A Practical Approach to an Audio-Amplifier Design

This Note discusses general considerations, design requirements, and performance for a 20-watt, hi-fi amplifier.

There are many different kinds of audio amplifiers, including hi-fi amplifiers, musical-instrument amplifiers, public-address amplifiers, phonograph amplifiers, portable-radio amplifiers, and auto-radio amplifiers. The design considerations vary with each of these. The hi-fi amplifier has the most stringent specifications because it must reproduce, more closely than the other amplifiers, an original input sound. A musical-instrument amplifier requires slightly less stringent distortion and bandwidth specifications, but it must be very conservatively designed in terms of its ability to drive a highly reactive, low-impedance load. A public-address-system amplifier can have a narrower power bandwidth than a hi-fi amplifier, although hi-fi amplifiers are being used more and more as public-address amplifiers. The remaining group of amplifiers has varied, but generally less stringent, specifications. This Note describes the specifications and a simplified practical construction procedure for a hi-fi amplifier design.

The power amplifier is only one of the elements in a signal-processing chain, and if it could be an ideal amplifier, the signal at its output would be only of a larger magnitude than that at the input; in all other aspects the output signal would be identical to the input signal. However, since there is no ideal amplifier, quality is judged on the basis of a number of parameters: THD (total harmonic distortion), IMD (intermodulation distortion), hum and noise, phase shift, transient intermodulation distortion (TIM), total bandwidth, rise time, and slew rate.

It is a generally accepted fact that, in order to provide good performance, the open-loop response of an amplifier should be at least 20 kHz,[1,2,3] that is, the −3 dB point should be above 20 kHz, a condition often difficult to achieve. One method used for improvement of the open-loop response employs local feedback with no compensation. If there is a need for compensation in the forward loop, a lead compensation should be used whenever possible[3]; the lead compensation increases the bandwidth and enables the amplifier to achieve a higher slew rate. The slew rate is a very important parameter; if the slew rate is too low, drastic transient intermodulation distortion (TIM), also called slew-induced distortion[4], can occur. The TIM takes place when an input causes the amplifier to run into slew-limited operation, a condition that occurs when the input signal changes faster than the amplifier can follow. This condition indicates that the amplifier is operating in an open-loop mode and that the feedback is trying to correct the output waveform, and very high overshoots occur within the loop. During the recovery, considerable TIM results.

One way in which the TIM can be minimized is by using a small amount of feedback.[3,5] However, it is possible to achieve low TIM even if the feedback is moderate to high, provid-

Printed in USA/3-79

ed the resulting slew rate of the amplifier is high enough.[4,6] In such a case, the only difference between a design using a small amount of feedback and one using a large amount of feedback could be in the recovery from clipping.

The slew rate necessary to essentially eliminate transient intermodulation distortion has not been firmly established; one author suggests 50V/μs for a 100W/8Ω amplifier,[7] while another suggests 0.5 to 1V/μs per peak output volt.[4] It is clear that the higher the power of the amplifier, the higher the slew rate should be, and that this parameter, slew rate, should be normalized, i.e., the specification should be given as V/μs/V. This is the only meaningful way of comparing amplifiers capable of delivering different amounts of power.

Circuit Description

There are a number of ways to design an amplifier for best possible performance; the 20-watt circuit described below uses only one of them. The circuit diagram of the amplifier, Fig. 1, is almost identical to the diagrams of the 100- and 300-watt circuits previously published by RCA[8]. The circuit comprises an integrated-circuit input stage and a power stage consisting of discrete transistors; it can be treated as two cascaded gain blocks with one common feedback loop. The discrete gain block has its own local feedback provided by R_{SD} and R_{FD}. The CA3140B integrated circuit was chosen as the input stage because it offers high input impedance, low noise, low offset, excellent thermal stability, direct coupling, and a very good slew rate. The power stage is designed to stress wide bandwidth and ultra-low THD. The input stage of the discrete section is a class A common-base stage that serves as a voltage translator. It translates the single-ended input voltage into the higher-level voltage needed to drive the rest of the circuit. The next stage, also class A, is mainly for voltage amplification. The upper and lower portions of this stage are connected to the V_{BE} multiplier, which provides the bias for the output section, the driver stage, and the output stage. Both the driver stage and the output stage are emitter followers, and are used to achieve the needed current gain. There is also a current-limiting circuit connected across the input of

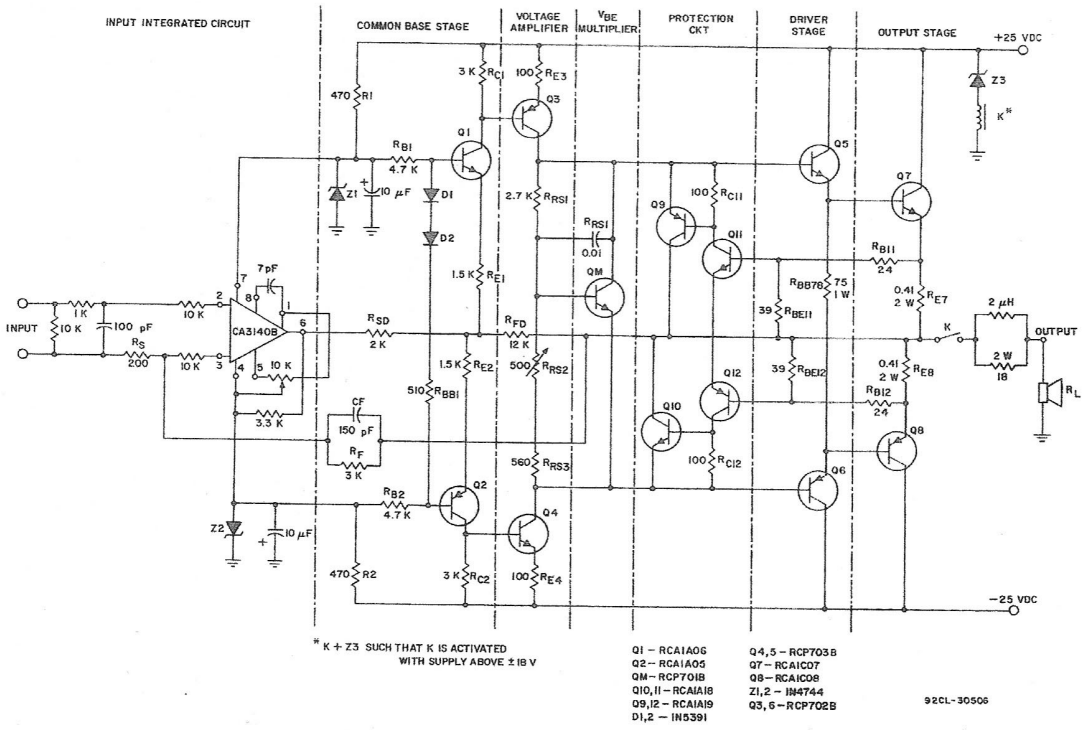

Fig. 1—Circuit diagram of the 20-watt amplifier.

the driver stage. Current limiting is necessary to protect the amplifier against excessive dissipation and possible destruction, particularly should the speaker leads become shorted.

