Transistor Amplifier DESIGN

4 NEGATIVE FEEDBACK

By A. Foord

NEGATIVE feedback occurs when a proportion of the output voltage of an amplifier is fed back to the input in such a way as to reduce the overall gain, so that the gain with feedback is less than the gain without feedback. Feedback can be used to:

(1) Give a predictable mid-band gain, the greater the amount of feedback the less sensitive the amplifier is to changes of transistor characteristics.

(2) Increase bandwidth or to give a shaped frequency response curve which depends almost solely on the passive components forming the feedback network, and does not depend on an accurate knowledge of transistor parameters (which may vary between one specimen and the next of the same type).

(3) Increase or decrease input or output impedances; by using different feedback arrangements it is possible to obtain input or output impedances higher (or lower) than those normally associated with transistor stages. In particular it is possible to arrange for a high input impedance and a low output impedance, so that amplifier stages can be cascaded without interaction.

(4) Reduce the distortion which normally occurs in the final stages of an amplifier, where current and voltage swings are highest.

BASIC PRINCIPLE

Any study of negative feedback begins with a functional block diagram, Fig. 4.1.

The circuit has two signal paths: the forward path, which is usually an amplifier and contains all the active devices, is marked with its voltage gain A; the feedback path B which has a gain less than unity. The bar above the symbol A indicates that there is a phase reversal in the amplifier, while B represents the fraction of the output voltage fed back to the input.

The phase reversal to obtain negative feedback occurs in the amplifier; the arrows in the block diagram reassure us that the feedback is in fact negative. In a simple case this is obvious, but for more complicated arrangements a check will ensure that we do not accidentally use positive rather than negative feedback.

Take a simple example, Fig. 4.2. Without feedback we require an input of 0·1V to the amplifier for an output of −10V. If negative feedback is added making B equal to 0·1 times, 0·1V is still required at the input of the amplifier itself to obtain an output of −10V.

The input to the addition point needs to be 1·1V, so that when the −1·0V is added to the input we are left with 0·1V to give an output of 10V. The overall gain with feedback is now

$$G' = \frac{-10}{1 \cdot 1} = -9 \cdot 1 \text{ times}$$

The negative feedback has reduced the gain from its open loop value of 100 times to a closed loop gain of 9·1 times, therefore the closed loop gain is approximately $1/B$ times ($1/B = 10$).

A more detailed examination will show that the gain with feedback is approximately $1/B$ provided the closed loop gain is much less than the open loop gain, Fig. 4.3.

Starting at the input to the amplifier (call this V) then the output of the amplifier is $\bar{A}V$, the voltage fed back is $\bar{A}VB$, and the input is $V + AVB$. Then the overall gain with feedback is given by

$$G' = \frac{V_o}{V_i} = \frac{\bar{A}V}{V + AVB} = \frac{\bar{A}}{1 + AB} = \frac{-1}{\frac{1}{A} + B}.$$

If B is much greater than $1/A$ (i.e. A is much greater than $1/B$) then $G = 1/B$.

