A High-Level Single-Sideband Transmitter*

OSWALD G. VILLARD, JR.†, ASSOCIATE, IRE

Summary—Double-sideband suppressed-carrier signals can be generated with relatively high efficiency at high power levels in grid-modulated balanced modulators so biased that very little plate current flows in the absence of audio input. When tubes having a constant-current characteristic are used, and the necessary quadrature phase shift between pairs of audio- and radio-frequency excitation voltages exists, two high-power balanced modulators of this type may be connected to a common output circuit to produce suppressed-carrier single-sideband signals. The peak efficiency of the combination for sinusoidal modulation is $\pi/4$ times or 78 per cent that obtainable when the same tubes are used as linear amplifiers. Operating adjustments are shown to be straightforward. The simplicity and power economy of this circuit make it attractive for those applications where moderately low distortion and reasonably good rejection of the undesired sideband are satisfactory.

THE ADVANTAGES of single-sideband radiotelephone transmission in terms of bandwidth and power requirements have been understood for many years. However, difficulties of single-sideband generation and reception have so far prevented the general use of this system except in cases of absolute necessity, such as radiotelephony below 100 kc, carrier-current communication, or long-distance multichannel short-wave telephone transmission.

The classical method of producing single-sideband signals, employing sharp filters, balanced modulators, and frequency changers, is undeniably complex, and does not lend itself readily to incorporation in radiotelephone equipment where flexibility in operation, as well as ease in adjustment and maintenance, are important.

The phase-rotation method of single-sideband generation, in which two sets of carrierless double sidebands are combined in such a way that the undesired upper or lower sideband is cancelled out, is fundamentally simpler (now that practical 90-degree audio phase-shift networks are available[1]), but in the systems so far disclosed[2–4] linear amplification is required in order to achieve the desired output power level.

It is the purpose of this paper to describe a transmitter of the phase-rotation type in which the single sideband is generated at high level and good efficiency directly in the final stage. The rf circuitry and circuit adjustments have been reduced to a minimum. The result is a transmitter which is scarcely more complicated or more difficult to adjust than the conventional plate-modulated one, which produces its output at an efficiency closely approaching that of a linear amplifier, and which gives reasonably good suppression of the undesired sideband together with moderately low distortion.

A block diagram of the basic method is shown in Fig. 1. Two balanced modulators, excited by rf voltages 90° out of phase, and by af voltages in phase quadrature, are connected directly to a common output circuit. Four tubes having a substantially constant-current characteristic are used, so that the amplitude and phase of their plate-current pulses is determined only by excitation and modulation voltages, and is independent of the loading and the tuning of the output circuit. The tubes are so

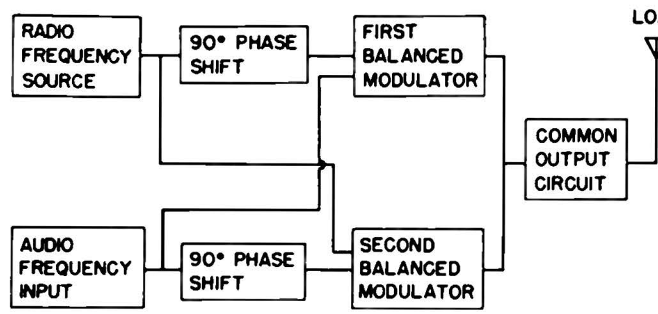

Fig. 1—Block diagram of sideband-cancellation method.

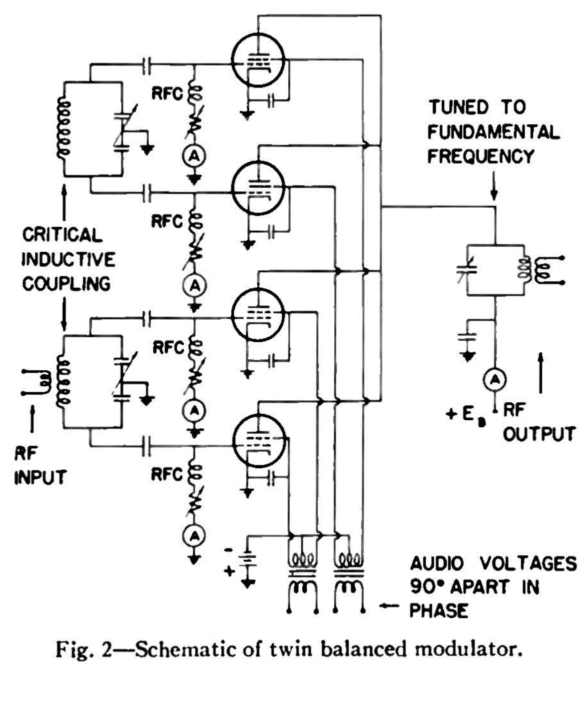

Fig. 2—Schematic of twin balanced modulator.

