed-paper dielectric components are now, thankfully, seldom found.

TRANSISTORS

Transistors are the other main source of non-ideal behaviour in electronic circuitry, in that they are strongly temperature, current, voltage, and frequency dependent in nearly all of their characteristics. Bipolar (NPN/PNP) junction devices are bad in all these respects, though manufacturing techniques have lessened the effects of some of these and circuit layouts have been evolved to reduce the influence of others.

A major residual problem with bipolar junction devices is that of 'hole storage' which prevents a clean current switch-off following a high-current pulse. This can be minimized by ensuring that the device is never driven into saturation, but hole-storage effects are always present. These defects are at their worst in power-output stages because of the high peak currents involved and it is in this position that fets and mosfets offer their greatest advantages.

The mosfet is a particularly attractive device to use in this application in that, since the conduction mechanism is that of an electrostatically induced charge layer in a relatively lightly doped substrate, it does not promote hole-storage effects. It also has a better HF response, which facilitates the design of stable negative-feedback systems, and their greater independence of gain on output current improves circuit linearity. When optimally biassed, their quiescent characteristics can also be less temperature-sensitive.

Power mosfets are available in several forms, as shown in **Fig. 2**, of which the two most common are 'U' and 'T', named after the shape of the active region or the nature of the current flow, and shown in 2(b) and 2(c).

Various manufacturers have introduced their own versions of these topologies, to optimize advantages or lessen disadvantages but in general the 'V' or 'U' mos types are faster, but less rugged and less well suited to complementary polarity than the 'T' mos forms. They all suffer from a high gate source capacitance, particularly in the higher current versions where multiple parallel channels are employed to lower the impedance of the conducting path, and this factor must be born in mind in designs employing them.

They are also prone to gate/source breakdown – causing device failure – if the permitted gate/source potential is exceeded, and this also must be guarded against in the design. This problem exists because, unlike small-signal (RF) mosfets, or –mos logic elements, protective zener diodes cannot be incorporated within the diffusion structure without introducing the possibility of thyristor action.

The remaining design problem is that, because of their excellent HF response, it is possible that RF oscillation may occur, in the tens or hundreds of MHz range, due to the unwise layout of external connecting wiring. Some care should be taken to avoid parallel paths for gate and source or drain leads, and gate stopper resistors should be employed where necessary, especially in the output stages. These should not be too large because of the presence of the fairly substantial gate source capacitance, which can be at least 1nF, in the case of power devices.

AN ALL-MOSFET AUDIO POWER AMPLIFIER

With the various design considerations discussed above in mind, and since small-signal U-mos transistors are now available in both P– and N– channel versions at a reasonable price, it seemed to be an interesting exercise to design an audio power amplifier using only mosfets. The objects of the circuit design were to limit the need for capacitors in the signal path, and to adjust the circuit component values so that the capacitor/s in the negative feedback path, where their imperfections could have a direct influence on the performance of the circuit, could be of a non-polar type.

My original intention was to use mosfets throughout, but these are more expensive than bipolar devices. In places, such as in the constant current sources, where there was little or no signal voltage and no particular advantage seemed to be offered by the use of a mosfet transistor, I have therefore opted for the less expensive bipolar component.

The final circuit layout chosen for the amplifier is shown in **Fig. 3** and is of fairly conventional form. A pair of P-channel mosfets, (Tr_3/Tr_4), is arranged as an input long-tailed pair, fed from a constant-current source, (Tr_1/Tr_2), driving a single N-channel, small-signal U-mosfet gain stage

