Audio power amplifier design — 2

Negative-feedback concepts

The best result of mathematics is to be able to do without it — OLIVER HEAVISIDE

by Peter J. Baxandall, B.Sc.(Eng), F.I.E.E., F.I.E.R.E.

In the January issue the concept, and possible consequences, of slew-rate limitation were discussed, with particular reference to one cause, in which the first stage of an amplifier is unable to supply the current demanded by the collector-to-base feedback-stabilization capacitor in the second stage. With suitably modified circuit designs such effects may be made insignificant. Before specific circuits are discussed in later articles, the present article will deal with some basic ideas about negative feedback and transfer functions.

Feedback terms: definitions

Fig. 1 represents the general case of an amplifier with overall feedback. The + and − signs against the symbols for voltages indicate the polarities that exist when the instantaneous values are called positive. V_{out}/V_{in} is the gain with feedback, or closed-loop gain. A is the forward gain, or open-loop gain. From the diagram it is evident that

$$(\beta V_{out} + V_{in})A = V_{out}$$

(Except at middle frequencies, the + sign must be taken to mean addition taking account of phase angle.)
From the above

$$V_{out}(1 - A\beta) = AV_{in}$$

$$\text{or } V_{out}/V_{in} = \frac{A}{1-A\beta} \qquad (1)$$

This formula may be regarded as the universal feedback formula, and is just as relevant to positive-feedback applications such as Q-multipliers and some active filters as it is to negative-feedback amplifiers. At medium frequencies, where it will be assumed there are no unwanted phase shifts, A should be taken as a simple negative number if the amplifier phase inverts, β should be taken as negative if the output from the β network is subtracted from V_{in} instead of being added as shown. For a negative-feedback amplifier $A\beta$ will be negative at medium frequencies.

Sometimes the denominator of (1) is given as $1 + A\beta$, and then only the magnitudes and not the signs of A and β are to be inserted in the formula. The formula is specifically a *negative-feedback* formula, and the corresponding formula for *positive* feedback then has a denominator $1 - A\beta$. This is surely an unnecessary complication, which can lead to confusion in some applications where it is not immediately obvious whether the feedback is to be treated as positive or negative.

The loop gain is the gain right round the feedback loop, and is $A\beta$ in Fig. 1. This concept is simple enough in the ideal context of Fig. 1, but in many practical circuits some care must be taken when calculating or measuring the loop gain. For example, how do we calculate the loop gain in Fig. 2? If the loop is broken by removing the connection between P and Q, and a test voltage V_t is applied between P earth, then this would produce, at the junction of R_2 and R_3, with Tr_1 removed, a voltage of $V_t\beta$. This voltage is effectively applied to the emitter of Tr_1 in series with a resistance of $R_2R_3/(R_2+R_3)$, which appears in series with $1/g_m$, reducing the effective mutual conductance of the stage. Alternatively we may calculate the value of R_2 and $1/g_m$ in parallel, and use this value in place of R_2 for calculating the actual feedback voltage appearing at the emitter due to the test voltage V_t. In obtaining the relevant output voltage from Tr_2, knowing its collector current, it is necessary to add a load resistor between Q and earth of the same value as that previously provided by the feedback network.

Fig. 3 illustrates the meaning of the terms series, shunt, current and voltage feedback. It will be seen that the convention is that 'series' and 'shunt' relate to the way the feedback is injected into the input circuit, whereas 'voltage' and 'current' relate to the manner in which the feedback is derived in the output circuit. Voltage feedback causes the load to be fed as from a generator whose internal impedance, or output impedance as it is often called, tends to zero as the amount of feedback is increased, whereas current feedback causes the output impedance to tend to infinity with increasing feedback.

