Negative feedback and non-linearity

Exploring the fallacy that n.f.b. reduces all harmonics equally

by 'Cathode Ray'

MR M. G. SALEM has recently called attention (in the July issue, pp. 59-60) to the fallacy, apparently not yet extinct, of supposing that negative feedback reduces all distortion harmonics equally, by the same factor as it reduces the gain of the amplifier to which it is applied. In doing so he mentioned, in effect, that I had put an explosive charge under this particular fallacy quite a long time ago – actually April 1961 under the above title, repeated as Chapter 19 in *Essays in Electronics*. Having myself forgotten doing so, I feel confident that few other than Mr Salem are so familiar with the said work that the following revised version of it will be greeted with widespread cries of protest against excessive repetition.

Undoubtedly the first thing to learn about negative feedback is that it is never so simple as it looks. Superficial study gives one the impression that it reduces undesirable things such as distortion of all kinds and noise, mains hum, etc., dividing them by $(1\text{-}AB)$ or $(1 + AB)$, depending on the conventions adopted. Also that the input and output impedances of the amplifier to which it is applied are either decreased or increased – in the same ratio? The truth is that, even if such complications as phase shifts are excluded, none of these things is necessarily so. The effects on impedance will *not* be in the same ratio. In general, noise reduction won't be, either. Some kinds can even be increased by negative feedback. In simple cases the reducing effect on distortion is more dependable, but even there one can easily go wrong, as in the example Mr Salem pointed out. That example concerned non-linearity distortion, the effect of which is to introduce signal frequencies (harmonics and intermodulation) not present in the original. Reducing non-linearity is usually the main object of negative feedback, because that causes the most objectionable kind of distortion. No amplifier with any claim to be suitable for high-quality sound reproduction would be without negative feedback.

So let us start with a reminder of how it is commonly said to reduce non-linearity distortion. Fig. 1(a) shows an amplifier with an A-fold voltage gain. For every millivolt (say) applied to the input it gives A millivolts out. To simplify things later, we assume that the amplifier is a phase-reversing one, as indicated by the gain being shown as $-A$. Now feed back a fraction B of this output, as at (b). The voltage fed back is thus $-AB$. From the point of view of the input terminals of the amplifier the $-AB$ fed back is in opposition to the signal required between those terminals ($=1$). The signal needed between XX to maintain the amplifier signal level as before is therefore $1 + AB$, of which the $+AB$ offsets the $-AB$ fed back, leaving a net input of 1*.

Fig. 1 thus shows that negative feedback reduces the overall or gross gain of the amplifier from A to $A/(1 + AB)$ – often denoted by A'. At this point all the books mention that if the design makes AB so much larger than 1 that 1 can by comparison be neglected, A' becomes (as near as makes no matter) $1/B$. The great significance of this is that B usually depends solely on something like a potential divider that is perfectly linear, so the non-linearities involved in A are more or less removed. These and other advantages are paid for by the extra amplification needed to make AB very much larger than 1 and at the same time to ensure enough net input.

We now switch attention to the distortion created inside the amplifier by its non-linearity. It can be considered as if due to an additional input, say d millivolts; or, hopefully, microvolts. At first we might suppose that because applying negative feedback reduces the gain from A to $1/(1 + AB)$ then the legitimate signal and the distortion would both be reduced in the same ratio, so the percentage distortion would be unchanged. However, comparing (a) and (b) in Fig. 1 we see that the signal level inside the amplifier, and therefore the amount of distortion, is the same in both cases, whereas the gross signal input is much greater in (b). Therefore distortion as a percentage of the signal has been reduced by feedback in the same ratio as the gain.

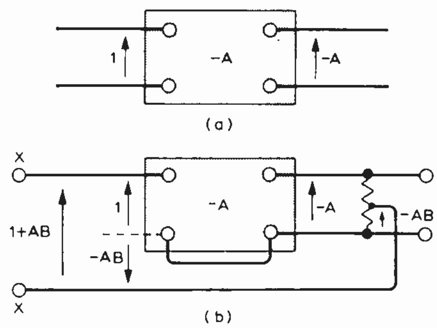

Fig. 1. *(a) represents an amplifier without feedback, and (b) the same amplifier with feedback, a fraction B of the output voltage being tapped off and returned to the input. In this case the amplifier is a phase-reversing one, so its voltage gain is shown as $-A$. The voltage fed back is therefore in opposition to the input voltage, which has to be increased accordingly; i.e., the feedback is negative.*

That is the point at which writer or reader (or both) tend to suppose that this important feedback law has been duly established and they can go on to something else. As an optional extra it may have been noted that if the gain of the amplifier is assumed to be (near enough) the same at all audio frequencies – as of course it ought to be – then the distortion harmonics and intermodulation products are all equally reduced by negative feedback.

But before hurrying on let us consider precisely what we have been meaning by A. We defined it – or, to be fair to you, *I* defined it – as the number of signal millivolts received at the output (Fig. 1(a)) for every millivolt applied at the input. But I didn't insist on millivolts, or on any particular signal level. The same A was assumed to hold good for the (presumably) much lower level of the distortion. In other words, A was assumed to be linear. That being so, it wasn't very clever to use it in a calcula-

*If no assumption is made about the polarity of the amplifier output being negative with regard to the input, the gain being called just A, then the gross input works out as $1 - AB$. This is correct for positive feedback, but for negative feedback either the amplifier or the feedback arrangement has to be phase-reversing, represented by making the value of either A or B negative, thus cancelling out the minus and giving $1 + AB$ as in Fig. 1(b). As we are considering only negative feedback, it seems rather pedantic and unnecessarily confusing to have to remember to use a double negative every time. In practice there are only (usually) two frequencies at which AB is simple plus or minus; for all the rest one has to consider other phase angles than 0 and 180 , using 'complex' algebra. But it is a *very* simple recap we are having, with a view to making just one point, not an exhaustive treatise on negative feedback.