Convention Record of the I·R·E

1953 NATIONAL CONVENTION

Part 4—Broadcasting and Television

SESSIONS ON . . .

- Television — I
- Television — II
- Symposium: Television Broadcasting
- Symposium: Television Broadcasting and UHF
- Broadcast and Television Receivers — I
- Broadcast and Television Receivers — II

SPONSORED BY

IRE PROFESSIONAL GROUPS ON . . .

- Broadcast and Television Receivers
- Broadcast Transmission Systems

Presented at the IRE National Convention, New York, N.Y., March 23 - 26, 1953

The Institute of Radio Engineers

CONVENTION RECORD OF THE I.R.E.

1953 NATIONAL CONVENTION

PART 4 - BROADCASTING AND TELEVISION

TABLE OF CONTENTS

Session 35: Broadcast and Television Receivers I
(Sponsored by the Professional Group on Broadcast and Television Receivers).

Session 41: Broadcast and Television Receivers II
(Sponsored by the Professional Group on Broadcast and Television Receivers).

THEORY OF SYNCHRONIZATION APPLIED TO NTSC COLOR TELEVISION

Donald Richman
Hazeltine Corporation

Introduction

The field test standards for color television specified by the National Television System Committee add color a monochrome picture by means of a frequency-interleaved carrier color signal which carries one component of the color information in its phase and another component in its amplitude. A color sync signal at a reference phase is provided in the transmitted signal to permit measurement of the instantaneous phase angle of the carrier color signal in receivers.

The color sync signal carries phasing information. The adequacy of this phasing information has been questioned.

At the request of the NTSC a study was made to find out how much phasing and synchronizing information is contained in the color sync signal and how it may be used. The study showed that the phasing and synchronizing information inherent to the color sync signal is sufficient to provide completely automatic operation. This means that a customer-operated control relating to color sync should be unnecessary on NTSC color television receivers. This paper will outline the facts upon which this conclusion is based.

I. The Color Burst

Figure 1 shows the NTSC synchronizing waveform during one line retrace interval. The color synchronizing signal consists of a burst of approximately 9 cycles of sinusoidal waveform at the color carrier frequency of 3.58 mc. It occurs during the line blanking pulse following each horizontal sync pulse. It is omitted during the 9 lines in each field in which the field synchronizing information is transmitted.

The color burst is used in the color television receiver to provide a control signal for the generation of a local continuous wave signal at the nominal burst frequency and locked to it in phase.

II. Burst Sync Integration System

Figure 2a shows a typical burst sync integration system. It consists of a Time Gate, (normally keyed from line flyback), an Amplitude Limiter, and an Integrator which averages the synchronizing information obtained over some period of time.

The input to the integration system consists of color sync plus undesired noise components. Figure 2b indicates how the sync can be phase modulated by noise.

The output of the integration system is a steady sync reference signal, as indicated by Figure 2c. The rms phase error due to input noise in the output reference signal is held within selected limits by the integration.

The narrow-band integrator can produce a static phase error. The static phase error may be made as small as desired by the use of degenerative feedback within the integrator. This long-time control may be achieved independently of the other basic properties of the sync system.

III. Integration Time for Phase Accuracy

Thermal noise is the most serious form of noise interference because it can only be discriminated against by averaging. It has therefore been used as the measure of interference which must be overcome by burst sync circuits.

Figure 3 relates the video signal-to-noise ratio, S_o/N_W, of a receiver to the rms phase error, Φ_{rms}, of the output reference signal coming from the integrator. The integration time, T_M, required to hold the error to the selected level, is a parameter. N_W is the rms noise in a 4.3 megacycle video bandwidth. A receiver having a 6 db tuner noise figure produces a signal-to-noise ratio of 1 with a 30 microvolt signal applied to a 300 ohm antenna.

The point A shown on Figure 3 corresponds to the approximate level of performance desired from the system. The rms phase error is to be held to within 5° at a signal-to-noise ratio of unity. This shows the required integration time to be of the order of magnitude of 1/200th of a second.

Circuits with this measure of noise immunity have been found experimentally to reduce color sync noise to about the point where it disappears in the video noise when color pictures are viewed.

IV. Passive Integrator

Figure 4a shows the block diagram for an integrator in which the required integration is obtained by use of a passive high-Q filter.

The amplitude response of the filter as a function of frequency is shown in Figure 4b. The equivalent noise bandwidth of the filter is f_N. The noise bandwidth of the filter should not exceed approximately 200 cps for the level of performance indicated on the previous slide. The necessary Q is then approximately 28,000; this requires the use of a crystal filter. Practical crystals in the frequency range of the color subcarrier can achieve the required Q but apparently cannot as yet exceed it by a large factor.

Figure 4c shows the phase characteristic associated with the crystal selectivity. Undesirably large static phase shifts could occur with reasonable tolerances on crystal frequency stability of tuning, but are prevented in the circuit shown by the use of feedback. The circuit includes the high-Q filter, a variable phase shifter, a 90° phase shift unit, and a low-pass (DC) filter, connected in a feedback loop.

The static phase may be held as closely as desired by putting enough DC gain in the feedback loop. The signal-to-noise ratio at the output of the system will not be impaired if the bandwidth of the DC filter is sufficiently narrow.

Rapid phase stabilization can occur when channels are switched if the frequency stability is comparable to the noise bandwidth. The stabilization time for small detunings will be of the order of a few times the transient time constant of the phase feedback loop. This means that completely automatic performance is possible with this circuit.

The ultimate limitations of passive integrators are: the limitations on Q, and the degradation of performance when the filter bandwidth is narrow compared to tuning stability.

V. Standard Automatic Frequency and Phase Control Locked Integrator

Figure 5a shows the block diagram of a standard automatic frequency-and-phase control loop. It includes a local reference oscillator, a phase detector which measures the phase difference between the sync signal and the oscillator, a filter, and a reactance tube which controls the oscillator frequency.

The filter has low uniform transmission in a frequency band up to half line frequency, and very high DC transmission.

The oscillator is held at sync frequency by a DC control voltage generated in the phase detector by a static phase shift $\Delta\phi$, and the amount Δf, by which the oscillator is pulled in frequency. The static phase error can be made as small as desired by making the DC loop gain high; this makes the holding range, f_c, much larger than the normal operating range.

Figure 5c shows (to the same scale) the modulation response curve for the APC loop as a function of modulation frequency. The APC system obtains its selectivity by heterodyning the sync signal against the oscillator and passing the phase difference information through a low pass filter with feedback. The response is determined as much by the feedback as by the AC transmission of the filter. The noise bandwidth of the APC loop is indicated on the figure as f_{NN}. Since an APC loop phase detector does not distinguish between noise components above or below the oscillator frequency, the noise bandwidth of the APC loop should not exceed approximately 100 cps for the desired level of performance. The APC loop is not limited in effective Q, and is automatically peak tuned. However, reducing the noise bandwidth impairs the pull-in performance.

Figure 5d is a sketch of pull-in time for this loop as a function of the initial oscillator detuning Δf. The time increases with Δf at a rate not slower than square law to the limit of the pull-in range, at which point it becomes infinite. For most designs the pull-in range is substantially smaller than the limit of one-half line frequency.

VI. Standard APC Optimum Pull-In

There is an upper limit to the pull-in performance obtainable with the Standard APC loop; this limit is shown in Figure 6 which relates the pull-in time T_F in seconds, to the noise bandwidth f_{NN} in cycles per second.

The relationship between noise bandwidth and pull-in time is shown for several values of Δf. For example, point A shows that an optimum design unit of this type having a noise bandwidth of 100 cycles can pull in from a value of Δf = 1000 cycles in 4 seconds. In the limit, the pull-in time is proportional to $(\Delta f)^2$, and inversely proportional to the cube of the noise bandwidth. This indicates that the Standard APC loop is adequate for completely automatic color synchronization.

Not all designs of the APC circuits will achieve the limits of performance shown in Figure 6. In fact the majority of past designs have fallen short of the limit. A method for obtaining all of the performance of which this circuit is capable, along with a more detailed description of its properties, is presented in the paper "Automatic Phase Control Synchronization for NTSC Color Television", later this afternoon.

The systems presented thus far permit a level of performance which appears to meet but not exceed the requirements for completely automatic burst synchronization. The signal itself permits substantially better performance.

VII. Response to Frequency Difference

Sync systems have two separate and distinct modes of performance. These are: maintaining phase stability after synchronization, and achieving synchronization. Each mode has its own limitations. Figure 7 shows how these are related in the standard APC system.

Figure 7a shows the modulation passband of the APC loop. Figure 7b shows the control voltage generated by this loop in response to an oscillator detuning, Δf, plotted to the same horizontal scale.

Within the linear region, frequency pull-in is instantaneous. A DC voltage proportional to the initial value of Δf is generated in the first cycle of beatnote. For larger values of Δf the circuit generates a control voltage which varies inversely with Δf as shown.

The inverse shape is not undesirable in itself. However, the relationship shown between noise transmission and generated voltage for pull-in for the standard APC Loop causes the expensive exchange of pull-in time for noise immunity shown earlier.

The pull-in performance for a specified noise bandwidth may be improved by use of an auxiliary frequency detector.

Figure 7c shows a form of frequency control curve which may be ideal for the auxiliary frequency difference detector. It approaches linear frequency detection when Δf is large but is inactivitated after synchronization has been achieved because of the flat region near the center of the curve.

VIII. Improved Synchronization System.

For best performance the sync system should make the pull-in and hold-in modes of operation as independent as possible of each other's limitations.

Figure 8 shows the block diagram for a sync system which does this. It supplements the Standard APC system shown earlier with an auxiliary frequency difference detector to improve the pull-in performance.

The frequency detector must be able to discriminate against noise.

The real limitation of the sync system with respect to frequency pull-in is the ability of the system to recognize a frequency difference and distinguish it from noise.

This sets the real upper limit of performance. If the frequency measurement is linear, then, after a time delay during which the frequency difference is measured, the reference oscillator may be switched instantaneously by the proper amount to provide synchronization. A system for doing this may be called an ideal sync system.

The shortest frequency pull-in time for reliable performance is the integration time necessary to measure the frequency difference with enough accuracy so that after switching, the oscillator will be within the range of instantaneous pull-in shown previously. For example, in a burst sync system designed for 7-1/2 kilocycles peak pull-in range, the integration time required for accurate frequency measurement at a signal-to-noise ratio of unity corresponds to approximately one cycle at a frequency equal to the noise bandwidth.

The ideal, switched, system has equal pull-in times for all frequency differences.

In the sync system of Figure 8, an essentially direct connection for the auxiliary frequency difference detector has been found practical. Such a system provides an intermediate level of performance which appears commercially attractive.

IX. Quadricorrelator: Frequency Difference Detector

Figure 9 presents one form of the frequency difference detector called a Quadricorrelator. The quadricorrelator can produce the frequency control curve shown earlier, with the necessary noise immunity. It includes a pair of synchronous detectors which heterodyne the sync signal against two reference signals in quadrature with each other. When a frequency difference exists, beatnotes appear at the outputs of the synchronous detectors and pass through the identical bandpass filters shown.

A differentiating circuit in the I-channel shifts the phase of the I-signal by 90°. The resulting signals are applied to a third synchronous detector, the output of which is filtered as shown. The output voltage is proportional to frequency difference for those frequencies which are passed by the band pass filters.

The vector diagrams (b) and (c) indicate how the polarity of the frequency difference is determined. R_I and R_Q are vectors in quadrature. The rotating vectors represent the sync signal. The relative direction of rotation depends on whether the frequency difference is positive or negative. When $\Delta\omega > 0$, the rotating vector comes in phase with R_Q 90° before it comes in phase with R_I. When $\Delta\omega < 0$ the opposite is true. The 90° phase advance caused by differentiation of the I-channel beatnote makes the signals applied to the final synchronous detector be in phase or opposing, depending on the sign of $\Delta\omega$.

The quadricorrelator obtains its noise immunity as a frequency detector by heterodyning a beatnote against the only independent signal available at identically the same frequency, which is a beatnote derived from a quadrature sample of the same signal. Thus, as indicated by Figure 9d, it effectively permits an automatic tracking heterodyne filter to be applied to the beatnotes.

One additional feature of the quadricorrelator has proven useful; its component parts are already present in NTSC color television receivers.

X. Comparative Performance

Figure 10 compares the upper limit of burst sync performance to the best previous performance. It shows the pull-in time in seconds plotted against a figure of merit which is the product of input signal-to-noise ratio and output rms phase error; for best performance these are both small.

Two groups of curves are shown. One group relates to the Standard APC system; point A corresponds to a signal-to-noise ratio of one and an rms phase error of 5°. For this case the Standard APC system can pull in from 1 Kc in four seconds.

The second group indicates the approximate limit of performance for the burst sync signal. For the same degree of noise immunity, and a 7-1/2 Kc pull-in range, point B shows a pull-in time of .05 seconds.

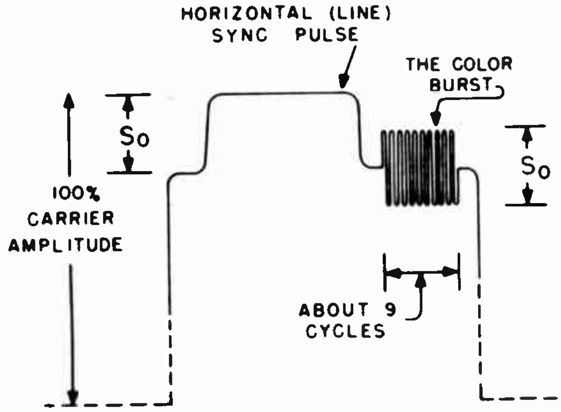
It is interesting to see what would happen in the unlikely case that the required phase error were reduced to one degree, making the abscissa unity. Limiting pull-in performance for the burst sync signal would then be about one second of pull-in time, while the Standard APC loop with optimum design would require seventeen hours.

Conclusions

The preceding discussion has shown that completely automatic operation of burst sync circuits is possible, simultaneously providing high noise immunity and negligible pull-in time. Two forms of well-known circuits are capable of meeting these severe requirements. The signal, and simple new circuits permit performance far in excess of what is required.

Bibliography

The mathematical derivations of formulae presented in this paper appear in Technical Monograph #7 of the National Television System Committee. This monograph will appear in the IRE Proceedings later this year.



$$F_{SC} = \frac{455}{2} F_H = 3,579,545 \pm 11 \text{ cps}$$

Fig. 1 - Burst and line sync waveforms.

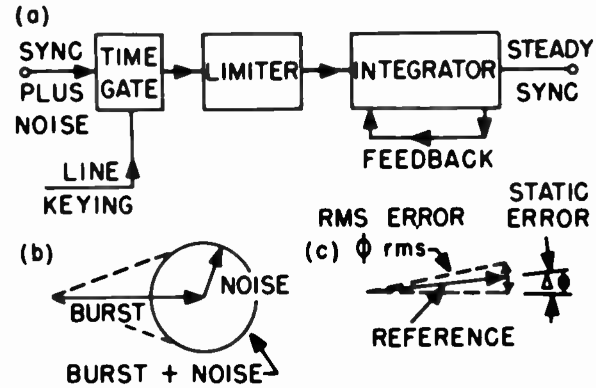


Fig. 2 - Burst sync integration system.

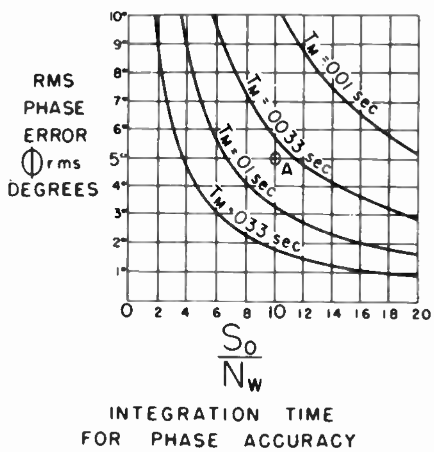


Fig. 3 - Integration time for phase accuracy.

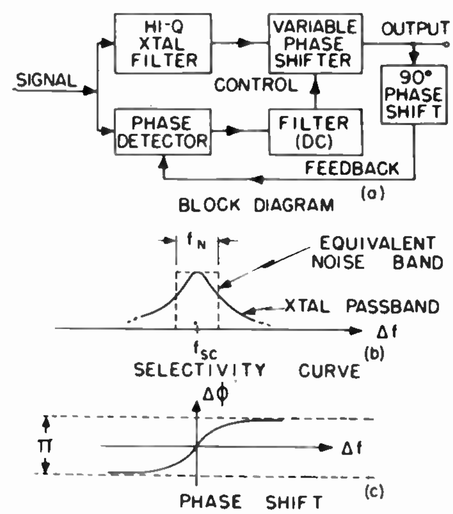


Fig. 4 - Passive integrator.

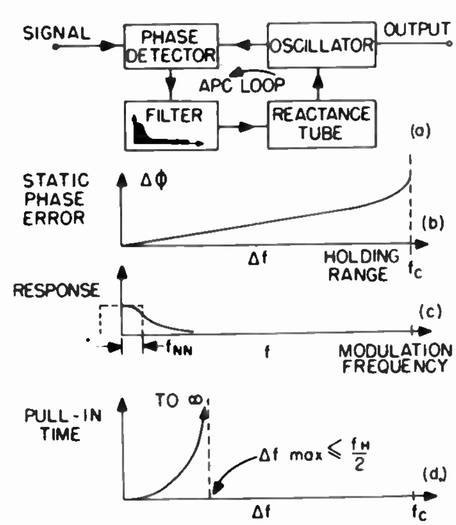


Fig. 5 - Standard APC locked integrator.

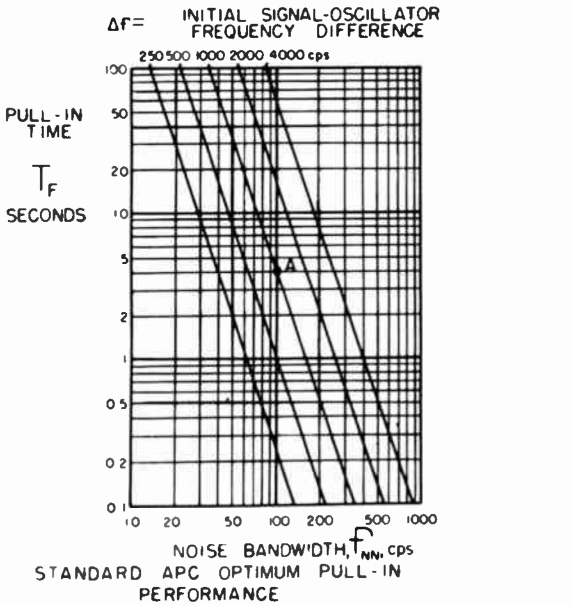


Fig. 6 - Standard APC optimum pull-in performance.

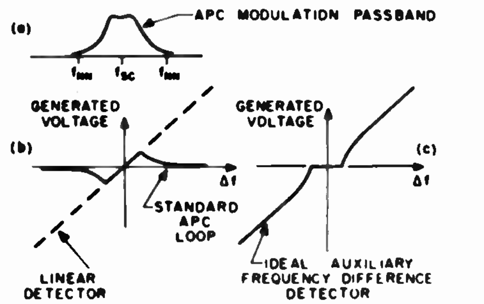


Fig. 7 - Response to frequency difference.

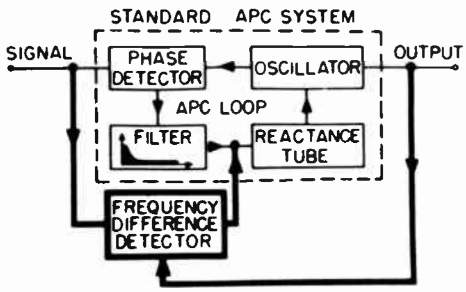


Fig. 8 - Improved synchronization system.

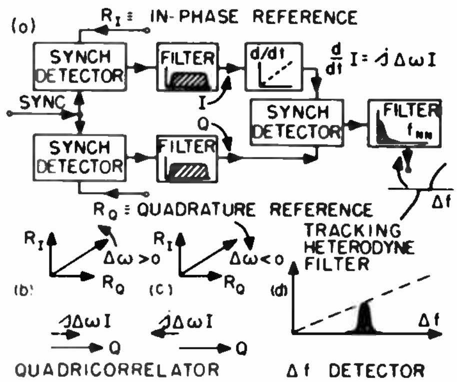


Fig. 9

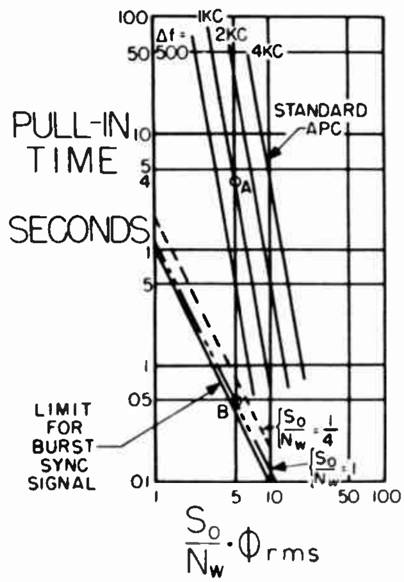


Fig. 10

COLOR SYNCHRONIZATION IN THE NTSC COLOR TELEVISION RECEIVER BY MEANS OF THE CRYSTAL FILTER

William E. Good
General Electric Company
Syracuse, New York

Summary---The quartz crystal filter or ringing circuit shows promise as a color synchronizing system for the NTSC color television receiver. Its performance is comparable to the A.F.C. type circuit and it responds in a passive fashion. A typical circuit is given, the design parameters are discussed and some of the difficulties are noted.

The commonly used technique for color synchronization in the NTSC color television receiver is the use of standard or modified Automatic Frequency Control circuits.[1,2,3,5] Another technique that shows promise is the use of the passive crystal filter.[1,3,4] This paper will cover the operation of the crystal filter and compare its performance with that of the A.F.C.

The function of the color synchronization circuit in the NTSC color receiver is to develop or generate a continuous wave 3.58mc voltage that is synchronized in frequency and phase with the reference "burst." This burst is transmitted after each horizontal synchronizing pulse and consists of eight cycles of the 3.58mc color subcarrier. (Figure 1a) The phase of this reference burst is that of -(B-Y) as shown in Figure 1(b). The continuous wave R.F. voltage is used to switch the synchronous detectors or color demodulators. Of course, the phase of this reference voltage must be properly set so that the demodulators detect along the correct vectors, the phases of which are indicated in Figure 1(b). Figure 2 shows the arrangement of the circuits in this part of the color receiver, from the second detector through the synchronous detectors.

The fidelity of the final colors on the TV screen is determined by the phase accuracy of the reference voltage at the synchronous detectors. Experimental evidence has shown that a phase inaccuracy of 5° to 10° may result in perceptible color degradation. This amount of phase error may thus be considered to be the tolerance that the system must maintain for satisfactory performance. This tolerance, of course, must be divided throughout the whole color system. If we allow the receiver the lion's share of say 5°, then we might reasonably allow the color synchronization portion up to say $\pm 3^\circ$ of phase error due to any one cause.

With this tolerance in mind we may now go ahead with the operation of the crystal filter and the criteria for its design. The crystal filter is basically a narrow bandpass filter, centered at the color subcarrier frequency of 3.58mc, which filters out the fundamental or carrier component from the burst spectrum. A burst gate is introduced ahead of the filter circuit to prevent picture information from influencing the phase of the resulting sine wave. Perhaps a simpler way of thinking of this operation is that the burst shock-excites the crystal and it continues to vibrate or "ring" for the rest of the horizontal picture interval. For this decaying wave to remain of appreciable amplitude during the 220 cycles of the horizontal picture interval, the Q of the circuit must be relatively high. Even with the highest usable Q's there is some decay during this interval, so that it is necessary to follow the filter with a limiter to obtain constant output for the synchronous detectors. It will be noted that the operation of such a filter is purely passive and depends entirely on the presence of the burst for any output to appear.

The appropriate bandwidth or Q of the filter is determined by evaluating two factors. One of these is the performance of the system under the influence of noise.[1,3] As might be expected, the narrower the filter, the less the phase of output voltage will be affected by a given amount of noise in the input signal. Thus good performance under noise conditions dictates a very narrow bandpass or a high Q. The other factor is the static phase shift or error due to the variation of some parameter in the color synchronizing system. The first of these is the variation of the color subcarrier frequency at the transmitter. The NTSC specifications[6] allow a tolerance of $\pm$ 0.0003% in the 3.579545mc subcarrier frequency which amounts to approximately $\pm$ 11 cycles. Considering the filter as a series tuned circuit, the phase shift of 3° for an 11 cycle frequency change necessitates a bandwidth of 440 cycles or a Q of approximately 8,000. This is an upper limit on the Q if we restrict ourselves to the $\pm 3^\circ$ phase shift. With a Q of 8,000, the wave will decay to approximately 90% of its initial value during one line. The other main source of variation would probably be due to a temperature change of the crystal in the receiver with a resulting change in its own natural frequency. The amount of power in the crystal filter circuit is small so that self heating of the crystal is probably negligible and, therefore, the ambient temperature in the vicinity of the crystal is the important factor. The phase change produced by a drift in crystal frequency is the same as that produced by an equal change in transmitter subcarrier frequency; so if we want again a 3° phase drift due to crystal frequency drift, a Q of 8,000 allows the frequency drift to be the same as the transmitted subcarrier drift or .0003%. Practical crystals can be made with a temperature coefficient of 1 ppm/$^\circ$C. in which case the allowable temperature tolerance is $\pm 3^\circ$C., after the equipment reaches its operating temperature. It would again appear that a Q of 8,000 represents a reasonable upper limit.

The absolute frequency to which the crystal must be ground can be determined by figuring the phase shift that would occur during one horizontal picture interval. If it is assumed that the output reference phase is correct at the beginning of the line, then it should be off by, say, no more than 3^o at the end of the line. This would be accomplished if the resonant frequency of the crystal differed by less than 140 cycles/second from the correct frequency. This gives a grinding tolerance of $\pm$ 0.003%. This tolerance could be easily enlarged by providing some tuning for the crystal in the filter itself by use of a variable capacitor in series with the crystal. Of course, the static phase error would be large for a crystal off frequency by 140 cycles ($\sim 30^o$ for this case) and the overall circuit must have enough range of phase control to make up for it.

A typical circuit for the crystal filter is shown in Figure 3. The burst gate is the first tube and the step-down transformer in its plate circuit reduces the impedance of the source to a value that will efficiently drive the crystal. The second tube is an amplifier which brings the signal up to a sufficient amplitude for driving the limiter, which is tube number three. The crystal filter circuit is quite similar to that in a communication receiver in which the crystal is used in its series resonant mode. The total resistance in the crystal circuit determines the Q. As the inherent Q of the crystal is much higher than the desired Q, this resistance is the sum of the values of the termination and of the source. The internal series resistance of an AT cut quartz crystal may be 50 ohms while the total external resistance may be 500 ohms for a Q of 8,000. Figure 4 shows typical values for one type of AT cut quartz crystal, as well as the equivalent circuit and the required external resistance as a function of bandwidth. The bridge type circuit is used to neutralize the shunt capacitance of the crystal and thus to permit the device to be an effective filter, particularly at frequencies away from resonance. The neutralization condenser may be adjusted by minimizing the transient that occurs during the burst interval as shown in Figure 5. If the neutralization is not used, striations may appear in the picture and also the noise bandwidth will be wider than necessary. The voltage developed in the termination must be amplified to a suitable value for effective limiting. Figure 5(a) shows the gated burst while the 5(b) shows the output of the un-neutralized filter. The neutralized filter output is shown in 5(c). Looking at the filter output at vertical rate shows the decay during the vertical synchronizing interval when no burst is being transmitted. An interesting side light appears in 5(d) in which a beat note shows up in the transient if the crystal is off frequency.

The limiter that has been used most is the simple grid current type with a time constant in the neighborhood of $2\mu s$. This limiter must be fairly effective in keeping the amplitude of the output wave at a constant value for driving the synchronous detectors. Otherwise, the variation in the amplitude of the reference wave may appear in the final output and give rise to shading of the color picture. For good noise performance, the limiting should be effective enough to prevent wide variations in amplitude of the driving signal due to noise, from appearing at the output of the limiter. If amplitude variations do appear in the output, then colored streaks may appear in the picture each time the limiter fails. An amplifier (V_2) is shown in the circuit of Figure 3 between the filter and the limiter to assure sufficient signal amplitude for effective limiting. With care, it appears feasible to eliminate V_2 and substitute an impedance matching or voltage step-up network and still obtain effective limiting.

No attempt is made here to incorporate a "color killer" which would cause the color portion of the set to become inoperative if no burst is present. However, it should be pointed out that the DC voltage that is developed at the grid of the limiter could be used for this purpose. In fact, this same DC voltage could be used for color A.G.C.

The transformer at the plate of the limiter is used to furnish the reference signals at 0^o and -90^o to the two detectors. Fortunately, even if the phase does come out wrong with all circuits tuned to resonance, it is possible to obtain the correct values by switching leads of the transformers involved. Of course, each tuned circuit is a potential source of phase shift or error and as such should be designed to minimize this possibility. A master phase control (perhaps a panel control) can be had by tuning any one of the tuned circuits, preferably ahead of the limiter. The circuit in the plate of the burst gate is convenient for this purpose.

Preliminary tests at the General Electric Company in Syracuse show that noise bandwidths of 500 to 1000 cycles should give satisfactory performance of the color synchronizing system.

This filter has been used in several experimental field test receivers with good results. In a number of practical field tests such as receiving color transmissions from WHAM about 70 miles away, it has been possible to obtain effective color synchronization even though the burst was barely visible in the noise. Saturated color bars were discernable at this signal to noise ratio but color slides were not.

One exasperating difficulty that arose early in this work was with vertical striations appearing across the screen which could not be tuned out, and even though the output voltage was effectively limited in amplitude a resulting phase shift occurred as pictured in Figure 6. Figure 6(c) is the output of a quadrature detector showing the phase error of the reference signal. This effect was traced to a spurious mode of the crystal which was exactly on the same frequency as one of the Fourier series components of the burst spectrum, i.e., separated from the color carrier by an integral multiple of the 15,750 cycle line rate. Spurious modes that gave the appearance of striations or vertical bars occurring at 3 and 5 times line rate away from the desired frequency appeared most often. A number of crystals were measured and found to have spurious modes at or near these frequencies plus several others - all on the high frequency side. The problem has now been solved by contour grinding the crystals in such a way as to keep the spurious modes in between these critical values or moved out far enough so they can do no harm. We are indebted to Mr. Roy Lewis of the General Electric Company for his assistance in this problem.

Another cause of phase error could be due to the action of the grid current type of limiter. The phase shift due to the decaying wave itself would be quite negligible but it is conceivable that the action of the limiter might be such as to cause the phase to shift as the wave decays. Several preliminary experiments seem to indicate that this effect is small, especially if the Q is high, and may amount to an error of less than 2°. The general slope in Figure 6(c) is apparently due to a combination of phase change due to limiting and phase change due to slight off frequency operation of the crystal. By tuning the crystal with a series capacitance, it is possible to keep the overall effect quite small.

To summarize, the crystal filter can be designed which has only a ± 3° phase error for transmitter frequency variations, a reasonable stability with temperature change and a passive type of performance. With the 3db bandwidth of 440c (Q = 8,000) or a noise bandwidth of $\frac{\pi}{2}$ x 440c or about 700 cycles, it is capable of good performance in noise right down to a point where the color picture is no longer usable.

The A.F.C., when properly designed, can, for the same noise bandwidth, have a very small static phase error with frequency change - much less than a degree for tolerable subcarrier frequency changes. Its output is quite constant. It has the problem of pull-in range and that of sufficient inherent stability to prevent side-lock to one of the other sidebands in the burst spectrum. When side-lock occurs, vertical color bars appear in the picture and as the oscillator is locked in, it is perfectly content to remain side-locked. The noise performance will be the same as the crystal filter for equivalent noise bandwidths, except possibly for the superior amplitude limiting characteristic of the A.F.C. circuit. The A.F.C. may lose control with a noisy signal and then the colors will appear to roll in a vertical direction, as the oscillator runs free and slightly off frequency.

In general, then, the two circuits are capable of equivalent performance under the same conditions with the A.F.C. having a slight edge, once it is locked in. With weak signals the A.F.C. may drop out of synchronization and give rolling colors while the crystal filter just fades out due to the passive nature of its operation. The A.F.C. may side-lock while the crystal cannot.

The work on the crystal filter for color synchronization has been greatly assisted by the efforts of three G.E. Advanced Course Students, namely, S. M. Garber, O. J. Jacomini and T.T. True.

References

[1]T. S. George - "Analysis of Synchronizing Systems for Dot-Interlaced Color Television," Proc. I.R.E. Vol. 39, page 124, February, 1951.

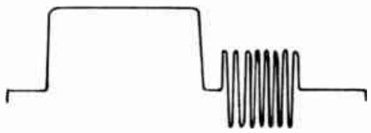
[2]K. E. Farr - "Compatible Color TV Receiver," Electronics, Vol. 26, page 98, January, 1953.

[3]D. Richman - "Color-Carrier Reference Phase Synchronization Accuracy in NTSC Color Television," NTSC Technical Monograph No. 7, NTSC-P12-302.

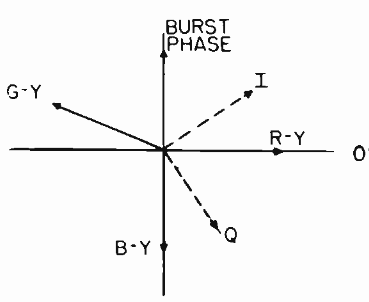
[4]W. E. Good - "Color Synchronization," Radio Fall Meeting, October, 1952 at Syracuse, New York.

[5]W. J. Gruen - "Theory of A.F.C. Synchronization," 1953 I.R.E. Convention Record.

[6]NTSC-G-306 Publication of February 4, 1953.



(a)



(b)

Fig. 1
(a) the NTSC color synchronizing waveform showing the horizontal synchronizing pulse on the left and the 8-cycle 3.58-mc burst on the right or back porch. (b) vector diagram showing the relative phases of the burst, the color difference signals and I and Q.

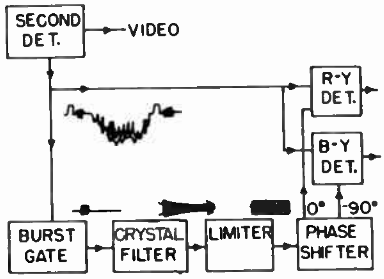


Fig. 2
Block diagram of the NTSC color television receiver between the second detector and the synchronous detectors showing the function of the crystal filter. Waveforms are depicted for one horizontal time interval.

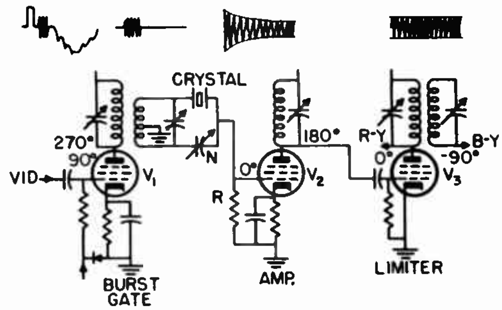


Fig. 3
A typical diagram of the crystal filter as used for color synchronization in the NTSC television receiver. V_1 is the burst gate. The crystal filter is followid by one stage of amplification and then by the limiter stage.

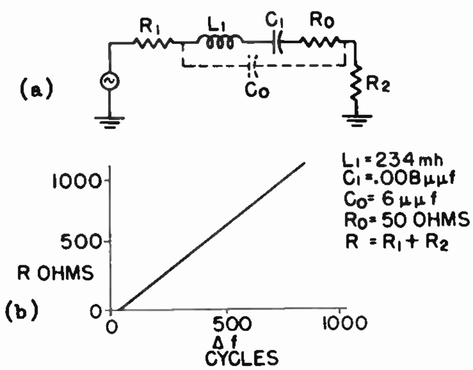


Fig. 4
(a) equivalent circuit of the crystal filter which operates in its series resonance mode. (b) typical values for one type of AT cut quartz crystal with graph of bandpass vs. the total external resistance in the crystal circuit.

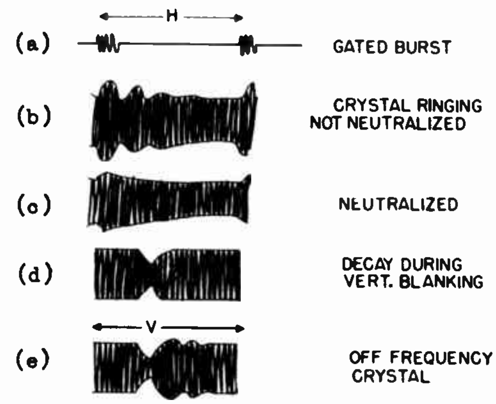
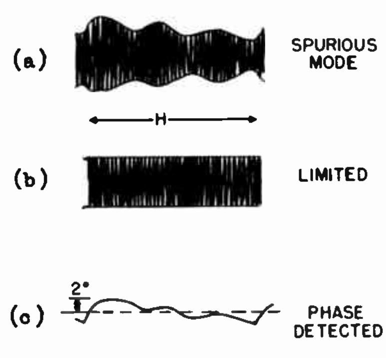


Fig. 5
(a) appearance of the burst at the plate of the burst gate tube. (b) decaying wave at output of crystal filter, not completely neutralized. (c) same as (b) with neutralizing condenser properly set. For a Q of 8000 the wave decays to approximately 90% during one line. (d) decay during vertical synchronizing interval when no burst is transmitted. (e) appearance of transient during buildup when crystal is considerably off frequency.

Fig. 6
(a) output of crystal filter (before limiting) showing the effect of a strong spurious mode approximately 3 X 15.75 kc away from the fundamental frequency. (b) wave in (a) after limiting. (c) output of a quadrature synchronous detector showing the phase error that exists in signal (b). The sine wave is due to the spurious mode, while the sloping part is due to a combination of a slightly off frequency crystal and possible phase distortion from the limiter.



APC COLOR SYNC FOR NTSC COLOR TELEVISION

Donald Richman
Hazeltine Corporation

Introduction

The complex synchronization systems used in modern television receivers are a result of the need for integration of the synchronizing information for improvement of signal-to-noise ratio. In the case of horizontal sync this problem has been universally solved in this country by the use of automatic phase control circuits such as that shown in Figure 1 in which the average frequency and phase of a local oscillator are controlled by the long-time-average phase of the synchronizing signal. The new synchronization problem presented by color sync in the NTSC system has therefore raised the problem of how these APC circuits may be applied for color sync. This paper will show how to obtain the most performance per unit cost from the Standard APC Loop.

I. "Standard" APC Loop Block Diagram

Figure 1 shows the block diagram of a standard automatic frequency and phase control loop. It includes a local reference oscillator, a phase detector which measures the phase difference between the sync signal and the oscillator, a filter which partly determines the integration properties of the APC loop and a reactance tube for controlling the oscillator frequency. The ratio, m, of AC to DC transmission through the filter is an important design parameter.

II. Measurable Parameters

Figure 2 shows two of the four basic parameters which determine the operating characteristics of the APC loop. Figure 2a shows the reactance tube sensitivity. It relates the control voltage E, applied to the reactance tube, to the frequency shift Δf, which it produces in the oscillator. It is normally nearly linear, having a slope $\beta = \frac{\partial f}{\partial E}$.

Figure 2b shows phase detector gain. It relates the phase difference, $\Delta\phi$, between the burst signal and the reference oscillator, to the voltage, E, which it produces. The curve (b) has a sinusoidal shape.

The slope of the phase detector gain curve at the operating point, indicated by $\mu = \partial E/\partial\phi$, is the gain of the phase detector for in-sync operation. μ is numerically equal to the peak value of the sine-wave. μ and the curve (b) must be measured statically.

The product $\mu \cdot \beta = f_c$ is the DC loop gain of the APC loop. f_c is numerically equal to the frequency holding range.

For normal in-sync operation at signal-to-noise ratios of interest the phase detector is essentially linear because it operates near the balance point shown.

However, during pull-in the system goes through many revolutions, each covering the entire range of phases. A phase deviation of many thousands of radians often occurs. Thus, while we can use linear analysis to determine the integration properties of the APC system after synchronization, we must resort to a solution of the existing non-linear system in order to find out how the system pulls in. Linear servo theory gives completely erroneous answers about pull-in.

III. Static Characteristics

Figure 3 shows the hold-in characteristics of the APC loop. It shows how much static phase error, $\Delta\phi$, is required to generate enough control voltage to shift the oscillator frequency by Δf. This curve is obtained from the curves for μ and β shown in Figure 2.

The DC loop gain, f_c, is equal to the holding range. By making f_c large enough, the maximum static phase shift may be held within any desired limits independently of the noise integration properties of the APC loop for in-sync operation.

Fortunately, making f_c larger makes the pull-in range larger and reduces the pull-in time. To illustrate this, the equation for the pull-in range is presented in Figure 3. The range is determined by the DC loop gain f_c, and the parameter m which is the ratio of AC to DC loop gain.

The term mf_c is the AC gain of the loop. The pull-in range is closely equal to the square root of the product of $2f_c$, the peak-to-peak holding range, and mf_c, and AC loop gain. For some designs, adequate pull-in performance requires a larger value of f_c than is required for tight static phase control.

IV. Dynamic Characteristics

Figure 4 presents the second pair of measurable loop parameters necessary for design. A simplified diagram of part of the APC loop shows the phase detector, the reactance tube, and the filter connecting them.

The filter consists of a large series resistance, R, often including the phase detector output impedance as a major portion, and the shunt time constant xT consisting of the resistor xR and the capacitor C.

For most frequencies in the pull-in range the capacitor is effectively a short circuit; the AC loop gain can be measured by observing the beatnote across the shunt time constant with the loop open. If the capacitor were shorted the DC and AC loop gains would be equal. For this simplified circuit the hold-in and pull-in ranges would both be reduced to mf_c and the noise bandwidth would be $\pi/2$ times mf_c.

The capacitor C permits the DC gain to exceed the AC gain, which is necessary for good pull-in performance. However, the reactance of the capacitor causes increased loop phase shift at some frequencies and hence a resonant rise in the transmission characteristic of the APC loop. This increases the noise bandwidth.

The noise bandwidth, f_{NN}, is a measure of how severely noise can phase modulate the oscillator. The noise bandwidth is very closely represented by the simple equation at the bottom of Figure 4. The added term due to capacitive reactance is the frequency at which the transmission of the APC filter is 3 db up from the minimum value m. The 3 db down frequency for the filter is much lower, often at a small fraction of a cycle.

V. APC Loop Modulation Response

Figure 5 shows the response characteristic of the APC loop as a function of modulation frequency. The frequency scale has been normalized by dividing by noise bandwidth. The curves shown represent the relative response of the system to small signal phase modulation by noise during in-sync operation. Curves are shown for five values of a parameter K.

Later in this paper, it will be shown that values of K near unity give a good deal better pull-in performance at any noise bandwidth than values of K differing substantially from unity. The parameter K provides a second relationship between the AC gain mf_c, and the shunt time constant xT, which have previously been shown to determine the noise band. K is $\pi/2$ times the product of mf_c and xT.

VI. Pull-In Control Effect

Figure 6 relates to the mechanism of pull-in for the standard APC loop.

Figure 6a shows the beatnote waveforms appearing at the output of the phase detector when the system is out-of-sync, for two values of oscillator detuning, Δf. The beatnote period in each case is one horizontal unit.

Figure 6b shows the DC component or average value of the beatnote. This DC component is fed back through the resistance divider and causes the beatnote frequency to be less than Δf. When $\Delta f < mf_c$ a voltage is generated within the first cycle of beatnote which is large enough to cause pull-in. When Δf exceeds mf_c a DC component is generated which varies inversely with Δf as shown.

The DC component charges the capacitor through the long time constant filter. This capacitor bias then reduces the effective value of Δf. This in turn shifts the average frequency difference, and the voltage toward which the capacitor is charging. If the DC loop gain is high enough, pull-in will eventually occur.

The gated nature of the burst sync signal limits the pull-in range for this circuit to one-half line frequency.

Figure 6c shows a simplified expression for the pull-in time which holds when the pull-in is limited by loop gain. The frequency pull-in time T_F is proportional to the product of the shunt time constant xT with the fraction shown. The numerator $(\Delta f)^2/(mf_c)^2$ defines the performance when only a small fraction of the pull-in range is used. The denominator contains a term which is the square of the ratio of Δf to Δf_{max}. The denominator goes to zero at the limit of the pull-in range, where T_F becomes infinite.

This equation can be written in several forms which are useful for design work. It is a good approximation to the curves shown on the next slide.

VII. Pull-In Characteristics

Figure 7 shows two graphs which relate pull-in time to the frequency difference Δf over which the oscillator must be pulled. Both the ordinate and the abscissa are normalized in terms of noise bandwidth. Introduction of the noise bandwidth makes the curves be functions of the two parameters m and K. The left hand curve shows performance for K = 1/2 and values of m of .1, .01, and .001 respectively. In this case, where K is constant, the DC loop gain f_c varies inversely with m. Thus, increasing the DC loop gain decreases the pull-in time represented by the approximate limit curve shown as a dotted line. The limit curve represents a square law relationship. The value of K = 1/2 was chosen because it makes this limit curve have its lowest value.

Figure 7b shows the performance for three values of K, with m constant at .01. The curve for K = 1/2 is seen to have the lowest linear position representing the limit curve. The curve for K = 1/2 requires longer pull-in times and has less pull-in range. The curve for K = 4 requires a longer pull-in time than the curve for K = 1/2 for small frequency deviations but has almost 50% more pull-in range than the curve for K = 1/2. These curves suggest that there is an optimum value for K.

VIII. Optimum Value of K

Figure 8 shows the optimum value of K. Figure 8a, on the left, represents relative pull-in range as a function of K for constant values of the noise bandwidth and the DC loop gain, f_c. For these values of f_{NN} and f_c the limit to the pull-in range is obtained with high values of K. For K = 1, 90% of this range is obtained, while for K = .1, about one-quarter of this range is obtained.

The pull-in range is not the real requirement on pull-in performance. The real requirement is that for some practical selected time interval, possibly not exceeding a few seconds under worst conditions, pull-in shall occur for as wide as possible a frequency range.

Figure 8b, on the right, shows a clear optimum design in this regard. The parameter λ^2 which is plotted against K is equal to the pull-in time T_F in seconds, divided by the ratio of Δf^2 to f_{NN}^3. For example if the oscillator detuning is Δf = 1 Kc and the noise bandwidth is 100 cycles, then λ^2 equals pull-in time in seconds. For this case pull-in time cannot occur in less than 4 seconds.

This limit curve applies when only a small fraction of the pull-in range is used. The other curves show what happens when a large fraction of the maximum pull-in range permitted by the DC loop gain is used. The limit curve has its lowest value at K = 1/2. However, when three-fourths of the relative pull-in range is used the curve moves up and has a minimum near K = 1. This means that for any pull-in time T_F, there will be much more pull-in range for values of K near 1 than for values of K not near from 1. Tests have been run in which K alone was varied. These tests show a clear optimum where these curves indicate it.

IX. Design Chart

Figure 9 shows a useful design chart. The chart can be used to select the best design consistent with any degree of oscillator stability. The relations between Δf and f_c are plotted for several different conditions.

Lines of constant slope passing through the origin represent constant values of phase angle, $\Delta\phi$, after pull-in from Δf in a circuit having some specified f_c. Values of $\Delta\phi$ equal to 10° and 5° are shown.

The three other curves shown indicate pull-in performance for a noise bandwidth of 100 cycles and K equal to 1. The curves, from top down, represent the maximum pull-in range, and the pull-in range obtainable for five seconds or one second of pull-in time. It can be seen that if five-second pull-in time is required at 100 cps noise bandwidth with this circuit, the oscillator stability including the effect of the reactance tube must be held within ± 1000 cycles.

Subjective tests have been run with weak signals using designs having varied noise bandwidths. Noise bandwidths up to 400 cps still provide good thermal noise immunity. A DC loop gain of fc = 30 to 40 Kc seems desirable. An APC loop of this type, with a noise bandwidth of 400 cycles, can pull in from ± 2-1/2 kilocycles in one second under the optimum design condition of K = 1.

The point at which color sync noise is masked by video noise is near 100 cps. The preferred noise bandwidth is 100 cps.

Conclusion

The analysis of operation and design conditions presented for the standard APC loop shows it can be designed to meet the requirements of high quality automatic color synchronization.

Acknowledgment

The valuable assistance of Mr. Honer E. Webb and Mr. J. R. White on computations and experimental work is gratefully acknowledged.

Bibliography

The mathematical derivations of formulae presented in this paper appear in Technical Monograph #7 of the National Television System Committee. This monograph will appear in the IRE Proceedings later this year.

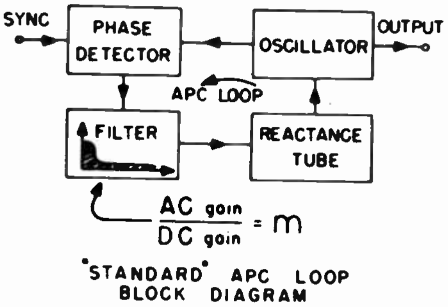


Fig. 1

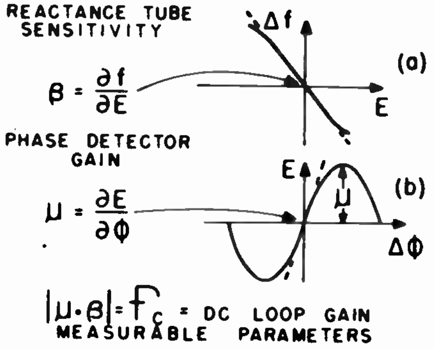


Fig. 2

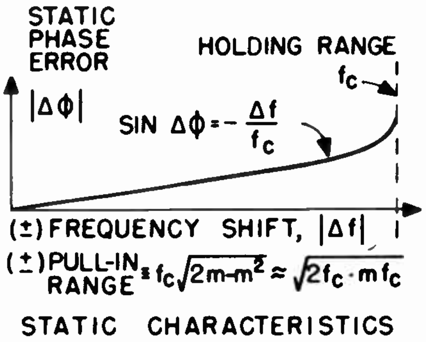


Fig. 3

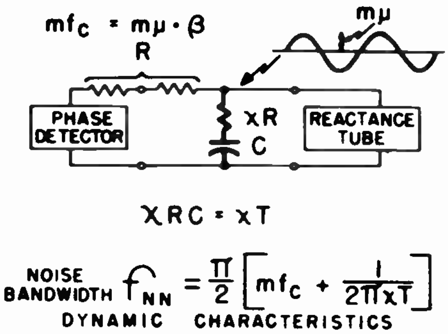


Fig. 4

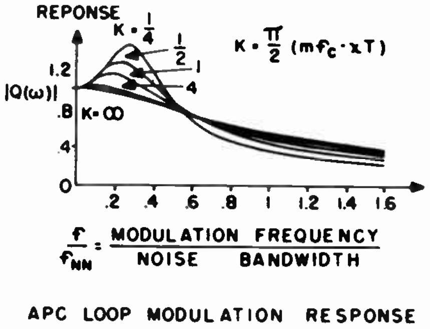


Fig. 5

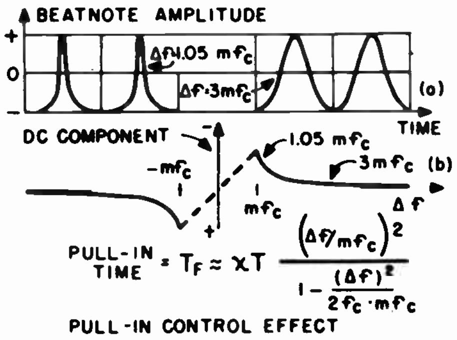
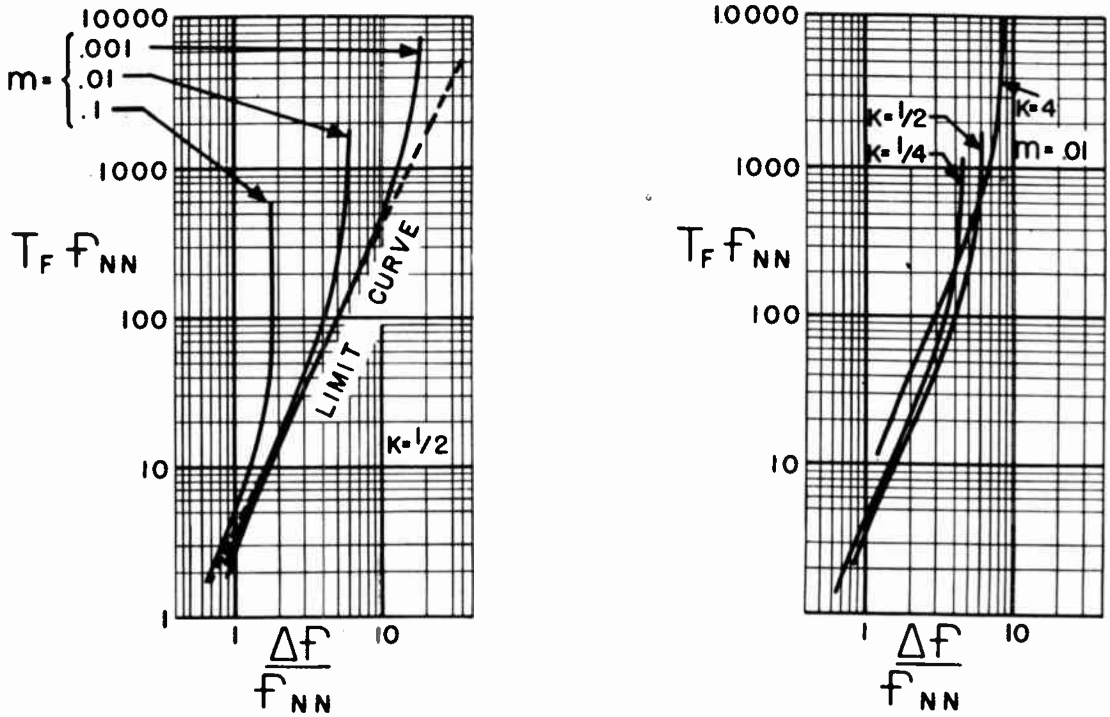


Fig. 6



(a)

(b)

Fig. 7

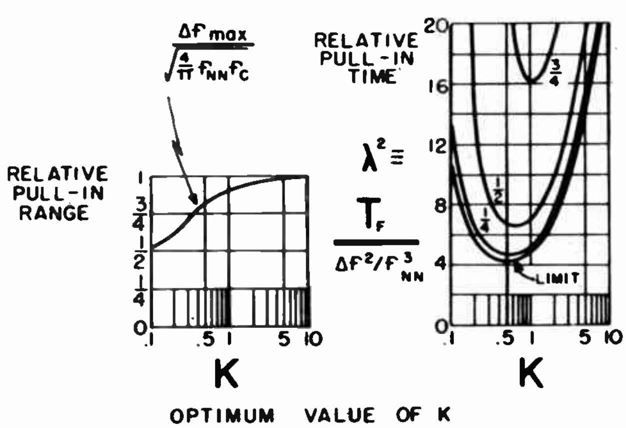


Fig. 8

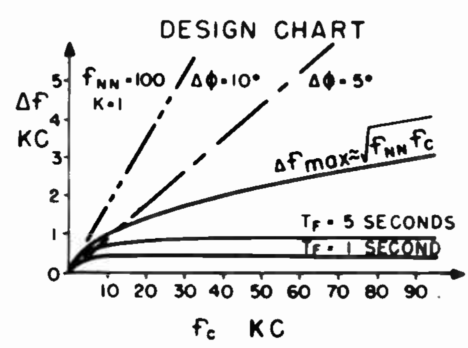


Fig. 9

TRANSIENT RESPONSE IN A COLOR CARRIER CHANNEL WITH VESTIGIAL SIDE BAND TRANSMISSION

J.S.S. Kerr
Electronics Laboratory
General Electric Company
Syracuse, New York

Introduction

In the proposed color television standards currently being field tested by members of the National Television Systems Committee, the color signal consists of two parts. One of these parts is the luminance information (this is the information presented in a monochrome receiver), and the other part of the color signal is the chrominance information. The chrominance information consists of two independent signals which modulate a sub-carrier in quadrature. This modulated sub-carrier is added to the luminance signal at video, and then the composite signal modulates an r.f. carrier. If the transfer functions which operate on the signal between sub-carrier modulation and sub-carrier detection satisfy certain conditions, the two original chrominance signals may be recovered independently.

In this paper, attention will be restricted to considerations regarding the transfer of the chrominance information to the receiver. The results are based on the equivalent video filter characteristics which Nyquist derived in one of his papers[1].

Derivation of Equivalent Video Transfer Characteristics

Let us examine first the process of modulation in quadrature and passing the resultant time function through a network with a transfer function which does not alter the side band structure, and second, passing the same time function through a network with a transfer function which does alter the side band structure. In general, a carrier, $\cos \omega_c t$, may be modulated by a video function $f(t)$, and then the resulting function synchronously detected by multiplying it with an auxiliary carrier, $\cos \omega_c t$. After the detected signal has passed through a low pass filter which removes the double carrier-frequency terms, only the original function of time, $f(t)$, will remain. If this same modulated $\cos \omega_c t$ carrier is synchronously detected with an auxiliary $\sin \omega_c t$ carrier and passed through the above low pass filter, there will be no signal output. Thus it appears possible to modulate a carrier in quadrature with two independent signals and recover them without cross talk. However, the foregoing assumes that the modulated carrier has passed through a network with a transfer function which will not affect the side bands of the carrier. This is somewhat idealized, since transmitted signals are always passed through filters which affect the side bands both after modulation and also before detection. How does this change the situation?

Consider a case where the side bands are affected by a transfer function. Let a carrier, $\cos \omega_c t$, be modulated by a function of time $f(t)$ with a complex video spectrum $K(\omega)$ with highest frequency ω_h. Let the resulting spectrum be $G(\omega)$. Then

$$G(\omega_c+\omega) = \overline{G(\omega_c-\omega)} = K(\omega) \qquad (1)$$

where the bar on top denotes complex conjugate. Now let this spectrum, $G(\omega)$, be operated on by the transfer function $g(\omega)$ through which the modulated carrier passes before detection. Then the resulting real function of time before detection will be given by

$$\begin{aligned} F(t) &= \mathrm{Re}\,\frac{1}{2\pi}\int_{\omega_c-\omega_h}^{\omega_c+\omega_h} G(\omega)g(\omega)\exp(j\omega t)d\omega \\ &= \cos \omega_c t \left\{\mathrm{Re}\,\frac{1}{2\pi}\int_0^{\omega_h} K(\omega)\left[g(\omega_c+\omega) + \overline{g(\omega_c-\omega)}\right]\exp(j\omega t)\,d\omega\right\} \\ &\quad + \sin\omega_c t \left\{\mathrm{Re}\,\frac{1}{2\pi}\int_0^{\omega_h} K(\omega)j\left[g(\omega_c+\omega) - \overline{g(\omega_c-\omega)}\right]\exp(j\omega t)\,d\omega\right\} \end{aligned} \qquad (2)$$

See Figure 1 for a schematic diagram of the above process. If this signal were synchronously detected with an auxiliary carrier, $\cos \omega_c t$, and passed through a low pass filter, the final result would be the expression in the first set of $\{\ \}$. This same result would have been obtained if the original video signal $f(t)$ had been passed through a video filter with a transfer characteristic $R_I(\omega)$ where

$$R_I(\omega) = \left[g(\omega_c+\omega) + \overline{g(\omega_c-\omega)}\right]. \qquad (3)$$

Also if $F(t)$ is synchronously detected with an auxiliary carrier, $\sin \omega_c t$, and the result passed through a low pass filter, in general, there will be some video output although originally only a cosine carrier was modulated. This detected signal will be referred to as the quadrature component. A video signal identical to the quadrature component would have been obtained if the original video signal $f(t)$ had been passed through a video network with a transfer characteristic $R_Q(\omega)$ where

$$R_Q(\omega) = j\left[g(\omega_c+\omega) - \overline{g(\omega_c-\omega)}\right] \qquad (4)$$

Thus the process of modulation, operation on the modulated carrier, and synchronous detection in quadrature may be replaced by a video network of the form shown in Figure 2. (To be perfectly general, synchronous detection of $F(t)$ with a carrier $\cos(\omega_c t - \theta)$ is equivalent to passing the original video function of time $f(t)$ through a video network with a transfer function $R(\omega)$ where

$$\begin{aligned} R(\omega) &= \left[g(\omega_c+\omega)+\overline{g(\omega_c-\omega)}\right]\cos\theta \\ &\quad + j\left[g(\omega_c+\omega)-\overline{g(\omega_c-\omega)}\right]\sin\theta . \end{aligned} \qquad (5)$$

Except for notation, this is the same result as that obtained by Nyquist in an article on telegraph transmission theory[1]. The equivalence of results (3) and (4) with those of Nyquist is demonstrated in the Appendix*.)

Notice that the lattice structure shown in Figure 2 is not one of the conventional four terminal lattice networks widely used in circuit theory. The connecting lines are channels. Also the restriction is imposed, that the networks in the boxes are unilateral.

If no cross talk is to appear between the two channels shown in Figure 2 for a particular video frequency ω_1, then it is necessary that the transfer function of the quadrature component network ($R_Q(\omega)$) be zero at $\omega = \omega_1$. Therefore,

$$g(\omega_c + \omega_1) = \overline{g(\omega_c - \omega_1)} \tag{6}$$

If no cross talk is to occur for video frequencies from d.c. up to, say, ω_ℓ, then the requirement that $g(\omega_c + \omega) = \overline{g(\omega_c - \omega)}$ for $0 \leq \omega \leq \omega_\ell$ is equivalent to specifying that the amplitude and phase of the transfer characteristic must be even and odd respectively about the carrier out to a frequency deviation of ω_ℓ.

The in-phase and quadrature video response to any frequency ω may be visualized by means of vector diagrams such as those given in Figure 3. The amplitudes of the component vectors are $|g(\omega_c + \omega)|$ and $|g(\omega_c - \omega)|$. Since the conjugate of the transfer function for the lower sideband is required, its phase angle is measured off in a clockwise direction. If the amplitude of the transfer characteristic is not even about the carrier, then the amplitude of the quadrature will be least when the two vectors subtract algebraically from each other. This occurs when the phase is odd about the carrier. It might be noted that this condition (odd phase) also gives the largest in-phase amplitude. The largest in-phase amplitude and the smallest quadrature amplitude (occurring when the phase about the carrier is odd) are indicated on the R_I and R_Q vectors respectively in Figure 3.

Analysis Using Equivalent Video Transfer Characteristics

The in-phase and quadrature response to a step input has been analyzed for a few different vestigial side band transfer characteristics using video transfer characteristics determined by the method of the previous discussion. Video low pass networks were included in all calculations. (At this point it should be mentioned that in systems proposed by the NTSC the quadrature response is eliminated either by Color Phase Alternation or by proper video filtering.) In the calculations vestigial side band filter amplitude characteristics were assumed which might be approached in the laboratory. Then the phase response of these filters was determined assuming that the networks being used were of the minimum phase shift type[2]. The minimum phase of the video networks was calculated from the analytical expressions for the networks. Then using the expressions developed earlier the equivalent video in-phase and quadrature transfer characteristics were determined. This was done twice; first for linear phase, and second for minimum phase.

Figure 4 shows the first transfer characteristics which were analyzed. The two transfer characteristics a and c on left are for the same vestigial side band filter which is assumed to be flat below 3 mc. This VSB filter has a gradual linear slope about a 3.89 mc sub-carrier and is down 6 db at the sub-carrier. Since this analysis was carried out, the proposed sub-carrier position has been moved down to 3.58 mc. The transfer characteristics b and d on the right are those of the video low pass filters. Transfer characteristic b is that of a two-stage RC low pass filter down 6 db at .5 mc, while transfer characteristic d is that of a two-pole Butterworth low pass filter down 6 db at 1 mc.