* Decimal classification: R423.52. Original manuscript received by the Institute, April 6, 1948; revised manuscript received, June 28, 1948.

† Stanford University, Stanford, Calif.

[1] R. B. Dome, "Wideband phase shift networks," *Electronics*, vol. 19, pp. 112–115; December, 1946.

[2] Paul Loyet, "Experimental polyphase broadcasting," Proc. I.R.E., vol. 30, pp. 213–222; May, 1942.

[3] M. A. Honnell, "Single sideband generator," *Electronics*, vol. 18, pp. 166–168; November, 1945.

[4] B. E. Lenehan, "A new single sideband carrier system for power lines," *Elec. Eng.*, vol. 66, pp. 549–592; June, 1947.

biased that very little plate current is drawn in the absence of an audio signal. The combination might be called a "twin balanced modulator." The schematic of a practical circuit is shown in Fig. 2.

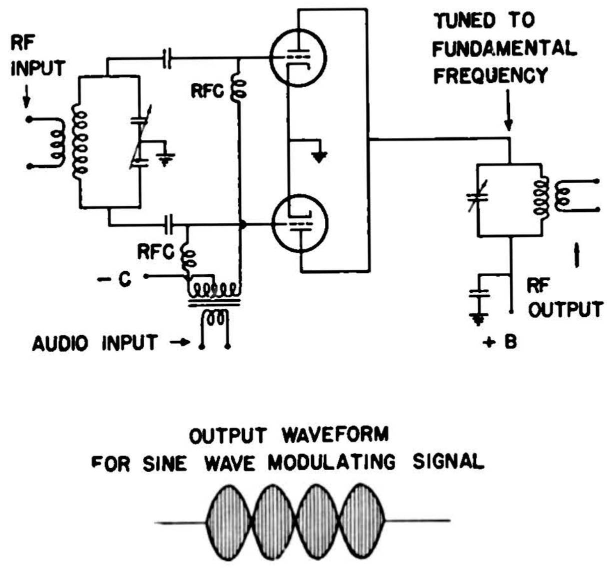

Fig. 3—Schematic of high-level, grid-bias-modulated balanced modulator.

The operation of the circuit may be more easily visualized after a review of the mathematical and vector relationships involved. The equation of the output of an ordinary balanced modulator (see Fig. 3) may be written as:

$$e = m \sin 2\pi ft \cos 2\pi f_a t \qquad (1)$$

where

e = instantaneous amplitude of the wave
m = a constant proportional to the amplitude of the modulating wave, i.e., degree of modulation
f_a = modulating (or audio) frequency
f = radio (or carrier) frequency.

This equation may be thought of as representing a wave of carrier frequency whose instantaneous amplitude is proportional to the cosine of the angle of the modulating wave. When this angle is in the second and third quadrants, its cosine is negative, and the phase of the wave of carrier frequency changes by 180°. We therefore have an rf wave which starts out from zero with given rf phase, builds up sinusoidally to a maximum, and then returns to zero again in the first half of the af cycle. This process is repeated in the second half, but the phase of the rf wave is reversed.

Cancellation of the undesired sideband in a double balanced modulator may be shown as follows. Equation (1) can be expanded to:

$$e = \frac{m}{2}\left[\sin 2\pi(f + f_a)t + \sin 2\pi(f - f_a)t\right]. \qquad (2)$$

Now the output of a second balanced modulator, fed by a rf voltage 90° out of phase with that of the first, as well as an af voltage in quadrature with that of the first, has the form:

$$e = m \cos 2\pi ft \sin 2\pi f_a t \qquad (3)$$

which, when expanded, is

$$e = \frac{m}{2}\left[\sin 2\pi(f + f_a)t - \sin 2\pi(f - f_a)t\right]. \qquad (4)$$

It will be seen that, when (2) and (4) are added, the difference-frequency sidebands cancel, while the sum-frequency sidebands add. The opposite result may be obtained by reversing the phase of either the rf voltage or the af voltage fed to one of the balanced modulators.

