Estimates of Nonlinear Distortion in Feedback Amplifiers*

EDWARD M. CHERRY, *AES Member*

Department of Electrical and Computer Systems Engineering, Monash University, Clayton, Vic. 3168, Australia (Retired)

A method is proposed for characterizing nonlinear distortion in feedback amplifiers by means of two generators at the input—one voltage and one current. These generators are the negatives of the voltage and current antidistortion which, added to voltage-source and current-source inputs, respectively, would reduce the output distortion to zero. Their values can be calculated easily from known variations of small-signal parameters such as transistor g_m, β, and the capacitances. The distortion-to-signal ratio is given by formulas similar to the offset-to-signal ratio or noise-to-signal ratio, and depends only on the distortion generators, the signal source, and the source impedance. The method correctly predicts harmonic and intermodulation distortion for practical amplifiers. It points toward techniques for reducing distortion, and it explains certain experimental results which hitherto have been problematical.

0 INTRODUCTION

Designers of low-distortion amplifiers must often address questions such as: what are the relative contributions of various nonlinearities to overall distortion? or, what is the effect of adding a local feedback loop? or, what is the effect of shifting a pole, to "tailor" the response of a feedback loop? or, does it matter whether the pole concerned comes before or after a nonlinearity? Definitive answers to these and other questions would be forthcoming if there were an easy method for calculating numerical values of distortion.

Fig. 1(a) shows a simplistic way of representing nonlinearity in a feedback amplifier. Distortion is considered as an extra signal-like component injected at the output of the forward path. Provided the input signal is adjusted so that the output amplitude remains constant as the feedback is varied, any nonlinearity in the forward path is exercised to the same degree, and the distortion generated in the forward path remains constant. Then, using the results of linear feedback theory, conjecture is that

$$y = r\left(\frac{G}{1 + GH}\right) \tag{1}$$

$$z = d\left(\frac{1}{1 + GH}\right) \tag{2}$$

where r and y are the input and output signals, d is the distortion in the forward path, and z is the resulting distortion at the output. In the somewhat more realistic representation of Fig. 1(b) distortion is injected at the output of the nonlinear block, the dependencies on complex frequency s are included, and

$$y(s) = r(s)\left[\frac{G_1(s)G_2(s)}{1 + G_1(s)G_2(s)H(s)}\right] \tag{3}$$

$$z(s) = d(s)\left[\frac{G_2(s)}{1 + G_1(s)G_2(s)H(s)}\right]. \tag{4}$$

Eqs. (2) and (4) do not account for the observed facts. They do not predict the frequencies of the distortion products which are generated, and they do not predict the way the amplitudes of these distortion products change when the loop gain is varied.

For example, a simple quadratic nonlinearity generates nothing but second-harmonic distortion when fed with a sinusoid. Therefore, when feedback is applied around a square-law device such as a FET, the distortion should contain only second harmonic and, from Eq. (4), the harmonic content should decrease with increasing feedback. In fact, a FET with feedback generates a spread of harmonics, and many of these harmonics initially increase in amplitude as the feedback is increased [1]. The seventh harmonic (which some audiophiles regard as particularly offensive) reaches its maximum when the loop gain is around 5 (see Appendix, Sec-

* Manscript received 1998 August 17; revised 1999 September 16.

tion A.4).

Formally, the error in Eqs. (2) and (4) comes about through the misapplication of linear circuit theory to a nonlinear system. An intuitive explanation can be seen in Fig. 2, in the context of the FET example:

1) The quadratic nonlinearity is normalized such that its small-signal gain (the gradient at the origin, see Appendix) is unity and the output saturates at $y = -0.25$, as shown by the heavy curve.

2) The feedback factor H is chosen equal to $(1 - 1/G)$, so that the overall small-signal gain A remains constant at unity when the loop gain is varied by changing the forward-path gain G.

Therefore, if the system input is held constant as G

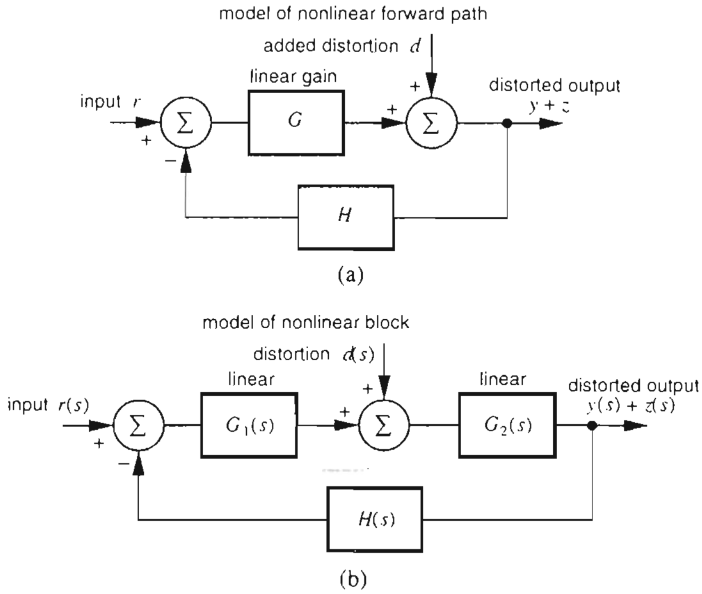

Fig. 1. Simplistic representations of nonlinear distortion in feedback amplifier. It is assumed that the input signal is adjusted so that the output amplitude remains constant when any change to the circuit is contemplated. (a) Basic. (b) Somewhat more realistic.

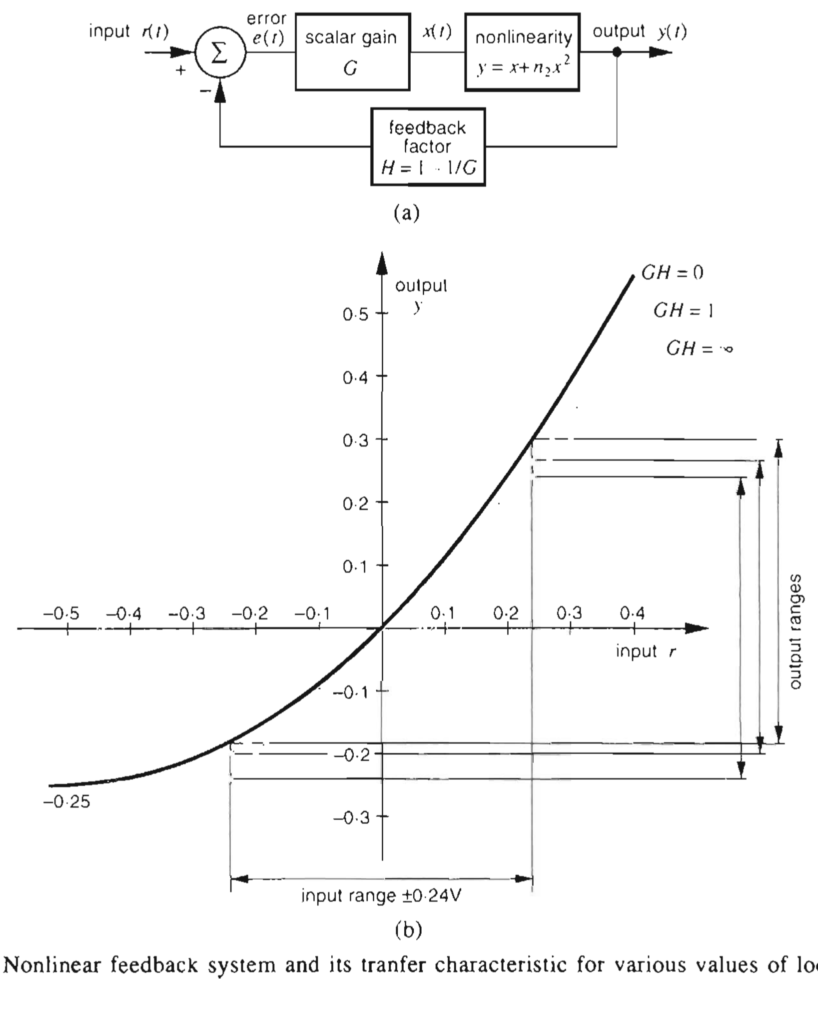

Fig. 2. Nonlinear feedback system and its tranfer characteristic for various values of loop gain.

is varied, the changes to the output waveform fairly reflect the effect of varying amounts of feedback on distortion.