As mentioned above, local feedback for the power stage of the 20-watt amplifier is provided by R_{SD} and R_{FD}; the feedback enables the power section to maintain the −3 dB point at 166 kHz. The amount of the feedback and the achieved corner frequency depend upon the f_T of the transistors used. It is desirable that the transistors used have as high an f_T as possible. The common-base stage, the common-emitter stage, and the driver stage (common-collector stage) all employ transistors with an f_T of 50 MHz or greater; the output stage uses 5-MHz devices. This choice, in conjunction with the high-speed integrated circuit and properly selected feedback,[9] results in a typical slew rate of 30V/μs as well as a very low distortion throughout the audio range. To keep the harmonic distortion low, every stage of the power section is designed to be as linear as possible and, for that reason, the output stage as well as both class A stages are biased for fairly high idling current. The quiescent current in the output stage is set to 150 milliamperes.

Calculations

OUTPUT STAGE

Power-amplifier calculations begin with the output stage, and are based on the power to be delivered to a given load. For 20 watts into an 8-ohm load, the peak voltage and the peak current across the load are;

$$E_m = \sqrt{2\,PR} = \sqrt{2 \times 20 \times 8} = 17.9\ V$$

$$I_m = \sqrt{\frac{2P}{R}} = \sqrt{\frac{2 \times 20}{8}} = 2.24\ A$$

Based on these values, the supply voltage can be determined and the type of output transistor chosen. Should a supply voltage of 25 volts be chosen (this choice will be explained below), the output transistor would have to have a V_{CEO} rating equal to the maximum power-supply voltage. The safe operating area of the output device can be determined initially by using a voltage equal to one half of the total supply voltage and a current value somewhat greater than the peak current the transistor must handle. In the case at hand, the minimum safe operating area is determined by values of 25 volts and 3 amperes. As an additional requirement, the output transistor must have good current gain at the peak required current, and it should be chosen so that the peak current falls on the linear portion of the dc current gain curve whenever economically feasible; the f_T should be as high as possible. Transistors that reasonably fulfill these requirements are the RCA1C07 and the RCA1C08.

The emitter resistance of the output device is determined based on the idling current and the necessary amount of thermal compensation. Obviously, the higher the value of emitter resistance, the better the thermal stability. However, at the same time, the voltage drop across the resistor lowers the efficiency of the amplifier. The final value is, then, a compromise between the two requirements cited. Since the idling current in this design was selected as 150 milliamperes, a value of 0.41 ohm was chosen as a suitable value of emitter resistance. This value provides a 61.5-millivolt drop and an adequate thermal compensation for the output stage.

The base-to-emitter resistance of the output device should be as low as possible to improve both the V_{CER} of the output device and the switching speed. On the other hand, too low a value would demand an additional current from the driver transistor. In the design described in this paper, a 78-ohm resistor was connected between the bases of the output transistors. Such a connection provides two benefits: no current flows through R_{BB78} when the current through the output device is peaked because the opposite transistor is not conducting, and the switching speed of the output section is im-

proved because the opposite transistor is providing reverse base drive through R_{BB78}.

Safe-Area and Heatsink Requirements

The safe-area rating of the output device is one of the most important aspects of an audio-amplifier design because most speakers are reactive rather than resistive loads. The reactive load causes a phase shift between the voltage and the current and varies with the frequency. The largest phase shift usually occurs at the low end of the frequency spectrum. The output transistor must, therefore, operate under difficult conditions; it must conduct a high current when there is a high voltage across it. There is a further complication in that all of this happens near the lower end of the audio spectrum. The difficulty derives from the fact that the highest amplitude of the musical signal is in the low-frequency region, and that the speakers usually demonstrate their lowest impedance and largest phase shift in the same region. In addition, the lower the frequency, the more difficult are the operating conditions for the output transistor, based on its thermal time constant. The maximum phase shift possible is 90°, but since only the ideal inductor or capacitor can cause a 90° shift, it is not likely to occur. However, a 60° shift is not unusual for some speakers. If a complex load is driven, the average dissipation per output transistor based on the sine wave input signal is given by:

$$P_{av} = \frac{E_o I_m}{\pi} - \frac{E_m I_m}{4} \cos \phi$$

where E_o is one half of the total supply voltage and E_m and I_m are the peak voltage and the peak current across the load. It is obvious that the highest dissipation occurs at cos $\phi = 0$, that is, when the phase shift is 90°. The voltage and current waveforms for the resistive and reactive loads are shown in Figs. 2 and 3.

At a 60° shift, the output transistor will experience approximately one

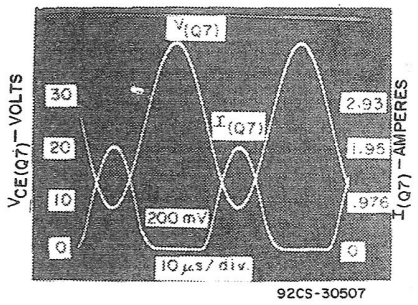

Fig. 2— Voltage and current waveforms for transistor Q7 with resistive load. (Note: Current I_{Q7} is read across R_{E7} (0.41 Ω); thus I (0.41 Ω) = 200 mV with I = 0.488 ampere/division.)

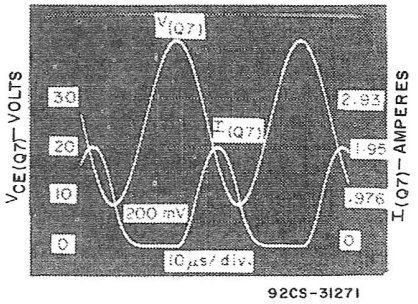

Fig. 3— Voltage and current waveforms for transistor Q7 with complex load, 60° shift. (Note: Current I_{Q7} is read across R_{E7} (0.41 Ω); thus I (0.41 Ω) = 200 mV with I = 0.488 ampere/division.)

half the supply voltage and 0.866 I_m at the same time. The most conservative design for an amplifier allows it to deliver full rated power into an ideal inductor or capacitor. This design mandates that the peak current through the transistor in the case of a shorted output be slightly above the peak current at full rated power. It is for this reason that the criteria for the safe operating area mentioned above, i.e., current somewhat higher than the required peak current and voltage one half of the supply voltage, was established. The current limit for the 20-watt amplifier is set at 2.91 amperes. The peak dissipation with the output shorted is:

$$P_{pk} = I_m E_o - I^2_m R_E$$
$$= 2.91 \times 25 - 2.91^2 \times 0.41 = 70.9 \text{ W}$$

The heatsink should be selected so that the junction temperature of the output transistor remains below the maximum allowed temperature under the worst possible conditions. It is

advisable, however, to keep the junction temperature as low as possible since the operating life of the transistor approximately doubles for every 10 °C reduction of the junction temperature.

The output devices selected have maximum thermal resistances of 1.67 °C/W. With the devices mounted directly on the heatsink, and with the use of a thermal compound, such as Dow Corning 340, and a mounting torque[10] of four to eight inch-pounds, the total thermal resistance between the junction and the heatsink is 1.87 °C/W in the worst case.

