A Power Amplifier "Improver"*

EDWARD M. CHERRY

Department of Electrical Engineering, Monash University, Clayton, Victoria 3168, Australia

An add-on facility for audio power amplifiers is described, by which nonlinear distortion products in the audible range of frequencies are reduced by a typical 25–30 dB. The improver uses the principle of nested differentiating feedback loops to increase the feedback around the amplifier. The improver also corrects phase nonlinearity in amplifiers, improves loudspeaker damping, increases isolation between stereo channels, and reduces hum, noise, and dc drift.

0 INTRODUCTION

A recent paper [1] describes a new feedback structure, a nest of differentiating feedback loops, by means of which Bode's limiting value of loop gain can be exceeded. Increased feedback around an amplifier results in reduced nonlinear distortion. A second paper [2] describes an audio power amplifier which uses the new structure; power transistor f_T is only 2 MHz, yet loop gain is at least 25 000 (88 dB) over the range 20 Hz to 20 kHz, and distortion products in the same range are less than 20 ppm.

The present paper describes an add-on "improver" for existing audio power amplifiers, again using nested differentiating loops to increase the feedback. Harmonic distortion and intermodulation products between sinusoids in the audible frequency range are reduced by a typical 25–30 dB, down to a lower limit below 10 ppm. Transient intermodulation distortion is reduced by a similar factor, provided the rate of change of the input signal is not so great that the amplifier is driven to its limiting slew rate.

The amplifier improver also reduces linear distortion. System response is flat and phase linear from 20 Hz to 20 kHz, so that step waveforms are reproduced without overshoot and low-frequency square waves without tilt. Loudspeaker damping and crosstalk between channels of a stereo amplifier are also improved by increased feedback. Drift at the output of a dc amplifier is reduced through the use of a feedback factor which increases with decreasing frequency and reaches unity at zero frequency. Noise is reduced for all but the very lowest noise amplifiers; the spectral density referred to the system input is below 20 nV/$\sqrt{\text{Hz}}$ with 2 kHz flicker corner, corresponding to −112 dB relative to 1 V, A-weighted dc to 20 kHz.

1 LOW-PASS THEORY

Readers should refer to Cherry [1] for the complete theory of nested differentiating feedback loops, and to Cherry [2] for a simplified but more readable account.

Fig. 1 shows a block diagram of the improver in a system. One of the nested differentiating feedback factors is the block sC_F, and others are provided within typical amplifiers by lag-compensating networks. The following design constraints are necessary between elements in the improver:

$$G_T R_T = 1 \tag{1}$$

$$G_T \beta \tau_X = C_F. \tag{2}$$

The high-frequency response of the amplifier to be improved is modeled as a two-pole function—surely more realistic than the single-pole model used in so many papers on feedback and distortion in amplifiers. The amplifier midband gain is A_0, the undamped natural frequency and damping ratio of its poles are $1/\tau_0$ and ζ, respectively. Network analysis shows that the voltage

* Manuscript received 1980 April 15; revised December 2.

  0004-7554/81/030140-08$00.75

gain of the system, ignoring the dashed network β, is

$$A(s) = \frac{\text{Output}}{\text{Input}}$$

$$= \frac{1}{\beta}$$

$$\times \left\{ \frac{1}{1 + s(\tau_X + \tau_L) + s^2\tau_X^2 \left[\dfrac{1 + 2\zeta s\tau_0 + s^2\tau_0^2}{A_0\beta(1 + \frac{1}{2}s\tau_X)^2} \right]} \right\} . \tag{3}$$

Ideally the parameters of the amplifier and improver should be related:

$$A_0 = 1/\beta \tag{4}$$

$$\tau_0 = \tfrac{1}{2}\tau_X \tag{5}$$

$$\zeta = 1. \tag{6}$$

If we choose

$$\tau_L = (\sqrt{3} - 1)\tau_X \tag{7}$$

then

$$A(s) = \frac{1}{\beta} \left[\frac{1}{1 + \sqrt{3}s\tau_X + s^2\tau_X^2} \right] \tag{8}$$

This response is a phase-linear function with cutoff frequency

$$\omega_{3dB} = \frac{0.786}{\tau_X} . \tag{9}$$

It can be verified using analysis or simulation that the system response can be made approximately phase linear by adjusting τ_L is any one of A_0, τ_0, or ζ departs from its ideal value in Eqs. (4)–(6) by as much as a factor 2, or if the amplifier has additional nondominant poles.

The sensitivity of $A(s)$ to changes in A_0 gives an approximate measure of the factor by which nonlinear distortion in the amplifier is reduced:

$$S[A_0] = \frac{A_0}{A(s)} \left(\frac{\partial A(s)}{\partial A_0} \right) . \tag{10}$$

If the ideal relations are satisifed,

$$S[A_0] = \frac{s^2\tau_X^2}{1 + \sqrt{3}s\tau_X + s^2\tau_X^2} . \tag{11}$$

Distortion products at all frequencies up to two octaves inside the system cutoff are reduced by at least 28 dB.