Fig. 4 shows how a combination of voltage and current negative feedback may be used to produce an amplifier with a prescribed value of resistive output impedance, such as might be required, for example, when feeding into a telephone line. This technique is less wasteful of available output power capability than is the alternative of

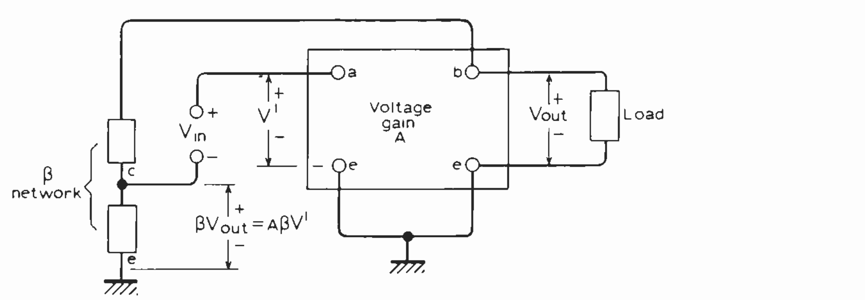
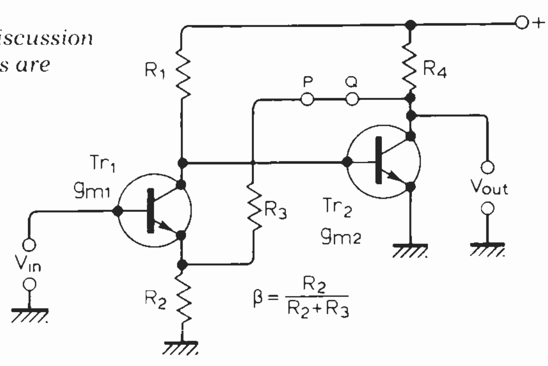


Fig. 1. Basic feedback-amplifier circuit.

Fig. 2. Circuit relating to discussion of loop gain. Biasing details are omitted for clarity.



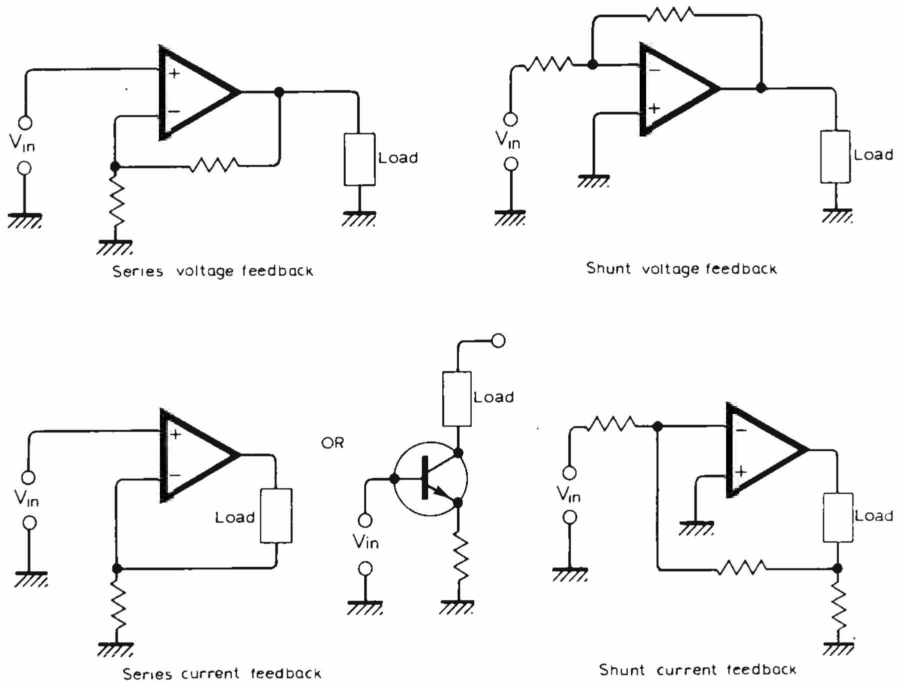


Fig. 3. Four different types of negative feedback.