Figure 5 shows the equivalent video in-phase and quadrature transfer characteristics of the vestigial side band filter plus the low pass video filters shown in Figure 4. The transfer characteristics on the left (a and c) result from a linear phase analysis while those on the right (b and d) are the result of a minimum phase analysis. The characteristics at the top are for the two-stage RC video filter, while those at the bottom are for the two-pole Butterworth case. It should be noted that in all cases the peak of the quadrature amplitude characteristic, $|R_Q|$, for minimum phase shift characteristics is larger and occurs at slightly lower frequencies than for linear phase shift characteristics. Also the in-phase amplitude characteristic, $|R_I|$, for the minimum phase case drops below that of the linear phase case (this is more noticeable in Figure 5 in the two-pole Butterworth case between .25 and .5 mc). Also notice that the quadrature amplitude characteristic for the two-stage RC case is much lower than that for the Butterworth case.

Figure 6 shows the transients obtained by putting a unit step through the equivalent video transfer characteristics shown in Figure 5. Again the curves on the right are for minimum phase, while those on the left are for linear phase; the curves on the top are for the two-stage RC case while those on the bottom are for the two-pole Butterworth case. It may be seen that the low cutoff video filter (two-stage RC) has a smaller quadrature component than the higher cutoff video filter. This might have been predicted by the quadrature amplitude characteristics of Figure 5 as pointed out earlier. Unfortunately, the use of a lower cutoff video filter to reduce the quadrature component has slowed the rise time of the in-phase transient considerably. While the amplitude of the quadrature component has not changed much in going from the linear phase case to the minimum phase case, the shape of the in-phase transient has been materially affected.

Having seen that changing the video cutoff to reduce the quadrature component is expensive from the point of view of rise time, consideration will now be given to changing the shape of the vestigial side band filter while keeping a two-pole

---

* Alternatively the intermediate steps of result (2) may be derived from first principles.

Butterworth network for the video low pass filter. Figure 7 shows some different shaped vestigial side band filter amplitude characteristics. The VSB filter shown in the top left is flat from below 3 mc up to and beyond the sub-carrier. A video boost network shown in the top right, b, of Figure 7 is used in conjunction with this so that the resulting equivalent video in-phase transfer characteristic will be flat from dc up in the linear phase case (before multiplying with the Butterworth characteristic). Two different step VSB filter characteristics are shown in the lower left, c, of Figure 7. First consider the one with the narrower step about the subcarrier. This transfer characteristic has been so designed that it gives the same equivalent video in-phase and quadrature transfer characteristics as the flat VSB filter with the video boost in the linear phase cases. The wider step has been put in to reduce the quadrature component if possible. Certainly in the linear phase case, the amplitude of the quadrature transfer characteristic will be less than that of the narrow step. However, the question arises as to how much the effect of phase distortion will be increased when one has to make the cutoff sharper because of the wider step.

It was found that there was an almost negligible difference between the equivalent video transfer characteristics of the narrow step and the flat VSB filter with the video boost for the minimum phase case even with the large percentage bandwidth involved (they were designed to be equal for linear phase). Therefore the transfer characteristics shown at the top of Figure 8 may be considered to be for either of the two VSB filters just mentioned. The transfer characteristics at the bottom of Figure 8 are those for the wide step. As before, the characteristics for linear phase are shown on the left, while those for minimum phase are shown on the right.

Notice in the linear phase cases that the quadrature transfer characteristics are almost zero for frequencies below 200 kc. and that the wide step has a lower quadrature characteristic out to the frequency at which the signal is being transmitted single side band. The in-phase characteristics in the case of linear phase are the same as for the first Butterworth network shown. In the minimum phase cases the effect of phase distortion about the subcarrier on the equivalent video transfer characteristics has now become more important. The quadrature characteristics start to leave zero quite rapidly near zero frequency, and the in-phase characteristics have dips in them. Notice that now the amplitudes of the quadrature characteristics are roughly the same for minimum phase so that perhaps not very much has been gained by using the wider step since then the VSB filter has to be brought down that much faster with resultant increased phase distortion. In addition, the dip in the in-phase characteristic is much more severe in the case of the wide step, so its in-phase transient will probably be worse than that of the narrow step. (Actually, the cutoff of the wide step was made even sharper than necessary at small amplitude levels in order to exaggerate the phase effects.)

The transients shown in Figure 9 bear out the above analysis. In the linear phase cases the wider step has less quadrature than the narrow step (or flat VSB filter) while the narrow step, in turn, has less quadrature than the sloping VSB filter with the same video filter shown in Figures 4, 5, and 6. In the minimum phase cases shown in Figure 9, the quadratures are roughly the same, but the in-phase transient of the sharp cutoff (wide step) VSB filter has deteriorated considerably. The shape of the transient above .6 units will give a picture with "smear."

Conclusions

Some conclusions to be drawn from the previous examples follow. Using a lower cutoff video filter will reduce the quadrature, but it does not appear to be practical because of the resultant increased rise time. Apparently it makes little difference whether one uses a step VSB filter or a flat VSB filter with a video boost in the minimum phase case, although large percentage bandwidths are involved. A step VSB filter gives less quadrature component than a sloping VSB filter. In the minimum phase case the size of the step should be a compromise between the effect of the difference of the amplitudes of the transfer function about the carrier, and that of phase distortion due to the cutoff.

Closing Comments

In the first field test specifications of the NTSC of November 26, 1951, the unwanted quadrature components were cancelled out over a frame by the use of Color Phase Alternation (CPA) which reverses the polarity of the quadrature transient with respect to the in-phase transient from one field to the next. The present modified NTSC proposals of January 15, 1953 eliminate the quadrature components by proper filtering. The subcarrier frequency has been lowered to 3.58 mc. and advantage is taken of the decreased acuity of the eye along certain directions in the chrominance plane. Thus one chrominance signal is transmitted double side band while the other signal has a higher video cutoff and is transmitted vestigial side band. The double side band signal contributes no quadrature component to the vestigial side band channel if the phase angle is made odd about the subcarrier. Although the vestigial side band signal introduces a quadrature component into the double side band channel, it only consists of frequency components above the cutoff of the video filter in the double side band channel and hence is rejected.

Acknowledgement

The author would like to acknowledge suggestions from P. W. Howells on the shapes of the VSB filters which were analyzed.

References

1. H. Nyquist: "Telegraph Transmission Theory" Transactions of the A.I.E.E. February, 1928, pp. 636 - 639

2. D. E. Thomas: "Tables of Phase Associated with a Semi-Infinite Unit Slope of Attenuation" Bell System Technical Journal, Volume XXVI, October, 1947 p. 870

Appendix I

Equivalence of Video Characteristics With Those of Nyquist

This appendix demonstrates the equivalence of the video transfer characteristics developed by Nyquist and those used in this paper. Nyquist works with the complex function $Y(\Omega)$ where $\Omega = \omega - \omega_c$ is a video frequency. He splits the symmetrical and skew symmetrical parts of $Y(\Omega)$ into real and imaginary components so that by definition

$$\tfrac{1}{2}\left[Y(\Omega) + Y(-\Omega)\right] = g_1(\Omega) + jb_1(\Omega)$$

$$\tfrac{1}{2}\left[Y(\Omega) - Y(-\Omega)\right] = g_2(\Omega) + jb_2(\Omega)$$

The present paper works with the complex transfer function $g(\omega)$ throughout and, incidentally, measures the argument from the frequency origin instead of from the carrier frequency ω_c. Thus

$$g(\omega) = g(\omega_c + \Omega) = Y(\Omega)$$

The equivalent video transfer characteristic given by Nyquist for cosine detection is

$$\begin{aligned} g_1(\Omega)+jb_2(\Omega) &= \tfrac{1}{2}\Big\{ \mathrm{Re}\left[Y(\Omega)+Y(-\Omega)\right] \\ &\qquad + j\mathrm{Im}\left[Y(\Omega)-Y(-\Omega)\right]\Big\} \\ &= \tfrac{1}{2}\left[Y(\Omega)+\overline{Y(-\Omega)}\right] \\ &= \tfrac{1}{2}\left[g(\omega_c+\Omega)+\overline{g(\omega_c-\Omega)}\right]. \end{aligned}$$

The equivalent video transfer characteristic given by Nyquist for sine detection is

$$\begin{aligned} -b_1(\Omega)+jg_2(\Omega) &= j\left[g_2(\Omega)+jb_1(\Omega)\right] \\ &= \tfrac{j}{2}\Big\{ \mathrm{Re}\left[Y(\Omega)-Y(-\Omega)\right] \\ &\qquad + j\mathrm{Im}\left[Y(\Omega)+Y(-\Omega)\right]\Big\} \\ &= \tfrac{j}{2}\left[Y(\Omega)-\overline{Y(-\Omega)}\right] \\ &= \tfrac{j}{2}\left[g(\omega_c+\Omega)-\overline{g(\omega_c-\Omega)}\right]. \end{aligned}$$

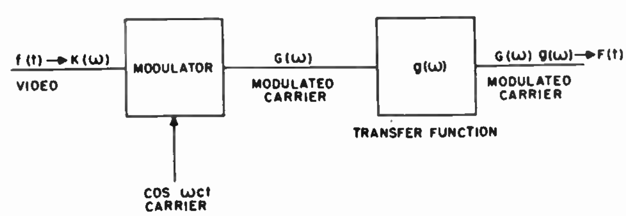
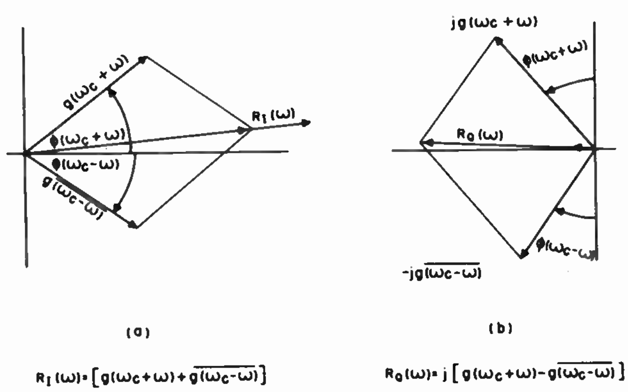


Fig. 1 - Schematic diagram showing signal path.



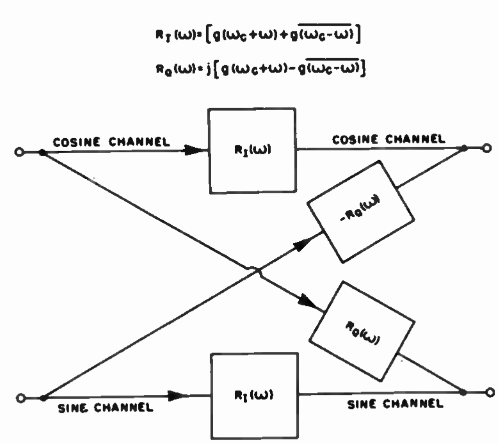


Fig. 2
Equivalent video lattice network.

Fig. 3
Vector representation of in-phase and quadrature transfer functions.

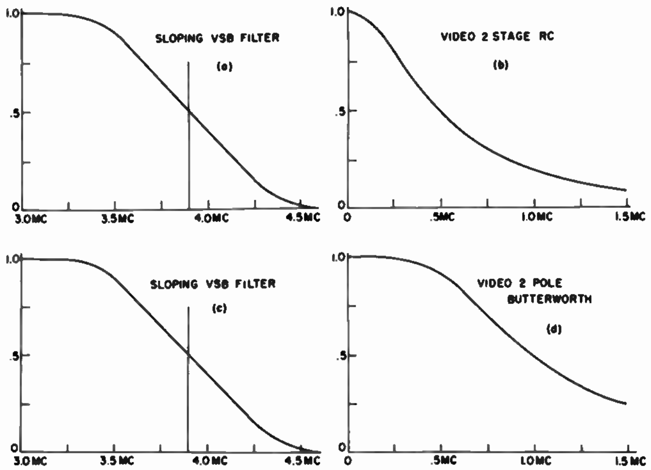


Fig. 4
Sloping VSB filter with different
video low-pass filters.

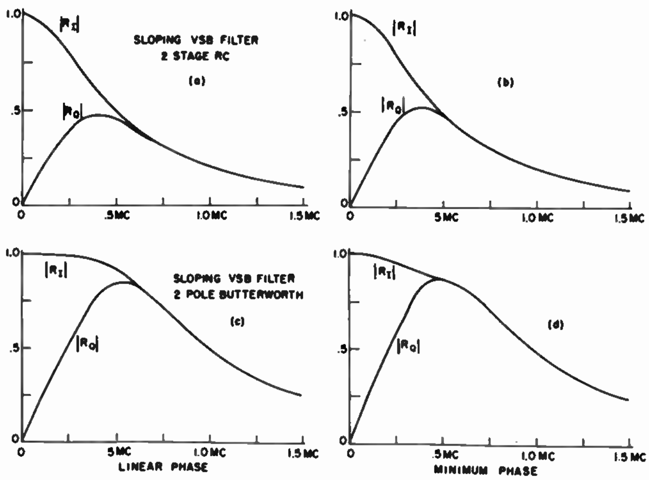
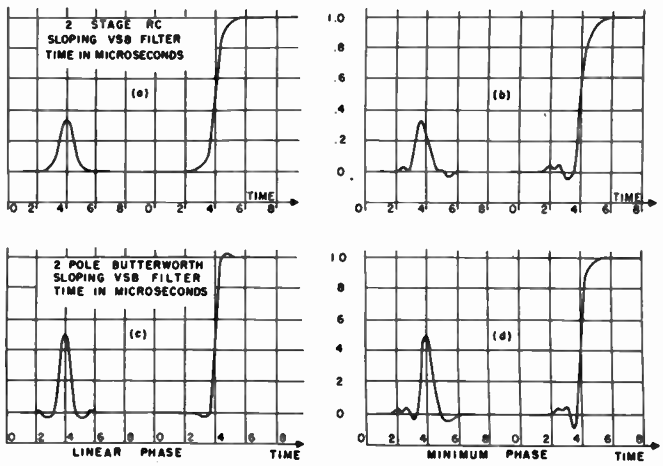


Fig. 5
Equivalent video transfer characteristics for sloping
VSB filter with different video low-pass filters.

Fig. 6
Transient responses to a step input
for sloping VSB filter with
different video low-pass filters.



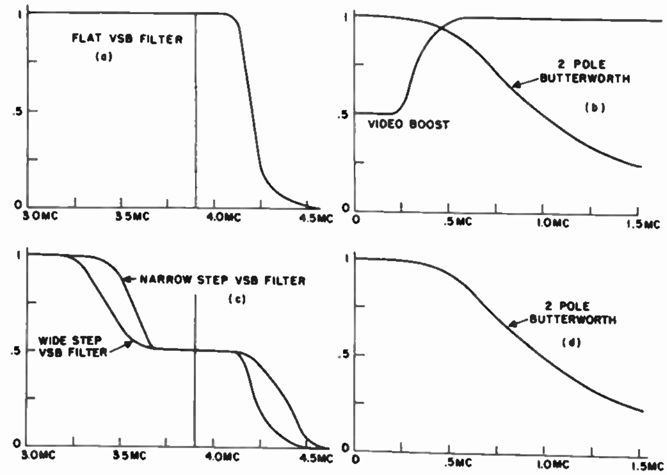


Fig. 7
Different VSB filters with a 2-pole Butterworth video low-pass filter.

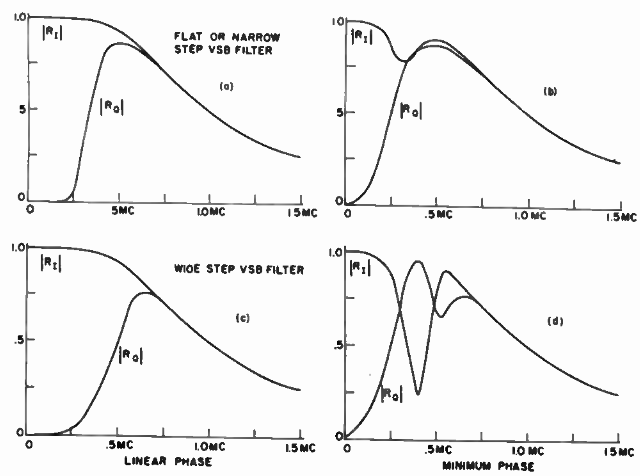
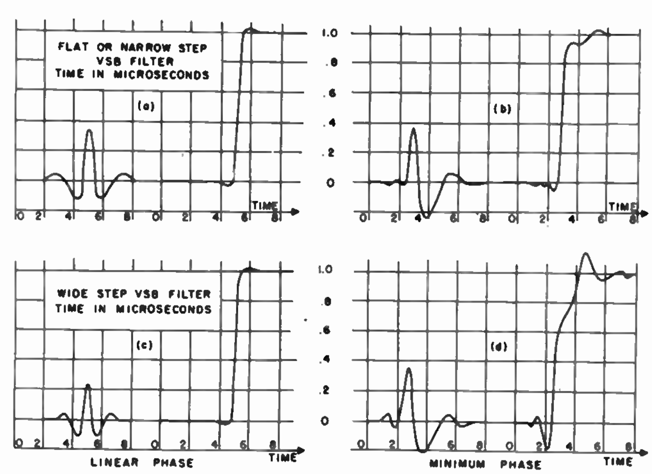


Fig. 8
Equivalent video transfer characteristics for different VSB filters with a 2-pole Butterworth video low-pass filter.

Fig. 9
Transient responses to a step input for different VSB filters with a 2-pole Butterworth low-pass filter.



TRANSIENTS IN COLOR TELEVISION

Paul W. Howells
Electronics Laboratory
General Electric Company
Syracuse, New York

Introduction

The main difference between monochrome and color television lies in the amount of information needed to specify an element of the picture being transmitted. In monochrome transmissions, the picture element is completely specified when we know its luminance. In color transmissions, the picture element is completely specified when we know three quantities, such as the amounts of three primary lights required to match its color. Thus, we may say that in monochrome television the information transmitted is the location of each picture element on a one-dimensional luminance scale while in color the information transmitted is the location of each picture element in a three-dimensional color space. (These statements ignore the information, carried by the line and field sync signals, about the spatial location of a picture element in the complete picture.)

In understanding the behavior of a color system, this concept (that the color signal specifies the location of a picture element in color space) is a very useful one. In fact, it does not require a great stretch of the imagination to speak of the location of the electrical signal itself in color space. This idea is respectable because the electrical signals at any point in the system are related to colorimetric quantities through transformations occurring at the transmitter and receiver. To the designer of color equipment, the conversion factor relating a colorimetric quantity to its electrical counterpart is important. For general analysis, however, it is usually not important to distinguish, either by the inclusion of conversion factors or by the use of different symbols, between colorimetric quantities and the electrical signals representing them. This paper will, therefore, use colorimetric symbols for both. It should be clear from context which usage is intended.

Using the concept of electrical signals in color space, the operation of the color system might be described as follows: The camera, in scanning the scene, makes a colorimetric analysis of the picture element under scan and produces three output voltages. Let us say that the camera is so designed that these three voltages represent the amounts of the NTSC Panel 7 primaries, R, G, and B. These voltages determine a point in color space, the camera color point. In the process of coding the color signals for transmission, the R, G, B voltages undergo a linear transformation which produces three related voltages representing the amounts of the NTSC transmission primaries Y, (R-Y), and (B-Y). (Let us ignore gamma precorrection for the time being.) These transmission primary voltages still determine the location of the camera color point, but in terms of a different set of cartesian axes in the RGB color space.

Between the point at which they are formed and where they are recovered in the receiver, the transmission primary signals undergo the usual difficulties with noise and interference, and the different bandwidth limitations necessary for simultaneous transmission through the 6 mc. channel. At the receiver, an inverse transformation is performed on the transmission primary voltages to regain the Panel 7 primary voltages R, G, B, for application to the tricolor display.

The transmission primary signals (or the R, G, B signals) at the receiver determine the location of the receiver color point in color space. For perfect reproduction, this receiver color point should coincide with the camera color point. In practice, the system should be so designed that the colorimetric deviations of the receiver color point from the camera color point due to noise, interference, and band limiting are such as to produce the least perceptible effect in the picture.

The Color Transient

As the color camera scans the scene, the scanning aperture encounters areas of different color, and the camera color point responds by moving about through the color space. As long as its motions are not too rapid, the receiver color point is able to follow them exactly. However, when the scanning aperture crosses a boundary between areas of different color, the camera color point may change position too rapidly for the receiver color point to follow. Since the transmission primary signals have different bandwidths, the path taken by the receiver color point in response to this shock is a complex one. Being able to move more rapidly in some directions than in others through the color space, it traces out a curving path having several more or less abrupt changes in direction. This three-dimensional path through color space, plotted as a function of time, may be called the transient response of the color system. It is analogous to the curve of luminance versus time which represents the overall transient response of a monochrome system and, in a similar way, its shape affects the appearance of transitions in the color picture.

Given the transient responses of the individual signal channels, it is the purpose of this paper to develop means of determining the color transient response of the system. Such a four-dimensional figure is difficult to display as a whole, so, for convenience, it will be analyzed into its component luminance and chromaticity transients.

Red-Green-Blue Color Space

The color transient response of the system currently being field tested by the NTSC is most easily visualized in an RGB color space, i.e., one having cartesian coordinate axes representing the amounts of the red, green, and blue primaries. This is true because in this system the signals represent either these primaries, or colorimetric quantities which are simply related to them. Pertinent characteristics of this color space are: first, all color points of the same luminance lie in a plane, and second, all points of the same chromaticity lie on a line passing through the origin. The unit amounts of the primaries are so chosen that an equal number of units of the three primaries are required to form an Illuminant C white. In this space, then, the illuminant C vector is the line making equal angles with the R, G, B, coordinate axes. Since unit amounts of the Panel 7 R, G, B, primaries have luminances of 0.299, 0.587, and 0.114 respectively, the total luminance of a color mixture is given by the expression:

$$Y = .299\,R + .587\,G + .114\,B \qquad (1)$$

Planes of constant luminance are obtained by setting Y equal to a constant in this expression. Such a plane of constant luminance in RGB space is shown in Figure 1.

NTSC Transmission Primaries

For present tricolor displays, a color is most conveniently described in terms of the red, green, and blue primary lights needed to produce it. However, a considerably better match between the transmission system capabilities and the requirements of the observer's eye is achieved when the color signal as transmitted is resolved along axes corresponding to primaries other than R, G, and B. For this purpose, the color signal is resolved into a luminance and a chrominance component. (Chrominance is a two-dimensional vector quantity which affects chromaticity but not luminance. It is defined as the colorimetric difference between the color in question and an illuminant C white of the same luminance.) Chrominance information is transmitted by the color subcarrier, the magnitude of the subcarrier being related to the length and its phase to the direction of the chrominance vector[1]. The chrominance vector does not produce amplitude and phase modulation of the color subcarrier directly; instead it is resolved into components lying along two "color difference" axes to obtain color difference signals (R-Y) and (B-Y) which independently modulate the color subcarrier in quadrature.

Figure 1 shows the location of the NTSC transmission primary axes in RGB color space. The luminance (Y) axis lies along the illuminant C vector, i.e., it is utilized at the receiver to produce an illuminant C picture of the proper luminance. The chrominance vector and the two color-difference axes, (R-Y) and (B-Y), lie in a plane of constant luminance in the RGB space and have their origin on the illuminant C vector. This means that when the color being transmitted is illuminant C, the chrominance signal is zero. When the color is other than illuminant C, the luminance signal establishes an illuminant C picture of the correct luminance, and addition of the chrominance signal moves the color point from illuminant C to the correct chromaticity without changing the luminance.

From considerations of noise and maximum demand, the two color difference signals are not given equal weight. In terms of actual voltage of the subcarrier produced, the two components of the chrominance signal are:

$$S_R = \frac{R - Y}{1.14} \qquad (2)$$

and

$$S_B = \frac{B - Y}{2.03} \qquad (3)$$

Transformation of color data from the RGB to the NTSC coordinate system is performed by the linear matrices:

$$\begin{bmatrix} S_R \\ Y \\ S_B \end{bmatrix} = \begin{bmatrix} .615 & -.515 & -.100 \\ .299 & .587 & .114 \\ -.147 & -.289 & +.436 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad (4)$$

and

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.14 & 1 & 0 \\ -.581 & 1 & -.394 \\ 0 & 1 & 2.03 \end{bmatrix} \begin{bmatrix} S_R \\ Y \\ S_B \end{bmatrix} \qquad (5)$$

The theory of such color coordinate transformations has been set forth in some detail by Bingley[1].

Subcarrier Transients

An important part of the overall transient response of the color system is the transient response of the color subcarrier itself. Two types of transmission will be considered:

(1) The case where both components of the chrominance signal are transmitted in a vestigial sideband fashion with the result that the quadrature component of the S_R signal is detected in the S_B channel, and vice versa.

(2) The case where the color difference signal requiring the least bandwidth for good subjective quality of the image is restricted in bandwidth so that it may be transmitted as a double-sideband signal. Quadrature components in both color difference signals may be eliminated by this scheme, since both signals are double-sideband over the frequency band they share[2]. [This system has been called the acuity matching system, since the bandwidths of the three color signals are so proportioned as to match the acuity of the eye.]

Subcarrier Transient Path - Case 1

Let us assume that at a color transient the camera color point changes its position in the subcarrier (S_R, S_B) plane, moving H units along a line making the angle θ with the S_B axis. This change in position is so rapid that for the subcarrier channel it amounts to a perfect step input. The center of the step is located at point (C_R, C_B). A sketch of this transient is shown in Figure 2(a). The signals applied to the S_R and S_B channels are the components of this step resolved along the S_R and S_B axes, i.e., $H \sin \theta$ and $H \cos \theta$. As has been shown by Kerr[3], these signals are detected at the receiver as though they had been passed through the network shown in Figure 2(b). Since the two channels are identical in this case, we may call $I(t)$ the in-phase unit step response of either channel and $Q(t)$ the quadrature unit step response. The two signals detected at the receiver are therefore:

$$S_R = C_R + H \sin \theta\, I(t) + H \cos \theta\, Q(t) \qquad (6)$$

$$S_B = C_B + H \cos \theta\, I(t) - H \sin \theta\, Q(t) \qquad (7)$$

Equations (6) and (7) are the parametric equations for the transient path of the receiver color point in the subcarrier plane. These equations are simplified if we translate the axes to the center of the transient and then rotate them θ° to line them up with the direction of the transient. Translation of the axes to the point (C_R, C_B) changes equations (6) and (7) to:

$$S_R' = H \sin \theta\, I(t) + H \cos \theta\, Q(t) \qquad (8)$$

$$S_B' = H \cos \theta\, I(t) - H \sin \theta\, Q(t) \qquad (9)$$

To rotate the axes by θ°, we use the transformation equations:

$$S_R'' = S_R' \cos \theta - S_B' \sin \theta$$

$$S_B'' = S_R' \sin \theta + S_B' \cos \theta$$

which reduce equations (8) and (9) to

$$S_R'' = H\, Q(t) \qquad (10)$$

$$S_B'' = H\, I(t) \qquad (11)$$

Equations (10) and (11) show that the shape of this subcarrier transient path at the receiver is always the same, regardless of its angle or the location of its center point. The position of the receiver color point in the direction of the transient is given by $H\, I(t)$ while its excursion to the side is given by $H\, Q(t)$.

Figure 2(c) shows a subcarrier transient for a case analyzed by Kerr[3]. The conditions are 1 mc. bandwidth for each color difference signal, 3.89 mc. subcarrier, step-type vestigial sideband filter, and linear phase. If the transient occurs in the opposite direction, or if color phase alternation is used, the sign of the quadrature component is reversed and the transient path is reversed about the line between its end points, as shown by the dotted lines in Figure 2(d). The timing dots shown on the transient path are 1/8 microsecond apart (approximately the rise time of the luminance transient).

Subcarrier Transient Path - Case 2

In the acuity matching system, currently being field tested by the NTSC, the color subcarrier frequency is reduced to 3.58 mc. and one of the color difference signals is restricted to 0.5 mc. bandwidth so that it may be transmitted as a double sideband signal. The other color difference signal is allowed a bandwidth of 1.5 mc. and is transmitted as a vestigial sideband signal. With proper design of the transmission path, quadrature components in both the S_R and S_B signals as detected may be eliminated.

Good results may be obtained by making S_R the vestigial sideband component of the subcarrier. However, it has been found[4] that greater subjective sharpness of some color transients is obtained when the vestigial sideband component (S_V) and the double sideband component (S_D) are located 33° in advance of the phase of the S_R and S_B components of the subcarrier, respectively.

Figure 3 shows a typical shape for the subcarrier transient path of the acuity matching system. The camera color point (Figure 3a) jumps H units along a line making the angle θ with the S_D axis, so the heights of the steps applied to the S_V and S_D channels are $H \sin \theta$ and $H \cos \theta$. Let the unit step responses of the S_V and S_D channels (Figure 3b) be $I_V(t)$ and $I_D(t)$ (Figure 3c). Since there is no quadrature response between the color difference channels in this system, the two signals detected at the receiver, neglecting the DC terms, are simply:

$$S_V = H \sin \theta\, I_V(t) \qquad (12)$$

and

$$S_D = H \cos \theta\, I_D(t)\ . \qquad (13)$$

Figures 3c and 3d show the S-shaped transient path parametrically determined by equations (12) and (13). The S-shape is characteristic of a system having a higher speed of response in one direction than in the other. Unlike the transient path of Case 1, the shape of this transient does depend upon the angle it makes with the S_D axis.

For instance, when $\theta = 0^o$ the path is a straight line parallel to the S_D axis and when $\theta = 90^o$ the path is a straight line parallel to the S_V axis. When θ becomes greater than 90^o, the curve again has an S-shape, but the S is reversed. Reversal of the direction of the transient, however, does not reverse the S-curve.

Complete Color Transient

The complete color transient path is the result of both the chrominance transient response just described and the luminance transient response. If the luminance signal remains constant, the chrominance transient takes place entirely in the corresponding plane of constant luminance. If, however, at some value of chrominance, we vary the luminance, the color point moves along a straight line parallel to the illuminant C vector. This means that the chrominance transient alone specifies, not a line path in RGB space, but a surface generated by moving a line parallel to the illuminant C vector through the chrominance transient path plotted in any plane of constant luminance. The complete color transient follows this surface at a height above the zero luminance plane specified by the luminance signal. Figure 4 shows a typical shape for a color transient path through RGB color space. The transient starts at a somewhat de-saturated green of high luminance (upper luminance plane) and ends at a saturated magenta of lower luminance (lower luminance plane). The subcarrier transient is the S-curve type shown in Figure 3, so the chrominance component of the transient is described by the S-shaped surface. At the start of the transient, the luminance signal remains at its initial value for a time, allowing the color point to follow the subcarrier transient path in the upper luminance plane. Near the center of the transient, the luminance transient begins dropping the color point rapidly in the illuminant C direction to the lower luminance plane. Here it continues its more leisurely progress along the remainder of the subcarrier transient path.

In appearance, this color transient consists of a sharp luminance transition about which are centered two successively less sharp chrominance transitions in the S_V and S_D directions. At normal viewing distance, the subjective sharpness of the complete transition approaches that of the luminance transition alone. Although the S_D component of the chrominance transient is considerably less sharp than either the S_V component or the luminance transient, the resultant chromaticity error at the transition occupies such a small area and has such colorimetric quality that its visibility is extremely low. For normal color pictures, rather close inspection is necessary to distinguish between such a picture and one in which the color signals have been transmitted at full bandwidth.

The Effects of Gamma Precorrection

In the preceding discussion, the assumption has been made that the electrical signals are linearly related to the colorimetric quantities they represent. In a practical system, this is not the case. The color display device has an approximately exponential transfer characteristic with an exponent of γ. The correct primary signals for this tube, then, are not R, G, and B, but

$$R' = R^{1/\gamma}$$

$$G' = G^{1/\gamma}$$

and

$$B' = B^{1/\gamma}$$

(Note: Expressions involving electrical or colorimetric quantities in non-linear relationships may not be dimensionally correct unless supplied with the proper conversion factors. These conversion factors are omitted in the interest of brevity.)

Precorrection of the R, G, and B primary signals for the non-linear characteristic of the cathode ray tube is performed at the transmitter. Subsequent linear operations to convert these gamma corrected signals to the transmission signals and, at the receiver, to convert them back, are performed exactly the same as in the linear case (Equations (4) and (5)). Where the transmission primary signals are formed in this way from gamma corrected primary signals, they are primed to indicate this fact.

It is interesting to see what the subcarrier surface in RGB space becomes when gamma precorrection is introduced. The expression for the luminance signal then is:

$$Y' = .299\,R' + .587\,G' + .114\,B' \qquad (14)$$

A subcarrier surface is the surface obtained by setting the luminance signal, Y', equal to some constant. It is obvious that this is no longer the expression for a luminance plane. At the illuminant C point, this surface is tangent to the luminance plane, but with increasing saturation it tends to depart more and more from this plane in the direction of increasing luminance. A sketch of this subcarrier surface in RGB space is shown in Figure 5. The lines describing the surface are the S_R', S_B' grid. In dotted lines underneath the subcarrier surface is shown the constant luminance plane to which the surface is tangent at illuminant C.

An important effect of this warping of the subcarrier surface is the fact that the transmitted signals are no longer completely separated into luminance and chrominance packages. At illuminant C, and for moderate saturations, the subcarrier does carry a good approximation to chrominance information, but at high degrees of saturation, its effect on luminance may become important, and its effect on chromaticity reduced. This effect of the subcarrier signal on luminance makes the control of subcarrier transients somewhat more important than it would be in a linear system.

Luminance Factor Map of Subcarrier Surface

When gamma precorrection is performed at the transmitter, the S_R', S_B', Y' coordinate system is warped into a set of curved lines running through RGB color space. Determination of the luminance and chromaticity of the receiver color point, given the values of S_R', S_B' and Y', is, therefore, a laborious process. However, it can be simplified by mapping the color space to show the effect of these quantities on luminance and chromaticity. Such maps can be made in terms of the normalized coordinates S_R'/Y' and S_B'/Y.

Figure 6 shows the contribution of the color subcarrier to the luminance of the receiver color point. The subcarrier luminance factor, K_S, is the ratio of the actual (correct) luminance of the receiver color point to that part of its luminance produced by the luminance signal alone. For example, where $K_S = 2$, the luminance signal contributes only half of the total luminance, while the color subcarrier contributes the other half. When the values of K_S and the luminance signal, Y', are known, the luminance of the receiver color point may be determined from the expression, (see Appendix I),

$$Y = K_S \, Y'^{\gamma} \qquad (15)$$

Figure 6 may be considered as a luminance contour map of the subcarrier surface of Figure 5. Assuming that this surface is tangent to the $Y = 1$ plane at illuminant C, the contours for the different values of K_S correspond to its intersections with planes of higher luminance. The value of luminance at each contour is in this case equal to the value of K_S.

The subcarrier luminance factor, K_S, is related to the chromaticity factor, K_c, introduced by Applebaum[5]. One is the inverse of the other. Since they are evaluated in different coordinate systems, however, a derivation of the expression used to evaluate K_S is given in Appendix 1.

Chromaticity Map of Normalized Subcarrier Surface

Figure 7 shows how the coordinate lines of the chromaticity diagram appear in the normalized subcarrier surface. The family of solid lines correspond to constant values of y while the dotted lines represent constant values of x.

The reason for the warping of this chromaticity grid will be better understood if we visualize how it is obtained. In RGB space, we construct the unit plane $X + Y + Z = 1$ and on it draw the grid of lines $x = \text{const}$, $y = \text{const}$. Then we project this grid onto the bowl-shaped subcarrier surface, projecting from the origin. Near the illuminant C point, the subcarrier surface is nearly parallel to the unit plane, so not much distortion of the chromaticity grid results, but as we approach the edges of the color triangle we are projecting to a surface which has turned nearly parallel to the direction of projection. In these regions, the chromaticity grid is greatly expanded.

The expressions relating x and y to S_R'/Y', and S_B'/Y' are derived in Appendix 2.

Evaluation of Transients

To evaluate the complete color transient response of the system, two questions must be answered:

First, assuming the transient responses of the individual signal channels are known, how does the receiver color point move through color space in response to a sudden jump in position of the camera color point, and

Second, given this path through color space, what are the resulting chromaticity and luminance transients?

The answer to the first question is given by Figures 2 and 3. These show the subcarrier transient paths resulting from two different methods of transmitting the color subcarrier. For a system having different transfer characteristics than those assumed, the subcarrier transient path may be determined by the same methods. Then, if the transient response of the luminance (Y') channel is known, the transient path in the normalized subcarrier diagram (S_R'/Y', S_B'/Y') may readily be plotted.

We may then proceed to the second question with the aid of Figures 6 and 7. If we plot the normalized subcarrier transient path in Figure 6, we may read off the values of the luminance factor, K_S, which, together with the actual luminance signal, Y', give us the value of the luminance, Y, reproduced by the receiver. (See equation (15).) The same transient, plotted in Figure 7, enables us to read off directly the values of chromaticity occurring during the transient.

Specifically, the method of determining a color transient, given its end points, is as follows:

If end points are $(x_1\ y_1\ Y_1)$ and $(x_2\ y_2\ Y_2)$ -

(a) Plot end points on chromaticity diagram of Figure 7 and read off the normalized color difference values S_R'/Y', S_B'/Y', corresponding to these end points.

(b) Plot these end points on the luminance factor diagram of Figure 6 and read off the subcarrier luminance factors K_{S1} and K_{S2} for the end points.

(c) Compute the values of Y' for the end points, by substituting the values of K_S and Y into equation (15). Use these values, together with the values of S_R'/Y', S_B'/Y' [step (a)], to determine the end points of the transient in the S_R', S_B', subcarrier surface.

(d) Using the subcarrier transient path shapes of Figures 2 or 3, plot the subcarrier transient between the end points determined in (c).

(e) Transfer the subcarrier transient to the normalized subcarrier surface by dividing the coordinates of each timing dot by the corresponding value of Y'. Since the interval between timing dots shown on the subcarrier transients of Figures 2 and 3 is equal to the rise time of the luminance transient, Y' is near its initial value up to the time t = -1/2 interval, and its final value at any time later than t = +1/2 interval. At t = 0, the value of Y' is the mean of its initial and final values.

(f) Plot the normalized subcarrier transient in Figure 7, and read off the values of x and y corresponding to each time point. These determine the chromaticity transient.

(g) Plot the normalized subcarrier transient in Figure 6 and read off the values of K_S for each time point. These values, together with the corresponding values of Y' may be substituted into equation (15) to determine the luminance transient.

Analysis of Color Transients

Using the method outlined, specific color transient responses may be evaluated for any case in which the transient responses of the individual signal channels are known. Several transients have been so evaluated for the two systems previously described:

Case 1, the case where both color difference signals are allowed 1 mc. bandwidth and modulate a 3.89 mc. subcarrier, and

Case 2, where one color difference signal is restricted to 0.5 mc., the other is allotted 1.5 mc. bandwidth, and both modulate a 3.58 mc. subcarrier.

In Case 1, both color difference signals have a short rise-time but each contributes a sluggish quadrature component to the other. In Case 2, the system currently under field test by the NTSC, these quadrature components have been eliminated. The rise time of one of the color difference signals has been decreased, while that of the other is increased.

For comparison of the systems, two main types of color transient are shown: transients from saturated colors to de-saturated colors, and transients between saturated colors. The first type cuts across the constant luminance contours of the subcarrier surface (Figure 6) while the second type runs more or less along these contours. The effect of the quadrature component of Case 1 is quite different in these two instances.

Figures 8 and 9 show the transient responses with Case 2 and Case 1 transmission, respectively, for the transitions from a saturated color to a flesh color. In the chromaticity transients, the timing dots are spaced at intervals of 1/8 microsecond. The crosses on the transient path indicate the start and finish of the luminance transient.

The chromaticity transients for Case 2 (Figure 8) show how this system takes advantage of the color perception characteristics of the eye. McAdam's data[6] on equally noticeable chromaticity differences at constant luminance indicate that in the central region of the chromaticity diagram the direction of minimum perceptibility lies more nearly along the y than the x axis. Note how, for all three transients, the relatively slow approach to (or departure from) flesh color is made in this direction of low sensitivity. This feature enhances the subjective sharpness of these transients. If we assume that, for the small area of the color transition involved, a change in y of 0.03 is just perceptible, we see that the elapsed time from the center of the transient to its end at a chromaticity not noticeably different from flesh color is 2 units of time for the transient from red, and four units each for green and blue.

The chromaticity transient of Figure 9 shows the response of the Case 1 system to the transient from flesh color to red. The dotted curve shows the path taken when the transition occurs in the reverse direction, (or the path taken on odd fields when color phase alternation is used).

Note here that from the center of the transient to the final chromaticity, three units of time have elapsed, but that the color point overshoots by a perceptible amount and does not return until after six units of time. This overshoot is produced by the trailing negative peak of the quadrature component. Depending on the direction of the line between the end points of the transient, the overshoot of Case 1 may occur in any direction; not necessarily in the direction of least perceptibility.

The luminance transients shown in Figures 8 and 9 show a minor contribution from the subcarrier transient. A small leading white and trailing black may be seen in all four transients due to this effect.

Figure 10 shows, for Case 2, a transient between a fairly saturated red and blue, in which the subcarrier transient exerts a greater effect on the luminance transient. In actual color pictures, such a transition would rarely occur, but it is included to show an interesting difference between the two systems. The chromaticity transient in this case is very good; the elapsed time from the center of the transient to either end is only two units of time. The luminance transient, however, shows an appreciable anticipatory drop contributed by the subcarrier transient. The effect appears as though it had been produced by phase distortion.

Figure 11 shows the response of the system of Case 1 to the same input transient. Here the effect of the subcarrier transient on luminance is pronounced. The reason is this: the direction of this transient is along the luminance contour lines of the subcarrier surface (Figure 6) so the excursion to the side produced by the quadrature component is across these contour lines in a direction affecting luminance. The impress of the quadrature component waveform on the luminance transient is clearly seen. Note that the two luminance transients (affected by quadrature components of opposite sign) cross at the instant at which the quadrature component is zero.

Conclusions

Comparison of the color transients resulting from the two methods of transmitting the chrominance information yields some interesting conclusions. In Case 2, the reduction in bandwidth of one color difference signal has actually resulted in greater sharpness of the chromaticity transients as well as in a cleaner luminance transient for transitions between saturated colors. Both of these results are due to the elimination of the quadrature components. When a strong quadrature component is present, its effect on luminance and chromaticity is such as to require its cancellation by means of color phase alternation. This feature is not necessary in a system where quadrature components have been eliminated by proper filtering as in Case 2.

While tentative conclusions may be drawn directly from a knowledge of the color transient path, its final evaluation must be made subjectively. This may be done either by direct experiment with a color system or by reference to data on color perception, such as have been published by McAdam[6]. The final test is, of course, the appearance of color transitions in an actual color television picture. The excellent results which have been obtained with the relatively simple system of Case 2 are the most convincing evidence so far of the quality of the color transient response of this system.

APPENDIX I

Luminance Factor Map of Subcarrier Surface

The concave subcarrier surface shown in Figure 5 is the surface obtained by setting the luminance (Y') signal equal to some constant. The gamma precorrected primary signals R', G', B' which are applied to the display device of the receiver are then obtained from the transmission primary signals by means of the linear transformation of equation (5)

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 1.14 & 1 & 0 \\ -.581 & 1 & -.394 \\ 0 & 1 & 2.03 \end{bmatrix} \begin{bmatrix} S'_R \\ Y' \\ S'_B \end{bmatrix} \qquad (5)$$

Given these input voltages, the display device produces Panel 7 primary lights R, G, and B proportional to the γth power of R', G' and B'. The total luminance (Y) on the screen of the display device is obtained by adding the primary lights, each weighted by the proper luminosity coefficient. That is:

$$Y = 0.299\ R + 0.587\ G + 0.114\ B$$
$$= 0.299(R')^{\gamma} + 0.587(G')^{\gamma} + 0.114(B')^{\gamma}$$

Substituting the values of R', G' and B' obtained from equation (5), we obtain

$$Y = 0.299\ (1.14\ S'_R + Y')^{\gamma}$$
$$+ 0.587\ (-.581\ S'_R + Y' - .394\ S'_B)^{\gamma}$$
$$+ .114\ (Y' + 2.03\ S'_B)^{\gamma}$$

If we divide both sides of this equation by Y'^{γ}, we obtain the expression

$$Y/Y'^{\gamma} = .299\ (1.14\ S'_R/Y' + 1)^{\gamma}$$
$$+ .587\ (-.581\ S'_R/Y' + 1 - .394\ S'_B/Y')^{\gamma}$$
$$+ .114\ (1 + 2.03\ S'_B/Y')^{\gamma}\ . \qquad (16)$$

We may write this as

$$Y/Y'^{\gamma} = K_S \qquad (15)$$

where K_S is a factor dependent only on the normalized subcarrier voltages S'_R/Y', and S'_B/Y'. K_S is the ratio of the reproduced luminance to the γth power of the luminance signal, assuming that all constants of proportionality between electrical and colorimetric quantities are unity. The luminance factor contours of Figure 6 were obtained by evaluating the right-hand side of equation (16) for a number of values of S'_R/Y' and S'_B/Y'.

APPENDIX 2

Chromaticity Map of Subcarrier Surface

The relation between the Panel 7 primaries and the CIE non-physical primaries is given by the linear transformation[5]

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.91 & -0.532 & -0.288 \\ -0.985 & 1.999 & -0.028 \\ 0.058 & -0.118 & 0.898 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \qquad (17)$$

Since the chromaticity coordinates are given by

$$x = \frac{X}{X + Y + Z} \quad y = \frac{Y}{X + Y + Z} \quad \text{and } z = \frac{Z}{X + Y + Z}$$

where $x + y + z = 1$,

we may substitute into equation (17) the values

$$X = \frac{x}{y} Y \quad Z = \frac{z}{y} Y \quad Z = \frac{(1 - x - y)}{y} Y$$

obtaining

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = Y \begin{bmatrix} 1.91 & -0.532 & -0.288 \\ -0.985 & 1.999 & -0.028 \\ 0.058 & -0.118 & 0.898 \end{bmatrix} \begin{bmatrix} \frac{x}{y} \\ 1 \\ \frac{(1 - x - y)}{y} \end{bmatrix} \tag{18}$$

The coefficients of the matrix are the same as in equation (17).

From equation (5), we may write the amounts of the Panel 7 primary lights at the receiver as

$$R = Y'^{\gamma} (1 + 1.14\, S_R'/Y')^{\gamma} \tag{19}$$

and

$$B = Y'^{\gamma} (1 + 2.03\, S_B'/Y')^{\gamma} \tag{20}$$

Equating the values of R and B given by equations (19) and (20) to those given by equation (18), and solving for S_R'/Y' and S_B'/Y' yields

$$\frac{S_R'}{Y'} = \left[\frac{1}{y}\frac{Y}{Y'^{\gamma}} \quad (1.530\, x - 0.170\, y - 0.201)\right]^{1/\gamma} -0.877 \tag{21}$$

$$\frac{S_B'}{Y'} = \left[\frac{1}{y}\frac{Y}{Y'^{\gamma}} \quad (-0.120\, x - 0.145\, y + 0.128)\right]^{1/\gamma} -0.493 \tag{22}$$

Note that the factor (Y/Y'^{γ}) in the above expressions is simply K_S. Since contours of this factor in the chromaticity diagram have been evaluated by Applebaum[5] and Livingston[7], work may be saved by reference to their data. Figure 7 was plotted using equations (21) and (22).

BIBLIOGRAPHY

1. Bingley, F. J. - "Colorimetry in Color Television." Paper presented at IRE Radio Fall Meeting, Syracuse, New York, October 20-22, 1952.

2. Kerr, J. S. S. and Howells, P. W. - "A Proposal for a Modification of the Chrominance Signal Specification." NTSC Report (NTSC - P13 - 289), August, 1952.

3. Kerr, J. S. S. - "Transient Response in a Color Carrier Channel with Vestigial Sideband Transmission." Paper presented at IRE National Convention, New York, March 23-26, 1953.

4. RCA Laboratories - "Tests Relating to the Choice of Narrow and Wide Band Components for a Balanced Color Gamut System." NTSC Report (NTSC - P13 - 286), October, 1952.

5. Applebaum, Sidney - "Gamma Correction in Constant Luminance Color Television Systems," Proceedings of the IRE, Vol. 40, No. 10, pp. 1185-1195, October, 1952.

6. McAdam, D. L. - "Quality of Color Reproduction," Proceedings of the IRE, Vol. 39, No. 5, pp. 468-485, May, 1951.

7. Livingston, Donald C. - "Colorimetric Analysis of Gamma Corrected Color Television Systems." Paper presented at IRE Radio Fall Meeting, Syracuse, New York, October 20-22, 1952.

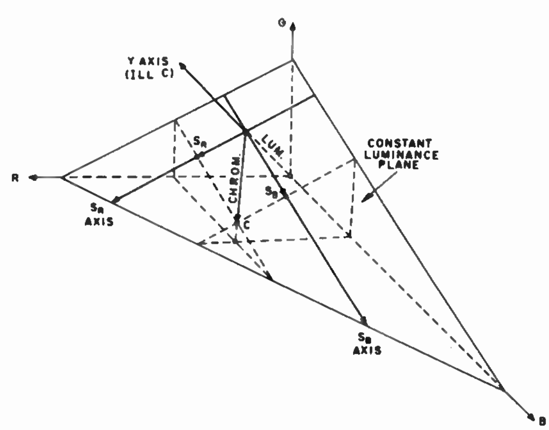
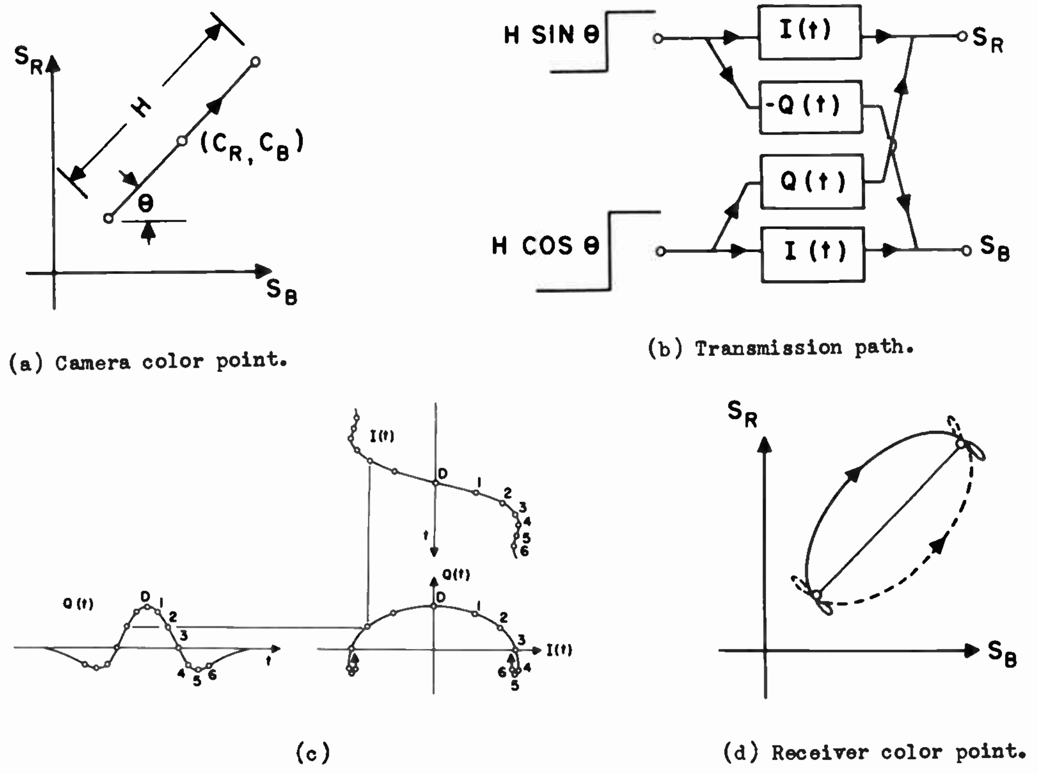


Fig. 1
Transmission primary axes in RGB color space.



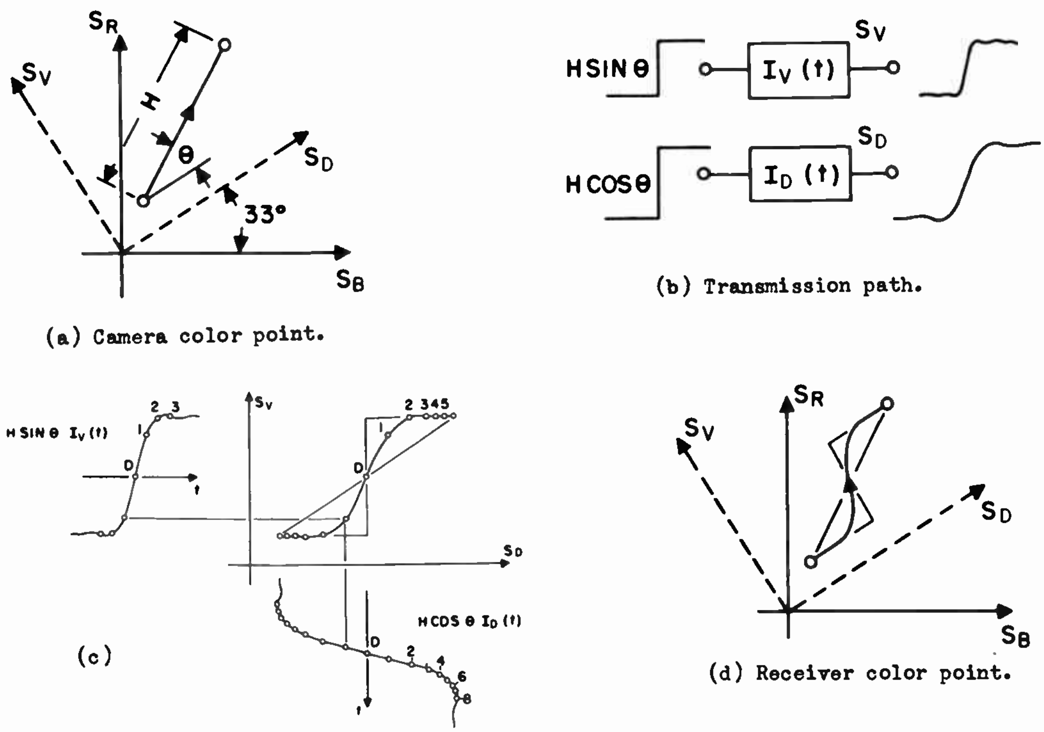
(a) Camera color point.

(b) Transmission path.

(c)

(d) Receiver color point.

Fig. 2
Subcarrier transient path, case 1.



(a) Camera color point.

(b) Transmission path.

(c)

(d) Receiver color point.

Fig. 3
Subcarrier transient path, case 2.

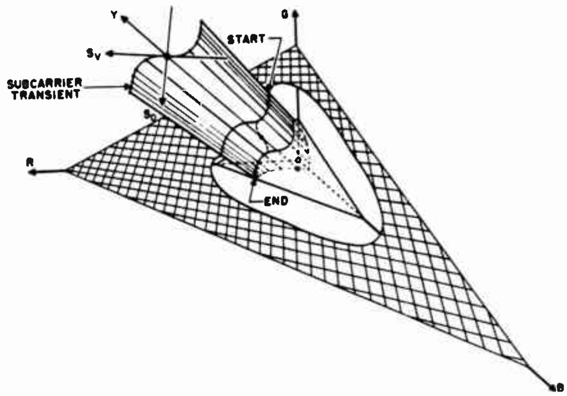


Fig. 4
Color transient path in RGB color space.

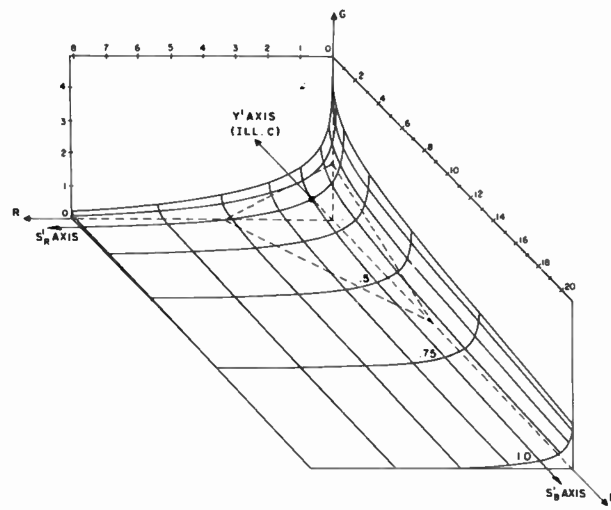


Fig. 5
Subcarrier surface in RGB color space.

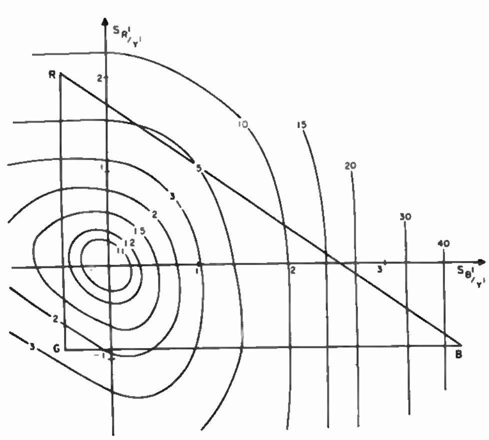


Fig. 6
Luminance Factor (K_S) contours in normalized subcarrier surface.

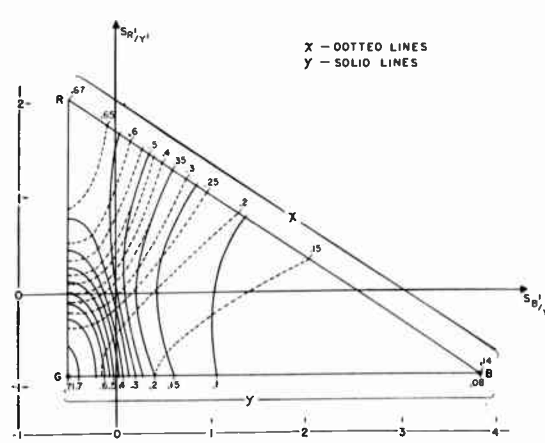
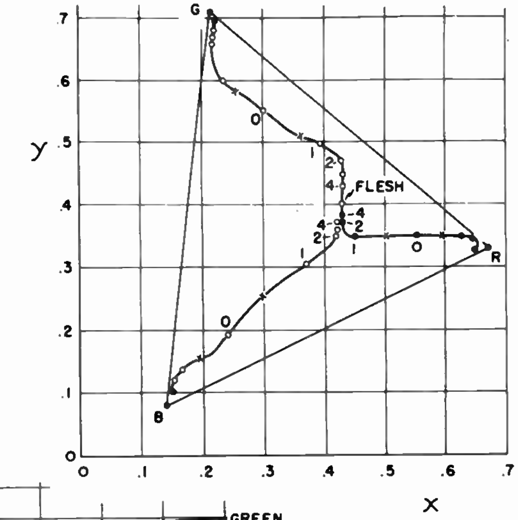


Fig. 7
Chromaticity coördinates in subcarrier plane.



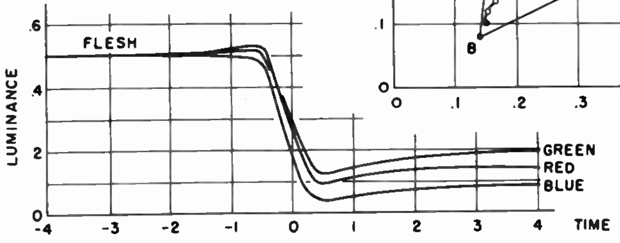
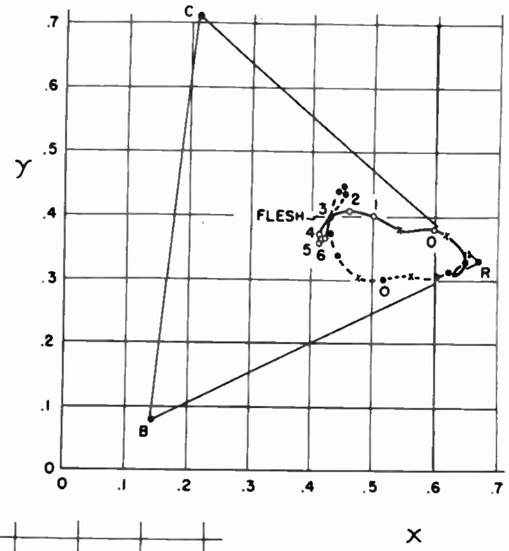


Fig. 8
Color transient in luminance and chromaticity, case 2.
Note: 1 unit of time = 1/8 microsecond.



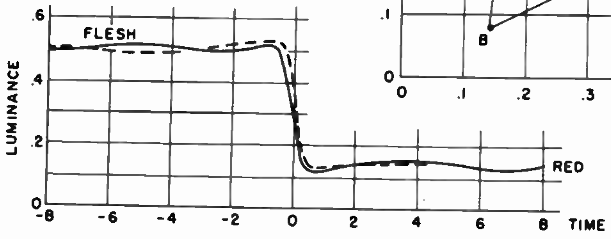
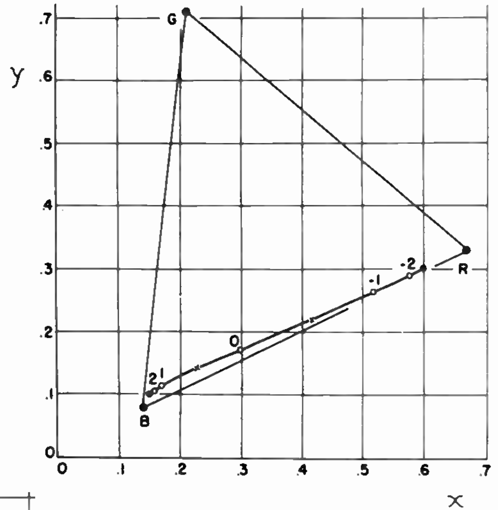


Fig. 9
Color transient in luminance and chromaticity, case 1.
Note: 1 unit of time = 1/8 microsecond.



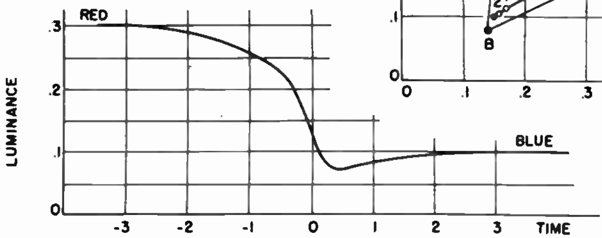
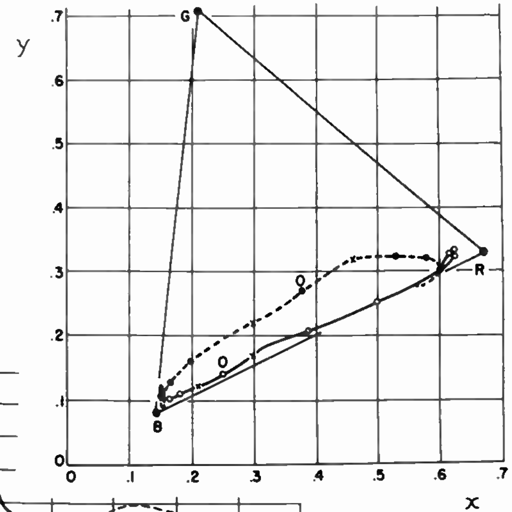


Fig. 10
Color transient in luminance and chromaticity, case 2.
Note: 1 unit of time = 1/8 microsecond.