The foregoing analytical approach demonstrates beyond question that the "improved" amplifier is stable and is of reduced sensitivity, but gives little insight into how the improver actually works. The whole improver may be considered as a kind of active feedback network around the amplifier to be improved, and it is the feedback around the resulting loop which leads to reduced sensitivity to changes in the amplifier. [The relation between Fig. 3 of [2] and Fig. 1 is as follows: The blocks G_T and $R_T(1 + \frac{1}{2}s\tau_X)^2/(s\tau_X)^2$ serve to increase the gain of the amplifier to be improved, and the combination constitutes the forward path of the complete system. The blocks sC_F and $\beta(1 + s\tau_L)$ provide two nested feedback loops.]

The transfer function of the improver considered as a feedback network is

$$B(s) = \left. \frac{\text{to amplifier input}}{\text{from amplifier output}} \right|_{\text{input} = 0}$$

$$= - \frac{[\beta(1 + s\tau_L)G_T + sC_F]R_T(1 + \frac{1}{2}s\tau_X)^2}{(s\tau_X)^2} \tag{12}$$

and, using Eqs. (1), (2), and (7),

$$B(s) = - \frac{\beta(1 + \sqrt{3}s\tau_X)(1 + \frac{1}{2}s\tau_X)^2}{(s\tau_X)^2} . \tag{13}$$

Fig. 2 shows the asymptotes of $|1/B(j\omega)|$ on log-log scales.

Also shown in Fig. 2(a) are the gain asymptotes of an amplifier having the ideal parameters of Eqs. (4)–(6).

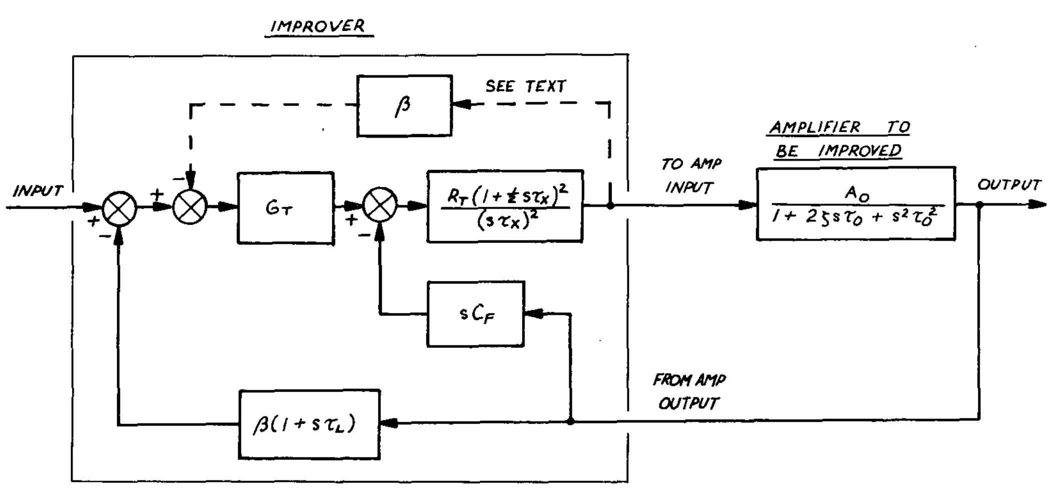


Fig. 1. Block diagram of the improver in a low-pass system.

The gain around the loop is the separation between the curves in Fig. 2(a), and is plotted separately in Fig. 2(b). Bode's rule [3] for stability of a feedback system can be paraphrased that the loop gain should not fall faster than a single-pole rate in the vicinity of the frequency of unity loop gain. The asymptote in Fig. 2(b) satisfies this rule, confirming that the system should be stable. The effects of changes in A_0, τ_0, ζ, and τ_L can easily be assessed from Bode plots of the type of Fig. 2.

Fig. 3 is the outline circuit of an improver designed with

$$1/\beta = 31.6$$

$$\tau_X = 2\ \mu s$$

$$G_T = 1\ mA/V.$$

β is defined by 68 kΩ and 2.2 kΩ; $\frac{1}{2}\tau_X$ is defined by 470 Ω and 2.2 nF in the second stage and 220 Ω and 4.7 nF in the third stage; G_T is defined by 2 × 1.0 kΩ in the first-stage emitters. The topology of the second stage is perhaps unfamiliar, but the circuit is actually a long-tailed pair using one *n-p-n* transistor and one *p-n-p*.

Note the 68 kΩ between improver output and feedback point, which corresponds to the dashed feedback network β in Fig. 1. This network has two purposes:

1) It allows the improver to bias correctly when the amplifier to be improved is not dc coupled, while having negligible effect on $A(s)$.

2) It increases $S[A_0]$ at low frequencies to $1 + 1/\beta$ (that is, it reduces the improvement), but very much increases the improvement at frequencies around $\sqrt{\beta}/\tau_X$, which corresponds to the top of the audible frequency band.