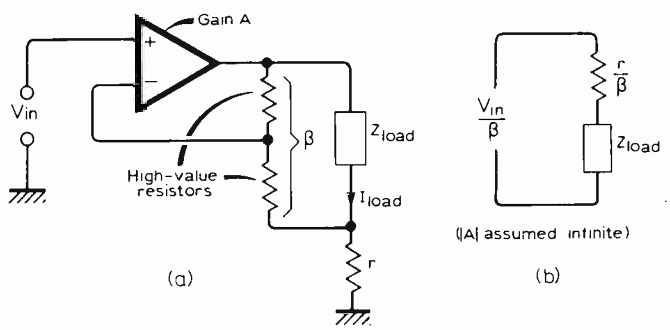


Fig. 4. (a) Feedback circuit with combined voltage and current feedback; (b) equivalent circuit as seen by load.

using an amplifier with simple voltage or current feedback, in association with a resistor equal in value to the required output impedance.

Considering Fig. 4(a), and assuming the ideal case of an infinite-gain amplifier, it is evident that

$$\beta V_{load} + r\,I_{load} = V_{in}$$

$$\text{or } \beta(Z_{load}\,I_{load}) + r\,I_{load} = V_{in}$$

which gives

$$I_{load} = \frac{V_{in}}{r + \beta Z_{load}}$$

or

$$I_{load} = \frac{V_{in}/\beta}{r/\beta + Z_{load}} \qquad (2)$$

This shows that the equivalent circuit must be as in Fig. 4(b). By arranging for the voltage drop across r to provide positive instead of negative feedback, a negative resistive output impedance may be obtained.

Amplifiers are often said to have x decibels of negative feedback at a specified frequency, and such a statement is open to more than one possible interpretation. It is sometimes taken to mean that $20\log_{10}|\text{loop gain}| = x$, but the normal and preferred meaning is that the amount of negative feedback is such as to reduce the amplifier gain by x dB, due precautions being taken to maintain equal loading conditions before and after closing the loop, as already explained. A little thought in relation to equation (1) will show that these two definitions of the amount of negative feedback are not precisely equivalent, and differ quite significantly when the amount of feedback is small. With the preferred definition, feedback is negative at a given frequency if it reduces the gain and positive if it increases the gain. Frequently a practical negative-feedback amplifier will exhibit a peak in its frequency response at high frequencies, near the unity-loop-gain frequency. In the region of the peak, the gain may be higher with feedback on than without it, so that the intended negative feedback has here become positive feedback.

It is sometimes said that feedback is negative if the real component of the feedback voltage, βV_{out}, is in antiphase with V'_{in} Fig. 1, V' itself being taken as purely real, and that feedback is positive if the real component of βV_{out} is in phase with V'. This, however, is a popular misconception, and is quite inconsistent with the distinction between positive and negative feedback given above – as will become evident from the discussion of phase relationships later in this series.

Stability considerations

The subject of stability in feedback systems is a vast one, on which many learned and highly mathematical treatises have been written. The most famous are probably those of H. Nyquist[1] and H. W. Bode[2], both of Bell Telephone Laboratories. Though old, these contributions deal with the fundamentals of the subject thoroughly and in depth, and are still regarded as absolutely sound. Many electronic engineers such as myself, particularly those lacking any formal training in feedback theory, are liable to feel rather overwhelmed by the amount and complexity of the available literature, and concepts such as complex frequency, poles and zeros, contour integration, the Heaviside operator, Laplace transforms and signal-flow graphs seem like insurmountable barriers to some people. However, I believe that the vital thing is to acquire sufficient theoretical understanding to be able to appreciate vividly the reasons for the various effects that occur, and what the available possibilities are for modifying the circuit design as first conceived to give optimum performance. The amount of detailed theoretical background necessary to achieve this is in fact surprisingly small – though some of the mathematical enthusiasts will probably deny this!