Observe how the shape of the transfer characteristic changes as the loop gain is increased, from a simple parabola to (ultimately) two straight lines. Between these extremes the curve is neither parabolic nor linear, and must have a power-series expansion of order greater than 2. Harmonics other than the second must be generated by a sinusoidal input (see Section A.1).

Included in Fig. 2(b) is an input signal of ±0.24 V peak. Observe also how the range of the output signal changes as the loop gain is varied, even though the overall small-signal gain remains constant. The nonlinearity is not exercised to a constant degree and the additive distortion in Fig. 1 does not in fact remain constant. Therefore, the simplistic approach of dividing the distortion without feedback by the loop gain is fundamentally flawed.

1 DISTORTION REFERRED TO INPUT

The method proposed in this paper for estimating nonlinear distortion in a feedback amplifier is that it should be referred backward to the input rather than forward to the output [2], [3].

Fig. 3 is the block diagram of a feedback amplifier, in which the forward path is split in two—after the style of Fig. 1(b). Between the sections is a static nonlinearity which has unity small-signal gain. In the absence of this nonlinearity the undistorted output $y(s)$ would be

$$y(s) \underset{\text{linear}}{\Rightarrow} r(s)\left[\frac{G_1(s)G_2(s)}{1 + G_1(s)G_2(s)H(s)}\right] \underset{\text{large } GH}{\Rightarrow} r(s)\left[\frac{1}{H(s)}\right]. \tag{5}$$

The problem is to find the distortion $z(s)$. Its value depends both on the detail of the amplifier blocks and on the input signal $r(s)$.

There exists an antidistortion $c(s)$ which, when injected at the input of the nonlinearity, as in Fig. 4, would reduce the output to the undistorted $y(s)$. Because the output $y(s)$ is known precisely, and because the linear transfer function $G_2(s)$ is known, the intermediate signal $u(s)$ can be found. Then $x(s)$ can be found by taking a reversion of the nonlinearity, and finally $c(s)$ is obtained as the difference between $u(s)$ and $x(s)$.

Because $G_1(s)$ is linear, the antidistortion in Fig. 4 can be moved to the input of the complete system. Fig. 5 is exact.

The new approach depends on one key approximation: the complete amplifier with feedback is almost linear. Certainly, it is far more nearly linear than is the forward path—if the feedback is doing its intended job. Therefore the principle of superposition applies. Then if the antidistortion is multiplied by the overall gain and moved from input to output, as in Fig. 6, the output from the complete system is very nearly the undistorted $y(s)$ given by Eq. (5). By subtraction, the distortion at

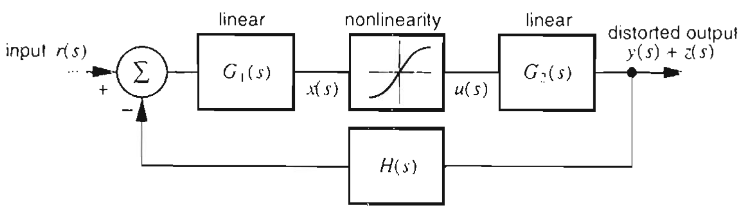

Fig. 3. Feedback amplifier with static nonlinearity in its forward path.

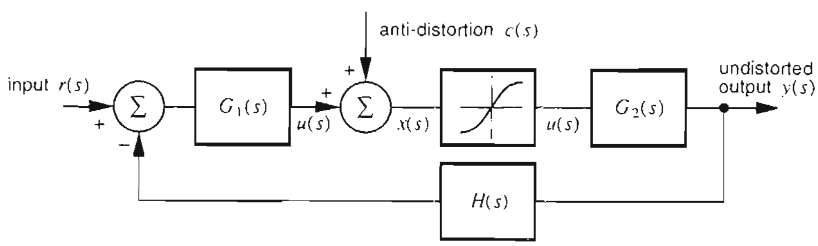

Fig. 4. Amplifier with antidistortion injected at the input of a static nonlinearity to give undistorted output.

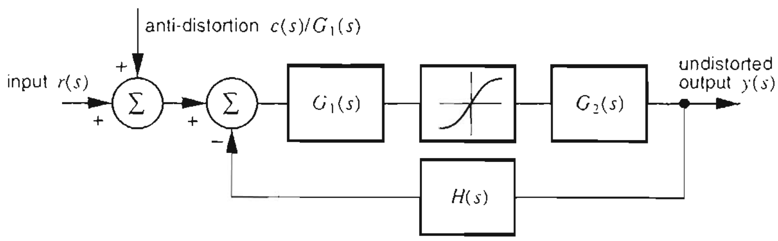

Fig. 5. Amplifier with antidistortion moved to its input.

the output from the amplifier itself must be very nearly.

$$z(s) = -\frac{c(s)}{G_1(s)}\left[\frac{G_1(s)G_2(s)}{1 + G_1(s)G_2(s)H(s)}\right] \underset{\text{large } GH}{\Rightarrow} -\frac{c(s)}{G_1(s)}\left[\frac{1}{H(s)}\right]. \tag{6}$$

The same distorted output would be produced by the linear system shown in Fig. 7.

1.1 Distortion-to-Signal Ratio

Said differently, the distortion-to-signal ratio of a feedback amplifier (tantamount to the percentage distortion) is determined by the ratio of a fictitious distortion input to the actual signal input, the former being the negative of the antidistortion which, when added to the input, would produce an undistorted output,

$$\text{distortion referred to input} = -\frac{c(s)}{G_1(s)} \tag{7}$$

$$\frac{\text{distortion at output}}{\text{signal at output}} = -\frac{c(s)/G_1(s)}{r(s)}\left(\frac{\text{gain at distortion frequency}}{\text{gain at signal frequency}}\right). \tag{8}$$

The whole is reminiscent of the offset-to-signal ratio for a dc amplifier, or the noise-to-signal ratio (in this case mean-square quantities must be used and correlation included). Offset referred to the input of an amplifier is the negative of the dc input that would reduce the output offset to zero. Noise referred to the input is the negative of the noise input that would reduce the output noise to zero.

Observe that the nonlinearity is exercised to almost the same degree in all of Figs. 3 through 6. Because the overall distortion $z(s)$ is small, the output waveforms in Figs. 3 and 6 are very close to the undistorted waveform in Figs. 4 and 5. The fundamental flaw in the simplistic approach has thus been circumvented. Observe too that the method applies even if the forward-path nonlinearity is gross. The only requirement is that there be sufficient feedback for the complete amplifier to be "almost linear" (see Section 1.2). Finally observe that, although the inclusion of local feedback loops may change the detail of a calculation, it does not change the principle.

In the usual case, where both distortion and signal are in the midband frequency range of the complete amplifier (which is wider than the midband range of the forward path), the gain cancels in Eq. (8) and

$$\frac{\text{distortion at output}}{\text{signal at output}} \underset{\text{mid-band}}{\Rightarrow} -\frac{c(s)/G_1(s)}{r(s)}. \tag{9}$$

Eq. (9) is a somewhat surprising result, as the feedback factor and the loop gain have dropped out entirely. The antidistortion $c(s)$ depends only on the nonlinearity and the signal amplitude at it [that is, on the output signal divided by $G_2(s)$]. Therefore the distortion referred to the input depends on the nonlinearity, the output signal amplitude, and the forward-path gains $G_1(s)$ and $G_2(s)$. $H(s)$ is not involved. Distortion as referred to the input of an amplifier is invariant with feedback.

The explanation is similar to that for offset or noise referred to the input, both of which are invariant with feedback. When feedback is applied to an amplifier, its overall gain is reduced. Therefore a larger input must be applied to produce the same output. The improvement in the distortion-to-signal ratio comes about not from a change in the distortion, but from a necessary change to the input.

