Amplifiers and Superlatives

An Examination of American Claims for Improving Linearity and Efficiency

By D. T. N. WILLIAMSON* and P. J. WALKER[§]

Recent articles, particularly in the United States, have shown that some confusion of thought exists about the "goodness " of an amplifier for reproducing sound. This has been accompanied by a debasement of the terms used to describe the properties of an amplifier; words that have a precise meaning when standing alone become meaningless when qualified by superlatives. This article, which starts with a. discussion of what constitutes a good amplifier, attempts to clear up some misconceptions and generally to clarify the position.

THE most commonly specified parameter of an amplifier is its harmonic distortion or inter-modulation content, and this has tended to be regarded as the primary standard of "goodness." In 1944, one of the writers suggested[1] that a standard of 0.1 per cent total harmonic distortion at maximum output was a good level of performance to aim at. Such a low distortion content is readily achievable by modern design methods, and is, in fact, undetectable by listening tests, however refined. This arbitrary figure was chosen, not because it was the maximum permissible, but because, being rather lower than was absolutely necessary, a good margin is left for any deterioration of the equipment when in service.

The requirements for a good amplifier were listed in the article already mentioned, but are worth repeating and extending. These are :-

(1) Negligible non-linearity distortion up to the maximum rated output. (The term " non-linearity distortion " includes the production of undesired harmonic frequencies and the intermodulation of component frequencies of the sound wave.) This requires that the dynamic output/input characteristic be linear within close limits up to the maximum excursion of any waveform, with frequency components inside and outside the working range, which is likely to be fed to the amplifier. This should hold good under conditions of varying load impedance such as are likely to be encountered in practice.

(2) Linear frequency response within the audible frequency spectrum of 10-20k c/s

(3) Negligible phase shift within the audible range. Although the phase relationship between the component frequencies of a complex steady-state **sound** does not appear to affect the audible quality of the sound, the same is not true of sounds of a transient nature.

(4) Good transient response. In addition to low phase and frequency distortion, other factors which are essential for the accurate reproduction of transient waveforms are the elimination of changes in effective gain due to current and voltage cut-off in any stages, the utmost care in the design of iron-cored components, and the reduction of the number of such components to a minimum. Changes in effective gain during " low-frequency " transients occur in amplifiers with output stages of the self-biased Class AB type, causing serious distortion which is not revealed by steady-state measurements. The transient causes the current in the output stage to rise, and this is followed, at a rate determined by the time-constant of the biasing network, by a rise in bias voltage which alters the effective gain of the amplifier.

(5) Low output resistance. This requirement is concerned with the attainment of good frequency and transient response from the loudspeaker system by ensuring that it has adequate electrical damping. The cone movement of a moving-coil loudspeaker is restricted by air loading, suspension stiffness and resistance, and electromagnetic damping. The efficiency of a baffle-loaded speaker is rarely higher than 5-10 per cent, and the air loading, which determines the radiation, is not high. Electromagnetic damping is therefore important in controlling the motion of the cone. This effect is proportional to the current which can be generated in the coil circuit, and is therefore proportional to the total resistance of the circuit. The output resistance of the amplifier therefore should, in general, be much lower than the coil impedance.

(6) Adequate power reserve. The realistic reproduction of orchestral music in an average room requires peak power capabilities of the order of 15-20 watts when the electro-acoustic transducer is a baffle-loaded moving-coil loudspeaker system of normal efficiency. The use of horn-loaded or other efficient loudspeakers may reduce the power requirement to the region of 10 watts. In an amplifier for the best possible quality it is, of course, assumed that output peaks required in practice will be comfortably below the maximum available. Even so, good design should ensure that the overload characteristic will not suffer temporary paralysis or " blocking " on momentary overload.

(7) The level of hum and noise should be at least 80dB below maximum output.

This is a formidable specification, and by no means every amplifier styled as " high quality " will meet it. However, provided that these requirements are met in every respect and that extraneous components occurring **within the audio** range (produced either by components inside or outside that range) do not

---

exceed a small fraction of 1 per cent, any amplifier will sound as good as any other amplifier and it becomes impossible to "improve " a power amplifier in the sense of producing better sound.