The two voltages of (1) and (3) (the rf outputs of the two balanced modulators) may be represented vectorially as in Fig. 4. The dotted vectors E_A and E_B represent

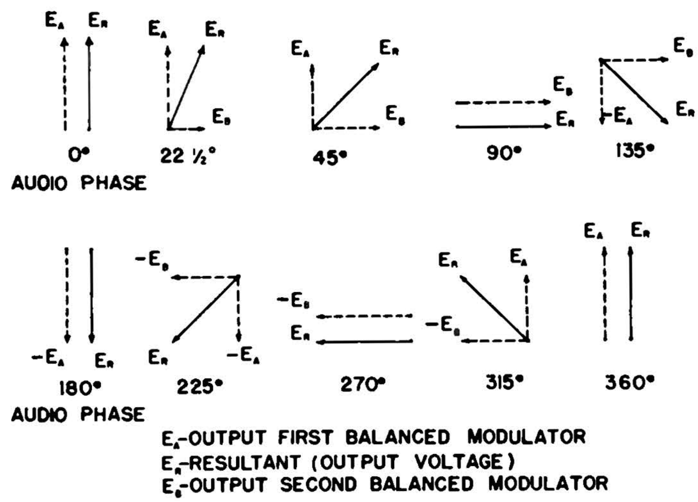

Fig. 4—Vector diagram of twin balanced-modulator action (sine-wave modulating voltage).

the magnitude and phase of the individual balanced-modulator output currents as viewed at intervals during one sinusoidal audio cycle. These two currents are in quadrature, and their amplitudes vary in accordance with af modulating voltages 90° apart. The heavy solid line (E_R) represents their vector sum; i.e., the voltage developed across the common tank circuit of the twin balanced modulator, assuming the tubes to be constant-current generators. It will be seen that the resultant vector maintains a constant amplitude while making one complete revolution for each cycle of audio frequency. The resultant is a new radio frequency different from that of the carrier—the upper, or the lower sideband, as the case may be.

The efficiency obtainable with a twin balanced modulator of this type can readily be determined by examining the efficiency of a high-level balanced modulator by itself. (Throughout this paper it will be assumed that the modulating voltage is a single-frequency wave.) Consider the circuit of Fig. 3. It is assumed that a steady rf input signal is applied, and that the tubes are so biased that no plate current flows in the absence of modulation. In the course of one cycle of modulation, the tubes alternately deliver their maximum power output to the common tank circuit, under conditions exactly the same as those obtaining at the positive crest of the modulating cycle in an ordinary grid-modulated amplifier. Let the peak efficiency at this point in the cycle be called η. The envelope of the balanced-modulator output is shown in Fig. 3. The average rf output is exactly 0.5 times the peak output power P_m, since the output voltage has the form of a series of half sine waves, and the average of a series of half sine waves squared is one-half the peak value. The peak dc input power P_{in} equals the peak output power P_m divided by the peak efficiency, namely, P_m/η. If the dc plate voltage is E_b, the peak dc plate current must be $P_m/\eta E_b$. Now the waveform of the dc input current is exactly the same as that of the rf output envelope—a series of half sine waves whose average is $2/\pi$ or 0.636 times their peak value. Therefore, the average dc input current is $2/\pi$ times the peak current, or $2P_m/\pi\eta E_b$. The efficiency, or average output power divided by average input power, is then $0.5\,P_m/(2P_mE_b/\pi\eta E_b)=\pi\eta/4$. For $\eta=0.66$, the efficiency is 52 per cent.

This is, of course, only true at full output, corresponding to complete modulation with a single-frequency tone. As with any grid-modulated or linear amplifier, the efficiency of a single or twin balanced modulator for lesser amplitudes is proportional to the amplitude of the modulating signal.

It is interesting to note that the efficiency of the twin balanced modulator is exactly equal to that of one high-level balanced modulator by itself. Since the dc plate current drawn by the tubes is independent of the rf plate voltage in the common tank circuit, the input power drawn by the two balanced modulators is exactly twice that drawn by one. The combined average output power, on the other hand, is equal to the peak power developed by each component balanced modulator (as may be seen in Fig. 4), or twice the average power output of each balanced modulator alone. Therefore, the efficiency of the pair is equal to that of its components, or 0.785η.

Since the circuit is symmetrical, each tube dissipates on its plate one-fourth the total power lost. This means that, for full utilization of tube plate dissipation capacity, somewhat higher operating voltages may be used than are customary for a single tube in class-C amplifier service.