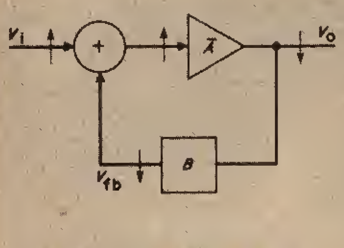

Fig. 4.1. Functional diagram of an amplifier with negative feedback path

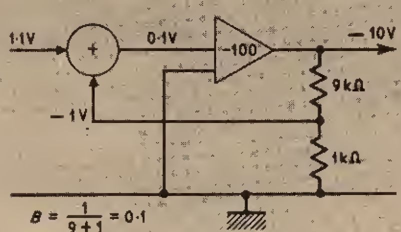

Fig. 4.2. A simple example of negative feedback

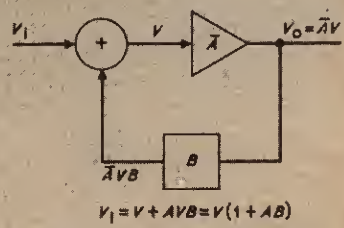

Fig. 4.3. Voltage relationships between the amplifier and the feedback circuit

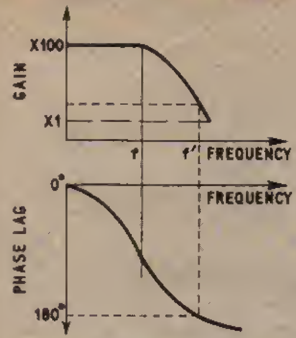

Fig. 4.4. Gain and phase characteristics which would cause instability

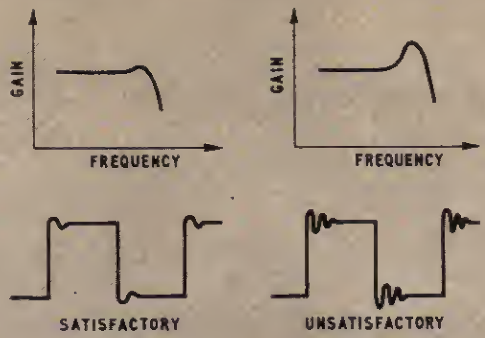

Fig. 4.5. Instability shown by "ringing" cycles

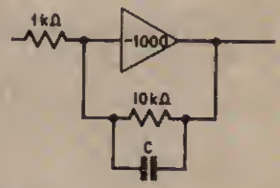

Fig. 4.6. The practical example of a high gain amplifier with feedback resistor and "anti-ringing" capacitor C

Therefore, the ideal gain with feedback (G') is equal to $1/B$, if the gain without feedback (A) is much greater than $1/B$.

To confirm this, compare the results using the accurate formula and the approximate formula; drop the minus phase sign since we are interested in magnitudes rather than the phase reversal we know occurs in the amplifier.

Actual closed loop gain $$G' = \frac{A}{1 + AB}$$

Ideal closed loop gain $$G = \frac{1}{B}$$

Calculation will show that, if the required gain with feedback is one tenth of the gain without feedback, then we do not need to use the accurate formula, since errors in assuming G equal to $1/B$ are small enough to be discounted.

STABILITY

Having assumed that the design of the amplifier is such that the feedback will always tend to reduce the gain, but unfortunately this will not always be so, any practical amplifier will contain reactive elements which will introduce a phase shift in the signal as it passes through the amplifier (quite apart from the 180 degree mid-band phase shift required to obtain mid-band negative feedback). The gain and phase characteristics of the amplifier might appear as Fig. 4.4.

Above a certain frequency f, gain falls and an extra phase lag is introduced. If we applied 100 per cent negative feedback to an amplifier with this characteristic, to give an overall gain of unity, the amplifier would oscillate.

While there is still greater than unity voltage gain around the loop, there is an extra 180 degrees of phase shift to cause the feedback (which was negative below f) to become positive at f'. The system would therefore oscillate at the frequency f'.

Designing for stability is complicated when a considerable amount of feedback is applied. Instability in a feedback amplifier is shown by a peak in the frequency response curve and ringing on a square wave signal (see Fig. 4.5).

Feedback over one or two stages is normally safe, although later on, when considering the use of 100 per cent feedback to raise input impedance, a non-mathematical approach will be applied to the stability problem. For most purposes, it is in order to see that the frequency response curve does not have a peak of more than a couple of decibels in it, and that the square wave response is satisfactory, i.e. free from ringing effects.

The photographs show results obtained with a high gain amplifier, this had a gain of 1,000 times (60dB) without feedback, and a gain of 10 times (20dB) with feedback (see Fig. 4.6).

Photo A shows the leading edge of the 1kHz square wave input; photo B shows the ringing on the output waveform without the capacitor C, and this was considered unsatisfactory. The capacitor was adjusted in value to obtain the acceptable response of photo C. Since for clarity the photographs only show the leading edge of the square wave, the time scale was extended to show the leading edge more clearly. The capacitor was increased to reduce the bandwidth to 20kHz which increased the rise time to that shown in photo D.

