tion concerning amplifier non-linearity. True, we guarded against complete absurdity by making the signal voltage in the amplifier the same in both Fig. 1 diagrams. But if the non-linearity is considerable, so that the distortion is a significant part of the total output, that safeguard isn't good enough. For, when the feedback is applied and reduces the distortion, the total output will be different.

The correct procedure, now that an element of doubt has been found to exist in the basis of the argument, would be to embark on a comprehensive and rigorous mathematical analysis that would cover every case. But you know me too well to expect that. Anyway, the higher the level of maths the greater the risk of going wrong or of the truth being obscured. (Mathematicians, don't bother to write to me on this, for I shall decline to answer.)

The 'line' in 'linearity' is the graph of output against input. These come in two kinds. One of them could be plotted by connecting a calibrated signal generator to the input of the amplifier and varying the signal strength there while measuring the corresponding peak or r.m.s. voltages at the output. It might look something like Fig. 2. There would

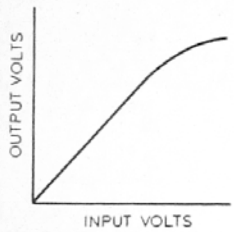

Fig. 2. *This is one kind of output/input graph, in which the voltages are peak or r.m.s. values.*

be no point in reversing the connections with the idea of extending the curve into the negative region, for its shape would necessarily be the same in reverse. The other kind, which is the one we are going to study, is seen by substituting the Y plates of a cathode ray oscilloscope for the output voltmeter, and connecting the X plates (with suitable distortionless amplification) across the input. The positive and negative half-cycles obviously swing the curve in both directions from the origin as their *instantaneous* values are shown on the screen, and their shapes are not necessarily the same.

A perfectly linear amplifier would yield a perfectly straight 'curve', as in Fig. 3(a). In the case of a power amplifier this would merely show it was being uneconomically under-driven. In a commercial world it is necessary to work up to some distortion, even though it be limited to less than 0.1%. Most amplifiers, so long as they are not over-driven, tend to show curves of two main shapes (or combinations of both), as in Fig. 3(b) and (c). The first has a

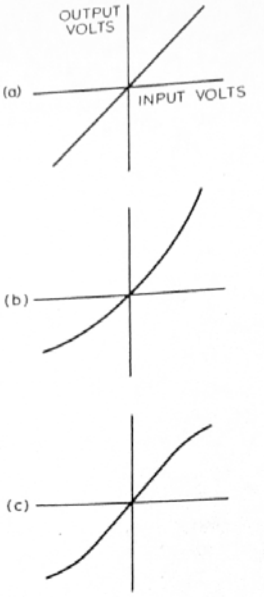

Fig. 3. *In this kind of output/input graph, instantaneous voltages are plotted. (a) is a linear (distortionless) characteristic; (b) and (c) are non-linear curves, representing respectively second and third order distortion.*

square-law term in its output/input equation, which generates a second harmonic of the signal, and second-order intermodulation. The second has a cubic term and generates third-order distortion, which sounds worse.

Now *A* (being output/input) is represented in these Fig. 3 diagrams by the slope of the curve. In (a) the slope is the same throughout, so *A* is constant and (assuming, as we usually can, that *B* is likewise) there need be no question as to exactly what 1 + *AB* means. But in such a situation there is no need for feedback! In (b) and (c), *A* is varying all the time, so one doesn't know what figure to insert for it when using the formula. We can say that Fig. 3(b) indicates a smaller *A* at the negative peaks than at the positive, so presumably the negative part of the curve is straightened out less by negative feedback than the positive part, but the effect on the distortion is difficult to assess without a large-scale mathematical operation. Let us see what we can do without that.

In order to find out whether the harmonic structure of the distortion (as distinct from its amount) is affected by feedback there should be no need to consider any particular practical amplifier. That is just as well, because it would be quite tricky to represent typical crossover distortion mathematically. A single transistor is easier, because it does have a Fig. 3-type graph that is a good approximation to an exponential curve, and (with suitable assumptions) the corresponding array of harmonics in the output can be derived as a basis for calculation. But why bother? Things will be much easier

and clearer if (at least for a trial) we assume we have a hypothetical amplifier with a pure square-law characteristic, like Fig. 3(b), and plotted quantitatively as Fig. 4, using the equation.

$$v_o = 100v_i + 1000v_i^2$$

where v_o is the instantaneous output voltage and v_i the sinusoidal input voltage. This gives the amplifier a gain of 100 as regards the fundamental.

A simple calculation shows that with a peak v_i of 0.04V the $1000v_i^2$ term causes 20% second-harmonic distortion. We can do it graphically by drawing a straight line joining the tips of the curve, noting how far up the v_o axis it comes (1.6V in this case) and lowering the line half the distance. It is then the linear part of the characteristic responsible for the fundamental, shown (dotted) as a pure sine wave in Fig.5 (a). The actual amplifier curve I have plotted in Fig. 4 is 0.8V lower at zero v_i and 0.8V higher at posititve and negative peaks. The points can be transferred to Fig. 5(a), and when joined up by the full line show what comes out of the amplifier when 0.04V peak is put in. The difference between this and the fundamental has been plotted below, (b), and is clearly a second harmonic. Both Fig. 4 and Fig. 5 show that its peak value is 0.8V, which in relation to the fundamental's 4V is 20%.

Anyone with the most elementary knowledge of the differential calculus will realize that the easiest way of finding the slope (which is *A*) at any point on the Fig. 4 curve is to differentiate its equation, thus:

$$A = \frac{dv_o}{dv_i} = 100 + 2000v_i$$

So at zero v_i it is 100, which is what one would expect, since an input confined to very small values of v_i would yield neligible distortion, and 100 is the slope of the fundamental line, corresponding to an amplification of 100. At the

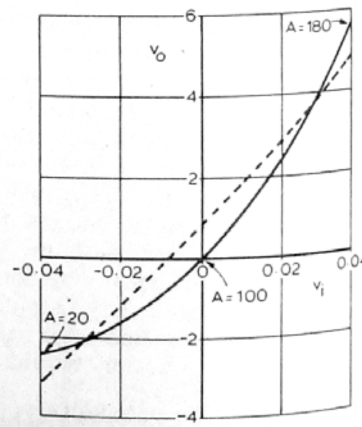

Fig. 4. *The full line is a graph of the Fig. 3 (b) type. The broken line shows its fundamental part; the vertical difference between the two represents second-harmonic distortion, as shown in Fig. 5.*