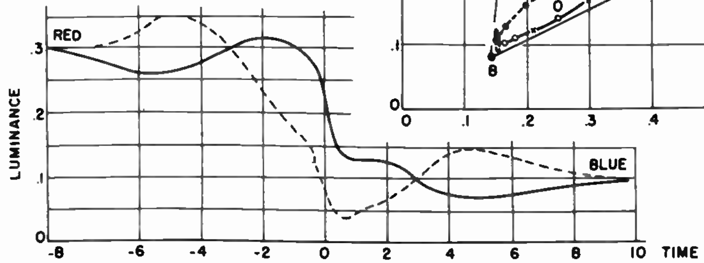


Fig. 11
Color transient in luminance and chromaticity, case 1.
Note: 1 unit of time = 1/8 microsecond.

PROBABILITY DISTRIBUTION MEASUREMENTS OF TELEVISION SIGNALS

William F. Schreiber

Cruft Laboratory, Harvard University

Cambridge, Massachusetts

Summary

A method is described of obtaining an upper bound on the information content of television signals by means of measuring the second-order probability distribution of video signals. Apparatus for generating the signals and presenting the distributions on an oscilloscope was built. The information content of four subjects was measured and found to range from .24 to 1.88 bits in a 32-level system.

Theoretical Preliminaries[1]

If a message is made up from an alphabet of s characters or symbols, each one of which is independent of its neighbors, then the information per symbol is

$$H = -\sum_{i=1}^{s} p(i)\log p(i) \qquad (1)$$

If now we consider messages constructed from the same alphabet as above, but having successive symbols dependent, the information is reduced considerably. If $p_i(j)$ is the probability of the j^{th} symbol occurring after the i^{th}, then $-\sum_{j=1}^{s} p_i(j) \log p_i(j)$ is the information carried by the next symbol after the i^{th} and by averaging over all the i's, we arrive at the average information conveyed by a symbol when the previous one is known:

$$H = -\sum_{i,j=1}^{s} p(i,j) \log p_i(j) \qquad (2)$$

Note that $p(i,j)$, the probability of the occurrence of the i^{th} and the j^{th} symbol in a pair, is equal to $p(i)p_i(j)$.

If statistical influence extends over more than two successive symbols, as is often the case, a similar formula holds but i then means not the i^{th} symbol, but the i^{th} state, which defines the entire past of the message having a statistical influence on the next symbol. The index i is then summed over all possible states.

A striking characteristic of the formula for information content is that H is larger when the $p_i(j)$ are more equal and H is smaller when the $p_i(j)$ are less equal. In the one extreme case of all the $p_i(j)$ being equal to $1/s$, $H = \log s$ while in the other extreme case of some one $p_i(j) = 1$, all the others being zero, $H = 0$.

Application to Television

Pictures are transmitted electrically by dividing them into a multitude of tiny elementary areas, measuring the brightness of each area, and transmitting the brightnesses in rapid succession. Within limits, the larger the number of elementary areas, and the larger the "alphabet" of allowable brightnesses, the more like the original will the reproduced picture seem.

The present television system provides the channel capacity to transmit 30 symbols per second for each elementary area. Furthermore, each of these symbols may be completely independent of the others. We could, for example, transmit a random signal derived from a picture in which the brightness of each elementary area was unrelated to the brightness of the other areas and in which each of the thirty pictures per second was unrelated to each of the others. In fact, however, the pictures which are transmitted are not like that at all. The nature of a meaningful picture is that, for the most part, each area is very much like its neighbors. If it were not, the picture would not look like anything. In addition, in order to produce the illusion of continuous motion, each entire picture must be very much like the previous one. If thirty completely pictures were transmitted each second, the human observer would not comprehend them.

This strong correlation between successive symbols, symbols separated by the line duration, and symbols separated by the frame duration, has the effect of reducing the information of the video signal. This can be seen by reference to (2), since high correlation means that some of the $p_i(j)$ (i near j, indicating little change) will be large, while most of the $p_i(j)$ (i not near j, indicating a large change) will be small.

The first step of a long-range program aimed at improving the transmission efficiency of television, then, is to determine just what the information reduction is. If it is sizeable, then it will be worth while to proceed to the next step, which is to devise new systems of higher efficiency.

Which Parameter to Measure

The purpose of this research was to obtain statistical information about the kind of pictures which are transmittable by electrical means. The complete statistical description of a signal is the joint amplitude probability distribution of n successive symbols, where n is at least as high as the range of statistical influence in the signal. By range of influence we mean the symbol separation beyond which knowledge of one of the symbols tells us "almost" nothing about the other symbol. As an example, in the English language it is said that statistical influence extends over about eight letters, and hence a knowledge of the eighth-order probability distribution, i.e., the probability of combinations such as "ATIONALS" or "XPQTZSPO," would enable us to find out everything about the language that could be found out, at least on a probability basis. We might calculate the information content of the language, which would permit us to state the maximum efficiency of any possible code, and we could also calculate the efficiency of any particular code.

In television the range of influence in any direction is at least one-quarter of the picture size, or some eighty symbols. If we are considering pictures with thirty-two brightness levels, then a table with 32^{80} (or about 10^{120}) entries would be required to write down the distribution. It can readily be seen that this is not the way to go about the problem. It is necessary instead to use some less exact but more economical procedure. Other workers in this field (Kretzmer[2] at Bell Telephone Laboratories, Elias at Harvard, Kohlenberg and Cheatham at Boston University, and presumably others) have measured the correlation function of pictures. It was decided to measure the joint probability density of two symbols, spaced apart by any amount in any direction. This parameter, which is intermediate both in complexity and usefulness between the complete probability description and the correlation function, requires, for a thirty-two level picture, a table of 32^2 (= 1024) entries for each measurement. Usually, five different separations were used for each picture.

One thing that can be done with the parameter measured is to compute an upper bound on the information content of the signal. Such a computation can also be performed by using the correlation function, but since the second-order probability density intrinsically contains more statistics than the correlation function, the bound computed ought to be a better one.

Knowledge of $p_i(j)$ enables an exact computation to be made of $H_x(y)$, which is the information generated by the transmission of one symbol (y) when a previous symbol (x) is known. If correlation in the signal extends no further than the pair x and y, then $H_x(y)$ is the true information content of the signal. If correlation extends further, then $H_x(y)$ is an upper bound on the true information content. This comes about because in the case of long-range correlation, the second-order distribution which is measured must be an average of all the second-order marginal distributions of the higher-order distribution which is the complete statistical description of the signal. Elias proves (Ph.D Thesis, Harvard University, 1950) that the entropy of an average distribution is equal to or larger than the entropy which is the average of the entropies of the separate distributions, the equality holding if the higher-order distribution can be obtained from the lower-order distribution. This can be seen physically as well, since if long-range correlation exists, knowledge of a number of previous symbols tells us more about the present symbol than knowledge of just one previous symbol, and hence the information conveyed by the present one must be reduced.

The correlation function, on the other hand, supplies an upper bound on $H_x(y)$. (Peter Elias, "A Note on Autocorrelation and Entropy," Proc. I.R.E. 39, p. 839, July 1951). For a given correlation function, a normal distribution has the highest entropy, which is calculable from the correlation function. The entropy so calculated is thus equal to $H_x(y)$ only if the signal is in fact normally distributed. The correlation function thus supplies an upper bound on the true information content which is looser than that obtained by the procedure used in this paper.

Experimental Arrangement

Measurement of the second-order probability density requires first the generation of two video signals, appropriately displaced in space or time, and then the counting of the number of times per frame each pair of brightnesses, i and j, occur. There are many ways in which each of these requirements may be met. The techniques used were selected for their simplicity and rapidity.

Producing the Two Signals

Since the displacement between the signals can be looked at as either geometrical or temporal, there are at least two methods of attack. One is to scan with one aperture and delay the signal by the desired interval to obtain the second signal and the other is to scan with two apertures suitably displaced. It is desired, of course, that the two signals be as nearly alike as possible, except for the delay, and that the delay be variable over a wide range. Delays of several microseconds are readily obtainable without serious degradation of 10 Mc/s video signals, by means of electrical delay lines. An arrangement of several fixed and variable-length lines does indeed suffice for delays long enough to measure the distribution of signals along the same horizontal scanning line. However, in order to deal with vertical aperture separation, delays of a multiple of a line scanning time (63.5 microseconds) are necessary, in which case acoustic delay lines are needed. In addition, an amplifier to make up the sixty or seventy db attenuation in the line must be included, as well as an equalizer to compensate for the inevitable phase and amplitude distortion. Constructing such a system is a task nearly as complicated as this entire project.[3] Finally, even if such acoustic delay systems were easily available, the presence of interlace in the standard scanning system would result in a comparison, not between vertically adjacent picture elements, but between elements spaced apart by an even number of lines. In addition, variable acoustic delay lines, which would allow a convenient adjustment of vertical separation, are even more difficult to make than fixed length lines. For all these reasons the two-aperture geometrical separation method was used instead.

Performing the measurement

Previous investigators have used various electronic methods to measure first-order probability distributions. Kretzmer[4], in his device called a Probabiloscope, applied the signal to be measured to one set of deflection plates of a cathode-ray tube and by use of an extremely ingenious photographic method measured the brightness of the trace as a function of deflection; Nienburg and Rogers[5] applied the signal to one set of deflection plates of an electrostatic storage tube for a certain length of time, and then measured the charge stored as a function of deflection. A fairly standard method is to pass the signal through slicing circuits and measure the proportion of the total scanning time when the signal is within a certain amplitude interval, and slowly scan the interval through the full range of signal amplitudes.

All of these methods could be used for the measurement of second-order probabilities by applying the second signal to a slicing circuit which is arranged to keep the first signal turned off except when the second signal is within the proper amplitude interval. Then a family of curves would result, one for each amplitude interval of the second signal. Kretzmer has done this.

The method used in this research to measure the second-order probability is considerably simpler than other proposed methods, and is thought to be unique. The basic idea is to apply the two signals to the vertical and horizontal deflection plates, respectively, of an oscilloscope. A pattern results which is a representation of the desired probability density p(x,y) where x and y, the amplitudes of the two signals, are represented by horizontal and vertical deflection, and where p(x,y) is proportional to the brightness of the screen at the point (x,y). This comes about since the brightness is proportional to current density, which in turn is proportional to the fraction of the scan time during which the electron beam is in the vicinity of the point. The beam is so deflected whenever the two signals have the pair of instantaneous values corresponding to the deflection. The pattern is recorded by measuring the brightness distribution in some appropriate manner, say by photographing the pattern, or by actually measuring the brightness at each point with a scanning photometer. With proper equipment design, this method can be made as accurate as the slicing circuit technique, and it has the further great advantage that qualitative results can be obtained very quickly by merely observing the pattern or by photographing it. This is an advantage of great importance for the purpose of this research, since what is desired is a large quantity of approximate information about a variety of pictures, rather than very precise information about a particular picture.

Description of the apparatus

The operation can be understood by reference to the block diagram, Fig. 1. The optical system is the principal part of the equipment.

The flying-spot cathode-ray tube has a very bright, short persistence phosphor, A standard television raster is produced thereon by means of synch and sweep generators, which for convenience are synchronized with the power line. On account of the short persistence, only one spot, roughly the size of the electron beam, is illuminated at one time. The raster is focussed by the projection lens on two identical 35 mm transparencics. This is done by means of the beam splitter, a half-aluminized, very flat mirror on a thin transparent base, which transmits and reflects about equal portions of the incident light. The light which passes through the transparencies is collected by the photomultiplier tubes, assisted by condensing lenses. Since only one spot of the CRT phosphor is illuminated at one time, light passes through only one point of each slide at one time. Consequently, the photomultiplier current is a video signal.

In the apparatus, one of the transparencies is mounted on slides so that it can be moved by measured amounts perpendicular to the optic axis, and the other is mounted so that it can be rotated about the optic axis. In addition, provision is made for the second transparency to be moved along the optic axis. By making use of all these adjustments, it is possible to project rasters of the same size and focus on each transparency, and to have the transparencies oriented identically with respect to the projected rasters, so that the same point of each picture is scanned at the same time. The pictures are then in register, and presumably, identical video signals will be produced. To scan different points simultaneously, the slide-mounted transparency is shifted horizontally or vertically by means of tenths micrometers.

The two photomultiplier outputs are amplified and equalized (to compensate for screen persistence of the flying spot tube) in the preamplifiers and again amplifier in the video deflection amplidier to a level high enough to produce about 1 1/2 to 2 inches deflection in the recording scope, which is operated at a total acceleration voltage of about 20,000 volts. For this purpose, it need not be a short persistence tube, but this characteristic makes the device more adaptable to other uses.

By applying one signal to the vertical and the other to the horizontal deflection plates of the recording scope, a pattern is produced on the screen whose brightness is proportional to the second-order probability distribution of the two signals, as already explained.

The brightness of the screen is measured with a Photovolt 520M Multiplier Photometer. The search unit of the photometer is mounted over the center of the screen, and is shielded from it by two opaque masks with small holes at the center so that only light emitted by a small circular area near the center of the screen falls on the photocathode of the search unit. To measure the various parts of the pattern, the entire pattern is moved past the holes by rotating the oscilloscope centering controls. Stepping relays facilitate rapid translation of the pattern past the holes by moving the centering controls a fixed amount each time the relay is energized. The steps are adjusted so that the entire pattern is covered in about thirty-two steps each way, or about 1000 readings in all. A second pair of solenoids causes the controls to return to extreme counterclockwise position when energized.

Monitoring

As a check on the operation of the system, a picture monitor is used. This is a commercial television receiver with a ten-inch tube, modified so that its sweep generator is synchronized by the same pulses that trigger the scanner sweep generator, and so that the output of the monitor video amplifier is applied to the grid of the picture tube. This signal is ordinarily the difference between the two video signals but it may be one or the other or the sum, and of either polarity depending on the setting in the subtractor unit. When the two signals are mixed, as is usually done, their relative amplitude may be adjusted, so that if the two pictures are in register, a null can be observed. This is a sensitive means of putting the pictures in register and of focussing.

Figure 2 consists of photographs of the probability pattern as it appears on the recording scope. The subject was the standard RMA Test Pattern. Distributions are shown for the two pictures in register, and also displaced horizontally, vertically, and diagonally. Notice that there is no significant difference in rate of spread for the various directions of displacement. This holds true for all ordinary pictures.

Figure 3 is a general view of the equipment.

Results of the Measurements

The oscilloscope presentation of the second-order probability distribution was observed for eleven still pictures and the RMA test pattern, as well as for five sequences selected from motion picture film. All these presentations were photographed, and in the case of the stills, a number of separations, horizontally, vertically, and diagonally, were used. The patterns of four of the stills and one sequence were also measured with the photometer. Five separations were used for each still and five pairs of pictures for the sequence. Each of these twenty-five patterns was used to compute the conditional uncertainty of one of the signals when the other one is known. The results are plotted in Fig. 4 as a function of separation expressed in Nyquist intervals for the stills (the reciprocal of twice the bandwidth) and in frames for the sequence.

When the two pictures under test are in register and if the apparatus were perfect, the recording scope pattern would be a 45° line, because the video signals would be identical. In this case the conditional entropy would be zero, which is reasonable since no information is conveyed by a message known to be identical to a previous message. In our case, however, the pattern corresponding to zero displacement is a spread-out 45° line, and the entropy calculated for this case is not zero, but is something over two bits (the maximum conditional entropy for a pattern of 32x32 points is $\log_2 32 = 5$ bits). This is a measure of the difference in the video signals.

When the pictures are out of register, the probability that the two elementary areas being scanned at any instant have equal brightnesses decreases, and the 45° line spreads out farther. This causes the measured entropy of the pattern to increase since the change is in the direction of equalizing the density in the various parts of the pattern.

It is apparent that the measured entropy is a function of both the signal and the noise. There is no convenient way to use the data to calculate exactly the entropy of the signal itself, but bearing in mind that the pictures themselves exhibit a wide range of complexity (and hence of entropy) and that as a result exact measurement of the entropy of a particular picture is not necessary, it is possible, using certain plausible assumptions about the signal and the noise, to arrive at an approximation to the entropy of the signal. The assumptions are:

1. The noise is normally distributed.

2. The conditional distribution of the signal is normal.

3. The signal and the noise are uncorrelated.

The first assumption is very nearly true, since the principal source of noise in photomultiplier tubes is thermal emission from the photo cathode, temperature-limited as in a temperature-limited diode, when operated under the conditions used in this experiment. The second assumption is not so obvious. Certainly the first-order distribution of the signal is not gaussian, since some pictures are mostly black, some mostly white, and some have a fairly even distribution of grays. However, the conditional distribution is similar to a normal distribution Except for some unusual cases, such as scenes with periodicities of cell size (very fine checkerboard) the conditional distribution will at least be unimodal. Obviously, in ordinary pictures, it is most probable that two adjoining cells are equally bright, and the greater the brightness difference, the less likely the occurrence of such a pair of cells. Figure 5 shows the similarity between a measured distribution taken at random and a normal curve. The protrusion indicates some deviation from normalcy, but the effect on the result is small, since for every curve with an upward protrusion, another can be picked out with a downward protrusion.

The third assumption is good, since although the rms value of the noise does in fact depend on the level of the signal, the two nevertheless have zero correlation since the instantaneous value of the noise is independent of signal amplitude.

On this basis the computation is straightforward:

If $$D = S + N \qquad (1)$$

then $\sigma_D^2 = \sigma_S^2 + \sigma_N^2$, where D, S, and N refer to the data, the signal, and the noise respectively, and σ is the standard deviation. For any first-order normal distribution

$$H = \log_2 \sqrt{2\pi e}\,\sigma \text{ bits per sample}^{6}; \qquad (2)$$

hence $$2^{2H_D} = 2\pi e(\sigma_S^2 + \sigma_N^2) = 2^{2H_N} + 2^{2H_S} \qquad (3)$$

and $$H_S = \tfrac{1}{2}\log_2(2^{2H_D} - 2^{2H_N}) \qquad (4)$$

This relation, which applies to any first-order distributions D, S, and N, conditional or unconditional, and to the corresponding entropy H_S, thus enables us to calculate the entropy of a signal when we know the entropy of the accompanying noise, and the entropy of the sum of the signal and noise, subject to the assumptions made. It remains to find H_N, which is done by measuring the entropy of the data when the pictures are in register. In this case, $H_S = 0$ and

$$0 = \tfrac{1}{2}\log_2(2^{2H_D(0)} - 2^{2H_N}), \qquad (5)$$

giving

$$2^{2H_N} = 2^{2H_D(0)} - 1 \qquad (6)$$

It is not necessary to calculate H_N from (6) since 2^{2H_N} can be substituted directly in (4), giving

$$H_S = \tfrac{1}{2}\log_2(2^{2H_D} - 2^{2H_D(0)} + 1) \qquad (7)$$

If H_D is taken at one Nyquist interval separation, then H_S becomes $H_x(y)$ since x and y are then adjacent elementary areas.

In the computations, a smooth curve was drawn through each plot of H_D vs separation, the value of the ordinate at zero separation being taken as $H_D(0)$ and the value at one Nyquist interval as $H_D(1)$ in (7). A similar procedure was followed for the sequence, except that as no data were taken for zero separation the value was obtained by extrapolation. The results are:

Subject	$H_x(y)$, bits
1	.77
2	1.88
6	.57
13	.24
S2	.90

Notice that the order of the results is the same as a subjective estimate of the order of complexity of the pictures. (See Fig. 6.)

A conservative upper bound on the entropy of systems with more quantization levels is obtained by assuming that the extra information is uncorrelated, and adding one bit for each binary digit above 5. This would be the case if the signal-to-noise ratio were 32:1, for example.

Kretzmer[7] obtains for $H_x(y)$ in 6 bit system, an average value of about 2 1/2 bits. This is of the same order of magnitude as one plus the result reported here and so the experiments may be said to be in consonance.

Applicability to Practical Systems

The results of this and similar

experiments based on the second-order probability density indicate that full use of this statistic would enable the bandwidth of television transmissions to be reduced by a factor of two or three. While there are immediate commercial applications of such a reduction if it could be instrumented in a simple manner, workers in the field are almost all of the opinion that very much larger reductions are possible, and that these reductions would be made evident by measurement of the higher-order probability distributions. For example, a reduction factor two or three is indicated by knowledge either of the previous sample on the same line or by the sample on the line above. Very likely, knowledge of both these samples simultaneously would produce a larger saving. This requires measuring a third-order probability density and if even larger savings are desired, the corresponding sample in the previous frame could be included by measuring a fourth-order density. The labor involved in making the second-order measurement reported in this thesis is so large that automatic methods would have to be developed for more involved measurements.

Acknowledgment

This paper is an abstract of the thesis, "Probability Distributions of Television Signals," submitted by the author to the Faculty of Arts and Sciences of Harvard University in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The work was performed under the direction of Professor Harry R. Mimno, and was supported by Harvard University and by the General Electric Educational Fund. At present the work is being continued under ONR Contract N5ori-76, T. O. 28, made possible through support extended Cruft Laboratory, Harvard University, jointly by the Navy Department (Office of Naval Research), the Signal Corps of the U. S. Army, and the U. S. Air Force.

References

1. C. E. Shannon, "A Mathematical Theory of Communication," Bell Syst. Tech. J. 27, pp. 379-423, 623-656, October 1948.

2. E. R. Kretzmer, "Statistics of Television Signals," Bell Syst. Tech. J. 31, pp. 751-763, July 1952.

3. A. L. Hopper, "Storing Video Information," Electronics 24, p. 123, June 1951.

4. Kretzmer, op. cit.

5. R. E. Nienburg and T. F. Rogers, "A Storage Tube as an Amplitude Distribution Analyzer," presented at the I.R.E. National Convention, New York, March 19, 1951.

6. Shannon, op. cit.

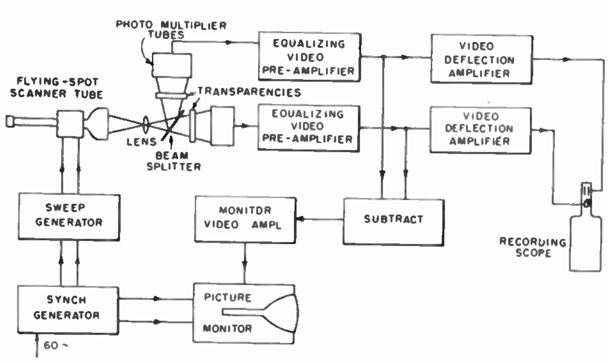
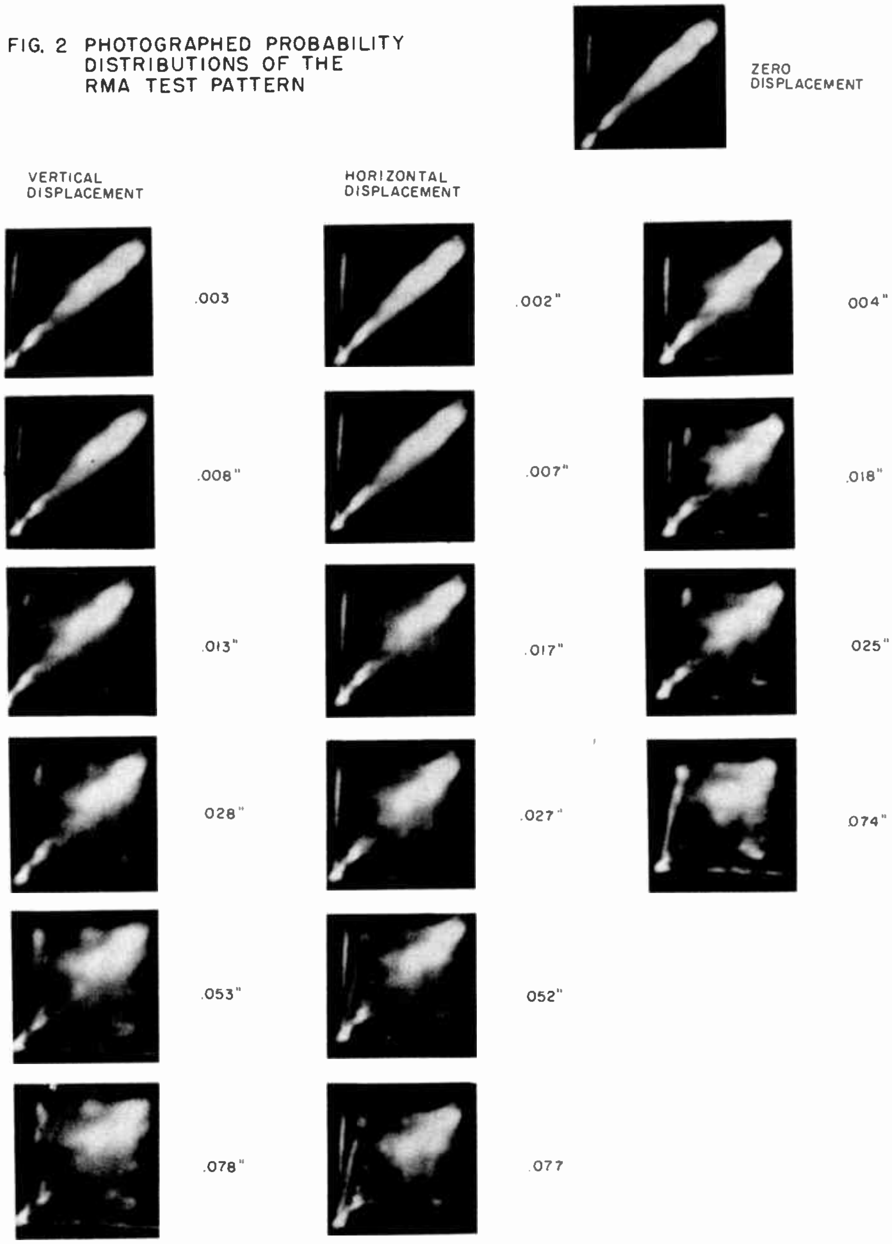


Fig. 1 - Block diagram of the probability machine.

FIG. 2 PHOTOGRAPHED PROBABILITY DISTRIBUTIONS OF THE RMA TEST PATTERN



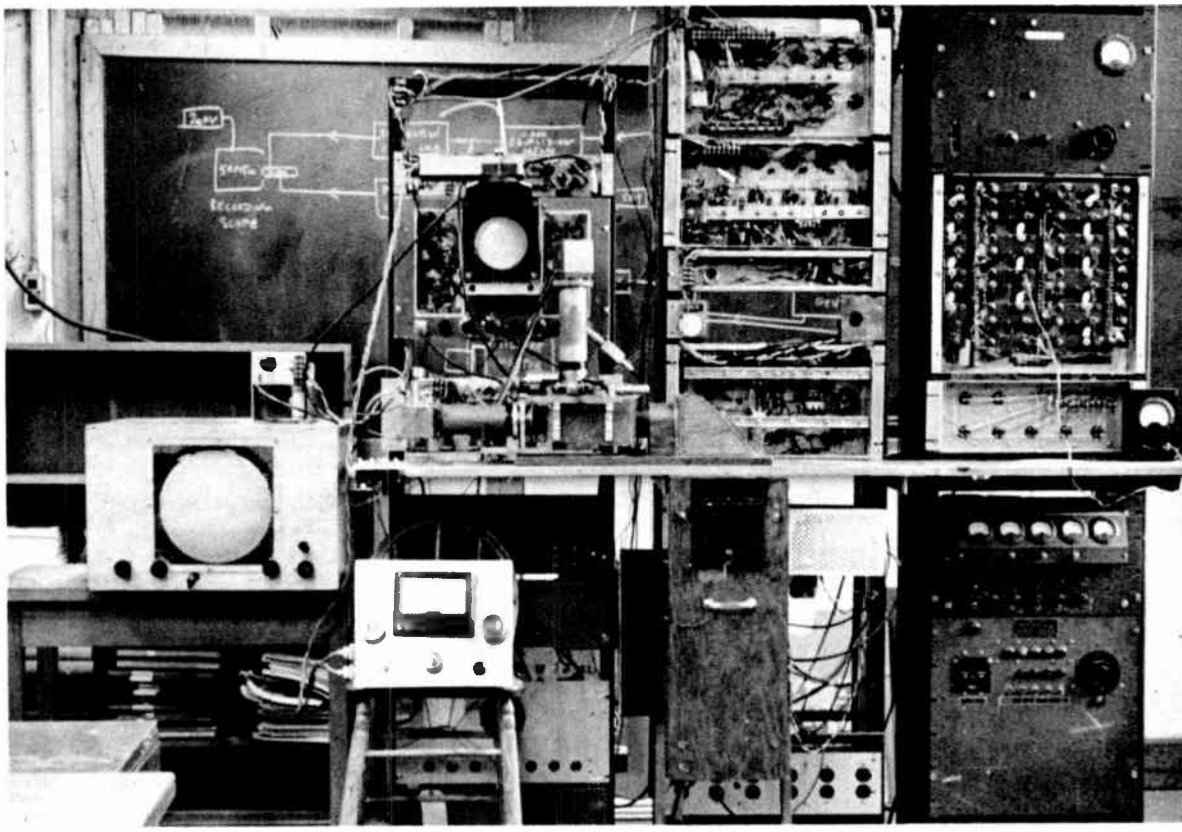


Fig. 3 - General view of the equipment.

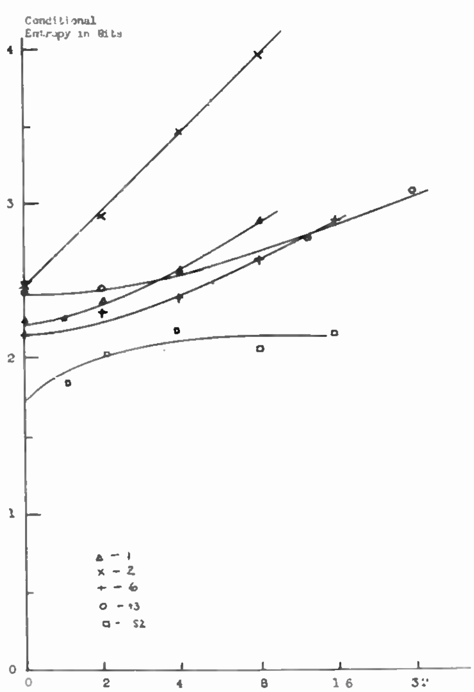


Fig. 4
Conditional entropy vs. separation.

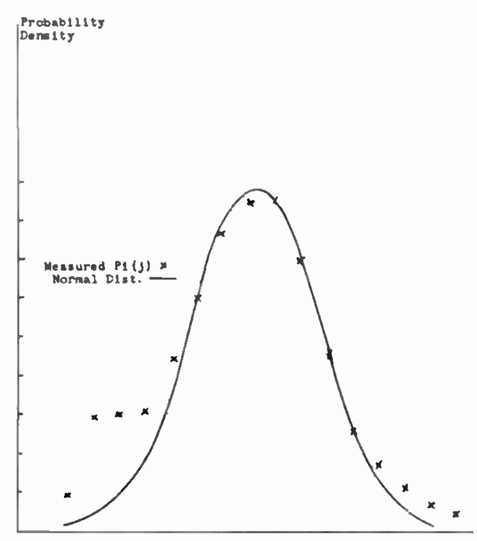


Fig. 5
Similarity of the conditional distribution to a normal distribution.

SUBJECT #1

SUBJECT #2

SUBJECT #6

SUBJECT #13

Fig. 6 - Photographs of the subjects used.

A Precision Line Selector for Television Use

I. C. Abrahams and R. C. Thor
General Electric Company
Syracuse, New York

Introduction

The instrument to be described in this paper is in reality a much more versatile one than the title implies. Indeed, its versatility is at least as important a recommendation for the instrument as its ability to perform its primary function as a line selector extremely accurately and conveniently. It is the purpose of this paper to describe the principles and design of the instrument as well as a number of its applications.

Basically, the approach used is to produce a variable phase shift at a frequency of 157.5 kilocycles, corresponding to the tenth harmonic of the standard television line rate. Phase shifts at any other frequency may then be obtained through the use of suitable frequency dividers or multipliers. The instrument, therefore, may be used wherever variation and/or measurement of time delays or phase shifts are desired in a television system. The frequency divider and multiplier chains are also of great utility.

General Description

The basic unit consists of three parts: (1) the phase-shifter; (2) the frequency divider; and (3) the frequency multiplier. The first two items are fundamental to the functioning of the line selector; the third is included as an added convenience, as will be described. The interconnection of these three parts, as well as the frequencies of their inputs and outputs, are shown in the block diagram of Figure 1. The phase shifter unit has one input, into which may be fed periodic signals having frequencies corresponding to that of the line frequency (15.75 kilocycles) or its second or tenth harmonics. Three outputs, having variable phase shifts with respect to the input, are available; the frequencies of these outputs are the same as mentioned above, i.e., 15.75 kilocycles, 31.5 kilocycles, and 157.5 kilocycles. The frequency multiplier chain takes the 157.5 kilocycle output and multiplies it by a factor of 128 in order to produce an output at approximately 20 megacycles (20.16 megacycles, to be precise). The frequency divider chain, on the other hand, takes the 31.5 kilocycle output of the phase shifter unit and divides it by factors of 35, 525, and 1050, to produce outputs of 900, 60, and 30 cycles, respectively. As will be shown, all of the abovementioned frequencies are of use in the analysis of a television system.

The Phase Shifter Unit

A block diagram of the phase shifter unit is shown in Figure 2. The input circuit of the first multiplier stage has an impedance level of 75 ohms and is aperiodic. The plate circuit of the first stage and the second amplifier are tuned to 157.5 kilocycles. Hence, any submultiple of this frequency may be fed into the input of this unit, provided it has sufficient 157.5 kilocycle content and/or there is sufficient amplitude to produce the required amount of currents of this frequency in the first amplifier tube. The 157.5 kilocycle output of the second amplifier is split into three symmetrical phases by means of appropriate R-C circuits, for application to the phase-shifter proper. This is the type which has often been used for radar purposes, and is described by Blackburn in Volume 17 of the Radiation Laboratory Series. Essentially, it consists of a capacitor-type phase-shifter in which a three-phase signal is applied to its input; the output is single-phase, having a phase-angle, with respect to any given one of the input phases, which is numerically equal to the angular rotation of the shaft. It is of the continuous type; i.e., the shaft may be rotated through 360 degrees, and on around again. A more complete description of the phase shifter capacitor is given in the accompanying paper entitled "Phase Measurements at Subcarrier Frequency in Color Television." Angular accuracies of $\pm$ 1% (about $\pm$ 4 degrees) are easily attainable. The 157.5 kilocycle sinusoidal output of the phase shifter capacitor, after suitable amplification, becomes one of the outputs. In addition, the same wave is put through limiter-amplifiers and used to synchronize a 31.5 kilocycle blocking oscillator, whose waveform is used as a second output. Finally, the 31.5 kilocycle wave is divided by two to form the third output, at 15.75 kilocycles. The limiters are of such a design as to give phase stability between the 157.5 kilocycle sine wave and the two lower frequency outputs. The net result of this unit is to produce a precision phase shift or delay between input and output at any of the three mentioned frequencies or at any combination of them.

Provisions are also made for easy alignment of the phase shifter. The problem here is one of producing three sinusoidal inputs for the capacitor, having equal amplitudes, and phases which are 120 degrees apart. The first condition is attained by the use of a slide-back voltmeter having an "electric eye" type of indicator tube. With the "Calibrate Selection Switch" of Figure 2 in the upper position, the back-bias on the "eye" is balanced against the rectified output of one on the input phases. The amplitude of the other two phases are each in turn so adjusted as to give the same shadow-angle of the "eye," using the same back-bias setting. The phase adjustments are made by throwing the "Calibrate Selection Switch" to the downward position and the "Use-Calibrate" to the "Calibrate" position. This connects the "eye" indicator to the output of the dummy phase shifter. This latter is simply a phase shifter capacitor whose rotor has been removed; its three-phase input is connected in parallel with that of the main phase shifter. The output of the dummy phase shifter will be, to a high degree of precision, the vectorial sum of the three inputs. It is preferred to any other type of adder because of this high precision, which is possible due to its mechanical construction. The slide-back voltage is set to zero, so that the "eye" now becomes a null indicator. The phase of two of the three-phase inputs is, therefore, varied until the output of the dummy phase shifter, i.e., the vectorial sum of the three phases, becomes zero. Since their amplitudes have previously been set to be equal, their phases will be symmetrical if their vectorial sum is zero. It should be noted that this lineup procedure would not be sufficient for a four-phase capacitor; hence, the choice of the three-phase type.

The shaft of the phase-shifter capacitor is fitted with a dial for manual operation; this dial is calibrated from zero to one hundred for 360 degrees of rotation. Hence, one division corresponds to one-tenth of one percent of a horizontal period. For most measurements, this is the desired unit; a conversion to, say, microseconds can easily be made, of course. A motor drive is also provided, together with a control switch having two speeds in each direction; the maximum speed is 300 revolutions per minute, which is the greatest amount which the bearings of the phase shifter are capable of standing.

It can be seen that, if the accuracy of the phase shifter is plus-or-minus one percent of one revolution, this is equivalent to plus-or-minus one-tenth of one percent of one line repetition period or $\pm$ .001 H. The amount of jitter was too small to be measured.

The physical layout of the phase-shifter unit is shown in Figure 3.

The Frequency Divider

A block diagram of the frequency divider unit is shown in Figure 4. Binary scalers are employed throughout for reliability and simplicity of operation. The General Electric model 4SNA1A1 was used for convenience. The 31.5 kilocycle pulses from the phase-shifter unit are fed into the first group of three scalers, which are arranged so as to divide by seven. This output (which is 4500 cycles) is fed to a second group of three scalers, which are designed to give a further division by five, resulting in a 900 cycle signal, which becomes one of the outputs of the unit. In addition, this signal is fed to a group of four scalers, which are arranged to give a division ratio of fifteen, producing a sixty-cycle signal, which is used as a second output of the unit. Finally, this frequency is divided by two, by means of a single binary scaler, to produce the thirty-cycle signal. This is used to synchronize a blocking oscillator, whose waveform forms the third output.

When this unit is driven by the phase shifter output, it requires 5250 revolutions of the phase shifter shaft to cause the thirty-cycle output to go through one cycle. Since its maximum speed is 300 revolutions per minute, it would, therefore, take 17 1/2 minutes to go through one complete frame; a time of up to one quarter of this value would still be necessary, even if one did not care which field was to be taken, and if he proceeded in the nearest direction. To obviate this difficulty, a "Divider Change Switch" was provided. Its function is to close, at the press of a button, a feedback loop which changes the over-all divider ratio from 1050 to 1049. The output frequency is thus raised by the ratio of 1050/1049, which causes the output pulse to advance through the frame at the rate of 30 lines per second, which is 60 times as fast as that obtainable by rotating the phase shifter at 300 revolutions per minute. Hence, rapid setting of the phase of the 30 cycle output pulse is made possible, in an extremely simple manner.

The Frequency Multiplier

This unit consists of three stages which multiply the frequency by four, and one which multiplies by two. Hence, the input frequency of 157.5 kilocycles is multiplied by 128 to give an output of 20.16 megacycles. This may be used to produce markers having a time spacing of approximately .05 microseconds, for measuring rise times of transients. In some cases, it might be convenient to build a multiplier having a ratio of 100, thus giving markers which are spaced .001 H apart. In the present instance, it was deemed more useful to deal in microseconds and megacycles, for the purpose of measuring rise-times and bandwidths; furthermore, for measurements in terms of horizontal periods, the calibration on the phase shifter dial may be used.

Assembly of Units

Figure 5 shows the units assembled into one rack. At the top is the phase-shifter unit; next below is the divider unit; below this is the multiplier unit. The power supply at the bottom furnishes regulated d-c to all units. The normal interconnection is shown in Figure 1, although some particular use of the apparatus may require other interconnections.

Use of Instrument as Line Selector

Figure 6 shows the manner in which the instrument may be used as a line selector. The line selector input is shown as consisting of the

15.75 kilocycle driving pulses, since these are usually most readily available. The 30 cycle output is used to trigger the sweep of the oscilloscope. The writing speed of the sweep should be that which is appropriate for the waveform to be examined. The use of the 20-megacycle timing markers is optional; if they are used, they should be fed to the Z-axis of the oscilloscope. The television picture monitor is used to indicate what part of the picture is being examined. This is done by placing the 30-cycle pulse on one of the signal electrodes of the picture tube in any convenient manner. Thus, the timing of the beginning of oscilloscope sweep with respect to the picture is clearly indicated. The approximate position of the 30-cycle marker is then set by depressing the "Divider Change Switch" on the divider unit. The exact position is then selected by means of the phase shifter knob.

Rise time of transients in the picture may be measured by the use of the 20-megacycle markers. Pulse widths or delays may be measured by means of the calibration of the phase shifter dial. This is done by positioning one of the points, between which the measurement is to be made, on the oscilloscope vertical cross-hair, by means of the phase shifter. The dial reading is then noted. When the other point is similarly positioned, the dial reading is again noted. The time difference is thus obtained to an accuracy of $\pm$ .001 H. This is particularly useful in measuring synchronizing pulses for width and position, to ascertain conformance to standards.

In summation, it may be said that the high stability and accuracy of the instrument allows it to perform its function as a line selector in an extremely useful manner.

Other Uses of Instrument

By synchronizing the oscilloscope from the 15.75 kilocycle output of the phase shifter, the same precision measurements as above may be made on phenomena occurring at this rate, with the added advantage that more brightness is obtainable on the oscilloscope, due to the higher repetition rate.

If the horizontal sweep of a picture monitor is synchronized from the 15.75 kilocycle output of the instrument, and the vertical sweep from the 60-cycle output, they may be conveniently phased with respect to the incoming video signal. This is useful, for example, for examining the video during the retrace of the monitor sweeps. Conversely, the video blanking and synchronizing pulses may be made to occur during the forward trace of the monitor; hence, familiar "pulse-cross" pattern may be simply obtained with the aid of this instrument. In summation, the instrument may be used wherever continuous delay of the horizontal and/or vertical synchronizing pulses is desired.

A specialized application of the instrument is shown in Figure 7. In this instance, it is used as part of a cross-hatch generator for checking geometric distortion of camera sweeps. The 900-cycle output of the divider unit is used to synchronize pulses which are roughly 2 1/2 lines or about 150 microseconds in length. Similarly, the 157.5 kilocycle output is multiplied by 2 and used to synchronize pulses which are approximately one-half microsecond in length. These two waveforms are then added together and clipped. When the resulting waveform is displayed on a picture monitor, a cross-hatch pattern is seen, which would consist of 15 horizontal lines and 20 vertical lines, were it not for blanking and retrace times. Figure 8 shows the proposed standard RTMA linearity chart, which will soon appear in an IRE Measurement Standard. This chart is placed in front of the camera, and the resulting video is displayed, along with the cross-hatch pattern on a picture monitor. The geometric distortion is less than two percent if all the cross-hatch intersections fall within the outer circles. In order to make such a measurement, it is necessary to be able to move the cross-hatch pattern electrically, in order to effect an optimum "fit" between it and the test pattern image. This is easily done by means of the phase shifter control. The horizontal lines are first positioned properly, after which the vertical lines are adjusted.

Variations in the velocity of the horizontal deflection of a picture monitor or receiver may be measured by comparing electrical displacement, as measured on the phase shifter dial, with the displacement of one of the lines of the cross-hatch, as the latter are moved across the screen by means of the phase shifter. Since the calibration of the latter is linear and accurate to 0.1 percent of a line, such a measurement will detect and measure accurately very small velocity errors. These might otherwise be greatly underestimated, using methods of measurement involving only a stationary cross-hatch pattern.

The frequency-divider unit itself has been put to great use around the laboratory, in a number of specialized applications. For example, it has been used in conjunction with a 31.5 kilocycle oscillator, as a simple synchronizing generator.

The instrument described has not only served its purpose well as a line-selector, but has found a number of other useful applications. No doubt, still further different uses will be found, which will further increase its utility in the television laboratory.

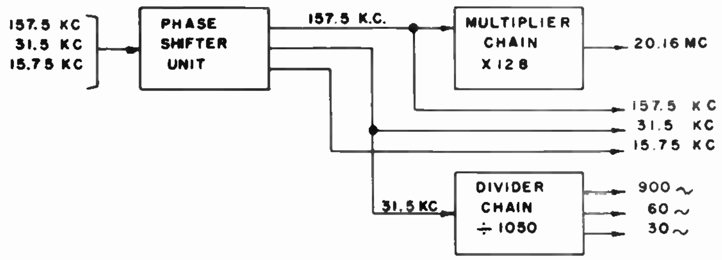


Fig. 1 - Over-all block diagram.

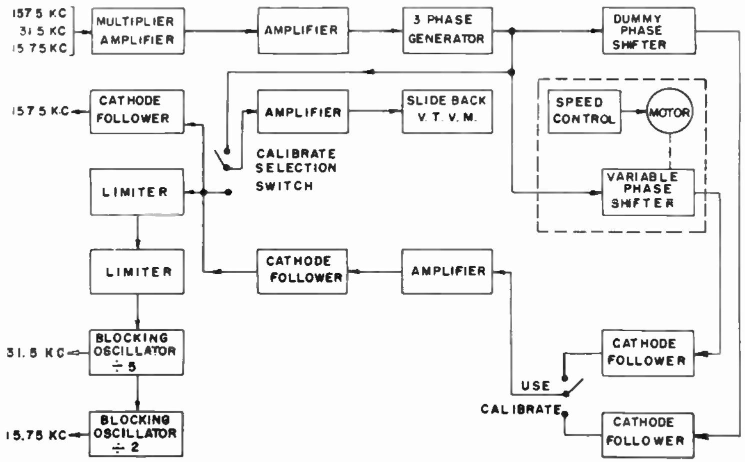


Fig. 2 - Block diagram, phase shifter unit.

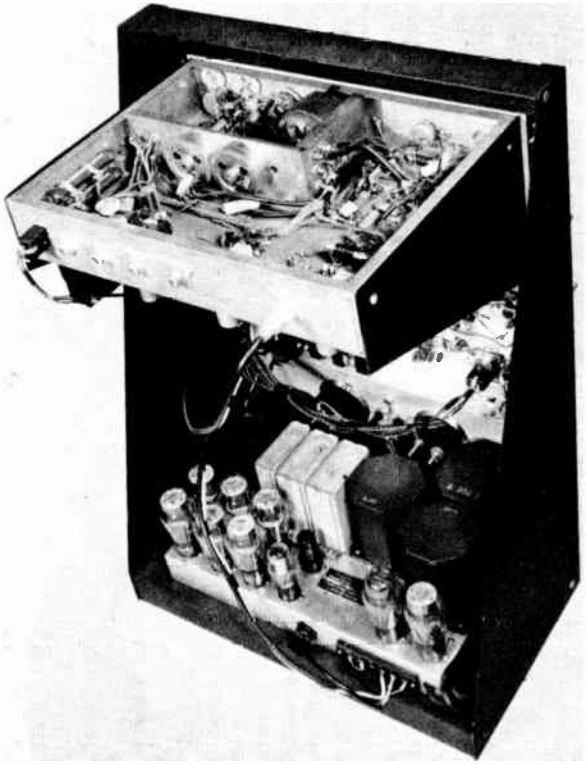


Fig. 3 - Rear view, showing phase shifter unit.

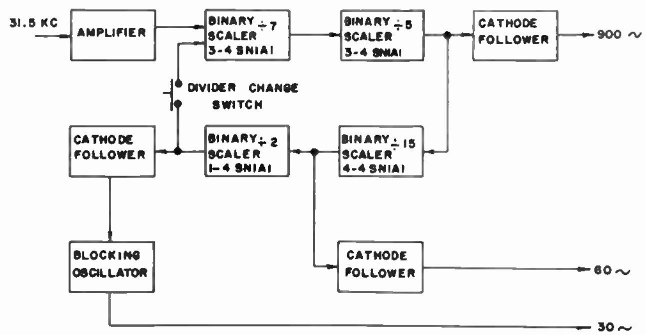


Fig. 4 - Block diagram, divider chain.

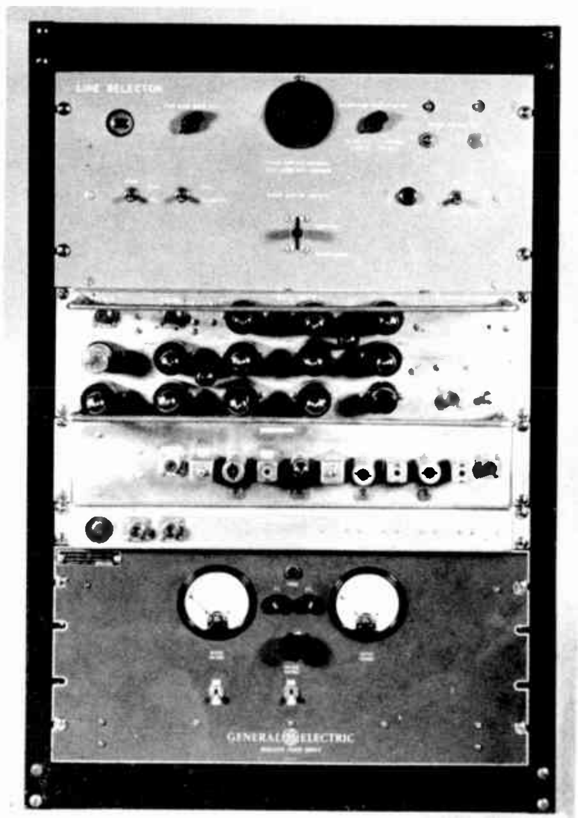


Fig. 5 - Front view, line selector.

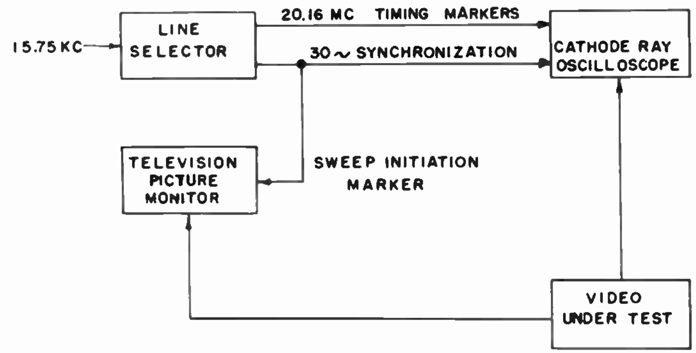


Fig. 6
Block diagram, showing use
of unit as line selector.

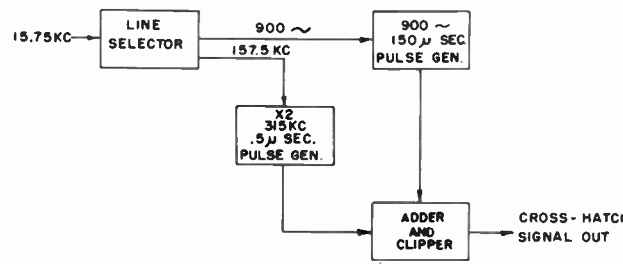


Fig. 7
Block diagram, showing use
of unit as cross-hatch generator.

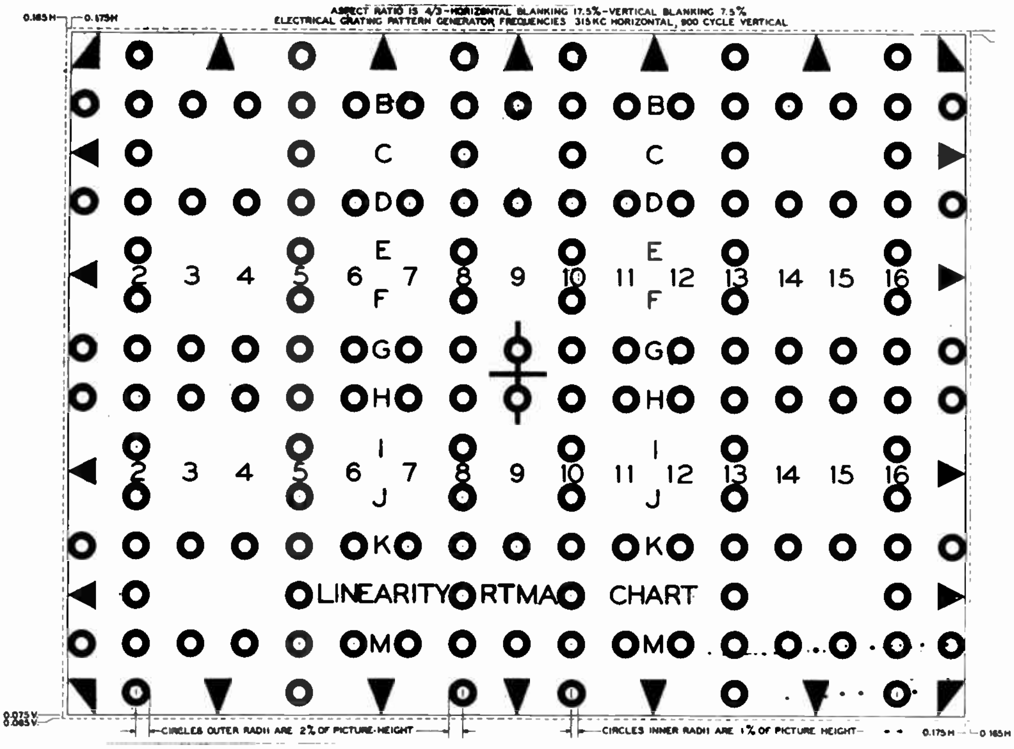


Fig. 8 - Proposed RTMA linearity chart.

COLORIMETRIC PROPERTIES OF GAMMA-CORRECTED COLOR TELEVISION SYSTEMS

Donald C. Livingston
Sylvania Electric Products Inc.
Physics Laboratories
Bayside, New York

This paper will present a detailed comparison between two different color television systems both of which fall within the broad framework of what might be called the generalized NTSC system. It will be shown that one of these systems is appreciably superior to the other in overall performance. The paper opens with s brief description of the generalized NTSC system. This is followed by the actual comparison of systems, which is carried out with the aid of a number of "system parameters" whose numerical values reveal the differences in performance between the two systems. These differences will indicate the superiority of one system over the other. Finally, some considerations on perceptibility of camera and receiver noise will provide additional evidence of the superiority of the better system.

The Generalized NTSC System

The general form of the color television system to be considered in this paper is shown in Fig. 1. At the extreme left, indicated symbolically by a tree, is the subject being scanned. It is regarded as furnishing a light signal with tristimulus values[1] R,G,B, these values being defined relative to the primaries characterizing the phosphors on the kinescope of the color television receiver in the right half of the diagram. Immediately to the right of the subject is the color camera, consisting of light receptors with three different spectral responses together with electrical transducers which develop color signals E_R, E_G, E_B which, to an acceptable approximation, are linear functions of R,G,B. The relative levels of these color signals are defined with respect to each other so that $E_R = E_G = E_B$ represents reference white.

A distinctive characteristic of the National Television System Committee (NTSC) color television system is the manner in which the color signals E_R, E_G, E_B are transformed into signals suitable for modulation onto an rf carrier. This transformation is carried out in the coder, shown as the element following the camera in Fig. 1. The coder uses the three color signals to produce a luminance signal E_L and two color-difference signals S_R and S_B. The color-difference signals are modulated at quadrature onto s subcarrier to form a chrominance signal. Finally, the chrominance signal and the luminance signal are added together to form a color picture signal which is then amplitude modulated onto an rf carrier in the usual manner. The luminance signal is equivalent to the ordinary video signal in monochrome television and will be interpreted as such when the color picture signal is displayed on a monochrome receiver.

In a color receiver, the color picture signal is demodulated from the rf carrier, and the color-difference signals are demodulated from the chrominance signal. Thus, the luminance signal and the color-difference signals are recovered except for certain small errors which can be neglected. These signals are then operated upon by the decoder in Fig. 1 to yield signals which are of such forms that, after adjustment of their levels by passage through video amplifiers, they sre capable of eliciting from the color kinescope a luminous response with tristimulus values R',G',B' which are equal to R,G,B or to some fixed constant times R,G,B.

Actual designation of forms for the various signals must be made in accordance with the characteristics of system components. In particular, the kinescope transfer characteristics must be considered. It is known that color kinescopes follow power-law characteristics of the form

$$Y = aE^{\gamma} \tag{1}$$

over a range $E_{min} \leq E \leq E_{max}$ which is sufficiently wide that signal excursions beyond these limits rarely occur. In eq. (1), Y is the reproduced luminance and E is the grid drive of the kinescope measured with respect to cutoff; a and γ are constants, the latter being called the gamma exponent for the kinescope. The numerical value of γ has been found experimentally to be approximately 2.2 for a number of color kinescopes.

It follows that the decoder output signals E_R'', E_G'', E_B'' in Fig. 1 should be arranged to be equal, as nearly as possible, to $E_R^{1/\gamma}, E_G^{1/\gamma}, E_B^{1/\gamma}$. There appears to be one simple way to bring this about rigorously, viz., by letting E_L, S_R, S_B be linear combinations of $E_R^{1/\gamma}, E_G^{1/\gamma}, E_B^{1/\gamma}$ so as to require the decoder to perform only linear operations such as addition and subtraction in order to yield the desired output signals. The best choice of signals meeting this requirement can be shown to be[2]

$$E_L = .299\, E_R^{1/\gamma} + .587\, E_G^{1/\gamma} + .114\, E_B^{1/\gamma} \equiv E_Y' \tag{2a}$$

$$S_R = E_R^{1/\gamma} - E_Y' \tag{2b}$$

$$S_B = E_B^{1/\gamma} - E_Y', \tag{2c}$$

in which numerical constants have been arranged so that the luminance signal carries all or nearly all of the luminance information when the subject chromaticity is in the neighborhood of reference white, the receiver primaries and the particular reference white used being those specified[3] by NTSC Panel 13. Actually, it will be seen later that, in spite of the fact that the system described by eqs. (2) yields the correct colors, this system nevertheless possesses some characteristics which render it somewhat inferior to another system even though the second system permits small errors in color reproduction to occur.

This second system is one in which the luminance signal has the form[4]

$$E_L = E_Y^{1/\gamma}\,, \tag{3}$$

wherein

$$E_Y \equiv .299\,E_R + .587\,E_G + .114\,E_B \tag{4}$$

is a quantity exactly equal to the subject luminance. The color-difference signals are the same as before. The method of decoding in this system is exactly the same as in the first; but since E_L is no longer a linear combination of $E_R^{1/\gamma}, E_G^{1/\gamma}, E_B^{1/\gamma}$, it is not possible to recover these quantities exactly. However, $E_Y^{1/\gamma}$ as given in eqs. (3) and (4) differs only slightly from E_Y' in eq. (2a) except for saturated colors, so relatively little error in recovery of the desired signals occurs. It will soon be shown, in fact, that the very presence of this error actually does enough good in various ways to more than compensate the little harm which it does to the accuracy of color reproduction.

The principal purpose of the remainder of this paper will be to examine the performance of the two systems whose signal specifications have just been described and to determine which of the two systems appears to offer the better combination of performance characteristics. It will be assumed that both systems employ a decoder operating in accordance with the relations

$$E_R'' = E_L + S_R \tag{5a}$$

$$E_G'' = E_L - \frac{.299}{.587}S_R - \frac{.114}{.587}S_B \tag{5b}$$

$$E_B'' = E_L + S_B\,. \tag{5c}$$

It is readily established[2] that when $E_L = E_Y'$,

$$E_R'' = E_R^{1/\gamma} \qquad E_G'' = E_G^{1/\gamma} \qquad E_B'' = E_B^{1/\gamma}\,, \tag{6}$$

whereas when $E_L = E_Y^{1/\gamma}$,

$$E_R'' = E_R^{1/\gamma} + \Delta E_Y \tag{7a}$$

$$E_G''' = E_G^{1/\gamma} + \Delta E_Y \tag{7b}$$

$$E_B'' = E_B^{1/\gamma} + \Delta E_Y\,, \tag{7c}$$

in which

$$\Delta E_Y \equiv E_Y^{1/\gamma} - E_Y'\,. \tag{8}$$

The analysis of the systems will still require that the proportionalities between R,G,B and E_R, E_G, E_B be known and that those between E_R'', E_G'', E_B'' and R',G',B' be known. It can be shown that the desired relations are[2]

$$E_R = \frac{R}{.286} \qquad E_G = \frac{G}{.261} \qquad E_B = \frac{B}{.453} \tag{9}$$

and

$$R' = .286nE_R'' \qquad G' = .261nE_G'' \qquad B' = .453nE_B'' \tag{10}$$

The analysis can now proceed. It will involve the definition of a number of quantities to be called system parameters, each of which will be introduced as the need arises.

Color Fidelity

The most important property of a color television system is clearly color fidelity. Since color is itself a three-dimensional quantity, being describable in terms of luminance, hue, and saturation, it is necessary to measure color fidelity in a way which takes into account distortions in any one of these three components. It will be convenient to introduce two system parameters for this purpose, one being a scalar and the other a two-dimensional vector. The former will be defined as

$$F \equiv \frac{Y'}{nY} \tag{11}$$

and called the Luminance Fidelity. The luminance of the subject being scanned is Y, and that of the reproduced image point on a color kinescope is Y'. Since Y and Y' are intended only to be proportional rather than equal, it is useful to designate the desired relation as $Y' = nY$, with n a constant whose value can be adjusted by the viewer. It expresses the effect of the contrast control. It can be seen from eq. (11) that F should be unity for correct luminance reproduction.

The vector will be called the chromatic distortion vector and will be most conveniently described through the three linearly dependent components

$$\Delta r \equiv r' - r \qquad \Delta g \equiv g' - g \qquad \Delta b \equiv b' - b, \tag{12}$$

in which r,g,b and r',g',b' are trichromatic coefficients of the scanned color and the reproduced color, respectively. If these components are all identically zero, then it follows that the system reproduces the hue and saturation, or the chromaticity, of a color correctly even though the luminance might still be in error.

To apply eqs. (11) and (12) to the analysis of a given system, it will be necessary to deduce relations for Y and Y' in terms of R,G,B. These can be shown[2] to be

$$Y = .33R + .71G + .08B \qquad (13)$$

and

$$Y' = .316n\left\{.299\, E_R''^{\gamma} + .587\, E_G''^{\gamma} + .114\, E_B''^{\gamma}\right\}, \qquad (14)$$

wherein E_R'', E_G'', E_B'' are related to R,G,B through various of the preceding equations. The quantity n in eq. (14) is the same n appearing in eq. (11).

Now with E_R'', E_G'', E_B'' given by eqs. (6) for the $E_L = E_Y'$ system and by eqs. (7) for the $E_L = E_Y'^{1/\gamma}$ system, the Luminance Fidelity F in eq. (11) is immediately calculable. The results, for $\gamma = 2.2$, are shown in Fig. 2 for the $E_Y'^{1/\gamma}$ system; it is seen that F = 1 at Illuminant C, that F is not much larger than unity over a considerable range of chromaticities around Illuminant C, but that it seems to become excessive at saturated chromaticities. On the other hand, F is exactly unity over the entire chromaticity gamut in the E_Y' system.

To determine the Chromatic Distortion Components for the two systems, one must use eqs. (6) through (10) in eqs. (12). The result for the E_Y' system turns out[2] to be that $r \equiv g \equiv b \equiv 0$; that is, there is no chromatic distortion. Thus, it is seen that the E_Y' system appears to have perfect color fidelity. Actually, however, since it is necessary in practice to band-limit the chrominance signal before adding it to the luminance signal, the demodulated color-difference signals in a color receiver will not be able to duplicate the most rapid variations in the original color-difference signals at the transmitter, and the result is that the reproduced color picture fails to resolve horizontal chromatic detail as finely as it does luminance detail. Thus, the above considerations on color fidelity are applicable only to large-area color detail.