The dependence of the distortion-to-signal ratio on $G_1(s)$ in Eq. (9) is obvious; the dependence on $G_2(s)$ is more subtle. If $G_2(s)$ is changed at constant input $r(s)$ and constant $H(s)$, the signals at the nonlinearity change and $c(s)$ must be reevaluated.

1.2 An Improved Approximation

This section addresses the fact that when the loop gain is finite, the output from the amplifier is not completely undistorted. Signals at the nonlinearity are therefore not exactly as postulated.

Fig. 5 can be rearranged as Fig. 8 by moving the

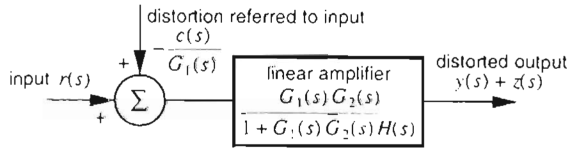

Fig. 7. Amplifier with distortion referred to its input.

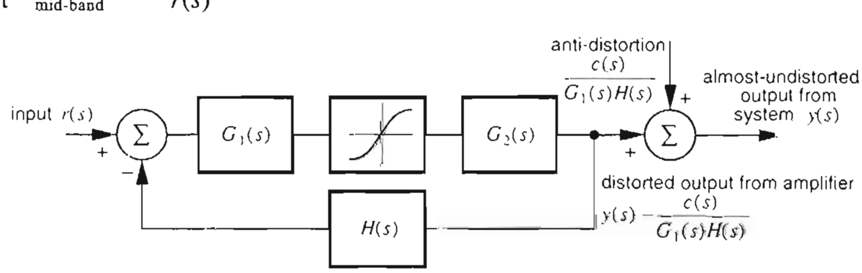

Fig. 6. Approximate equivalent of Fig. 4, based on the fact that amplifier with feedback is almost linear. Distortion values shown are based on large loop-gain approximation $A(s) \approx 1/H(s)$.

antidistortion to the output side of the feedback factor. Fig. 9 is a further rearrangement, in which the antidistortion is subtracted and then added again to achieve the same end result; note the signs. Both Figs. 8 and 9 are exact.

Somewhat more precisely, the effect outside the loop of the antidistortion injected inside the loop can be "guesstimated" using linear feedback theory and Eq. (2). The resulting term is only a correction to the total antidistortion, so any error in its value is a correction within a correction:

$$\begin{pmatrix}\text{equivalent antidistortion}\\ \text{outside the feedback loop}\end{pmatrix} \underset{\text{"guesstimate"}}{=} \frac{\text{antidistortion inside loop}}{1 + \text{loop gain}}$$

$$= -\frac{c(s)}{G_1(s)H(s)}\left[\frac{1}{1 + G_1(s)G_2(s)H(s)}\right]. \qquad (10)$$

Adding this to the antidistortion that is injected outside the loop,

$$\begin{pmatrix}\text{total antidistortion}\\ \text{at system output}\end{pmatrix} = \frac{c(s)}{G_1(s)H(s)}\left[1 \underset{\text{"guesstimate"}}{-} \frac{1}{1 + G_1(s)G_2(s)H(s)}\right]$$

$$= c(s)\left[\frac{G_2(s)}{1 + G_1(s)G_2(s)H(s)}\right]. \qquad (11)$$

The distortion at the output of the amplifier under real operating conditions can be found by removing the antidistortion from Fig. 9. The component of antidistortion injected outside the feedback loop corresponds very nearly to the distortion predicted by Eq. (6)—exactly so if the loop gain is large. The component inside the feedback loop is more difficult because the loop contains the nonlinearity. Certainly, removal of this component will affect the signals at the nonlinearity. However, if the feedback is doing its intended job, the contribution of this component to the system output should be very much attenuated, and can perhaps by neglected. If so, the amplifier distortion reduces to that predicted by Eq. (9).

By subtracting this antidistortion from the undistorted system output, the amplifier distortion under real operating conditions must be

$$z(s) \underset{\text{"guesstimate"}}{=} -c(s)\left[\frac{G_2(s)}{1 + G_1(s)G_2(s)H(s)}\right]. \qquad (12)$$

Eq. (12) differs from Eq. (4) of the simplistic approach in two important regards:

1) $c(s)$ in Eq. (12) is obtained from the reversion of the nonlinearity which would have produced $d(s)$ in Eq. (4).

2) Eq. (4) is thoroughly suspect. In contrast, the main term in the first form of Eq. (11) is exact; only the correcting term is approximate.

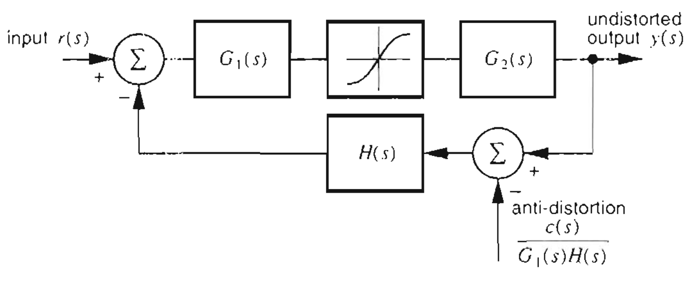

Fig. 8. Feedback amplifier with antidistortion moved to output side of feedback factor.

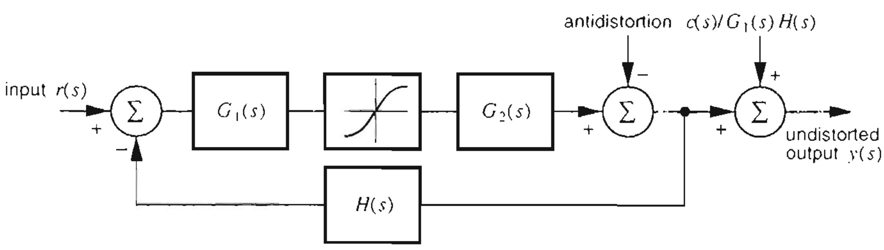

Fig. 9. Feedback amplifier with antidistortion inserted at two points.

1.3 Experimental Comparison

In an amplifier based on Fig. 3,

$$G_1(s) = \frac{3.16\text{M}}{s}$$

$$G_2(s) = \frac{316\text{k}}{s}$$

$$H(s) = 0.1\left(1 + \frac{s}{224\text{k}}\right)$$

$$u = x - 0.014x^3 \underset{\text{reversion}}{\Rightarrow} x \approx u + 0.018u^3 + \cdots .$$

In combination these give a small-signal midband gain of 10, with a 50-kHz maximally flat bandwidth. The cubic nonlinearity corresponds to a smooth saturation at ±5 V peak input; the given reversion matches within ±1% up to ±3 V peak input.

The amplifier was fed with two sine waves,

0.35V peak at 12 kHz
0.25V peak at 17 kHz

and the third-order intermodulation products at the output were measured at 7 kHz, 22 kHz, 41 kHz, and 46 kHz. Table 1 records the gains of the various blocks at the signal frequencies. Table 2 records the measured signal amplitudes at the input and output, and the calculated amplitudes at other points.

In Table 2 the signal amplitude $u(\mathrm{j}\omega)$ at the input of the G_2 block is calculated on the basis of undistorted output and the known $G_2(\mathrm{j}\omega)$. Then the antidistorted input $x(\mathrm{j}\omega)$ to the nonlinearity is calculated from the reverted nonlinearity (see Section A.1) using the signal amplitudes just found. The third-order intermodulation products in $x(\mathrm{j}\omega)$ [corresponding to $c(\mathrm{j}\omega)$] are divided by $G_1(\mathrm{j}\omega)$ to give the antidistortion as referred to the input, and division by $H(\mathrm{j}\omega)$ predicts the output distortion in accordance with Eq. (6).

At 7 kHz the loop gain is large and the measured distortion agrees with predictions to the limit of experimental precision. At 22 kHz, and more especially at 41 and 45 kHz, the loop gain is quite small as these frequencies approach the overall cutoff. There is significant distortion at the output of the nonlinearity, and significant differences exist between the measured output distortion and that predicted by Eq. (6). However, the distortion predicted by the improved Eq. (12) is almost exact.