A *simplified* thermal model of the transistor consists of the parallel RC network shown in Fig. 4.[1] The heat-

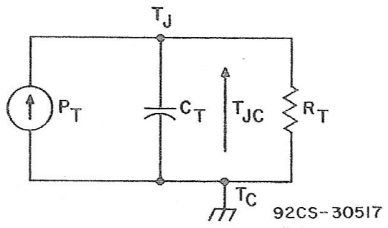

Fig. 4—Simplified thermal model of a transistor.

sink temperature may be regarded as constant because its thermal time constant is very long in comparison with the thermal time constant of the transistor. The junction temperature of the transistor will vary as a function of the input signal and the thermal time constant of the transistor. The thermal time constant depends mostly upon the size of the silicon chip and the construction of the package. It is desirable that the package have low thermal resistance and high thermal capacitance so that the ac signal will produce a minimal junction temperature change and the temperature will be as low as possible. If the input signal is a square wave (as is the case when the output is shorted), the approximate junction temperature rise is given by:

$$\Delta T_J = P_{pk} \times \theta_{JHS} \frac{1 - e^{\frac{-t}{\tau}}}{1 - e^{\frac{-2t}{\tau}}} \qquad (1)$$

where P_{pk} = peak power
θ_{JHS} = thermal resistance, junction to heatsink
t = duration of on time of the pulse
τ = thermal time constant of the transistor pellet and package combination

With the current limit set to 2.91 amperes and the supply voltage at ±25 volts, peak power is 70.9 watts. The thermal time constant of the transistor must be determined by measurement, and in the case of the RCA1C07 and RCA1C08, the shortest time constant measured was 30 milliseconds. For the purposes of calculation, including margin, 25 milliseconds was selected as a safe number.

The lowest musical note usually encountered is low A on the piano, 27.5 Hz. It should be noted, however, that this frequency does not represent the absolute worst-case condition, but an arbitrarily chosen criteria. The lowest possible musical note is 16 Hz, while other conditions, such as turntable rumble, low-frequency interference from an FM receiver, or turn-on, turn-off transients could be worse for the output section than the 27.5-Hz signal selected.

The 27.5-Hz frequency yields a pulse duration, t, of 18.2 milliseconds and, from Eq. 1:

$$\Delta T_J = 70.90 \times 1.87 \frac{1 - e^{\frac{-18.2 \times 10^{-3}}{25 \times 10^{-3}}}}{1 - e^{\frac{-36.4 \times 10^{-3}}{25 \times 10^{-3}}}}$$

$$= 89.4^\circ C$$

With a T_J(max) of 150 °C, the maximum transistor case temperature is:

$$T_C(max) = T_J(max) - \Delta T_J$$
$$= 150 - 89.4 = 60.6^\circ C$$

If the maximum ambient temperature (T_A(max)) is 40 °C, the maximum ther-

mal resistance between the transistor case and the ambient must be:

$$\theta_{CA} = \frac{T_C(max) - T_A(max)}{P_{av}} = \frac{60.6 - 40}{35.45} = 0.58^\circ C/W$$

per output transistor. $P_{av} = \frac{1}{2} P_{pk}$ because of a square-wave pulse.

With the heatsink as determined above, the junction temperature never exceeds the maximum allowed 150°C. On the other hand, a case-to-ambient thermal resistance of 0.58 °C/W per output device indicates the need for a very large heatsink, certainly not an economical item for a 20-watt amplifier. A smaller heatsink used in conjunction with a thermal cutout is a much better solution. If a heatsink such as a Wakefield No. 421 is used (one per output device), the thermal cutout must disconnect the power at:

$$T_S(max) = T_C(max) - P_{av} \times \theta_{CHS} = 60.60 - 35.45 \times 0.2 = 53.5^\circ C$$

or sooner to prevent $T_J(max)$ from exceeding 150 °C. A 52 °C cutout should be used to provide maximum flexibility under the described conditions.

It should be noted that the criteria described is very conservative and not commonly used in practice because it mandates costlier design. If the lowest note of the base guitar were considered to be the lowest input frequency (40 Hz), Eq. 1 would yield a junction temperature rise of:

$$\Delta T_J = 70.90 \times 1.87 \frac{1 - e^{\frac{-12.5 \times 10^{-3}}{25 \times 10^{-3}}}}{1 - e^{\frac{-25 \times 10^{-3}}{25 \times 10^{-3}}}} = 82.5^\circ C$$

which would permit the use of a smaller heatsink and a higher cutoff temperature.

It is important that each heatsink be electrically isolated so that the output transistor can be mounted directly on it. This arrangement provides a very low θ_{CHS}, as mentioned earlier. If a 0.002-inch mica washer is placed between case and heatsink, θ_{CHS} becomes much higher,[10] approximately 1.3 °C/W, which is a much worse condition than previously encountered.

The worst musical condition for an amplifier is rock music received from an FM station (as determined empirically).[11] This condition causes an average power dissipation in the output stage of 15 percent of the maximum power. Taking 26 watts as maximum power, with a 60° phase-shift load, the dissipation per output device, P_{avM}, is approximately:

$$P_{avM} = \frac{E_o I_m}{\pi} - \frac{E_m I_m}{4} \cos\phi$$

$$P_{avM} = \frac{25 \times 0.987}{\pi} - \frac{17.9 \times 0.987}{4} \times 0.5 = 6.88\ W$$

It should be noted that this dissipation is very close to the maximum dissipation into an 8-ohm resistive load:

$$P_{max} = \frac{25 \times 1.61}{\pi} - \frac{12.9 \times 1.61}{4} = 7.62\ W$$

In both cases 26 watts is taken as maximum power because a typical current-limiting effect, which produces 26 watts into an 8-ohm load, occurs at 2.56 amperes.

With a Wakefield No. 421 heatsink per output device (approximately 1.67 °C/W at 7 watts dissipation) the highest heatsink temperature under the worst musical input condition is:

$$T_S(max) = T_A(max) + P_{avM}\theta_{HSA} = 40 + 6.88 \times 1.67 = 51.5^\circ C$$

which just approaches the cutoff temperature of 52 °C. This result indicates that the transistor case temperature will not reach 60 °C under normal operating conditions, and that the design described guarantees a high number of thermal cycles as well as long operating life.

With the procedure described above, the selection of a safe operating area for the transistor is based on the peak allowable current

and one half the supply voltage. However, the slope of the Is/b curve of the selected transistor must not be too steep, and it must at least satisfy the load line for a 60° complex load. The voltage-current relationship for the output transistor (for example, Q7) loaded with a 60° complex load is given by:

$$e_{CE7} = E_o - e_L - i_C R_E$$

$$e_L = I_m Z_L \sin \omega t$$

$$i_C = I_m \sin (\omega t - \pi) \Big|_{\frac{\pi}{3}}^{\frac{4\pi}{3}}$$

This equation, in the worst case when $I_m = 2.91$ A, is shown plotted in Fig. 5. The safe operating area curves for the RCA1C07 and RCA1C08 at 25°C and 60°C case temperature are also shown in Fig. 5. The idling cur-

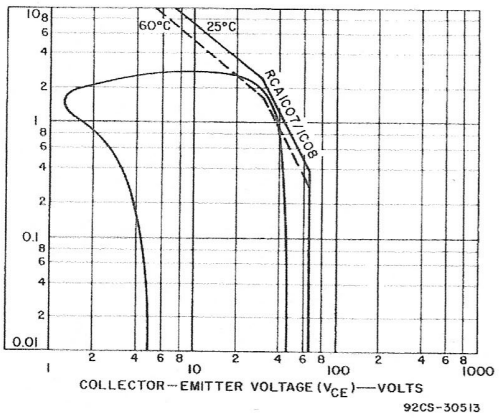

Fig. 5—Load line for transistor Q7 with 60° load and Im = 2.91 A; $V_{CC} = \pm 25$ V.

rent of as much as 150 milliamperes can be neglected and is not included in the plot.