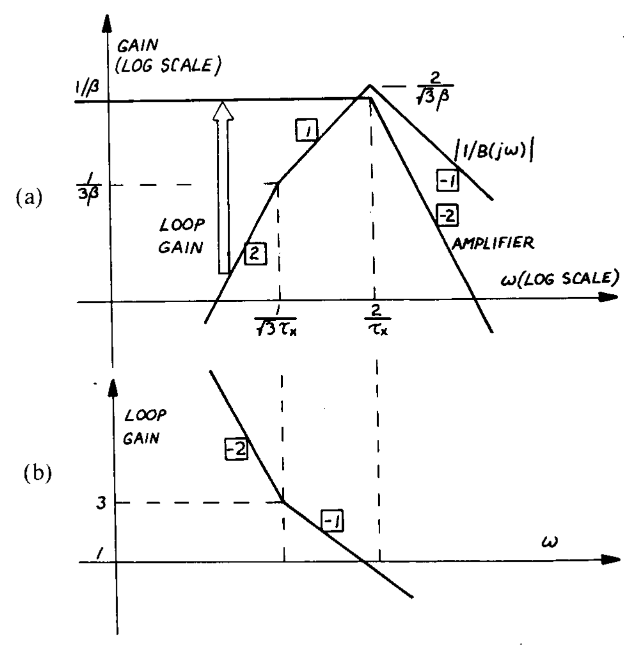


Fig. 2. The improver considered as a feedback network around the amplifier to be improved (low-pass case). (a) Gain asymptotes of the forward path (the amplifier to be improved) and feedback factor (the improver). (b) Asymptotes of loop gain.

2 TRANSIENT INTERMODULATION DISTORTION

In recent years there have been many discussions of the interplay between feedback and transient intermoduation distortion (TIMD). Sadly, some of this material contains misquotations and/or is wrong [4]–[9].

The necessary condition [10] for avoiding gross TIMD in a feedback amplifier with broadband input is that the stage or stages before the forward-path dominant pole should not clip on a signal input twice the amplitude of full rated input to the complete amplifier with feedback. This condition is essentially independent of the amount of feedback applied. When the input is band-limited, this condition can be relaxed by a factor approximately equal to the ratio of signal bandwidth to amplifier closed-loop bandwidth.

The first stage in Fig. 3 has transfer conductance G_T = 1 mA/V and is capable of 8 mA peak-to-peak output. The improver itself is therefore free from gross TIMD for broadband input signals up to 4 V peak to peak.

Clearly the improver cannot increase the limiting slew rate of an amplifier, although it can precondition the input to the amplifier (by restricting its bandwidth) so that the demanded slewing rate is reduced. In the quasi-linear region of amplifier operation the improver reduces all distortion products, including those associated with transients, in accordance with Eq. (11).

3 HIGH-PASS THEORY

Fig. 4 is a block diagram of the improver in a high-pass system. The amplifier to be improved is modeled by its midband gain A_0 and low-frequency response function $\psi(s)$.

The forward path of the improver has midband gain $1/\beta$ and approximates a single-pole low-frequency cutoff of time constant τ_A. The overall feedback network has midband transmission β with phase-linear low-frequency cutoff also of time constant τ_A. The present improver is designed with

$$1/\beta = 31.6$$

$$\tau_A = 30\ ms.$$

Fig. 5 shows in detail how these transfer functions are realized. The necessary design constraints are

$$R_1C_1 = R_2C_2 = \tau_A \tag{14}$$

$$\frac{R_1 + R_2}{R_1} = \frac{1}{\beta}\,. \tag{15}$$

If the amplifier to be improved has a typical low-frequency response, the response of the system will approach the demanded gain set by the overall feedback network in the improver, namely,

$$A(s) \to \frac{1}{\beta}\left[\frac{s + 3/2\tau_A}{(s + 1/\tau_A)(s + 1/2\tau_A)}\right] \times \left[\begin{matrix}\text{other singularities}\\ \text{close to origin}\end{matrix}\right] \tag{16}$$

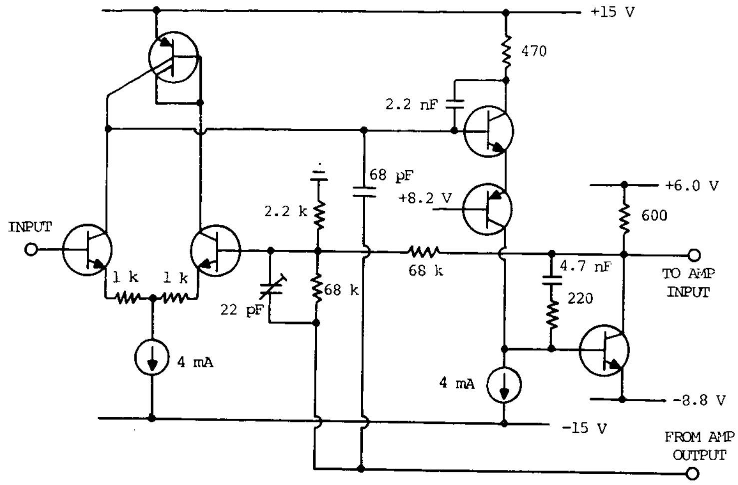


Fig. 3. Outline circuit of the improver (low-pass components only).