Efficiency.- On the other hand it is, of course, absurd to say that amplifiers cannot be improved in any sense. Efficiency, compactness, and reliability are important and it is with regard to these factors that improvements are possible and in many cases desirable. The size and purpose of an amplifier will determine how much weight should be given to each of these factors.

In small amplifiers with power outputs below twenty watts, power efficiency is not usually of prime importance, and other considerations such as ease of construction and certainty of results may easily outweigh it. With large amplifiers, efficiency becomes a very important factor. Sometimes it is not possible to produce a high output unless the effiency is high, because of valve dissipation limits.

Controlled and Uncontrolled Production.-The designer's aim is (or should be) to produce the best possible sound for those who will ultimately make use of his efforts. If he designs for construction not under his control, he must, as far as possible, ensure that every amplifier made will meet his performance figures without undue difficulty and with the employment of limited measuring equipment. In such a case, the avoidance of circuitry which is not straightforward, or in which deviations from specification in the values and construction of components is liable to produce poor results, is paramount.

The designer who has control over the production of the complete equipment has a different set of problems. He has much greater freedom of choice, and it is likely that he will arrive at different circuitry.

The amplifier which was described in *Wireless World,* April-May, 1947, is an example of the first kind of circuitry. The success it has achieved since then, in the hands of people with widely varying degrees of skill, is indicative that trouble-free performance is readily obtainable with only simple adjustments. This amplifier was based upon a triode output stage with a good output transformer, the performance of which was improved by overall negative feedback.

The only valid criticism which can be made about its performance is on the score of efficiency. The output of about 15 W is produced for an h.t. consumption of 56 W, an efficiency of only 27 per cent. By the use of tetrodes as output valves this could be raised to the order of 35-40 per cent, which means that, for the same power consumption, the output could be increased to 22 W, or, alternatively, that the power consumption could be reduced by about 20 W.

Whether this is worth doing or not is problematical. In the designer's opinion, the additional risk of trouble in unskilled hands outweighs the advantage of higher maximum power out-
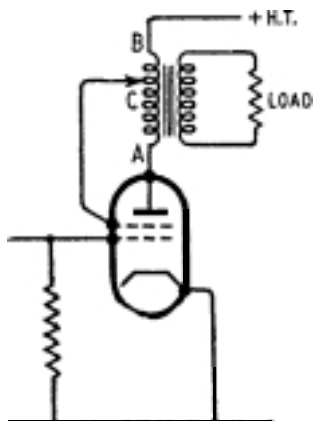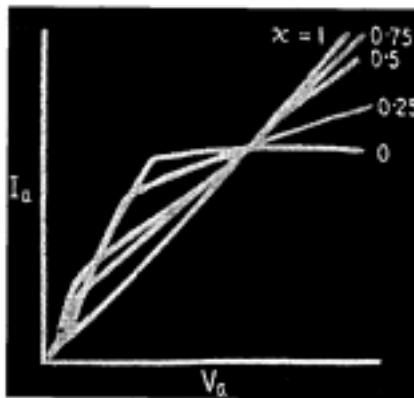

Oscillogram showing transition of Ia-Va curves from tetrode to triode form us the tapping point in Fig. 1 is moved from B to A.

put, as in the vast majority of cases even the present level of output cannot be fully utilized. The reduction in power consumption and ht. voltage would not greatly affect the cost, although it would give a higher factor of safety for the capacitors in the circuit. There is, however, no evidence that this is at present inadequate.

Circuits with Distributed Loads.-Articles [2,3] have recently been published in the United States claiming the superiority of a so-called " ultra-linear " output circuit in which the output valves are used as tetrodes, with negative feedback applied non-linearly by connecting the screens to a tap on the primary of the output transformer. It is stated that the performance is audibly improved over that of triodes with similar degrees of negative feedback.

The present writers do not believe this claim. The circuitry which forms the basis of these American claims for " ultra-linearity " and higher efficiency has, in fact, been familiar in this country for several years, and the technique has been further developed and used in a commercially produced high-quality amplifier. [4,5] It consists of the distribution of the load impedance between the electrodes of each output valve in order to obtain the optimum performance from that valve.