There are several reasons why it is unnecessary for a good amplifier designer to know as much mathematical feedback theory as is sometimes supposed. Firstly, much of the fundamental analysis was originally done to *find out* what the stability criteria were, and how they could be expressed in forms convenient for engineers to use. This having been done, and being well established, the engineer can use the results without needing to be able to prove them. Secondly, provided there is a proper qualitative understanding of the problem, the precise optimum values of some components are often best determined experimentally. This is largely because, at the quite high frequencies involved – which may extend up to several MHz – some degree of approximation to the true transistor behaviour would inevitably have to be adopted in a purely theoretical, perhaps computer-aided, design approach. Some people may say that arriving at optimum values for some components by trial and error does not constitute a respectable modern design technique, but I cannot agree with this outlook. One way to regard such a trial-and-error approach is to say that one is using the actual amplifier circuit itself as an analogue computer – changes are made to the circuit values and the results are displayed in analogue form on an oscilloscope. If carried out in an intelligent manner, this seems to me to be a much more direct, economical and generally sensible technique than that of forming a mathematical model of the circuit for processing by a digital computer, but I recognise that what is best done depends a good deal on the background and preferences of the designer.

In some quarters there is a belief that the circuit designer himself should spend his time in an office with paper and a computer, and leave the practical work to others, but I do not think that this philosophy is the most effective one. Experimental work is very stimulating – some unexpected effect is observed, and in a flash one may see that a modification to the circuit would be an improvement. This can often be tried immediately, and may lead to prolonged thought and further ideas. At some point a theoretical analysis may be called for, followed by more experimenting. It is this continuous alternation of experimental and theoretical activity that leads, in my experience, to the evolution of novel and improved designs. Of course, an almost inevitable result of such activity is often that what started off as a neat experimental board tends to have become a somewhat untidy bird's nest at a later stage. However, I think most amplifiers having any real originality of design have probably evolved through such a stage before reaching that of an elegant printed-circuit board.

A very real danger is that if an engineer becomes too absorbed in advanced mathematical techniques, he may fail to give enough attention to other more down-to-earth, but very important, aspects of the overall design work. In a contribution some years ago[3], I said "whilst it is virtuous to be able to analyse a circuit, it may be even be more virtuous to be able to see that a detailed analysis is unnecessary, or to invent a better circuit whose behaviour is more easily predicted."

The aim in what follows will be to present the minimum theoretical background which is thought to be necessary for anyone undertaking to design the feedback stabilization aspects of an audio amplifier with understanding and in a properly optimized manner. Little more than the j-notation[4] will be employed. However, some readers will doubtless wish for a rather broader background of theory, since much published literature on amplifier design uses the concepts of complex frequency, poles and zeros etc. At a fairly elementary level, the excellent series of articles by "Cathode Ray" (M. G. Scroggie) in this journal in 1962 may be recommended[5, 6, 7, 8]. A more advanced and complete treatment of feedback theory and practice will be found in a very good book "Amplifying Devices and Low-Pass Amplifier Design" by Cherry and Hooper[9]. Though they do not hesitate to use determinants etc. when thought to be appropriate, a true engineering outlook is evident and the book contains much very enlightened practical advice on design aspects.

In a.c. coupled amplifiers, stability problems arise at both low and high frequencies. Only the high-frequency problems will be considered here, i.e. all circuits will be treated as d.c. coupled amplifiers, but the principles discussed are very easily adapted, in common-sense ways, to the low-frequency situation when necessary.

Some simple notions about transfer functions will first be considered, because understanding these helps one to appreciate better how the whole negative-feedback story fits together. A transfer function for a feedback amplifier, or any other circuit, is simply an equation giving V_{out} as a function of V_{in}. It is normally assumed that the amplifier is free from non-linearity distortion, but apart from this reservation, the transfer function contains all the necessary information about the frequency response, phase response, transient response and stability margins of the amplifier. The snag is that, except in quite simple cases, deriving and simplifying the transfer function for a feedback amplifier is exasperatingly tedious, even for those with a natural aptitude for such things, which I certainly do not have! The Nyquist diagram, and Bode amplitude and phase plots considered later, represent a vastly more convenient and practicable approach for most amplifier design purposes.