Evaluation of eqs. (12) for the $E_L = E_Y'^{1/\gamma}$ system leads to values for the Chromatic Distortion Components which are not zero in general. The results of this evaluation are shown in Fig. 3, in which the tail of each arrow represents the subject chromaticity and the head of each represents the corresponding image chromaticity. It appears that, in general, the chromatic distortion has the character of a desaturation which is greatest for the most saturated colors. In particular, it has its maximum effect at saturated red. However, it is noteworthy that the eye is known to be very insensitive to the forms of chromatic distortion which occur in this system, so the distortion can be expected to be much less noticeable than the diagram would suggest. This seems, in fact, to be the case; for in actual laboratory tests, trained observers had difficulty in distinguishing color differences between pictures formed by the E_Y' and $E_Y'^{1/\gamma}$ systems.

Monochrome Luminance Fidelity

Closely related to color fidelity as described above is the response of a receiver to the luminance signal alone. This refers, for example, to the reproduction of luminance by a conventional monochrome receiver when receiving color picture signals. What is more important, however, is that it also refers to the performance of a color television receiver when, because of band-limiting of the chrominance signal, the demodulated color-difference signals are unable to vary rapidly enough to duplicate the original color-difference signal waveforms in the transmitter. Under this condition, the color receiver can respond to detail only to the extent that the transmitted luminance signal varies.

The system parameter to be introduced to measure this property of the system will be called the Monochrome Luminance Fidelity F_m, and it will be defined as

$$F_m \equiv \frac{Y_{Lm}'}{nY}, \qquad (15)$$

in which Y_{Lm}' denotes the luminance displayed on a monochrome receiver when the subject luminance is Y. Ideally, Y_{Lm}' should equal nY, so that proper monochrome luminance fidelity is indicated by $F_m = 1$.

It can be shown[2] that

$$Y_{Lm}' = .316nE_L^{\gamma}, \qquad (16)$$

so that eq. (15) can be evaluated with the aid of eqs. (9), (13), and (16) together with eqs. (3) and (4) for the $E_L = E_Y'^{1/\gamma}$ system or with eq. (2a) for the $E_L = E_Y'$ system. The result of this evaluation is shown in Fig. 4 for the E_Y' system. It is seen that F_m falls considerably as the chromaticity departs from Illuminant C; the decrease is slow at first, then becomes very rapid. This situation seems serious for two reasons. The lesser of these in importance pertains to the behavior of monochrome receivers; it appears that a monochrome receiver will display much less than the desired amount of luminance when the subject color is highly saturated red or blue. Thus, the diagram shows that only 7.3% of the desired luminance will appear at saturated blue when $\gamma = 2.2$.

The significance of Fig. 4 for a color receiver is due to band-limiting of the chrominance signal. When a subject is finely detailed in both luminance and chromaticity, both the received luminance and color-difference signals must vary rapidly in order to reproduce the color detail correctly. If, however, the received color-difference signals are prevented by band-limiting from varying sufficiently fast, the only detail that will be reproduced will be that described by the luminance signal. But it has been seen that the luminance signal is much too small at saturated colors for proper representation of subject luminance. Its variations, too, will be too small if $E_L = E_Y'$; hence, the luminance detail will be suppressed. Consequently, the reproduced picture will be lacking in both luminance detail and chromatic detail even though, due to the action of the slowly-varying color-difference signals, the mean displayed luminance will be correct (since $\overline{F} = 1$).

Evaluation of eq. (15) for the $E_L = E_Y^{1/\gamma}$ system leads to the result $F_m = 1$, which means that color picture signals using this luminance signal will not only lead to correct luminance display on a monochrome receiver but to better rendition of luminance detail on a color receiver than for the $E_L = E_Y'$ system.

Constant-Luminance

There are a number of reasons for the desirability of designing the NTSC color television system so that variations in the chrominance signal amplitude will not affect the luminance displayed by a color receiver. Thus, when variable multipath transmission conditions cause periodic reinforcement and cancellation of the subcarrier, it would lead inevitably to periodic fluctuations in the saturation of the displayed colors. This would be sufficiently objectionable by itself, but it would be considerably more so if the displayed luminance also varied periodically. Moreover, it has been demonstrated that noise and spurious signals in the chrominance channel are less perceptible if they affect only the displayed chromaticity than if they affect both displayed chromaticity and luminance. These conclusions are expressed by the Constant-Luminance Principle, which is fully discussed elsewhere[5].

It is convenient to introduce two system parameters to measure the degree of adherence of a given system to the constant-luminance principle. One of these is the Constant-Luminance Index K, defined by

$$K \equiv \frac{Y_L'}{Y'}, \qquad (17)$$

in which Y_L' is the luminance reproduced by a color television receiver in the absence of the chrominance signal and is equal to Y_{Lm}' for the system being considered. Thus, through comparison of eqs. (11), (15), and (17), one can write

$$K = \frac{F_m}{F}, \qquad (18)$$

permitting K to be computed from previous system parameters.

In a color television system, K is a measure of the effect of the chrominance signal on the displayed luminance. It is desirable that this signal have no such effect, so it is desired that $K \equiv 1$. Although K can equal unity in a system wherein γ is unity, no satisfactory way appears to be available for arrangement of a system wherein $\gamma \neq 1$ so that K can be unity for all chromaticities. When eq. (18) is evaluated for the E_Y' and $E_Y^{1/\gamma}$ systems with $\gamma = 2.2$, F and F_m values being obtained as previously described, the results are as shown in Figs. 4 and 5. It is seen that $K = 1$ at Illuminant C but that its value decreases as the chromatic saturation increases, particularly toward blue. It is also evident that the $E_Y^{1/\gamma}$ system approaches somewhat closer to the ideal than does the E_Y' system.

The second parameter to be introduced to measure the adherence of a system to the constant-luminance principle will be the Subcarrier Luminance Sensitivity L, defined by[4]

$$L \equiv \frac{\Sigma^{1/\gamma}}{A_o}\left(\frac{\delta Y'}{Y'}\right)_{max}, \qquad (19)$$

in which $\Sigma \equiv R+G+B$ is the total tristimulus value of the subject color, A_o is the amplitude of a spurious signal in the chrominance channel, Y' is the displayed luminance in the absence of the spurious signal, and $\delta Y'$ is the change in displayed luminance due to the effect of the spurious signal. The ratio $(\delta Y'/Y')_{max}$ in eq. (19) is the value which results when the phase of the spurious signal is such that $\delta Y'$ is maximized at fixed Y'. In practice, it is of greater interest to compare L values from one system to another and at various chromaticities than to fix a significance to the actual numerical value of L. In general, the greater the value of L, the more perceptible will be the effect of a spurious signal in the chrominance channel on the reproduced picture.

Values computed for L by eq. (19) are shown in Figs.6 and 7 for the E_Y' and $E_Y^{1/\gamma}$ systems, respectively, the value $\gamma = 2.2$ being used for the gamma exponent. It appears that the $E_Y^{1/\gamma}$ system is considerably less susceptible to spurious signals in the chrominance channel than is the E_Y' system.

Perceptibility of Luminance Noise

It remains to consider one further respect in which the systems with $E_L = E_Y^{1/\gamma}$ and with $E_L = E_Y'$ may be compared. In the course of subjective experimental tests of the two systems, it was noted that the picture produced on a monochrome receiver by a color picture signal was considerably more noisy when $E_L = E_Y'$ than when $E_L = E_Y^{1/\gamma}$. It can be shown quite readily that the observed behavior is logically to be expected, whether the displayed noise originates in the receiver itself or in the camera. These same conclusions are also applicable to noise reaching a color picture via the luminance channel of a color receiver.

When the displayed luminance Y_L' due to the luminance signal alone is perturbed by an amount $\delta Y_L'$ through the effect of noise from all possible sources, it results that the perceptibility fractions $P[E_L] \equiv \delta Y'_L/Y'_L$ associated with use of the luminance signals E'_Y and $E_Y^{1/\gamma}$ are[6]

$$P[E'_Y] = .299\, E_R^{\frac{1}{\gamma}-1} \frac{\delta E_R}{E'_Y} + .587\, E_G^{\frac{1}{\gamma}-1} \cdot \frac{\delta E_G}{E'_Y} + .114\, E_B^{\frac{1}{\gamma}-1} \frac{\delta E_B}{E'_Y} + \gamma \frac{\delta E_L}{E'_Y} \tag{20a}$$

and

$$P[E_Y^{1/\gamma}] = .299 \frac{\delta E_R}{E_Y} + .587 \frac{\delta E_G}{E_Y} + .114 \frac{\delta E_B}{E_Y} + \gamma \frac{\delta E_L}{E_Y^{1/\gamma}}, \tag{20b}$$

respectively. In eqs. (20), the quantities δE_R, δE_G, δE_B, and δE_L represent spurious perturbations in the red, green, and blue camera output signals and in the received luminance signal, respectively. Several specific conclusions may be drawn from eqs. (20). First, suppose that only receiver noise is present. Then

$$\frac{P[E'_Y]}{P[E_Y^{1/\gamma}]} = \frac{E_Y^{1/\gamma}}{E'_Y} \tag{21}$$

It follows that $P[E'_Y] > P[E_Y^{1/\gamma}]$ in this case since $E_Y^{1/\gamma} > E'_Y$. For saturated colors, this ratio can reach rather significant proportions, as can be seen in the extreme case of saturated blue, for which case the ratio become 3.3 for $\gamma = 2.2$.

Next, suppose that only red and blue camera noise are present, and consider a color close to the red primary, so that $E'_Y \cong .299\, E_R^{1/\gamma}$ and $E_Y = .299\, E_R$. Then

$$\frac{P[E'_Y]}{P[E_Y^{1/\gamma}]} = \frac{\frac{\delta E_R}{E_R} + \frac{.114}{.299}\left(\frac{E_B}{E_R}\right)^{1/\gamma} \frac{\delta E_B}{E_B}}{\frac{\delta E_R}{E_R} + \frac{.114}{.299}\left(\frac{E_B}{E_R}\right) \frac{\delta E_B}{E_B}}, \tag{22}$$

from which it is obvious that the perceptibility ratio approaches unity if the red noise predominates over the blue noise, while it approaches $(E_B/E_R)^{\frac{1}{\gamma}-1} > 1$ if blue noise predominates. Analogous conclusions are readily reached for other combinations of signals and noise sources, but space does not permit elaboration on details here. The reader is referred to reference 6. It will suffice to say that examination leads, in nearly every case, to the conclusion that noise from any given source is more perceptible when E'_Y is used as luminance signal than when $E_Y^{1/\gamma}$ is used. There exist no cases in which the noise susceptibility of E'_Y is less than that of $E_Y^{1/\gamma}$.

Summary and Conclusions

The preceding discussion has demonstrated that it is possible to predict theoretically the properties of color television systems using luminance signals of the forms E'_Y and $E_Y^{1/\gamma}$, respectively. It has been indicated that the system with $E_L = E_Y^{1/\gamma}$ is superior to that with $E_L = E'_Y$ in respect to constant-luminance adherence, monochrome luminance fidelity, and perceptibility of noise in the luminance signal. It is slightly inferior in color fidelity, but it has been judged through experimental tests with the two systems that the color degeneration is of negligible magnitude while the accompanying benefits are readily appreciable.

References

1. Definitions of colorimetric terms employed in this paper may be found in Wintringham, Color Television and Colorimetry, Proc. IRE, 39, 1135-1172 (1951).

2. Livingston, Colorimetric Analysis of Gamma-Corrected Shunted Monochrome Simultaneous Color Television Systems - Part I, prepared for NTSC Panel 13; July 1, 1952. This report and Parts II, III, and IV of the same series can be obtained by writing to Donald C. Livingston, Physics Laboratories, Sylvania Electric Products Inc., Bayside, New York. Most of the material in the present paper is condensed from Parts I, II, and III of the colorimetric analysis series.

3. NTSC Panel 13, Report on Color Video Standards; Oct. 23, 1951.

4. Livingston, op. cit., Part II.

5. Livingston, op. cit., Part III.

6. Livingston, Perceptibility of Noise in the Luminance Signal prepared for NTSC Panel 13; January 5, 1953. Copies may be obtained by writing to the author. See reference 2 for address.

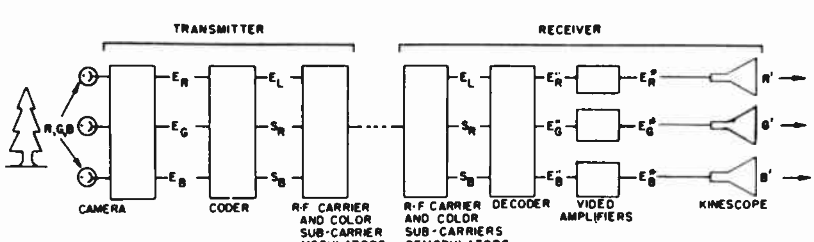


Fig. 1
Generalized form of the NTSC color television system.

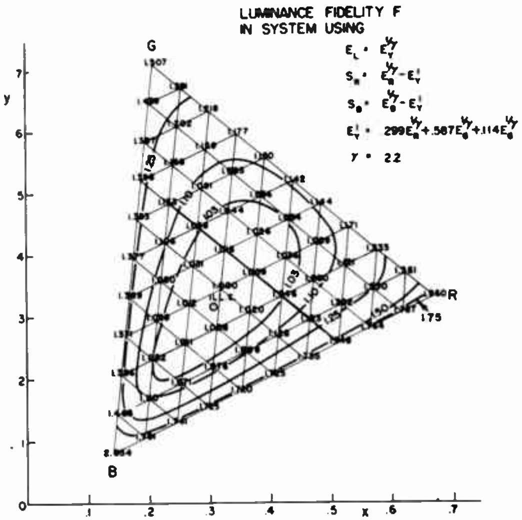


Fig. 2
Luminance fidelity F in the $E_Y^{1/\gamma}$ system.

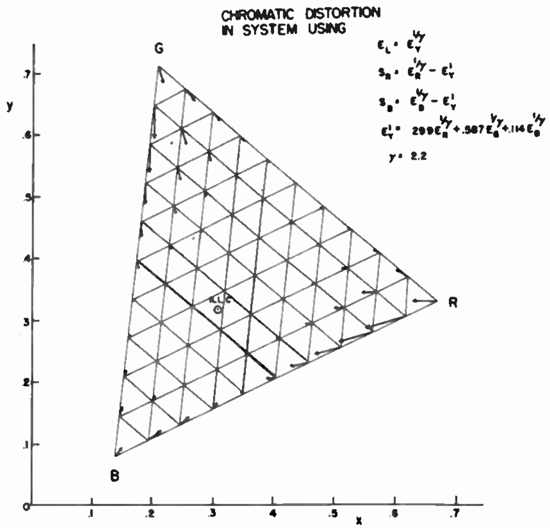


Fig. 3
Chromatic distortion in the $E_Y^{1/\gamma}$ system.

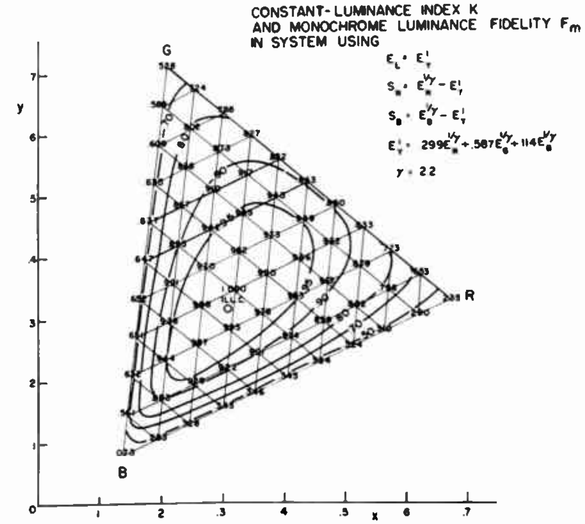


Fig. 4
Monochrome luminance fidelity F_m and constant-luminance index K in the $E_Y^{\prime}$ system.

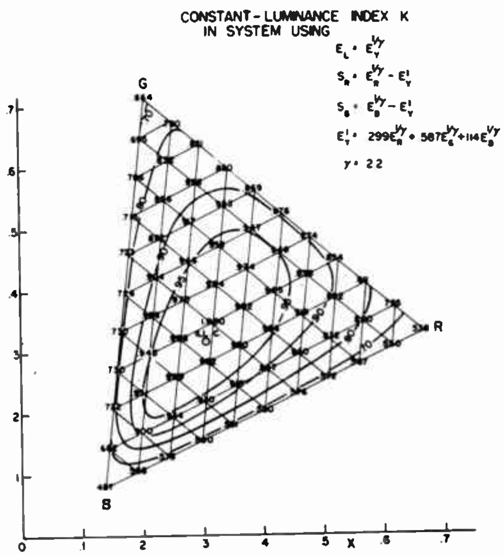


Fig. 5
Constant-luminance index K in the $E_Y^{1/\gamma}$ system.

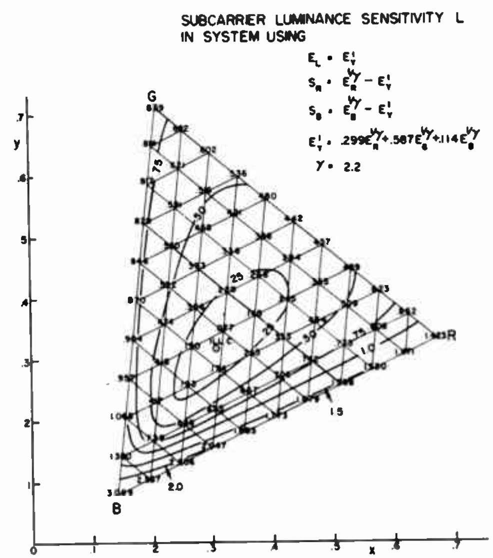


Fig. 6
Subcarrier luminance sensitivity L in the $E_Y^{\prime}$ system.

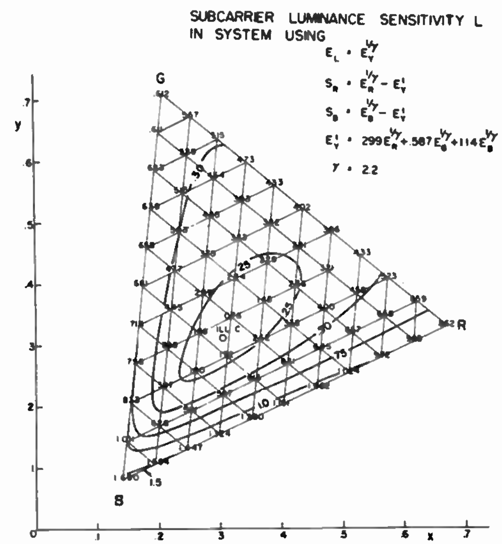


Fig. 7
Subcarrier luminance sensitivity L in the $E_Y^{1/\gamma}$ system.

PHASE MEASUREMENTS AT SUBCARRIER FREQUENCY IN COLOR TELEVISION

Arthur P. Stern
General Electric Company
Syracuse, New York

Introduction

In the Color Television System studied by the National Television Systems Committee (NTSC), the chrominance information modulates two subcarriers of the same frequency, but 90° out of phase relative to each other. The subcarrier frequency is located approximately 3.58[1] mc. away from the main carrier. The chrominance information is recovered in the receiver by synchronous detection.

In order to secure high quality color reproduction, it is essential that the components of the color signal and the detecting voltages at the receiver satisfy accurately the NTSC signal requirements. For reliable adjustment of transmitter and receiver and for experimental investigations, the possibility of making accurate phase measurements is of great importance.

The purpose of this paper is to describe the principles and the operation of phase measuring equipment built in the Electronics Laboratory of the General Electric Company, Syracuse, New York.

Purpose of the Measurements

The objective of phase measurements in Color Television is generally either of the following:

1. The phase difference of two subcarrier voltages must be measured, the angle assuming any arbitrary value. The subcarrier voltages can either come from separate sources and exist simultaneously, or they might be sequential parts of the same signal. This is the case, for example, when an NTSC-signal carries information corresponding to two vertical bars of different colors and the phase of the corresponding subcarrier pulses must be determined relative to each other or relative to a reference axis.

2. Two signals must have a certain prescribed phase difference relative to each other. The objective is rather "alignment" than "measurement." In this case, special (possibly simple, depending on the expected angle) instruments can be designed to allow an easy alignment.

The purpose of the equipment to be described is mainly to permit measurements, but naturally it can also be used for alignment purposes.

Principles of Measurement

Equipment for measuring phase at a given frequency can be designed on the basis of various principles. Elegant solutions permit the display of the phase angle to be measured on the screen of an oscilloscope. However, it was felt that the procedure to be described is capable of giving most accurate results, the error being less than a degree.

The method of measurement consists of comparing the phase of each of the signals or signal portions, the phase difference of which is to be measured with the phase of the output of a calibrated phase shifter.

This phase shifter permits the continuous variation of the phase of a subcarrier voltage S_{var}. The signals (or signal portions) the phase difference of which is to be determined be S_{x1} and S_{x2}. The output S_{var} of the phase shifter is first added to S_{x1} and the sum is displayed on an oscilloscope (or measured in some other way). Phase and amplitude of S_{var} are varied until cancellation occurs and the position of the phase shifter α_1 noted. The same procedure is repeated for S_{x2} with a reading α_2 on the phase shifter dial. The phase difference between S_{x1} and S_{x2} is then given by $(\alpha_1 - \alpha_2)$.

A convenient continuous phase shifter can be built with the aid of a multiphase "phase shift condenser." The design of this device (if perfectly built) is such that the actual angle of rotation of the condenser shaft is equal to the electrical angle.[2]

In order to eliminate errors due to inevitable imperfections of the condenser and its associated circuits, calibration is necessary. It is more convenient to calibrate and it is also easier to avoid imperfections in the circuitry at low frequencies. Therefore, phase shifting is accomplished at a comparatively low frequency and the output of the phase shifter is then heterodyned to subcarrier frequency. This operation does not affect the relative phase and, therefore, the calibration of the low frequency phase shifter is valid for the phase shifted subcarrier voltage obtained by heterodyning.

The Phase Measuring Equipment

Figure 1 represents a block diagram of the phase measuring apparatus. The components of this equipment are:

1. Phase Shifter (Figure 2) - The phase of a continuous low frequency voltage with constant phase, F_{Lcon}, is varied by means of a 3-phase condenser. The three inputs to the phase shift

[1] Equipment described in this paper has been built for the "old" subcarrier frequency 3.89 mc.

[2] Blackburn, J. F. - Components Handbook, McGraw Hill, New York, 1949.

condenser are supplied by a phase splitting network and must be 120° out of phase and equal in amplitude. The fine adjustment of amplitude and phase of the inputs is done with the help of another phase shift condenser, the rotor of which has been removed. If the three inputs have proper phase and amplitude, the output of this condenser is zero.

The output of the phase shifter is a voltage, F_{Lvar}, the phase of which varies continuously with the setting of the condenser. To make accurate readings possible, the dial of the phase shifter is provided with a 1:5 gear ratio. Readings of a fraction of a degree can be made. (The low frequency at which phase shifting is accomplished may be chosen arbitrarily. In the equipment described, 157.5 kc = 10 x line frequency was used.)

2. Converter (Figure 3) - The output F_{Lvar} of the phase shifter heterodynes a frequency F_H of constant phase, so that after several amplifying and filtering stages, we obtain a subcarrier voltage S_{var}, the phase of which varies according to the phase of F_{Lvar}. The frequency F_H is obtained by mixing a subcarrier wave of constant phase, S_{con}, with a low frequency wave of constant phase F_{Lcon}.

3. Adder (Figure 4) - In the adder, the output S_{var} of the converter is passed through an amplifier with variable gain and added to the reference signal S_o. The sum is displayed on the screen of an oscilloscope. Amplitude and phase of S_{var} are adjusted to obtain cancellation. S_o and S_{var} are 180° out of phase relative to each other. This sets the reference angle α_o of the phase shifter. The same procedure is applied to the signal S_x the phase of which relative to S_o is to be measured and the difference of readings $(\alpha_x - \alpha_o)$ on the dial of the calibrated phase shifter is the phase difference between S_x and S_o.

It has already been mentioned that in practical cases, the reference signal S_o and the unknown S_x are not continuous waves, but sequential parts of the NTSC signal which do not occur at the same time, but are repeated periodically at horizontal or vertical rate. In this case, cancellation is obtained only for a certain portion of the signal. The above described method, however, is still applicable, as S_{var} is a continuous wave.

As reference signal in the above procedure, one uses very often the "color synchronizing signal": a "burst" of subcarrier frequency located on the back porch of the horizontal blanking pulse.

4. Calibrator (Figure 5) - The amplitude of the phase shifter output F_{Lvar} is theoretically independent of its phase and the phasor representing F_{Lvar} describes a uniform rotation over the whole range. It has been mentioned above that due to inevitable inaccuracies in the construction of multiphase condensers and in the circuitry this is not realized too well.

To achieve very accurate measurements, the phase shifter must be calibrated. This is done in the following way: The input to the phase shifter F_{Lcon} and the output F_{Lvar} are each put through multiplying channels. The outputs of the two multiplier chains are respectively constant and variable phase voltages of n times the phase shifter frequency. (In the equipment described n = 10. The output of the multiplier chains is 1.575 mc.) Multiplication of frequency means also multiplication of phase angles.

The two waves are connected respectively to the horizontal and vertical deflection channels of an oscilloscope. We obtain straight lines for phase differences of $k\pi$ (k being any integer). These phase differences correspond to $k\pi/n$ phase differences of the phase shifter frequency. With n = 10, we have 20 check points per revolution. The check points are 18° apart: as the phase shifter deviates only slightly from the ideal linear behavior, linear interpolation does not introduce noteworthy errors.

Comments on the Operation of the Equipment

The above described method of phase measurements is very accurate. Instabilities in the system were minimized by keeping the Q's reasonably low and detuning the resonant circuits. The low frequency was generated by a crystal oscillator. The stability of the phase shifter and other circuit elements is satisfactory.

The accuracy of the measurements depends considerably on another factor. In spite of several filtering stages, some amount of S_{con} works its way through from C to D in Figure 3. Figure 6 illustrates the resulting phase error. By careful design, the ratio of S_{con} to S_{var} at the output has been reduced to approximately 0.3%. This corresponds to a maximum phase error of approximately 0.2°. (See Figure 6.)

Gain variation in the S_{var} channel must not influence its phase. The gain is varied by varying the grid bias of the first amplifier tube of the adder. This results in a variation of the dynamic input capacity of this tube which changes the phase of the preceding cathode follower output of the converter. By applying a capacitive divider at the input, the change in capacitance has been minimized.

A high-pass filter eliminates the low frequency components of the signal, so that different subcarrier frequency components appear on the scope with common base line. The positions of the various signal portions can be identified by using the switch S_2, so that the input is directly displayed. Figure 7 illustrates the pictures obtained on the screen of the oscilloscope.

The overall error of measurements with the described equipment is estimated to be of the order of .5°.

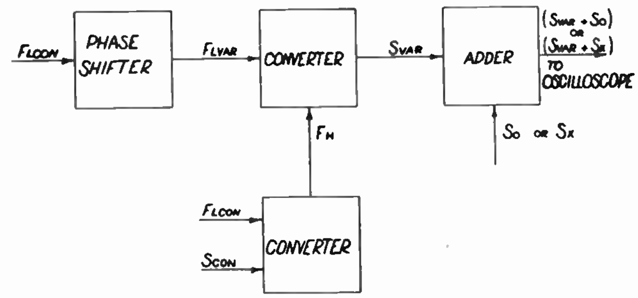


Fig. 1 - Block diagram of phase measuring equipment.

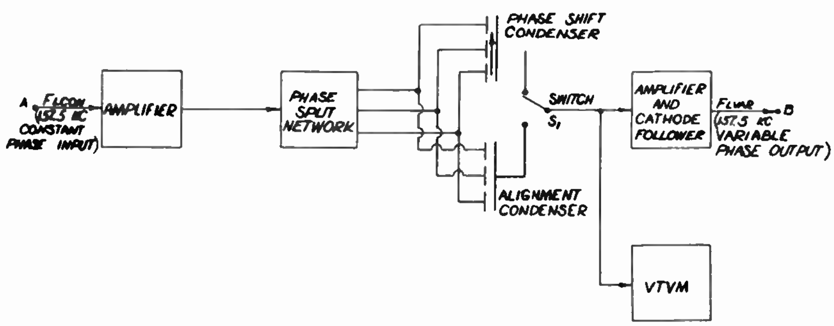


Fig. 2 - Block diagram of low-frequency phase shifter.

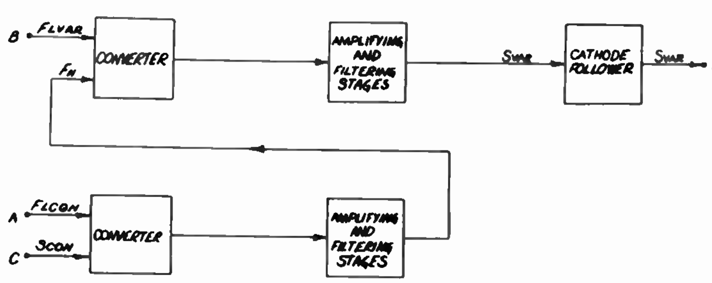


Fig. 3 - Block diagram of converter.

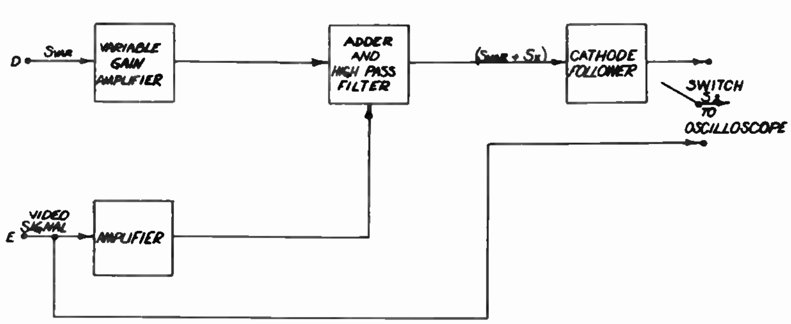


Fig. 4 - Block diagram of adder.

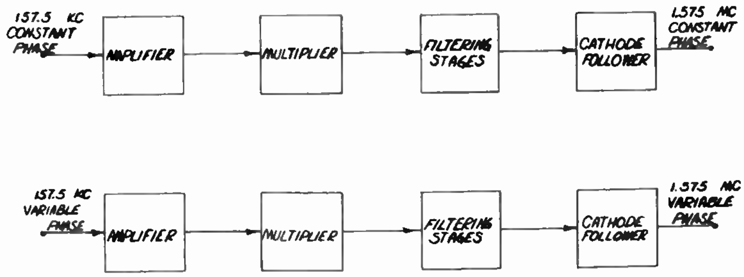


Fig. 5 - Block diagram of calibrator.

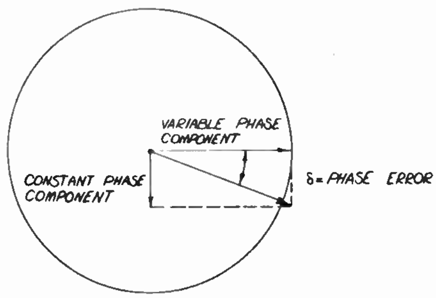


Fig. 6
Maximum phase error caused by constant phase subcarrier feedthrough from C to D in Fig. 3.

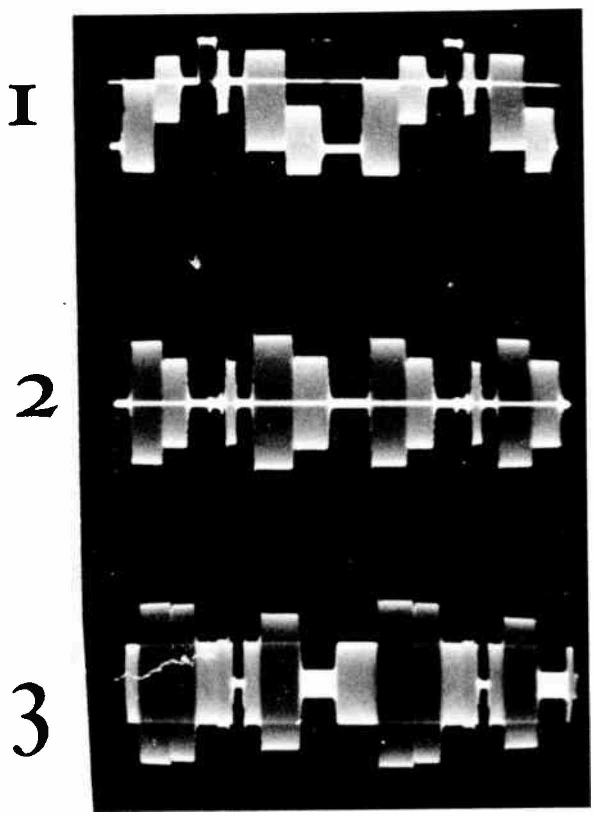


Fig. 7
(1) video signal; (2) video signal through high-pass filter (= chrominance signal); (3) picture after adding subcarrier to (2), in order to cancel the "burst".

A MONITORING SYSTEM FOR NTSC COLOR TELEVISION SIGNALS

Charles E. Page
Hazeltine Corporation
IRE National Convention
March 24, 1953

The color television signal presently specified by the National Television System Committee is composed of two major parts, a monochrome or luminance signal which carries essentially all of the brightness information, and a 3.579545 mc color subcarrier which is modulated by the chrominance or color difference signal information. The color subcarrier is modulated by the chrominance information in such a fashion that a change in the amplitude of the color subcarrier may correspond to a change in either the saturation or the brightness of a colored area or a combination of both. A change in the relative phase of the color subcarrier always corresponds to a change in the hue of the colored area.

Figure 1 shows a color bar pattern photographed from a dichroic monitor, and below it the conventional amplitude versus time display obtained from the cathode ray oscilloscope synchronized at line rate. Here we see at the left the horizontal synchronizing pulse followed by the burst and then by the luminance and chrominance or color subcarrier components of the various bars. We can see from an examination of this display that it is quite easy to measure the relative amplitudes of the color subcarrier components in the various bars and to compare these amplitudes to the amplitudes of their respective luminance components. However, it is difficult to get even a rough idea of the relative phase of the color subcarrier corresponding to the various bars from this sort of a display. It is therefore evident that this type of display can provide only a fraction of the information required for monitoring NTSC color television signals, and that additional information regarding the relative phase of the color subcarrier at various points in the picture is necessary. There are many methods available for the measurement of relative phase but many of these, while they are capable of giving results of a high order of accuracy, are rather slow and laborious and thus not well suited for rapid visual monitoring purposes. What would really be desirable is a vector display of amplitudes and phases of the color subcarrier components in different sections of the picture. One form of vector display has been described by Dr. Schlesinger and Mr. Nero of Motorola. The vector display to be described in this paper is a somewhat different arrangement.

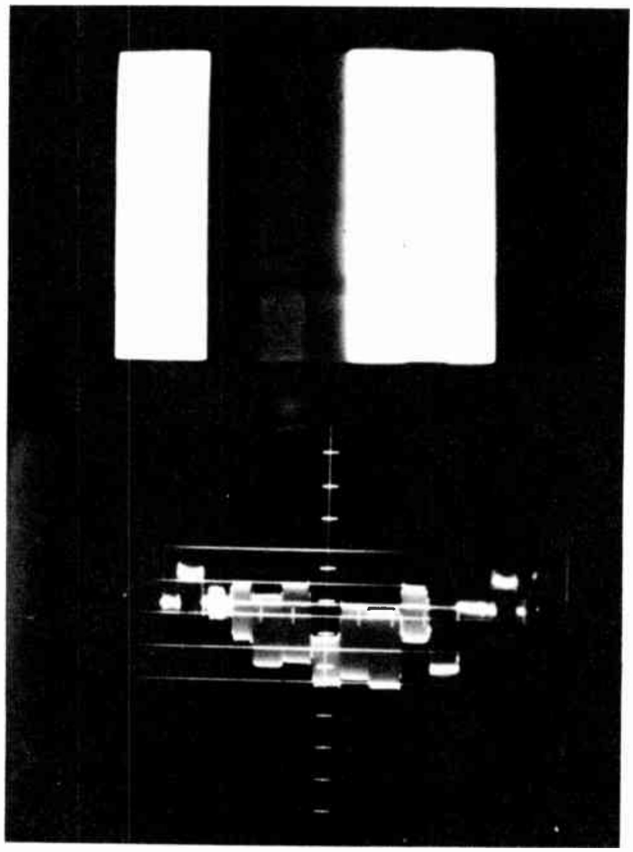


Figure 1

Figure 2 shows the specification of the color subcarrier component of the National Television System Committee signal. Examination of this specification shows that the color subcarrier is specified as the vector sum of two amplitude modulated carriers which are in time quadrature. This suggests that we might be able to obtain the desired vector display by recovering the amplitude modulation components of the two carriers and then applying them to the horizontal and vertical plates of an oscilloscope.

$$E_{SC} = \left[E_Q' \quad \sin(\omega t + 33°) + E_I' \cos(\omega t + 33°) \right]$$

FOR COLOR DIFFERENCE VIDEO FREQUENCIES BELOW 500 Kc, THE SUBCARRIER SIGNAL MAY ALSO BE REPRESENTED BY

$$E_{SC} = \frac{1}{1.14} \left[\frac{1}{1.78} (E_B' - E_Y') \sin \omega t + (E_R' - E_Y') \cos \omega t \right]$$

Figure 2

Figure 3 shows an arrangement which we tried for accomplishing this result. The color video signal first passes through a band-pass amplifier which removes the low frequency monochrome components and permits full double sideband transmission of the color subcarrier components which then pass to the pair of synchronous demodulators which operate in time quadrature. The outputs of the demodulators then go through low pass filters which remove residual color subcarrier components to the vertical and horizontal plates of the cathode ray oscilloscope.

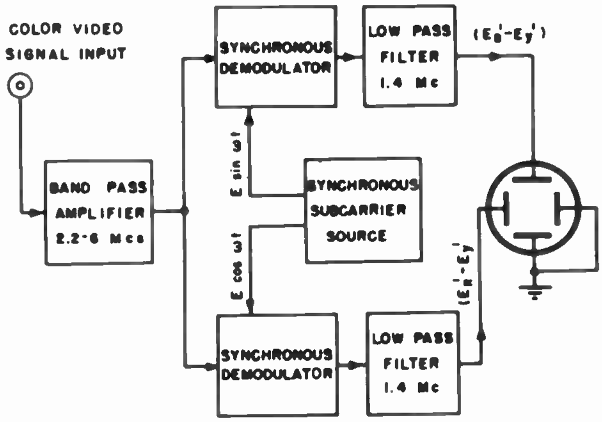


Figure 3

Figure 4 shows a colored bar pattern and below it the phasor display which we obtained with the arrangement which we have shown in the previous slide. If we start at the center of the phasor display we can see that the color subcarrier phase went first to blue at the top of the phasor diagram, then counterclockwise to magenta, then to red, then to yellow, then to green, then to cyan, back to blue, and then to white which carries it to the center for the display. We also see the burst pointing directly downward in the -(B-Y) direction.

Well, this sort of a display seemed to be the sort of a thing we were looking for, and we felt that if we had some assurance that the two demodulators were really in quadrature and that the horizontal and vertical gain settings of the oscilloscope were correct, that we might then have some confidence in the thing as giving us a proper representation of the amplitude and phase of the various color subcarrier components.

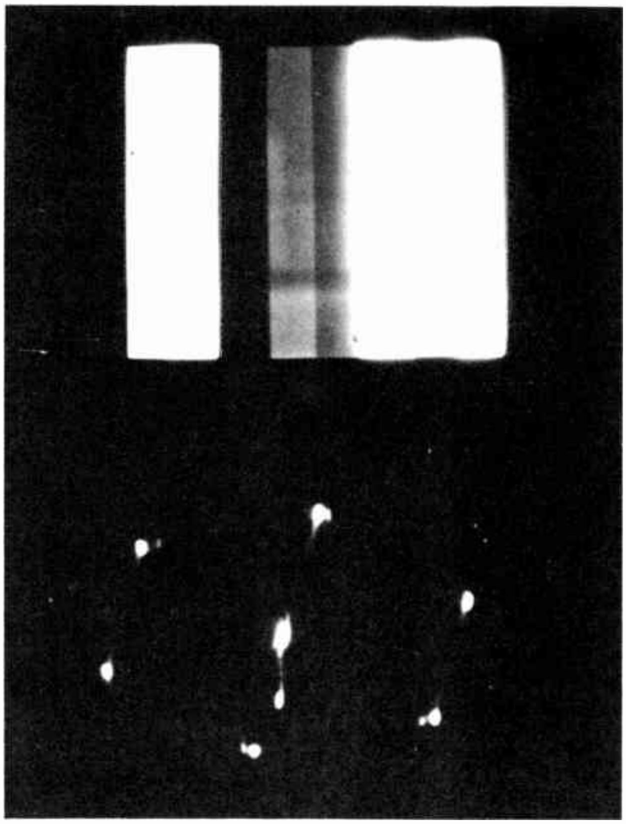


Figure 4

One common method for checking the quadrature relationship of a pair of sinusoidal signals is to connect them to the horizontal and vertical plates of an oscilloscope, adjust the gain settings for equal horizontal and vertical deflection, and to then check that the resultant pattern is a circle. By injecting a test signal whose frequency is 10 Kc different from that of the color subcarrier frequency into the demodulators via the bandpass amplifier, we obtain from the outputs of the demodulators a pair of 10 Kc sinusoidal beatnotes which will be in quadrature if the demodulators are in quadrature. If we then connect the outputs of the demodulators to the horizontal and vertical plates of the oscilloscope and properly adjust the gains, the resultant pattern will be a circle. If, we in addition periodically reverse the phase of the signal input to one of the demodulators by 180° at a low frequency audio rate such as 30 cycles per second, the oscilloscope display will be a circle where the demodulators are in quadrature but will degenerate into a pair of crossed ellipses whenever the quadrature relationship is not correct.

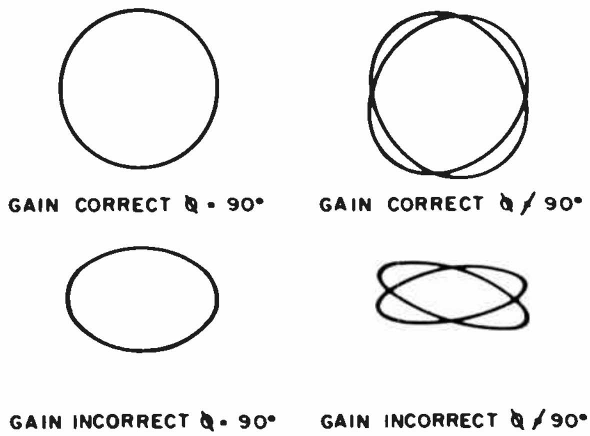


Figure 5

Figure 5 shows the appearance of this test pattern for various conditions of adjustment when we are using the 180° phase reversing scheme. In the upper left hand corner we have the case where the quadrature relationship is correct, the oscilloscope gain has been properly adjusted so the resultant pattern is a nice single circle. In the upper right hand corner we have the case where the oscilloscope gains have been adjusted properly, but the quadrature relationship is not correct, and our test pattern has degenerated into a pair of crossed ellipses. In the lower left hand corner we have the case where the 90° phase relationship is correct so that we have a single pattern, but the oscilloscope gain settings are not correct so that the pattern is an ellipse with a horizontal major axis. Here in the lower right hand corner we have the case where everything is misadjusted and the pattern has degenerated into a pair of rather skinny crossed ellipses. This then is a nice simple and accurate check for correct quadrature relationship independent of oscilloscope nonlinearities and gain settings and in addition allows us to get the oscilloscope gain settings correct. However, this test is an accurate test only if we assure that the 180° phase reversing amplifier is reversing by exactly 180°. Fortunately it is easy to check the 180° reversing requirement.

If we feed both of the demodulators with the reference subcarrier signal, a demodulating subcarrier signal in phase, we obtain the pattern shown in Figure 6. In the upper left hand corner we have the case where the relative phase shift between the two channels is zero producing the straight line which goes from lower left to upper right and the 180° phase reversing amplifier is reversing by exactly 180° producing the straight line going from lower right to upper left. In the upper right hand corner of the slide we have the case where the relative phase shift between the two channels is still correct and producing a straight line but the phase reversing amplifier is not reversing 180° thus producing an ellipse instead of a straight line in the opposite direction. In the lower left hand corner of the slide we have the case where the relative phase shift between the two channels is not correct producing an ellipse going from lower left to upper right, but we know that the phase reversing switch is reversing by 180° because the ellipse which goes from lower right to upper left is identical with the other one. Over here in the lower right hand corner we have the case again where everything is misadjusted. The relative phase shift between two channels is not zero and the phase reversing switch is not reversing by 180°. Thus we can see from these figures that it is quite easy to provide a self-checking arrangement for this equipment.

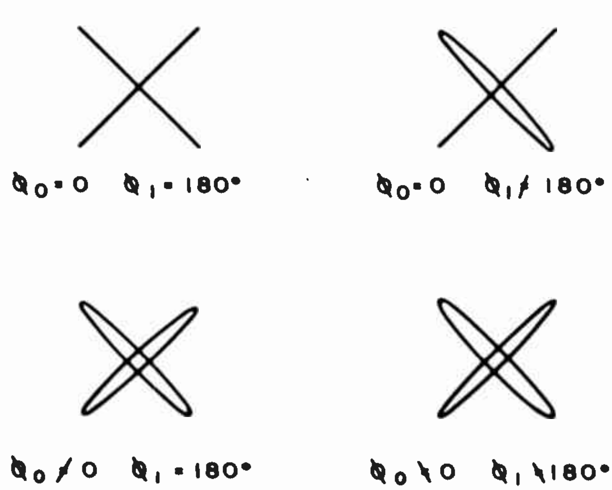


Figure 6

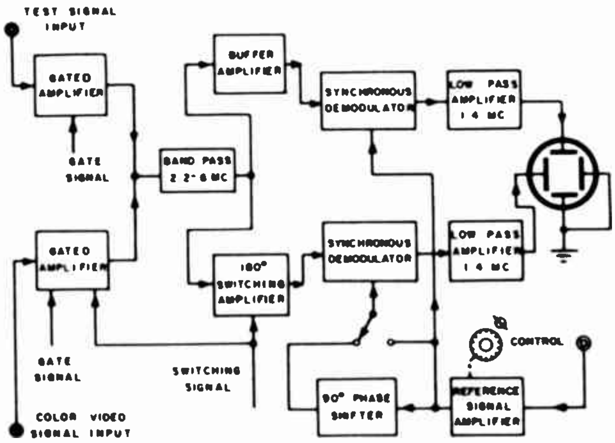


Figure 7

Figure 7 shows the final arrangement of the equipment which incorporates the self-checking features which we just described. The color video signal passes through an amplifier stage to the band-pass filter, thence in one channel to the 180° switching amplifier, in the opposite channel to a buffer amplifier which is electrically similar, then to the synchronous demodulators which may be operated either in phase or in quadrature, and then through the low-pass amplifiers to the cathode ray oscilloscope. The test signal passes through a second amplifier to the band-pass filter and then through the same channels which the video signal followed. In order to provide a continuous presentation on the cathode ray oscilloscope the signals are gated as follows. A gating signal is applied to the video signal amplifier which turns it on during 50% of the time. A second gating signal is applied to test signal amplifier which turns it on during 50% of the video signal amplifier off time. In addition the switching signal is applied to the video signal amplifier so that it is turned off whenever the switching amplifier is in the 180° position. The oscilloscope then presents on a time-shared basis the output due to the video signal input which is the desired phasor diagram which was previously shown, the output due to the test signal which will be a single circle when the adjustments are proper, and the output for no signal input which will be a dot at the center of the display. In addition we thought that it would be nice to have some 90° electrical phase markers incorporated in the system for checking possible quadrature errors in the oscilloscope itself. To accomplish this we take the 10 Kc beatnote between the reference signal and the test signal, multiply it by 4 and pass it to a gated pulse-forming amplifier which then produces pulses at a 40 Kc rate during the time when the test signal amplifier is also gated on. These pulses are used to produce 90° blanking markers on the oscilloscope display.

Figure 8 shows the complete display which we obtain from the arrangement shown in the block diagram of the previous slide when the input signal is the signal from an NTSC encoder modulated by the color bar pattern which was shown in the previous figures. We see the calibrating circular pattern, which at the moment is a single circle since the quadrature relationship is properly adjusted, and the phasor diagram of the signal which is similar to the one which we saw previously. You will note that on the cursor about the dot corresponding to the tip of each vector there is a square which indicates the tolerances set by the NTSC for the amplitude and phase of the various color subcarrier components. In addition you will note that the center of the phasor diagram coincides with the center dot of the display indicating that no spurious subcarrier signals are being radiated during white or black.

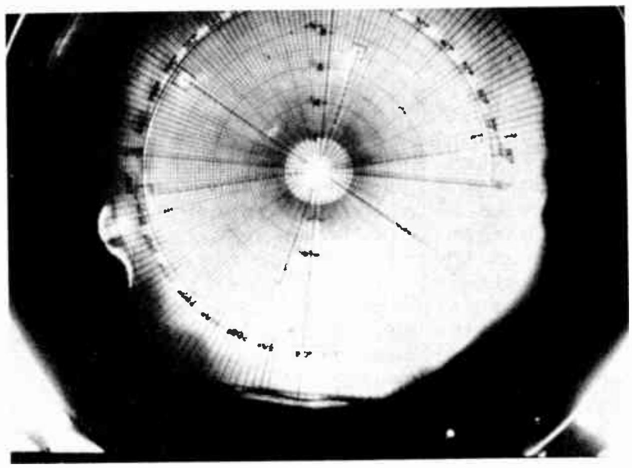


Figure 8

Figure 9 shows the phasor diagram produced by the complete equipment when things are not so good. The calibrating or test circle has degenerated into a pair of crossed ellipses indicating that the quadrature relationship between the two demodulators is not correct and the dot at the center of the phasor diagram does not coincide with the dot at the center of the display indicating that a spurious subcarrier signal is being radiated during black in this case in the -(R-Y) direction.

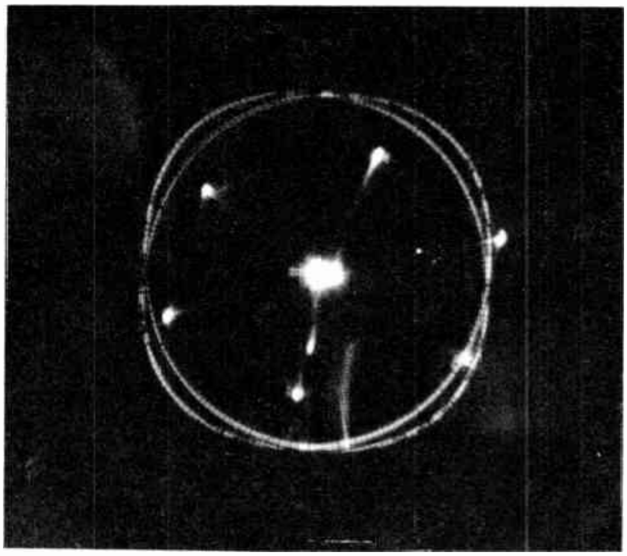


Figure 9

Since this equipment requires a number of auxiliary signals and since it was felt that the equipment might often be used for monitoring purposes at locations remote from the signal generating source, it was thought desirable that the equipment provide its own signals from the color video signal input and here shown in figure 10 we have the arrangement for the final form of the equipment. The synchronous amplifier, the synchronous demodulators etc. are the ones that were shown in the previous slide. In the blocks below we have a sync separator which strips sync from the composite video signal input, uses it to operate the switching signal generator for the 180° phase reversal and generates a gating

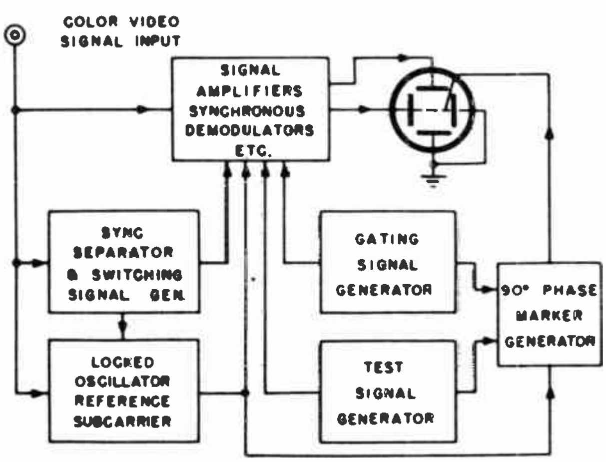


Figure 10

pulse for the locked oscillator reference subcarrier generator which is synchronized to be burst. Then we have the gating generator which provides the necessary gating signals, the test signal generator, and finally the 90° phase marker generator. In addition to its use for signal monitoring purposes this sort of an arrangement is of some interest in examining the chrominance content of various colored subjects.

Figure 11 shows a colored slide which was photographed again from the dichroic monitor and below it the corresponding display produced by this equipment. The things which are of prominence as far as the color subcarrier vector or chromaticity are concerned are the orange mail box which is producing the vector which extends from the center toward the lower left hand corner of the slide and the next most saturated thing in the picture, the red mail box which is producing a vector which goes nearly horizontally from the center toward the left hand side of the slide.

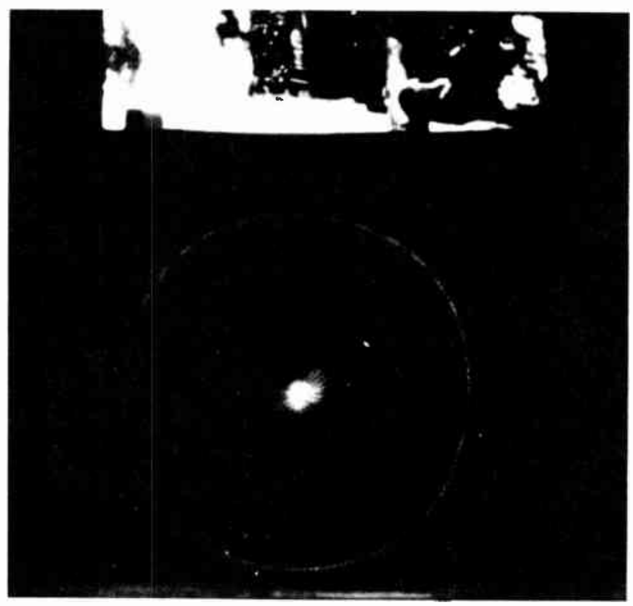


Figure 11

In closing I would like to express my appreciation to Mr. B.D. Loughlin who conceived the idea and arrangement of this equipment, to Mr. R.P. Burr who provided the colored slides, and to Mr. Charles J. Hirsch without whose counsel and encouragement this paper would not have been presented. Thank you.

THE DESIGN OF AUDIO CONSOLES FOR TELEVISION

by
Robert H. Tanner, B.Sc. (Eng.)
Audio Engineer
Northern Electric Company Limited
Belleville, Ontario, Canada

During the past year there has appeared a number of articles and papers relating to the facilities required for handling the audio component of a complete television programme, and some of these have dealt in considerable detail with the design of specific T.V. audio consoles. In presenting yet another paper in this series, it should be explained that the equipment to be described has been designed entirely independently and with an eye to the peculiar needs of Canada's Radio and Television set-up, which as most of you will know differs considerably from that of the United States. Like so many other things in Canada, radio reflects features of both the British and the American method of operation. It combines a public corporation entrusted with the task of bringing radio to as high a proportion as possible of our scattered population with a system of private radio stations comparable with those familiar to you on this side of the border. Television broadcasting recently launched in Canada is at the moment being channelled exclusively through C.B.C. outlets but there are indications that in the near future privately owned transmitters will be authorized. Needless to say, private stations are considerably more concerned with the economic aspects of broadcasting than is a public corporation, and this fact is reflected in the design of equipment. The first console I am going to describe to-night was designed by us to meet a reasonably detailed functional specification drawn up by the Canadian Broadcasting Corporation, for use in their television studios in Montreal and Toronto. The second is our own idea of a console flexible and versatile enough to meet all the needs of the private broadcaster, whether for AM, FM, or TV.

Before describing the CBC Television Console, I would like to take this opportunity of expressing our thanks to Messrs W. A. Nichols and W. C. Little of the CBC Plant Department for their cooperation and assistance in its design.

The Console is illustrated in figure 1, which shows that the front panel is divided into three sections. The left hand section comprises a small jack-field to which are connected all the programme inputs, which by means of patch cords or normal connections may be extended as required to the console's ten preamplifier inputs. For flexibility, all incoming programme material is brought to microphone level so that all inputs may be treated alike. Jacks are also provided on the inputs and outputs of the 3 programme amplifiers to allow rapid changeover in the event of an emergency. It has frequently been stated that more input facilities are required for television work than for normal sound broadcasting; this console makes provision for 24 inputs, apart from 4 remote lines, and 10 pre-amplifier mixing channels. These facilities it is thought will be sufficient for even the most complicated of shows.

The centre section is occupied by the control panel proper with the 10 input faders mounted along the bottom in groups of 3, 4, and 3, respectively. The two right-hand channels are equipped with relays for announcer control. Each group is connected via its own mixing bus to a sub-master amplifier and attenuator, a feature which is of very great use, as for example, in musical programmes in which the balanced output from several microphones picking up an orchestra, may in turn be balanced with the output from the soloist's microphone. The 3 sub-master outputs are again mixed and fed to the input of the main programme amplifier, the control for which will be seen immediately to the right of the volume indicator. This net-work comprises the main programme chain of the console. However, a feature which is believed to be unique is the provision of two auxiliary programme amplifiers, which, by means of selector switches (mounted on the right-hand panel) may be connected across the output of any pre-amplifier, any sub-master amplifier, or the main programme output of the console. This connection is made by means of what is virtually a hybrid circuit, so that, if for example, one of the auxiliary amplifiers is switched to pre-amp number 2 there is negligible cross-talk from pre-amplifiers one and three, even though the outputs of all three pre-amps are connected to the same mixing bus. It is thought that this feature will be extremely useful in many ways, two of which may be outlined here. It is generally agreed that television operation frequently requires that a programme originating outside the studio be fed to a studio loudspeaker so that it may be heard by those taking part in the production. Typical examples of such programme are records, sound tracks of motion pictures, or programmes incoming on a remote line. With the arrangements incorporated in this console, it is only necessary to switch one of the auxiliary amplifiers to a pre-amplifier or sub-master handling this programme material and connect the output of the auxiliary amplifier to a monitor amplifier feeding a studio loudspeaker. The signal may simultaneously be fed to the main programme output, or not, as desired, but the isolation provided in the console is sufficient to ensure that no howl-back can result. Another interesting use of this feature arises on occasions when a programme is being transmitted both on sound and T. V. simultaneously, as is now the case, especially with some musical shows. It is well known that the sound accompanying the T. V. picture of, let us say an orchestra should for maximum realism vary in accordance with the picture; for example, if the cameras are focused on the French horns then the sound balance should emphasize these instruments. On the other hand, the listener with no picture to guide him desires the optimum musical balance at all times. With this console the correct balance of the orchestra may be established on one group of microphones, and fed via a sub-master to an auxiliary amplifier and thence to the sound net-work line. A moving microphone mounted on a boom and connected to another sub-master group, can be added to the main programme output, which of course will be fed to the T. V. sound transmitter. It is our belief that experience will evolve many other ways in which these two auxiliary amplifiers, together with their selection arrangements will prove of great use to the operator.

The console is equipped with a variable sound effects filter connected to the output of pre-amplifier No. 1. As is common for such devices two controls are provided, one affecting the treble end of the spectrum, the other the bass. In each case five different frequency responses are available. In a projected re-design of this equipment, it is proposed to make this sound effects filter work at microphone input level and impedence, thus allowing it to be patched in as required on any pre-amplifier channel.

As already mentioned the right hand panel of the console mounts the two selector switches associated with the two auxiliary amplifiers. In addition, controls are provided for the operator's head set and monitor loudspeaker, allowing him to select the programme material fed to each of these devices, and control their volume.

With the exception of the circuits used to provide isolation on the auxiliary amplifier feeds, the circuitry of the console is entirely conventional. Figure 2, which shows a rear view, illustrates the mounting of the unitized amplifiers as well as such units as repeating coils, filter coils, etc. Only two types of amplifiers are used, thirteen of one type for pre-amplifier and sub-master service, and three of the other for the main programme and auxiliary amplifiers. In the case of the two latter a change in one feed-back resistor gives the additional gain required in this position. The associated monitor amplifiers and power supplies are mounted elsewhere in order to avoid any heavy AC fields within the console itself. Perhaps as a result of this and of the design of the pre-amplifiers, the console exhibits a very good performance, especially as regards noise, with a signal to noise ratio under normal working conditions of from 75 to 80 db, with AC on all the amplifier filaments. If these latter are fed from a DC source a signal-to-noise ratio of 80 db can be obtained with the minimum attention to system grounding, etc.

Figure 3 illustrates the type of amplifier used in this equipment. Each of these little units is approximately 2-1/4" wide by 8" long, and is equipped with an input transformer of the plug-in variety, so that the input impedence may be changed as required, to suit the conditions under which the amplifier is working. Contrary to the practice of several other manufacturers the amplifiers themselves do not plug in, but are equipped with soldered terminals. It is felt that with amplifiers as reliable as they are to-day, the disadvantages of the plug in feature outweigh its advantages. A view of the front of the console with the panel opened is given in Figure 4, which illustrates the excellent accessibility of the design.

We may turn now to the second console, illustrated in Figure 5 which will be seen to bear many mechanical resemblances to the one already described. In its basic form, however, it is of course considerably smaller, although the design possesses the unique feature that two similar units may be joined together in such a way that they are integrated both mechanically and electrically to form the large two channel console illustrated in Figure 6. With this in mind let us return to the consideration of the small basic console, which is equipped with 5 pre-amplifiers, each with a two-way input selector (the keys to the left of the colume indicator) thus allowing as many as ten low level inputs to be connected permanently to the console. In addition, the two keys on the extreme left allow any one of four remote lines to be connected to the left hand fader. The outputs of the six faders may be connected as required to either of two mixing busses by means of the key immediately over each control (with this key in the central position the corresponding fader is connected to a third auxiliary bus which may be used for cueing purposes). In the small basic console, only the channel 1 bus is provided for further amplication in the form of a booster amplifier, a master gain control (on the right of the bottom row) and a line amplifier, although if desired, a similar amplifying chain for channel 2 can be provided externally. Even without this, however, the embryonic second channel has a number of important uses. For example, the monitor amplifier which is supplied with this equipment possesses sufficient gain to work direct from the mixing bus level, so that channel 2 may be used for auditions, cueing, etc. Again by connecting the second channel to an additional amplifier feeding a loudspeaker in the studio, and at the same time connecting it through a small pad to one of the low level inputs "fold back" operation, as already described in connection with the first console, may readily be achieved. The partial block schematic of Figure 7 will help to clarify this point.

Very complete facilities are provided on this console for carrying out remote broadcasts with only one line between the studio and the remote point. At the same time the control of these facilities has been kept extremely simple, with the addition of only one key (immediately to the right of the volume indicator). In the central position, remote cue, that is to say, the output of the monitor amplifier, is fed to all four remote lines, regardless of the position of the line selector keys, unless of course one line is actually transmitting the programme. With the key in the upper position the direction of transmission is reversed, and all four lines are connected to the input of the monitor amplifier giving a "multiple over-ride" condition. Placing the key in the lower position gives "selected over-ride" from whichever line has been selected by the line keys. Under these circumstances, the programme input to the monitor may be cut by turning down the monitor gain control on the front panel of the console, thus enabling a test to be taken from the line, without interference from any other source. In addition to all this, operation of the talk-back key adjacent to the "over-ride remote cue" key enables the operator to talk back down the line to the remote operator. Should one or more of the remote line inputs be normally associated with net-work, it is readily possible to disconnect any or all of these facilities from the input or inputs concerned. Although this arrangement is somewhat tricky to describe in a few words, it is thought that the functional layout of the controls is so simple that an operator will have no difficulty in mastering it very readily.