The load line exceeds the rated curve between 20 and 39 volts. Based on the continuous rating, this condition would not be allowed. However, when the lowest frequency signal input is 27.5 Hz, the condition would only exist for 5 milliseconds and would be allowed since the thermal time constant of the output transistor is at least 25 milliseconds. If the thermal time constant is not known, the safest procedure is to reduce the case temperature so that the load line is inside the safe operating area curve.

Protection Circuit

The protection circuit is necessary to prevent excessive dissipation in case of a short-circuited output stage or too low a load impedance. There are many ways of protecting the output stage and several different circuits are commonly used. The two in most general use are the current-limiting and the load-line limiting circuits. The current limiting circuit simply limits the current; the load-line limiting circuit varies the current limit as a function of the voltage across the output transistor of the amplifier. This feature makes the load-line limiting circuit attractive because it allows the use of a smaller heatsink or a transistor with less safe operating area under the same operating conditions. On the other hand, the load-line limiting circuit usually does not allow high phase shift to occur, which means that the full rated power can be utilized only if the speaker load is mostly resistive. The current-limiting circuit allows maximum power at any load, but this means more difficult operating conditions for the output stage and a larger heatsink for the same power output.

The 20-watt amplifier employs the current-limiting circuit, as shown in Fig. 1. Two transistors are used for each half of the output section. The advantage to using two transistors instead of one is the sharper "knee" of the limiting circuit, which assures better utilization of the output transistors. In practical terms, the peak allowed current (the current at which the protection circuit is activated) can be very close to the peak current for the rated power output. The limiting circuit functions as a voltage-operated circuit; i.e., the current is sensed as a voltage drop across R_{E7} and R_{E8}. This method of operation mandates that the values of R_{B11} and R_{BE11} be as low as possible. Transistor Q9 must be capable of sinking any excess current coming from Q3 (the same holds true for Q10 and Q4 because of the symmetry of the circuit). In the worst case, with the input overdriven, the

highest current to be shunted away from Q5 is approximately 106 milliamperes. At that current, with a 2-volt V_{CE}:

$$I_{B9} = \frac{I_{Q9}}{h_{FEQ9}} = \frac{106}{20} = 5.3 \text{ mA}$$

$$I_{C11} = I_{B9} + I_{RC11} = I_{B9} + \frac{V_{BE9}}{R_{C11}}$$

The value of resistor R_{C11} must be kept low to minimize any possible leakage of transistor Q9, but the value must be high enough so that I_{C11} can create sufficient voltage for the operation of Q9; 100 ohms was chosen as a suitable value. From Eq. 2:

$$I_{C11} = 5.3 \times 10^{-3} + \frac{0.75}{100} = 12.8 \text{ mA}$$

$$I_{B11} = \frac{I_{C11}}{h_{FEQ11}} = \frac{12.8}{20} = 0.64 \text{ mA}$$

The current through the combination of R_{B11} and R_{BE11} should be much higher than I_{B11}.

$$V_{RE7} = I_{Q7} \times R_{E7} = 2.91 \times 0.41 = 1.19 \text{ V}$$

V_{BE11}, the voltage at which Q11 will just start to conduct, is 0.74 volt in the worst case. Thus:

$$\frac{V_{BE11}}{V_{RE7}} = \frac{R_{BE11}}{R_{BE11} + R_{B11}} = \frac{0.74}{1.19} = 0.622$$

If:

$$\frac{I_{B11}}{I_{RB11}} \leqslant 25$$

then:

$$(R_{B11} + R_{BE11}) \leqslant \frac{V_{RE7}}{25\ I_{B11}}$$

$$= \frac{1.19}{25 \times .64 \times 10^{-3}}$$

$$= 74.4 \text{ ohms}$$

And:

$$R_{BE11} = 0.622 \times 74.4 = 46.3 \text{ ohms}$$

$$R_{B11} = 74.4 - 46.3 = 28.1 \text{ ohms}$$

Since these are not standard values, R_{BE11} was selected as 39 ohms and R_{B11} as 24 ohms.

$$I_{RB11} = \frac{V_{RE7}}{R_{B11} + R_{BE11}}$$

$$= \frac{1.19}{63} = 18.9 \text{ mA} >> I_{B11}$$

$$V_{BE11} = V_{RE7} \times \frac{R_{BE11}}{R_{B11} + R_{BE11}}$$

$$= 1.19 \times \frac{39}{63} = 0.737 \text{ V}$$

It is important to note that this calculation was made with the highest V_{BE} possible for transistors Q9 and Q11. If a typical V_{BE} is taken into consideration, the peak amplifier current would be lower. With a V_{BE} of 0.65 volt for the RCA1A18 and the RCA1A19, peak current is:

$$V_{RE7} = V_{BE11} \frac{R_{B11} + R_{BE11}}{R_{BE11}}$$

$$= 0.65 \times \frac{63}{39} = 1.050 \text{ V}$$

$$I_m = \frac{V_{RE7}}{R_{E7}} = \frac{1.050}{0.41} = 2.56 \text{ A}$$

These values indicate that, in a typical case, the load-line would be a bit safer than that shown in Fig. 14. It should also be noted that V_{BE11} must never be low enough to allow the limiting current to approach the peak rated current of the amplifier. In practice, the V_{BE} cannot be much lower than 0.65 volt for a transistor such as the RCA1A18. For an I_m of 2.25 amperes:

$$V_{BE11} = (2.25 \times 0.41) \frac{39}{63} = 0.57 \text{ V}$$

a value too low to initiate conduction in Q11. The significance of this result is in the implied admonition that the protection circuit should not be allowed to interfere with amplifier operation up to 20 watts into an 8-ohm load.

The above design of the 20-watt circuit is based on the use of a

±25-volt regulated power supply. If an unregulated power supply is used, the full-load condition should not drop the voltage below ±25 volts. The no-load voltage depends upon the compromise between the price and the desired regulation of the power supply. Based on the no-load voltage, a new load-line for the output stage must be plotted and the operating conditions changed to guarantee operation within the ratings of the transistors used.

DRIVER STAGE

The minimum current gain for either of the output transistors is not specified at 2.24 amperes, but it can be extrapolated from the published curves of the 2N6488 and the 2N6491 transistors. At 2.2 amperes and 4 volts, the minimum current gain is 35; and the driver must be capable of delivering at least:

$$I_{c5} = \frac{I_{c7}}{h_{FEQ7}} = \frac{2.24}{35} = 64 \text{ mA}$$

Based on operation under the short-circuit condition, a maximum heatsink temperature of 60°C, the 1.3°C/W thermal resistance of a 0.002-mil mica washer, and a peak output current of 2.91 amperes, the safe operating area of the driver must be at least 118.6 milliamperes at 25 volts. The RCP702B and the RCP703B both satisfy this requirement easily.

Voltage-Gain and Common-Base Stages

The minimum gain of both drivers at 6 milliamperes and 4 volts is estimated at 30; and the drivers require a base current of:

$$I_{B5} = \frac{I_{c5}}{h_{FEQ5}} = \frac{64}{30} = 2.13 \text{ mA}$$

It is desirable that the class A stage preceding the driver stage be biased with a current much higher than the required 2.13 milliamperes minimum. This bias makes the stage more linear and helps to keep the distortion low. In addition, it improves the operation of the V_{BE} multiplier (Q_M). The higher the current through the V_{BE} multiplier, the easier it becomes to keep the bias voltage constant.[12] The idling current was chosen as 9.2 milliamperes; too high an idling current is undesirable because it causes higher quiescent dissipation. In the case under discussion, the quiescent dissipation is:

$$P = I_{c3} \times \frac{1}{2} V_{CC} = 9.2 \times 10^{-3} \times 25$$
$$= 230 \text{ mW}$$

which is well below the maximum rating for the RCP702B.