However, it is always theoretically possible simply to use the j-notation to calculate the currents and voltages everywhere in the amplifier circuit due to V_{in} and V_{out}, and thus to form the transfer-function equation. Purely as an illustration of the ideas involved, consider the simple and somewhat idealized circuit of Fig. 5. Using the j-notation gives the current in C_2 as $j\omega V_{out}C_2$. The current in R_4 in the direction shown is V_{out}/R_4. The current in R_3 is the sum of these currents, enabling one to calculate V'. Continuing on these lines leads to the result:

$$V_{in} = -V_{out}R_{in}/R_1[1 + j\omega C_2R_3 + R_3/R_4 + j\omega C_1R_2\,(1 + j\omega C_2R_3 + R_3/R_4) + j\omega C_2R_2 + R_2/R_4 + R_1/R_4] \qquad (3)$$

This as it stands is not much use, for one cannot easily see the physical significance of it. The vital thing when deriving transfer functions is to continue until they have been got into a nice tidy, recognisable form. By collecting terms and rearranging, equation (3) can be got into the form:

$$V_{out}/V_{in} = K \times \frac{1}{1 + j\omega T_1 - \omega^2 T_2^2} \qquad (4)$$

K in this is given by:

$$K = \frac{R_1R_4}{R_{in}(R_1 + R_2 + R_3 + R_4)} \qquad (5)$$

T_1 and T_2 are time constants, each given by a somewhat cumbersome expression with several terms in. One can, moreover, very usefully go a stage further than (4), and get it into the form:

$$V_{out}/V_{in} = K \times \frac{1}{1 + (1/Q)j\omega T - \omega^2 T^2} \qquad (6)$$

Here T is obviously equal to T_2 of equation (4), and we also must have $(1/Q)T = T_1$, giving $Q = T/T_1$, i.e.:

$$Q = T_2/T_1 \qquad (7)$$

Now the physical significance of (6) is instantly apparent if one knows how to "read" it. Q is the Q of a tuned circuit arranged as in Fig. 6(a), having a resonance frequency given by $\omega_0 = 1/T$.

Sometimes transfer functions such as

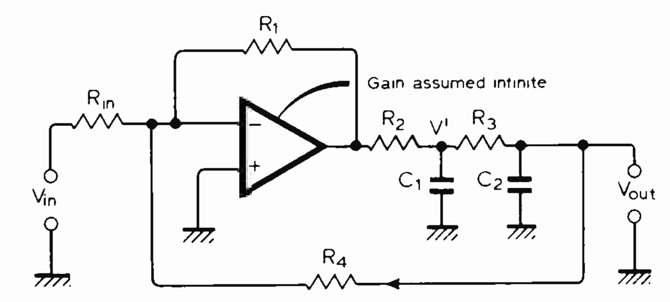


Fig. 5. Circuit relating to discussion of transfer functions.

(6) are given in the form:

$$V_{out}/V_{in} = K \times \frac{1}{1 + (1/Q)pT + p^2T^2} \quad (8)$$

Comparing (6) and (8) it is evident that p = jω. Though it is perfectly all right, in a sine-wave context, to regard p simply as a convenient abbreviation for jω, its full significance is much deeper, for it is Heaviside's operator and means d/dt. Equations such as (8) are thus applicable not only under sine-wave conditions, but also for any other kind of input waveform. Mathematical techniques are available whereby, given the amplifier transfer function, the output waveform resulting from a voltage step or other transient input may be calculated. But in view of the ease with which such responses may be obtained using an oscilloscope, the actual need for such mathematical techniques seldom if ever arises in normal amplifier design work, in my experience. Sometimes when the transient response of an experimental amplifier circuit is under consideration, it is convenient to make up a little simulator circuit, in which all time-constants have been increased by a factor of, say, a thousand compared with the real circuit. The idealized response can thus be obtained, and the relationship between this and the response of the original circuit may shed light on the significance of stray capacitance or other overlooked effects in the latter. The ready availability of type 741 operational amplifiers makes it very quick and easy to do such tests.