In its simplest form, and as used by Hafler and Keroes, the circuit arrangement is as shown in Fig. 1. The circuit is normal, except that the screen of the tetrode can be tapped on the output transformer primary winding, thus coupling it to the anode in any ratio, and rendering a section of the load impedance common to both electrodes. If the tap is affixed at point A, giving a coupling factor of unity, the stage behaves as a triode, its performance being determined by the dynamic characteristic of the screen and control grid, and if the tap is shifted to point B, the coupling factor is zero, and the stage behaves as a simple tetrode. If now the screen is tapped at intervals between point B and point A, there will be a progressive inclusion of the load impedance in the screen circuit and a progressive change from tetrode characteristics at B, to triode characteristics at A

It is convenient to consider the stage as being a tetrode with negative feedback applied to the screen It differs from a tetrode stage with feedback applied to the control grid, first in that the screen introduces a non-linear element in the feedback loop, and secondly, because as the degree of feedback is in-

Fig. 1. Basic circuit used by Hafler and Keroes.

[2] Hafler and Keroes: "An Ultra-linear Amplifier Audio Engineering November, 1951, pp. 15-17.
[3] Hafler and Keroes : "Ultra-linear Operation of the Williamson Amplifier." Audio Engineering, June 1952, pp. 26-27.
[4] First introduced in 1945; see advertisement of the Acoustical Manufacturing Company in Wirless World, July 1945, P. 10.
[5] Moir : "Review of British Amplifiers," FM TV, October 1951 pp. 30-32.

creased, the permissible anode swing is decreased, due to anode current cut-off on the negative excursions of screen voltage, until finally the Ia-Va curves become those of a triode. A more complete account of the behaviour of the circuit is given in the Appendix.

The curves of Fig. 2, which apply to a single valve, show the effect of changing the tapping point. The output at the point of overload, the relative gain and the output resistance decrease as the degree of feedback is increased, the output resistance decreasing very rapidly because it is dependent, not on the gain reduction, but on the amplification factor, which undergoes a large change.

The distortion at constant output, which is predominantly second harmonic, rises slowly at first until the point of overload is reached. Since the maximum power output is progressively reducing, the distortion curves for maximum output or a percentage of maximum output will show a progressive reduction from tetrode to triode operation.

In view of the increase in distortion, as shown in Fig.2, it would appear that there is little to be gained by the use of this circuitry, since a simple tetrode, with feedback to the control grid, would give a similar reduction of output resistance, and this would be accompanied by a progressive reduction in distortion. This is indeed true in the case of a single valve.

When push-pull connection is considered we obtain a different state of affairs, because of the cancellation of second harmonic components, and we arrive at the curves shown in Fig. 3. The general shape of the curves is maintained under varying loads. With up to 30 per cent of the winding common to screen and anode the reduction in distortion is greater than the reduction in gain. This " something for nothing " is small, however, and can be lost or even reversed if there are appreciable departures from precise balance at any frequency.

It is, of course, not essential to have the screen tapped directly on the primary winding, and an auxiliary winding tightly coupled to the primary (ideally in the form of a bifilar winding) may be used. This has the advantage that the screen supply voltage may differ from that of the anode to give optimum operating conditions.

The " Acoustical " Circuit.-A further development of this circuit (" super-ultra-linear," perhaps?) used in the Acoustical " Q.U.A.D." amplifier and its predecessor, takes the common portion of the winding and inserts it in the cathode, giving the circuit of Fig. 4. As far as the anode and screen circuits are concerned, this arrangement is identical to that of Fig. 1 (neglecting winding resistance and leakage reactance). The only difference lies in the grid circuit, since that portion of the voltage appearing across the common winding is now applied in the grid circuit as overall negative feedback. It is, of course, feedback of the most desirable kind, since it is the most practical method of applying voltage feedback over a single stage without either throwing gain away wastefully or increasing the load on the previous valve.