2 VOLTAGE AND CURRENT

Section 1 is in terms of blocks. The signals in real systems are of two kinds: across and through variables, corresponding to voltage and current in electrical instances. Therefore distortion must be characterized by two generators at the input, and

$$\text{distortion voltage at input} = -\begin{pmatrix}\text{antidistortion voltage which, in conjunction with a signal voltage}\\ \text{source, would reduce the output distortion to zero}\end{pmatrix} \quad (13a)$$

$$\text{distortion current at input} = -\begin{pmatrix}\text{antidistortion current which, in conjunction with a signal current}\\ \text{source, would reduce the output distortion to zero}\end{pmatrix} . \quad (13b)$$

Fig. 10 shows the nonlinear amplifier model with Théve-nin and Norton equivalent signal sources.

Table 1. Magnitude of gain versus frequency for Fig. 3.

Frequency (kHz)	$\lvert G_1(\mathrm{j}\omega)\rvert$	$\lvert G_2(\mathrm{j}\omega)\rvert$	$\lvert H(\mathrm{j}\omega)\rvert$	$\lvert 1 + G_1G_2H\rvert$
12	41.9	4.19	0.106	17.6
17	29.6	2.96	0.111	8.8
7	71.9	7.19	0.102	51.7
22	22.9	2.29	0.118	5.3
41	12.3	1.23	0.153	1.8
46	10.9	1.09	0.163	1.6

Table 2. Predicted signal amplitudes (volts peak) in Fig. 3.

Frequency (kHz)	Measured Input	Measured Output	$u(\mathrm{j}\omega)$	$x(\mathrm{j}\omega)$ [Eq. (53)]	Distortion at Input [Eq. (7)]	Distortion at Output [Eq. (6)]	Improved [Eq. (12)]
12	0.35	3.5	0.83	0.84	—	—	—
17	0.25	2.5	0.84	0.84	—	—	—
7	—	0.0010	0	0.0079	0.00011	0.0011	0.0011
22	—	0.0035	0	0.0080	0.00035	0.0030	0.0035
41	—	0.0053	0	0.0079	0.00064	0.0042	0.0054
46	—	0.0055	0	0.0080	0.00073	0.0045	0.0056

The distortion-to-signal ratios in Fig. 10 are

$$\frac{\text{distortion output}}{\text{signal output}} = \frac{v_{dist} + i_{dist}Z_S}{v_{sig}} \left[\frac{\text{gain } v_o/v_{sig} \text{ (or } i_o/v_{sig}) \text{ at distortion frequency}}{\text{gain } v_o/v_{sig} \text{ (or } i_o/v_{sig}) \text{ at signal frequency}}\right]$$

$$\underset{\text{midband}}{\Rightarrow} \frac{v_{dist} + i_{dist}Z_S}{v_{sig}} \tag{14a}$$

$$\frac{\text{distortion output}}{\text{signal output}} = \frac{i_{dist} + v_{dist}Y_S}{i_{sig}} \left[\frac{\text{gain } i_o/i_{sig} \text{ (or } v_o/i_{sig}) \text{ at distortion frequency}}{\text{gain } i_o/i_{sig} \text{ (or } v_o/i_{sig}) \text{ at signal frequency}}\right]$$

$$\underset{\text{midband}}{\Rightarrow} \frac{i_{dist} + v_{dist}Y_S}{i_{sig}} \tag{14b}$$

As with the offset-to-signal ratio and the noise-to-signal ratio, gain, input impedance, and the other parameters of the amplifier usually cancel out. The distortion voltage dominates when the source impedance is small; the distortion current dominates when the source impedance is large. Notice that the "gain" in Eqs. (14) is defined in terms of the Thévenin or Norton equivalent generator within the source, not the signal at the input of the amplifier. The same consideration applies to offset-to-signal ratio and noise-to-signal ratio.

2.1 Interpretation of "Gain"

When performing the equivalent of dividing the antidistortion $c(s)$ by the gain $G_1(s)$, four pairs of rather similar quantities must be clearly differentiated.

The small-signal voltage gain A_V of an amplifier (or any stage within an amplifier) is the ratio of the small-signal component v_o of the output voltage to the small-signal component v_i of the input voltage under actual operating conditions. Alternatively, A_V is the total derivative of the instantaneous output voltage V_o with respect to the instantaneous input voltage V_i,

$$A_V \equiv \frac{v_o}{v_i} = \frac{dV_o}{dV_i} . \tag{15}$$

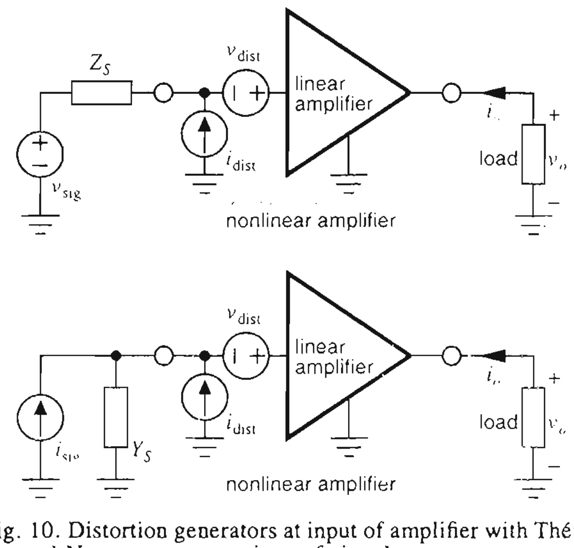

Fig. 10. Distortion generators at input of amplifier with Thévenin and Norton representations of signal source.

In contrast, the voltage amplification factor μ is the ratio of v_o to v_i when the signal output current i_o is zero (in other words, with the load open-circuited). Alternatively, μ is the partial derivative of V_o with respect to V_i when I_o is held constant,

$$\mu \equiv \left.\frac{v_o}{v_i}\right|_{i_o=0} = \left(\frac{\partial V_o}{\partial V_i}\right)_{I_o} . \tag{16}$$

Similarly, the small-signal current gain A_I of an amplifier is the ratio of i_o to i_i under actual operating conditions or, alternatively, is the total derivative of I_o with respect to I_i,

$$A_I \equiv \frac{i_o}{i_i} = \frac{dI_o}{dI_i} . \tag{17}$$

The current amplification factor β is the ratio of i_o to i_i when the signal output voltage v_o is zero (the load is short-circuited) or, alternatively, the partial derivative of I_o with respect to I_i when V_o is held constant,

$$\beta \equiv \left.\frac{i_o}{i_i}\right|_{v_o=0} = \left(\frac{\partial I_o}{\partial I_i}\right)_{V_o} . \tag{18}$$

Mutual conductance g_m is defined strictly as the ratio of i_o to v_i under short-circuit-load conditions, or as the partial derivative of I_o with respect to V_i,

$$g_m \equiv \left.\frac{i_o}{v_i}\right|_{v_o=0} = \left(\frac{\partial I_o}{\partial V_i}\right)_{V_o} . \tag{19}$$

To the author's knowledge there is no universally agreed nomenclature for the corresponding normal-operating-conditions total-derivative quantity. The author prefers transfer conductance G_T (and deprecates the use of transconductance because this is often used synonymously with mutual conductance), and

$$G_T \equiv \frac{i_o}{v_i} = \frac{dI_o}{dV_i} . \tag{20}$$

Finally, mutual resistance r_m is the ratio of v_o to i_i under open-circuit load conditions, or the partial derivative of V_o with respect to I_i,

$$r_m \equiv \left.\frac{v_o}{i_i}\right|_{i_o=0} = \left(\frac{\partial V_o}{\partial I_i}\right)_{I_o} . \tag{21}$$

Transfer resistance R_T is the corresponding normal-operating-conditions total-derivative quantity,

$$R_T \equiv \frac{v_o}{i_i} = \frac{dV_o}{dI_i} . \tag{22}$$

More generally, the midband quantities μ, A_V, β, A_I, g_m, G_T, r_m, and R_T are replaced by functions of complex frequency: $\mu(s)$, $A_V(s)$, $\beta(s)$, $A_I(s)$, $y_m(s)$, $Y_T(s)$, $z_m(s)$, and $Z_T(s)$.