An additional facility for which provision is made although the extra equipment is only installed as required, provides announcer control of one pre-amplifier channel. Since the changes involved are very simple, it is possible for a station engineer to add this feature at any time.

The console uses the same amplifiers already described, and their mounting is shown in Figure 8. In this case to deal with situations where the console may be mounted with its back close to a wall or window, the frame work mounting the amplifiers is pivoted and may be swung up to allow maximum accessibility to all the components. Figure 9, which shows a front view of the console with the panel opened, again demonstrates the accessibility of the panel components and of the telephone type terminal strips, to which all external connections are made.

Returning now to the double version of this unit (Fig. 6) it will be obvious that this makes provision for as many as 20 low level inputs and 8 remote lines, any or all of which may be switched to either channel; in this case the second booster and line amplifier combination is connected to channel 2, thus giving full double channel working. Two monitor amplifiers are also provided, one of which is used exclusively for the control room, while the other feeds speakers in the studio or studios. Separate keys for talking back to lines on the one hand and studios on the other are provided, and while the studio operator enjoys all the same facilities for remote cue and over-ride as are given by the single console, the additional facility is provided of feeding selected over-ride to the studio speakers. By connecting the channel 1 and channel 2 mixing busses through pads to two of the low level inputs, "fold-back" operation may very readily be set-up, without the use of any additional equipment, and by similar means it is possible, as in the case of the CBC console, to feed the two outputs with differently balanced versions of the same programme. In the double console provision is made for adding as required announcer control to either 1 or 2 pre-amplifier channels.

While it is electrically and mechanically possible to couple three of these units together to form an integrated whole, no situation has as yet been encountered sufficiently complicated to warrant such a set-up.

Once again monitor amplifiers and power supplies are mounted external to the console. The monitor amplifier is capable of delivering an output of 15 watts, with less than 2% distortion, and possesses sufficient gain to work direct from microphone level. The regulated power supply is rated to give a maximum of 100 milli-amps at 300 volts, with a no-load to full-load regulation of $\pm$ 1/5 volt, at any input voltage between 105 and 125 volts. Although the use of a regulated supply of this kind is by no means essential, it has been found to be of considerable assistance in preventing very low frequency motorboating when several amplifiers with good low frequency responses are connected to the same power-supply.

It is hoped that this short paper will be sufficient to give at least an outline of the trend of thinking in Canada regarding the design of audio equipment for an industry which may well symbolize a new and exciting phase in the growth of our Country.

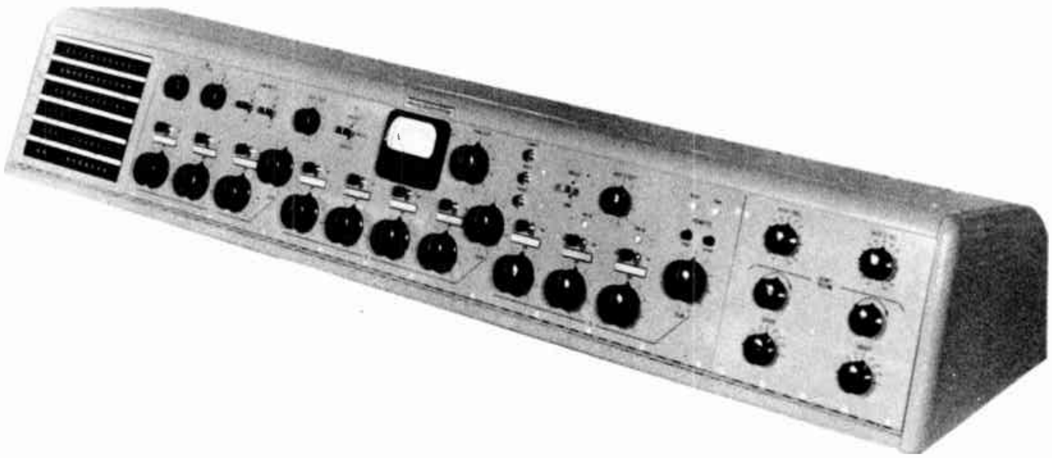


Fig. 1

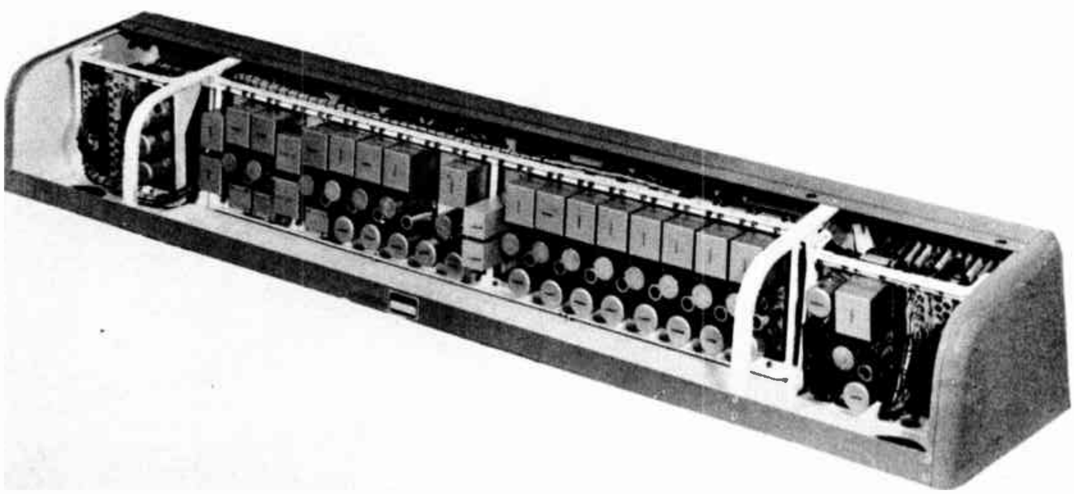


Fig. 2

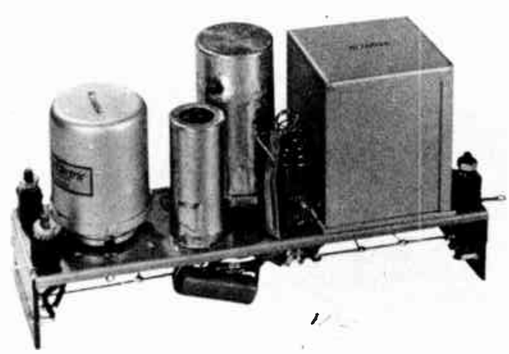


Fig. 3

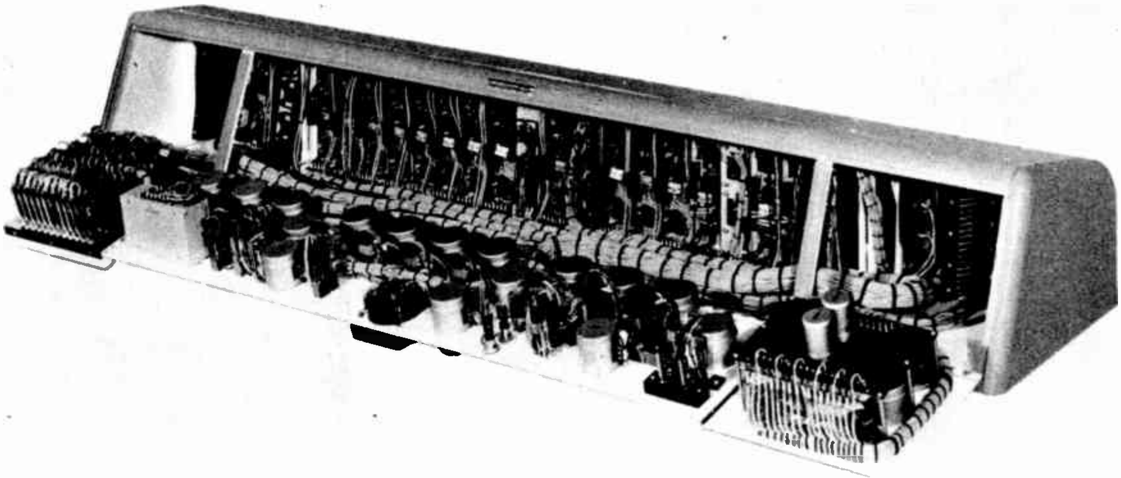
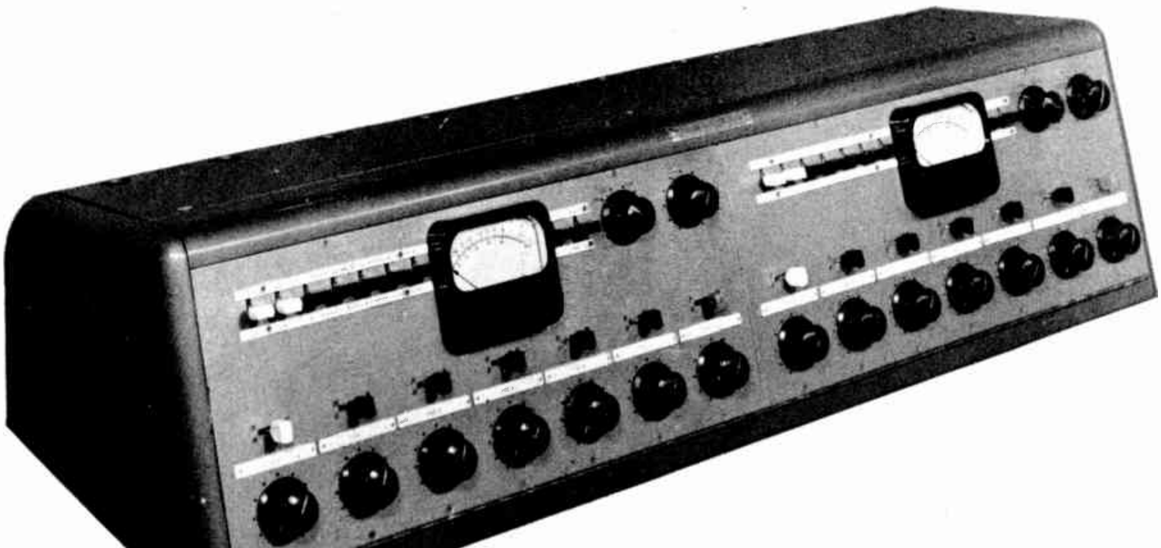


Fig. 4

Fig. 5

Fig. 6



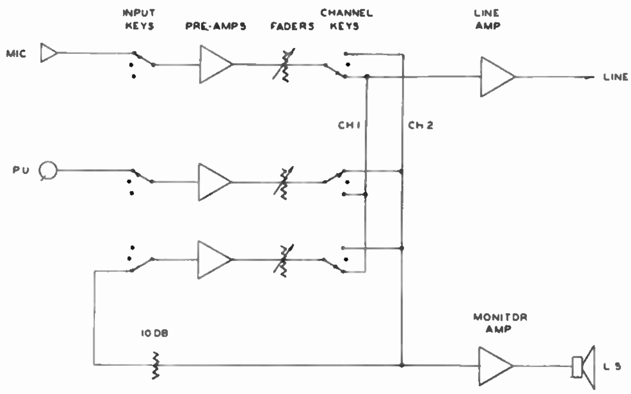


Fig. 7

Fig. 8

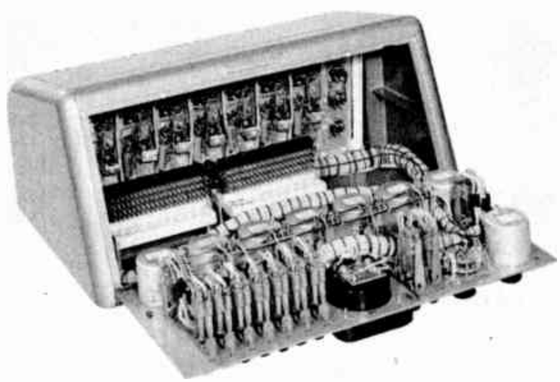


Fig. 9

BUILDING TV BROADCAST FACILITIES
FOR
GROWTH, FLEXIBILITY AND ECONOMY

Allen R. Kramer and Edwin R. Kramer
Kramer, Winner and Kramer
Architects and Broadcast Engineers
33 W. 42nd St.
New York 36, N.Y.

SUMMARY- An analysis of a practical, master plan for the organized development of a typical radio station, located in a medium-sized community, which has entered lately the field of TV and wishes to combine both facilities in the same building.

ASIDE from certain legal and financial considerations, the first step in the design of a building should be the selection of a site. Site selection is noted as a design step because the site determines as much about the form of the building as any other single factor. The size of the site may determine whether a single or multi-story scheme may be used. Conformation of the land, the presence of rock or water, availability of utilities, relation to adjacent buildings, and sources of noise are a few of the physical factors to be considered. No site should be purchased without the help of the architect who is going to design the building. It is also important to consider the following: accessibility, advertising and public relations value, and tax rates.

Site requirements are much more complex than one is inclined to accept, before an analysis has been made. If it is financially possible, it is wise to buy at once a piece large enough for all future requirements. The attempt to buy an adjacent piece of land at a later date may be unsuccessful or may cost two or three times the original land cost. It is generally believed by laymen that TV is a magic world where limitless funds are available and any price will be paid. It is wise to buy once, buy enough, and buy anonymously.

The subject of antenna and transmitter site location in relation to the broadcast area and to the station itself is a distinctly separate problem which is so dependent upon local conditions that it will not be considered at this time.

There is one word which has become rather commonplace in a discussion about the future of TV. That word is <u>expansion</u>. Everyone agrees that expansion <u>will be</u> necessary, but there is a concept which is mentioned less frequently and which is even more important. That concept is <u>flexibility</u>, and it implies not only the <u>possibility</u> of a variety of expansions, but the rearrangement of existing facilities to form totally new plans. This brings us to the question of predictability of the future of TV. We cannot be certain about such factors as the proportion of film to live shows that will be required to arouse regional interest and satisfy regional pride.

Our views are based on the assumption that live shows will be required to provide local interest in a station and also that some public service programs must be live; that some form of studio audience will always be required for certain types of entertainment; that radio will depend to an increasingly greater extent upon recorded material.

Because one's entire thinking in the planning of a TV station should be oriented toward the ultimate scheme, it seems appropriate that we should start with a description of the final stage of growth and work back to the earlier stages.

In Fig. 1 we have an overall view of the completed radio-TV building.

A ground level plan for this final stage is illustrated in Fig. 2. The arrows show the points of entry of the various types of people who use the building. Control of this complicated traffic is essential and to a large extent this has determined the form of the building.

All facilities for the production of TV shows are located on one level and no other traffic is permitted to cross this area. Sets and heavy equipment are brought into the storage and workshop area from the left and eventually are moved onto the studio floor from that direction.

Performers enter the building from a rear parking lot. They go directly to the dressing rooms and feed into the studio area from the right. Costume storage is located with the dressing rooms. The layout also reveals that nowhere can performers and control room technicians interfere with each other. All the control rooms are located above.

Administrative and clerical personnel are restricted to the one-story wing on the far right, and the few who are authorized to enter the production or control sections can do so at the intersection of the wings.

The public and business callers are admitted to the lobby which provides access to the studio audience auditorium and to the general business offices. Control room personnel would enter through the lobby and go directly to the upper level by way of the stair at the hub of the plan.

In every stage of expansion the mobile units are housed in the rear of the building. They adjoin one of the TV studios and might conceivably be used to control a broadcast from the studio.

The upper level of the fully-expanded scheme is shown in Fig. 3. This second level occurs only over the central core of the building. Two TV control rooms with announcers' booths have direct visual control over the large studios. Immediately adjoining is the TV master control and shop. Directly across the hall is the film bloc. It consists of a fireproof film storage room, preview room, processing and editing group and the actual film projection room. Control for the multiplexers occurs in the master control room.

Control of the two smaller TV studios is handled blind. It is becoming increasingly obvious that direct visual control of a program is unnecessary except in the case of quite complex assignments. At a station visited recently, we found a very large studio with an adjacent control room located on the same floor level and equipped with a large window over the control consoles. The control personnel had long since found that from their position it was usually impossible to see large areas of the studio because of intervening sets and equipment. They had further found that these obstructions did not handicap them and had eventually permitted a permanent set to be erected directly in front of the window. Control windows are shown only in the largest studios and in each case, well above the top of the highest sets.

At one end of the TV control area are located the art room, offices for programming personnel, engineers, and a lounge for off-duty hours.

The far end of the upper level is devoted to radio broadcasting facilities. The master control room overlooks one of the large TV studios for those occasions when a program is broadcast by both TV and radio sections. This studio is also available for certain radio broadcasts such as a large choral group or an orchestra. In addition, there are three other broadcast studios including one small one for interviews and news broadcasts. Radio performers would enter the building at the performers' entrance in the rear and arrive at the studio doors via the stair without having passed through the building.

Across the hall from the control room is the recording studio which immediately adjoins the record library. The news room which will dispense material to both radio and TV is located directly between those departments.

The cross-section drawing, Fig. 4, shows graphically the vertical relationships among the various sections of the building. All production traffic is on one level and the control rooms overlook the large studios and operate undisturbed on a separate floor.

One of the advantages inherent in this plan is that the distinct separation of functions into wings and levels permits the building to be shut down except for a few essential rooms during the night hours. At the same time, the elements are brought together in such a manner that the areas requiring most wiring and those which produce the greatest heat load are concentrated for minimum cable length and economical duct work. Since the building is elaborately zoned, only those areas which call for heating, cooling, or ventilation, are supplied with such facilities. Heat drawn from lights and equipment is used in cold weather and exhausted in summer.

While mechanical air conditioning will be a necessity in most climates, it is possible to reduce materially the cooling load on this building by the employment of a few simple devices. Accordingly, this building features large overhangs over glass areas, heat absorbent glass, light-colored wall surfaces, and a roof covered with brilliantly reflective white marble chips. This last device alone is capable of reducing roof temperatures by 30 degrees. Additional cooling can be obtained by the use of water spray nozzles on the roof. The water used is that which is recirculated by the air conditioners. Shade-producing planting aids in the reduction of outside noise levels as well as in cooling the building.

Let us now trace the stages of growth by which the building may evolve. Depending upon the local demand, the studio audience may or may not be included, and the plant operates perfectly without it. At a slightly earlier stage, the building could exist without the administrative wing. This is shown in Fig. 5. During another period, only the two large studios would exist without the set storage area; one large studio would be active and the other would serve as storage area for all other studios. The administrative area is also capable of being expanded or reduced or even eliminated. It can, therefore, be seen that there are very many possible combinations during the period of growth. Given a plan which properly begun, it is not even necessary to foretell the future with complete accuracy, either in regard to the type of facilities which may be needed or the sequence in which the need may arise.

This brings us to the first stage of construction, which is shown in Fig. 6. As far as function is concerned, it has been assumed that we are dealing with a station that operates almost exclusively on network, film, and remote broadcasts. It would attempt only simple live broadcasts. An inspection of the plan shown in Fig. 7 indicates that the principles of traffic separation are already in operation. Performers enter from the rear parking area and use the dressing rooms and TV studio or go directly up to the radio studios. The large room opposite the TV studio, which becomes a dressing room in a later stage, is here shown in use as the prop storage area. Mobile units are housed adjacent to the TV studio. The upper level, shown in Fig. 8, can be recognized as the core of the expanded stage described previously. The film bloc is almost identical. The TV control room is more modest in size but already positioned where it can overlook the large studio which is to come. Art work and programming offices are in the same relative position they will hold later, but reduced in size.

The radio facilities hold the closest resemblance to the final scheme, the only difference being that here there is a large studio in place of the small news broadcast studio seen previously. It will be noted that the radio control room gains the use of the very large TV studio when this change is made. The studio which separates the radio and TV control rooms could conceivably be used for simultaneous broadcasts of the simple type such as panel discussions, a lecture or a political speech.

Radio engineers are accustomed to precision, but in the main precision is not a characteristic of the building industry. Compared with almost any other modern enterprise, building is archaic. Buildings are usually one of a kind and they are painstakingly erected out of large numbers of small but heavy units such as brick and plaster which are bonded together so inseparably that they cannot be moved without being destroyed. Since most building materials are at least in part fabricated on the site, the general construction area becomes for a very long time a noisy and uninhabitable place. This is always a serious disadvantage, but it is particularly difficult when an alteration or addition to an existing building is involved. In the case of a TV station, attempting to maintain scheduled broadcasts, the complete disruption of facilities would be intolerable.

Expansion and flexibility are weasel words unless they are accompanied by a construction method as advanced as the building's planning.

If we were to try to designate the characteristics of an ideal construction material and method, we would ask for: (1) Strength; (2) lightness; (3) moderate first cost; (4) economy of future maintenance; (5) attractive appearance; (6) speed, silence, and cleanliness during first installation and later rearrangement; (7) good sound absorption characteristics; (8) good sound insulation characteristics; (9) possibility of carrying wiring and cables within the thickness of walls and floors; (10) 100% salvage when walls are to be removed; (11) compliance with most building codes; and (12) availability in stock form anywhere in the country.

A material which satisfies alomost all the requirements we have mentioned already exists. It is a building panel manufactured of steel or aluminum by about a dozen companies, and has been in use for several years, mainly in industrial applications. Its real capabilities have never been exploited and its use seems to us to be particularly appropriate in radio and TV broadcasting buildings.

The use of one of the several panel types now available is illustrated in Figs. 9 to 12a.

Fig. 9 illustrates the ease of erection of the panels, here being employed as an interior partition.

A cut-away section of a wall panel is shown in Fig. 10. Here the type shown has been perforated on one side and filled with three inches of fiberglas. As such, it makes an extremely efficient sound absorber which protects permanently the soft absorbent material. This is far superior to fiberglas contained by wire netting which is often used. Any degree of sound absorbency can be achieved in a room built of these panels, i.e., a wall which required 50% treatment would be composed of perforated panels alternating with solid faced ones. Where careful sound insulation of adjoining spaces is necessary, this can be accomplished by using two walls slightly separated and isolated both from each other and from the floors and ceilings.

Fig. 11 illustrates the ceiling and floor panels which may also be perforated or not as the occasion demands. Here we have one material brought to the site completely assembled and acting both as the structual floor and as the sound absorber. These panels are now in use in a large recreation building in Binghamton, N.Y. In this building, an auditorium is located directly over a bowling alley, and insulation is so effective that no disturbance reaches the auditorium.

Fig. 12 shows the cellular floor in its role as a conveyor of wiring and cable. A predetermined grid is established on the floor with knockouts at frequent intervals. It becomes possible to tap power, telephone, and intercom lines at any point on the floor. The system is unaffected by the rearrangement of walls.

In Fig. 12a, we have an office using the perforated panels. These are painted steel. In areas in which aluminum would be appropriate, maintenance costs are reduced to zero.

Our attempt here has been to suggest to prospective builders of radio and TV broadcasting stations a means of employing the benefits of prefabrication and standardization which are difficult to find in the building industry today.

We have also tried to establish some principles for the functional arrangement of broadcasting stations. We have chosen as our means a theoretical example, and it would be well to emphasize that no specific case could ever be solved in the same way. Irregularities of the site, relationships to other buildings, the necessity of building within an existing structure, use of existing electronic equipment, and the operating characteristics of each organization will mold each building into a different outward form.

The direct cooperation of station executives, production personnel, engineers, and the architect will be required to produce a plan capable of solving its present problems and of meeting all of the surprises that the growing industry will inevitably produce.

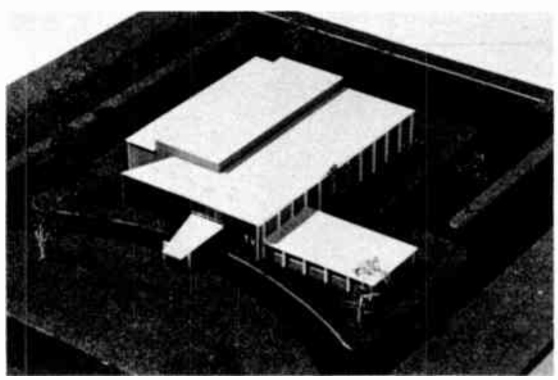


Fig. 1 - Over-all view of completed radio-TV building.

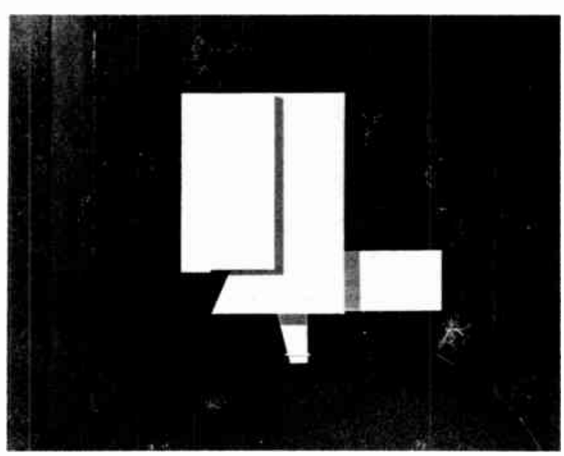


Fig. 1(a) - Top view of radio-TV structure.

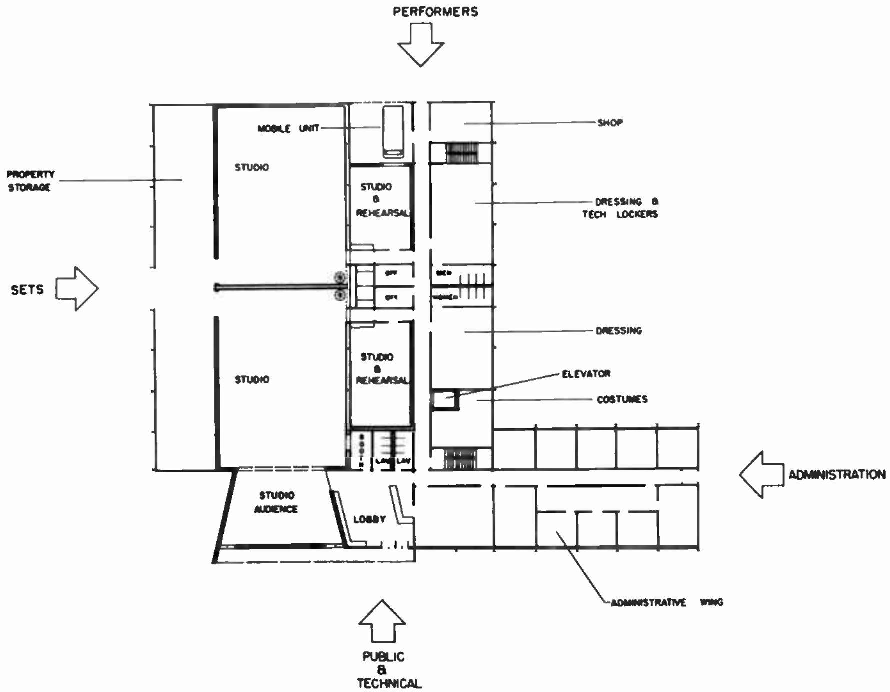


Fig. 2 - Ground level plan for final stage of building.

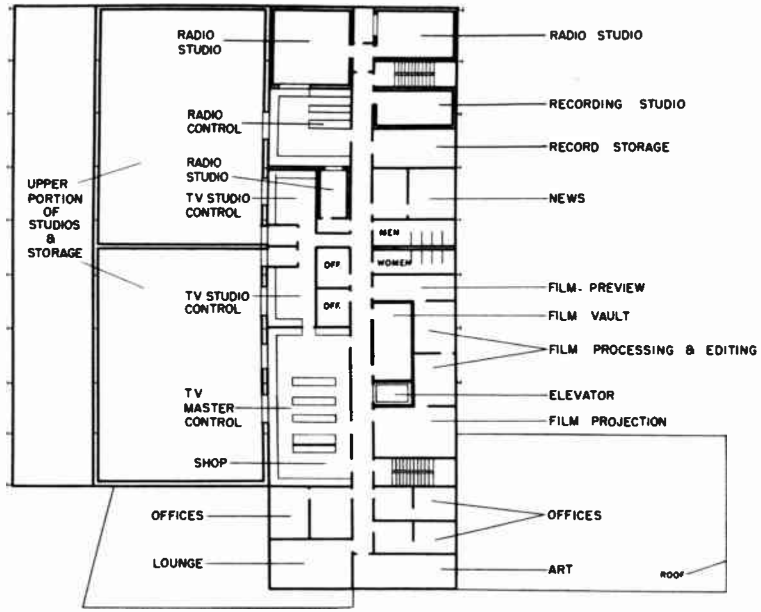


Fig. 3 - Upper level plan of fully expanded scheme.

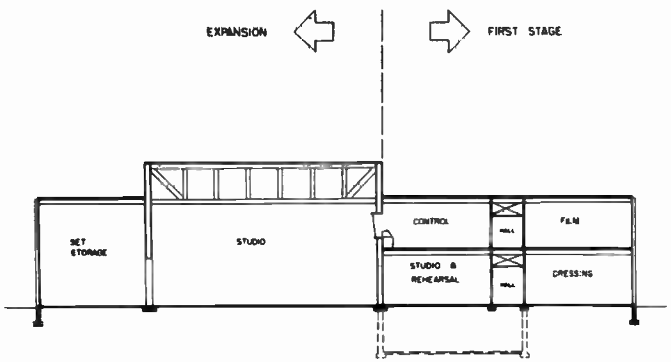


Fig. 4
Cross-sectional drawing which shows vertical relationships among various sections of building.

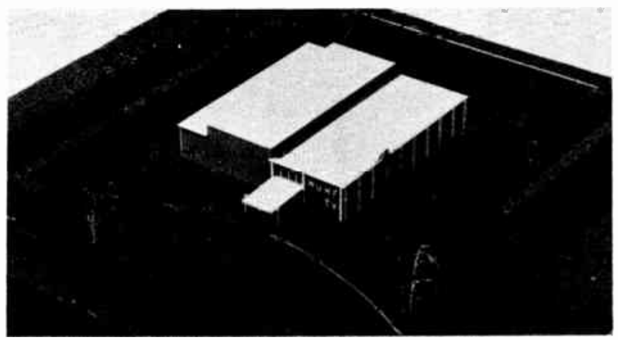


Fig. 5
View of radio-TV building without administrative wing.

Fig. 6 - Building in first stage of construction.

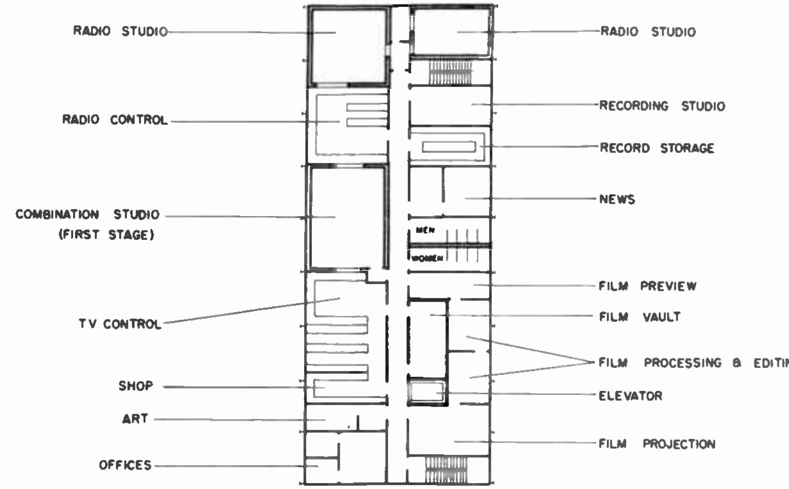


Fig. 7 - Plan for ground level, in the first stage.

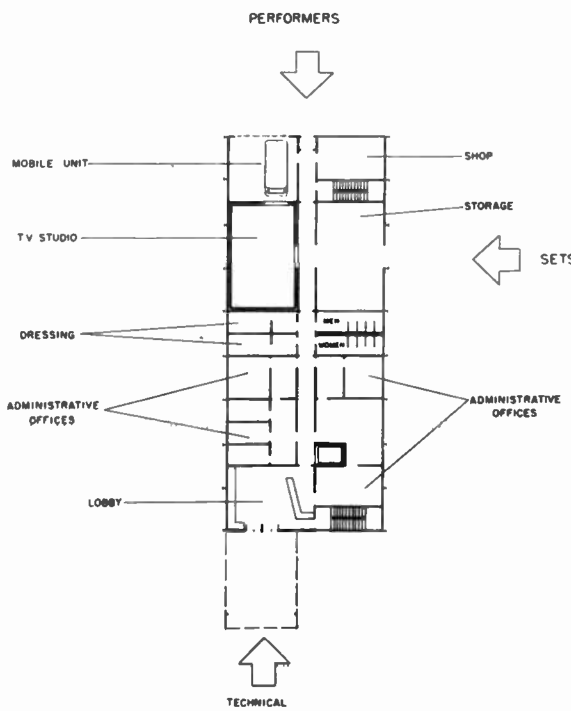


Fig. 8 - Upper level plan, during first stage.

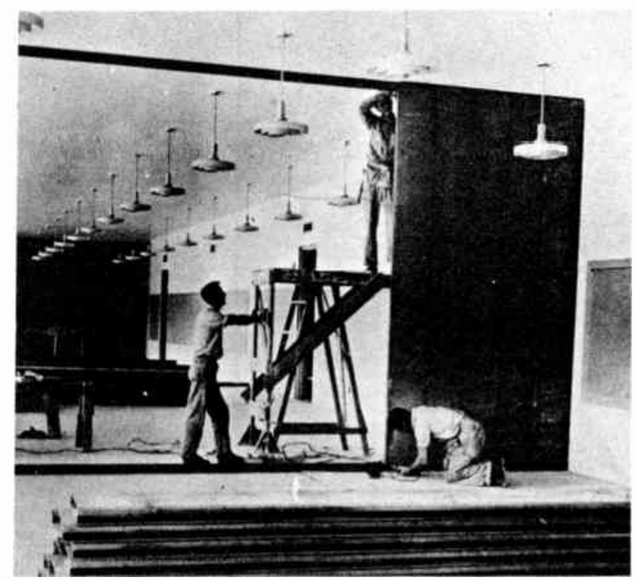


Fig. 9 - Erecting removable steel panel as an interior partition.

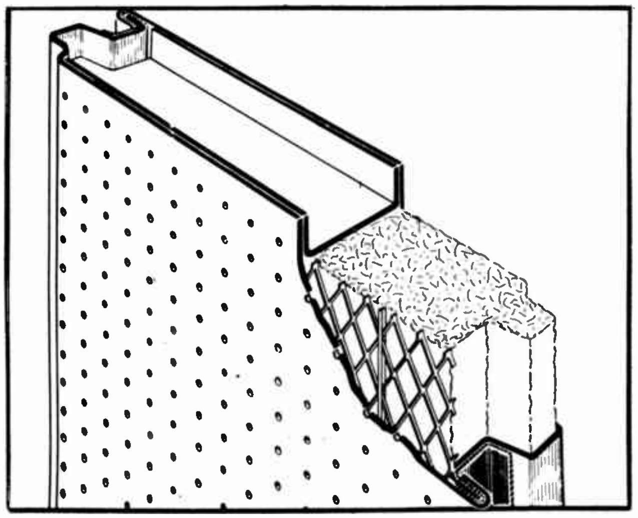


Fig. 10
Cut-away section of perforated panel, filled with fiberglas for sound absorption.

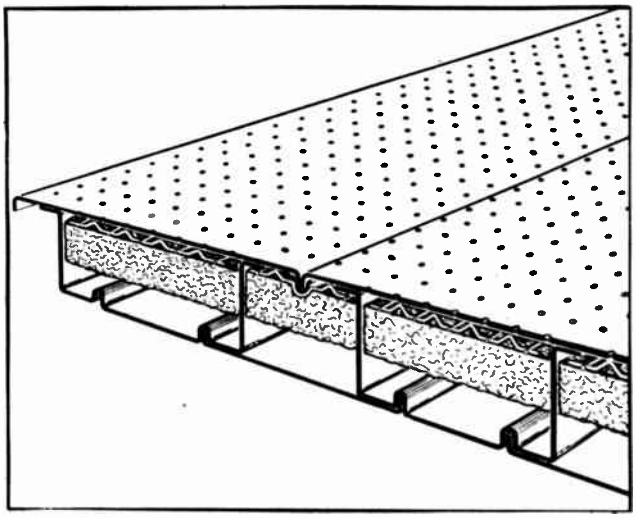


Fig. 11
Ceiling and floor panels which may be perforated or not as requirements dictate.

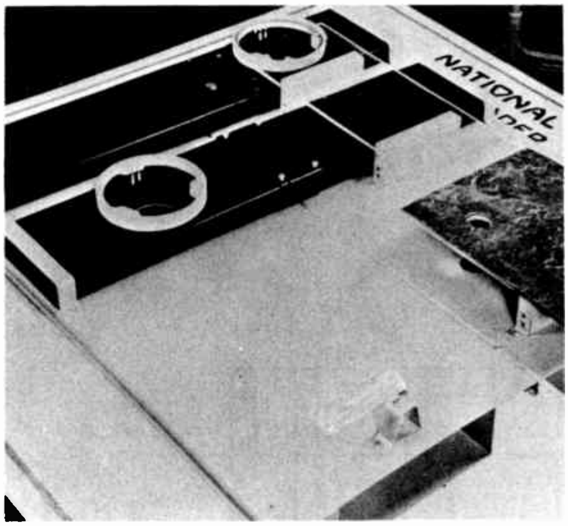


Fig. 12
Cellular floor which serves as conveyor of wiring and cable.

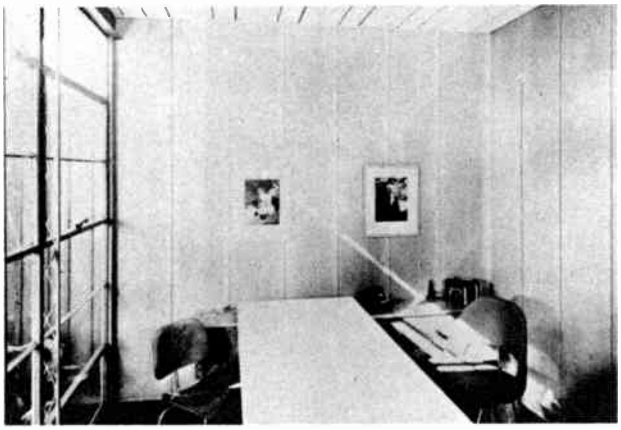


Fig. 12(a)
Office featuring use of perforated removable panels.

FASHIONS IN TV TRANSMITTING ANTENNAS

Frank G. Kear
Kear and Kennedy
Washington, D. C.

and

John G. Preston
American Broadcasting Company, Inc.
New York, New York

Part I

With the advent of nation-wide commercial television operation during the Mid-Forties and its immediate acceptance by the Public, it became the problem of the prospective television broadcaster to get his station on the air as rapidly as possible and with the equipment then available. The industry was prepared for this and produced the necessary quantity of television transmitters, antennas and studio equipment to meet this demand. By the time of the so-called "Freeze", equipment had been standardized and in particular, the antennas in use were limited to those having from three vertical bays up to a total of six vertical bays. The power output of the transmitters was limited but there was likewise a limit on the effective radiated power and an even greater limit was imposed when the height of the antenna exceeded 500 feet. Because of these limiting factors, the service provided by the available equipment and antenna system was generally satisfactory and in the cases where television coverage appeared inadequate it was generally ascribed to the transmission characteristics of the very high frequencies.

Although practically all of the 100-odd stations operating at the time of the "Freeze" had antenna heights of 500 feet or less, there were a few exceptions, notably those on Mount Wilson, near Los Angeles, and in particular, KECA-TV, the Los Angeles operation of the American Broadcasting Company. Operation from this point gave what was probably the first definite indication of the need for control of the radiation pattern of the transmitting antenna. The studies undertaken during the "Freeze" on the need for correcting the Mount Wilson situation laid the groundwork for the solution of the problems presented at the termination of the "Freeze" when increases in allowed power and increases in antenna height emphasized the necessity for controlled radiation.

The first television antennas were designed to augment the effective power of the relatively low-powered transmitters then available by concentrating the energy along the horizon. This of course was only accomplished by reducing the energy radiated in other directions but with the limited gains then contemplated and the limited antenna height of 500 feet or less, this introduced no problem.

Figure 1 shows a portion of the vertical directivity pattern of a typical three-bay superturnstile antenna. It will be noted that from a maximum along the horizon the field tapers off gradually until it reaches the value of ten per cent of the maximum field at some 15° below the horizon. The field does not drop off to a null until an angle of 19° below the horizon is reached. If we refer now to Figure 2 we can interpret the effect of this drop-off in field in terms of the distance from the transmitter at which serious effects may be noted. It will be seen that for an antenna height of 500 feet the field does not decrease to the ten per cent value until one is within 0.3 mile of the transmitter. For most transmitter locations this would not create a noticeable area of low signal. However, if we increase the antenna height to 1000 feet the drop-out occurs at 0.6 mile and for 2000 feet it is 1.2 miles from the transmitter. At these distances the area of low signal is more likely to lie in populated districts.

If now we wish to obtain more effective radiated power by employing an antenna of greater gain, we might use a six-bay or even a twelve-bay antenna. This would provide quite high values of effective radiated power with relatively low investment so far as transmitter costs are concerned. However, if we refer now to Figure 3 which shows the vertical pattern of the six-bay antenna and of the twelve-bay antenna superimposed on that of the three-bay antenna, it will be seen that the rapid decrease in field as we go below the horizon might well introduce serious problems, particularly at antenna heights of 1000 feet or more. The twelve-bay antenna, for example, has three complete nulls, occurring at roughly 5°, 9° and 14°, before reaching the first null of the three-bay.

Figure 4 is a plot of the calculated field intensity against distance for a three-bay antenna and for a six-bay antenna with an effective radiated power of 100 kw in the frequency range 54-88 mc, and an antenna height of 500 feet. The effect of the nulls in the pattern is very apparent; however, they are still very close to the transmitter and in many cases would be unnoticed. If we refer to a situation more typical of present planning we should consider Figure 5, which shows the calculated field intensity versus distance with 100 kw effective radiated power (i.e., along the horizon) and with an antenna height of 1000 feet for the three-bay antenna and for the twelve-bay antenna. It should be noted particularly that the major limitation of the twelve-bay antenna is not so much the existence of the additional nulls but rather is directly related to the rapid decrease in radiated field at very small angles below the horizon. Since complete nulls are rarely present in practice due to slight inaccuracies in tuning and adjusting the antenna, it has frequently been maintained that there will be no adverse effect on coverage if the twelve-bay antenna were employed rather than, for example, the three-bay. Inspection of Figure 4 reveals the fallacy of this argument. Even if we were to assume the null to be filled in to a value equal to the peaks of the lobes, the field from the twelve-bay antenna remains substantially constant from a point five miles away from the transmitter to within a half mile of the transmitter, whereas that from the three-bay antenna continues to rise (under the same assumptions as to null fill-in) in to the transmitter site. From a competitive viewpoint this is an important factor which cannot be overlooked. If a ratio of signal intensities between two otherwise comparable transmissions is of the order of 5 to 1, viewer preference will invariably select the strong signal. When, in addition, the field drops out in the presence of the nulls to the extent normally found in practice, this ratio may exceed 20 to 1.

Figure 6 shows a similar comparison to that of Figure 5 with the exception that the antenna height has been increased to 2000 feet. Here again a rapid fall off in signal as well as the presence of nulls, even though those nulls may be partially filled in, produces relative field intensities within the first ten miles of the transmitter which strongly favor the three-bay antenna.

If now we consider the UHF situation, where available transmitter power is relatively low and admittedly inadequate at the present time, we have an even more difficult problem. Here we are considering antennas with gains of the order of 15 to 25. Such gains are required at present in order to meet reasonable effective radiated power values, but they are misleading in that no appreciable field is radiated at other than angles very close to the horizon; consequently, the coverage problem becomes even more acute, particularly with antennas of substantial height.

Fortunately, we have at hand a means for correcting this situation; that is, the application of directional antenna design such as is commonly used in standard broadcasting. By application of the same formulas and principles presently used to compute directional antennas where the directivity is in the horizontal plane, we may adjust the directivity in the vertical plane until we arrive at an optimum distribution of such radiation. It is, of course, impossible with a physically realizable structure, to abstract all of the energy ordinarily radiated above the horizon and direct it below; however, it is possible to change the shape of the directive pattern materially and approximate this condition.

The procedure to be followed in order to arrive at a suitable antenna configuration is similar to that employed in determining upon a suitable pattern for standard broadcasting. Profiles are prepared in the various directions in which service is important or those directions possessing physical characteristics which indicate necessity for special study. From these profiles a plot of distance versus angle below the horizon can be prepared. The minimum field which would be acceptable at the important points along each profile must next be determined, basing this determination upon the nature of the terrain, density of population and the type of building construction to be encountered.

Particular attention must be paid to the point beyond which the coverage becomes less important in order to determine the extent, if any, to which the main lobe of the antenna may be depressed. These data permit calculation of the desired unattenuated field which should be radiated at each of the pertinent angles below the horizon. By suitable adjustment of the field ratios and phasing between the various elements of the antenna, the vertical directional pattern is then made to conform as nearly as possible to the values previously determined to be necessary. Obviously, the more elements employed, the more flexible the pattern; however, in general it has been found from weight, size and complexity considerations that six elements are satisfactory under many conditions, and that twelve represents for VHF a reasonable maximum number.

Part II

Let us consider, for example, a six-bay antenna as modified for use by KECA-TV, of the American Broadcasting Company, on Mount Wilson. The standard six-bay antenna divides the power equally above and below the horizon and has its first null 9.5° below the horizon. As located on Mount Wilson this null coincides with the location of the city of Pasadena, an important service area. Furthermore, the extreme limit of service to the west is at the land's end, some 31 miles from Mount Wilson, and this is 2° below the point of maximum radiation of the standard six-bay antenna. Figure 7 is a typical profile from Mount Wilson with the areas to be served and depression angles marked thereon.

By various devious means we determine upon the following field ratios and phase relations between the elements of a modified six-bay antenna for Mount Wilson:

TABLE I

MODIFIED 6-BAY SUPERTURNSTILE ANTENNA

Bay	Field Ratio	Phasing-Degrees
1 (Bottom)	1	-35
2	1	- 5
3	2	0
4	2	+30
5	1	+35
6	1	+65

With this configuration the antenna produces a pattern shown in the solid line in Figure 8. Here the dotted line represents the conventional three-bay antenna with unity field ratios and excited in phase. Inspection of the two curves discloses the fact that the first null of the modified six-bay antenna now occurs at approximately the same point as that of the three-bay antenna and furthermore, that the peak of radiation from the array occurs at 2.5° below the horizon, so that a maximum amount of energy is directed over the service areas desired and the region of very low fields in the Pasadena area has been eliminated. Figure 9 shows the computed fields westward from Mount Wilson for the standard six-bay antenna and the modified antenna. The degree of improvement is apparent. It should also be noted that with the antenna modified in this fashion the power radiated below the horizon is 2.1 times the power radiated above the horizon which is a marked improvement in efficiency.

The redistribution of energy results in decreasing the gain of the six-bay antenna structure to approximately 80 per cent of the gain figure for the in-phase antenna; however, it is adequate to permit the use of a reasonably powered transmitter and still obtain the maximum effective radiated power permitted by the FCC.

A different type of problem is encountered when we consider operation in the high frequency portion of the VHF band with antenna heights lower than the value at which a reduction in power is required. In this case it is desirable to maintain a figure for the overall gain which will permit achieving an effective radiated power of 316 kw with a transmitter power of 50 kw. Taking into account transmission line losses, harness losses, etc., this means that a pattern gain of at least 7.5 must be maintained. A twelve-bay antenna was selected as possessing sufficient elements to have a reasonable degree of flexibility and at the same time having a gain in the in-phase condition high enough so that the rephasing and redistribution of element power would not reduce it below the 7.5 figure.

Let us consider a specific case. It is proposed to operate KGO-TV, the American Broadcasting Company station in San Francisco, California, with an effective radiated power of 316 kw and to locate the antenna on a 700-foot tower which is in turn located on top of Mount Sutro. This gives an effective height for the antenna of 1550 feet. Profiles were drawn in the various pertinent directions from the Mount Sutro site and the unattenuated field required to establish a minimum field of 150 mv/m along each profile to the point where attenuation takes control was determined. This value of 150 mv/m was selected as representing a value high enough to provide a good service signal throughout the heavily built-up areas of San Francisco and Oakland.

The equations for the pattern of a twelve-section antenna were then adjusted until the required unattenuated field was obtained. The following table represents the combination finally arrived at:

TABLE II

MODIFIED 12-BAY SUPERTURNSTILE ANTENNA

Bay	Field Ratio	Phasing-Degrees
1 (Bottom)	1	0
2	1	0
3	1	0
4	1	0
5	2.4	-30
6	2.4	-30
7	2.4	+30
8	2.4	+30
9	1	0
10	1	0
11	1	0
12	1	0

Figure 10 shows a portion of the vertical plane pattern of a standard twelve-bay in-phase antenna and the modified pattern selected for Mount Sutro. The gain of the antenna thus modified is approximately 7.6, which will permit an effective radiated power of 316 kw with a 50 kw transmitter. The field produced by these two antennas was calculated and is shown on Figure 11. It will be noted that the field does not drop below the 150 mv/m value until one is within 0.5 mile of the antenna, whereas the standard antenna shows numerous areas in which the field would be far below this value.

When we examine the problem of UHF coverage, the factor of total available power assumes an important position. Although the Federal Communications Commission authorizes effective radiated powers of up to 1000 kw for UHF stations, present-day transmitters will not permit us even to approach this figure without absurdly high antenna gain figures. The problem at the present time is to find the best possible distribution for a total available power of 1000 watts to 12 kw in order that UHF stations going on the air in the immediate future will establish a field intensity coverage pattern which will permit them to approach competition with VHF.

As an example of one such problem, we have WAKR-TV in Akron, Ohio. Here the transmitter is located on a rooftop in the center of the business district. This is desirable in that all of the high values of field intensity cover areas which need these values, but it imposes an additional problem in that the presence of nulls or sharp decreases in field close in to the antenna would make the coverage ineffectual.

Several companies collaborated in devising a UHF antenna which would provide service for WAKR-TV. The one finally selected was a TFU-24-BM (slot-type radiator). This antenna is made up of sixteen sections and by dividing the power unequally between the upper ten and the lower six sections and suitably phasing the driving system, a directional pattern was achieved as shown in Figure 12. It will be noted that the main beam is depressed one degree below the horizon and the nulls are filled in down to at least fourteen degrees. The gain of this antenna computes to be 79% of the standard untilted antenna; that is to say, 79% of 24, or 19.

The computed field intensity produced by this antenna with an effective height of approximately 350 feet and a total power input to the antenna of 1000 watts, is shown in Figure 13. The close-in field out to the point of noticeable attenuation is maintained very close to 100 mv/m. While this is still a marginal field in a competitive market, it represents what is believed to be the best distribution of the available energy. It is the more acceptable since there is reason to believe that transmitter powers of 10 to 12 kw will be available within the next twelve months, which if used in connection with this antenna would increase the close-in field to an average of 300 mv/m. This compares favorably with the value of 100 to 200 mv/m used as a design minimum in one of our VHF illustrations.

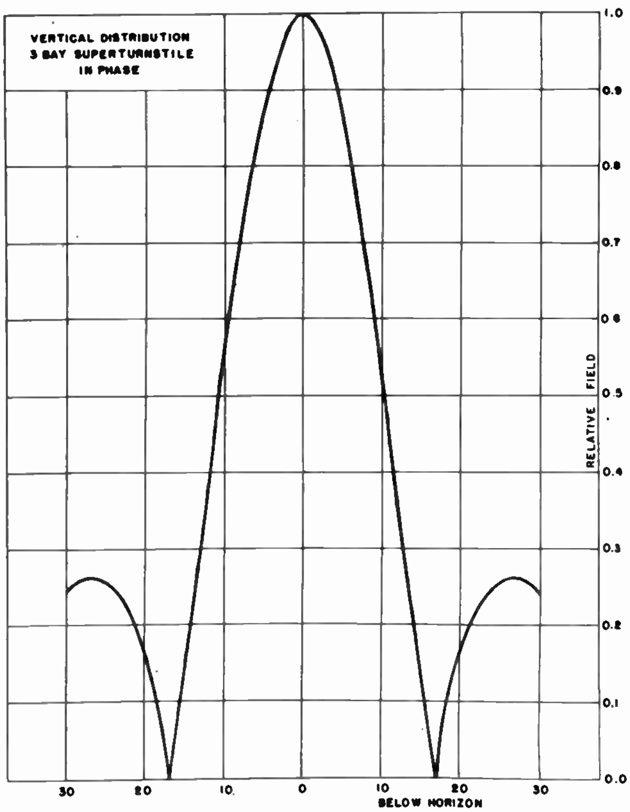


Fig. 1
Vertical plane field distribution.
Typical 3-bay superturnstile.

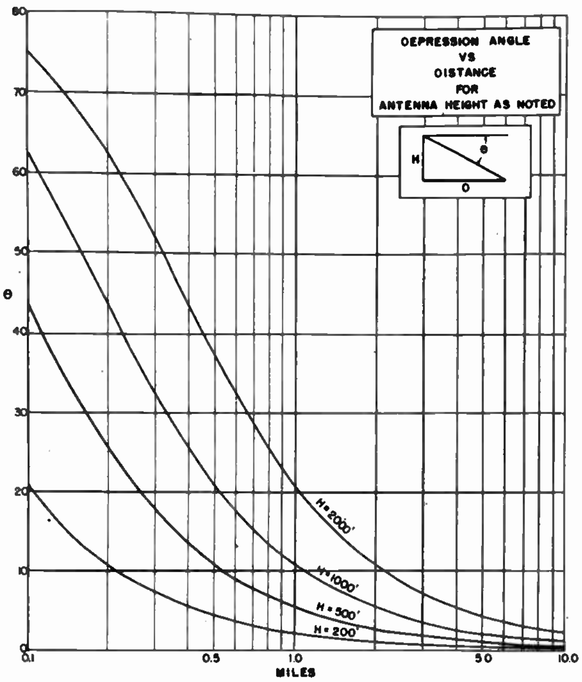


Fig. 2
Depression angle vs. distance for
various antenna heights.

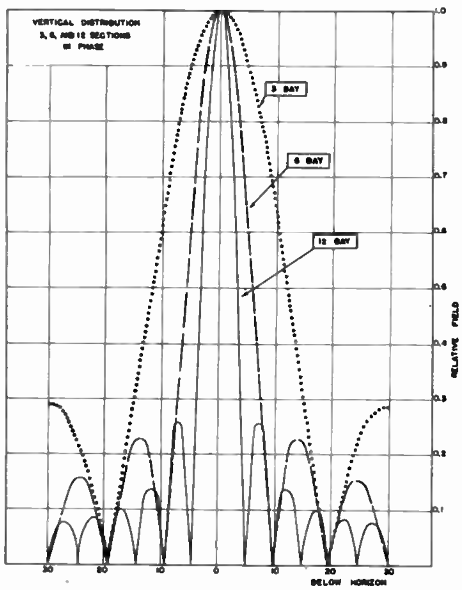


Fig. 3
Vertical plane field distribution.
Typical 3-, 6-, and 12-bay
superturnstiles.

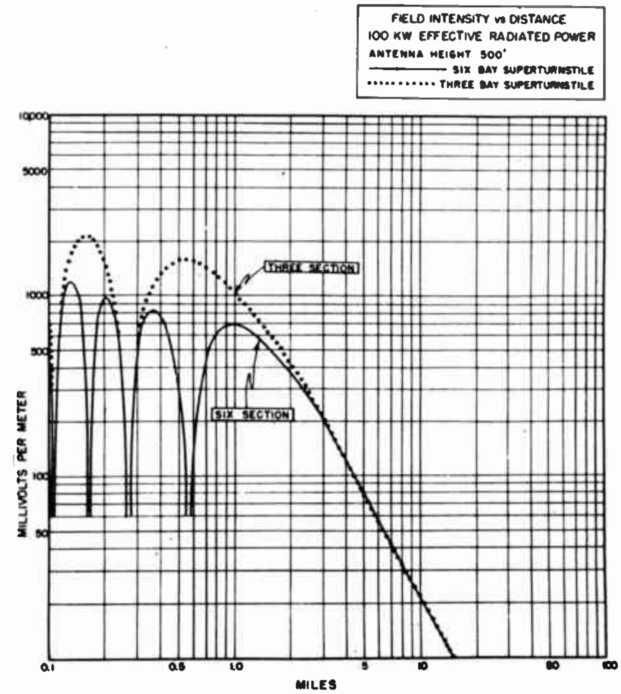


Fig. 4
Calculated field intensities. Typical
3- and 6-bay antennas
for 500-foot effective height.

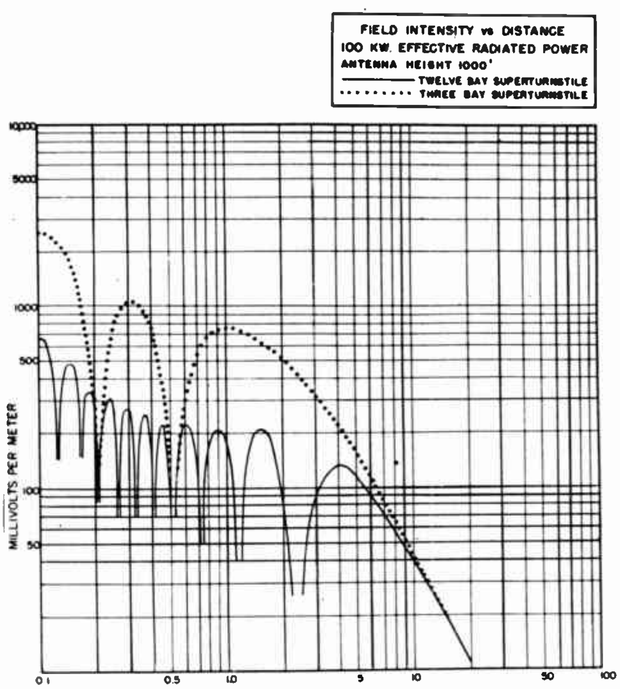


Fig. 5
Calculated field intensities. Typical
3- and 12-bay antennas for
1000-foot effective height.

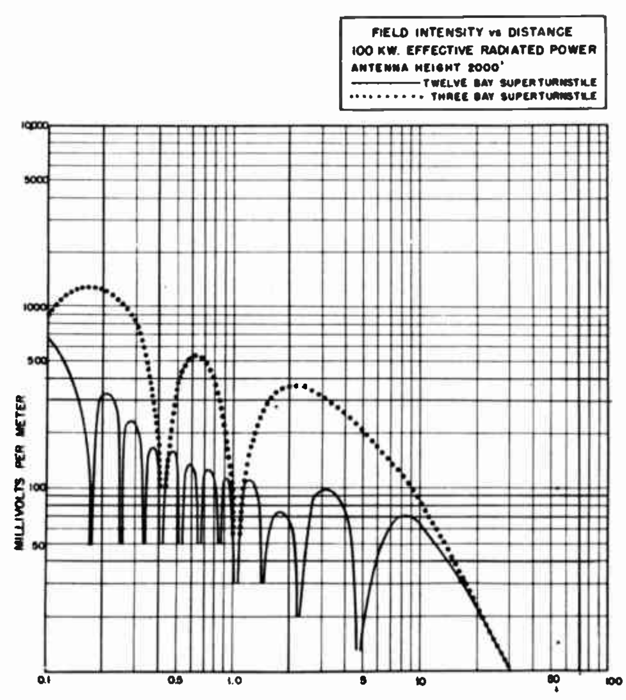


Fig. 6
Calculated field intensities. Typical
3- and 12-bay antennas for
2000-foot effective height.

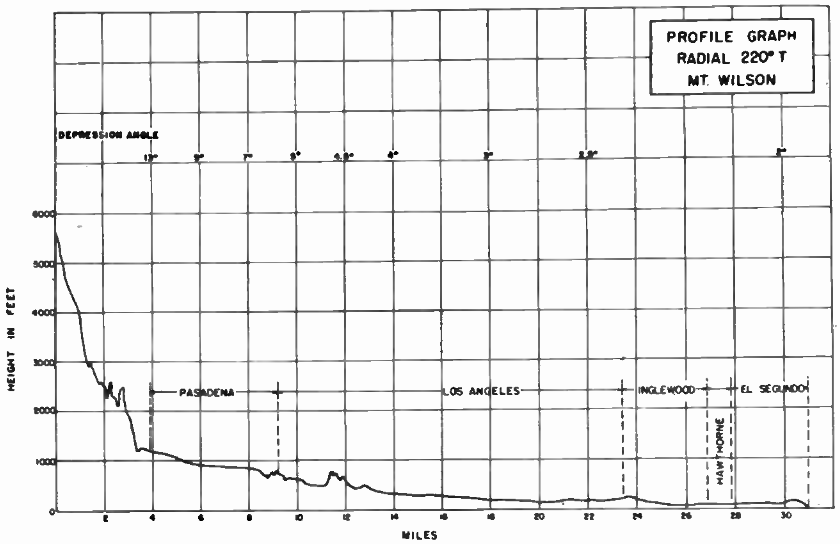


Fig. 7
Profile from Mount Wilson - azimuth 220° true.

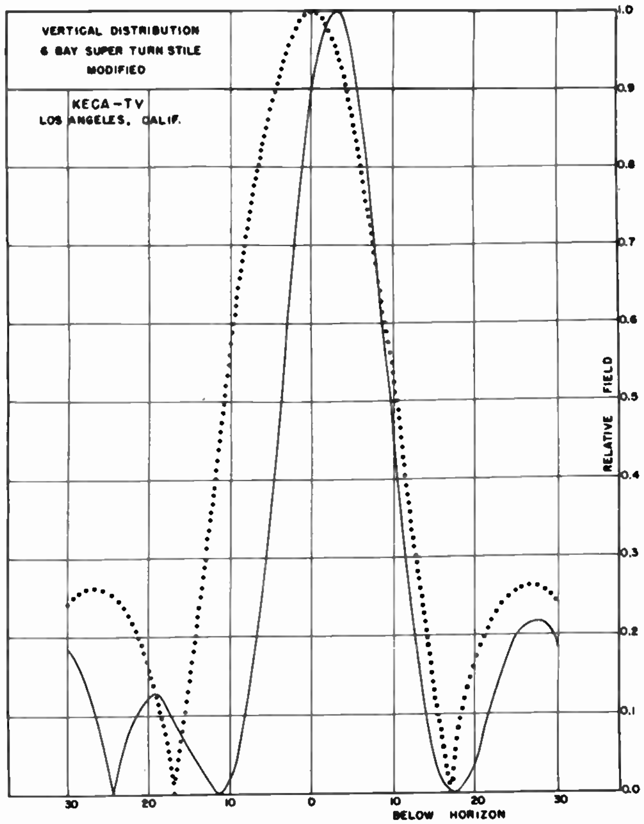


Fig. 8
Vertical plane field distribution.
Standard 3-bay and modified
6-bay antennas.

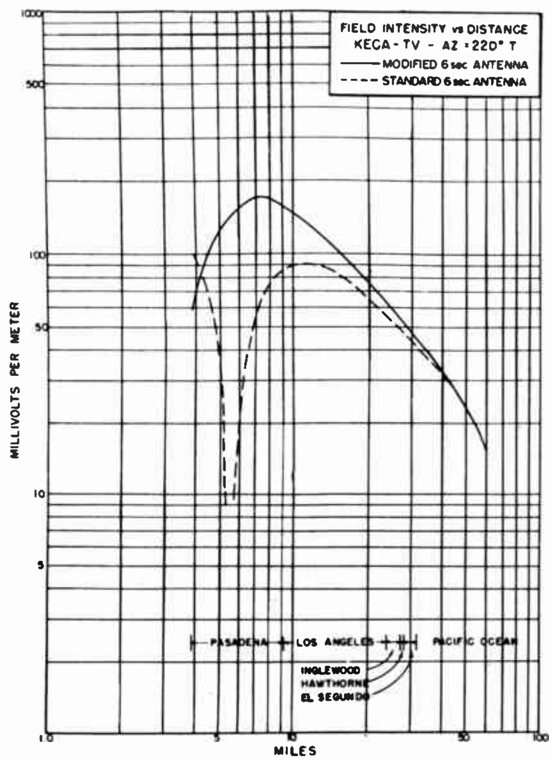


Fig. 9
Calculated field intensities for KECA-TV
on Mount Wilson with standard 6-bay
and modified 6-bay antennas.