The RCA1A06 and RCA1A05 transistors were selected for the input common-base stage because of their excellent f_T and good low-current gain. The idling current was selected to be no less than 0.5 milliampere since the current gain of Q3 (RCP702B) is a minimum of twenty. The biasing network for the RCA1A06 and RCA1A05 was determined as follows:

$$V_{BB12} = V_{BE1} + V_{BE2} + I_{c1,\,c2}\,(R_{E1} + R_{E2})$$

The emitter resistors were chosen as 1.5 kilohms since 1.5 kilohms and 0.5 milliampere produce 750 millivolts, a voltage which is slightly greater than the V_{BE} of the RCP702B and RCP703B. V_{RE1} and V_{RE2} provide thermal compensation. Compensation is improved when the voltage across the emitter resistor is as high as possible. In the case under discussion, 0.75 volt was judged to be adequate. Therefore:

$$V_{BB12} = 0.65 + 0.65 + 0.5 \times 10^{-3} + (1.5 + 1.5) \times 10^{3} = 2.8 \text{ V}$$

The biasing network comprises diodes D1 and D2 and the 510-ohm resistor R_{BB1}. The diodes provide thermal compensation for V_{BB}. Ideally,a diode such as the RCAD1300A should be used because it is designed specifically for such applications. Moreover, diodes D1 and D2 should be thermally connected with transistors Q1 and Q2. The diodes used in

the 20-watt circuit do not match perfectly the temperature characteristic of the V_{BE} of Q1 and Q2, but, because of a very low power dissipation in Q1 and Q2, and a low total temperature change, they do provide adequate compensation.

The minimum gain of Q2 is estimated at 30 at 0.5 milliamperes, which means that the base current does not exceed 17 microamperes. The current through diodes D1 and D2 should be one hundred times that value. Thus, R_{BB} is determined as:

$$R_{BB12} = \frac{V_{BB12} - V_{D1} - V_{D2}}{100\, I_{B1}}$$

$$= \frac{2.8 - 0.7 - 0.7}{10^2 \times 17 \times 10^{-6}} = 823 \text{ ohms}$$

The final value of 510 ohms was selected, which yields:

$$I_{D1D2} = \frac{V_{BB12} - V_{D1} - V_{D2}}{R_{BB12}}$$

$$= \frac{2.8 - 0.7 - 0.7}{510} = 2.74 \text{ mA}$$

To obtain the 2.8-volt bias, the 30-volt (±15 volt) regulated supply voltage already used for the operational amplifier was utilized. A quantity R* is defined as:

$$R^* = R_{B1} + R_{B2}$$

$$R^* = \frac{V_{OAS} - V_{BB12}}{I_{D1D2}} = \frac{30 - 2.8}{2.74 \times 10^{-3}}$$

$$= 9{,}927 \text{ ohms}$$

and $R_{B1} = 4{,}963$ ohms. The closest standard value is 4.7 kilohms, with which the final value of I_{D1D2} is:

$$I_{D1D2} = \frac{V_{OAS} - V_{D1} - V_{D2}}{R_{B1} + R_{B2} + R_{BB12}}$$

$$= \frac{30 - 0.7 - 0.7}{4700 + 4700 + 510} = 2.88 \text{ mA}$$

$$V_{BB12} = V_{D1} + V_{D2} + I_{D1D2} \times R_{BB}$$

$$= 0.7 \times 0.7 \times 2.88 \times 10^{-3} \times 510$$

$$= 2.87 \text{ V}$$

Resistor R* is split into two equal resistors in order to set the signal input point (the junction of R_{E1} and R_{E2}) as close to the zero dc level as possible. With the resistor values selected as described, the idling current becomes:

$$I_{C1} = I_{C2} = \frac{V_{BB12} - V_{BE1} - V_{BE2}}{R_{E1} + R_{E2}}$$

$$= \frac{2.87 - 0.65 - 0.65}{1500 + 1500}$$

$$= 0.523 \text{ mA}$$

It should be noted that V_{D1} and V_{D2} are not exactly 0.7 volt and that resistor tolerances were not taken into account, but for all practical purposes, this calculation is accurate.

The emitter resistor of Q3 was chosen as 100 ohms, since this value yields a voltage drop that is close to V_{BE3} if I_{C3} is approximately 7.7 milliamperes.

$$V_{RC1} = R_{E3} \times I_{C3} + V_{BE3}$$

A typical V_{BE3} at 10 milliamperes is 0.65 volt and:

$$V_{RC1} = 100 \times 7.7 \times 10^{-3} + 0.65 = 1.42 \text{ V}$$

$$R_{C1} = \frac{V_{RC1}}{I_{C1}} = \frac{1.42}{0.523 \times 10^{-3}}$$

$$= 2715 \text{ ohms}$$

To ensure a sufficiently high idling current in Q3, a 3-kilohm resistor was selected for the value of R_{C1}; and

$$I_{C3} = \frac{I_{C1} R_{C1} - V_{BE3}}{R_{E3}}$$

$$= \frac{0.523 \times 10^{-3} \times 3 \times 10^{3} - 0.65}{100}$$

$$= 9.19 \text{ mA}$$

The value of emitter resistor R_{E3} could have been made much lower, but was intentionally selected as 100 ohms. When the amplifier is driven near or into clipping, the V_{CE} across the output transistor is very low, and in some cases the output transistor will saturate. This condition must be avoided if the best possible high-frequency performance for the given output transistor is to be achieved.

Whenever the collector-to-emitter voltage is low and the transistor is

conducting fairly high current, the base-widening effect[13] takes place. This effect not only lowers the current gain of the transistor, but also drastically reduces its f_T. Base widening can cause instability in the amplifier, and even if it is properly designed and optimized for maximum efficiency, the best possible performance at high frequencies will not be achieved. The base-widening effect is not very pronounced in low-voltage transistors, but in high-voltage devices used in high-power amplifiers, it becomes very important. The lowest possible collector-to-emitter voltage of the output device in the 20-watt circuit is:

$$I_{C7} = 2.5 \text{ A}$$

$$V_{CE7}(\text{min}) = I_{C3} \times R_{E3} + V_{CE(sat)3} + V_{BE5} + V_{BE7}$$

$$V_{CE7}(\text{min}) = 9.2 \times 10^{-3} \times 100 + 0.1 + 0.7 + 0.8 = 2.52 \text{ V}$$

At the same time:

$$V_{RE7} = I_{C7} \times R_{E7} = 2.5 \times 0.41 = 1.025 \text{ V}$$

The minimum supply voltage is:

$$1/2\ V_{SS} = E_m + V_{CE7} + V_{RE7} = 17.9 + 2.52 + 1.025 = 21.45 \text{ V}$$

For the best possible performance and for a conservative power output rating, a ±25-volt supply is used.