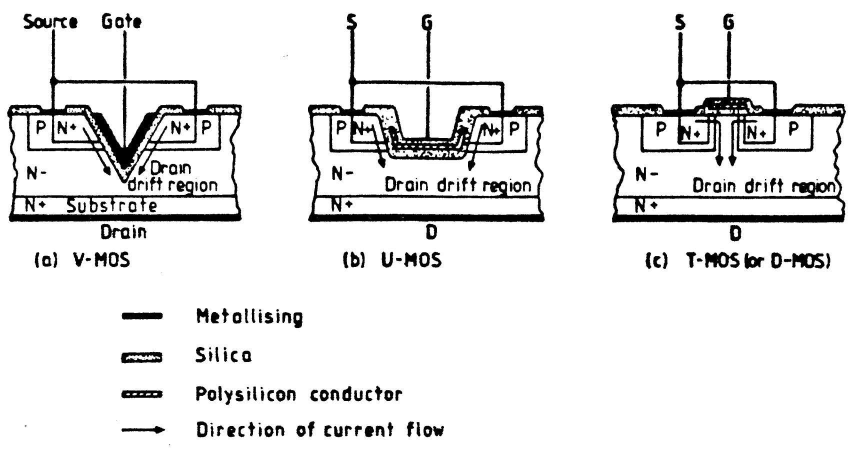


Fig.2. Three forms of the power mosfet

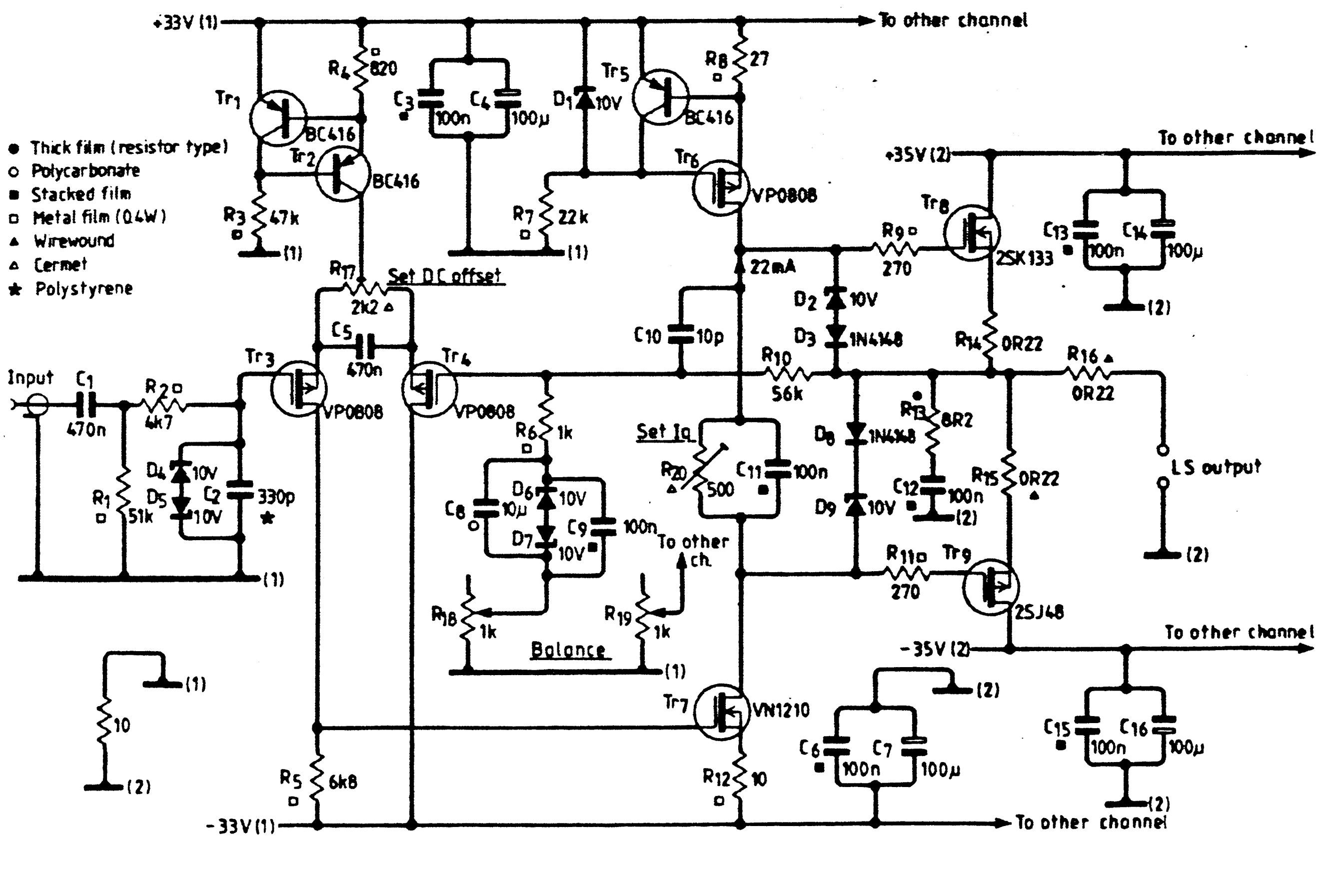
(Tr_7). Since it was intended that the output stages of the amplifier should operate largely in class A, in which the residual harmonic distortion of the circuit would be very low, it was not thought necessary to use a 'current mirror' as the load for Tr_3/Tr_4. This use of a current mirror is a conventional technique for increasing both circuit gain and available negative feedback for a given overall loop gain, as a means for 'cleaning up' a less-good performance.