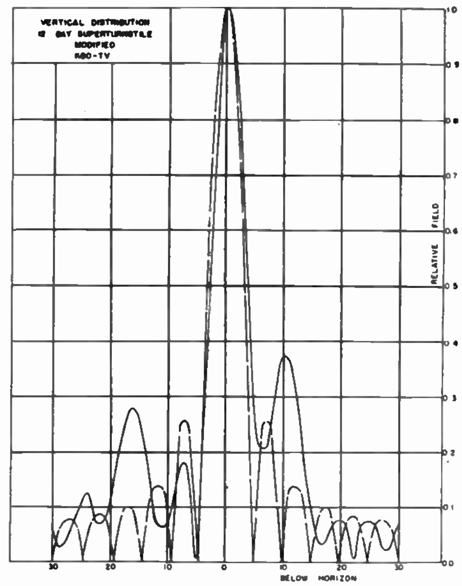


Fig. 10
Vertical plane field distribution.
Standard 12-bay and modified
12-bay antennas.

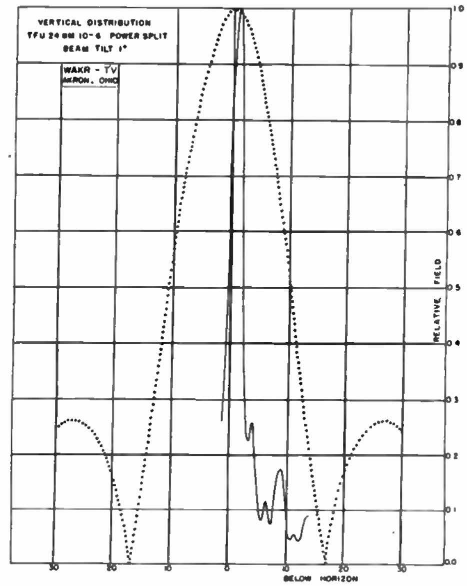


Fig. 12
Vertical plane field distribution.
Standard 3-bay and modified
24-bay antennas.

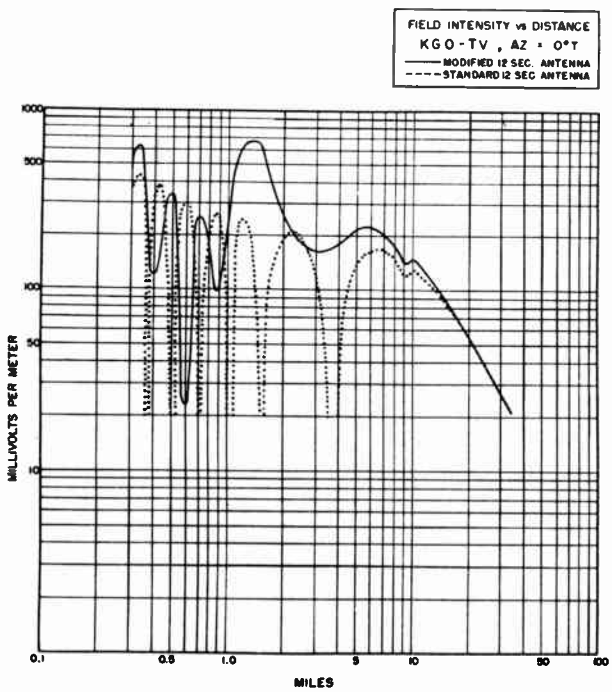


Fig. 11
Calculated field intensities for KGO-TV
on Mount Sutro with standard 12-bay
and modified 12-bay antennas.

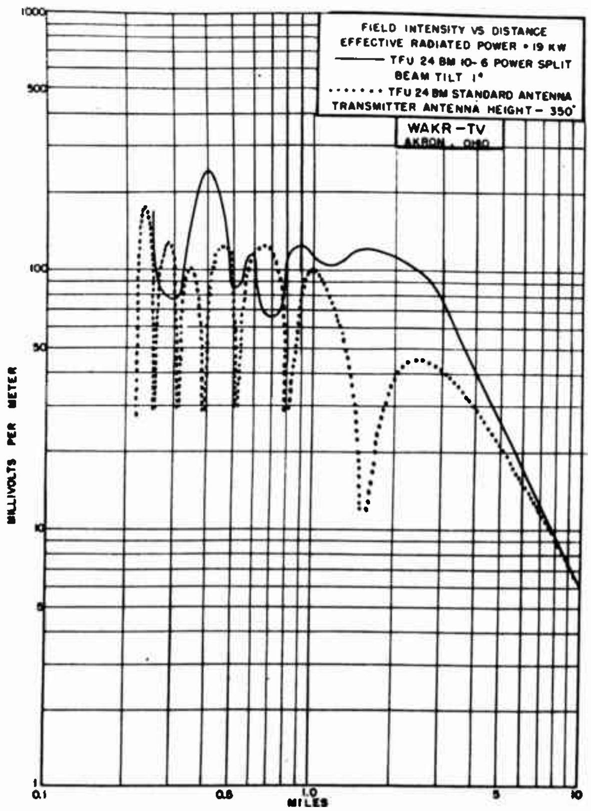


Fig. 13
Calculated field intensities for
WAKR-TV in Akron with standard
uhf and modified uhf antennas.

HIGH GAIN AMPLIFIERS FOR
HIGH POWER TELEVISION TRANSMITTERS

John Ruston
Allen B. Du Mont Laboratories, Inc.
Clifton, N. J.

Introduction

The design of broadband power amplifiers for TV transmitters is restricted by the availability of tubes suitable for operation in the very high frequency bands, it being necessary to design an amplifier around one of a very limited number of suitable types. Among the main requirements of a suitable tube are very low inductance connections to the electrodes and very good shielding between the input and output electrodes. These two requirements are met by the ring seal type of construction which is now used almost universally in the higher power tubes intended for VHF operation. This construction enables one electrode to be used as a shield between the input and output circuits, complete continuity with an external shield being obtained by connecting it to the ring terminal of the tube. High power tubes of this construction heretofore available have all been triodes in which the grid is used as the shield. This necessitated the use of the grounded grid circuit with its inherently low power gain. Now that high power ring seal tetrodes have become available, it is possible to obtain very much higher power gain since the shielding between input and output circuits can be provided by the screen thus enabling a grounded cathode circuit to be used.

In this paper, it is proposed to examine the theoretical and practical potentialities of the new high power tetrode, particularly as regards power gain, when used in broadband linear amplifiers for TV transmitters.

Theoretical Performance

Simplified Circuit

The simplified grounded cathode tetrode amplifier shown in fig. 1 has the tube input capacitance C_1 and output capacitance C_o resonated by external inductances L_1 and L_o respectively. DC isolating capacitors, etc. have been omitted so that only the basic rf circuit is shown. It is well known that in order to obtain a bandwidth Δf in the input circuit, its loading must be equivalent to a shunt resistor R_1 where:

$$R_1 = \frac{1}{2\pi \Delta f C_1} \quad \cdots\cdots\cdots\cdots (1)$$

The effective output load resistance R_o is usually the value which gives maximum signal power output from the tube (without exceeding its rated anode dissipation) as determined by drawing a suitable load line on its constant current characteristics[1, 2]. The resistance determined in this way will generally be less than $\frac{1}{2\pi \Delta f C_o}$ so that the output circuit bandwidth will be adequate. If this is not so, then R_o must, of course, be reduced to give the desired bandwidth and a lower power output accepted.

From the load line, the peak rf grid voltage E_g required to give the maximum power output W_o can be determined in the conventional way. The power gain at bandwidth Δf is then W_o divided by $\frac{E_g^2}{2R_1}$

Maximum Tube Capability

An indication of the power gain possible in a practical high power TV amplifier can be obtained by examining the constant current characteristics of the Eimac type 4W20000A tetrode as shown in fig. 2. The load line OS has been drawn to show the downward swing of the anode voltage at sync peak level, i.e., the half cycle during which anode current flows in a class B amplifier. Point O corresponds to the maximum rated dc anode voltage of 8000 and to a dc grid voltage of -180 which is suitable for class B operation. Point S corresponds to an anode potential of 1500 volts below which the characteristics bend upwards and would result in sync compression. The peak rf anode voltage at sync peak level is thus 6500 volts and consequently at black level it is .75 x 6500 = 4875 volts. The minimum anode potential at black level is then 8000 - 4875 = 3125 volts, corresponding to point B. The slope of line OB, representing the downward swing of the anode potential at black level, is adjusted so that the anode dissipation is 20 Kw, the maximum rating for the tube. The slope of line OS is then adjusted so that it represents the same load impedance as that determined by OB. The lines do not quite coincide because of the somewhat non-uniform spacing of the constant current curves. This results in some slight sync stretching which is compensated by reducing the relative sync amplitude of the transmitter video input signal.

From the load lines OS, OB the following data can be computed[1, 2].

	Sync Level	Black Level
Anode current - amperes		
Peak rf	23	17
Peak rf fundamental	10	7.6
DC	6.4	4.78
Anode power - kilowatts		
Input	51.2	38.2
Output	32.5	18.2
Dissipation	18.7	20
Peak rf grid voltage	350	280

Output load impedance $= \frac{6500}{10} = 650$ ohms

The tube input and output capacitances are 125 uuf and 23 uuf respectively and hence:

Output circuit bandwidth = 10.8 Mc

For an input circuit bandwidth of 5 Mc:

Effective input load R_1 = 255 ohms

RF driving power at sync level = $\frac{350^2}{2X255}$ = 240 watts

Power gain = $\frac{32500}{240}$ = 136

Input Circuit Loading

The above data shows that with a bandwidth suitable for TV transmission the maximum peak power capability of the tube is 32.5 Kw at the unusually high power gain of 136. The design of a practical high power amplifier then becomes a mechanical problem of devising a physical layout which represents the simple electrical circuit of fig. 1. For the output circuit, an adjustable inductor can readily be designed and connected to resonate the tube output capacitance. A loading resistance can then be coupled to the inductor. The input circuit appears at first sight to present two major problems: (a) to connect a loading resistance capable of dissipating 240 watts without changing the circuit reactance and (b) to obtain an rf ground on the cathode which is inside the tube and separated from the cathode terminal by an appreciable inductance. Fortunately, these two problems mutually resolve themselves since an inductance between cathode and ground effectively loads the input circuit.

In fig. 3, the circuit of fig. 1 is elaborated to show the grid and cathode lead inductances L_G and L_K respectively so that external connections are made to the tube at points G and K. The screen to grid capacitance is neglected for the moment and the input capacitance C_1 is then the grid-cathode capacitance of the tube.

If an rf ground is connected to the input circuit such that the total inductance between cathode and ground is L, then it can be shown[3] that the input circuit loading is equivalent to a shunt resistance R_1 where:

$$R_1 = \frac{1}{\omega^2 L C_1 g_m} \quad \cdots\cdots\cdots\cdots\cdots (2)$$

From (1) and (2)

$$L = \frac{\Delta\omega}{\omega^2 g_m}$$

Where: ω = angular operating frequency
gm = effective tube transconductance
$\Delta\omega = 2\pi\Delta f$

When suitable numerical values for the type 4W20000A tube are inserted in this formula, it is found that for a 5 Mc bandwidth the inductance L is equal to the cathode lead inductance L_K at about 90 Mc. Hence, at low band TV frequencies (57-85 Mc), additional series inductance L_S is required external to the tube, and so the rf ground is placed at point B in fig. 3. At high band TV, frequencies, L is less than L_K, and so the rf ground is placed at A, and the effective inductance between cathode and ground is reduced by adding the series capacitor C. This tends to reduce the bandwidth, and when the loading is increased to offset this, the power gain is reduced by a factor $\frac{C}{C + C_1}$. However, the theoretical power gain obtainable at 200 Mc is still of the order of 100.

It is interesting to note that the type of operation depicted in fig. 3 is intermediate between grounded cathode and grounded grid operation. If the rf ground is moved from the cathode, round the inductive part of the circuit to the grid, the bandwidth progressively increases and the power gain decreases. By controlling the position of the ground, the bandwidth or gain can be adjusted to the optimum for the particular application. This approach also makes apparent the effect of the grid-screen capacitance C_{GS} which was previously neglected. It is evident that the effective input capacitance to be used for computing bandwidth is the grid-cathode capacitance plus some fraction of C_{GS}. This fraction decreases from unity when the ground is at the cathode to zero when the grid is grounded.

A Practical Amplifier

Construction

At low band TV frequencies, it is quite practicable to construct an amplifier having a circuit like that in fig. 3 and with a performance approaching that previously calculated. The type of construction employed in a commercial high power TV amplifier presently being produced[4] is shown diagramatically in fig. 4. The rf circuit is contained in a shielded compartment which is divided into input and output sections by a horizontal grounded partition. The type 4W20000A tube is mounted with the anode downwards and with the grid and cathode terminal rings projecting up through a hole in the partition. Spring fingers connect the screen terminal ring to a plate which forms a bypass capacitor with the underside of the partition, the plate and partition being separated by sheets of teflon. The anode tuning inductor, consisting of two pieces of tubing joined by a movable shorting bar, is connected between the anode terminal and another plate which forms a bypass capacitor with the under side of the screen plate. The anode circuit is inductively coupled to an output circuit consisting of a similar type of inductor resonated by an air dielectric capacitor on the underside of the partition. Coupling can be adjusted by swinging the output inductor as indicated in the diagram. An adjustable capacitor couples the output circuit to the coaxial output transmission line. The anode and output circuits are adjusted as a pair of slightly over coupled tuned circuits to give a substantially uniform response, an adequate bandwidth and at the same time, provide the optimum load impedance for the tube.

In the upper section of the shielded compartment, the grid ring terminal is connected by spring fingers to a circular plate supported on insulators from a grounded bracket. The input tuning inductor connected between grid and ground consists of two parallel strips joined by a movable shorting block. A dc isolating capacitor is built into the lower strip, and the bias voltage is applied through an rf choke. RF drive is fed in through a flexible coaxial line which is tapped on to the input tuning inductor as shown. Spring fingers connect the cathode terminal of the tube to a plate supported on insulators above the grounded bracket. This plate is connected to ground through an adjustable inductor consisting of another pair of parallel strips joined by a movable shorting block.

The inductors all have sufficient range of adjustment to cover the low band TV channels 2 to 6, the shorting bars being preset for the channel on which the amplifier is to operate. Fine tuning of the input, anode and output circuits is obtained by means of small variable capacitors (not shown) operated by control knobs on the front of the unit. The position of the shorting block on the cathode inductor is determined by the amount of loading required on the input circuit.

Input Impedance

The tapping point on the input circuit for connection of the input cable is chosen so that the latter is approximately terminated by its surge impedance. However, since the input circuit loading is not very heavy, the terminating impedance will vary considerably over the passband which means that the driver unit supplying the input power will not be operating into a constant resistive load. The output of the driver unit with which this amplifier was designed to operate[5] consists of a pair of slightly overcoupled tuned circuits coupled to a 51.5 ohm coaxial output line. Hence, when a short length of coaxial cable is connected between the driver output and the amplifier input, the system becomes virtually a triple tuned circuit in which the coupling between the second and third circuits is a low impedance link. The system can then be aligned by using the conventional procedure for triple tuned circuits.

Performance

The performance obtained from the amplifier can, by careful adjustment, be made to agree quite closely with the theoretical figures. A peak power output somewhat greater than 30 Kw can be obtained with an all black picture without exceeding 20 Kw anode dissipation; so, in order to allow a reasonable safety factor the amplifier is rated for a maximum peak power output of 25 Kw. It was not readily possible to measure the power gain accurately because the driver was not operating into a resistive load, and so the driving power under normal operating conditions could not be measured directly; but, it was estimated that a power gain exceeding 100 could be obtained with adequate bandwidth for TV broadcasting. However, since the amplifier was designed to be driven by an existing driver unit having a rated peak power output of 500 watts, the amplifier is normally operated at a power gain of the order of 50.

Lower Sideband Re-insertion

It is an advantage if, in a TV transmitter, the necessary lower sideband attenuation can be obtained in the lower power stages because the filters required are then simpler and less costly. However, any non-linearity in the amplifiers following the filter can cause re-insertion of the lower sideband. It can, for instance, easily be shown that if an amplifier input signal contains only frequencies f_c and $f_c + f_m$ then a non-linear relationship between grid voltage and anode current can result in anode current components of frequency $f_c - f_m$. This will occur in the case of a single ended class B amplifier since non-linearity is inherent in the method of operation which produces an anode current waveform consisting of the grid voltage waveform with one half clipped off. The level of the lower sideband voltage present in the amplifier output is not readily amenable to calculation since the re-inserted voltage depends upon the anode circuit impedance, and it may augment or reduce that due to any lower sideband in the input voltage. However, it appears that in a grounded cathode amplifier the lower sideband voltage present in the output will generally exceed the maximum permissible irrespective of the amount in the input signal. The inclusion of an inductance between cathode and ground will evidently have a degenerative effect on the inserted sideband currents in the same way that it has the effect of loading the input circuit for the normal input signal. Here again, the amount of degeneration cannot be readily calculated since it depends upon the magnitude and phase angle of the section of the input circuit "seen" by the current. This in turn is modified by the coupling to the driver output circuit. It is, however, apparent that for a given set of conditions the degeneration will be proportional to the magnitude of the cathode inductance.

The existing driver unit used with the 25 Kw amplifier described above already had a built in lower sideband filter so it was most desirable that the lower sideband re-insertion in the amplifier should be kept to an acceptably low level. It was found that when the cathode inductance was such as to give the maximum power gain of about 100, then the overall lower sideband attenuation was not quite sufficient, being of the order of 18 db. A reduction of the power gain to about 50 increased the attenuation to the acceptable figure of 22 db. It was fortunate that this was the power gain required in the amplifier to give a comfortable operating power level in the driver.

Feedback and Stability

It is practically essential in a broadband TV amplifier that the feedback from output to input circuits be maintained at a very low level in order to avoid appreciable interaction between the tuning of the circuits when their passbands are being adjusted. If this requirement is satisfied, it will automatically preclude self oscillation of the tuned grid tuned plate variety at any frequency in the region of the operating frequency. Oscillation at any other frequency can of course be dealt with by the usual procedures for eliminating parasitic oscillations.

With the ring seal type of tube construction, coupling between input and output circuits external to the tube can be virtually eliminated. The remaining causes of feedback in the tetrode amplifier herein described are then: (a) anode to grid capacitance, (b) screen lead inductance, (c) anode current. If only the anode to grid capacitance is considered and the criterion for the maintenance of self oscillation is applied[6], it is found that when the power gain is 50, oscillations cannot be maintained below about 200 Mc. The amplifier would then be stable at low band TV frequencies, but the feedback voltage would be of the order of one third the input voltage and interaction between the tuning of input and output circuits would be excessive. However, it has been shown[7] that the screen lead inductance, i.e., the inductance between the screen electrode and rf ground, tends to neutralize the anode-grid capacitance, and at some frequency usually in the VHF band the tube is completely self neutralized. With the type 4W20000A tube mounted in the circuit depicted in fig. 4, the self neutralizing frequency is in the low band TV region, and the feedback is satisfactorily low over the whole band. No adjustment of screen to ground inductance is necessary. The final possible cause of interaction between the input and output circuit tuning, i.e., anode current, is negligible in a tetrode amplifier because the anode current is substantially independent of anode load impedance. This is in contrast to a triode amplifier in which considerable anode current variation occurs with change of anode loading. This results in variation of input circuit loading whenever the anode circuit is adjusted unless special precautions are taken to overcome it.

Conclusion

Both theoretical considerations and results obtained with a practical amplifier have shown that power gains up to 100, at a peak power output exceeding 30 Kw, can be obtained in the final amplifier of a TV transmitter using the new high power tetrode tube. Variable input circuit loading, obtained by means of a variable inductor connected between cathode and ground enables the bandwidth or power gain to be adjusted as desired. By reducing the power gain to about one half, the maximum obtainable, lower sideband re-insertion can be reduced enough to permit the inclusion of the lower sideband filter in the low power driver unit. As a result of the self neutralizing characteristic of the tetrode tube, the amplifier is very stable and there is negligible interaction between input and output circuit tuning.

References

1. R. I. Sarbacher, "Graphical Determination of PA Performance", Electronics; December, 1942.
2. "Tube Performance Computer", Eimac Application Bulletin #5, Eitel McCullough Inc., San Bruno, Cal.; 1952.
3. F. E. Terman, Radio Engineer's Handbook, McGraw Hill Book Co., New York, N. Y.; 1943, first edition; p. 472.
4. Du Mont, Model TA-183-A VHF Visual Amplifier and Model TA-182-A VHF Amplifier.
5. Du Mont, Acorn Series 5000 (Low Band).
6. I. E. Mouromtseff, "Tuned-Grid Tuned-Plate Oscillator", Communications; vol. 20, pp. 7-9; August, 1940.
7. W. G. Wagener, "500 MC Transmitting Tetrode Design Considerations", Proc. IRE; vol. 36, No. 5; May, 1948.

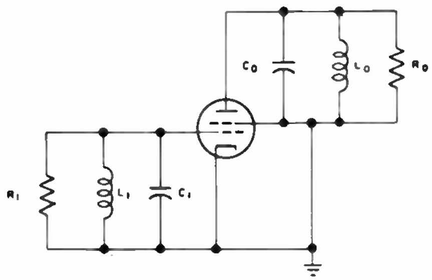
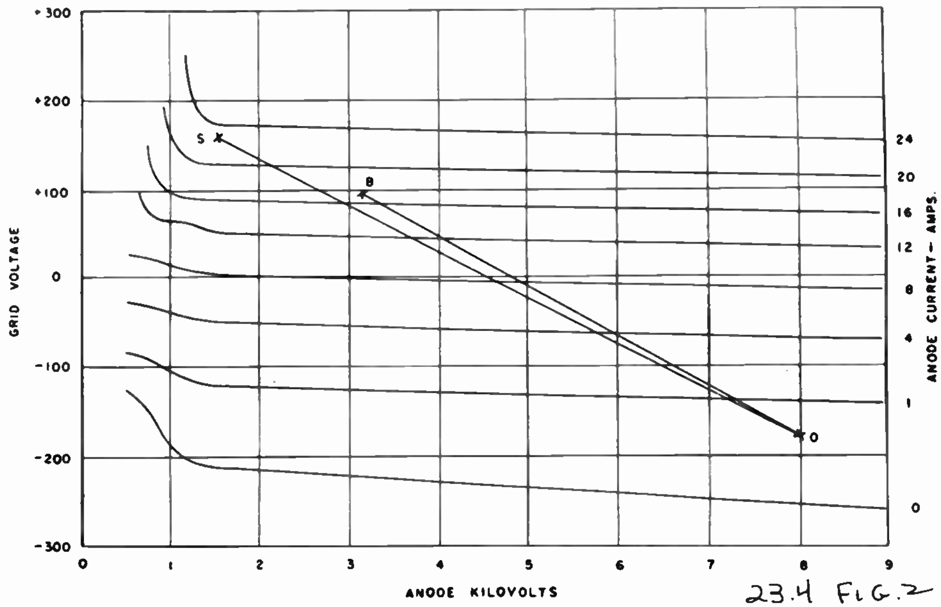


Fig. 1
Simplified grounded cathode tetrode amplifier.



23.4 FIG.2

Fig. 2
Eimac type 4W20000A constant current characteristics.

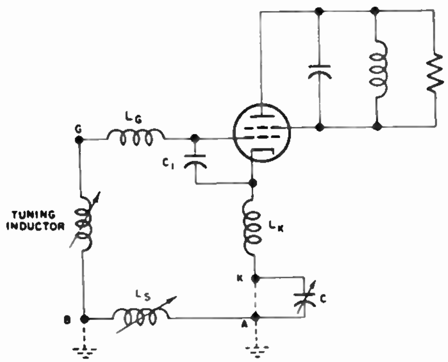


Fig. 3
Tetrode amplifier showing lead inductances.

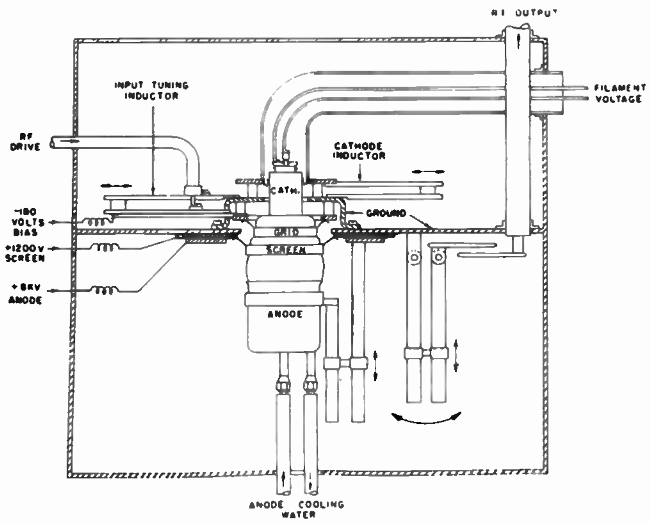


Fig. 4
Diagram showing type of construction.

OPTIMUM UTILIZATION OF THE RADIO FREQUENCY CHANNEL FOR COLOR TELEVISION

R. D. Kell and A. C. Schroeder
RCA Laboratories Division
Princeton, New Jersey

ABSTRACT

To produce a simultaneous television image in color, three communication channels must be available. The first of these may be used to transmit the scene brightness, the second the degree of color saturation, and the third the hue or color. For compatibility the brightness is transmitted as amplitude modulation in the usual way. A subcarrier is introduced to carry the other two pieces of information as amplitude and phase modulations. The optimum loading of these two auxiliary communication channels is the major consideration of this paper.

In order to have a proper understanding of some of the color television transmission standards now being field tested, it is well to review the basic thinking that has gone into their makeup. We can now transmit in the same radio frequency channel a picture in color having practically the same detail as the present monochrome pictures. To do this we have had to make the most of our knowledge of communication engineering and a new field of study involving the subjective characteristics of normal vision.

The basic characteristics of the color television system now being field tested may be summed up as follows:

First: The signal is so arranged as to obtain compatibility with the present monochrome receivers.

Second: The signal has a minimum of redundancy.

Third: The signal make-up takes full advantage of the characteristics of vision. Under this there are three sub headings:

(a) Three color vision is provided and needed in large areas.

(b) Two color vision along a preferred locus is provided and needed for intermediate area detail.

(c) No color information is provided or needed in small detail.

In the next few minutes I will review how these basic characteristics have been integrated into a practical color television system.

To produce a television picture in color, three video signals must be produced in the receiver. The first representing the detail and color content of the red components of the original scene. The second representing the the same information relating to green, and the third the information relating to the blue components of the scene. These three signals are the result of an analysis of the original scene into red, green and blue components. The communication problem is that of transferring this information from the studio to the receiver with the greatest possible efficiency.

There are two parts to this; the first relates to the transmission of the information with a minimum of redundancy; while the second concerns itself with assuring that only information useful to the eye is presented to the communication channel for transmission. Each of the red, green and blue components of the scene to be transmitted contain brightness information as well as color information. To satisfy the requirement of compatibility, the brightness of the scene must be transmitted as amplitude modulation of the carrier in the normal way. This brightness signal is made up by adding the red, green, and blue signals in such proportions as to produce a signal representing the visual luminance of the scene. When transmitting the brightness components of the three color separations in this way, it would be redundant and a waste of communication channel to also transmit the same brightness information combined with the color information. To remove the brightness components from the color separations the brightness signal is subtracted from each of the red, green and blue color signals. This produces three signals representing red minus brightness, green minus brightness and blue minus brightness, or stated simply the three signals represent R-Y, G-Y and B-Y. If these three signals were transmitted to the receiver along with the brightness signal, signals corresponding to red, green and blue could be recovered by simply adding the brightness signal to each of the three. However, the transmission of four pieces of information, when only three are required, would again represent the transmission of redundant information. The signal representing the green separation can be obtained at the receiver by subtracting the sum of the red and blue signals from the brightness signal, or by taking the sum of R-Y and B-Y to obtain - (G-Y). Therefore it is necessary to transmit only signals representing R-Y and B-Y along with the brightness signal.

Bedford has pointed out the fact that the eye can not see color in small detail and that this property of vision may be used to advantage in a color television system. Here then is another opportunity to further reduce the amount of information transmitted. It has been determined experimentally that the color information may be limited in bandwidth to approximately one and one-half megacycles without the loss of information being detected by the eye. There now remains the problem of transmitting these two pieces of color information with a minimum of visibility along with the brightness signal. An advantageous arrangement for transmitting color information would be to use two carriers of the same frequency displaced in phase by 90 degrees. To generate such a signal the subcarrier voltage is divided into two parts. The first part is amplitude modulated with a signal representing R-Y and the second part is shifted in phase by 90 degrees and modulated with a signal representing B-Y. These components are then combined to form the transmitted signal. If in the receiver, a voltage having a reference phase is available, the two signals may be detected without crosstalk since they are at right angles to each other. By transmitting a small sample of the subcarrier at a fixed phase during horizontal blanking time the necessary synchronous voltage can be made readily available in the receiver for synchronous detection.

Now that we have all of the color information on a single carrier, the problem is to choose the carrier frequency so that it will produce a minimum of spurious signal effects in the brightness channel. One obvious way to determine this frequency is to simply look at the kinescope and change the frequency of the subcarrier until it is least visible. This occurs when the positive half cycles of the carrier, dots of light, interlace. This frequency is always some integer plus 1/2 times the frame frequency and can be an integer plus 1/2 times the line frequency since line frequency is an odd multiple of frame frequency.

Another way of looking at interlacing is to consider the frequency spectrum and interlace the subcarrier and its sidebands with the harmonics of frame and line frequencies. This approach also gives a subcarrier frequency of an integer plus 1/2 times the frame frequency. When approached from this direction it is called frequency interlace. However, frequency interlace always gives dot interlace and dot interlace always gives frequency interlace, that is, they are just different names for the same thing and for the same purpose, namely the reduction of interference between the brightness video signal and the color subcarrier.

With the color subcarrier frequency so

chosen as to have a minimum visibility, there are reasons for wishing to choose a frequency as high as may be passed by the receiver circuits and other reasons for wanting the subcarrier at a lower frequency. The fact that the receiving kinescope is non-linear contributes to the visibility of the subcarrier. Also insufficient persistence of vision and of the kinescope screen material, contribute to the visibility of the subcarrier. For these reasons it is desirable to make the residual dot structure as fine as possible by selecting a high subcarrier frequency. The reason for wishing to choose a a lower frequency will be considered in more detail.

A choice of a nominal 3.58 mc as a subcarrier frequency represents a balance between the various factors involved. With this carrier frequency there is .6 mc between the subcarrier and the end of the pass band assuming the receiver passes frequency up to 4.2 mc. This means that we can transmit the two signals representing color information in quadrature without crosstalk up to a frequency of .6 mc. Beyond this frequency the missing sideband causes the carrier phase to shift, introducing spurious signals into each color channel from the other. These spurious signals appear as incorrect color on the edges of objects. To prevent this crosstalk between the two color signals, one of the signals is limited to .6 mc. In this way we are transmitting the two color signals in quadrature on the subcarrier only in the frequency range where double sideband transmission is possible. Beyond this point a single color signal is transmitted. The choice of 3.58 as a subcarrier frequency is the result of balancing the desire for reduced dot structure against the desire to transmit color signals in the range of detail where the eye can make use of the information.

Reference has been made to the signals R-Y and B-Y in quadrature on the same carrier. Obviously the carrier can have only a single instantaneous amplitude and phase determined as the resultant of the two color signal vectors in quadrature. In other words the phase of the carrier actually represents color or hue information with the amplitude of the vector representing color saturation. Conversely the single vector can be separated into other pairs of vectors at right angles to each other. By selecting pairs of vectors other than those representing R-Y and B-Y, it is possible to select various loci on the color triangle along which all colors are reproduced for frequencies between .6 and 1.5 mc.

There is considerable information in the literature indicating that the eye becomes progressively color blind as the viewed object size is reduced. First there is three color vision, then two color vision, and finally only brightness vision with no color sensation. To determine the preferred locus for the region of two color vision a series of tests were made using a complete color television system. The signal source was a studio camera. The viewers used were both a tri-color tube and dichroic viewer using three kinescopes. The three kinescope viewer had a highlight brightness of approximately 100 ft. lamberts which made possible critical viewing.

A series of Munsell colors covering the color gamut were viewed each in turn superimposed on each of the others. The evaluation of the quality of the transition from one color to the other was on the basis of 10; one being excellent with 10 representing an unsatisfactory reproduction. The circuitry was so arranged that the two quadrature vector signals, one of which was limited in bandwidth to .6 mc, could be made up of varying proportions of the three primary color signals. In this way the color locus for those frequencies between .6 mc and 1.5 mc could be selected. As a result of the tests a preferred pair of vectors were found approximately 33 degrees from the B-Y and R-Y pair.

The summarized data comparing this preferred pair of vectors which have been termed I and **Q** with R-Y and B-Y is shown in the first slide (Fig. 1). Here the average rating of quality of color transition from each color to all other colors is shown. It is seen that improvement is obtained on all color transitions with the greatest improvement coming in the region of flesh tones and blues.

The color locus for the limited bandwidth of B-Y and the vector composed of color information such as to produce the Q vector displaced 33 degrees from B-Y, is shown in the next slide. It is noted that the B-Y locus is a straight line while the 33 degree locus approximates a part of an ellipse. The curvature of the 33 degree locus is caused by the addition of color components to produce the desired vector after they have passed through the non-linear circuits required for proper gamma reproduction.

With this selected choice of color vectors a grid of transitions may be plotted on the color triangle. (No. 3) The solid lines represent the paths taken by a color transition in the direction having a frequency response limit of 1.5 mc. The dotted lines represent the path taken by color transitions in the direction having a frequency response of .6 mc. Stated in another way a color transition from g to p would have a steepness of rise corresponding to a band pass of .6 mc, while a transition from bg to y would have a steepness of rise corresponding to a band pass of 1.5 mc. The transition from g to r would contain components of steepness passed by both the .6 mc and 1.5 mc circuits.

There is material in the literature which shows how the eye tends to have color perception only along a line locus for small color detail. For example Willmer and Wright, No. 4, were able to color match all the colors of the spectrum with only a blue and red light source when the object was sufficiently small. This data may be presented on the color triangle, in the form of a series of diverging lines which pass through the test colors and the color mixture which gives a match. These lines would be parallel to the locus of low frequency transitions if there was perfect correlation between the television tests and those of Willmer and Wright.

The data of Middleton and Holmes is shown on the next slide, No. 5. Here may be seen the reduction of color sensitivity of the eye toward the color locus as experimentally chosen for our television system. Hartridge has made similar tests on color acuity in small detail. His data is shown in the next slide (No. 6). Again the same locus is indicated.

A different approach to the selection of the wide band color locus may be based on the years of color photographic experience. Here the assumption is that in two color photography a color locus has been chosen as a result of experience to produce the most satisfactory result. This means that the chosen two colors are such as to approximate most closely the results obtainable with a three color process. In other words the difference between the two color processes is subjectively a minimum. In our television system this minimum difference is the information selected to be transmitted at the reduced bandwidth.

The locus of colors as reproduced by a two color process is shown in Fig. 7. Here again from an entirely different approach it is seen that the color reproduction is along the locus chosen for our television system.

From this brief review of some of the broad considerations involved in the proposed NTSC standards, we see that we have evolved a compatible system and at the same time reduced to a minimum the transmission of redundant information and the transmission of information that, due to characteristics of vision, the eye can not utilize. From the standpoint of ceiling performance it is believed that the broad concept of the proposed NTSC standards have in their make-up, the most efficient possible utilization of a six megacycle channel for the transmission and reception of television in color.

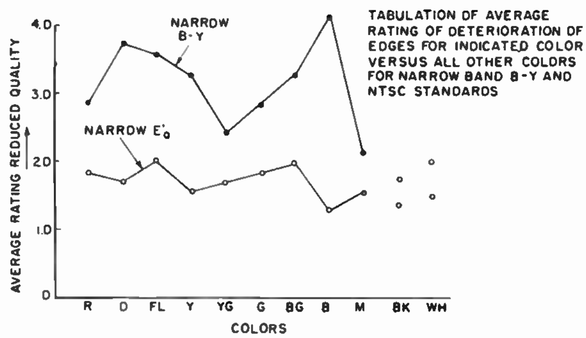


Fig. 1

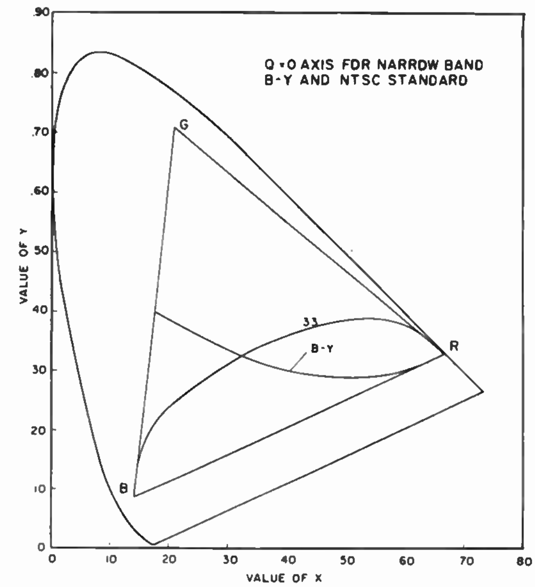


Fig. 2

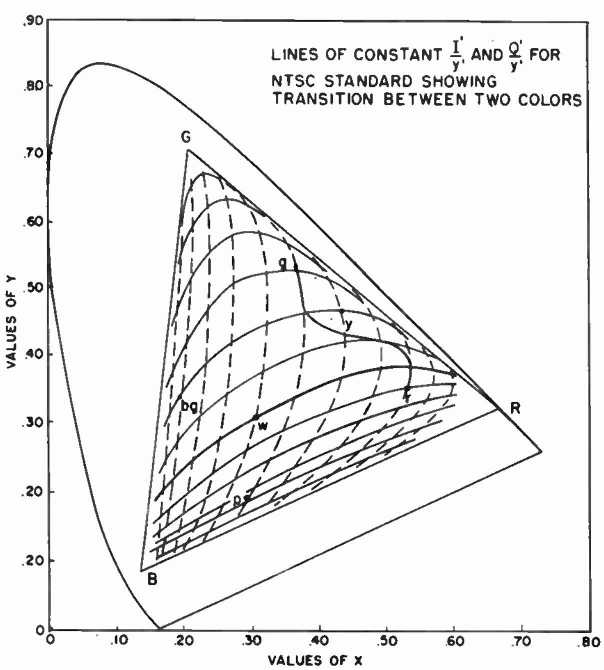


Fig. 3

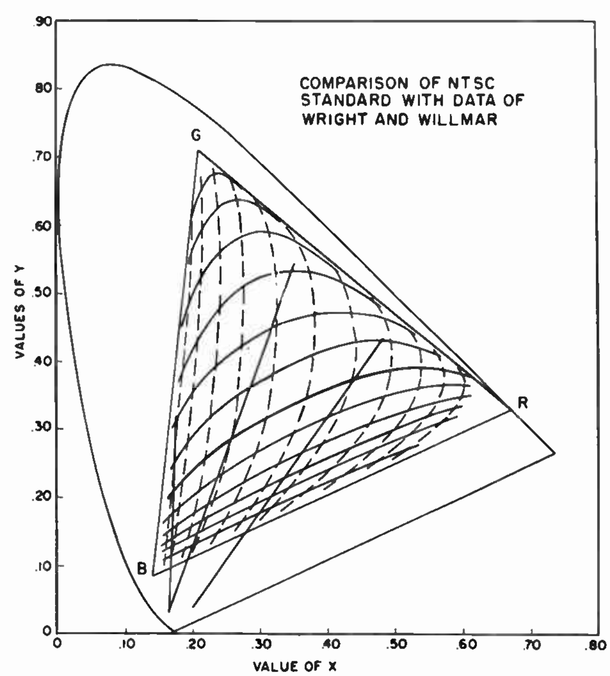


Fig. 4

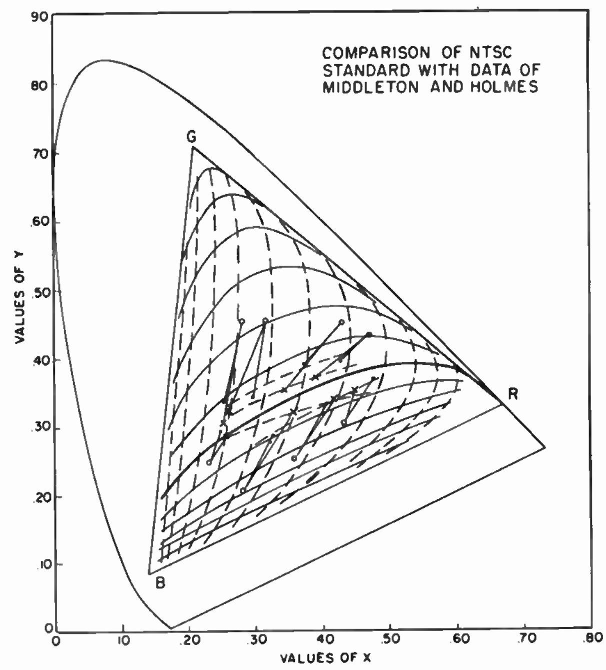


Fig. 5

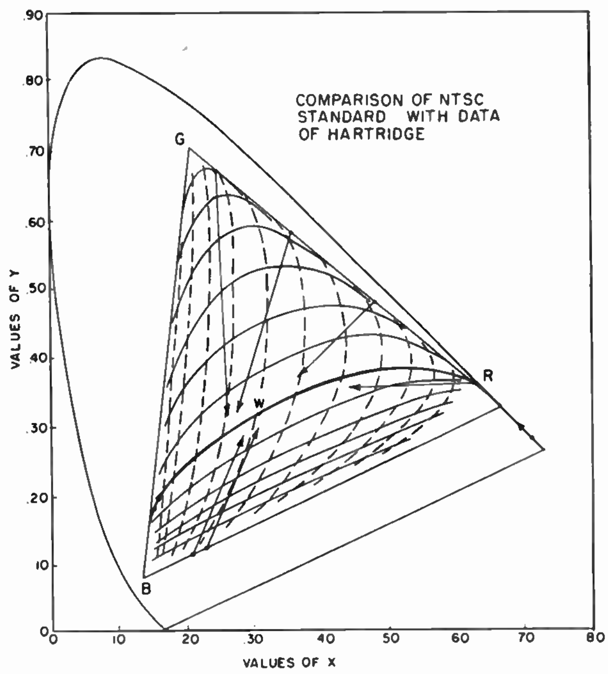


Fig. 6

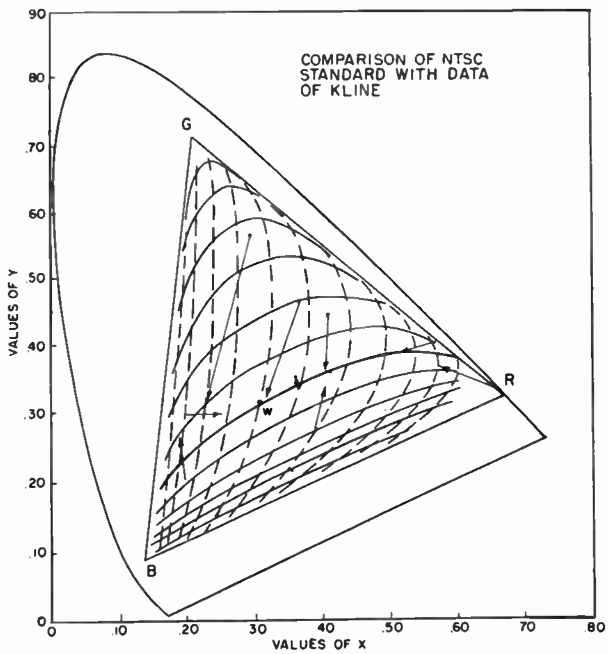


Fig. 7

A FLEXIBLE TV STUDIO INTERCOMMUNICATION SYSTEM

R. D. Chipp and R. F. Bigwood
Du Mont Television Network
New York, N. Y.

A television studio intercommunication system has many and varied requirements, which depend upon the complexity of the operation and the number of people involved in it. Typical of a simple program is a boxing or wrestling match utilizing perhaps two cameras and two fixed microphones. In this case, the simple carbon system that is built into television field equipment very adequately serves for intercommunication. Fig. 1 shows schematically such a system, which can provide two-way communication between all connected locations. The headset and microphone assemblies are plugged into jacks in the video control unit and the pickup camera. The intercom circuits are carried though as a part of the camera cable. The program director in most cases passes instructions to the cameramen through the switcher or technical director. If the audio man is not physically located adjacent to the program director, cues are given to him through this same circuit. If one-way communication is desired, it is obtained quite simply by providing a headset without a microphone at those locations that need to receive only.

Note the toggle switch shown at each video control unit. This switch isolates the control unit and its associated camera from the order circuit. It has been found to be extremely useful in the event that there is camera trouble while on the air or in the event there is need for camera repair or adjustment during rehearsal. By isolating the camera in this manner, the video control operator and the cameraman can talk to one another without disturbing those who must simultaneously use the regular order circuit for normal communication.

A more complex operation is one which might involve a studio operation combined with film facilities. Fig. 2 shows a floor plan of the present studio layout of the Du Mont station, WDTV, in Pittsburgh. Here the standard carbon circuit, with the addition of a three bus amplifier talk system has proved to be particularly useful. Fig. 3 shows this system in schematic form. Any one of the three busses can be selected manually by the technical director, the audio operator, and the video control operator. During rehearsals the program director, who sits next to the technical director, will use the studio address bus, while the technical director and the video control operator will use the amplified carbon system. The audio operator can cut into the carbon system if he wishes to talk to the mike boom operators. When cues are passed to the projection room, the technical director switches from the carbon system to the projection room bus. The projectionist has a separate talk back circuit to acknowledge instructions and report readiness to roll film, etc.

When on the air, the program director passes instructions or cues to the technical director sitting on his left. In turn, the technical director will select either the amplified carbon bus or the projection bus. The studio address bus is then used only for playback in the event that the program requires it. As is the case with the single system previously discussed, talk back from the studio to the control room is made available where needed.

As we progress to a still more complex operation, such as projected for the New York Du Mont Tele-Centre at 205 E. 67 Street, the need for increased flexibility is evident. Consider an operation of the sort indicated in Fig. 4. Here we have an elaborate studio program, portions of which may originate in the field and other portions of which may originate in a distantly located projection room. In the studio we may have an orchestra, back projectionist, numerous stage hands and stage electricians, and other special effects personnel. We have made a study of the flow of traffic required for intercommunication purposes. It is immediately apparent that the requirements during rehearsal and during the actual program are quite different. Fig. 5 shows in schematic form the flow of traffic during a rehearsal. The width of each line is approximately proportional to the volume of traffic, and the arrows indicate the direction of flow. Maximum volume is from the program director to the floor via the studio address system. Although the volume is light on other circuits, they are of great importance when needed. Fig. 6 shows the traffic flow during the program. Most instructions are now passed through the technical director, although there is still the same need for important supplementary circuits to other locations.

The problem is to design an intercom system that will meet these traffic requirements. Prime system considerations are basic simplicity with maximum flexibility. These requirements have been met by utilizing a type of cross bar arrangement, which is shown in generalized form in Fig. 7. Although this system design fits any type of studio layout, I will discuss its proposed use in the standard Du Mont split control booths at the New York Tele-Centre. Fig. 8 shows a floor plan of this standard control booth wherein the program director and assistants, plus the technical director, are located separately in the front section. Directly in back and raised two feet are the audio operator and the light control operator. These men can see all monitors over the heads of those in the director's booth. To the left is the video control section. This split booth arrangement has proved to be extremely satisfactory. Some of the reasons are:

(1) It separates functions - program direction, video operation, and audio control.

(2) It isolates audio monitoring which, if at a satisfactory level for the audio operator, often disturbs the director or video operators.

(3) The various instructions and sometimes prolonged discussions between the occupants of one room do not distract others.

(4) In case of video trouble, a cameraman and video operator may converse at low level without disturbance from audio monitoring or director's and switcher's orders.

Referring again to Fig. 7, there are five basic circuits that form a part of this system. What is called the "A" circuit carries non-interrupted program cue, the order circuit, and the microphone circuit for return talk. The "B" circuit, which appears on a separate set of receptacles in the studio, is similar to the "A" except that the program cue may be interrupted for instructions by the technical director, the program director, or the audio control man. A standard headset assembly may be plugged into either "A" or "B" circuit receptacles. The "C" circuit is the carbon system associated with the cameras. A "D" circuit, while not an individual system, is an extension of the audio operator's, technical director's, or director's source circuit, which appears in a patch field in the transmission bays at Master Control, and can be routed as required to external locations. Similarly, remote incoming order circuits can be patched into the intercom loudspeaker in the director's booth of any studio. The "E" circuit is the studio address system, which is normally used during rehearsal for instructions to the cast and floor personnel. The public address system for audience coverage and the playback system for film and turntable sound in the studio are separate from the intercom, and are not shown in Fig. 7.

The building blocks that make up this system are essentially sources and loads. The sources are microphones, pre-amplifiers, and keys or relays for selecting the desired load or loads. The loads are bridging amplifiers feeding either loudspeakers or the standard intercom headset assembly used by floor personnel. We have minimized the number of different components and have developed a number of standard units. For example, the pre-amplifiers are identical in construction to the program pre-amplifiers, with the exception that the input and output transformers have been modified to limit the frequency response and considerably lower the cost. The bridging amplifiers are also identical in construction to the program amplifiers with a similar transformer change.

The standard headset is an assembly, specially fabricated for this application, shown in sketch form in Fig. 9. The general construction follows lightweight aircraft practice and various headset and microphone combinations are interchangeable. The headbands will take a single or double headphone combination, and a microphone can be supported by either a right or left earphone. The control unit is intended to be supported at chest level, and contains transfer switches for connecting either earphone to either the program feed or the order circuit feed. The microphone for return talking may be disconnected by means of a switch located on this same chest plate. Volume controls are provided for each earphone. The order and program circuits are at a relatively high level, that is, loud enough, with the earphone controls at maximum volume, to overcome high ambient sound such as might exist near an orchestra.

The key switching arrangement consists of a standard panel which may be mounted in any location and can accommodate up to 15 keys. The director's panel may have as many as 15 load selector keys, whereas the audio and video operators may have only 4 or 5 keys. In the special case of the technical director, certain of these intercom keys are built into the video switching panel. For example, the order circuit keys for the film room are placed close to the appropriate video switching and projector control buttons. All keys are two way, and are so wired that one position continuously ties the source microphone to the selected load busses. The other position transfers the selector relay control to a foot pedal. In this latter case, only when the foot pedal is depressed will the source circuit be connected to the load. This arrangement is used, for example, to permit the technical director to route orders to the projection room only when pertinent, and exclude from the projection room the continuous chatter normally on the order circuit. In similar fashion, the audio operator can momentarily break into the boom operator's program circuit to issue special or emergency instructions.

Referring again to Fig. 7, let us review the functions of this intercom system during rehearsal and program. During rehearsal the director uses the studio address or "E" circuit for instructions to all floor personnel. This will be supplemented by the technical director's use of the "A" circuit to cameramen, floor manager, video booth, audio booth, etc. Talk back to the director's booth is frequently through the program microphones that are open, although those personnel having chest sets complete with microphones will automatically come up on a loudspeaker in the director's booth. In addition, during rehearsal the "C" circuit may be used by individual video control operators and cameramen to discuss camera adjustments and other similar matters without interfering with the progress of the rehearsal. At the same time, an assistant director may be discussing cues with a remote unit by means of the "D" circuit. In this way it is believed that lost time during rehearsals can be minimized.

When on the air the situation changes. Needless to say, the studio address system is disabled and any sound re-enforcing is done with the separate public address system. The majority of instructions channel through the technical director. The technical director sets keys in the up position to feed the studio order circuit and any other locations that require continuous cueing. These locations will usually use headsets. Keys in the down position permit cues to film control, transmission, etc. only when the foot pedal is depressed. These locations are usually set up on loudspeakers. The program director, the technical director, and the audio operator can also feed the "B" or interrupted program circuit. This circuit might be used by the program director to instruct the orchestra conductor to make a last minute music cut, or by the audio operator to pass emergency instructions to a boom man.

The audio operator has a single headset receiving the technical director's order circuit, and two loudspeakers for program monitoring and cue. He can talk to the director's booth, video booth, or those other locations provided on the available keys. These keys will normally be left in the unlocked position, and the press-to-talk foot pedal will be used.

The video operators will talk to cameramen on the carbon circuit and, as outlined before, can isolate a single camera in the event of trouble. In addition, they have a limited number of keys which function in the same manner as those in the audio control booth.

The generalized system shown in Fig. 7 can be assembled in a single standard rack. It can be either decreased or increased in size and scope, as required by the traffic load or by the size of the studio plant, without departing from the basic design philosophy, and without requiring modification of any existing installed equipment.

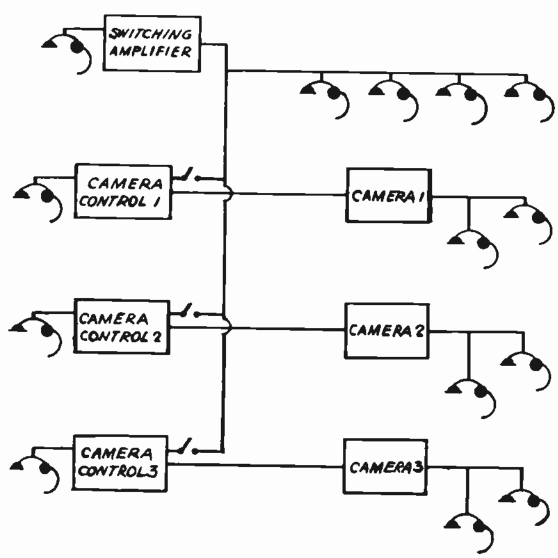


Fig. 1 - Standard camera chain intercom.

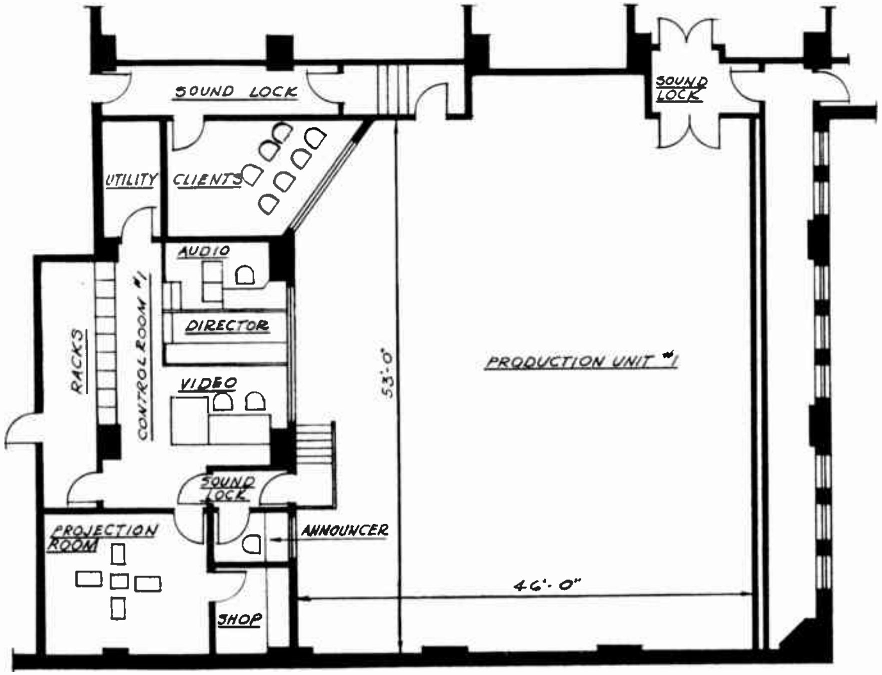


Fig. 2 - Existing studio layout, WDTV, Pittsburgh.

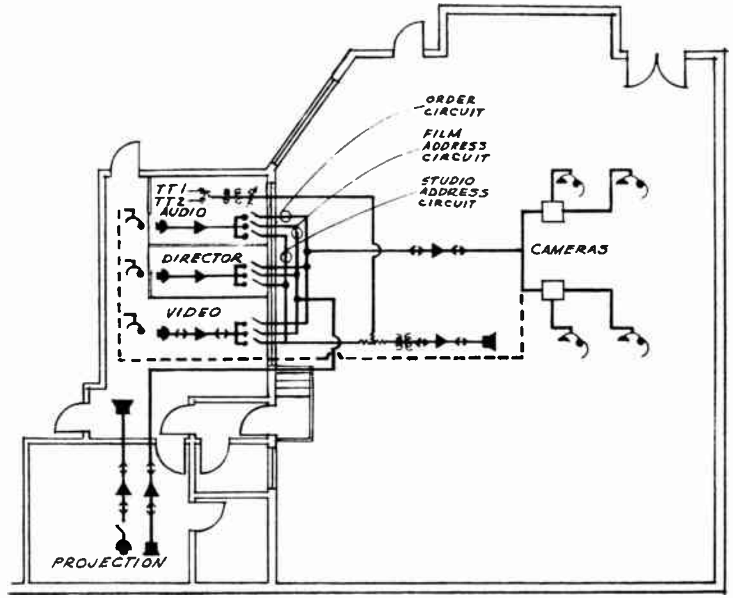


Fig. 3 - Studio intercom - type 2.

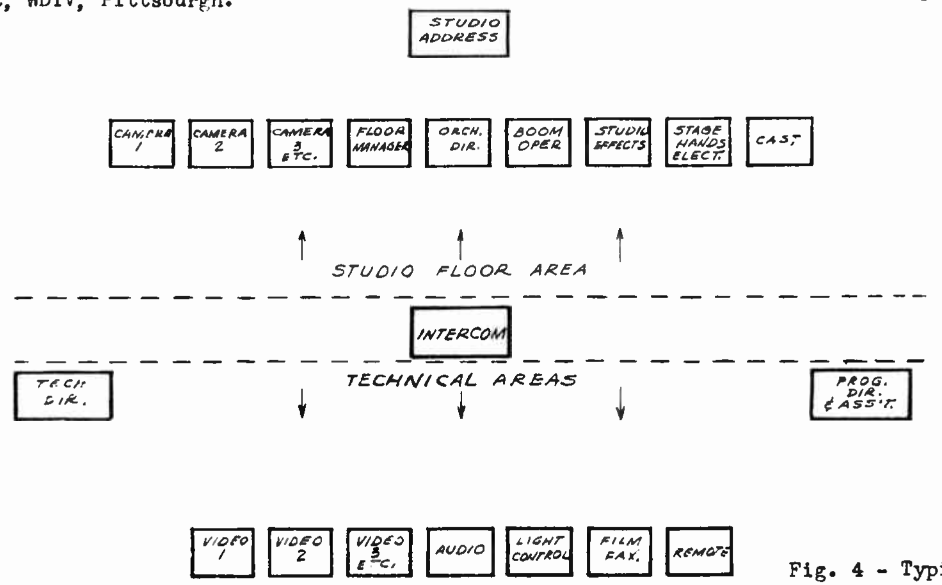


Fig. 4 - Typical coverage intercom system.

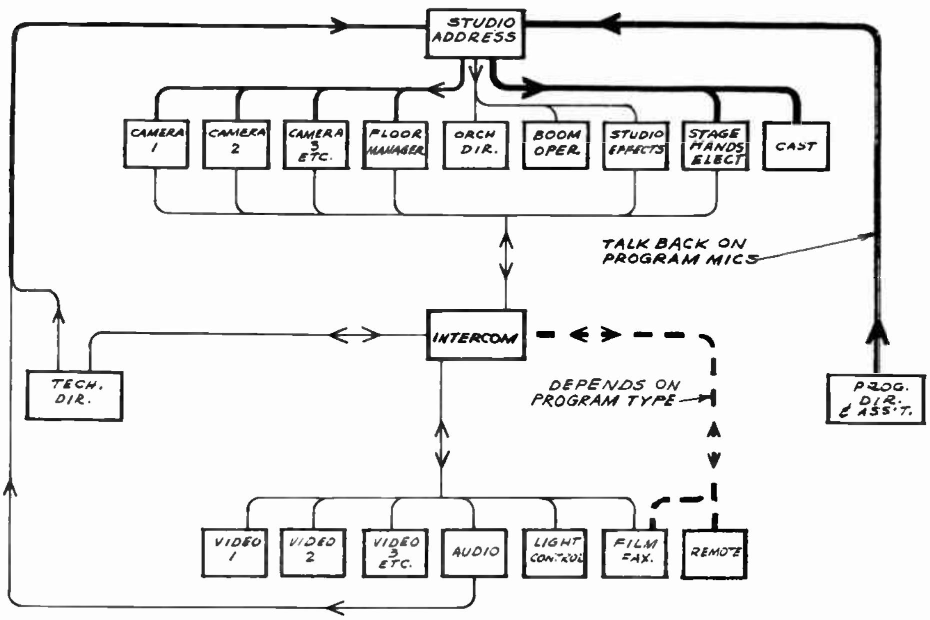


Fig. 5 - Rehearsal traffic flow intercom system.

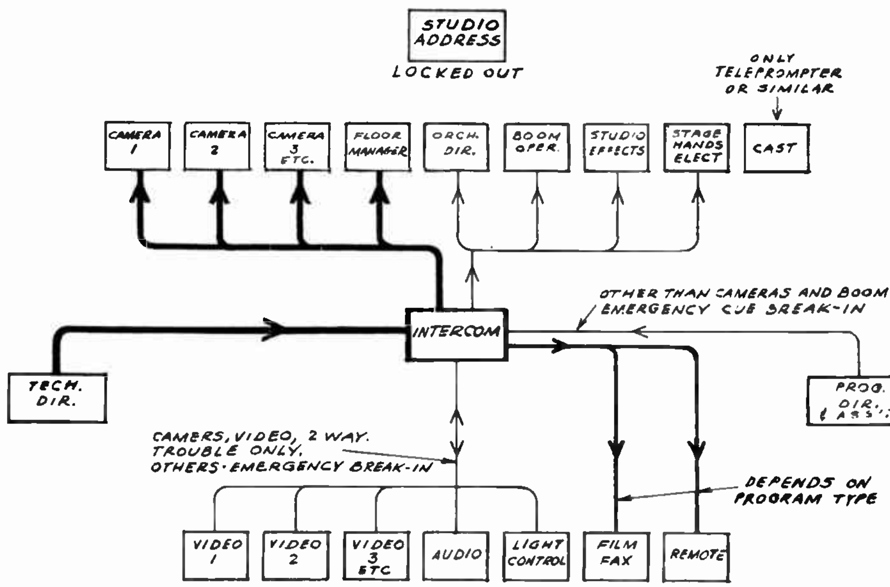


Fig. 6 - On air traffic flow intercom system.