Screen-grid modulation, with this circuit, offers the greatest ease in adjustment, and is therefore to be preferred. Where the saving in audio power or voltage required for screen modulation outweighs a considerable increase in the complexity of adjustment, control-grid modulation may be used. Plate modulation is not satisfactory, because at one point in the audio cycle the two tubes must operate with equal dc plate voltages of 0.707 of maximum, yet must produce a combined rf output of 1.0 times the maximum. To support this output, the instantaneous voltage at the tube plates would have to swing negative during the rf cycle at a time when plate current would normally be flowing, even when the various phase relationships are taken into account. Suppressor-grid modulation will also be found to be unsatisfactory; although it can be used, screen current and dissipation are high when the suppressor grid is biased to cutoff, thus preventing full utilization of tube capacity.

The steady dc grid-current flow, which is present with screen modulation, greatly facilitates setting the 90° phase shift between rf voltages applied to the two component balanced modulators of the twin balanced modulator. The two grid tuned circuits of the two pairs of tubes in Fig. 2 may be inductively coupled together, and excitation fed to one. When the secondary circuit is tuned to resonance, (as evidenced by maximum grid current) a 90° phase shift exists between the voltage across this circuit and that of the primary. By adjusting for critical coupling, the magnitudes of these voltages are made equal, so that all four tubes receive the same excitation. Somewhat less than the normal class-C grid excitation may be applied, the exact amount depending on the operating conditions chosen.

The screen grids must be given a slight negative bias in order to prevent excessive plate-current flow in the absence of an audio signal. The exact amount is dependent on the tube characteristics, the rf drive applied to the control grids of the tubes, and on the plate voltage. The correct value is not especially critical and may readily be found in practice. A no-signal plate current of roughly one-tenth the full-output plate current is generally satisfactory.

The question of sideband suppression, and the various circuit adjustments necessary to maintain it, are considered next. First, it should be observed that high-level balanced modulators are inherently "balanced," insofar as carrier suppression is concerned. Since the tubes are biased nearly to cutoff, virtually nothing will be radiated until an audio signal is applied. In cases where complete and automatic suppression of the carrier during quiet intervals is desired, this feature is very convenient; however, it is obvious that, if the circuit is not properly adjusted, the carrier will reappear along with any sideband output.

Suppression of the undesired sideband, in any sideband cancellation system, is dependent on the accuracy of setting of both the rf and the af phases. Since this suppression depends on the difference between large quantities, it will always be difficult to maintain completely in practice. The situation for a misadjustment of either rf or af phase by itself is illustrated in Fig. 5. The voltages e_1 and e_2 represent the two upper or the two lower sidebands, whichever are to be cancelled. If it is assumed that these voltages are equal in magnitude, R, their resultant, equals $2e_1 \sin \alpha/2$ where α is the angle by which either the rf or the af phase differs from 90°. The ratio of the desired sideband (very nearly $2e_1$) to this undesired resultant is approximately $2e_1/2e_1 \sin \alpha/2$, or $115/\alpha$, where α is small and is expressed in degrees. (α may, of course, be considered to be the algebraic sum of rf and af phase deviations.) It is seen that when $\alpha=1°$, the ratio of desired to undesired sideband is

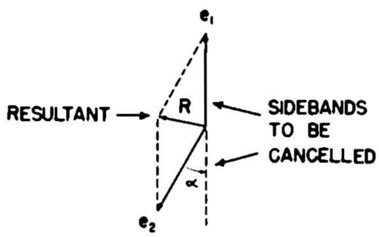

Fig. 5—Effect of rf or af phase angle on sideband cancellation.

roughly 40 db. This ratio drops to 20 db when α becomes approximately 10°. The necessity for close control of phase is clear.

A simple audio phase-shift network of the sort shown in Fig. 6 (taken from footnote reference 1) provides a

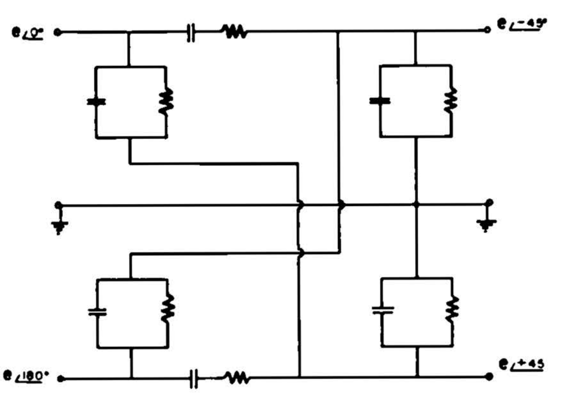

Fig. 6—Sample 90° wideband audio phase shift network (as shown in footnote reference 1).

phase shift departing from 90° by no more than 5° over the range 200–3000 cps. Over appreciable portions of the range, the phase characteristic is within 2°.