2.2 Types of Antidistortion

The type of antidistortion that must be injected at the input to a nonlinearity, in order to reduce the output distortion to zero, depends on the circumstances. In resistive circuits it may be a voltage or it may be a current. Where nonlinear capacitors or inductors are involved, it may be charge (as in Section 3.4) or flux.

For example, g_m of a FET falls at small currents, and (in the case of an n-channel type) the output waveform is "compressed" near the negative-going peak of the gate voltage. Therefore the required antidistortion is momentarily a negative voltage added in series with the gate. Similarly, β for a BJT falls at large currents, and (in the case of an npn type) the output is compressed near the positive-going peak of the base current. Therefore the required antidistortion is momentarily a positive current added in shunt with the base.

In order to express a voltage antidistortion in terms of a voltage at the input, the antidistortion must be divided by the composite voltage amplification factor of the preceding stages, not their composite voltage gain,

$$\text{distortion voltage at input} = -\frac{\text{antidistortion voltage at input to nonlinearity}}{\text{voltage amplification factor } \mu(s) \text{ of all preceding stages}} . \tag{23}$$

This is because the antidistortion voltage is in a mesh that includes both the input impedance of the following stage and also the output impedance of the preceding stage. The antidistortion adds not to the actual input voltage to the nonlinearity, but to the Thévenin equivalent voltage within the preceding stage. Exactly the same consideration applies to calculating the contribution to offset or noise from any stage other than the first in an amplifier: the offset or noise voltage at the input of that stage should be divided by the voltage amplification factor in order to refer it to the input of the complete amplifier. In the cases of offset and noise, the error introduced through dividing by voltage gain is often negligible, because offset and noise from the first stage usually dominate in the total. For nonlinearity, however, it is usually the last stage that dominates, so it is important that the calculation be done correctly.

For other situations the corresponding results are

$$\text{distortion current at input} = -\frac{\text{antidistortion voltage at input to nonlinearity}}{\text{mutual impedance } z_m(s) \text{ of all preceding stages}} \tag{24}$$

$$\text{distortion voltage at input} = -\frac{\text{antidistortion current at input to nonlinearity}}{\text{mutual admittance } y_m(s) \text{ of all preceding stages}} \tag{25}$$

$$\text{distortion current at input} = -\frac{\text{antidistortion current at input to nonlinearity}}{\text{current amplification factor } \beta(s) \text{ of all preceding stages}} . \tag{26}$$

3 AUDIO POWER AMPLIFIER

This section applies the concept of distortion generators at the input, to estimate the harmonic distortion associated with five nonlinearities in the audio amplifier shown in Fig. 11: first-stage g_m, third-stage β and g_m, second-stage collector capacitance, and first-stage common mode. The first nonlinearity is well known; an early analysis was by the late Peter Baxandall [4]. The second through fourth are known although the details are not widely recognized—especially the fourth. Voltage distortion dominates in all of these. The fifth nonlinearity has hitherto been problematical, but can be explained in terms of current distortion. The sensitivities of this circuit to changes in transistor parameters have been published [5] and make for interesting comparison.

In the analysis, transistors are represented by the equivalent circuit shown in Fig. 12. For the class B output stage, the transistors are usually Darlingtons, and numerical values are suitably defined combinations of the *npn* and *pnp* halves. Relevant transfer functions are the following.

Stage 1,

$$\frac{i_{o1}}{v_{i1(dif)}} = \left.\frac{i_{o1}}{v_{i1(dif)}}\right|_{v_{o1}=0} = \frac{2}{1/g_{m(left)} + 1/g_{m(right)} + 2R} \underset{symmetry}{\Rightarrow} \frac{1}{1/g_{m1} + R} \tag{27}$$

$$\frac{i_{o1}}{i_{i1}} = \left.\frac{i_{o1}}{i_{i1}}\right|_{v_{o1}=0} = 2\beta_1 \tag{28}$$

Stage 2,

$$\frac{v_{o2}}{i_{i2}} = -\frac{\beta_2 R_{i3}}{1 + s\beta_2 R_{i3} C} \approx -\frac{\beta_2\beta_3 R_L}{1 + s\beta_2\beta_3 R_L C} \tag{29}$$

$$\left.\frac{v_{o2}}{i_{i2}}\right|_{i_{o2}=0} = -\frac{1}{sC} \tag{30}$$

$$\left.\frac{i_{o2}}{i_{i2}}\right|_{v_{o2}=0} = \beta_2 \tag{31}$$

Stage 3,

$$\frac{v_{o3}}{v_{i3}} = \frac{g_{m3}R_L}{1 + g_{m3}R_L} \approx 1 \tag{32}$$

$$\frac{v_{i3}}{i_{i3}} = \beta_3(1/g_{m3} + R_L) \approx \beta_3 R_L \tag{33}$$

Overall voltage gain and 3-dB bandwidth,

$$\frac{v_o}{v_i} \approx \frac{R_{F1} + R_{F2}}{R_{F1}} \tag{34}$$

$$\omega_{3dB} \approx \frac{1}{(1/g_{m1} + R)C}\left(\frac{R_{F1}}{R_{F1} + R_{F2}}\right). \tag{35}$$

In Eq. (30) the mutual impedance of the second stage (into an open-circuit load) approaches ideal integration. In Eq. (31) the current amplification factor is independent of the lag compensating capacitor; short-circuiting the load disables the local feedback provided by C. Better equivalent circuits and approximations could of course be used, but one purpose of this section is to demonstrate that distortion can be estimated with good accuracy and little effort. Certainly, there is nothing approaching the effort of a formal sensitivity analysis.

Calculated distortions are compared with values obtained from a "variable-distortion" amplifier [6], used in the electronics teaching laboratory at Monash University. Parameters such as transistor β can be set to selected values by front-panel switches, with a tolerance around ±5%. Nonlinearities can also be set, with somewhat looser tolerances. The following data are used except where stated.

Stage-1 parameters,

$$I_1 = 1\ \text{mA}$$

$$g_{m1} = 40\ \text{mA/V}$$

$$\beta_1 = 500$$

$$R = 220\ \Omega$$

$$R_S = 600\ \Omega$$

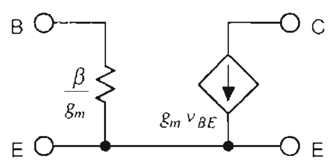

Fig. 12. Assumed small-signal equivalent circuit of BJT.

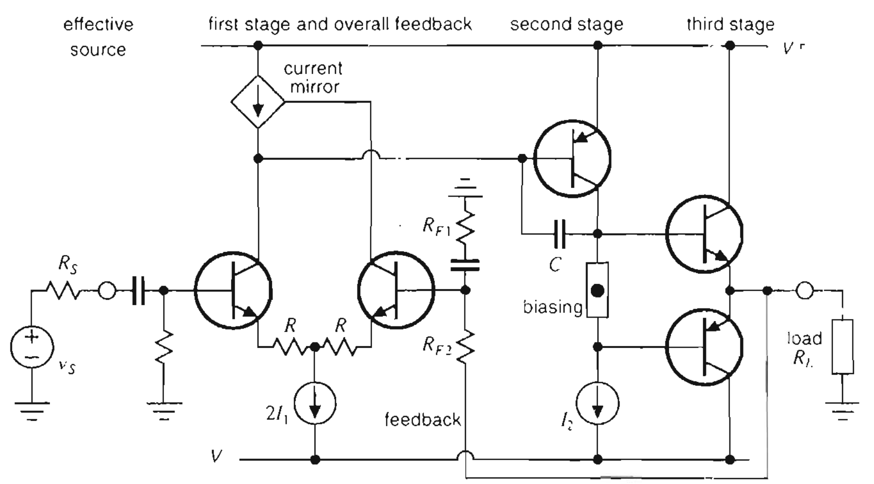

Fig. 11. Typical circuit topology for audio power amplifier.