Feedback around one stage only is called local feedback, and since only the common emitter stage provides a phase reversal of its output signal with respect to its input, it follows that local feedback can only be applied around the common emitter stage.

There are two basic ways of applying feedback to the common emitter stage; one arrangement is considered next, and the other is dealt with later, in the section on virtual earth amplifiers.

SERIES LOCAL FEEDBACK

Referring to Fig. 4.7, the resistor R_E in series with the emitter accounts for the applied feedback, this resistor enables a feedback voltage V_f proportional to load current, to be fed back in series with the input voltage V_i. The base-emitter voltage of the transistor is reduced by the feedback so that V_{BE} is less than V_i.

If R_E is small and of the same order of value as the internal emitter impedance r_E (say 100 ohms or so), then the amount of feedback is small. Distortion is slightly reduced and bandwidth is increased very slightly, at the expense of a small decrease in gain. If R_E is large then the gain is given by

$$G = \frac{V_o}{V_i} = \frac{R_L}{R_E}$$

The input impedance is given by

$$Z_i = h_{fe} R_E$$

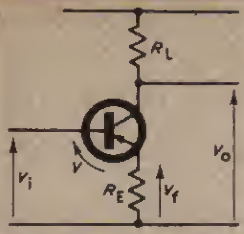

Fig. 4.7. Series local feedback by the unbypassed emitter resistor

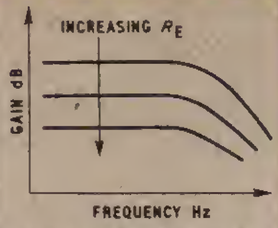

Fig. 4.8. The effect of R_E on gain and frequency response

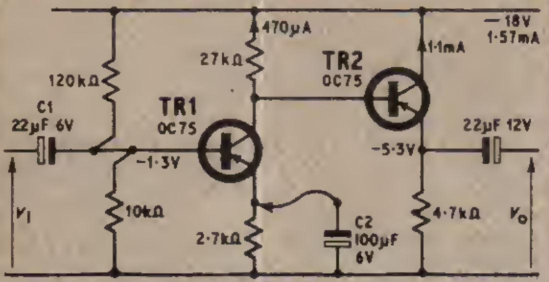

Fig. 4.9. Emitter follower output to prevent undue loading on the amplifier stage

In practice for a single stage amplifier biased in the normal way with a divider chain on the base, this increase in input impedance is masked to some extent by the shunting effect of the chain. The effect of R_E on gain and frequency response is as shown in Fig. 4.8.

To avoid loading R_L and to maintain a high gain without feedback, the output can be taken via an emitter follower, a practical circuit is shown in Fig. 4.9.

Since the transistor is used in common emitter we must work out approximately the bandwidth we would expect. For TR1 the collector current is of the order of 0·5mA. Suppose the transistor current gain is typically 50 at 0·5mA. The 3dB down point in common emitter is given by

$$f = \frac{f_T}{h_{fe}} = \frac{1{,}000}{50}$$

$$f = 20\text{kHz}.$$

One would expect the response to roll off somewhere at this frequency, the exact point depending on h_{fe} and f_T for the particular specimen of transistor. Since TR2 acts as an emitter follower the overall frequency response is limited by TR1 since an emitter follower has a frequency response far better than that of a common emitter stage.

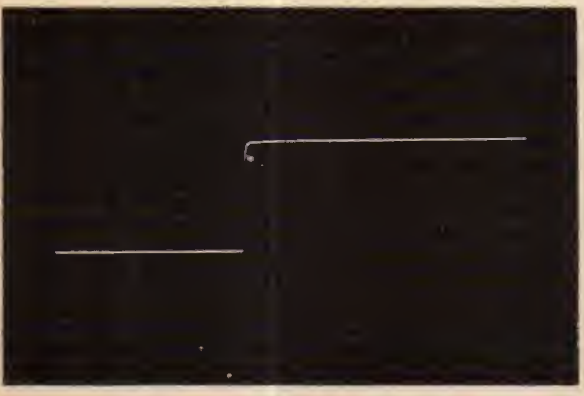

Photo A. *Leading edge of the 1kHz square wave*

Photo C. *Ringing is brought down to an acceptable level by selection of a parallel capacitor across the feedback resistor*