By way of comparison, assuming the phase settings to be perfect, if one of the sidebands to be cancelled is 11 per cent larger than the other, the ratio of desired to undesired sideband is nearly 25 db. Thus, a given inaccuracy in phase setting is slightly more serious than a given inaccuracy in relative magnitude.

Presence of any sizeable amount of the undesired sideband shows up as a distortion of the envelope of the combined output, when the audio input is a sine wave. The undesired frequency component in the output combines with the desired sideband and causes the envelope of the combination to have a "modulation," or ripple whose frequency is twice that of the audio modulating frequency. Small inaccuracies in either magnitude or phase setting (af or rf) produce ripples nearly identical in appearance, since in each case the effect of the misadjustment is to permit some undesired sideband to appear as a result of incomplete cancellation.

This ripple may conveniently be used to find the correct rf phase setting. Assuming the audio phase shift to be correct, the operator must first equalize the individual balanced-modulator output amplitudes, which may be done by adjusting the relative gain of the two af input channels. One balanced modulator should be operated at a time, and its output may conveniently be measured by means of a meter or oscilloscope. The two are set to produce rf outputs of equal magnitude for a given common modulating signal. The rf phase may then be adjusted until the ripple on the output envelope disappears.

This indication is quite sensitive, as may be seen with the aid of Fig. 7. Here E_A and E_B represent the instan-

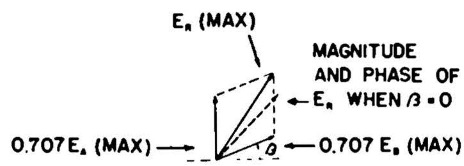

Fig. 7—Vector diagram illustrating sideband envelope modulation due to error in radio-frequency phase setting.

taneous rf outputs of each pair of balanced modulators. It may be shown that for a given radio-frequency phase error β, the resultant of these two voltages $E_{R(\max)}$ has its greatest value at the point in the audio cycle at which the two modulator outputs are equal, and 0.707 times the maximum; i.e., the condition illustrated. Since the angle β represents the deviation of the radio-frequency phase from quadrature, it follows that at this point in the audio cycle

$$
\begin{aligned}
E_{R(\max)} &= 2\cdot 0.707\cdot E_A \cos 1/2(90° - \beta) \\
&= E_{A(\max)}(\cos \beta/2 + \sin \beta/2)
\end{aligned}
\qquad (5)
$$

where $E_{A(\max)}$ represents the output of one balanced modulator at the crest of the audio cycle. The percentage ripple superimposed on the twin-balanced-modulator output envelope may be defined as follows:

$$
\text{percentage of ripple} = \frac{E_{R(\max)} - E_{A(\max)}}{E_{A(\max)}} \times 100. \qquad (6)
$$

$E_{A(\max)}$, the peak output of one balanced modulator alone, is also the amplitude of the single-sideband output in the absence of any ripple. The percentage of ripple then equals

$$\text{percentage of ripple} = 100(\cos \beta/2 + \sin \beta/2 - 1) \quad (7)$$
$$\cong 100 \sin \beta/2$$
$$\cong 0.8\beta,$$

where β is small and expressed in degrees.

Thus a 10° inaccuracy in phase setting results in an 8 per cent ripple "modulation" of the single-sideband envelope, which is quite readily visible on an oscilloscope.

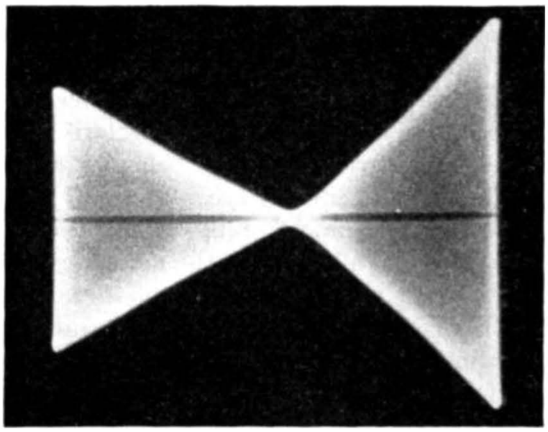

Fig. 8—Trapezoidal pattern: one balanced modulator, balancing resistors improperly set.

When highest accuracy of setting is desired, it is best to observe the output envelope with an oscilloscope, "touching up" the grid tuning (i.e., the rf phase setting) until the ripple is at a minimum. Alternatively, a simple rectifier connected to a pair of earphones may be excited from the output: correct phase setting will result in a disappearance of the audible tone caused by the ripple.