Stage-2 parameters,

$I_2 = 3.5$ mA

$\beta_2 = 50$

$C = 100$ pF

Stage-3 parameters,

$g_{m3} = 650$ mA/V (including emitter degeneration)

$\beta_3 = 1000$

$R_L = 10\ \Omega$

Feedback network,

$R_{F1} = 2.2$ kΩ

$R_{F2} = 47$ kΩ

Signal amplitudes,

$\hat{v}_S = 440$ mV peak

$\hat{v}_o = 10$ V peak

$\hat{i}_o = 1$ A peak .

3.1 Nonlinearity in g_{m1}

In order to correct nonlinearity in g_{m1}, an antidistortion voltage must be added in series with stage-1 input. Because there is no gain ahead of this point in the amplifier, Eq. (23) gives the distortion voltage referred to the input as

$$v_{dist} = -v_{anti} . \tag{36}$$

The distortion current is zero.

For given v_{anti} (that is, for a given nonlinearity and signal amplitude in the first stage) there is no way of reducing the distortion voltage referred to the input. However, the signal amplitude at the nonlinearity (and hence v_{anti} itself) depends on the gain $G_2(s) \Rightarrow v_{o3}/i_{i2}$, which follows the nonlinearity: strongly on the signal frequency and C, less so on β_2, β_3, and R_L, weakly on g_{m3} [Eqs. (29) and (32)]. (Eq. (38) of [5] adds formal support.)

In an experiment, first-stage parameter values are altered to increase distortion (for ease of measurement) without significantly changing the overall gain or bandwidth,

$I_1 = 100\ \mu$A

$R = 0$.

The voltage gain of the third stage is 0.87 [Eq. (32)], and at 10 kHz the transfer impedance of the second stage is 153 V/mA [Eq. (29)]. Therefore, at ±10-V output from the amplifier, the output from the first stage is ±75 μA peak. Taking account of 70-μA quiescent base current in the second stage, the instantaneous currents in the left/right members of the first-stage long-tailed pair are 135/65 μA, 173/27 μA, and 97/103 μA at the quiescent point and signal peaks, respectively.

From Eq. (27) the transfer conductance at the quiescent point is 3.51 mA/V. If the stage were linear with constant transfer conductance, the differential input voltage would be ±21.4 mV peak, undistorted.

The required antidistortion of this input is found by reverting the transfer conductance and using the method of differential errors (see Section A.3),

$$\frac{dV_{i1(dif)}}{dI_{o1}} = \frac{v_{i1(dif)}}{i_{o1}} = \frac{1/g_{m(left)} + 1/g_{m(right)} + 2R}{2} . \tag{37}$$

When the instantaneous currents are substituted into Eq. (37), the gradient of a graph of $V_{i(dif)}$ (vertical axis) versus I_{o1} (horizontal) is 285, 250, and 529 V/A at the quiescent point and signal peaks, and the differential error at the peaks is −0.123 and +0.856 [Eq. (56)]. Hence from Eqs. (59) the required antidistortion is

$D_2 = -12.3\% \Rightarrow -2.61$ mV peak in 21.4 mV fundamental

$D_3 = +3.1\% \Rightarrow 0.65$ mV peak in 21.4 mV fundamental

and from Eq. (36) the distortion voltages referred to the input are the negatives of these.

The input to the complete amplifier is ±440 mV peak. From Eq. (14a), the predicted second- and third-harmonic-distortion-to-signal ratios are 0.59 and 0.15%, respectively. The measured distortions at 10 kHz are 0.58 and 0.15%.

3.2 Nonlinearity in β_3

In order to correct the nonlinearity in β_3, an antidistortion current must be added in shunt with stage-3 input. From Eqs. (25) and (26) the distortion voltage and current referred to the input are

$$v_{dist} = -i_{anti}\left(\frac{1/g_{m1} + R}{\beta_2}\right) \tag{38}$$

$$i_{dist} = -i_{anti}\left(\frac{1}{2\beta_1\beta_2}\right) . \tag{39}$$

Observe that distortion referred to the input is independent of frequency. Distortion does not increase with increasing frequency, even though overall loop gain falls away. For given i_{anti}, distortion can be reduced (without changing overall gain and/or bandwidth) by increasing β_2. Or, i_{anti} itself can be reduced by increasing β_3. (Eq. (42) of [5] adds formal support.)

In an experiment, the amplifier output is ±1 A peak. Small-signal β_3 is selected as 1000, dropping to 500 and 333 at the positive and negative peaks, respectively. If the transistors were linear with constant $\beta_3 = 1000$, stage-3 input current would be ± 1 mA peak, undistorted.

The required antidistortion of this input is found by reverting the definition of transistor small-signal β,

$$\frac{dI_B}{dI_C} = \frac{i_B}{i_C} = \frac{1}{\beta}$$

For the selected nonlinearity, the gradient of a graph of I_B (vertical axis) versus I_C (horizontal) is 0.001, 0.002, and 0.003 at the quiescent point and peaks, respectively. When these are substituted into Eq. (56), the differential error at the peaks is +1.0 and +2.0 and, from Eqs. (59), the required antidistortion current is

$$D_2 = -12.5\% \Rightarrow -125\ \mu\text{A peak in 1 mA fundamental}$$

$$D_3 = +12.5\% \Rightarrow 125\ \mu\text{A peak in 1 mA fundamental}\,.$$

From Eqs. (38) and (39), the second- and third-harmonic-distortion voltages referred to the input are both 0.61 mV, the currents are 2.5 nA.

The input to the complete amplifier is ± 440 mV peak. From Eq. (14a) the predicted second- and third-harmonic distortion-to-signal ratios are both 0.14%. The contribution of i_{dist} is negligible because R_S is quite small. Measured distortions at 1 kHz are 0.15 and 0.15%. Measured distortions at 10 kHz are 0.23 and 0.22%, significantly different from prediction, but the hardware for the β nonlinearity is known to be imperfect at high frequencies.

3.3 Nonlinearity in g_{m3}

In order to correct a nonlinearity in g_{m3}, an antidistortion voltage must be added in series with the stage-3 input. From Eqs. (23) and (24) the distortion voltage and current referred to the input are

$$v_{dist} = -v_{anti}\left(\frac{1}{g_{m1}} + R\right)sC \tag{41}$$

$$i_{dist} = -v_{anti}\left[\frac{sC}{2\beta_1}\right]. \tag{42}$$

Observe that distortion referred to the input involves the derivative of antidistortion. It continues to decrease with decreasing frequency below the forward-path pole, where the overall loop gain becomes constant. For a given amplitude and frequency of v_{anti}, there is no way of reducing distortion without changing overall gain and/or bandwidth. (Eq. (41) of [5] adds formal support.)

In an experiment the class B output stage is underbiased to give 500 mV dead band, and v_S is reduced to 44 mV peak; β_2 is increased to 5000 to minimize other distortions. The required antidistortion is approximately a square wave, 500 mV peak to peak, which contains *inter alia* 106 mV, 63.7 mV, and 45.4 mV peak at the third, fifth, and seventh harmonics. For 1 kHz fundamental, all harmonic voltages referred to the input are 50 μV [Eq. (41)], or 0.12% of the reduced input. Once again the contribution of distortion current is negligible. The measured distortions are 0.14, 0.12, and 0.12%.

3.4 Nonlinearity in Second-Stage Collector Capacitance

In order to correct nonlinearity in a second-stage collector–base capacitance, an antidistortion charge must be added in shunt with the stage-2 input. (Strictly, an equal component must also be added at the stage-2 output, but this is divided by β_2 when referred back to the stage input.) Therefore, since charge is the integral of current,

$$v_{dist} = -q_{anti}\, s\left(\frac{1}{g_{m1}} + R\right) \tag{43}$$

$$i_{dist} = -q_{anti}\left(\frac{s}{2\beta_1}\right). \tag{44}$$

Observe that distortion referred to the input involves the derivative of the antidistortion charge, and therefore increases proportional to frequency. For a given amplitude and frequency of q_{anti}, the distortion (at constant overall gain and bandwidth) can be reduced by reducing R, but C must simultaneously be increased. (Eq. (43) of [5] adds formal support.)