Heaviside's operational calculus tends to be somewhat out of favour nowadays, but a very strong case in its favour is presented by two authors from the BBC Research Department in reference 10. It is argued that the technique gives a much better physical insight into the nature of the problem being investigated than do the alternative mathematical techniques available.

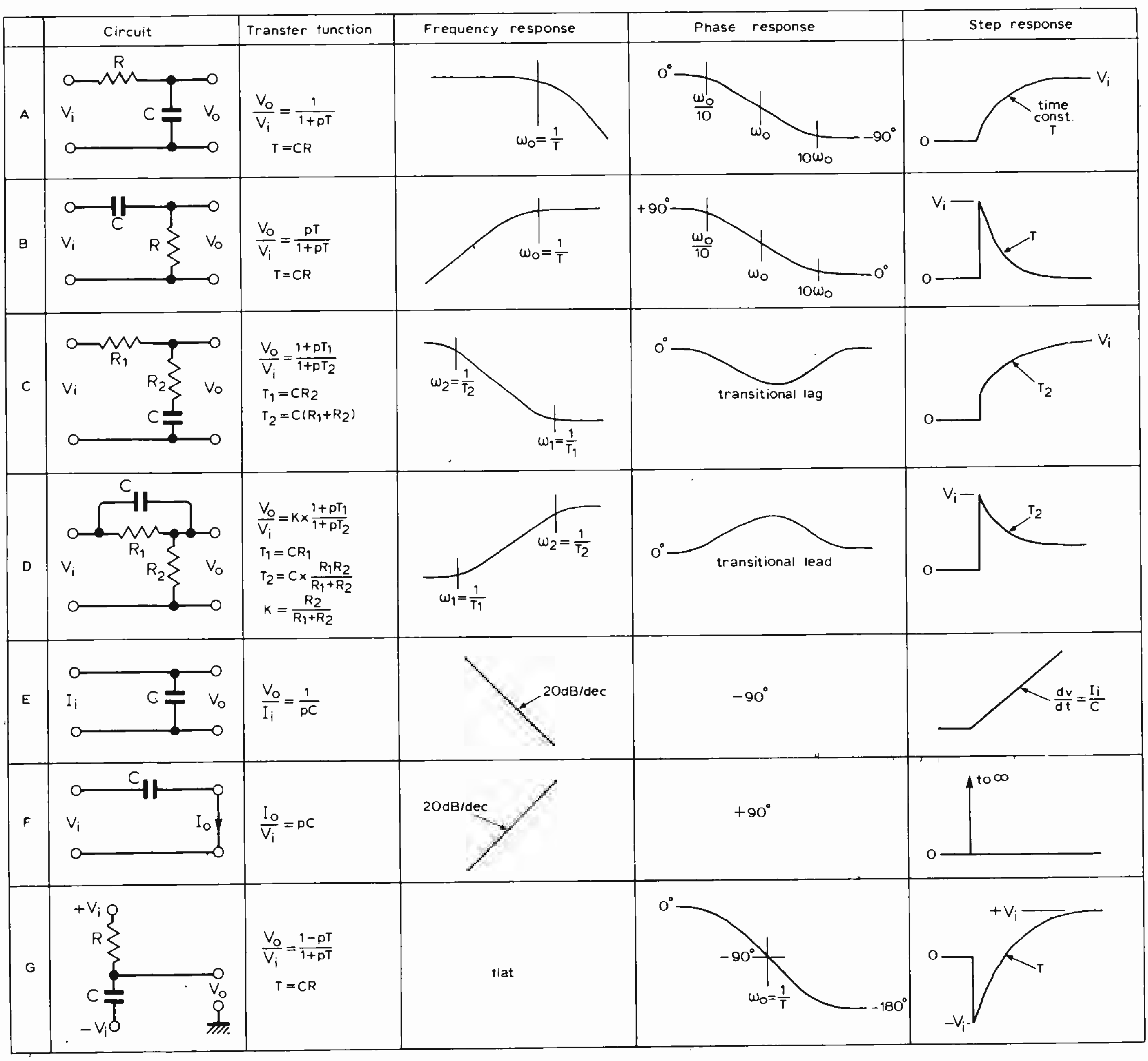
For amplifier designers, the important things to appreciate about transfer functions may be summarized as follows:

(a) Any linear network or amplifier has a transfer function.

(b) However complex the network or amplifier may be, the denominator of the transfer function — if you're clever enough — can be got into the form of a number of factors, which may be either quadratic ones as in equation (8), or simpler ones of the form (1 + pT).

(c) If any of the quadratic factors in the denominator have negative Q, i.e. negative damping, the system will be unstable.

(d) The numerator can take various forms according to whether the system has a low-pass, band-pass or high-pass type of response, and whether there are notches in the frequency response or not.



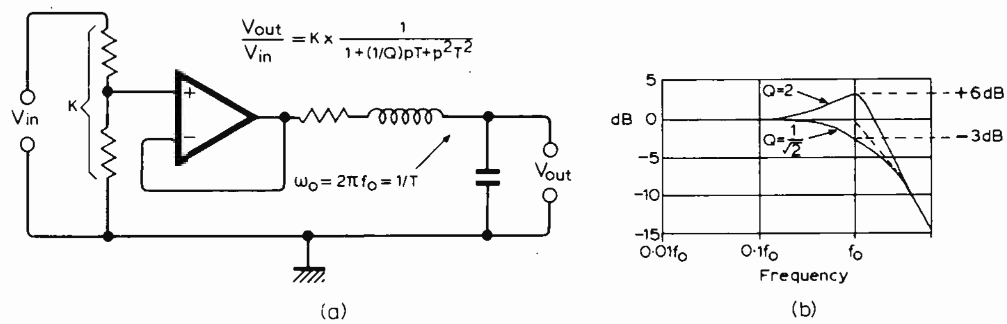


Fig. 6.(a) Circuit giving same response as Fig. 5; (b) and (c) show the frequency response and the step response respectively for two values of Q. $Q=1/\sqrt{2}$ gives second order Butterworth response.

(e) Any required response characteristic whatever can be obtained from a combination of suitably-designed feedback amplifiers, without the need for any inductors, this being the basis of the whole subject of active filters.[11]

Though it is seldom sensible to try to derive the overall transfer function of a complete feedback amplifier, except in the relatively simple cases which usually apply in active-filter design, it is quite important to be able to derive the transfer functions of parts of the circuit of a feedback amplifier, for this is really the basis of most practical design work on such amplifiers. The table gives some simple networks familiar to most readers, together with their transfer functions and frequency, phase and step-input responses. The relevance of the all-pass case G will become evident later. Though the transfer functions may be worked out using the j-notation, and p substituted for jω at the end, it is really more convenient to work with p from the beginning. Thus the impedance of a capacitor is $1/pC$ and the impedance of an inductor is pL. Suppose, for example, we have R and C in parallel. The total impedance is given by

$$Z = \frac{R \times (1/pC)}{R + (1/pC)}$$

Multiplying top and bottom by pC gives

$$Z = \frac{R}{1 + pCR} \qquad (9)$$

This is therefore the ratio V_{out}/I_{in} for the network, and as would be expected it has the same form of transfer function as network A in the table.