Again, since the output impedance of both Tr_6 and Tr_7 is very high, and is largely independent of operating voltage within the range employed, I did not consider it necessary to 'bootstrap' these devices to improve their linearity or to lessen the dependence of gate-drain capacitance upon gate-drain potential.

There is always a temptation for circuit designers to 'lily-gild', but experience suggests that more elaborate circuit structures aimed at further reducing already-low THD values also make the problems of loop stability more complex, and may impair the overall transient performance.

In the design of Fig. 3, the 'Zobel' network C_{12}/R_{13}, together with the small capacitor, C_{10}, is all that is needed to provide an adequate gain and phase margin in the feedback loop; C_{10} is employed in a position which greatly lessens the tendency to slew-rate limiting, in comparison with the more conventional and less satisfactory technique

Fig.3. Final circuit of the mosfet power amplifier.



in which C_{10} would be connected between drain and gate of Tr_7 to provide a 'dominant lag' form of HF compensation. This latter approach gives better THD figures at the upper end of the frequency passband, but impairs 'slew-rate' characteristics and transient behaviour.

As I have already said I do not feel that there is any particular virtue in striving for ultra-low THD figures – certainly not below the 0.01% level – at the expense of circuit complexity and cost, or with the possible penalty of impaired or more complex transient response. The design shown, though relatively simple in layout, has an excellent performance in respect of both THD, (better than 0.01% at all power levels, within the frequency range 20-5kHz, and less than 0.03% up to 20kHz) and step-function response which is quite free of ringing and overshoots.

Layout, and power supply. Circuit designers tend to assume that power supply lines will be pure DC, of a known and stable value and devoid of signal residues or mains frequency ripple, and tend to ignore the ill effects which might arise if this is not the case. While there are well known circuit techniques which improve the degree of supply-line signal rejection, it is more elegant to remove this problem at source by using properly stabilized DC supplies. With modern devices this approach offers no problems and any well designed supply circuitry will suffice.

I would also recommend that both the small-signal and the power output parts of the circuit are fed from separate supplies, to lessen the need for a very low source resistance from them. With the circuit shown, there will be no significant penalty in channel separation from operating both channels from the same low-power and high-power supply lines.

With conventional circuit-design procedures, it is quite easy to design stabilized power supplies with an output impedance which is only a small fraction of an ohm. To the 'subjective-sound' fraternity – among whose current fads is the employment of entirely independent power supplies for each channel, with massive and costly reservoir capacitors (but only in a crude rectifier/capacitor system), and filing cabinet sized mains transformers – I would observe that, to obtain a supply line impedance of 0.1 ohms at 5Hz would require a reservoir capacitor of 0.3F. Four of these would not appear to be a cost-effective (or space saving) alternative to a stabilized PSU.

In the case of the feedback-path DC-blocking capacitor, C_8, I would prefer that this should be of polycarbonate dielectric type and, if this is of spiral-wound rather than of stacked-foil type, it should itself be bypassed by a smaller stacked-foil component to lessen the impedance of this path.

Operation mode. I noted above that this design was intended to operate 'largely in class-A'. My experience and observation over a number of years suggests that the bulk of domestic listening, even with relatively inefficient loudspeaker units, is at peak output power levels in the range 0.1 – 3 watts.

For a nominal speaker impedance of 8 Ω, this could be met with an output stage quiescent current of 0.4 ampères/channel, set by R_{20}. On higher output-power demands, the circuit slides quite gracefully into class-AB operation.

Those quoted in Fig. 3 will allow a maximum output level of about 35-40 watts/channel, with a static thermal dissipation for each output device of some 14 watts, for which adequate heat sinking (3°C/watt for each device) should be provided. For higher power class-A operation, a higher quiescent current should be chosen, with more massive output device and power supply heat-sinking. Beyond I_q values of 1A, it would probably be helpful to parallel the output devices, together with their associated emitter and gate-stopper resistors.

Overload protection. I would prefer this to be provided by a simple re-entrant style of current limit in the power supply itself, which could be combined with some electronic sensing circuitry to shut down the PSU in the event of an unacceptably large DC offset appearing at the output terminals. The Hitachi output mosfets appear to be sufficiently rugged for simple gate-protection zener diodes to prevent device breakdown.

References

1. Self, D.R.G., *Electronics and Wireless World*, July, 1988, pp692-696.
2. Linsley Hood, J.L., *Electronics and Wireless World*, Letters, September 1988, pp860-861.

Hart Electronic Kits Ltd., of Penylan Mill, Oswestry, Shropshire, SY10 9AF, can supply all the components needed for this design.