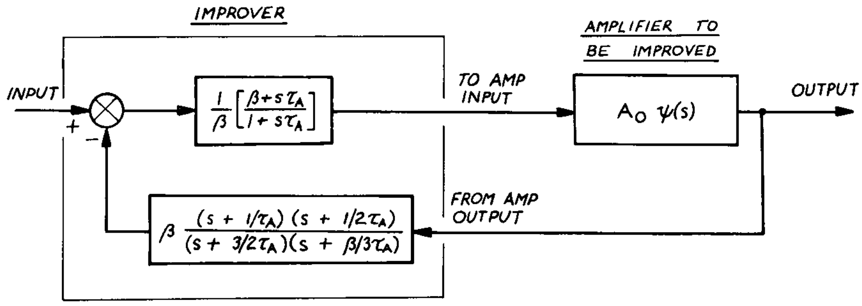


Fig. 4. Block diagram of the improver in a high-pass system.

This function is phase linear. For τ_A having its designed value 30 ms the tilt on a 20-Hz square wave is less than 5%.

As in the low-pass case, the improver may be considered as a kind of feedback network around the amplifier to be improved. The transfer function of this feedback network is

$$B(s) = \left.\frac{\text{to amplifier input}}{\text{from amplifier output}}\right|_{\text{input} = 0}$$

$$= -\frac{s + 1/2\tau_A}{s + 3/2\tau_A} . \tag{17}$$

There are two common forms of the low-frequency response $\psi(s)$ of the amplifier to be improved in Fig. 4:

1) The amplifier has a dc-coupled forward path. A capacitor in the feedback network contributes a closed-loop pole of time constant τ_F typically of the order 30 ms, and a coupling capacitor at the input contributes a second pole of time constant τ_C typically about 300 ms:

$$\psi(s) = \left[\frac{1/A_0 + s\tau_F}{1 + s\tau_F}\right]\left[\frac{s\tau_C}{1 + s\tau_C}\right] .$$

Fig. 6 is a typical outline circuit.

2) As for 1) above, with the addition of an output coupling capacitor which contributes a pole of time constant τ_D typically about 30 ms:

$$\psi(s) = \left[\frac{1/A_0 + s\tau_F}{1 + s\tau_F}\right]\left[\frac{s\tau_C}{1 + s\tau_C}\right]\left[\frac{s\tau_D}{1 + s\tau_D}\right] .$$

Fig. 7 shows the asymptotes of: the amplifier gain, type 1) above with the typical values for its time constants; the gain $|B(j\omega)|$ of the improver considered as a feedback network; and the gain around the resulting feedback loop.

Loop gain falls through unity at about 0.5 Hz, and loop phase shift is about 110° leading at this frequency. The system is stable, with a wide margin for amplifier parameter changes. Eq. (16) is a valid approximation for the system response.

For an amplifier of type 2) the extra pole brings the system to the point of instability. The improver provides an optional RC network for connection in series with the amplifier input (Fig. 8). This network has two effects:

1) It modifies the time constant of the pole associated with the input coupling capacitor.

2) It introduces a shelf in the loop frequency response.

The system can be made stable for any amplifier

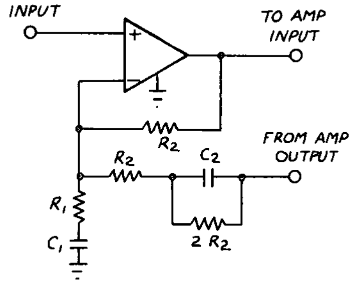
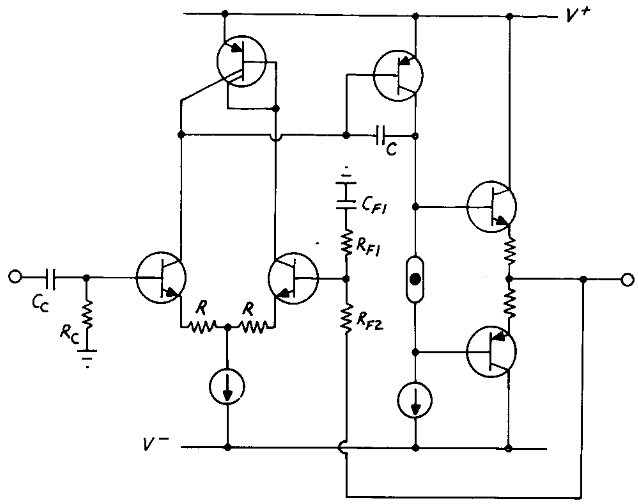


Fig. 5. Outline circuit of the improver (high-pass components only).

response by choosing R sufficiently large and C sufficiently small. The price of going too far is a departure of the system response from phase linearity. The values in the present improver, 220 kΩ and 150 nF, seem a reasonable compromise. Short circuiting the phase compensating capacitor C_2 in the improver also increases the stability margin but again degrades phase linearity.



$\tau_C = R_C C_C$

$\tau_F = R_{F1} C_{F1}$

$$\beta = \frac{R_{F1}}{R_{F1} + R_{F2}}$$

Fig. 6. Popular outline circuit topology of an audio power amplifier.