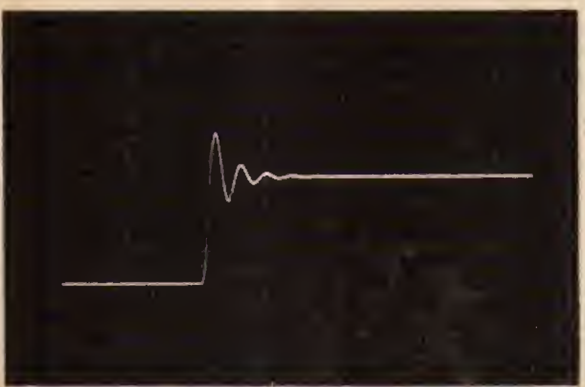

Photo B. *Ringing caused by non-selective feedback*

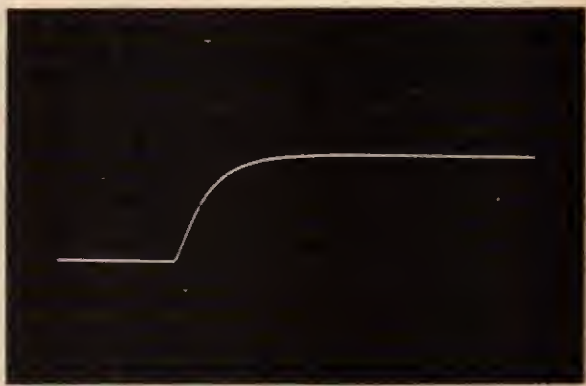

Photo D. *Larger value of capacitance increases the rise time*

The results without feedback (C2 connected) and with feedback are:

Gain without feedback	=	310 times = 50dB
Bandwidth	=	30kHz
Input impedance	=	1kΩ
Output impedance	=	300Ω
Maximum output	=	1·5V r.m.s. no load
	=	500mV r.m.s. into 1kΩ

The frequency response was measured under no load conditions at 100mV r.m.s. Up to 50kHz or so, the amplifier will provide 500mV r.m.s. into 1 kilohm, but above this frequency the emitter follower current gain starts to drop and the waveform distorts, so that 500mV would only be obtained without distortion into a load greater than 1 kilohm.

The gain without feedback appears high until we remember that the collector load is 27 kilohms rather than the 1 kilohm or so we would expect for another common emitter stage and a high gain transistor is being used.

Gain with feedback	=	$\frac{27}{2{\cdot}7}$ = 10 times = 20dB
Bandwidth	=	68kHz
Input impedance	=	7kΩ
Output impedance	=	300Ω
Maximum output	=	500mV r.m.s. into 1kΩ (up to 50kHz)

The actual measured gain was 19·5dB, which is probably an error in measurement or tolerances on the collector and emitter resistors. Since the open loop gain is 30 times the closed loop gain, one might expect the gain of 20dB to be independent of variations in characteristics between one OC75 and the next, although the bandwidth might alter slightly.

Although the gain has been reduced by a factor of 30 times, bandwidth has only increased twice. Series local feedback is often used inside another overall feedback loop, where a predictable stage gain is required rather than an unpredictable (though higher) gain.

With an input impedance of 7 kilohms and an output impedance of 300 ohms, these amplifiers can be cascaded without interaction. Bandwidth is also extended at the low frequency end, but since this is within our control (coupling and decoupling capacitors) the main benefit of this type of negative feedback is the predictability of mid-band gain rather than the small extension of bandwidth.