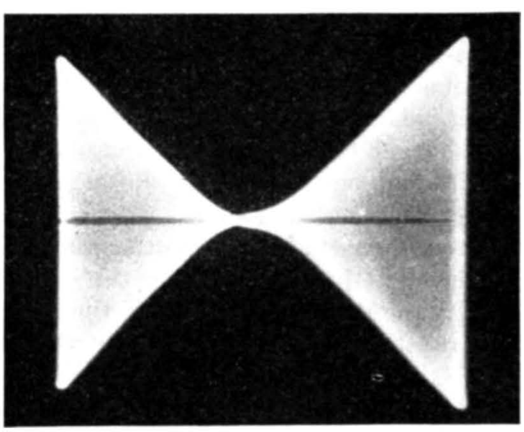

Fig. 9—Trapezoidal pattern: balancing resistors correct, but too high a negative bias on screen grids.

A similar procedure can also be employed to equalize the outputs of the two tubes in each pair of balanced modulators. One possible way to adjust their outputs is to vary the grid-bias resistors shown in Fig. 2, until each tube produces the same rf output for a given positive dc screen voltage. To set the resistors, it is best to turn off one balanced modulator. If the remaining two tubes are not equalized with respect to each other, a trapezoidal oscilloscope pattern will resemble Fig. 8 or Fig. 11.

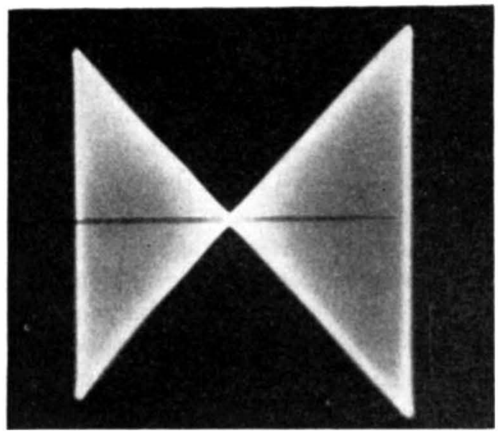

Fig. 10—Trapezoidal pattern: correct adjustment.

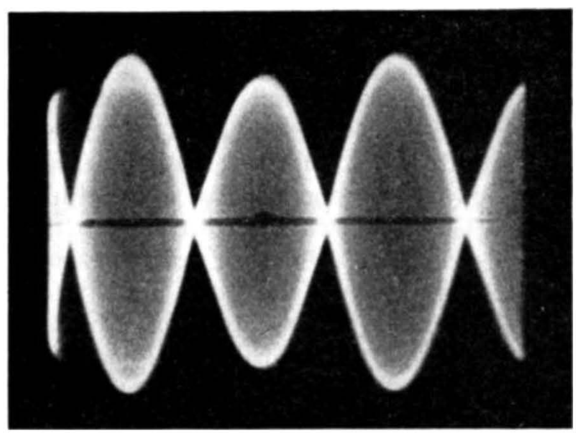

Fig. 11—Output envelope, same as Fig. 8.

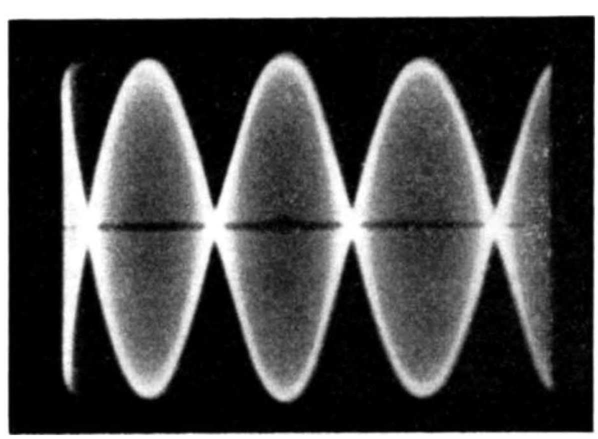

Fig. 12—Output envelope, same as Fig. 10.

Adjustment of the grid resistors for equal tube efficiencies will give a pattern like that of Fig. 10 or Fig. 12. The process is then repeated for the second pair of balanced modulators comprising the twin balanced modulator. If an oscilloscope is not available, it is possible with some practice to set the grid resistors by listening to the rectified output envelope of each balanced modulator in turn. The change in harmonic structure of the output envelope, when the two tubes of each modulator are equalized, may be detected by ear.