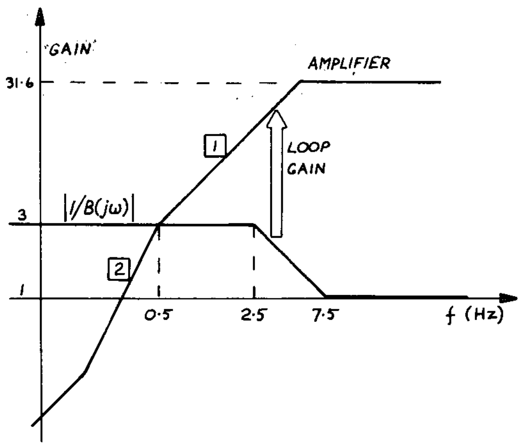


Fig. 7. The improver considered as a feedback network around the amplifier to be improved (high-pass case); gain asymptotes of the forward path (the amplifier to be improved) and feedback factor (the improver).

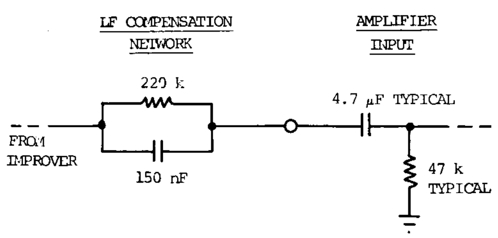


Fig. 8. Low-frequency compensation network for amplifiers of type 2).

4 EXPERIMENTAL RESULTS

Fig. 9 is the complete circuit of an improver designed for use with amplifiers whose upper 3-dB cutoff is above 80 kHz. Features not yet discussed include:

1) Separate outputs for use with noninverting and inverting amplifiers;

2) A loudspeaker relay, with fast opening and delayed closing, to prevent "thumps" at turnoff and turn-on;

3) A network to provide a well-defined load resistance for the amplifier at very high frequencies [11];

4) Long-tailed pair stages throughout plus catching diodes, to accelerate recovery from overdrive.

This improver has been used in conjunction with a number of amplifiers, and outline results follow for two cases.

The improver of Fig. 9 has not proved quite universal. Trouble has occurred with amplifiers that have inadequate slew-rate limits. Any multipole nonlinear feedback system will exhibit some form of jump resonance if it is driven hard enough at a high enough frequency. Whether or not this jump resonance remains as a sustained oscillation after the overdrive is removed depends on many factors. In the case of the improver, it appears that the amplifier should be capable of producing full sinusoidal output up to at least 30 kHz.

4.1 An Inexpensive Amplifier

This amplifier is of the basic topology shown in Fig. 6. The first stage operates at 1 mA per transistor, with 330 Ω in each emitter lead and 10 kΩ for the current source. The second stage is a single transistor operating at 5 mA, with C = 100 pF and a bootstrapped 3.9 kΩ for the current source. The third stage is a pair of Darlingtons operating at 100 mA quiescent, with 0.68 Ω in each emitter lead. Midband gain is 22.5, with lower and upper 3-dB frequencies 4.5 Hz and 185 kHz, respectively. This amplifier was selected to show what the improver can do for a simple circuit.

The combination of amplifier plus improver has a midband gain of 31.4, flat ± 0.5 dB from 20 Hz to 20 kHz. Figs. 10 and 11 compare 20-Hz and 10-kHz square waves for the amplifier without and with the improver; note the improvement in phase linearity. Fig. 12 shows detail of a 1-kHz sine wave for the amplifier with improver under overdrive conditions; this illustrates the unconditional stability of the system.

Figure 13 is the measured spectral density of equivalent noise voltage referred to the system input. System noise is almost entirely determined by the improver, as its gain precedes amplifier noise. At frequencies around 5–10 kHz noise is dominated by thermal contributions from resistors in the first stage and its current mirror. The circuit is proportioned so that each resistor contributes its open-circuit noise voltage to the equivalent noise referred to the input:

1.0 kΩ input stopper
2 × 1.0 kΩ in emitters

2.2 kΩ feedback resistor
2 × 1.0 kΩ in current mirror
1–2 kΩ total base resistance of four transistors.

The measured spectral density never quite reaches its horizontal asymptote, but at its minimum corresponds to 15 kΩ.

At low frequencies flicker noise on the base current of the input transistors becomes dominant; these transistors should be high-gain low-noise types (BC109 or 2N930 were used in prototypes). A small contribution to flicker noise is also made by the transistors in the current mirror; low-noise types could be substituted if desired.