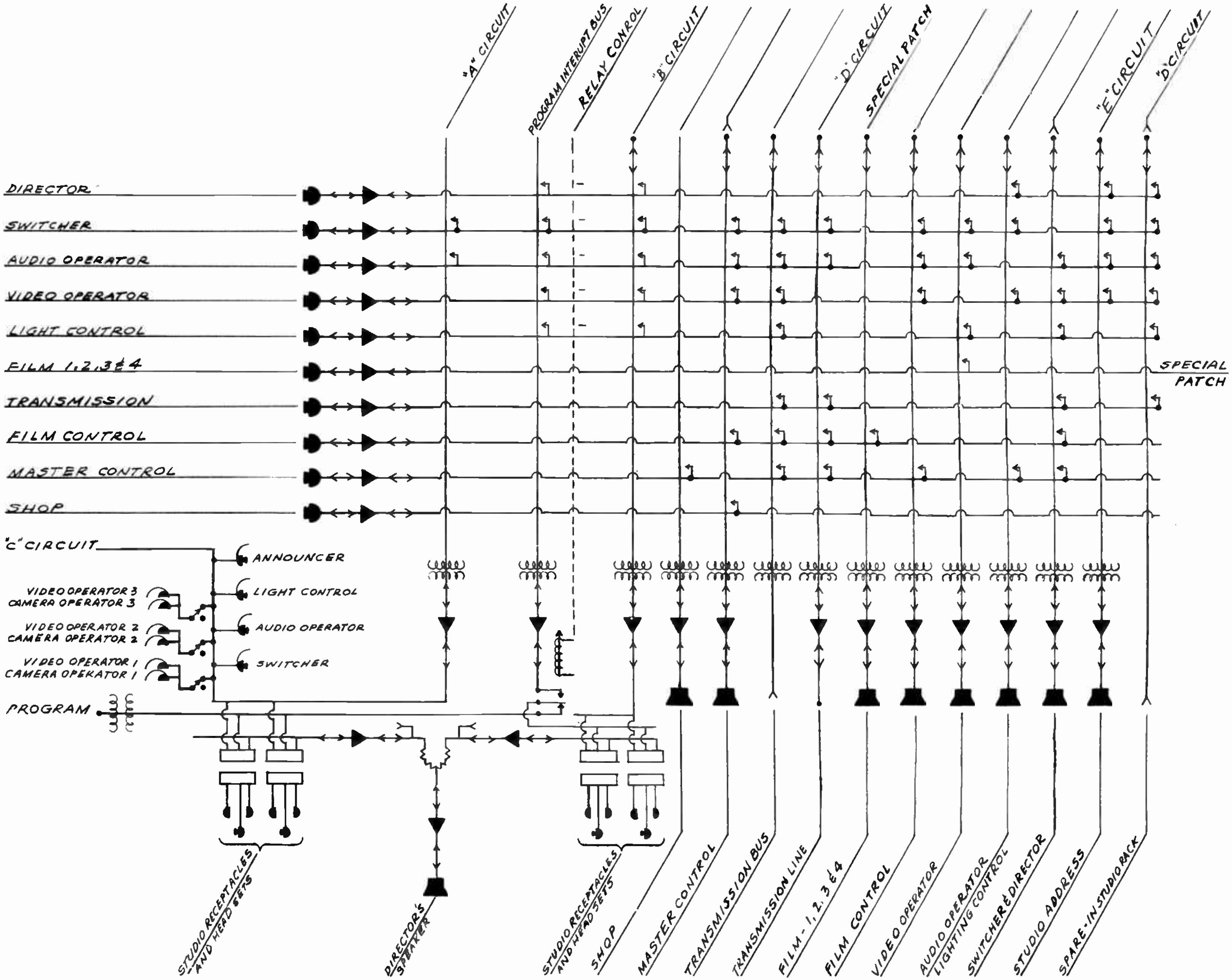


Fig. 7 - Generalized block diagram, intercom system.

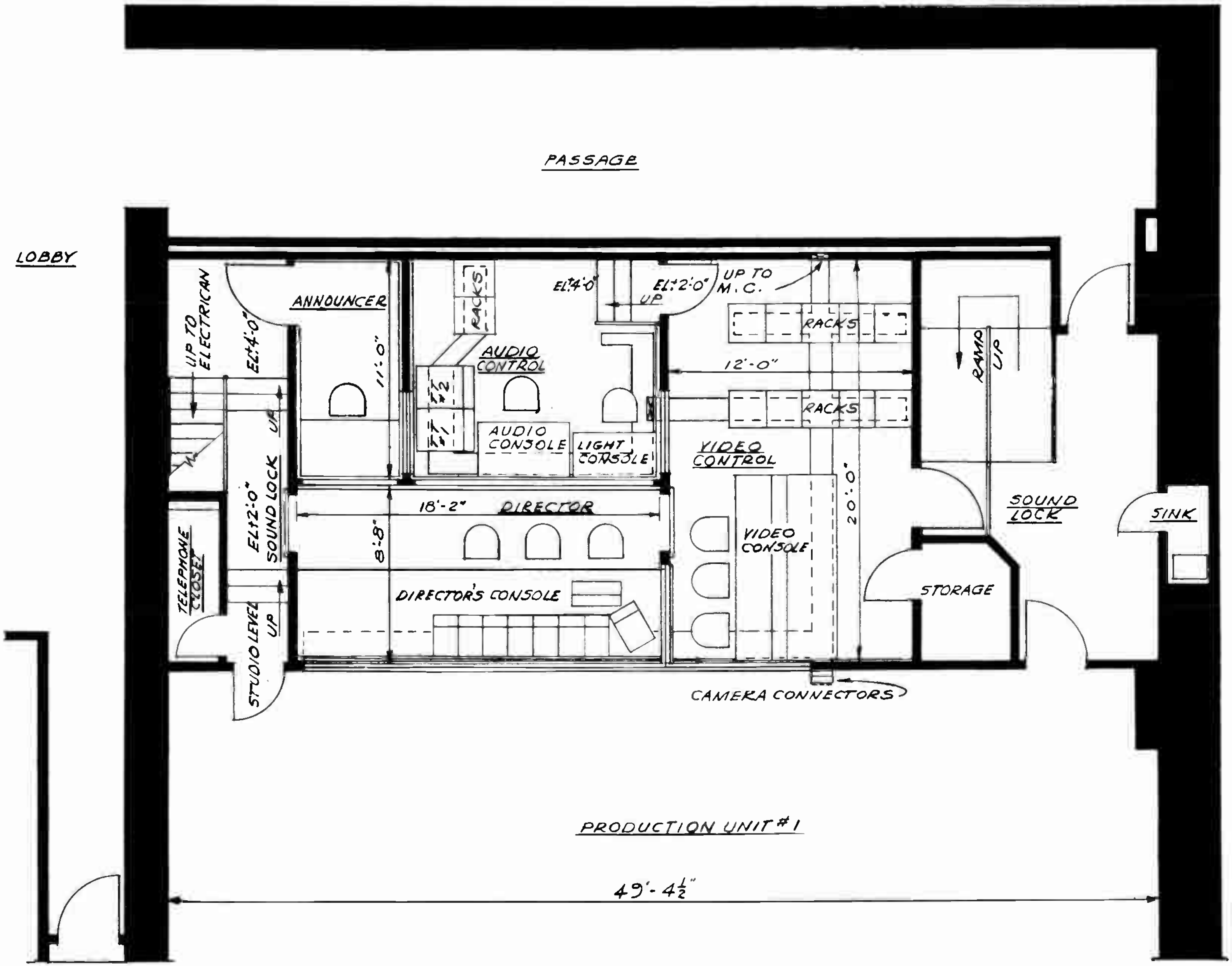


Fig. 8 - WABD Standard control room, 205 E. 67 St., N.Y.C.

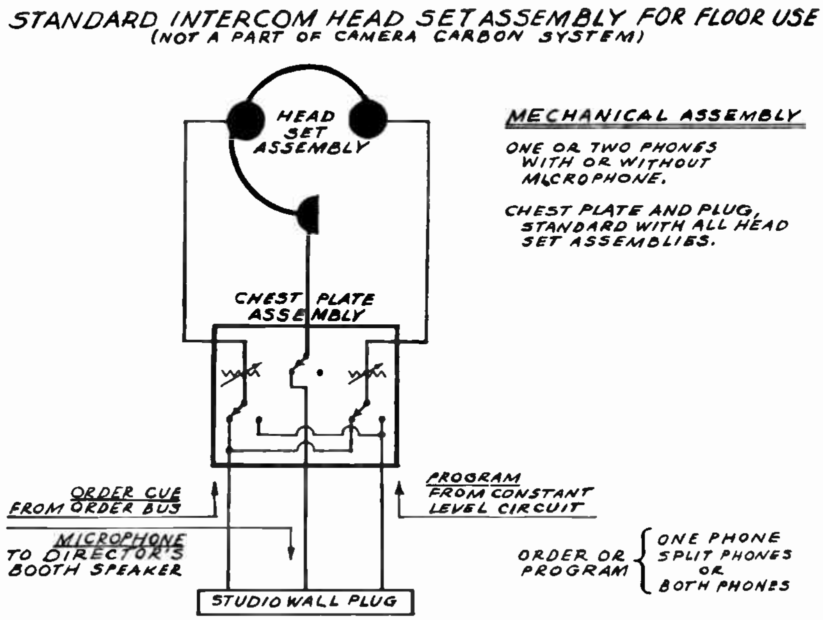


Fig. 9 - Standard head set assembly intercom system.

CBS TELEVISION'S HOLLYWOOD TELEVISION CITY:
VIDEO, AUDIO, AND COMMUNICATION FACILITIES

Richard S. O'Brien
Robert B. Monroe
Price E. Fish

Columbia Broadcasting System, Inc.
New York, N. Y.

Summary--Late in 1952 CBS inaugurated television service from its new Hollywood Television City headquarters. Located on a 25 acre site, Television City is constructed from a flexible and expansible master plan that permits ultimate expansion to 24 studio units.

The initial construction phase now completed provides two audience and two non-audience studios, each exceeding one-quarter acre in area, together with the necessary technical, production, scenery construction, and office facilities to completely support CBS-Hollywood television operations.

This paper describes the philosophy underlying the design of the video, audio, and communication facilities for this project. Emphasis is placed on description of methods and features that are new or novel.

Introduction

CBS Television City, in Hollywood, California, is a fully self-contained center expressly designed for the efficient production of television programs. When the 25 acre site is fully developed, the plant will include 24 studio units each providing over one-quarter acre of unobstructed working space. Four of these studio units, together with a generous amount of auxiliary service space, are included in the initial unit in which operations began late in 1952.

Figure 1 shows the modern, functional appearance of the initial unit of Television City. The four studios are housed in the building extending to the right, while scenery shops and production offices are in the adjoining three-story, glass-walled service building at the left. The buildings cover $3\frac{1}{2}$ acres of ground and contain over $8\frac{1}{2}$ acres of concrete floor, emphasizing the extent of even this first phase of the project.

As can be seen in Figure 2, the initial construction, indicated by the shaded area, is but a small fraction of the ultimate plant. Additional studios may be provided by extending the two initial wings or by adding new wings on either side of the service building. This architectural design allows great flexibility in the sequence and exact manner of plant expansion. The service building may be expanded in height as well as length to keep pace with studio additions. If required, office buildings may be built to a thirteen-story height along the boundaries of the property.

The interior arrangement of the initial buildings is indicated in Figure 3. The two studio wings in the studio building are separated by a forty foot wide passageway which allows free flow of traffic from the scenery shops and material receiving area to the studios. Three rehearsal halls, each measuring seventy feet square, are located in the space above this passage and other service areas adjacent to each studio. Two of the studios are arranged for non-audience performances but the two studios in the rear wing are fitted to accommodate studio audiences.

The spacious 110 foot wide stage area is shown in Figure 4 which is a view from the rear of one of the audience studios. The height is 28 feet from stage to overhead structure. The control room, in which the studio video and audio facilities are located, is at the left. In the foreground is the console from which the electronically controlled studio lighting system is operated. This system controls over 300 kilowatts of incandescent lighting in sixty 5 kilowatt circuits and can be preset to remember any five desired combinations of dimmer settings for all 60 circuits. The lighting combination for a particular program sequence is then selected by operation of a small push button selector switch. Thyratron tube banks associated with this unit are located in the basement area.

The basement level, shown in Figure 5, is also utilized for dressing rooms, scenery and properties storage, maintenance shops, and the central technical area. In the central technical area are grouped those facilities which serve all studios equally. Thus they are placed, logically, near the geographical center of the ultimate plant.

System Considerations

Figure 6 illustrates in a greatly simplified manner the functional relationship between the groups of facilities in the central technical area and a typical studio. The centralized facilities groups include master control, telecine, program control rooms, television recording, telephone company terminal equipment, and two-way microwave link with the local station transmitter site, this latter equipment being located on the service building roof level.

Master control contains both the first and the last elements in the video system, namely, timing pulse generation and distribution, and trunk switching of final program signals. The latter function is shared by the audio system, but timing pulse distribution whereby all film and studio cameras are scanned in synchronism is, of course, peculiar to video.

From the "telecine" (a contraction of television-cinema) room, video and audio signals obtained from film, or other recording media, are transmitted to studios for blending with live portions of programs. For programs which consist entirely of film material, or in which all live portions originate off the premises, the signals may be routed instead to a program control room. These program control rooms are equipped with switching and mixing facilities identical to those in a studio control room but they have no cameras or studio area associated with them.

Program signals originating at remote locations may reach the master control area by means of telephone company circuits or via micro-wave relay. From master control, such remote signals may be routed to either a program control room or studio control room for integration with other program material.

Of these groups of facilities, the initial installation includes the central pulse distribution section of master control, the telecine facilities, and the four initial studios.

An important feature of the Television City video system is the complete decentralization of the studio switching facilities to the respective control rooms. This is contrasted with earlier systems in which all video switching was done by relays located in master control. A method of compensating for signal transmission delay time has been evolved to make this system workable and is described below.

Decentralization of equipment to the control room follows well established audio design practice. This practice reduces the amount of centrally located space which must be allocated in advance for technical facilities and removes all restrictions on the choice of video switching equipment for the studios. This contributed greatly to the flexibility with which additional studio units may be added or existing units modified.

Studio Control Room

From the control room position at the rear of the audience area, Figure 7, the director can readily keep in mind the viewpoints both of the studio audience and, by means of picture monitors, that of the much larger network audience.

The excellent visibility of the stage is shown in Figure 8, a photograph taken from the directors position in the control room during a program rehearsal. The monitors below the window are associated with individual cameras except for the center unit which is a studio output line monitor. The large screen overhead monitors include one cue and two preview monitors.

In a plan view, Figure 9, the control room personnel are seen to be grouped about the program director who sits at the left end of the directors console. Program assistants are at his right, key technical personnel at his left. The technical director at his immediate left operates the video switching console and is responsible for all technical phases of the performance. The audio operating position is far enough to the left that a glass partition may be erected around this section to give a measure of acoustic isolation should this be desired. The camera control operators are placed on a lower floor level and sight lines carefully controlled so that all persons concerned with picture quality may have an unobstructed view of the high-quality camera-control monitors. Space is provided on the camera control level for additional control equipment which may be required, e.g., color television. The announcer is located where he may obtain visual cue signals from either the audio operator or the program director.

Auxiliary video, audio, and communications equipment is mounted in air-cooled equipment racks which form the rear boundary of the control room operating area. Cabinets built in to the rear wall of the control room contain terminal blocks on which all interconnecting runs from the studio to the central technical area are terminated.

Telecine

The telecine portion of the central technical area is shown in Figure 10. Telecine is divided into a projection room and a control room. The initial installation includes three 35mm, three 16mm, and two still projectors, space being provided for double this complement in both the projection and control room. A photograph of the telecine projection room is shown in Figure 11.

A feature of the projection arrangement is the provision of a separate film camera for each projector. In a plant the potential size of Television City, the peak film requirements can be met with a smaller total equipment investment, using a camera for each projector, than with the arrangement widely used in smaller installations, wherein one camera serves several projectors. In the latter, the use of one projector automatically ties up all of the others. With separate cameras, a much greater degree of flexibility in equipment assignment is achieved.

Assignment of the various projector-camera groups to studios is accomplished by means of the "patch cross" unit, Figure 12. This unit is located in the telecine control room along with the film camera controls and video distribution facilities. To achieve equivalent flexibility with a more conventional switching system would require a system with 70,000 or more switch or relay contacts. The unusual "cross" form of the upper section was dictated by the requirement of providing 128 receptacles in the most compact space.

Telecine operation is preparational in nature, the actual control of film material in a program being done, properly, in the studio or program control room to which a film channel has been connected through the patch cross unit.

Master Control

In master control, Figure 13, across the hall from telecine, space has been provided for the facilities required to select between forty or more input channels and to feed video and audio output signals to as many as twelve outgoing channels. The necessary monitoring and control facilities, patching and adjustment racks, and terminal equipment racks are to be arranged as indicated.

However, these facilities are to be installed at a later date, a temporary equipment installation serving to interconnect the four initial studios with an existing CBS Television master control in Hollywood. The interconnection is made via permanently installed telephone company transmission facilities. The central pulse distribution system which is described in detail below is installed in its permanent location in a corner of master control. Here the master control operator will have direct surveillance over this very important central facility.

A view of the central terminal frames adjacent to master control, Figure 14, provides an example of the provisions for expansion designed into Television City. Here all interconnecting cable circuits between studios and various parts of the central technical area appear on terminal blocks. Even coaxial cables are spliced on newly-designed solder-type terminal blocks. Easily accessible overhead wiring is used throughout the central technical area to further enhance flexibility for expansion.

Video Facilities

Pulse Generation and Distribution

Video signals originate with the driving and synchronizing pulse generators, two of which are permanently installed in the master control area. They are located together with sync distribution facilities adjacent to the central terminal frames and master control terminal facilities.

The pulse distribution block diagram, Figure 15, shows the sync generator selector and the allocation of pulses to the several buses. Pulse distribution amplifiers provide multiple feeds to each of the studios and other operating areas. Individual control of pulse levels is available on all channels. Distribution units are strategically placed at each stage of system expansion to provide access at critical junction points for the introduction of an emergency timing pulse feed, if for any reason continuity through distribution amplifiers should fail. As shown in the detail on Figure 15, a distribution unit is a two channel switching device which permits the selection of the regular or a spare input. A waveform monitor may be readily connected into the monitor output at any time. In addition, a visual monitoring indicator is provided for the pulse signals. This indicator, shown in Fig. 16, consists of a group of neon lamps driven by isolation amplifiers, each associated with a particular timing pulse. The neon lamps glow steadily in the presence of a normal signal, blink slowly in the absence of the signal, and, of course, extinguish in the event of power supply failure.

The distribution of driving pulses and synchronizing signals throughout a large plant such as Television City requires that attention be given to the time delay inherent in long transmission paths. For instance, the horizontal rate pulse signals to all studios are intentionally delayed in order that pulses reaching the studios directly from master control will arrive at the same time as pulses arriving by way of telecine. This delay corresponds to twice the distance between master control and telecine. Pulses from the sync generators thus arrive at a given studio at the same time, regardless of which path is taken. The output of film cameras may then be freely mixed and dissolved with studio cameras, without time displacement of the blanking and synchronizing pulses.

Similarly, in order that the video output signals from two or more studios may be freely mixed and dissolved at master control, the total traverse time to studio and return must be the same for all studios. This is accomplished by the introduction of delay sections in the video lines from each studio to master control, with the exception of the farthest studio. These lines are adjusted to obtain time delays equivalent to twice the difference between (a) the distance from master control to a given studio, and (b) the distance from master control to the farthest studio.

Delay adjustment at Television City resulted in total video transmission paths of 1100 feet from telecine via studios and return to master control. Transmission paths between the present studios and master control varied between 512 and 823 feet.

The upper portion of Figure 17 shows a blanking signal from a telecine film camera routed through one studio to master control. The lower oscillogram shows the same signal routed to all four studios and received simultaneously at master control.

Telecine Video Facilities

Figure 18 shows a typical 16mm film projector, its camera, and associated equipment rack. From his normal operating position, the operator may view the picture output of the camera or studio cue lines, monitor the audio output of the projector or studio, and communicate with the camera control operator as well as the studio with which he is working.

The projector may be assigned, by means of the patch-cross unit, to two separate studios. Remote control of the projector, however, is transferred from one studio to the other at the discretion of the operator.

The video facilities associated with a telecine projector, Figure 19, consist essentially of the film camera and associated control unit, picture and waveform monitors, distribution of timing pulses to the cameras, and distribution of camera output signals to the studios.

Transmission losses of the video lines to the studios are offset by equalization. Camera outputs are pre-equalized with special attenuation equalizers, expressly designed to compensate video cable losses. To standardize the amplitude and video response of camera output signals, as they appear at the studios, the equalization employed corresponds to the median length of cable runs between telecine and studios.

Video signals from each camera chain appear on two patch plugs at the patch-cross unit. From this point the signals are connected to the studios where the film material is integrated into the live portions of the program. The patch cross operation is described in more detail below.

Studio Video Facilities

Figure 20 is a view of a typical studio control room looking from the audio control area. It shows the video switching console in the center, and immediately beyond, the directors console. To the right, at the rear of the operating area, the power and video racks may be seen. The camera control console is just visible under the studio observation window.

The camera control console, Figure 21, accommodates four camera control units with their associated master monitors and a studio line monitor. Unit construction of the console will permit the integration of additional sections if required.

Five video and two audio racks form the rear boundary of the upper operating area of the control room. The video racks contain the monitor switching system, distribution amplifiers, jack fields, composite and non-composite relay switching systems, and mixing amplifiers. Ample rack space has been provided for future expansion of facilities.

The studio video switching system, Figure 22, includes three component sections for the relay switching of non-composite, composite, and monitor signals. The non-composite switching system has input positions for as many as eleven studio and film cameras. These camera signals may be selected directly, or as many as three may be superimposed on each other, through the use of two dual faders. The output of this non-composite section, after the addition of sync, appears as one of four program inputs to the composite switching section. Composite signals are selected by direct switching.

In addition to two preview channels serving individual preview picture monitors, three general monitor buses at the output of the monitor switching system serve lighting, utility, floor, sound effects, and audience picture monitors.

The video switching console, Figure 23, provides push-button control of the relay switching activities just outlined. Note the division of the panel into non-composite, composite, monitor, and projector control areas. Preview channels may select any of the twelve non-composite and four composite signals. The effects fader selector rows may choose any of four telecine lines and up to seven studio cameras. The line fader may select the same signals and, in addition, the output of the effects fader, thus providing considerable flexibility in the synthesis of multiple camera shots. The effects fader, or an additional fader in the spare position provided, may readily be connected to control special matting or electronic wiping amplifiers. Composite program selection includes the studio output or any of three remote signals. Three monitor buses may select either of the two preview channels, studio line, cue, or a spare input. Up to four projectors may, by remote control, be started, stopped, or "still-framed" (i.e., the continuous projection of a single frame of film).

In the studio installations, the video response-frequency characteristics are uniform up to 10 Mc. Figure 24 shows the over-all video response of a studio from a typical camera input and from a remote line input.

Transmission of such wide-band studio output signals to master control requires the use of attenuation equalizers and associated amplifiers to compensate for transmission losses. The effectiveness of this arrangement is illustrated in Figure 25 which shows the video response at the master control terminal facilities to a sweep signal introduced at a remote line input to a studio switching system.

Figure 26 is a view in an audience studio showing the seating and stage areas. Attention is called to the pantograph-mounted picture monitors serving the studio audience. These monitors may be easily positioned to provide the audience with the most advantageous view of both stage activity and the transmitted picture program material.

Audio and Communication Facilities

Figure 27 shows the general appearance of a Television City studio control room. At the extreme left are two transcription turntables with an associated control unit. To the right of the turntables is the audio control console followed by a utility desk, the video switching console, and the directors console.

The audio console is completely self contained and follows well known CBS audio design practices. It includes such design features as single plugs and jacks because of their smaller size and superior performance, the use of only two amplifier types, both plug-in, and two audio tube types to simplify maintenance, and standardization on a circuit impedance of 150 ohms to assist in achieving uniform response-frequency characteristics. As a matter of fact, these features have proven so successful they have been employed in the audio facilities throughout the entire Television City plant.

The turntable control unit is located to the left of the audio control operator and in this position the controls are within his reach for normal operation. However, the physical separation makes it possible to employ a second man to operate the turntables should this prove desirable.

Figure 28 shows all equipment units associated with the control room audio installation - console, turntables, and two racks of equipment. The rack on the left contains equipment associated with the studio communication facilities. The rack on the right contains all equipment required for audience sound reinforcement and audience reaction microphones. This rack is employed only in audience type studios. The division of the audio facilities into these four basic units results in an extremely flexible audio system, in which the basic units can be used alone, or in combination, to equip both radio and television studios of various types.

Figure 29 shows in simplified block diagram form the scope of the audio facilities contained within the console. An eight position mixer accommodates seven studio microphones and a circuit from a sound-effects console to be described later. A four position microphone sub-mixer, which permits a portion of a program to be pre-balanced, may be routed through any one of the regular microphone mixing channels by means of a single patchcord connection. Other mixing channels are provided for an announcers microphone, A and B positions for film sound and remotes, and a turntable channel that accommodates the two transcription turntables.

The output of each microphone passes through a preliminary amplifier and bridged-T mixer control and is then divided by means of a multiple-output resistance isolation network which features a high degree of isolation between each of its outputs. One of the network outputs continues on to the program channel. The other outputs are associated with reverberation and sound reinforcement channels.

The microphone and sound effects channels are combined in a suitable differential network, the combined output being controlled by a single mixer key switch. Other mixer key switches control the announce channel, the A channel, B channel, and the turntable channel.

Two program channels are provided. Program material from the mixer key switches can be connected to either channel one or channel two. Either channel can, in turn, be connected to feed the outgoing program line.

Two regular monitoring channels and loudspeakers are provided. One serves the audio control operator, the other serves those at the directors console.

Figure 30 supplements Figure 29 and shows the remainder of the control room audio facilities. Reverberation effects are sometimes required in the production of television programs. Reverberation is added to a microphone channel in the following manner. The reverberation output from the resistance isolation network in each microphone channel is brought to two reverberation selector switches as shown in this slide. This permits either one or two microphone channels to be selected for reverberation. The selected channels are transmitted to a reverberation chamber or other reverberation device. The output of the reverberation device returns to the studio where it is introduced into the program circuit by means of the A or B mixer channels.

The second output from the network supplies a program sample from each mixer channel to sound reinforcement selector switches. Any desired mixer channel can be reinforced on the studio loudspeakers merely by closing the desired key switch. All adjustments made in gaining the regular program channel are equally effective in controlling the gain of the sound reinforcement channel.

A six position mixer is provided for audience reaction microphones. The output of this audience mixer is patched into any one of the regular microphone channels.

Sound Effects Facilities

The sound effects station in the Television City studios is located on a balcony overlooking the studio as shown in Figure 31. It is equipped with a CBS sound effects console that provides all facilities normally required to handle the sound effects requirements of the most complex television programs. The sound effects console is equipped with three variable-speed turntables and four record reproducing arms. A six position mixer accommodates the four reproducer arms, as well as two utility mixing positions for sound effects microphones or special sound effect devices.

The sound effects console transmits the complete sound effects portion of the program to the control room audio console at standard line level. Another higher level output transmits the sound effects to studio loudspeakers.

Attention is called to the studio floor monitor immediately to the right of the console. This is a very useful equipment assembly combining in a single unit a picture monitor, loudspeaker, clock, and telephone station.

Studio Communication System

Most visitors to well designed television studios are impressed by the excellent coordination existing between the large group of people involved in the production of a program. This coordination results from the use of an extensive system of communication whereby each person involved is kept informed, and is cued and directed in his activities.

A studio communications system must distribute intelligence to all engaged in the production of the program. This intelligence takes three forms: first, spoken words, such as direct instructions from the director; second, audio program material from which many cues are obtained, for example, the orchestra leader usually gets cues for his music from the dialogue of the program; and third, picture program material which provides information to almost all concerned with the program, for example, the sound effects operator synchronizes his sounds with studio action by means of a picture monitor.

Aural communication is accomplished by five basic systems: a private telephone network known as the interphone system, a headphone cueing system, a low-frequency induction field radio channel, a loudspeaker talkback and call system, and a two-way intercom with telecine.

The interphone system is shown in Figure 32. A private telephone circuit is provided between each studio camera and its associated camera control. When a switch at the camera control position is closed, this interphone circuit is connected to a conference bus and the cameraman and camera control operator may then communicate with the program director, technical director, or any other camera or station also connected to the conference bus.

An amplifier reinforced feed from the directors desk microphone is introduced into the conference bus at a level slightly higher than other interphone conversations. The director can therefore be heard at all interphone stations even though he is not speaking into the microphone of an interphone headset. The reinforcement gives him an appropriate degree of priority over other interphone conversations at all times.

A second interphone system serves the studio lighting personnel. There are stations in the studio for the lighting director, lighting control console operator, and lighting patch panel operator. A second station is provided for the lighting director in the control room should program conditions require his presence there. Color coded calling lights are provided at each lighting interphone station.

Another interphone circuit connects the audio control man and the sound effects operator.

The interphone system constitutes only one portion of the studio communication facilities. It is necessary to provide communication circuits to others not served by the interphone system, such as the microphone boom operators, the announcer, and orchestra leader. This is accomplished by means of a headphone cueing system that employs conventional headsets which have been split to permit the reproduction of different information in each ear.

Headphone cue reproduces the voices of the program and technical director in one earphone and audio program material in the other. A second type of headphone cue specifically for the microphone boom operators permits the audio control man to break into the audio program headphone and talk directly to the boom operators at any time.

Communication to performers as well as others in the studio is accomplished during studio rehearsals by means of talkback loudspeakers in a manner similar to that employed in radio studios.

Figure 33 shows the circuits employed in the headphone cue, talkback, and other portions of the communication system. The output of the three directors microphones are amplified, limited, and pass through a volume control to the headphone bus. The volume limiter is a simple non-electron tube device employing thermistors which maintains a uniform output level for a wide range of microphone input levels. A similar channel is provided for the microphone on the audio console, and a third channel for audio program material.

Bridged circuits from each of the three headphone buses pass through volume controls, relays, and an amplifier to the studio and announce booth loudspeakers.

Because of the physical separation of telecine and studio, a two-way intercom system has been provided between the two points. The intercom employs the same line for transmitting and receiving. The switching arrangement normally aligns the channel for reception. When the switch is operated, the components are realigned for transmission.

An emergency circuit to telecine has been provided for use in the event of a failure of the regular channel. The same amplifier is used for a dressing room call system. A loudspeaker is installed in each of the twenty-four dressing rooms as well as in the associated corridors and in other places where performers may congregate.

Some persons involved in the television program production, notably the studio floor manager, must be free to move to any part of the studio without the encumbrance of a cable, yet must remain within the range of the voice of the program director at all times. An induction-field transmission channel is provided for these persons.

The receiver is a compact unit carried over the shoulder as shown in Figure 34. The induction-field transmitter employs amplitude modulation and operates in the low frequency range between 100 and 200 kilocycles.

<u>Telecine Audio Facilities</u>

A typical telecine projector and camera grouping is shown in Figure 18. Each such grouping of camera and projector has an identical rack of equipment associated with it. The rack houses the audio and communication facilities along with a picture monitor. Controls are grouped on the panel immediately below the picture monitor.

Other controls, as well as an intercom station, are located on the small control panel below the picture monitor at the camera control station associated with each film camera. This can be seen in Figure 12.

Figure 35 shows the circuits associated with a typical projector. The audio output of the projector is amplified and supplies a line level audio signal to each of the two associated patchcross plugs. Monitoring circuits are provided for loudspeakers at the projector and camera control.

An intercom is provided between the projector and its associated camera control. Circuits from this intercom extend to the patchcross plug. When the patchcross projector-to-studio connection is made, all audio and communication circuits, as well as video and projector remote control circuits, between telecine and studio are established.

Figure 36 shows a closer view of the patchcross unit. Two patch plugs on the lower section of the unit are associated with each projector. The receptacles on the upper section are studio circuits, four such channels being provided to each studio and six to each program control room.

Connection between any projector and any studio is made by patching the projector plug into the desired studio receptacle. The second projector patch plug makes it possible to establish a connection with a second studio on a preset basis while still working with the first studio. The transfer from one studio to the other can then be made conveniently and quickly by means of transfer switches on the projector rack.

As with all facilities in Television City, adequate expansion space has been provided on the patchcross unit for circuits associated with future projectors and future studios.

<u>Acknowledgment</u>

As is the case with all projects of this scope, a great many people were involved in the design and engineering of Television City. We regret that it is not practical to mention each person, and his contributions; however, the contributions of all are sincerely appreciated. The project was under the over-all direction of William B. Lodge, Vice-President, and A. B. Chamberlain, Chief Engineer of CBS, and under the immediate direction of Howard A. Chinn, CBS Chief Audio-Video Engineer.

Fig. 1
CBS Television City - the studio building at the right houses four one-quarter acre studios; the three-story, glass-walled structure to the left is a service building housing scenery shops, maintenance facilities, and offices.

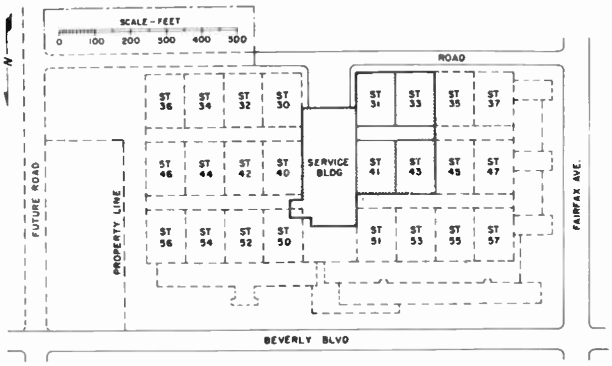


Fig. 2
The initial unit of Television City, shown in the shaded area, is only a small fraction of the ultimate plant. Future studio numbers have been preassigned to insure an orderly numbering system regardless of the sequence of expansion moves.

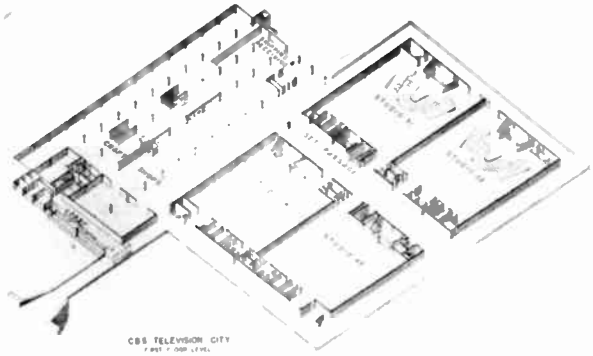


Fig. 3
A cut-away view of the studio floor level shows the easy accessibility from craft shops to studios by means of a 40-foot wide set passage.

Fig. 4
A general view of one of the audience studios. The control room is at the left while the lighting control console and preset panel is in the foreground.

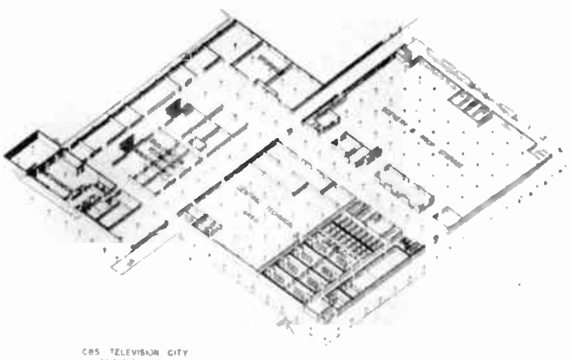


Fig. 5
A cut-away view of the basement level. Well over a half acre of floor space has been provided for scenery and property storage.

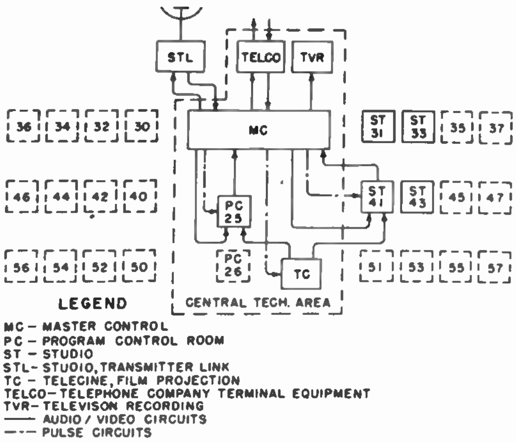


Fig. 6
This simplified block diagram shows the relationship between the technical facilities in the central technical area and a typical studio.

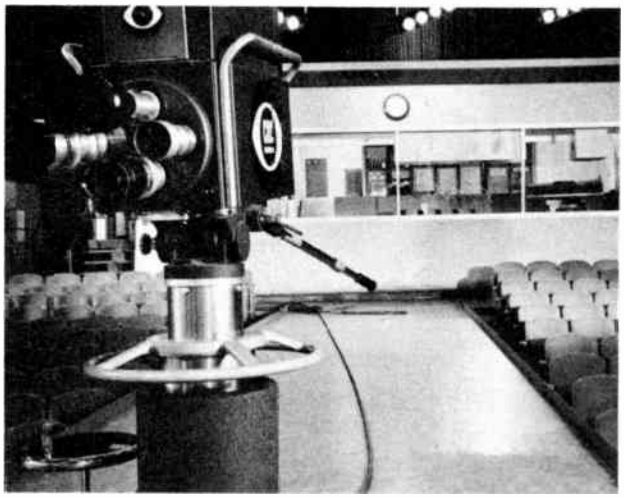


Fig. 7
A view of a studio control room from the stage. Note the removable camera ramp.

Fig. 8
View of a studio stage from the director's position in the studio control room.

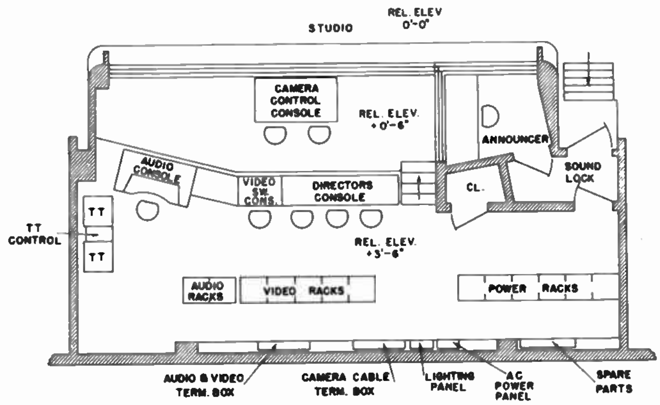


Fig. 9
Plan view of Television City studio control room arrangement.

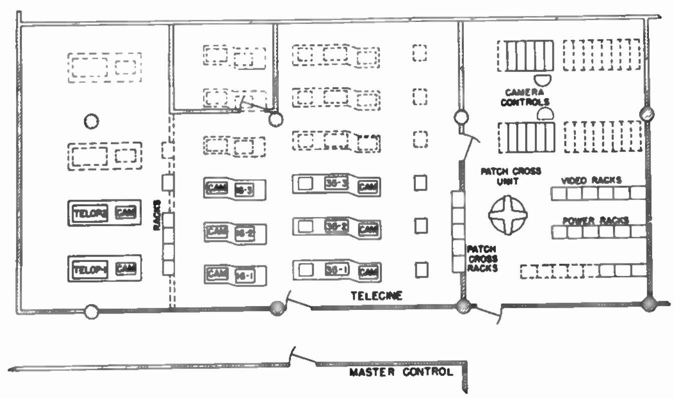


Fig. 10
Plan view of telecine room.

Fig. 11
The telecine room employs a separate film camera with each projector. A standardized rack of equipment associated with each projector-camera grouping contains all audio and communication facilities, as well as picture monitoring facilities.

Fig. 12
The telecine control room contains camera control units associated with film cameras, video distribution facilities, and the patch cross unit whereby the various film projectors are assigned to studios.

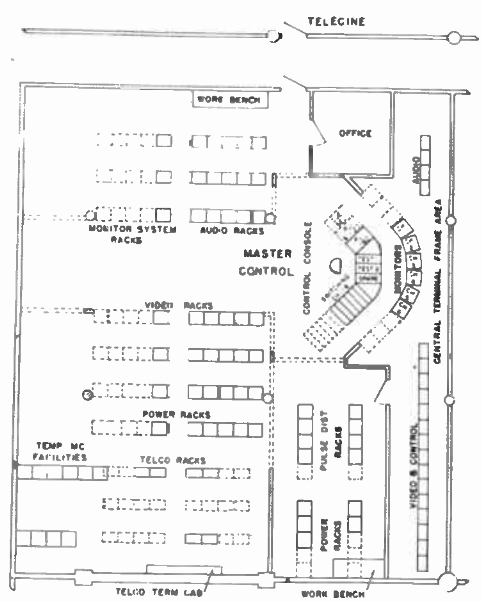


Fig. 13
A plan view of the Television City master control room.

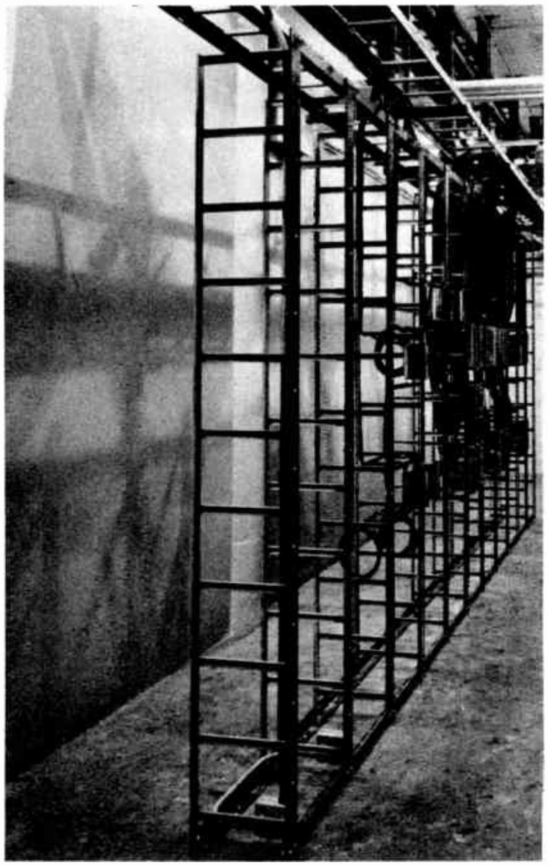


Fig. 14
The central terminal frames adjacent to master control. Space has been allocated on the frames for all audio, video, and communication circuits associated with future studios.

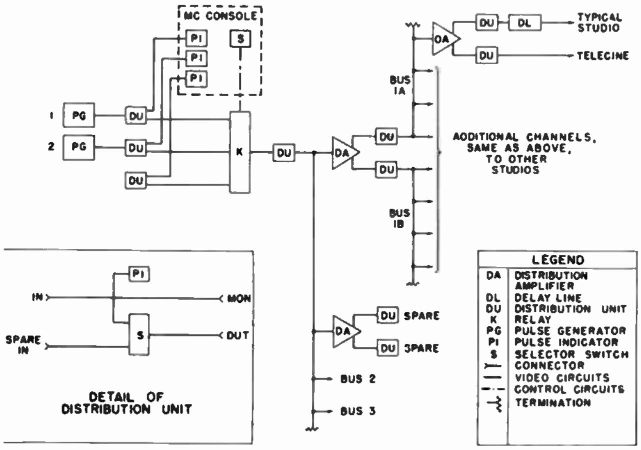


Fig. 15
Simplified block diagram of central pulse distribution system. Although this simplified diagram shwos the circuits for distribution of only one pulse signal, in reality there are four separate pulses distributed by the pulse distribution system, i.e., horizontal scanning drive pulses, vertical scanning drive pulses, blanking signals, and composite synchronizing signals.

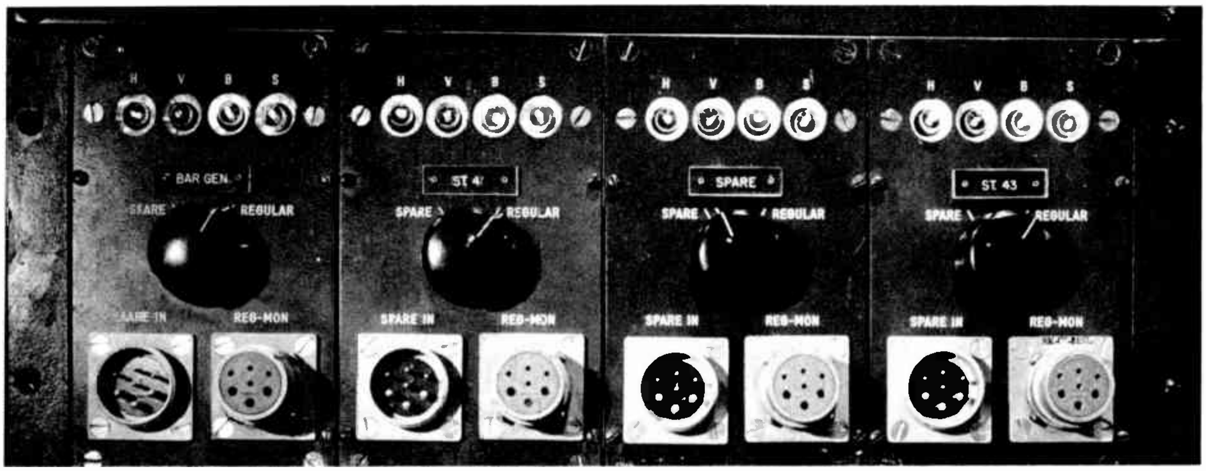


Fig. 16
Pulse distribution units, four of which are shown above, are provided at each important junction point of the pulse distribution system. They provide neon lamp visual monitoring indicators on each of the four pulse signals as well as provisions for switching over to a pre-patched spare pulse feed in an emergency.

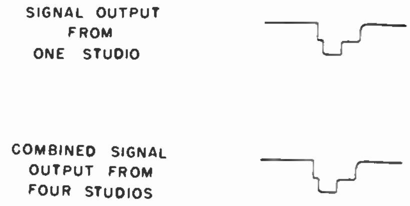


Fig. 17
Upper oscillogram shwos a blanking signal from a telecine film camera routed through one studio to master control. Lower oscillogram shows the same signal routed to all frou studios and received simultaneously at master control.

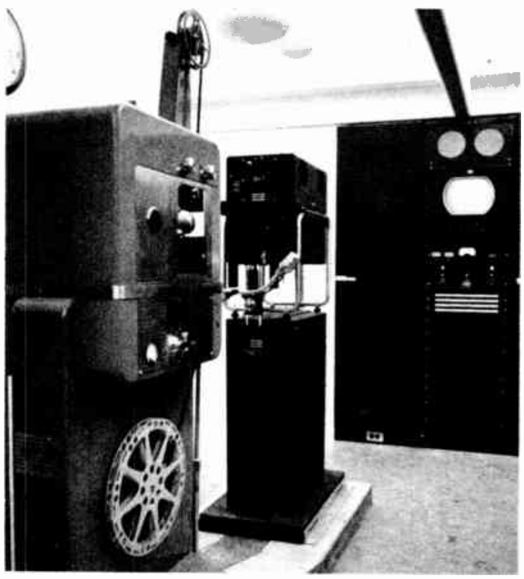


Fig. 18
A close-up of a 16mm projector and its associated film camera. Auxiliary facilities are mounted in the standardized equipment rack, one of which is provided for each projector-camera grouping.

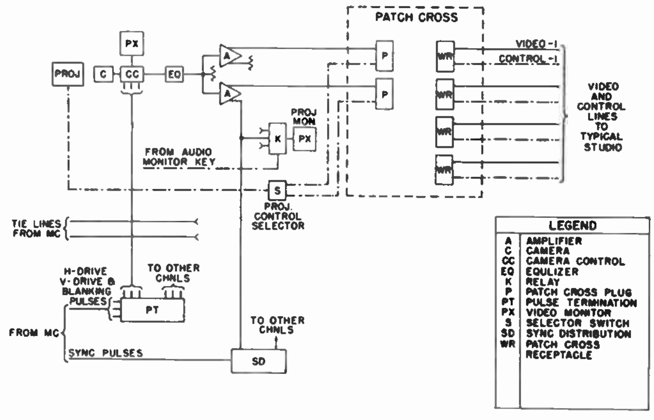


Fig. 19
Blook diagram of the video facilities associated with a typical film projector. The patch cross unit accommodates audio and communication circuits (see Fig. 35) in addition to the video and projector remote control circuits shown above.

Fig. 20 - Photograph of typical studio control room taken from the audio area.

Fig. 21
The studio camera control console located on the lower level of the control room. Ample space has been provided for additional control equipment.

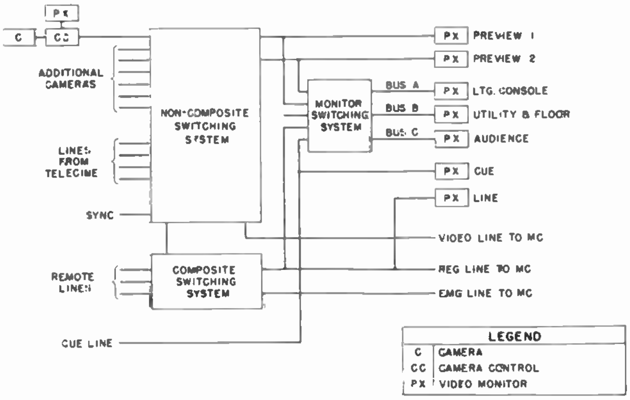


Fig. 22
Simplified block diagram of the studio video switching system.

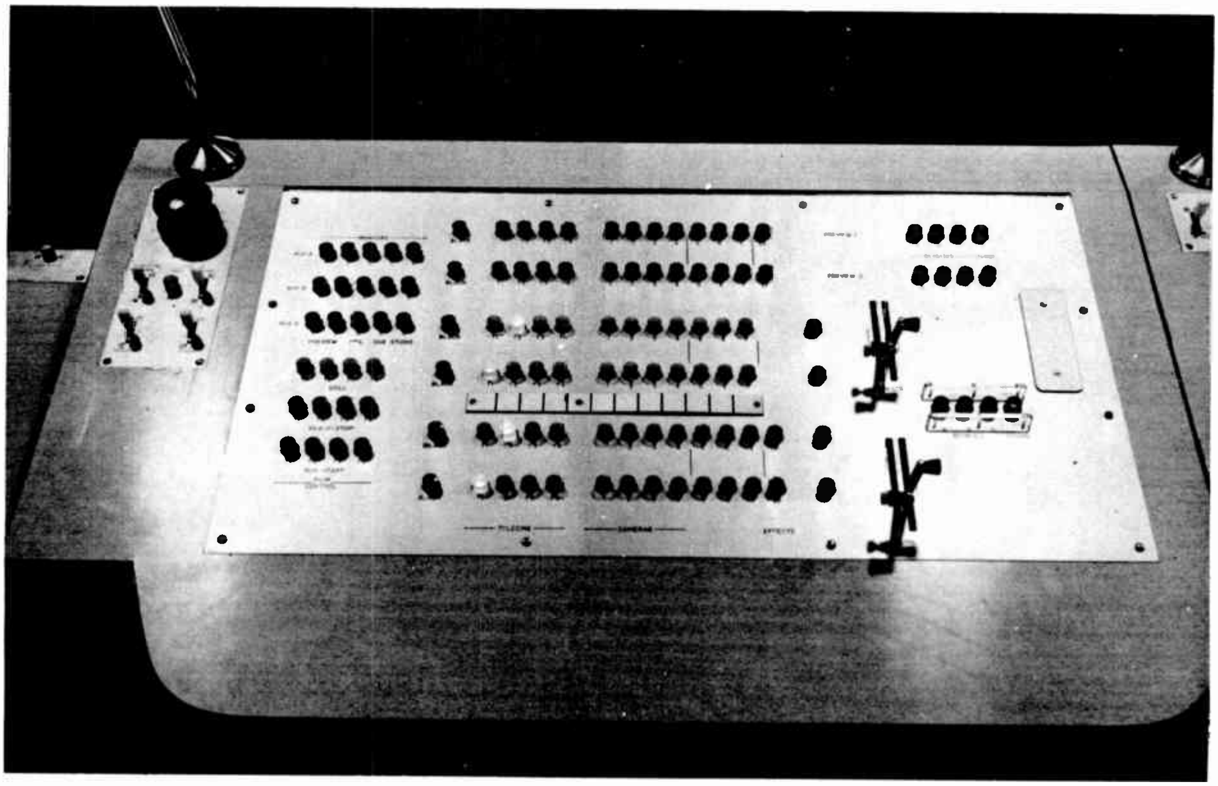


Fig. 23
A close-up of the studio video switching console. The small panel to the left contains intercom controls and a monitoring loudspeaker volume control.

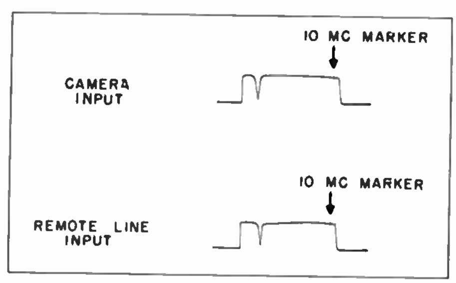


Fig. 24
Over-all response of Studio 33 video facilities from a typical camera input and from a remote line input.

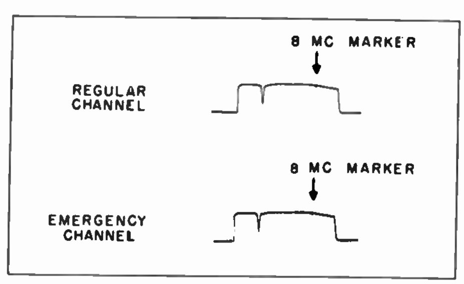


Fig. 25
Over-all video response of Studio 33 through master control.

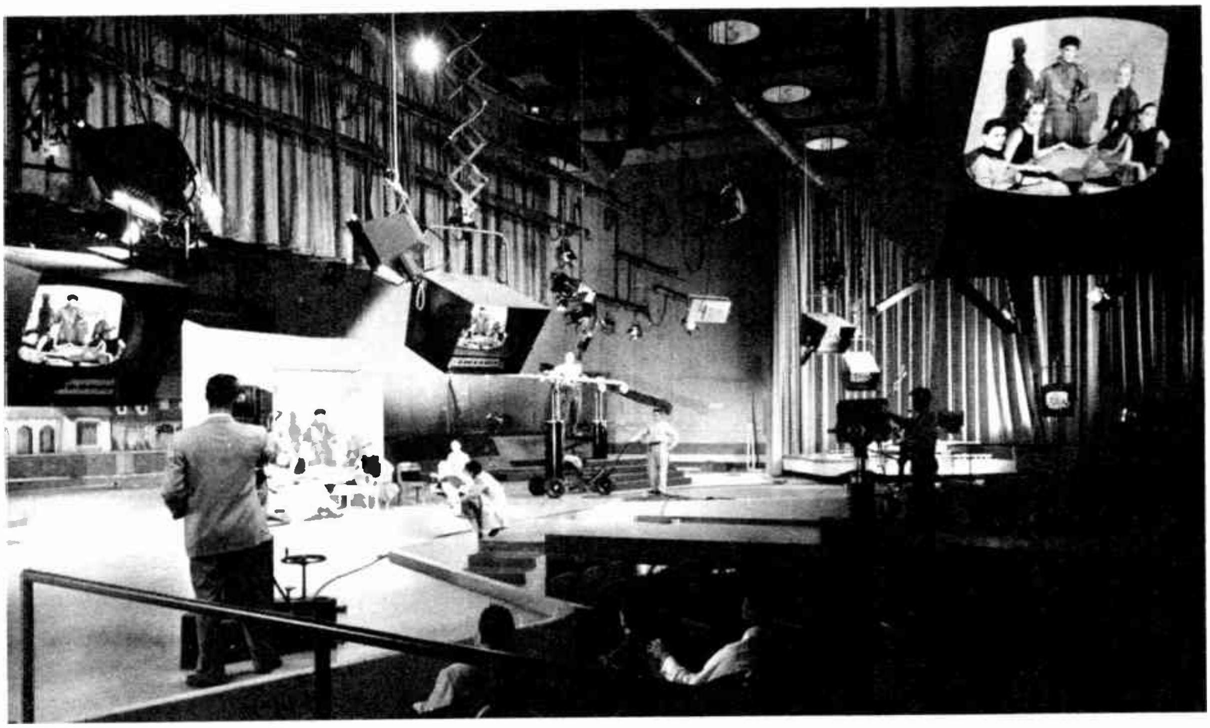


Fig. 26
A view of the stage and audience areas of an audience studio showing the pantograph suspended audience picture monitors.

Fig. 27
The general appearance of a Television City control room. The audio control console is in the foreground.

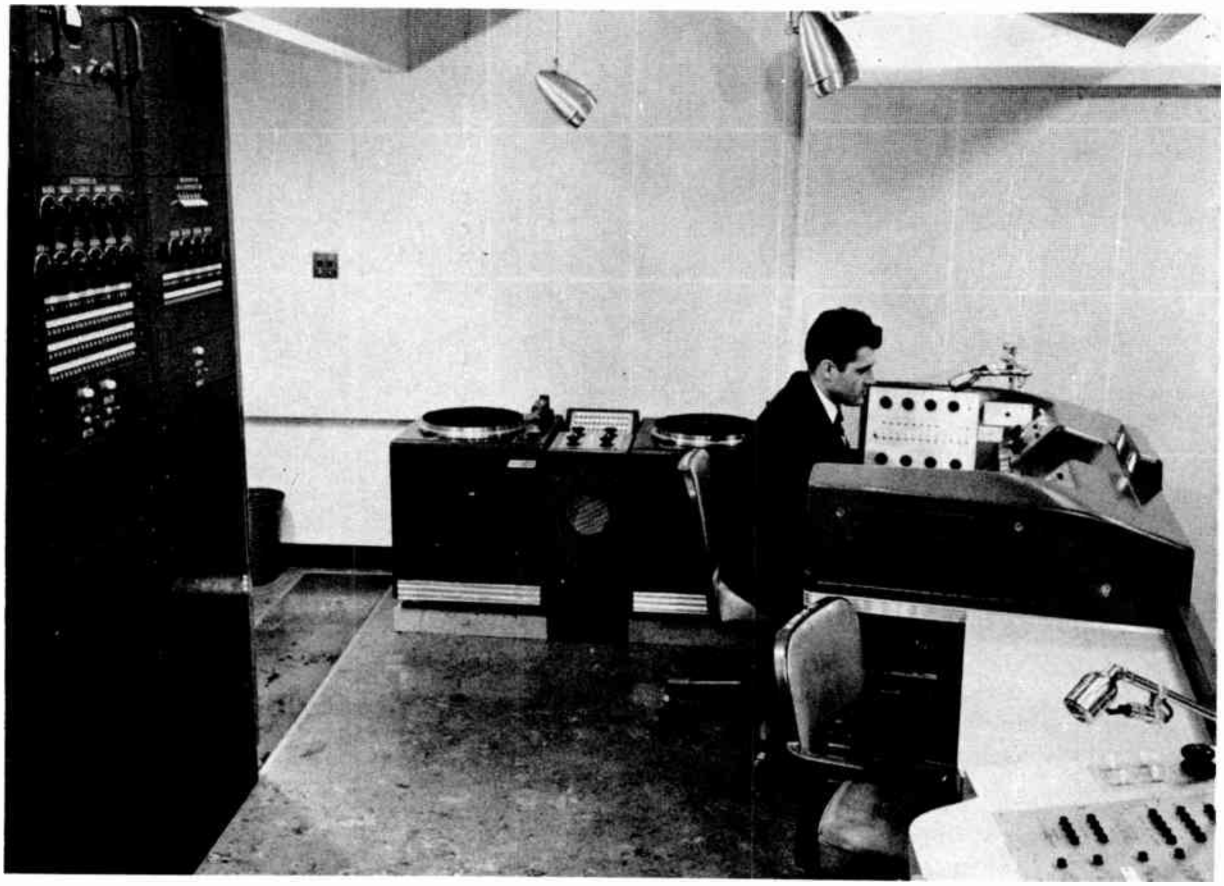


Fig. 28
A view of the audio area of a studio control room showing the audio console, transcription turntables with their associated control unit, and two equipment racks. The rack on the left contains studio communication facilities, the rack on the right contains equipment associated with audience sound reinforcement and audience reaction microphones.

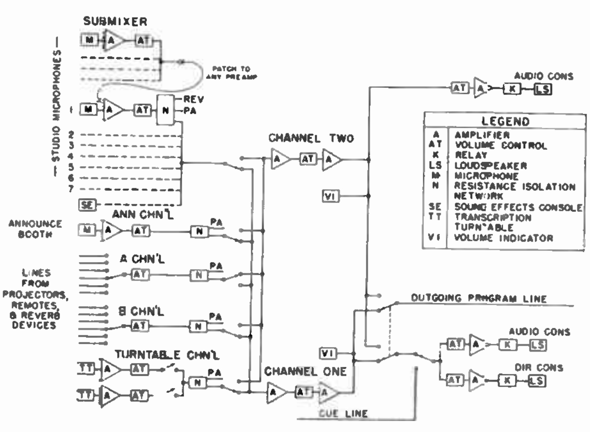


Fig. 29
This simplified block diagram shows the scope of the facilities provided by the studio audio console. The studio audio facilities are capable of simultaneously mixing 21 of 28 normally connected program sources.

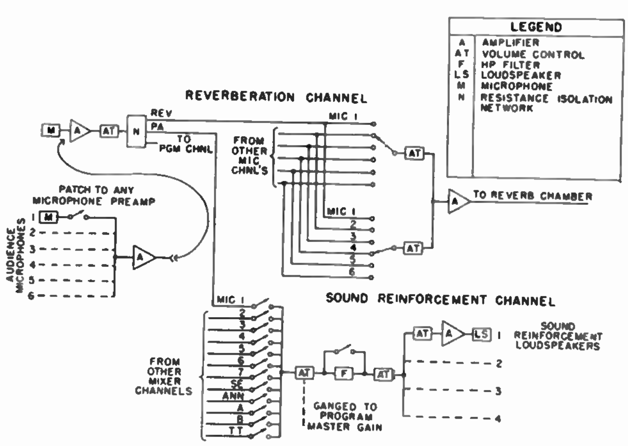


Fig. 30
This block diagram supplements Fig. 29 and shows studio reverberation facilities, sound reinforcement facilities, and audience microphone mixer.

Fig. 31
One of the studio sound effects stations located on a balcony overlooking the studio. The sound effects console is equipped with a clear plastic script and record holding rack that does not obstruct forward vision of the operator.

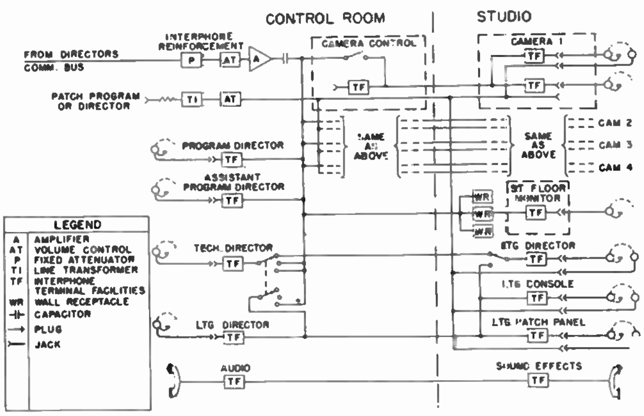


Fig. 32
The interphone system, one component of the studio communication system, provides telephone communication between many involved in the program production.