In an experiment the output from the amplifier is ± 10 V peak and the voltage gain of the third stage is 0.87, so the output from the second stage is ± 11.5 V peak; β_2 is increased to 5000 to minimize other distortions. The second-stage collector–base capacitance is selected as 20 pF at the quiescent point, 32 pF and 17 pF at the positive and negative peaks. If the capacitance were constant at 20 pF, the input charge would be ± 231 pC peak, undistorted.

The required antidistortion of this input is found from the definition of small-signal capacitance,

$$c_C = \frac{dQ_B}{dV_C}\,. \tag{45}$$

Thus, the gradient of a graph of Q_B (vertical) versus V_C (horizontal) is 20, 32, and 17 pC/V at the quiescent point and peaks, respectively. When these are substituted into Eq. (56), the differential error at the peaks is +0.60 and −0.15, and, from Eqs. (59), the antidistortion is

$$D_2 = +9.4\% \Rightarrow 21.6\text{ pC peak in 231 pC fundamental}$$
$$D_3 = +1.9\% \Rightarrow 4.3\text{ pC peak in 231 pC fundamental}\,.$$

From Eqs. (43) and (44), the second- and third-harmonic distortions at 10 kHz fundamental, referred to the input, are 0.67 and 0.20 mV, or 2.7 and 0.8 nA.

With 440-mV peak input to the amplifier, the predicted second- and third-harmonic-distortion-to-signal ratios are 0.15 and 0.05% [Eq. (14a)]; the contribution of i_{dist} is negligible. Measured distortions at 10 kHz are 0.16 and 0.07%.

3.5 First-Stage Common-Mode Distortion

In amplifiers of the topology of Fig. 11 there is a relatively large common-mode voltage between first-stage collector and emitter, equal to the actual input voltage. The collectors are more or less at signal ground, but the full input voltage appears at the emitters. In contrast, the differential-mode voltage between the bases is the actual input divided by the overall loop gain. Common-mode nonlinearities may be excited, and some form of common-mode distortion can occur.

One mechanism (there are others) involves the nonlinear collector–base capacitance. An antidistortion charge is required at both transistor bases, corresponding to distortion currents at both the input and the feedback point. The distortion-to-signal ratio becomes

$$\frac{\text{distortion at output}}{\text{signal at output}} = \frac{v_{\text{dist}} + i^{+}_{\text{dist}} R_{\text{S}} - i^{-}_{\text{dist}}(R_{\text{F1}} \| R_{\text{F2}})}{v_{\text{S}}} . \tag{46}$$

Amplifiers are known for which distortion goes through a minimum as the source resistance is varied. Eq. (46) provides the explanation: common-mode distortion associated with collector–base capacitance should be null when $R_{\text{S}} = R_{\text{F1}} \| R_{\text{F2}}$.

In the variable-distortion amplifier, the first-stage common-mode voltage is ±440 mV peak at near-full output; the collector–base capacitance is 2.0 pF at the quiescent point, 2.05 and 1.95 pF at the positive and negative peaks. If this capacitance were constant at 2.0 pF, the input charge at either transistor base would be ±880 fC peak, undistorted. The differential error at the peaks of $dQ_{\text{B}}/dV_{\text{C}}$ is +0.025 and −0.025 and, from Eq. (59), the antidistortion is

$$D_2 = +0.625\% \Rightarrow 5.5 \text{ fC peak in } 880 \text{ fC fundamental}$$

$$D_3 = 0 .$$

For a 15.9-kHz input, the second-harmonic distortion current (at 200 k.rad/s) at either base is 1.1 nA peak, and, from Eq. (46), the distortion should be 0.00025% for every 1 kΩ of unbalance between R_{S} and $R_{\text{F1}} \| R_{\text{F2}}$. The distortion voltage is zero.

In the case of the variable-distortion amplifier, the common-mode distortion is unmeasurably small. However, the common-mode voltage can be large in amplifiers for which the overall gain is small, and common-mode distortion can become significant.

4 CONCLUSIONS

Nonlinear distortion in a feedback amplifier can be modeled by two generators at the input, one voltage and one current, after the style of the generators that model dc offsets and noise. The values of these generators can be calculated. They are invariant with feedback. The method applies no matter how gross the forward-path nonlinearity may be, provided only that the amplifier with feedback is "almost linear."

The author is not aware of any other method that can predict distortion in a feedback amplifier with anything approaching the precision exemplified in Section 1.3 and 3, or with so little computational effort.

5 ACKNOWLEDGMENT

The anonymous reviewers of this paper made a number of helpful suggestions, which the author gratefully acknowledges.

6 REFERENCES

[1] 'Cathode Ray' (M. G. Scroggie), "Negative Feedback and Nonlinearity," *Wireless World*, vol. 83, pp. 47–50 (1978 Oct.).

[2] P. J. Baxandall, "Audio Power Amplifier Design: Part 6," *Wireless World*, vol. 84, pp. 69–73 (1979 Feb.).

[3] E. M. Cherry and K. P. Dabke, "Transient Intermodulation Distortion—Part 2: Soft Nonlinearity," *J. Audio Eng. Soc.*, vol. 34, pp. 19–35 (1986 Jan./Feb.).

[4] P. J. Baxandall, "Audio Power Amplifier Design: Part 1," *Wireless World*, vol. 83, pp. 53–57 (1978 Jan.).

[5] E. M. Cherry, "Feedback, Sensitivity, and Stability of Audio Power Amplifiers," *J. Audio Eng. Soc.*, vol. 30, pp. 282–294 (1982 May); also, "Further Thoughts an 'Feedback, Sensitivity, and Stability of Audio Power Amplifiers,'" *ibid. (Communications)*, vol. 31, pp. 854–856 (1983 Nov.).

[6] E. M. Cherry, "A 'Variable-Distortion' Audio Power Amplifier," in *Inst. Radio Elec. Eng. Australia Convention Dig. IREECON81* (1981 Aug.), pp. 346–347.

[7] B. M. Oliver, "Distortion and Intermodulation," Hewlett-Packard Application Note 15 (1963 Dec.).

[8] E. M. Cherry and D. E. Hooper, *Amplifying Devices and Low-Pass Amplifier Design* (Wiley, New York, 1968).

APPENDIX
MATHEMATICS OF NONLINEARITY

A.0 Introduction

A time-invariant nonlinear transfer characteristic [such as Fig. 2(b)] can be expressed as a power series,

$$\underset{\text{all } t}{u =} A_0[x + n_2x^2 + n_3x^3 + n_4x^4 + \cdots] \tag{47}$$

from which

$$\frac{du}{dx} = A_0[1 + 2n_2x + 3n_3x^2 + 4n_4x^3 + \cdots] \tag{48}$$

and, at the origin,

$$\left.\frac{du}{dx}\right|_{x=0} = A_0 . \tag{49}$$

In these equations, the gradient du/dx at any point is the *incremental gain*, the ratio of a small change in input to the resulting small change in output; and the gradient at the origin is the *small-signal gain* A_0.