Since the screen is now connected to a decoupled point, it follows that the designer is free to use different voltages for screen and anode if required, without the necessity of additional windings, with their attendant disadvantages Further, the arrangement makes it possible to provide the optimum degree of smoothing for the anode and screen supplies. The grid resistor may be taken to the cathode end of the common winding if desired, with the result that the

input resistance is increased, thus facilitating the design of the previous stage and reducing the value of the coupling capacitance necessary.

The use of all the electrodes in this way gives additional flexibility in design, so that the parameters of the output stage may be varied to suit the penultimate stage —a technique giving greater scope in arriving at a well-balanced ratio of stage distortions and hence an optimum final design. In the output stage of the Acoustical " Q.U.A.D." amplifier, using two KT66 valves, the values are so chosen that an output of 12
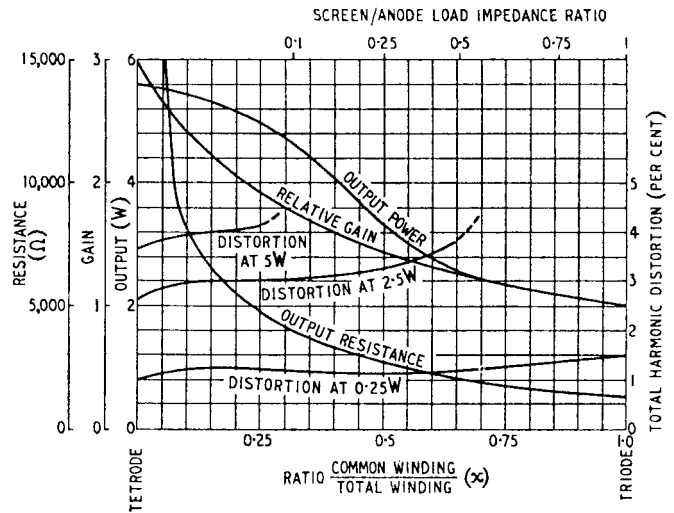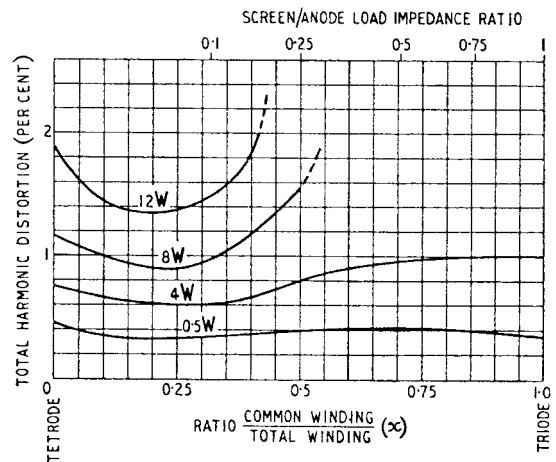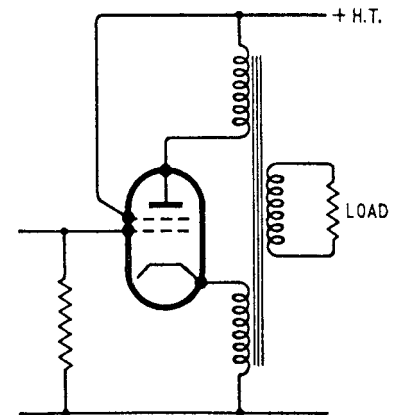
Fig. 2. Curves of output power, relative gain, output resistance and distortion for the circuit of Fig. 1. The distortion components are predominantly second and even harmonics.

Above : Fig. 3. Curves of distortion for a push-pull arrangement (predominantly third and odd harmonics).

Right : Fig. 4. Basic circuit used in the Acoustical Manufacturing Company's "Q.U.A.D." amplifier and its predecessor.

watts is obtained from a 320-V supply with a grid-to-grid input of 72 volts peak, the input resistance being 1 mehogm. The total harmonic distortion of the output stage alone is not more than 0.7 per cent, and the output resistance is approximately half the load resistance.

By comparison with the same valves used as triodes to obtain the same output, the following advantages are apparent:

(1) Distortion is less than half that of triode connection, due to the cathode and screen feedback.