A simple illustration of the practical utility of thinking of transfer functions in terms of p rather than jω arises if one considers the problem of determining the output waveform to be expected from network B in the table when the input waveform is a linear voltage sweep, or ramp. One simply "operates upon" the input waveform with bits of the transfer function in turn, chosen in the order that makes things easiest. Thus the ramp waveform multiplied by pT, i.e. differentiated, gives a step waveform. The step multiplied by $1/(1+pT)$ gives an exponential output waveform as shown at the top right-hand corner of the table. A particularly lucid and easy-to-understand paper dealing with topics such as this was written just after the war by Professor F. C. Williams[12]. Though the practical circuits are, of course, all valve ones, the lengthy discussion of the overall design philosophy is highly relevant to present-day problems. The aim was to evolve reliable circuits of precision performance, suitable for trouble-free production, using the minimum of mathematics. Acknowledgement is made to A. D. Blumlein for having provided much of the early inspiration for this work. Some of these pulse circuit ideas are of greater interest to audio engineers than in the past, even in the non-digital field, because of the increased attention now being given to transient response and impulse measuring techniques.

In planning the feedback stabilization details for most audio amplifiers, the normal practice is to think in terms of the rate at which the loop gain is attenuated with rising frequency, bearing in mind all along that the transient behaviour is closely related to this. The relevant techniques will be discussed in the next article.

Corrections to January 1978 article

In Fig. 1, a resistor should be inserted in series with Tr_4 emitter. The arrow in Tr_1 collector lead should be labelled "I_{dc}." In equation (6), the denominator should be "$2\pi\hat{V}_{in}$". The equation just below equation (6) is completely wrong and should be:

$$\frac{\text{slew-rate limit}}{\hat{V}} = 2\pi f_{crit} \qquad (7)$$

In Fig. 3(a), the top waveform was inadvertently cut off at the bottom and should be a complete sinewave. Apologies for the bad reproduction of these waveforms. In the fourth line of the footnote on page 55, the word "is" should be inserted before "approximately". On page 56, first column, 14 lines from the bottom, the word "amplifier" should be inserted between "the" and "slew-rate".

References

1. Nyquist, H., Regeneration Theory, *Bell System Tech. J.*, Jan. 1932, p.126.
2. Bode, H. W., Network Analysis and Feedback Amplifier Design. (van Nostrand 1945).
3. Baxandall, P. J., Papers for the Practising Designer, Letter to Editor, *J.I.E.R.E.*, Dec. 1968.
4. Cathode Ray, "j", *Wireless World*, Feb. 1948.
5. Cathode Ray, Transfer Functions, *Wireless World*, April 1962, pp.177-181.
6. Cathode Ray, Poles and Zeros, *Wireless World*, May 1962, pp.225-229 and June 1962, pp.289-294.
7. Cathode Ray, Differential Equations, *Wireless World*, July 1962, pp.333-337.
8. Cathode Ray, Excitations and Responses, *Wireless World*, Aug. 1962, pp.379-383, Sept. 1962, pp.447-450 and Oct. 1962, pp.507-511.
9. Cherry, E. M. and Hooper, D. E., Amplifying Devices and Low-Pass Amplifier Design. (John Wiley 1968).
10. Head, J. W. and Mayo, C. G., Unified Circuit Theory in Electronics and Engineering Analysis. (Iliffe 1965).
11. Girling, F. E. J. and Good, E. F., Active Filters, *Wireless World*, Aug. 1969 to Dec. 1970 inc. 16 parts; see particularly Sept. 1969, pp.403-408. (Note: In these articles *q* is used in place of Q in equations such as my eqn. (8), Q being reserved for bandpass filters, where it has a somewhat different significance.)
12. Williams, F. C., Introduction to Circuit Techniques for Radiolocation, J.I.E.E., Vol. 93, Part IIIA, No. 1, pp.289-308 (1946).

Advertisement correction

We have been asked by E & L Instruments U.K. to inform readers that there is an error in their current series of advertisements in *Wireless World*. This is an omission of the fact that Quarndon Electronics, Slack Lane, Derby, are also making the £12.50 special offer for the SK10, cash with order. Quarndon are also implementing, on behalf of E & L Instruments, the lifetime guarantee on the SK10 sockets.