For the purposes of this appendix, a *soft nonlinearity* is defined as one that can be approximated "adequately" by truncating its power-series expansion after the cubic term,

$$\underset{\substack{\text{soft}\\\text{nonlinearity}}}{u} \equiv A_0[x + n_2x^2 + n_3x^3] . \tag{50}$$

A.1 Soft Nonlinearity and Distortion

When an input consisting of two sinusoids

$$x(t) = B_1 \cos \omega_1 t + B_2 \cos \omega_2 t \tag{51}$$

is applied to the unity small-signal gain soft nonlinearity

$$u = x + n_2x^2 + n_3x^3 \tag{52}$$

the resulting output is [7]

$$\begin{aligned}
u(t) = {} & \frac{1}{2}n_2B_1^2 + \frac{1}{2}n_2B_2^2 \qquad \text{dc shift} \\
& + \left(B_1 + \frac{3}{2}n_3B_1B_2^2 + \frac{3}{4}n_3B_1^3\right)\cos \omega_1 t + \left(B_2 + \frac{3}{2}n_3B_2B_1^2 + \frac{3}{4}n_3B_2^3\right)\cos \omega_2 t \qquad \text{fundamentals} \\
& + \frac{1}{2}n_2B_1^2 \cos 2\omega_1 t + \frac{1}{2}n_2B_2^2 \cos 2\omega_2 t \qquad \text{second harmonics} \\
& + \frac{1}{4}n_3B_1^3 \cos 3\omega_1 t + \frac{1}{4}n_3B_2^3 \cos 3\omega_2 t \qquad \text{third harmonics} \\
& + n_2B_1B_2\,[\cos(\omega_1 + \omega_2)t + \cos(\omega_1 - \omega_2)t] \qquad \text{second-order intermodulations} \\
& + \frac{3}{4}n_3B_1^2B_2\,[\cos(2\omega_1 + \omega_2)t + \cos(2\omega_1 - \omega_2)t] \\
& + \frac{3}{4}n_3B_1B_2^2\,[\cos(2\omega_2 + \omega_1)t + \cos(2\omega_2 - \omega_1)t] \qquad \text{third-order intermodulations} .
\end{aligned} \tag{53}$$

The magnitudes of the various components in Eq. (53) are independent of any phase shift between the two input sinusoids, but the phases are not. Therefore, if the input frequencies ω_1 and ω_2 are such that two components of the distortion happen to be of the same frequency, due care must be exercised in their summation.

In Section 1.3, n_3 for the reverted nonlinearity is 0.18 and the signal amplitudes B_1 and B_2 are read from the fourth column of Table 2 as 0.83 and 0.84. The amplitudes of the resulting third-order intermodulations follow from the last two lines of Eq. (53), but note that the significance of $x(t)$ and $u(t)$ in this equation must be interchanged—in order to find the antidistorted input that would give undistorted output.

A.2 Harder Nonlinearities

When higher exponent terms are included in the nonlinearity, the algebra of the general case verges on unmanageability. However, the trend is illustrated by two cases—the simple quartic and the simple quintic, each with a single-frequency input $x(t) = B \cos \omega t$,

$$\underset{\text{quartic}}{u} = x + n_4x^4 \tag{54a}$$

$$\underset{\text{quintic}}{u} = x + n_5x^5 . \tag{54b}$$

The outputs are

$$\underset{\text{quartic}}{u(t)} = \frac{3}{8}n_4B^4 + B\cos \omega t + \frac{1}{2}n_4B^4 \cos 2\omega t + \frac{1}{8}n_4B^4 \cos 4\omega t \tag{55a}$$

$$\underset{\text{quintic}}{u(t)} = \left(B + \frac{5}{8}n_5B^5\right)\cos \omega t + \frac{5}{16}n_5B^5 \cos 3\omega t + \frac{1}{16}n_5B_5 \cos 5\omega t . \tag{55b}$$

Observe that the quartic nonlinearity generates four times more second harmonic than fourth harmonic, and that the quintic generates five times more third harmonic than fifth harmonic.

Any more complicated cases should be considered on its merits, but the pattern is that high-even-exponent and high-odd-exponent terms in the nonlinearity generate principally second-order and third-order distortion products, respectively. If some approximate analysis gets the seconds and thirds right, the total will not be far wrong.

A.3 Differential Error

The differential error γ at any point on a transfer characteristic is the normalized difference between the incremental gain at that point and the small-signal gain,

$$\gamma = \frac{\dfrac{du}{dx} - \left.\dfrac{du}{dx}\right|_{x=0}}{\left.\dfrac{du}{dx}\right|_{x=0}} = 2n_2x + 3n_3x^2 + 4n_4x^3 + \cdots . \tag{56}$$

Differential error provides a link between small-signal-like calculations, the coefficients in the power-series expansion of a soft nonlinearity, and hence distortion [8, ch. 9].

Incremental gain at any point on a transfer characteristic can be calculated via the ordinary formulas for small-signal gain, but using the numerical values for parameters like transistor g_m and β which pertain at that point. Hence γ at that point can be found.

If the values of γ at the peaks of a symmetrical input signal $\pm\hat{x}$ are written as γ' and γ'', respectively, then, from Eq. (56) with a soft nonlinearity,

$$\gamma' = 2n_2\hat{x} + 3n_3\hat{x}^2 \tag{57a}$$

$$\gamma'' = -2n_2\hat{x} + 3n_3\hat{x}^2 \tag{57b}$$

from which

$$n_2 = \frac{\gamma' - \gamma''}{4\hat{x}} \tag{58a}$$

$$n_3 = \frac{\gamma' + \gamma''}{6\hat{x}^2} . \tag{58b}$$

In the case where the input is a single sinusoid $x(t) = B \cos \omega t$, it follows from Eqs. (53) and (58) that the second- and third-harmonic distortions are

$$D_2 = \frac{\gamma' - \gamma''}{8} \times 100\% \tag{59a}$$

$$D_3 = \frac{\gamma' + \gamma''}{24} \times 100\% . \tag{59b}$$

Although derived here for the case of a mathematically soft nonlinearity, Eqs. (59) work well in more general situations; Section A.2 is a partial justification.

Section 3 uses Eqs. (59) to calculate the antidistortion required at the input of a nonlinear process in order to give undistorted output. Because the nonlinearity is reverted, the significance of $x(t)$ and $u(t)$ in the preceding development must be interchanged.

A.4 Quadratic Nonlinearity with Feedback

The power-series solution of Fig. 2 is

$$y = r + \left(\frac{1}{1+GH}\right)r^2 - \left(\frac{2}{1+GH}\right)\left(\frac{GH}{1+GH}\right)r^3 + \left(\frac{5}{1+GH}\right)\left(\frac{GH}{1+GH}\right)^2 r^4 - \left(\frac{14}{1+GH}\right)\left(\frac{GH}{1+GH}\right)^3 r^5 + \cdots . \tag{60}$$

By differentiation, the coefficient of r^k for $k > 2$ goes through a maximum when

$$GH_{\text{maximum}(r^k)} = k - 2 . \tag{61}$$

Thus, second-harmonic distortion is maximum when there is no feedback, but the offensive seventh harmonic initially increases when feedback is applied, reaches a maximum when $GH = 5$, and then drops away again. The maximum coefficient of r^7 is 8.8, surprisingly large compared with the maximum coefficient of r^2, which is a mere 1.0.

The constants +1, −2, +5, −14, . . . in Eq. (60) are the negatives of the coefficients in the reversion of the quadratic nonlinearity in the forward path. If

$$y \underset{x'= 0.5}{=} x + x^2$$

as assumed in Fig. 2, then

$$x = y - y^2 + 2y^3 - 5y^4 + 14y^5 - 42y^6 + 132y^7 + \cdots . \tag{62}$$

Ultimately it was the recognition of this fact which lead to the discovery of the new method for calculating distortion.

THE AUTHOR

Edward M. Cherry obtained the B.Sc. and M.Sc. degrees in physics and electronics in 1957 and 1959 from the University of Melbourne, Australia, and the Ph.D. degree in electrical engineering in 1962.

From 1963 to 1965 he was a Member of the Technical Staff of the Bell Telephone Laboratories, Murray Hill, NJ, and in 1966 he was a Temporary Research Associate of the U.K. Atomic Energy Authority, Harwell. He then took up an appointment with the Electrical and Computer Systems Engineering Department at Monash University, Australia, where he remained for 32 years. He spent 12-month sabbaticals at Bell Laboratories, New Jersey (with a Fulbright scholarship) and at Tektronix Inc., Oregon, and had a teaching exchange at Luleå University, Sweden. He retired in January 1999.

Professor Cherry's research interests are in electronic hardware. He has published two books, over 40 full-length papers in refereed journals, and a quantity of other articles. He is a Fellow of the Institution of Radio and Electronics Engineers, and has twice been awarded that Institution's annual medal for the most meritorious paper published in its *Proceedings*. He is at present living in the United States, working part-time for Maxim Integrated Products as a consultant on high-speed analog integrated circuits. Email: millend@teleport.com