(2) Efficiency is increased from 27 per cent to 36 per cent, resulting in the h.t. voltage being lower by about 100 V, thus simplifying reliability problems both in the amplifier itself and throughout the range of pre-amplifiers, tuners, etc., which may take their supply from it.

(3) Less smoothing is necessary for equivalent hum-levels.

Additional overall negative feedback can be applied to the complete amplifier and with the ratios used in the above example, 8 db less feedback is required for a given level of distortion than with triodes. This assists in maintaining a good margin of stability and reduces the effect of overloading due to the presence of frequencies in the input which are outside the effective feedback range.

There remains the question of output resistance to be considered, about which there appears to be some confusion. This is probably due to the general use, as a measure of the efficacy of damping, of the load resistance/output resistance ratio, sometimes called the damping factor, the scale of which becomes virtually meaningless at high values. It appears to be common practice to aim at as high a value as possible, with presumably infinity (zero output resistance) as the ideal. An output resistance of zero is, of course, only an arbitrary figure, which has ultimately to be added to the speech coil resistance.

Independently of the output-stage circuitry, zero-or any reasonable value positive or negative-can be obtained by a mixture of negative voltage and positive current feedback. It should be pointed out, however, that the optimum value is dependent upon the loudspeaker and particularly the intended performance of its enclosure, so that the doctrine " the more damping the better," is not always sound.

Comparison of Output Circuits—To summarise, Table I gives a comparison of the relative merits and demerits of various output circuits. It will be seen

of triodes and distributed-load tetrodes (especially the cathode-coupled variety), with the exception of efficiency, in which respect the tetrode circuits are superior to the triode.

In order to avoid misapprehension, it should be stated that a similar order of performance is obtainable from a conventional tetrode circuit, by the application of the appropriate degree of negative feedback, preferably in the form of multiple loops to ease the stability problems. The advantage of the distributed-load circuits is that, as a considerable amount of negative feedback is included in the output stage itself, the design of the remainder of the amplifier is simplified and the problems of stability and restriction of scope in design usually associated with large amounts of overall negative feedback are avoided.

Practical Difficulties.— So far we have only been considering the ideal case. In a practical transformer, however, the windings are not perfectly coupled, but are more loosely coupled by a complex network of leakage reactances formed by the distribution of leakage inductance and self-capacitance throughout the windings. This departure from the ideal may mean that, at high frequencies, the circuit is not at all as it would appear on paper, and the effective sense of the coupling may even be reversed, producing oscillation. In a less severe case, peaks and troughs in the frequency response characteristic may occur, accompanied by " ringing " and instability when the amplifier is excited by a transient.

These defects are serious, and can only be avoided by designing the output transformer carefully and by maintaining close control over its production, as even with transformers of the same nominal specification, wide variations in performance at high frequencies may occur due to minor variations in the quality and thickness of insulants, and in the positioning of windings.

The designer who has complete control over production can arrange that the materials and construction of the transformer do not depart from specification, and even if they should do so he has facilities to detect and correct the deviation at an early stage. He is therefore able to make use of circuitry of the type discussed, without the dangers which would attend its use in unskilled hands or where measurement facilities are absent. The advantages obtained have already been discussed.

The constructor with limited facilities cannot be too strongly advised to keep to proven circuits which

TABLE I

Parameter	Triode-connected tetrodes (Class A)	Tetrodes (Class A)	Tetrodes with load distributed between anode and screen (optimum value)	Tetrodes with load distributed between anode, screen and cathode (Q.U.A.D. arrangement)
Efficiency (per cent)	27	38	36	36
Relative power output	1	1.4	1.35	1.35
Relative distortion just below onset of grid current	1	2	1.5	0.5
$\dfrac{\text{Load resistance}}{\text{Output resistance}}$	2–4	0.05–0.1	0.5–1	2

keep to designs requiring the minimum number of coupled circuits in the output transformer, since the possibility of pitfalls is greatest in this component and increases rapidly with the number of windings when all these must be closely coupled.

It will be appreciated from the foregoing that there are a large number of solutions to the problem of designing a first-class amplifier, and no one of these solutions can be called the best solution. Each has its advantages and disadvantages, and the individual designer must choose that which most nearly meets his needs. The "goodness" of an amplifier is not shown by its circuit diagram. Circuits have no inherent magic properties, but are merely the tools with which the designer seeks to achieve a certain result, and different designers-provided always that they have the same high standards in view-may achieve the same result by different means.