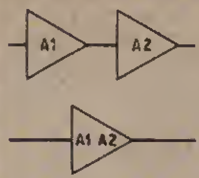

Fig. 4.10. Combining two identical amplifiers

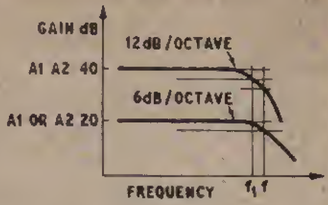

Fig. 4.11. The response of each amplifier is added to give overall response

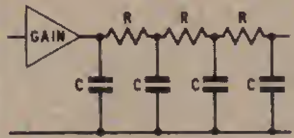

Fig. 4.12. Equivalent circuit of a transistor looks like that of a transmission line at high frequencies

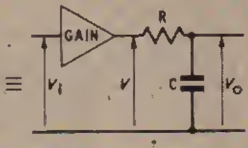

Fig. 4.13. A single CR conversion of Fig. 4.12

TRANSFER FUNCTIONS

The amplifier can be represented as a block, so that two amplifiers in series could be represented by adding the two separate gains in decibels (see Fig. 4.10).

A_1 = 10 times = 20dB $A_1 A_2$ = 100 times

A_2 = 10 times = 20dB   = 40dB

Working on the frequency response curve and adding decibels this would result in the response shown in Fig. 4.11.

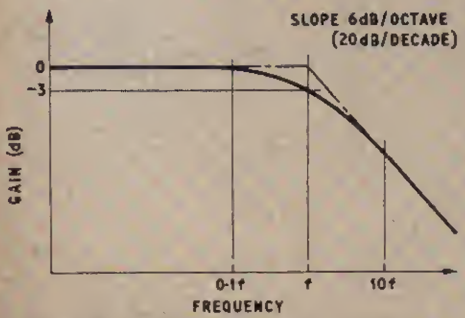

Fig. 4.14. Approximating the frequency response, then deriving the accurate response (solid line)

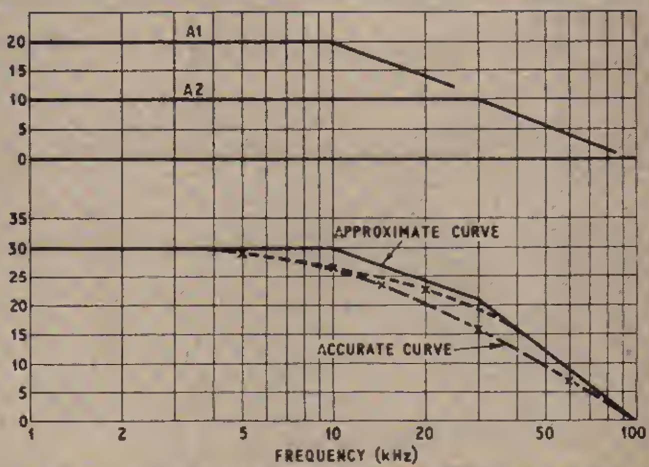

Fig. 4.15. Summing the response of two amplifiers in cascade

Where each amplifier was 3dB down (at f), the response is now 6dB down (for two identical amplifiers), and the new 3dB down point is lower down at f_1, as one would expect. The slope of the curve for a single amplifier is approximately 6dB per octave, and for two amplifiers in series this will be 12dB per octave.

In the equivalent circuit of a transistor (Fig. 4.12) it appears as a transmission line for high frequencies, but as a first approximation it can be considered as a single CR network, Fig. 4.13.

Taking the CR network or single time constant on its own, at low frequencies C has a high impedance and $V_o = V$. At a frequency when C has a reactive impedance equal to R, V_o is 3dB down with respect to V, and continues to fall at 6dB per octave (20dB per decade) with increasing frequency, as in Fig. 4.14.

The solid line curve is the accurate frequency response, while the dotted line is the straight line approximation. The point f where the impedance of the capacitor is equal to the resistor is called the "turnover" or break point. The maximum error between the accurate and straight line approximation is 3dB which occurs at the break point. In practice the approximate curve is drawn; f is 3dB down, $0{\cdot}5f$ and $2f$ are 1dB below the approximate curve; the accurate curve is drawn from this information.