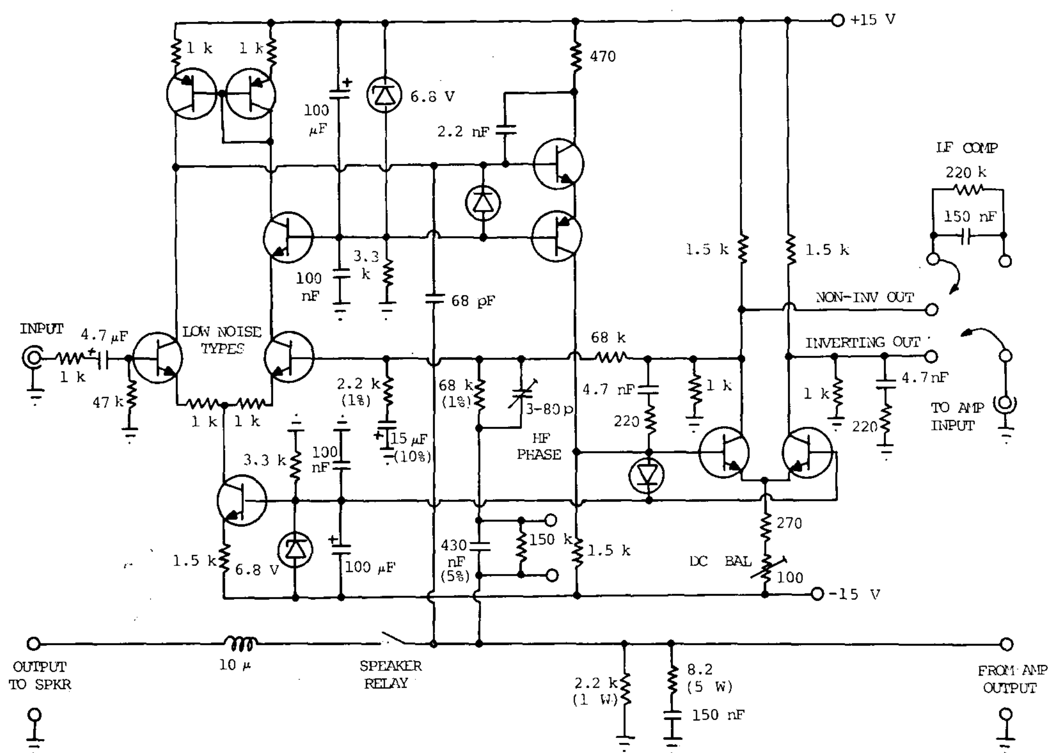
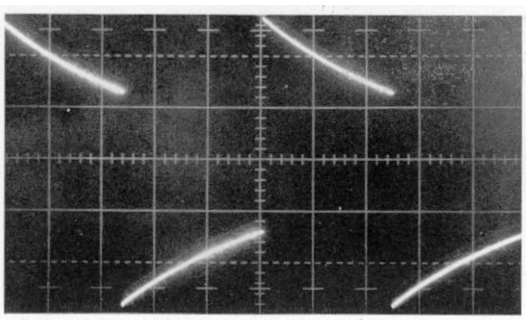
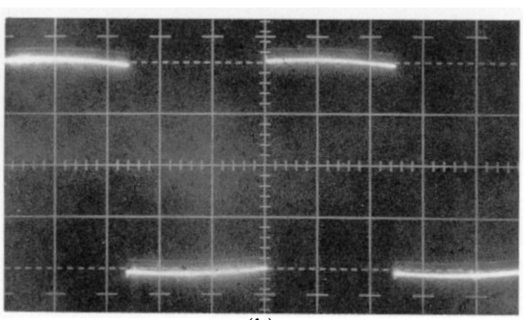


Fig. 9. Complete circuit of an amplifier improver. Unmarked transistors are general-purpose types.

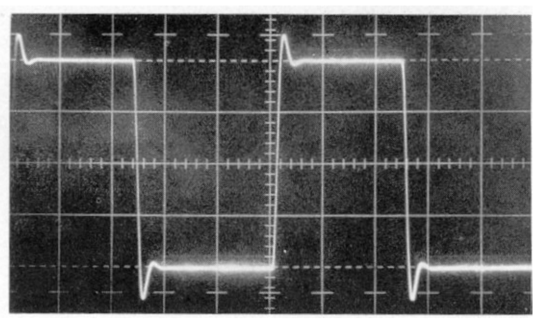


(a)

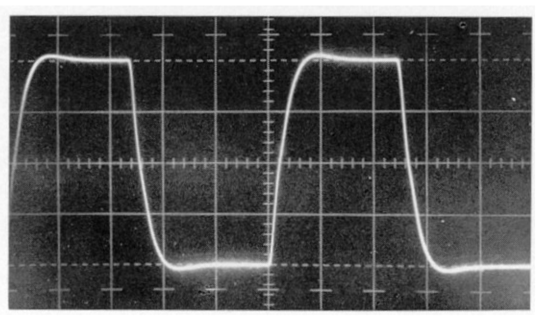


(b)

Fig. 10. 20-Hz square-wave response of a simple amplifier. (a) Without improver. (b) With improver. Vertical scale 2 V/div; horizontal scale 10 ms/div.



(a)



(b)

Fig. 11. 10-kHz square-wave response of a simple amplifier. (a) Without improver. (b) With improver. Vertical scale 5 V/div; horizontal scale 20 µs/div.

The gains of the second and third stages fall above 1 kHz, and noise sources other than in the first stage become significant. At high frequencies the dominant noise source in the given circuit is the −15 V supply rail; noise voltage on this rail appears between collector and base of the noninverting output transistor where it injects current into the base circuit via 4.7 nF, and this current divided by the falling combined gains of the first and second stages appears as an f^2 component in the input spectral density. The arrangement of Zener diodes and bypass capacitors is intentional.

Table 1 compares second- and third-harmonic distortion without and with the improver at four frequencies. Note the approximate 26-dB reduction of harmonics at midfrequencies, and the much greater reduction around 14–15 kHz associated with the dashed β network in Fig. 1. Note also that total noise-plus-distortion measurements might be misleading unless the supersonic f^2 component of noise is rejected.