APPENDIX

Let

$$\mu = \frac{E_a}{E_g} = g_m R_a \,, \quad m = \frac{E_s}{E_g}$$

$$\therefore \frac{E_a}{E_s} = \frac{\mu}{m}$$

x = fraction of output fed to screen.

E = output voltage in load R_L, produced by grid voltage e.

Neglecting the effect of screen current in the load, which is normally less than 10 per cent,

Current in load due to e

$$= e \; \frac{\mu}{R_a + R_L}$$

Current in load due to $xE = -xE \; \frac{\mu}{m(R_a + R_L)}$

$\therefore$ Voltage in load $= E = \left\{ e \frac{\mu}{R_a + R_L} - xE \frac{\mu}{m(R_a + R_L)} \right\} R_L$

$\therefore e \frac{\mu R_L}{R_a} = E \left(1 + x \frac{\mu R_L}{m R_a} \right)$
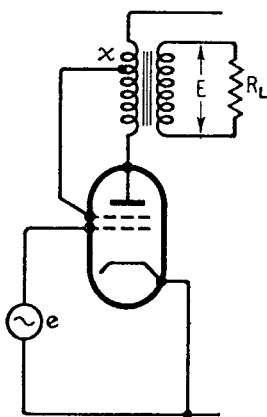
$\therefore$

$$\text{Gain} = \frac{E}{e} = \frac{\dfrac{\mu R_L}{R_a + R_L}}{1 + x \dfrac{\mu}{m} \dfrac{R_L}{R_a + R_L}}$$

When $x = 0$,

$$\frac{E}{e} = \frac{\mu R_L}{R_a + R_L}$$

when $x = 1$,

$$\frac{E}{e} = \frac{\dfrac{\mu R_L}{R_a + R_L}}{1 + \dfrac{\mu}{m}\left(\dfrac{R_L}{R_a + R_L}\right)}$$

Output resistance.

Let $R_L \to \infty$, then $\dfrac{R_L}{R_a + R_L} \to 1$

Then gain $(x = 0) = \mu$

$$\text{Gain } (x) = \frac{\mu}{1 + x\dfrac{\mu}{m}} = \mu'$$

Output resistance $(x = 0) = \dfrac{\mu}{g_m}$

$$\text{Output resistance} = \frac{\mu'}{g_m} = \frac{\mu}{g_m\left(1 + x\dfrac{\mu}{m}\right)}$$
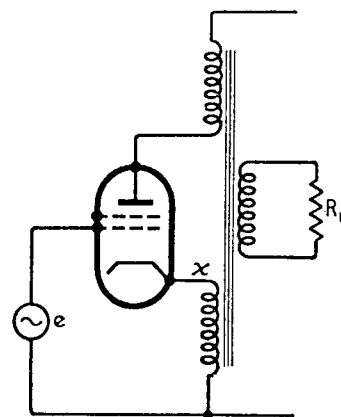
"Q.U.A.D." case

Gain is given by

$$\frac{\mu \dfrac{R_L}{R_a + R_L}}{1 + x\dfrac{\mu}{m}\dfrac{R_L}{R_a + R_L}} = A$$

Gain with feedback $= \dfrac{A}{1 + xA}$

$$\text{Gain} = \frac{\dfrac{\mu \dfrac{R_L}{R_a + R_L}}{1 + x\dfrac{\mu}{m}\dfrac{R_L}{R_a + R_L}}}{1 + x\dfrac{\mu \dfrac{R_L}{R_a + R_L}}{1 + x\dfrac{\mu}{m}\dfrac{R_L}{R_a + R_L}}}$$

$$\text{Output resistance with feedback} = \frac{\text{Output resistance without feedback}}{1 + x\mu}$$

$$\text{Output resistance} = \frac{\dfrac{\mu}{g_m\left(1 + x\dfrac{\mu}{m}\right)}}{1 + x\dfrac{\mu}{1 + x\dfrac{\mu}{m}}}$$