For two amplifiers in cascade the procedure just outlined is shown in Fig. 4.15. The bandwidth of the two amplifiers in series is almost identical, but the ultimate slope is now 12dB per octave rather than the 6dB per octave for a single common emitter stage.

RELATION OF RISE TIME AND BANDWIDTH

A square wave can be considered as the sum of a number of harmonically related sine waves; these include a fundamental sine wave at the basic repetition frequency and frequencies of three, five, seven times, and so on (odd harmonics of the basic frequency).

If a square wave of 1kHz is fed into an amplifier with a sharp cut off at 1·5kHz all the harmonics are filtered out, leaving only the 1kHz fundamental, Fig. 4.16.

If the response of the amplifier is extended to 20kHz the output would consist of the fundamental and harmonics up to 19kHz. Harmonics at 21kHz, 23kHz upwards would be attenuated according to the roll-off of the amplifier response curve. The square wave would hardly be degraded at all, since the amplitudes of these harmonics (relative to the fundamental) are small. The loss in harmonics increases the rise time of the square wave, Fig. 4.17.

Fig. 4.16. A square wave signal can be filtered to give the fundamental frequency only

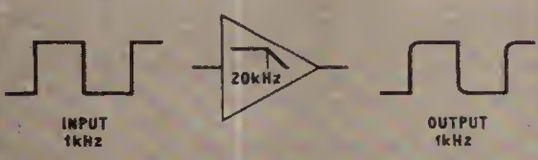

Fig. 4.17. Increase in the rise time of the square wave by loss of harmonics

To determine the bandwidth of an amplifier we would feed in a square wave with a rise time better than we would expect the amplifier to handle, and measure the degradation on the output.

Suppose our square wave had a rise time of 5μs and after passing through the amplifier this was degraded to 25μs, then our amplifier has a rise time of $\sqrt{(25^2-5^2)}$ or 24·5μs and its bandwidth is given by:

$$f = \frac{0{\cdot}35}{\text{rise time}} = \frac{0{\cdot}35}{24{\cdot}5} \times 10^6 = 14{\cdot}3\text{kHz}$$

This method is only an approximate means of determining bandwidth, it would tell us if our amplifier had a bandwidth of 20kHz or 10kHz, but we could not rely on discriminating between bandwidths of 20kHz and 17kHz.

However the edges of the square wave do represent the type of signals present in a transient, which simple sinewave testing cannot do, so that we can see immediately any instability or excessive overshoot or ringing in the amplifier. The disadvantages are that we do need a square wave of good rise time, and an oscilloscope capable of showing it.

Next month: Negative feedback applied to practical circuits.

BETTER SOUND

THE BBC announces that four programmes in a new series "Better Sound" will be broadcast on Fridays at 7.00–7.30 p.m. in Study Session, Radio 3 from May 3 to 24. Listeners will be invited to send questions of general interest, or requests for more information on particular topics covered in the series and these will be dealt with in two extra programmes which will follow the repeat of the series later in the year.

The series will be repeated on Radio 4 on Saturday mornings at 11.00–11.30 a.m. from August 17 to September 14. There will be no programme on August 31 (Bank Holiday weekend), but there will be two additional programmes on Saturdays, September 21 and 28.

Each programme will focus attention on one area of this wide field. A number of topics (e.g. microphones, loudspeakers, stereo) will therefore be treated in more than one programme. Advice on particular makes cannot be given and the construction and repair of equipment will not be dealt with.

Programme 1: Transmission and reception of radio, including stereophonic broadcasting. Explanation of AM and FM, etc.

Programme 2: The nature of sound, and room acoustics, with demonstrations of the effect of different placings of microphones and loudspeakers.

Programme 3: The reproduction of music in mono and stereo; hi fi equipment.

Programme 4: Tape-recording for the amateur.

The diagrams in the Study Notes (BBC Publications, 2/6 plus 5d postage) will be helpful in following the broadcasts and the explanations in the text of the basic principles of the transmission, recording and reproduction of sound in mono and stereo will be useful for reference, particularly for the less knowledgeable listener.