Fig. 14 shows the current output waveform for the amplifier first stage when the system is driven to half maximum output by a fast-rise square wave. Without the improver the peaks are clipped and TIMD is generated; clipping does not occur when the improver is in the system, even at full output.

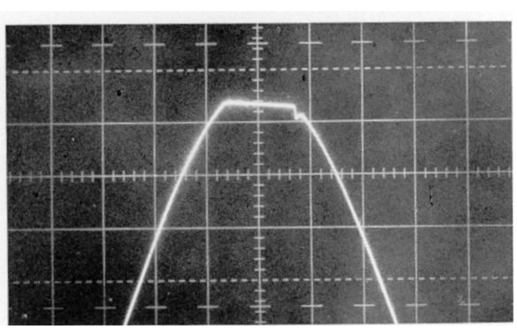


Fig. 12. 1-kHz sine-wave response of a simple amplifier plus improver. Detail of overdrive and recovery. Vertical scale 5 V/div; horizontal scale 50 μs/div.

4.2 A Very-Low-Distortion 50-W Amplifier

This amplifier was chosen in an attempt to set an upper bound on residual distortion in the improver. The amplifier has a midband gain of 66.6, ± 3 dB at 1.8 Hz and 160 kHz. Table 2 compares second and third harmonics without and with improver at four frequencies. Evidently improver residuals in the frequency range up to 20 kHz are below 10 ppm.

5 CONCLUSIONS

An amplifier improver of the type described in this paper appears a viable alternative to the purchase of an expensive new amplifier for the audio enthusiast who wishes to upgrade his system. The improver parameters β, τ_A, and τ_X can, of course, be changed to suit particular amplifiers, but the chosen values seem a reasonable compromise and cover most of the available range.

Patent application has been made for the nest of dif-

Table 1. Comparison of percentage second- and third-harmonic distortions for a simple amplifier without and with the improver. Measurements taken at 20 W into 8 Ω.

Fundamental	Harmonic	Without Improver	With Improver
1 kHz	D_2 (2 kHz)	0.028	0.0013
	D_3 (3 kHz)	0.11	0.0038
5 kHz	D_2 (10 kHz)	0.10	0.0042
	D_3 (15 kHz)	0.13	0.0029
7 kHz	D_2 (14 kHz)	0.12	0.0032
	D_3 (21 kHz)	0.14	0.012
10 kHz	D_2 (20 kHz)	0.14	0.014
	D_3 (30 kHz)	0.16	0.038

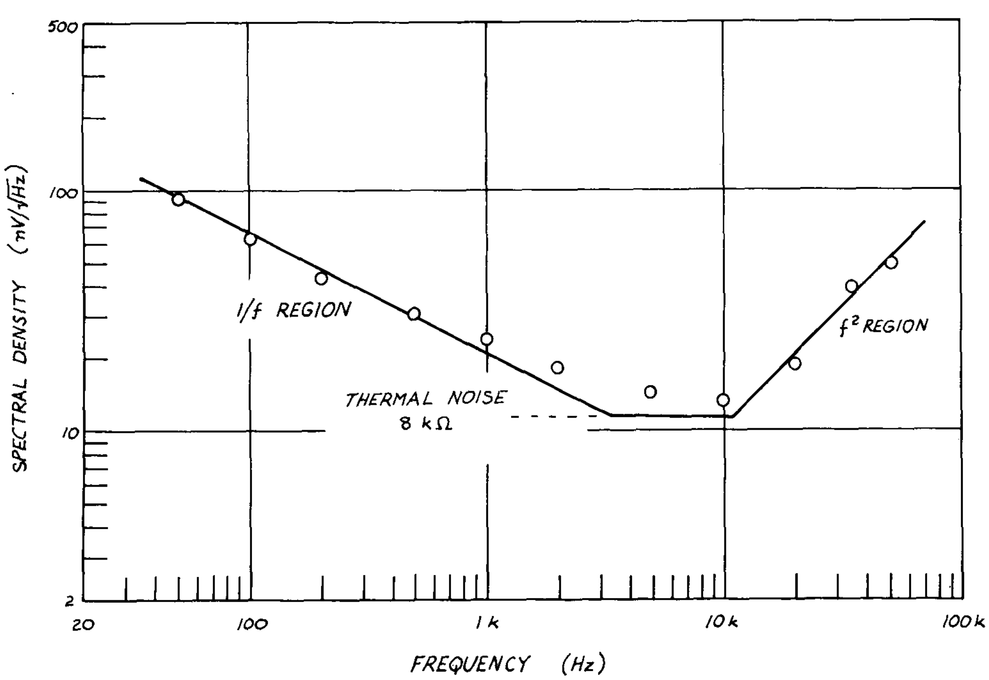
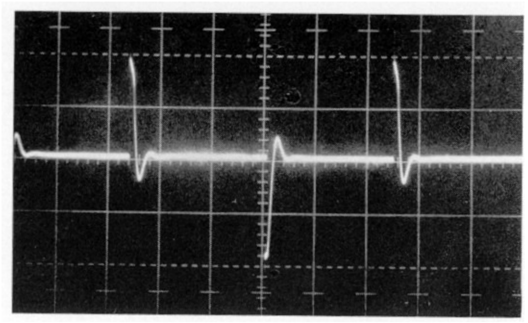
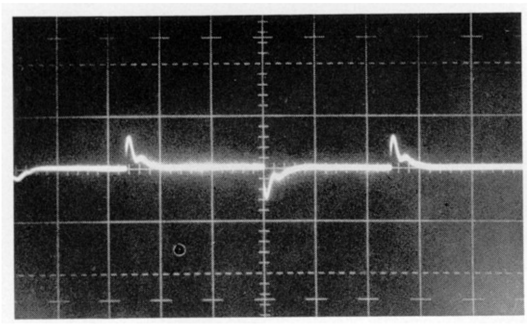