V_{BE} Multiplier

The quiescent current in the amplifier is set by the V_{BE} multiplier stage.[12] The RCP701A, Q_M, was selected for the V_{BE} multiplier because it has a minimum current gain of 50 at 7 milliamperes. With an idling current of 200 milliamperes, the voltage drop across R_{BB78} is:

$$V_{RBB78} = V_{BE7} + V_{BE8} + I_{C7} \times R_{E7} + I_{C8} \times R_{E8}$$

$$V_{RBB78} = 0.6 + 0.6 + 0.15 \times 0.41 + 0.15 + 0.41 = 1.323 \text{ V}$$

$$I_{B8} = \frac{I_{C8}}{h_{FEQ8}} = \frac{0.15}{50} = 3 \text{ mA}, \quad I_{B8} = I_{B7}$$

$$I_{C5} = I_{B7} + I_{RBB78} = I_{B7} + \frac{V_{RBB78}}{R_{BB78}}$$

$$I_{C5} = \frac{1.323}{75} + 0.003 = 20.6 \text{ mA}$$

This is the idling current of the driver, Q5, in the worst case. The h_{FE} of the driver at 20.6 milliamperes is approximately 30, the worst case, and:

$$I_{B5} = \frac{I_{C5}}{h_{FEQ5}} = \frac{20.6}{30} = 0.69 \text{ mA}$$

The idling current of Q3 (9.2 milliamperes), should, therefore, be adequate. The total voltage of the V_{BE} multiplier is:

$$V_M = V_{RBB78} + V_{BE5} + V_{BE6} = 1.323 + 0.65 + 0.65 = 2.62 \text{ V}$$

With the gain of 50 at 7 milliamperes, the base current of the V_{BE} multiplier in the worst case is:

$$I_{BM} = \frac{I_M}{h_{FEQM}} = \frac{I_{C3} - I_{B5} - I_{RS}}{h_{FEQM}}$$

The current through the shunt resistors, with R_{RS}(typical) = 3410 ohms is:

$$I_{RS} = \frac{V_M}{R_{RS}} = \frac{2.62}{3410} = 0.77 \text{ mA}$$

and:

$$I_{CM} = I_{C3} - I_{B5} - I_{RS} = 9.2 - 0.69 - 0.77 = 7.74 \text{ mA}$$

then:

$$I_{BM} = \frac{7.74}{50} = 0.155 \text{ mA}$$

With the above values employed, the V_{BE} multiplier provides adequate compensation.

Voltage and Current Gain Calculations

The current gain of the emitter follower is:[14]

$$A_I = 1 + h_{fe}$$

For the low-frequency signal it can be approximated as:

$$A_I \simeq h_{FE}$$

Therefore,

$$A_{I8} = 35 \text{ and } A_{I6} = 30$$

with the lowest gain transistors in the circuit. The total current gain of the driver and the output stage is:

$$A_{I68} = A_{I6}\ A_{I8} = 35 \times 30 = 1050$$

The input impedance of the common-collector stage is given by:

$$R_i = h_{ie} + (1 + h_{fe})\ R_L$$

In a real situation, since $h_{fe} >> 1$ and $h_{ie} << h_{fe} R_L$ for the low frequency signal:

$$R_i \cong h_{FE}\ R_L$$

Therefore:

$$R_{i6} = h_{FEQ6}\ h_{FEQ8}\ R_L$$

$$R_{i6} = 30 \times 35 \times 8 = 8{,}400 \text{ ohms}$$

if R_E is neglected. This is the effective load resistance appearing in the collector of Q4. The voltage gain of the Q4 stage is:

$$A_{V4} = \frac{A_{I4}\ R_{L4}}{R_{i4}} \simeq \frac{h_{FE}\ R_{i6}}{h_{FE}\ R_{E4}} = \frac{R_{i6}}{R_{E4}} = \frac{8400}{100}$$

The common-base stage at the input of the discrete section has a voltage gain of:

$$AV = \frac{\Delta I_C\ R_L'}{\Delta I_E\ R_E} \quad \Delta I_C \simeq \Delta I_E$$

with:

$$R_{L2} = \frac{R_{C2}\ R_{i4}}{R_{C2} + R_{i4}} \frac{3000 \times 3000}{6000}$$

$$= 1500 \text{ ohms}$$

$$A_{V2} = \frac{R_{L2}}{R_{E2}} = \frac{1500}{1500} = 1$$

The total voltage gain of the discrete section is:

$$A_{VD} = A_{V2}\ A_{V4}\ A_{V6}\ A_{V8} \simeq 84$$

$$= 38.5 \text{ dB}$$

A_{V6} and A_{V8} are approximately 1 since, for the common collector:

$$A_V = A_I \frac{R_E}{R_i}$$

With the local feedback loop closed ($R_{FD} - R_{SD}$), the gain of the discrete section is:

$$A_{VDC} \approx \frac{R_{FD}}{R_{SD}} = \frac{12000}{2000} = 6 = 15.6 \text{ dB}$$

This result also indicates that the local loop gain is:

$$A_{LL} = \frac{A_{VD}}{A_{VDC}} = \frac{84}{6} = 14 = 22.9 \text{ dB}$$

The above calculation is correct if all transistors have minimum gain. In a typical case, the gain is expected to be:

$$A_{VD} = A_{V2} \times A_{V4} = \frac{R_{L2}}{R_{E2}} \times \frac{R_{i6}}{R_{E4}}$$

$$= \frac{\frac{R_{C2}\ R_{i4}}{R_{C2} + R_{i4}}}{R_{E2}} \frac{h_{FEQ6}\ h_{FEQ8}\ R_L}{R_{E4}}$$

$$= \frac{\frac{3 \times 10^3 \times 60 \times 100}{3 \times 10^3 + 60 \times 100}}{1500} \frac{60 \times 100 \times 8}{100}$$

$$= 640 = 56 \text{ dB}$$

The maximum voltage required to drive the discrete amplifier section to the full power output is given by:

$$E_p in = E_m \times \frac{R_{SD}}{R_{FD}} = 17.9 \times \frac{2000}{12000}$$

$$= 2.98 \text{ V}$$

Assuming that the open-loop gain of the discrete section is high enough and that the error voltage is small, $E_p(in)$ is the maximum voltage that the integrated circuit must provide into a 2000-ohm resistor. This is a very light load for the CA3140B, and the best possible performance is obtained.

Frequency and Phase Response

With the CA3140B IC connected, the open-loop gain is greater than 100 dB at low frequencies. The closed-

loop gain depends upon the desired sensitivity and performance of the amplifier. It is obvious that, in the discrete section, the output transistors, with their associated capacitance and output impedance, will have the lowest corner frequency. The roll-off frequencies of the other stages are much higher, and it is expected that the whole discrete section will have a uniform 6-dB-per-octave roll-off. By measuring the open-loop frequency response and the loop frequency response, a graphical analysis necessary or needed for the completion of the design can be made.

As mentioned above, the open-loop −3-dB point for the discrete section is at 3.4 kHz, while the loop −3-dB point is at 166 kHz. A typical open-loop gain was measured at 51 dB, which is not far from the predicted 56 dB. The closed-loop gain was measured as 15.33 dB versus 15.56 dB calculated.

A typical open-loop gain of the CA3140B is 100 dB at low frequencies. The only corner frequency is at 50 Hz, and the roll-off is a uniform 6-dB-per-octave roll-off. Unity gain in a typical case is at 5 MHz and, in the worst case, at 4.5 MHz. With all of the above information available, a graphical analysis and overall feedback selection can be made, as shown in Fig. 6.