The proper antenna loading is that which just allows, at full output, the dc plate voltage E_b in Fig. 2 minus the peak rf output voltage E_{ac} across the common tank circuit, to equal the voltage at which a virtual cathode forms within the tubes. This is the point at which the tube plate current just begins to be appreciably affected by the instantaneous plate voltage and corresponds to the condition where saturation begins to occur.

The first step in designing a twin balanced modulator is the selection of tetrode or pentode tubes (preferably the former, because of their lesser output capacitance) whose rated plate dissipation is one-fourth the desired peak rf power output. (This follows because the full-output efficiency for a sine-wave modulating signal is approximately 50 per cent in practice, and each tube dissipates one-fourth the total power lost.)

The rated tube carrier power output for class-B linear or grid-modulated telephony is noted, as well as the rated plate and screen potentials for this class of service. In order to produce the greatest peak power for single-sideband service, it is desirable to scale the operating potentials upward by a factor equal to the square root of the ratio between the peak power desired and the peak power developed in linear or grid-bias modulated amplifier service. This will ordinarily be by a factor of the order of 1.5 times. The average dc plate current consumed at full output may be calculated on the basis of the assumed over-all efficiency, which is $\pi/4$ times, or 78.5 per cent, that of one tube operating as a linear amplifier of single-sideband signals.

The screen input power required by the four tubes is supplied by the two audio amplifiers. The audio power needed is, in point of fact, far less than what would be expected if it were assumed that the screen current drawn in this class of service was in direct proportion to that drawn in class-C telegraph operation. On the basis of such an assumption, it may be shown that the two audio amplifiers would each be required to supply an audio power output equal to the normal class-C telegraph dc screen input to one tube of the four-tube twin balanced modulator. This in itself is not a large amount of power, as the dc screen input to a tetrode tube is commonly of the order of one-twentieth the radio-frequency output power.

However in twin balanced modulator service, a considerable saving in screen modulating power results from the fact that virtual cathode formation is not permitted to occur at any point in the rf or af cycle. In normal class-C operation of tetrodes, the minimum plate voltage is allowed to swing down to the point where a large share of the space current is momentarily attracted to the screen grid, thus increasing the average screen current and input power. With aligned-grid tetrodes—now almost universally used—when the minimum plate voltage is not allowed to swing this low, an exceedingly small fraction of the space current is diverted to the screen grid. This is a result of the normal electron focusing action within the tube. In consequence, a screen-grid-modulated balanced modulator requires a remarkably small amount of audio power, of the order of 3 or 4 watts for a 500-watt output. The regulation of the voltage source must, of course, be good; a requirement which can readily be met by use of inverse feedback.

The linearity of the modulation characteristic obtainable with a screen-grid-modulated balanced modulator has been found to be entirely acceptable.

The output capacitances of the four tetrode tubes, which are all in parallel across the common tank circuit, may lead to lower-than-normal tank-circuit efficiency in the 15- to 30-Mc frequency range. With coils of ordinary Q, the inductance required for resonance, and therefore the parallel-resonant impedance, become small. It will be found that tubes of recent design tend to have much lower output capacitances than their older counterparts. The 125-watt RCA type 803 pentode, for example, has an output capacitance of 29 $\mu\mu$f; yet the 1000-watt Eimac type 4-1000A tetrode has an output capacitance of but 7.6 $\mu\mu$f.

The effects of tube output capacitance may, of course, be mitigated by designing the tank circuit as a pi network rather than as a simple resonant circuit, as is sometimes done in television practice.

Choice of the proper tank-circuit Q is conventional. It may be of help to recall here that, at one point in the audio cycle, one of the four tubes is delivering the entire output, under conditions equivalent to those obtaining in any class-C amplifier.

Oscillograms illustrating the operation of an experimental twin balanced modulator are shown in Figs. 8 through 14. Figs. 8 through 10 are trapezoidal patterns obtained with only one balanced modulator running. In Fig. 8, the grid-bias resistors are improperly set, so that one tube has a lower efficiency than the other. The correct adjustment is that which makes the slopes of the diagonal lines the same in each half of the pattern. In Fig. 9, the resistors are set correctly, but the negative screen bias is excessive in relation to the tube plate voltage and rf excitation. Both tubes are cut off during a portion of the audio cycle. Fig. 10 shows the correct adjustment. (The right-hand half of this pattern is somewhat larger than the left because of stray carrier voltage being picked up from a buffer stage.) Fig. 11 shows the appearance of the output envelope for the conditions of Fig. 8, while Fig. 12 shows the correct adjustment. In Fig. 13 both balanced modulators are running, but the undesired sideband is not completely