Fig. 13. Spectral density of equivalent noise voltage referred to the input.



(a)



(b)

Fig. 14. Current output from the first stage of a simple amplifier of the topology of Fig. 6, with 10-kHz square-wave output as in Fig. 11. (a) Without improver. (b) With improver. Vertical scale 1 mA/div; horizontal scale 20 μs/div.

Table 2. Comparison of parts-per-million second- and third-harmonic distortion for a very-high-quality amplifier without and with the improver. Measurements taken at 50 W into 8 Ω.

Fundamental	Harmonic	Without Improver	With Improver
100 Hz	D_2 (200 Hz)	7*	7*
	D_3 (300 Hz)	7*	7*
1 kHz	D_2 (2 kHz)	7*	7*
	D_3 (3 kHz)	2*	2*
5 kHz	D_2 (10 kHz)	7	4
	D_3 (15 kHz)	16	7
10 kHz	D_2 (20 kHz)	4	4
	D_3 (30 kHz)	37	12

* Readings not accurate; dominated by flicker noise in test set.

ferentiating loops on which the amplifier improver depends. Enquiries regarding its commercial exploitation, either as an add-on facility such as an improver or in complete amplifiers, should in the first instance be addressed to the Legal Officer, Monash University.

6 REFERENCES

[1] E. M. Cherry, "A New Result in Negative-Feedback Theory and Its Application to Audio Power Amplifiers," *Int. J. Circuit Theory*, vol. 6, pp. 265–288 (1978 July).

[2] E. M. Cherry, "A High-Quality Audio Power Amplifier," *Proc. IREE (Australia)*, vol. 39, pp. 1–8 (1978 Jan./Feb.).

[3] H. W. Bode, *Network Analysis and Feedback Amplifier Design* (van Nostrand, Princeton, NJ, 1945).

[4] M. Otala, "Transient Intermodulation Distortion in Transistorized Audio Power Amplifiers," *IEEE Trans. Audio*, vol. AU-18, pp. 234–239 (1970 Sept.).

[5] M. Otala and E. Leinonen, "Extension of the Theory of Transient Intermodulation Distortion," *Proc. IREE (Australia)*, vol. 37, pp. 53–59 (1976 Mar.); republished *IEEE Trans. Acoust., Speech, Sig. Proc.*, vol. ASSP-25, pp. 2–8 (1977 Feb.).

[6] W. M. Leach, "Transient IM Distortion in Power Amplifiers," *Audio*, vol. 59, pp. 34–41 (1975 Feb.).

[7] P. Garde, "Transient Distortion in Feedback Amplifiers," *Proc. IREE (Australia)*, vol. 38, pp. 151–158 (1977 Oct.); republished in *J. Audio Eng. Soc.*, vol. 26, pp. 314–322 (1978 May).

[8] E. M. Cherry, "Three Audio-Amplifier Dragons," *Proc. IREE (Australia)*, vol. 37, pp. 354–360 (1976 Dec.).

[9] E. M. Cherry and M. Otala, "Comments on 'Extension of the Theory of Transient Intermodulation Distortion,' " *IEEE Trans. Acoust., Speech, Sig. Proc. (Correspondence)*, vol. ASSP-27, pp. 653–654 (1979 Dec.).

[10] E. M. Cherry, "Transient Intermodulation Distortion," *IEEE Trans. Acoust., Speech, Sig. Proc.*, accepted for publication 1981.

[11] A. N. Thiele, "Load Stabilizing Network for Audio Amplifiers," *Proc. IREE (Australia)*, vol. 36, pp. 297–300 (1975 Sept.).

THE AUTHOR

Edward Moore Cherry was born in Melbourne, Australia, in 1936. The University of Melbourne awarded him Bachelor's and Master's degrees in science in 1957 and 1959, respectively, and a Ph.D. in electrical engineering in 1962.

From 1963 until 1965 Dr. Cherry was a member of the technical staff of Bell Telephone Laboratories in New Jersey, and in 1966 he was a temporary research associate of the United Kingdom Atomic Energy Authority in Harwell. In 1967 he returned to Australia to assume a position at Monash University, where he is presently associate professor of electrical engineering. While on a sabbatical in 1973, he returned to Bell Laboratories on a Fulbright scholarship.

Professor Cherry's research interest is in electronic hardware. He has published over 50 papers and a book in that area. A Fellow of the Institution of Radio and Electronics Engineers, Cherry has twice been awarded that Institution's annual medal for the most meritorious paper published in its Proceedings.