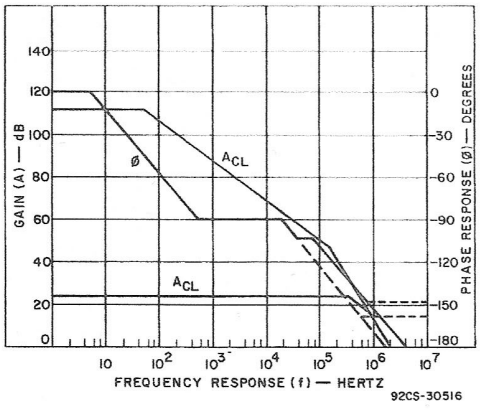

Fig. 6— Predicted open-loop and closed-loop frequency response (A_{OL}, A_{CL}) and phase response (ϕ).

With the integrated circuit and discrete sections connected, the total gain of the circuit is 115.5 dB from dc to 50 Hz. At 50 Hz, the gain starts to decrease at 6 dB per octave (20 dB per decade) until it reaches a second corner frequency at 166 kHz (the first corner frequency of the discrete section). From 166 kHz, the slope is 12 dB per octave all the way to unity gain at approximately 2 MHz. Selection of the overall feedback depends upon the maximum allowed distortion. The open-loop distortion at 20 kHz and 15 watts was approximately 4 percent; to bring it down to approximately 0.04 percent, 40 dB of negative feedback is necessary. The allowable distortion, then, determines the ratio of resistors R_F and R_S. If R_F is selected to be 3000 ohms, the value for R_S is calculated as follows:

Open-loop gain A_{OL} = 64 dB

Loop gain A_L = 40 dB

Closed-loop gain A_{CL} = 24 dB

$$A_{CL}\,(dB) = 20 \log A_{CL}$$

$$24 = 20 \log A_{CL}$$

$$A_{CL} = 10^{1.2} = 15.85$$

$$R_S = \frac{R_F}{A_{CL}} = \frac{3000}{15.85} = 189 \text{ ohms}$$

A standard value of 200 ohms was selected, yielding a final closed-loop gain of 23.5 dB.

The sensitivity of the amplifier must now be checked; it is given by:

$$E_{in} = \frac{R_S}{R_F} E_{out}$$

where E_{out} is the voltage at rated power. This equation yields:

$$E_{in} = \frac{200}{3000} \times 12.65 = 0.843 \text{ V}$$

which is acceptable.

It should be noted that the design can be based on the desired sensitivity. In such a case, the open-loop gain and with it the loop gain must be adjusted to provide the desired distortion. At this point, the horizontal line at 23.5 dB denoting the closed-loop gain is drawn into the diagram shown in Fig. 2.

In determining the desirable roll-off frequency of the overall feedback,

the phase diagram is plotted first. It is obvious that the higher the roll-off frequency, the better the rise time, slew rate, and bandwidth. As a general approximation, the phase starts to change at one tenth of the corner frequency, reaches 45° at the corner frequency, and 90° at ten times the corner frequency. Therefore, the phase change (lag) starts at 5 Hz and reaches −90° at 500 Hz. The next corner frequency is at 166 kHz, which keeps the phase at −90° until a frequency of 16.6 kHz is reached. Above 16.6 kHz, the phase changes toward −180°, which it would reach at 1.66 MHz if there were no frequency compensation in the feedback loop. Even in such a case, the amplifier would be stable, with a phase margin of approximately 22°.

The phase margin is the difference between 180° and the actual phase at zero-dB loop gain. Obviously, the amplifier is stable only if the phase is less than 180° at zero-dB loop gain. The gain margin is given by the actual loop gain at 180° phase. In a case without frequency compensation, the gain margin would be −130 dB and the amplifier would be stable. To improve the phase margin, a frequency of 350 kHz was selected as the corner frequency of the feedback network. This selection puts the A_{CL} line roll-off above 350 kHz at 6 dB per octave until it intersects the A_{OL} line. The expected closed-loop response of the amplifier is given by the A_{OL} line. The phase response is now modified because at 35 kHz the feedback network starts to change the direction of the phase change, and the resultant phase is a constant one at −104°. The change of the phase begins again at approximately 80 kHz since the unity loop gain occurs at 800 kHz, and the slope at the A_{CL} line is 12 dB per octave from that point on. The phase margin is now approximately 31.5°, which is not a large margin, but enough to guarantee stable operation.

The feedback capacitance can be calculated as:

$$C_F = \frac{1}{2\pi f\ R_F} = \frac{1}{2\pi \times 350 \times 10^3 \times 3 \times 10^3} = 151\ \text{pF}$$

A standard value of 150 picofarads was selected.

The Measurement of Frequency and Phase Response

The frequency response and the phase response of the loop were measured, as shown in Fig. 7. The feedback loop was broken and the dc feedback was provided by the 22-kilohm resistors while the ac feedback was eliminated by use of the 10-microfarad capacitor. The dc feedback is essential to prevent the output of the amplifier from shifting away from zero dc level. Without the dc feedback, and because of the very high dc gain of the integrated circuit, the smallest error voltage would keep the output of the amplifier at either supply voltage.

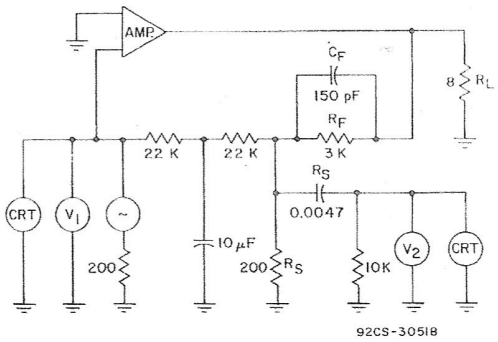

Fig. 7—Circuit used to measure frequency and phase response (A_L and ϕ_L).

The component values in the RC network involved in providing the dc feedback were chosen so that the measurement frequency would be much higher than the corner frequency of the low-pass filter. The input signal was injected through a 200-ohm resistor to simulate identical operating conditions. The output signal was monitored across another 200-ohm resistor, R_S. The output voltage was measured after the high-pass filter (3.4 kHz corner frequency). The filter makes the measurement easier by eliminating most of the low-frequency noise.

The loop gain is given by the ratio of V_2 over V_1. The input voltage V_1 is kept constant and the frequency increased until the gain is less than unity. The phase response can be

measured in several different ways; the most simple is by use of an oscilloscope. There are two oscilloscope methods: one involves the use of horizontal plates as one input and vertical plates as the other input; the second method makes use of a dual input in the alternating mode. In both cases, it is essential to assure that neither the oscilloscope nor the connecting leads cause any errors, which can easily occur at frequencies as high as 1 MHz.

It is important to note that the loop gain and the phase response should be checked at various output levels. The actual behavior of the amplifier will be very close to that predicted if all transistors are operating within their linear range, that is, on the flat portion of the gain-bandwidth curve. Fig. 8 shows the actual loop-gain

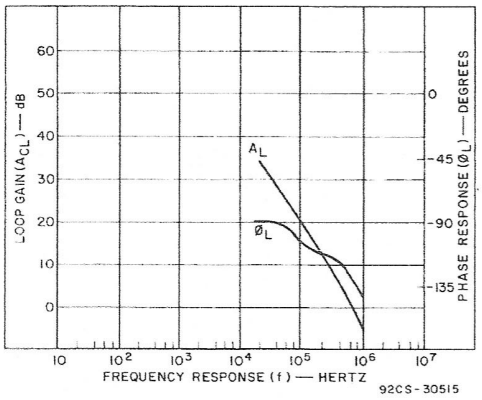

Fig. 8—Measured loop gain, A_L, and phase response, ϕ_L.

response and the phase response of the 20-watt amplifier.