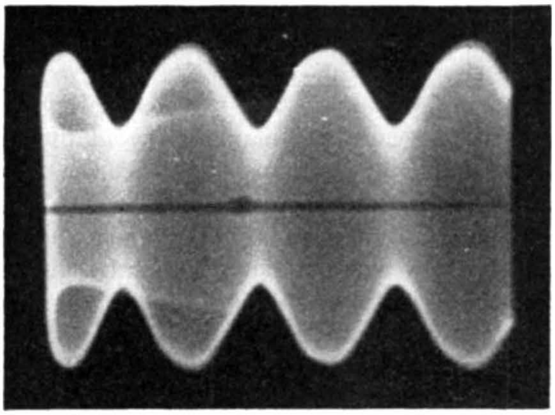

Fig. 13—Output envelope, both modulators: either amplitudes equal, rf phases incorrect, or amplitudes unequal, rf phase correct.

cancelled due to inaccuracies in either phase or amplitude setting, or both. The correct phase and amplitude settings produce an output envelope very nearly free from ripple, as will be seen in Fig. 14.

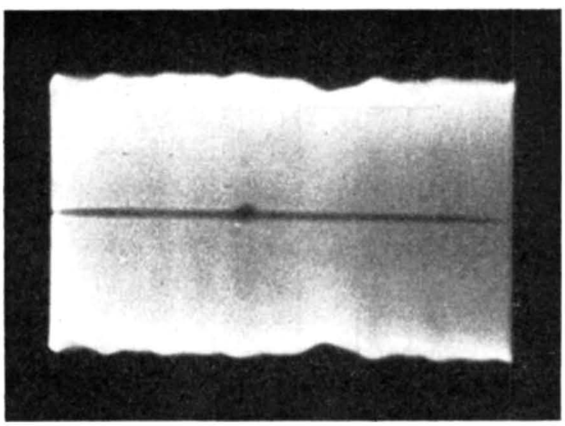

Fig. 14—Output envelope, both modulators: amplitudes and phases correct.

It is worth noting that the undesired-sideband rejection obtainable with the twin balanced-modulator circuit can be made as complete as desired by proper adjustment of phase and amplitude settings. (The rejection is to some extent dependent on audio level.) The amount of rejection achieved in practice over a band of modulating frequencies, then, is primarily determined by the phase and amplitude characteristics of the audio phase-shift network. Tests have shown that it is readily possible, at a single audio frequency, to attenuate the undesired sideband far below the level of the higher-order sidebands caused by the presence of harmonic distortion in the audio amplifiers and by nonlinearities in the modulation process.

An advantage of phase-rotation single-sideband transmitters, in general, is the fact that the sideband radiated may be changed by reversing the polarity of the audio input to one of the component balanced modulators. This may be an operating convenience, particularly when the receiver is one of the type deriving its effective selectivity from a combination of an audio low-pass filter and 90° wideband phase-shift networks.[5] In this instance, both transmission and reception can be changed from one sideband to the other merely by throwing switches at the transmitter and receiver, no retuning being needed. In certain cases, such flexibility is of great assistance in avoiding interference.

To radiate a reduced carrier with the twin-balanced modulator, a positive bias voltage may be applied to the screen grid of one of the four tubes.

Finally, it seems desirable to point out that one program may be transmitted on one side of the suppressed carrier, and another on the other side, by providing a second pair of audio inputs between which a 90° phase shift exists. The polarity of one of these channels is reversed with respect to that of its counterpart in the first pair. These two pairs of audio voltages, when added and fed to the input terminals of the twin balanced modulator, will produce two independent sets of single sidebands, one on either side of the carrier.

Conclusions

Two high-power balanced modulators, biased to cutoff in the absence of an audio input and using tubes which behave substantially as constant-current sources, may be connected to a common tank circuit for the generation of single-sideband signals by the phase-rotation method. The efficiency obtainable with this arrangement approximates that of a conventional linear amplifier.

Simplicity, ease of adjustment, and power economy recommend this circuit for applications where a certain amount of distortion and undesired sideband output can be tolerated.

Acknowledgment

The work described in this paper represents an outgrowth of research carried out under Contract W28-099-ac-131 between Stanford University and the Air Matériel Command, Watson Laboratories. The author is indebted to the Air Matériel Command for permission to publish, and to Frederick E. Terman, Dean of the Stanford School of Engineering, for inspiration, help, and guidance.

[5] O. G. Villard, Jr., "Simplified single-sideband receiver," *Electronics*, vol. 21, pp. 82–86; May, 1948.