Measurements were begun at 20 kHz because of the noise and the high gain limitations at low frequencies. The low-frequency portion of the curve is not important since it is known that the operational amplifier will guarantee stability and that the phase shift can occur only between zero and −90°, as there are no additional low-frequency nodes in the circuit. The loop response shows unity gain at 660 kHz and a phase margin of approximately 40°; these results are better than predicted. The gain margin is difficult to measure by increasing the frequency until the phase reaches 180°; consequently an alternate method was used. The value of R_F, Fig. 7, was lowered to 750 ohms and C_F changed to 610 picofarads to keep the 350-kHz corner frequency unchanged. The changes in R_F and C_F increased the overall feedback by 12 dB. Since there were no signs of instability in the square-wave response of the amplifier, it was determined that the gain margin was well above 12 dB. Typical performance curves for the 20-watt amplifier are shown in Figs. 9 through 16; typical performance characteristics are listed in Table I.

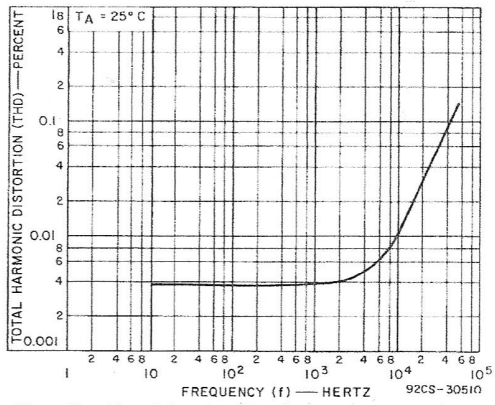

Fig. 9—Total harmonic distortion as a function of frequency at 20 watts.

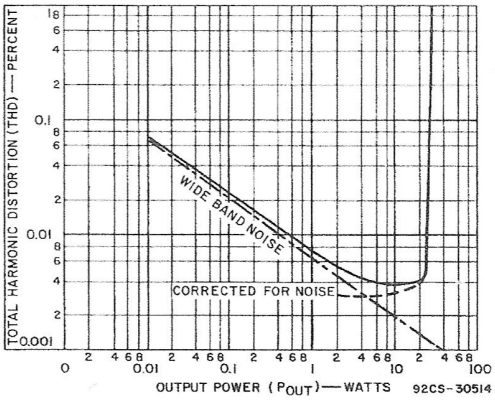

Fig. 10—Total harmonic distortion at 1 kHz as a function of power.

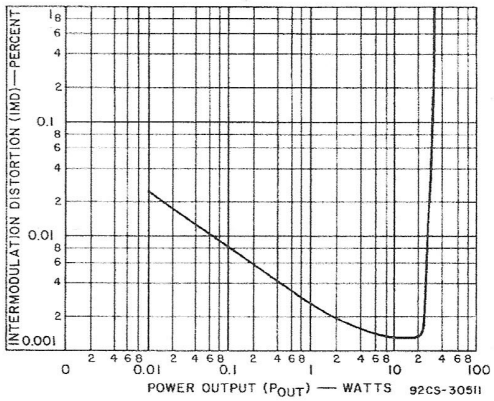

Fig. 11—Intermodulation distortion, 60 Hz and 7 kHz, 4:1, as a function of power.

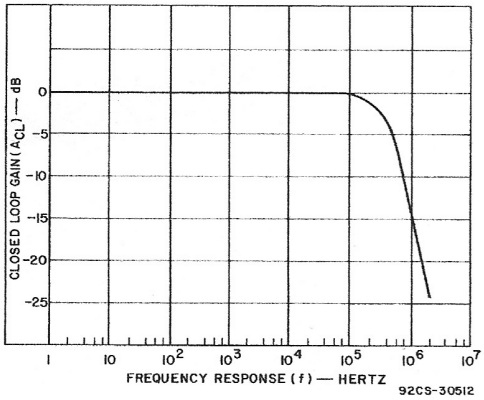

Fig. 12— 1-watt frequency response.

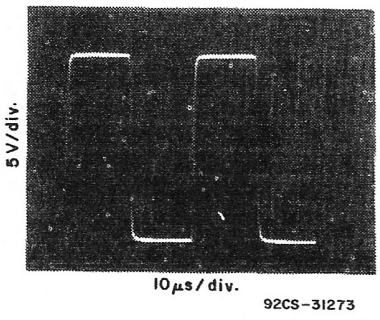

Fig. 13— 20-kHz square-wave response, 8-ohm resistive load, 35 V_{p-p}, 49 microsecond period.

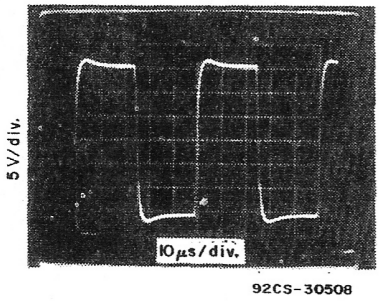

Fig. 14— 20 kHz square-wave response, 8-ohm, 60° capacitive load.

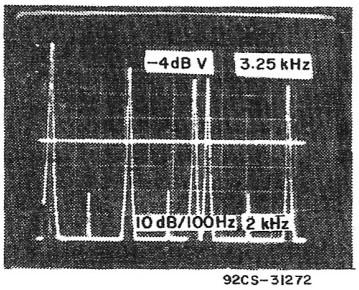

Fig. 15— Transient intermodulation distortion, input sine wave 15 kHz, square wave 3.25 kHz, no distortion at – 66 dB.

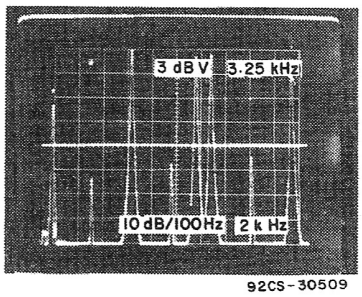

Fig. 16— Transient intermodulation distortion, 3.25 kHz notched out with twin-T filter, no distortion at – 77 dB.

TABLE I – Typical Performance

Rated Power:	**20 watts into an 8-ohm load**
Sensitivity:	**0.85 volt at 20 watts**
Input Impedance:	**10 kilohms**
Hum and Noise Below Rated Power Output	
Open Input:	**97 dB**
Shorted Input:	**98 dB**
Phase Shift:	**0 or 1.5° at 20 Hz, –5° at 20 kHz**
Slew Rate:	**30V/μs or 1.36 V/μs/V**
Rise Time:	**1.3 microseconds**
Damping Factor:	**280**

Acknowledgment

The author thanks C.B. Leuthauser, L.A. Kaplan and C.F. Wheatley for their helpful suggestions and comments.

References

1. Daugherty, D.G., "Design Considerations for Linear Transistor Audio Power Amplifiers", Ph. D. Dissertation, Univ. of Wisconsin, 1964.

2. Daugherty, D.G. and Greiner, R.A., "Some Design Objectives for Audio Power Amplifiers", IEEE Transactions on Audio and Electroacoustics Vol. AU-14 No. 1, March 1966.

3. Otala, M., "Circuit Design Modifications for Minimizing TIM Distortion in Audio Amplifiers", Journal of the AES, Vol, 20 No. 5, June 1972.