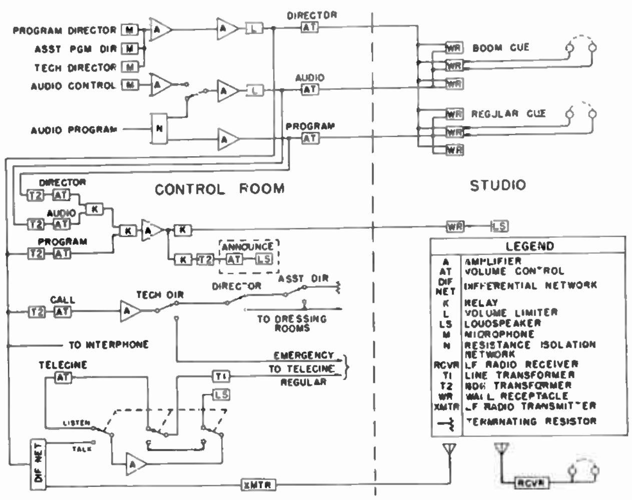
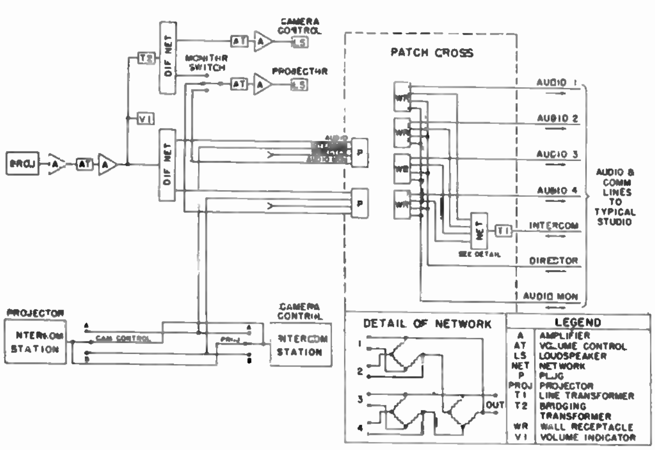


Fig. 33
A simplified block diagram of the Television City studio communication facilities. These facilities augment the interphone system shown in Fig. 32.

Fig. 34
Communication to the studio floor manager, as well as other studio personnel who must be free to move to any part of the studio, is accomplished by means of a low-frequency induction-field transmission channel. One of the compact receiver units is shown in this photograph.

Fig. 35
This simplified block diagram shows the audio and communication facilities associated with each telecine film projector. Connection of projector to any studio is made by the patch cross unit.



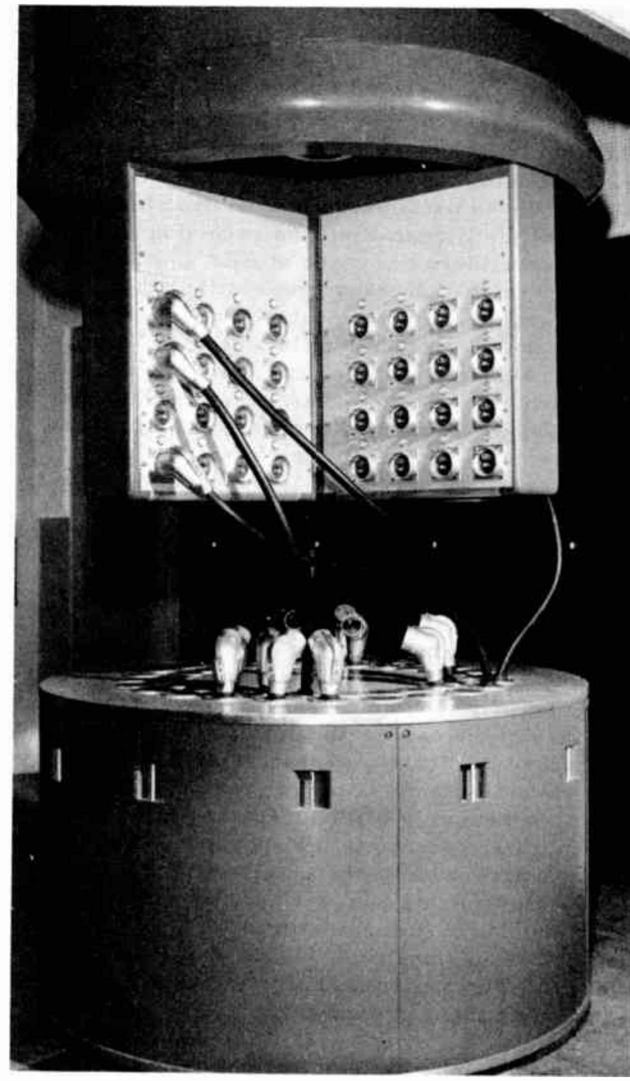


Fig. 36
A closer view of the patch cross unit. All video, audio, communication, and remote control circuits between a projector and studio are established by a single patching operation.

AN EXPERIMENTAL STUDY OF WAVE PROPAGATION AT 850 MC

Jess Epstein and Donald W. Peterson
RCA Laboratories Division, Princeton, N. J.

In establishing a television station the prediction of the service area is of prime importance. In general this involves an evaluation of the effects of such factors as wave refraction, earth reflection, diffraction, and attenuation. Most of these parameters depend upon both the height of the transmitting and receiving antennas. Plans were therefore made at RCA Laboratories for ultra-high-frequency propagation studies which would permit the evaluation of these factors in the prediction of field strength throughout a broadcast service area. The most suitable tower overlooking typical terrain was the 760-foot WOR-TV tower in North Bergen, N. J., and arrangements were made to use this location.

Field experiments were conducted in the summer of 1952 with apparatus installed on WOR's tower, on the Palisades just west of New York City. Four high-gain unidirectional antennas with beam widths which were narrow vertically and broad horizontally were equally spaced along the tower's height. These antennas were of a novel zig-zag type developed at the Laboratories. Because of the narrow vertical patterns, it was necessary to tilt the antennas vertically to enable measurements on pattern maxima. This was done remotely with selsyns to convey accurate angular position to the operator. The transmitters were 10 watt, self-excited oscillators also remotely controlled and mounted at each antenna in order to eliminate costly and lossy transmission line. Field observations were made in a car which was in radio communication with the station operator. The car was equipped with a motor-driven telescoping 35-foot mast which was installed periscope fashion in the roof of the car.

Measurements were made along two quite different radials, one flat and the other relatively hilly. The data thus obtained is sufficiently general to illustrate the effects and importance of the various propagation factors previously discussed.

A certain number of sites were selected which were line-of-sight from the transmitting antenna. Some of these measurements confirmed the existence of an interference pattern between a direct wave and a wave reflected from the ground, as the height of the receiving antenna was varied. However, the ratio of maximum to minimum signal was usually low indicating that the reflected power was small. Where an interference pattern existed no benefit was derived from increased transmitting antenna height since maximum signal could be obtained by an optimum adjustment of the receiving antenna height for any given transmitting antenna height. The maximum field strengths obtained under these conditions were close to calculated free-space values.

An effort was made to obtain a fairly homogenous sampling of measurements in a congested residential area and a similar measurement in open country. These were analyzed to obtain "experience factors" which correlate the median transmission losses with transmitting antenna height. These losses are a measure of the shadowing and attenuation effects of buildings and trees and are one of the factors which must be included in making field strength predictions.

The last class of measurements included sites that were shadowed by hills. These were of extreme importance in that they provided measurements against which prediction methods could be checked. The prediction involves the use of the "experience factors" and the calculation of the diffraction losses introduced by hills. Although subject to many practical limitations this theory predicts that transmitting antenna height is more valuable with moderate than with heavy shadowing. It also shows increasing shadow loss with frequency. Both of these theoretical trends have been corroborated as experimental tendencies. We have concluded that useful predictions of wave propagation can be made with free space theoretical field strength reduced by theoretical knife edge diffraction shadow loss and by suitable empirical experience factors.

HIGH POWER UHF KLYSTRON APPLICATION

Austin E. Rankin
General Electric Company
Schenectady, N. Y.

This paper discusses some of the application aspects of a group of 15 kilowatt tunable klystrons which have made possible ultra-high-frequency television broadcasting in any channel at power levels up to 300 kilowatts, effective radiation.

Six designs, type numbers GL-6237 to GL-6242, cover the entire band, 470 to 890 megacycles and offer many advantages to the equipment manufacturer and broadcaster.

High power transmitter design is greatly simplified as these types contain most of the radio-frequency final amplifier circuit including integral resonators and input and output coupling circuits. Driving power requirements are minimized by the high gain of these devices; better than 30 db when tuned for narrow band operation and on the order of 26 db when broad-banded for visual service.

Each klystron, pretuned before shipment for service in a specific channel, is ready for immediate installation in either the aural or visual transmitter with only minor tuning adjustments being required to optimize performance.

Klystron General Description

These tubes are three resonator designs having unipotential tantalum disc cathodes heated by bombardment, and a collector capable of dissipating 51 kilowatts. The cathode seals and output seal are forced-air cooled; the drift tubes and collector are water cooled. Electro-magnetic focusing of the electron beam is required.

Klystron Construction

Figure 1 shows the functional parts of the tube; tungsten filament, bombarded cathode, the three resonators, input and output coupling loops and water-cooled collector. Each drift-tube section is also water cooled. In use the output probe or antenna, enclosed in a glass bulb which forms the vacuum envelope, projects into and excites a section of waveguide which must seat on and make good radio-frequency connection to the output flange. The input connection consists of a 5/8" diameter 46-ohm line coupling, UG-46/U.

Built-In Portability

Because of the weights of these types, ranging up to 280 pounds for the lowest frequency design, special attention has been given the problem of safe handling, storage, and installation. Permanently mounted on the large diameter steel mounting plate, which also serves as a pole piece for the electro-magnetic beam focusing system, are four rollers which may be used to support and convey the klystron along suitably designed tracks or rails. The crate in which the tube is shipped will contain a shock-mounted frame incorporating a pair of such rails for loading or unloading a tube and a corresponding set of rails in the transmitter cubicle in conjunction with a suitable transfer dolly will eliminate any necessity to lift the tube.

General Circuit Requirements

Figure 2 illustrates the general circuit requirements for operating the klystron. The body of the tube is grounded directly. This eliminates the need for d-c voltage isolation in mounting the tube, in the water lines or between the tube and radio-frequency input and output cavities, however it requires that the cathode and bombarder supply operate below ground by the amount of the beam voltage. The filament and its supply must in turn be insulated for beam plus bombarder voltage.

Filament Supply and Cathode Bombarder

The diode consisting of the bombarder cathode and the filament is operated emission limited in the interest of keeping filament temperature at the lowest possible value for longest life. This encourages some instability in bombarder current. This is due in part to the back radiation effect of the cathode on filament temperature. For example; if a slight increase in bombarder current or voltage does occur, it results in increased input to the cathode and a corresponding temperature increase which in turn causes a slightly higher filament temperature and a further increase in bombarder current. Thus, the effects of slight fluctuations normally encountered in filament or bombarder supply voltages are greatly amplified and must be compensated for. This requires a circuit which will react on increases in bombarder current to decrease filament voltage. This circuit should start the filament at about 8 volts and gradually reduce the voltage to the normal operating value of about 6.3 volts as the filament heats and bombarder current increases to a normal value of about 460 milliamperes.

Magnetic Housing

The magnetic field required for collimating the electron beam is supplied by 4 independent magnet-coils each requiring a variable D-C supply, or if preferable, a common supply and separate variable resistors. These 4 magnet coils include one for each of the three resonators and a small one surrounding the cathode region which is reversed in polarity to cancel out stray field from the main coils. The separate control of each coil current allows maximum flexibility in adjusting for optimum operation with low drift-tube collection current and facilitates utilization of the best possible gains and efficiencies. The magnitude of the magnetic field required varies with the tube type and the channel to which it is tuned. Values will range between 300 and 400 gauss.

Beam Power Supply

The beam supply should have minimum capabilities of 3 amperes at 17 kilovolts for full 12 kilowatt synchronizing peak level operation and should be well filtered. An ammeter in series with the positive side of the beam supply reads total beam current. The negative terminal of this meter is connected to the collector and also to ground through a milliameter which reads drift-tube collection current. The meters should, of course, be shunted by low voltage breakdown devices to protect the meters and insure that the collector and positive side of the beam supply will remain near ground potential in the event of a meter burn-out.

R-F Output Circuit

The output circuit consists of a waveguide section having tunable plungers in each end. The klystron output probe is inserted into the guide near one end and a probe near the opposite end couples power out of the guide into 3-1/8", 50-ohm coaxial line leading to the load or antenna. The waveguide output section is one of two standard RTMA sizes. The WR 1500A, 15" x 7-1/2" (RTMA type) guide will be used for the three low frequency types and WR 1150A, 11.5" x 5.75" (RTMA type) guide for the three high frequency types. The length of the load probe and positions of the end plungers will vary with each channel and data on these parameters will be available for the equipment designer.

Special Tube Protective Precautions

There are a number of tube protective devices which should be utilized in applying the klystron in addition to conventional cooling interlocks and overload protection on all supplies. Mismatches in the waveguide section may cause over-heating or puncturing of the output seal and protection should be provided in the form of a directional coupler and appropriate additional circuitry to remove beam voltage when the reflected power exceeds a certain minimum value. The coupler illustrated is of the "Bethe-hole" coupler type. The VSWR should not be allowed to exceed 2:1 in the guide. A lengthwise slot in the guide with provisions for mounting a probe and detector provide a convenient means of measuring voltage standing-wave ratio.

It is important that the drift-tube collection current be maintained below 250 milliamperes and an over-load relay should be provided to monitor this current and remove beam voltage if it becomes excessive. As further protection it is recommended that the beam supply be interlocked with the magnet supplies to prevent application of beam voltage without the magnetic field.

A protective resistance of 50 ohms should be used in series with the beam supply to limit surge currents in the event of an electrical fault in the tube.

Maximum Ratings

I would like to review briefly the cooling requirements for these tubes and the electrical ratings.

Cooling Requirements

The collector requires a water flow of 15 gallons per minute for rated collector dissipation of 51 kilowatts. This is obtained with a pressure drop across the tube of about 65 pounds per square inch. The drift tubes require 2 gallons per minute at the same pressure drop.

Filament and cathode seals are adequately cooled by an air flow of 300 cubic feet per minute directed upward over the glass seals. The output seal requires 15 cubic feet per minute.

Electrical Ratings

At the normal operating filament voltage of 6.3 volts the filament draws approximately 38 amperes. The starting current with 8 volts on the filament should not be allowed to exceed 100 amperes.

The maximum bombarder input is 1200 watts. This may be obtained with various combinations of bombarder voltage and current within the limits of 3 kilovolts and 550 milliamperes. There is some advantage in operating at bombarder voltages near the maximum limit as this minimizes bombarder current requirements and permits the filament to operate at the lowest possible temperature with beneficial effects on life.

Maximum ratings for broad band television service include 18 kilovolts beam voltage and 3.25 amperes of beam current. The maximum beam input, however, is limited to 51 kilowatts by the dissipation rating of the collector which must absorb the entire d-c beam power in the absence of drive. This condition is approached when modulation is at white level.

Performance

The characteristics and performance of the klystron are well adapted for high power television service from the standpoint of gain, linearity and bandwidth.

Output Vs. Drive Characteristic

Figure 3 shows a typical performance characteristic taken on a GL-6241 operating at 750 megacycles. This was taken on a tube synchronously tuned, however a very similar characteristic is obtained on a broad-banded tube with the exception that power gain is considerably lower. It will be seen that the characteristic is very linear up through the region corresponding to black or pedestal level and is sufficiently linear above this level to permit operation at 12 kilowatts for synchronizing peaks with only a moderate amount of pre-emphasis. This particular characteristic saturates at about 14.5 kilowatts. Efficiency at 12 kilowatts is 24.4%.

Narrow Band Response

Figure 4 shows a typical narrow band response curve taken with all three cavities synchronously tuned. Power gain in the linear portion of the characteristic may be as high as 33 decibels. This response has a half-power bandwidth of about 2.25 megacycles.

Broad Band Response

Figure 5 is a response curve illustrating the effect of stagger-tuning the resonators for broad-band operation. There are a number of techniques that might be successfully used in obtaining the response needed for visual service.

One method involves isolating the klystron from a driver having the proper response with an attenuating pad of about 6 decibels. In this case the klystron within itself must also have the proper response. A second technique involves working the driver through a line-stretcher directly into the klystron and using the output-tuning and coupling of the driving stage as additional variables in obtaining the desired over-all response of the combination.

The latter method results in much better utilization of driving power since the attenuation is eliminated and the klystron gain itself is higher as it is not necessary to de-tune the cavities to the extent required by the first method.

The response curve in figure 5 is that of a klystron tuned within itself to have ample bandwidth for visual transmitter service. This response was obtained by synchronously tuning all three cavities to a frequency corresponding to visual carrier and then increasing the frequency of the center cavity. In other words, with respect to the center frequency of the response band, the input and output resonators are tuned on the low side with the center resonator on the high side. It will be noted that the power gain, although considerably lower than for narrow-band operation is still a very respectable figure, 400 to 1 or 26 decibels.

Visual Transmitter Operating Conditions

In order to operate on a characteristic that saturates at 14 or 15 kilowatts, a condition desirable for 12 kilowatt synchronizing peak level performance, it is necessary to use close to maximum input on the highest frequency types; i. e., 3 amperes at 17 kilovolts. This should normally be obtained with approximately 6.3 volts on the filament and with a bombarder voltage of 2500 volts and bombarder current of about 460 milliamperes. At lower frequencies somewhat higher gain and efficiency may be expected and required beam input ranges down to 2.5 amperes at 17 kilovolts permitting a corresponding decrease in cathode heating power.

Aural Transmitter Operating Conditions

The klystron used in the aural transmitter, having to deliver only 6 kilowatts, may operate with reduced beam power. The aural tube may also be operated closer to the saturation level since linearity is not a consideration. The reduced beam input may be obtained by dropping both voltage and current to operate at 2 amperes at about 13 kilovolts or it may be obtained at 17 kilovolts by reducing cathode input to emission limit the beam to a value on the order of 1.5 amperes.

High Continuous-Wave Power Capabilities

Although presently rated primarily for television applications at 12 kilowatts synchronizing peak level, these six klystrons do have continuous performance capabilities up to 15 kilowatts. Thus from any radio-frequency consideration the tubes are operated very conservatively in a 12 kilowatt television transmitter.

In conclusion, we feel that the 15 kilowatt klystron with integral cavities, essentially a complete radio-frequency amplifier in itself, is the logical choice for high-power television. Availability, exceptionally high gain, reasonable efficiency for these frequencies and long life and reliability inherent in the design are factors which have brought these types to leadership in the field of high-power ultra-high-frequency television.

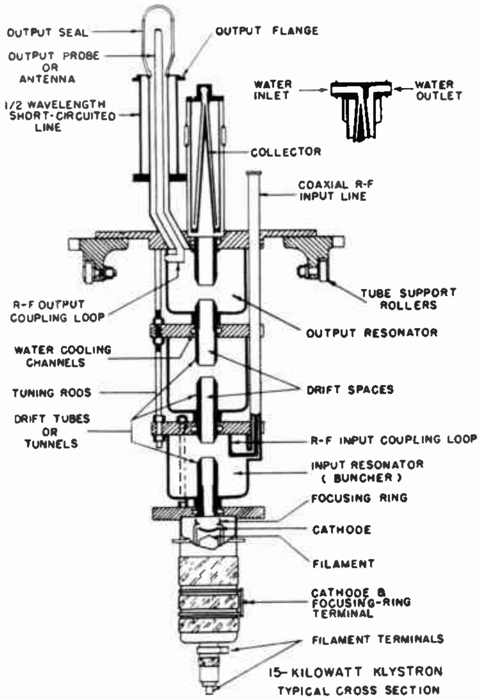


Fig. 1
15-kw klystron - typical cross section.

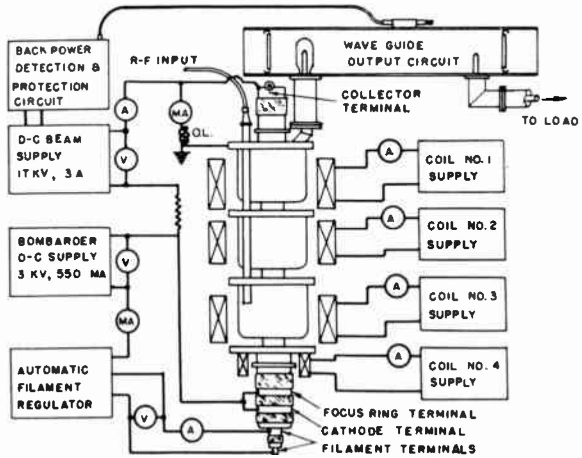


Fig. 2
Simplified circuit diagram - 15-kw klystron.

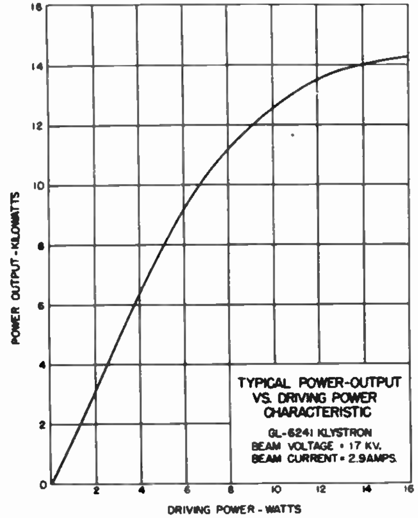


Fig. 3
Typical power output vs. driving power characteristic GL-6241.

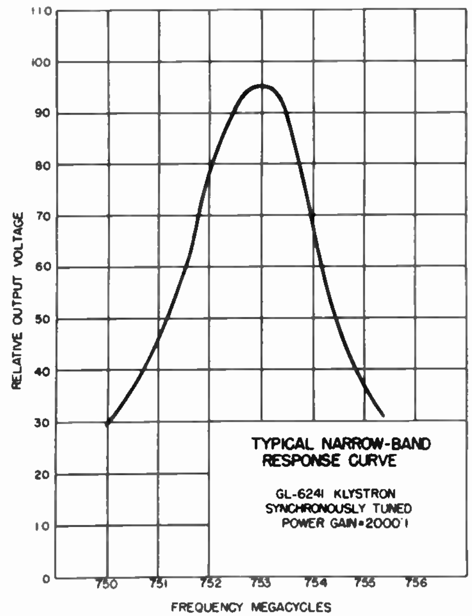


Fig. 4
Typical narrow-band response curve GL-6241.

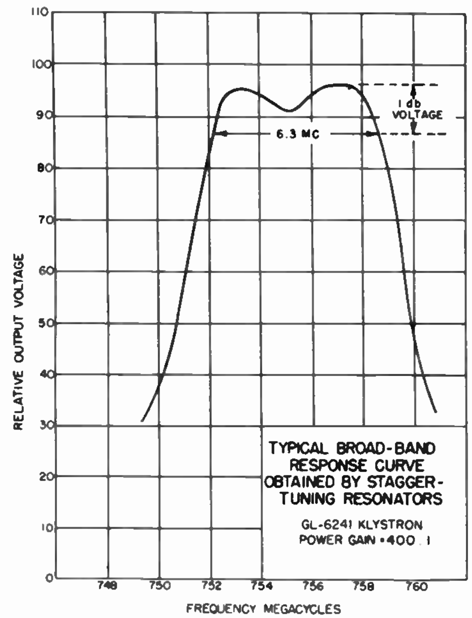


Fig. 5
Typical broad-band response curve GL-6241.

HIGH POWER UHF KLYSTRON AMPLIFIER DESIGN

Norman Hiestand
Varian Associates - San Carlos, California

In March 1951, Varian Associates announced the development of a klystron suitable for use as the final amplifier of UHF transmitters. Since that time there has been a great deal of interest and widespread acceptance of the klystron as the best means of obtaining high power in the UHF region. Today, tubes are in production to provide 15 kilowatts cw power over the entire UHF-TV band and new designs are under way which will increase this power by a factor of 5 (to 75 kilowatts) and expand the frequency range in which this type of tube is available.

The series of tubes that will be described was developed for the General Electric Company specifically for use in their high power UHF television transmitter and this has been their principal application to date. However, there are many other uses for such klystrons and application engineers are now considering their use in communications transmitters, dielectric heating units, linear accelerators and the like.

Conventional amplifiers begin to run into serious limitations which make their operation relatively poor at ultra high frequencies. The principal limiting factors are electron transit time, lumped electrical reactances and low Q resonant circuits. These limitations are so well known that there is no need to dwell on them here, except to point out how they are overcome. The klystron combines the electronic bunching principal, which simplifies the transit time problem, with the resonant cavity which largely eliminates lump reactances and permits the use of high Q circuits. Since spacings and size of the klystron are not limited entirely by wave length, it is possible to obtain very high power gain combined with high output power and relatively good efficiency. Nor do gain and efficiency fall off rapidly as frequency increases. Gain as high as 33 db can be obtained in the type of klystron to be described. In television service the tubes are operated with about 25 db gain. Efficiencies as high as 38% have been obtained when operating at full 15 kilowatt power output.

Linearity of the klystron is important, especially in the television application where the tube is used as a linear amplifier, since any distortion contributed by it would reduce the quality of transmitted signal. Fortunately, linearity is no problem. Power output is related to power input as a Bessel function of the first order. Almost perfect linearity is achieved up to about 80% of full power. Where great linearity is not required, as in synch pulses, the tubes are operated to full output with a resulting increase in efficiency.

Noise is held to a minimum in the power amplifier klystron. There are no grids, therefore, no partition noise. The cathode is outside of the r.f. interaction region so that no hum is introduced by a.c. heater operation. Random emission is not a problem at this power level. The main source of noise is beam voltage ripple and, only the simplest filtering will reduce this more than 60 db below the carrier. It is these and other advantages which will be pointed out that make the klystron a superior means of producing high power in a tube suitable for use as an amplifier.

The high power klystron will be considered here for a fairly general viewpoint but certain features of the Varian tube that will be of interest to the system designer will be emphasized. Fortunately, any klystron can be divided into three separate units for purposes of description and in the amplifier these are almost completely independent. These units are (1) an electron source or cathode, (2) the r.f. section and (3) an electron collector. Or, put another way, the klystron consists of an electron beam around which are disposed certain r.f. components in a manner that largely eliminates the high frequency effects commonly encountered in more conventional tubes.

First let us review, very briefly and qualitatively, the operation of a klystron power amplifier. The problem is to transfer a low level r.f. voltage to a high level r.f. current with as much gain as practical and without introducing undesirable components such as noise. This is accomplished by superimposing an r.f. component on a d.c. beam of electrons. The r.f. component of current is amplified by the familiar klystron drift action. The klystron differs from conventional amplifiers in that it makes use of the electrons to produce the velocity modulation or bunching. Furthermore, the r.f. control signed is applied after, rather than before the acceleration of the electrons by the beam voltage, thus removing the cathode from the r.f. interaction region. Incidentally, in the region between the cathode and the first gap the electrons are accelerated to the full beam voltage. The current in this region is constant. This is an important consideration to the power supply designer. Since the cathode current is constant, a low impedance or regulated supply is not required and, as already mentioned, very little filtering is required to keep power supply hum at least 60 db below the carrier level.

The r.f. signal to be amplified is introduced into the first or buncher cavity. The voltage appearing across the first gap causes electrons to be either speeded up or slowed depending upon its phase. In the drift space between the first and second cavities those electrons that are speeded catch up with the slower electrons ahead of them. Thus, the beam is broken up into groups or bunches. This bunched current excites the second cavity and in the case of multicavity klystron the beam is again modulated and further amplification occurs between the second and third cavities. In the last or output cavity the beam is debunched as the r.f. energy is extracted by the load. The series of 15 kilowatt tubes has three cavities, the 75 kilowatt tubes will have four. The desired gain and bandwidth are the principal determining factors. There is no theoretical limit to the number of cavities that can be used but practical considerations usually limit the number of three or four, although experimental tubes having even more have been successfully operated. Klystrons with more than two cavities are called cascade amplifiers and are analagous to a multi-stage low frequency amplifier.

Consider now the formation and disposition of the electron beam. The diameter and length of the beam are determined principally by r.f. considerations and physical convenience, the voltage and current by desired power output and expected efficiency. In the case of the Varian tube, a beam about 1/2" in diameter, 3 feet long, using 17 thousand volts and 3 amperes produces the desired results. Most klystrons use a conventional Pierce gun and these tubes are no exception. One of the main requirements of any tube is long life, so instead of a standard oxide cathode, a pure tantalum emitter .100" thick is used. Tantalum combines ease of handling with relative freedom from contamination and will give many thousand hours of operation. The cathode in any klystron is indirectly heated to prevent beam modulation by filament voltage. This cathode is heated by bombarding it with d.c. beam of electrons obtained from a tungsten filament. 2500 volts, at 1/2 a is required to heat the button. The cathode temperature is approximately 2300°C. The life of the heater cathode combination is designed to exceed 10,000 hours. One tube operated for over 7,000 hours and failed only after someone dropped a wrench on it, figuratively speaking. When a failure occurs in the cathode the tube can be returned to the factory repaired, reprocessed, and returned to the customer as good as new. This can be done several times. The cost of this operation is far less than the original price of the tube. It is this feature which makes the total life of a given tube exceptionally long and the cost per hour of operation exceedingly low. Actually, this compares very favorably with the operating cost of high power transmitting tubes in the broadcast region.

The beam terminates in a dissipating electrode, or collector that is capable of dissipating the entire beam power, which in the case of the 15 kilowatt tube can be as high as 50 kilowatts. The collector is simply a cylinder over which water is passed at a fairly high velocity - about 16 g/m at 65 psi pressure drop. It is insulated from the body or r.f. section of the tube to aid in monitoring current.

The beam is kept from spreading in the drift section by an external axial magnetic field which can be supplied by a long solenoid surrounding the tube. In practice, several coils are used to permit access to the tube for tuning and other adjustments. Individual control of the fields provides a variable which can often be used to optimize gain and power out and to regulate the amount of current that is collected by the drift tubes instead of the collector. This current is called stray or body current. By proper adjustment of the fields 100% of the beam will reach the collector, however, maximum power output does not occur when body current is a minimum. Generally, about 10% of the total current is permitted to remain in the drift section. The reason is simple enough. Coupling between the beam and the resonant cavities is obtained in the gap of the resonant cavity. The voltage gradient across the gap is maximum at the edge and diminishes toward the center, and coupling is greatest where the voltage is maximum. Therefore, as the beam diameter is made smaller, to minimize body current, coupling is reduced and performance falls off.

The resonant circuits used on a tube of this type may take one of several forms and there are many theoretical and practical considerations that effect their design. The first basic choice facing the designer to the type of cavity to use. External cavities, that is cavities which are not part of the vacuum envelope, offer certain advantages among which are flexibility, possibly greater tuning range and easily variable coupling. Integral cavities eliminate all sliding contacts and reduce the cavity losses. Both these factors are extremely important at high power. A practical tuning range is obtained in the integral cavity klystron by flexing a copper diaphragm to vary the gap spacing. No mechanical tuning problems are encountered, since tubes of this type are not required to be tuned repeatedly or even often. Once adjusted they will stay on frequency indefinitely. Perhaps the most important feature of the integral cavity tube from the point of view of the system engineer is that all of the high level r.f. circuitry is designed into the tube and is not subject to any type of deterioration during the life of the tube. Nor, is any maintenance every required.

The only circuit carrying high power, which is not part of an integral cavity tube is the transmission line to the antenna.

For maximum usefulness, any tube needs to operate over as wide a frequency range as possible, naturally, integral cavities tend to limit this range somewhat but, since it is felt that at high power levels the advantages far outweigh the disadvantages, integral cavities are used on the Varian tubes. The useful range of any klystron amplifier is limited by mechanical and electrical considerations. For maximum gain and efficiency the electrical length of the gaps is optimum at only one frequency. For a given gap spacing as frequency is raised the gap becomes electrically too long, as frequency is lowered it becomes too short. In either case performance falls off somewhat.

Cavities are designed with several important considerations in mind. For maximum tuning range the gaps must be as short as possible and at the center of a doubly re-entrant cavity. The gap should be short in terms of an r.f. cycle but not so short that multipactor can occur between the gap edges. This, of course, is particularly important in the high level cavities. A word about multipactor might be in order, for it is a very real problem at high power and unless carefully considered in the design can cause unsatisfactory performance at certain frequencies. Multipactor is the result of secondary emission taking place in an r.f. field, between two surfaces, such as the gaps. If the power level is high enough and the field has the correct phase relationship to accelerate the secondary electrons in phase then a regenerative r.f. current can be built up between the surfaces. The effect is to load the cavity and possibly, if there is enough power, to cause damage to the surfaces. With this in mind a minimum gap length is set and the length and diameter of the cavity adjusted to resonate at the desired frequency.

Energy must be coupled into and out of the tube through tuned circuits - generally, but not necessarily, loops in the input and output cavities. It is helpful if these circuits can be made tunable, since broadbanding of this type of circuit is difficult. The Varian tube uses a coax which couples directly to a type N line. The output is also coax and it terminates in a probe which is inserted into standard RTMA waveguide. Output coupling is adjusted by a moveable short in the end of the waveguide.

Coupling into the first cavity is adjusted to present a match to the input drive. Usually, a 50 ohm line is used. This, combined with known factors of beam loading and shunt resistance, determine the external Q of the cavity. This is measured and the coupling loop adjusted to give the desired value. The output coupling is not quite so simple but in general follows the same procedure except that it is designed for optimum debunching and not necessarily for a match. In all of this design the parameters are determined partially from theory and partially from past experience. There are too many variables to set up an all inclusive design formula. As is the case with most tubes, there is no substitue for a designer with years of experience.

Taking all of the various factors into account Varian engineers have designed a series of 6 tubes for the General Electric Company's UHF television transmitter. Three use 9" diameter resonators, 3 use 6" resonators. Each has a operating range of about 11%. They are tuned by varying the length of the gap. Input and output coupling is adjusted externally where necessary so that the tubes will meet the following minimum performance standards. Operating at not more than 18 kv, with beam power of 51 kilowatts the tubes will deliver 15 kilowatts saturation power with a gain of 24 db when the cavities are stagger tuned to give a bandwidth of 6 megacycles to the 1 db points.

A typical tube operating at 17 kv, 3 amperes at 750 megacycles has given 30 db gain at full power and 32 db low level gain. At 16 kv the low level gain is practically unchanged while the high level gain is somewhat lower. The fact that gain remains almost constant up to about 3/4 of full power illustrates very effectively why power supply filtering is not a serious problem.

When all three cavities are tuned to the visual carrier frequency of a television channel a bandwidth of about 1 1/2 megacycles is obtained. The tubes are broadbanded by adjusting the resonant frequency of the second cavity approximately to the aural carrier frequency. This results in a band at least 6 megacycles wide to the 1 db points with slight peaks at the carrier frequencies. Properly adjusted the tubes will transmit either the visual or aural signal.

The designs described here can easily be expaned to cover frequencies both above and below the TV band, since once determined, it is a simple matter to scale the design to a new center frequency.

With the FCC announcement that a maximum effective radiated power of 1 megawatt will be permitted in the UHF-TV band it has become apparent that still more powerful amplifier will soon be required. With the television in mind Varian engineers, at the request of the General Electric Co., have started the design of an amplifier klystron which will deliver a saturation power of 75 kilowatts. The design for this tube is still in the early stages but a few target specifications will probably be of interest. The tube will operate at between 30 and 35 kv with about seven amperes beam current. It will have a broadbanded cathode similar in design to that used on the 15 kilowatt tube. Since it will be advantageous to keep the r.f. drive power at a relatively low level tube is being designed with four cavities. This will insure a minimum gain of 40 db narrow band and 30 db when broadbanded. The addition of the fourth cavity will slightly improve efficiency so that a value of 40% may be expected at saturation power.

There are two big problems that will have to be solved before a successful tube can be built. These are the collector and the output window. There is no theoretical limit to the amount of power that can be generated by a klystron. One can certainly construct cathodes capable of supplying higher currents and operating at higher voltages than contemplated here. Focusing this beam in the drift tube will require only about 200 gauss because of the higher voltage being used. Actually, collecting the 225 kilowatt beam only becomes difficult when the overall length of the tube is considered. In order to be practical the tube must be held to a reasonable overall length which means that a clever design will be required to minimize collector length.

Almost as much power as is being contemplated here has been transmitted through a ceramic dome similar to that used on the smaller tubes. A version of this dome, probably water cooled, may be used, but it is felt that a waveguide output through a ceramic window holds more promise.

Design work is proceeding on this higher power tube and it is probable that the first model will be tested in 1954. Further technical information about the klystrons described here is available from either the General Electric Co. or Varian Associates.

HIGH POWER UHF TELEVISION BROADCASTING SYSTEMS

H. M. Crosby
Development Engineer, Electronics Division
General Electric Company, Syracuse, N.Y.

Introduction

Several years of television broadcasting experience in this country have shown the need for a considerable increase in transmission power of stations now operating in many localities. Several VHF stations have recently increased their power many-fold, and others are now making plans to do so in the near future.

When it became apparent that the UHF band might have to support a large part of our television service, several independently-conducted propagation surveys were made to determine the usefulness of this part of the spectrum. Results of these investigations have, in general, indicated that high power transmission would be of special advantage at these frequencies. It is undoubtedly true that in certain areas, well-suited topographically for UHF propagation, low and medium power stations will provide quite satisfactory service. However, in many localities now assigned UHF channels, there is likely to be no good substitute for high power. The high ceiling of 1000 KW assigned to UHF television by the Federal Communications Commission indicates recognition of this fact. Unfortunately, the difficulties of generating high power at these frequencies, as compared to operation in the VHF bands, increase at least in proportion to the frequency.

Two years ago the writer described the development and design of a successful 5 KW UHF television transmitter employing a newly-developed klystron amplifier tube.[1] Experimental on-the-air operation of this equipment, together with intensive tube and circuit development, has now resulted in the successful commercial design of a 12 KW transmitter. This transmitter, in conjunction with a new design of high gain antenna,[2] will provide up to 300 KW effective radiated power throughout the UHF television band. The transmitter and antenna are now in regular production and several commercial installations are underway. At least two of these are already in daily operation and many more will be before the end of the year.

Preliminary Considerations

In the preliminary stages of the usual transmitter design, the engineer often has available complete data on several types of tubes which might serve the purpose. Then by a series of comparisons and compromises, coupled with a certain amount of development, he tries to arrive at a balanced design of transmitter. For the initial planning of the high power stages in this transmitter, there were no suitable tubes available. It was apparent that the tubes and the transmitter would both have to be developed in parallel. A study was made of the type of tube which would be best suited to high power UHF operation. Low power versions of certain types, such as magnetrons and traveling wave tubes were tested for television service. Advantages and disadvantages were weighed. The circuit to be used with the tube had to be given careful consideration. All in all, it seemed to us that the klystron with built-in r-f circuitry offered the best possibilities. Looking at it today, we believe the choice was a fortunate one, and certainly one which has resulted in achieving high power UHF television much earlier than might otherwise have been the case. The development of the tube and its manufacture have been described elsewhere in the literature. Here we will consider mainly a brief description of a successful transmitter design using the tube.

General Description of Transmitter

The General Electric 12-KW UHF television transmitter includes a complete 100-watt transmitter, driving separate high power klystron linear amplifiers for the visual and aural signals. The transmitter, as shown in Figure 1, is housed in five cubicles forming one unified assembly. The more frequently-used controls and meters are located on the top front panels of each cubicle. Here, also, are located the operating sequence lights which permit the operator to quickly locate faults in functioning of the equipment. Other controls, relays, switches and auxiliary meters are located on the lower front panels behind non-interlocked doors where they may be inspected or adjusted anytime during operation if desired. The 100-watt driver is located in the center of the assembly, with the visual power amplifier and rectifier cubicles to the left. The aural power amplifier and its rectifier equipment are located to the right. Built-in wiring ducts in the cabinet floors permit simple inter-cubicle connections by means of the wiring harness furnished with the transmitter. The cabinet assembly is 21 feet 6 1/2 inches wide, 37 inches deep and 83 inches high. The high voltage plate transformers, filter reactors, water-cooling equipment and circuit breakers are located external to the transmitter cabinet. In a typical installation these are usually located in a separate room behind the transmitter. Also external to the cabinet are the combination side-band filter and diplexer as well as the dummy load. In this design, component accessibility has been kept in mind. All cabinets may be entered through full-height doors in the rear which are equipped for double safety with door interlocks that remove dangerous voltages and with door-operated grounding switches for all high-voltage circuits. Cooling of cabinets and small tubes is accomplished by filtered air, using built-in blowers in each cabinet.

The Driver Unit

A functional block diagram of the transmitter is shown in Figure 2. The dashed line separates the 100-watt driver from the high power amplifiers. In the 100-watt transmitter serving as the driver, standard commercial tubes are used in circuits, which, for the most part, are conventional up to one-fourth of the output frequency. In the visual channel, these consist of a highly stable crystal oscillator, followed by a buffer and three tripler stages. The visual signal is then doubled twice and amplified to the 100-watt level by means of Type 4X150A UHF tetrodes in cavity-type circuits. The visual crystal oscillator used in this transmitter has been developed to a high state of stability to meet the new rigid requirements of the FCC. This will insure successful off-set carrier operation for minimizing co-channel interference.

Visual Modulator

The output amplifier stage of the driver is grid-modulated by means of the video signal supplied by a clamp-type modulator. This modulator is designed to accept a standard composite video signal. This signal is then processed to remove waveform imperfections and noise, amplified to the proper level and applied to the grid of the r-f driver tube. Figure 3 will help to clarify the functions of this modulator. The composite video signal in passing through the video amplifier, is also applied to a sync extractor circuit, followed by a sync stretcher circuit which permits adding to the original sync amplitude. A pulse-forming circuit generates narrow pulses immediately following the trailing edges of the synchronizing signals. These narrow pulses, through the clamp driver circuit, key the white-clip clamp, the sync-amplitude clamp, and the pedestal-level clamp, which in turn, set the pedestal level of the proper stages in the video amplifier. The white-clip clamp may be adjusted to the correct level to prevent over-modulation in the white region. This is especially important in preventing hum in receivers which are of the intercarrier type. The sync-amplitude clamp provides a means of adjusting and maintaining the correct percentage of sync. The pedestal-level clamp permits adjustment of pedestal power to the desired operating level. The cathode-coupled output stage presents a low-impedance source of composite video signal to the modulated r-f amplifier.

Aural Modulator

Referring again to the block diagram, Figure 2, it will be noted that the frequency-modulated aural signal is generated by a phase-shift modulator, the center frequency of which is controlled by a crystal having a frequency in the range of 147 to 268 kilocycles, depending on the assigned channel. The audio signal, after passing through the proper network, phase-modulates the crystal-controlled center frequency in such manner as to result in pure frequency modulation having exceptionally low noise and distortion.

Aural Frequency Control

The signal from the aural modulator is now multiplied and filtered, after which it is applied to a balanced modulator or mixer where it is added to the considerably higher frequency derived from the visual crystal oscillator. The resulting sum signal becomes the controlling source of the aural carrier. The visual and aural crystal frequencies are selected in such ratio that the visual crystal produces about 85% of the frequency control of the aural carrier. This signal is then multiplied by a series of stages quite similar to those used in the visual channel. The result is a frequency-modulated aural carrier whose center frequency is 4.5 megacycles higher than the visual carrier, and essentially controlled by the visual carrier crystal. In other words, the visual and aural carriers are effectively locked together with a fixed separation of 4.5 megacycles, both carriers deriving their principal frequency control from the same quartz crystal. This method of control permits utilizing to its maximum advantage the intercarrier type of television receiver. There are two other advantages resulting from this unusual method of controlling the aural carrier. It results in fewer stages of frequency multiplication in the aural part of the transmitter. It also results in a lower noise level in the aural output, since any inherent f-m noise in the first stages of the aural modulator, instead of being multiplied all the way up to final frequency, is instead merely "added" in the mixer stage. This results in about 18 db less noise than would have obtained from separate crystal control and straight multiplication of the aural signal.

Driver As Emergency Transmitter

It may be emphasized again that the 100-watt driver just described is a complete television transmitter, and under emergency conditions, can give a very good account of itself when coupled directly into the antenna. It's effectiveness should not be underestimated. Surprisingly good results have been obtained from this low power in certain localities having terrain well-suited to UHF propagation. Figure 4 taken from a survey made recently by an independent consulting firm [3] illustrates what coverage has been obtained from this 100-watt transmitter operating temporarily on one-half power at a typical installation in the middle west. Equally good results are being obtained at other installations.

Klystron Power Amplifiers

The visual and aural power amplifiers of this transmitter employ multi-cavity klystrons as linear amplifiers, driven directly by the 100-watt driver. These tubes, developed especially for UHF television, and in particular for this transmitter, have been well described in other papers.[4] It should be pointed out here that their use greatly simplifies the high power stages of this transmitter, due to their high power gain and because of the fact that the r-f circuitry is a part of these tubes.

Klystron Tuning

While the klystrons are usually tuned for broadband television operation by the tube manufacturer before shipment to the station, tuning adjustments can also be easily made in the field if necessary. An r-f sweep oscillator, furnished with the transmitter, may be used to check the transmitter response when desired. This oscillator is provided with crystal frequency markers to permit adjusting the response to the correct value. Figure 5 shows a typical response curve taken on a 12 KW klystron properly adjusted for operation in this transmitter. This tube has been adjusted so that it is suitable for operation in either the visual or aural amplifiers. This is the normal adjustment, since each station is usually equipped with three tubes, two active and one spare.

Filament Control

Long life expectancy is one of the features of these klystrons. To obtain maximum service from each tube, the filaments are operated temperature-limited. While this can be accomplished by manual control if desired, an automatic filament control circuit has been included for each klystron used in this transmitter which is extremely simple and reliable in operation. This permits operating each klystron filament at the minimum temperature necessary to provide the desired r-f output power from the tube. The filament control equipment also automatically compensates for normal line voltage variations. This method of control will extend considerably the useful life of a tube over that to be obtained by the more common space-charge-limited operation usually employed in transmitter tubes.

Output Coupling

Since the klystron itself includes all the resonant r-f circuits, there are none of the usual transmission - line tanks, capacitors or cavities with sliding contact tuning fingers to flash over or burn out. A short section of waveguide is employed to couple the klystron to the output coaxial line. Once installed initially for a particular station, this waveguide needs no adjustment of any kind.

Handling the Klystron

The klystrons used in this transmitter are naturally larger and heavier than conventional low frequency tubes used with external circuits. However, they may be easily installed and replaced by average operating personnel. In this equipment the handling of the tube becomes a very simple procedure easily carried out safely and quickly by one person using a special carriage or dolly provided with the transmitter. Figure 6 shows a klystron mounted on one of these dollies, ready for installation in the transmitter. In this case, the spare tube has previously been fitted with its field coils and other attachments and stored in a spare tube rack so that it is always ready for quick installation in the transmitter. The dolly supports the tube on two rails similar to those installed in the transmitter cubicle. To install the tube in its cubicle, the rubber-tired dolly is rolled up to the rear of the cabinet, where its two rails are clamped to the extended ends of the tube-support rails in the transmitter. The tube then rides on its ball-bearing races into the transmitter cabinet to the operating position. Its r-f output flange is then fastened to the waveguide by means of thumbscrews. Water-cooling connections are quickly made by means of quick-connect snap-on fittings. Filament, cathode and field coil connections are plugged in and the cathode-cooling air boot fastened in place. The tube, having been previously tuned before storage in the spare tube rack, is now ready for operation as shown in Figure 7.

High Speed Overload Protection

The transmitter operates from a 208/230 volt 3-phase power supply. The switchgear consists of three electrically-operated air circuit breakers which serve as the primary contactors for the high voltage plate supplies. These breakers provide high speed protection not usually possible with ordinary contactors. Their high interrupting capacity of 15,000 amperes is a desirable feature too often overlooked in the design of high power transmitters. This transmitter also offers other time-tested and desirable features for minimizing lost-program time, such as multi-shot plate overload recycling, instantaneous power line failure reclosure, emergency start, and, of course, a complete supervisory light system.

Filter and Diplexer

To meet the RTMA and FCC specifications for vestigial sideband transmission, and to permit operation of both the aural and visual transmitters into a single transmission line feeding the antenna system, this transmitter is supplied with a combination sideband filter and diplexer network. This consists of a series of resonant cavities and transmission line sections so tuned and adjusted as to properly attenuate the lower frequency sideband, and at the same time, permit combining of the two transmitter outputs with negligible interaction. This network is of the constant impedance type and is mounted external to the transmitter, usually on a wall directly behind the cabinet.

High Voltage Rectifiers

The rectifiers used to supply high voltage d-c are usually one of the less-interesting parts of a new transmitter design. However, in the case of this transmitter, an unusual arrangement of rectifier circuits results in considerable saving of space, and what is even more important, a sizable reduction in operating tube cost actually amounting to several hundred dollars per year. A high power klystron requires, for efficient operation, a rather high voltage, as compared to conventional types of tubes. In this case, for normal full power operation of the visual amplifier, d-c potentials of around 17 KV are employed. Normally, this would require mercury vapor rectifiers of large size to safely withstand the high inverse voltages. Yet, the relatively low currents required fall well within the ratings of small rectifier tubes. Series operation of two separate rectifiers using the low cost Type 673 tubes proved to be the answer for this transmitter in the case of the visual output stage. The aural amplifier, requiring somewhat less voltage, is supplied by a single rectifier identical to either of those used for the visual channel. All three rectifiers use identical liquid-filled 3-phase plate transformers, provided with voltage-changing taps and arranged for either wye or delta operation. By this flexible arrangement, a wide range of d-c voltages is available. In the event of transformer failure, which is, of course, a rare occurrence with liquid-filled units, the transmitter may be operated indefinitely at a sizable percentage of its full power, using two of the three transformers. This is made possible because all three units are of identical design.

RF Transmission Line

The problem of efficiently transferring r-f power to the antenna system at these high frequencies is a real one. High losses occur in coaxial lines, and large sizes of line must, therefore, be used to keep these losses reasonable. Losses of 20 or even 30 per cent will not be unusual in the use of high towers with long line systems, especially in the higher channels. By the use of waveguide instead of coaxial line, the transmission losses may be considerably reduced. Of course, waveguide is more costly than even the largest sizes of coaxial line and the tower supporting it must be designed for the increased wind load. It is perhaps the best answer for transmitting maximum power in the case of antennas located 750 to 1000 feet above the transmitting equipment.

Antenna System

A very important part of any high power UHF television transmitting system is the antenna. Here we can employ really high gain antennas. Power gains of 20 or 25 are entirely practical in many locations. Since the number of wavelengths in the vertical aperture determines the vertical directivity, it is possible to realize high gain at these frequencies with structures quite reasonable in size. There are several designs of antennas that may be used for UHF television, and in general, the coverage and patterns are similar. But the type of antenna used with this transmitter offers several important advantages. This antenna is of the side-fire helical design. It has high gain per bay, thus minimizing the number of feed points. It is of low impedance design, making it remarkably free from the effects of ice and sleet, and its construction permits use of a simple system for deicing. A method of beam tilt and null "fill-in" is easily applied to this antenna in the field when required by the particular terrain.

A single feed point at the center of each bay drives two helices wound in opposite direction, thus canceling the vertical radiation. Due to the proper spacing from the supporting mast, sufficient radiation occurs per turn to result in almost no reflection from the ends of a helix of only a few turns. The distance around one turn is arranged to be an integral number of wavelengths, so that currents in like elements are in phase. A typical bay of this type would be five wavelengths long and would have a power gain of about five. A number of such bays may be stacked to obtain the desired gain, their feed points being an integral number of wavelengths apart for in-phase excitation.

A five-bay antenna of this type recently installed in the field is shown in Figure 8. The feed line is coaxial, the mast itself serving as the outer conductor. Each bay is probe-coupled to the inner conductor to obtain its proportion of the total power. Impedance match is maintained throughout the feed system, thus achieving maximum bandwidth. There are no points of high voltage in any part of the antenna system. This accounts for its freedom from the effects of ice. The inner conductor of the coaxial feed is shorted to the mast one-quarter wave below the point of line entrance and again one-quarter wave above the top feed point. Being hollow, the inner conductor offers a simple means of feeding the beacon light cables up through the center of the mast without disturbing the antenna field. The entire structure is at ground potential for protection from lightning. Since the top turn of the structure is connected to the mast, a simple deicing means is provided by merely passing 60-cycle current up through the copper-weld helical conductors, the return circuit being the mast itself. The helices are supported from the mast by insulators made of Kel-F compound which has a low dielectric constant, minimizing reflections at support points, and which also provides great strength and excellent weathering properties. The entire antenna is designed to withstand extremely high wind velocities.

A five-bay antenna of this type has a power gain of twenty-five, giving a beam width of 2.1 degrees in the vertical plane, and a substantially circular horizontal pattern. In any antenna having high vertical directivity, nulls will occur at only a few degrees below the horizontal. In the case of the five-bay assembly described, these nulls would occur at approximately 2.2, 4.4 and 6.6 degrees below horizontal. In cases where antenna heights of more than 1000 feet are desired, this might result in low signals at certain distances from the antenna well within the normal service area of the transmitter. In such cases, some beam tilt or null "fill-in" may be required. These adjustments are easily accomplished in the field with this antenna design by use of swivel flanges between each of the different bays, permitting rotation with respect to each other.

Figure 9 is a plot of the vertical directivity determined by actual measurements taken at the factory on the antenna shown in Figure 8. Because it was to be installed on top of a 1000 foot tower, this antenna was adjusted to provide a downward tilt of about 0.6 degrees to minimize nulls. The horizontal pattern of this antenna, plotted from actual factory measurements, is shown in Figure 10.

First Installations

The country's first commercial high power UHF television station began daily operation in February of this year. This station, WHUM-TV, located near Reading, Pennsylvania, uses a 12 KW transmitter and a 5-bay helical antenna of the type described in this paper to radiate more than 250 KW effective power. The antenna is mounted on a 1000 foot tower at the top of a hill about 1600 feet above sea level. Power is fed from the transmitter to the antenna through 7 1/2" x 15" waveguide—probably the longest run of large waveguide installed to date. Figure 11 shows the transmitter as installed at this station. The 5-bay helical antenna, at the top of the very impressive 1000 foot tower, is visible in Figure 12. While field strength measurements have not been completed at this early date, preliminary spot checks indicate good results, particularly to the east and south where the country is of more even contour. As expected, some cases of poor reception were found to clear up by proper installation of suitable receiving antennas and transmission lines. In all UHF localities, one of the most important jobs is certain to be the training of dealers and servicemen in the proper installation of UHF receivers. The managements of these two stations are to be congratulated on being the first two commercial installations to use high power in the UHF band.

Higher Power Development

Now, may we take a brief look into the future. Having reached effective radiated powers of 250 to 300 KW, we must consider the possibilities of attaining the maximum power now allowed by the FCC for UHF television; i.e., 1000 KW. With antennas now available, having power gains of 20 to 25, this would require 40 to 50 KW at the antenna feed point, using transmitters rated 50 to 60 KW. We believe such power is possible, using methods quite similar to those now providing 12 KW in our present transmitters. Developments along that line are already under way, and with a reasonable amount of luck, coupled with a lot of hard work, we may realize our goal in the not too distant future.

References

(1) H.M. Crosby, Five-KW Klystron UHF Television Transmitter, ELECTRONICS, June, 1951.

(2) L.O. Krause, Sidefire Helix UHF TV Transmitting Antenna, ELECTRONICS, August, 1951.

(3) McIntosh & Inglis, Washington, D.C., Engineering Report, February, 1953.

(4) N.P. Hiestand, High Power UHF Klystron Amplifier Design, UHF Symposium, IRE National Convention, 1953.

A.E. Rankin, High Power UHF Klystron Application, UHF Symposium, IRE National Convention, 1953.

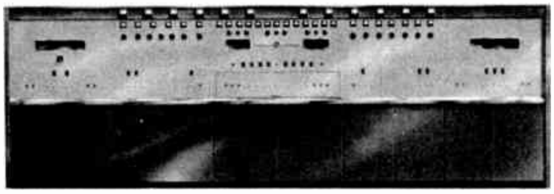


Fig. 1
General Electric 12-kw uhf television transmitter, type TT-25-A.

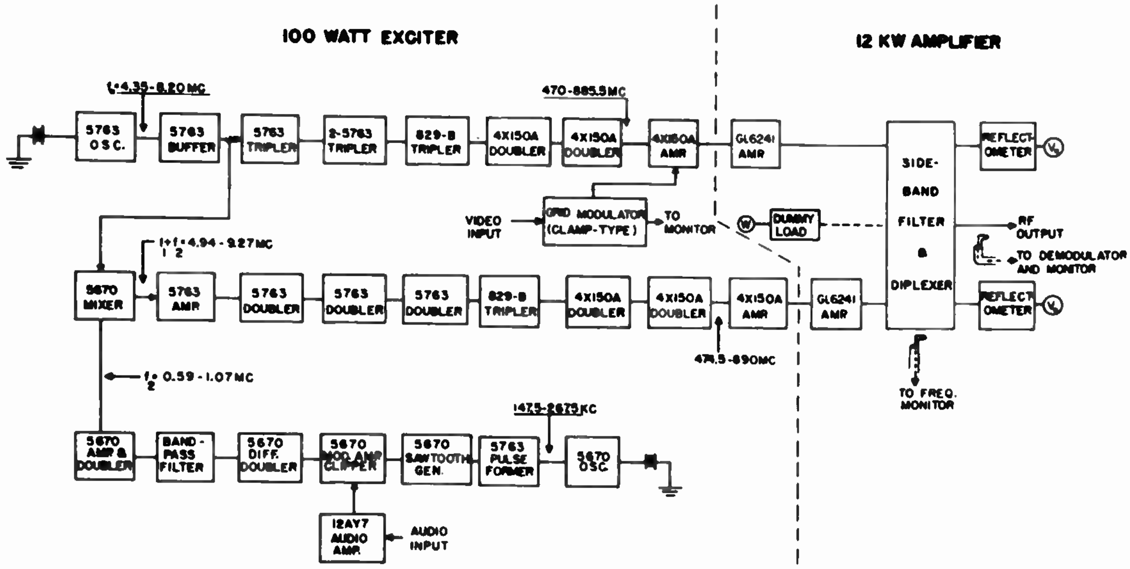


Fig. 2 - Block diagram of the visual and aural stages.

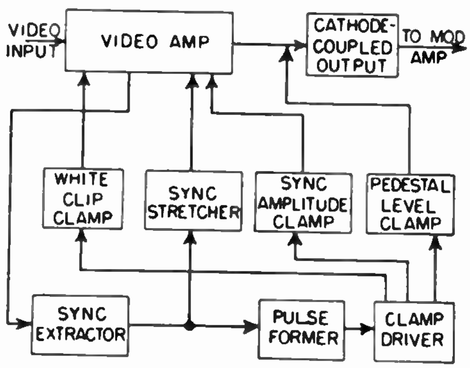


Fig. 3
Functional diagram of video modulator.

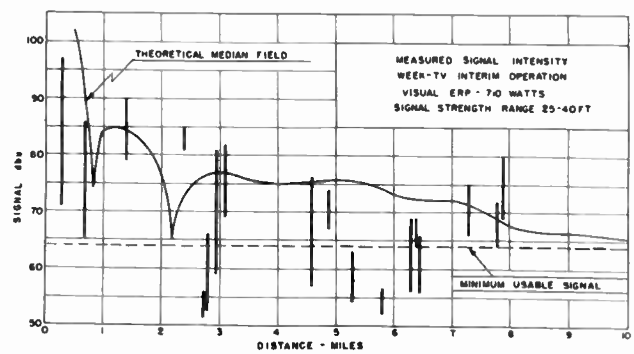


Fig. 4
Measured field intensity obtained by use of driver at typical installation.

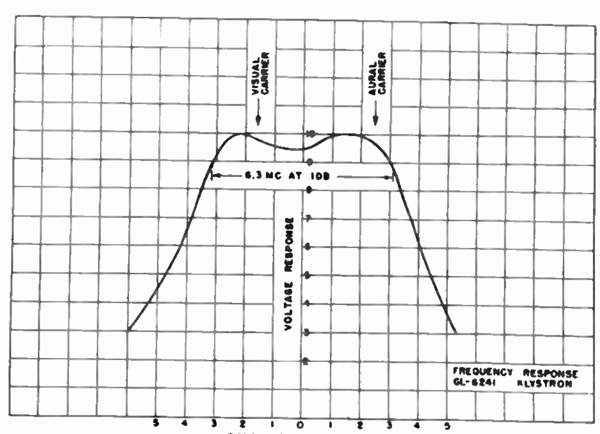


Fig. 5
Measured frequency response of 12-kw klystron adjusted for broadband operation.

Fig. 6
Klystron mounted on dolly, ready for installation in transmitter.

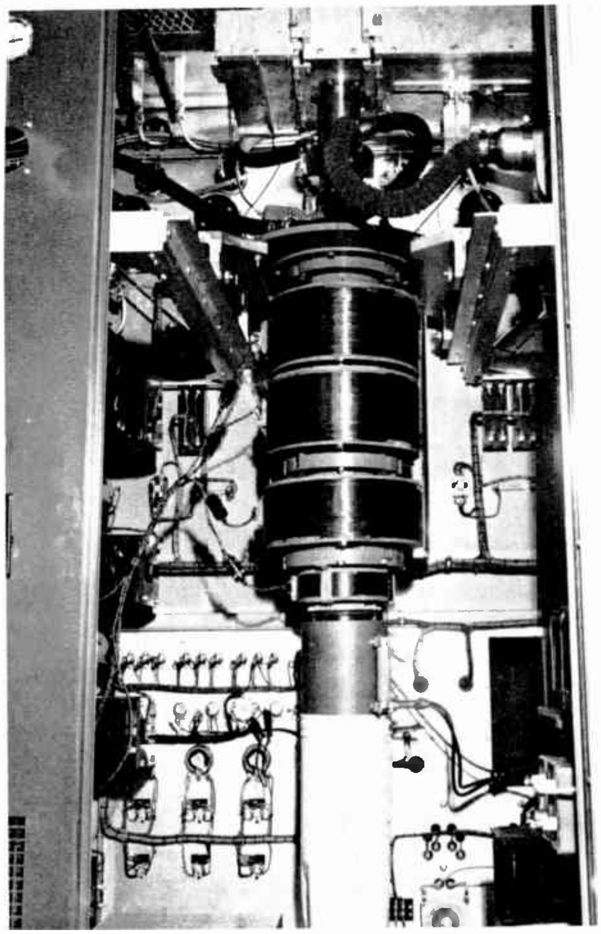


Fig. 7
Klystron in transmitter, ready for operation.

Fig. 8 - Five-bay helical uhf television antenna for channel 61.

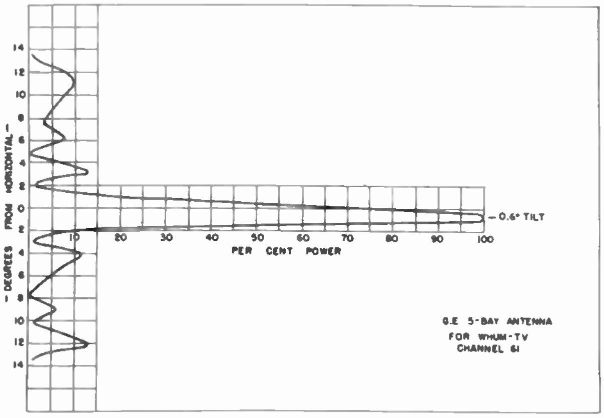


Fig. 9
Measured vertical pattern of 5-bay helical antenna, channel 61.

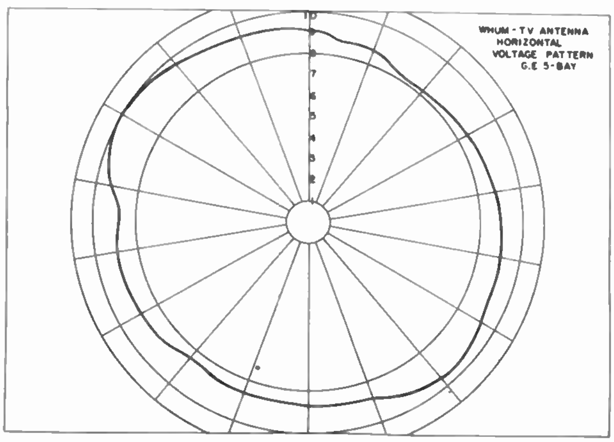


Fig. 10
Measured horizontal pattern for 5-bay helical antenna, channel 61.

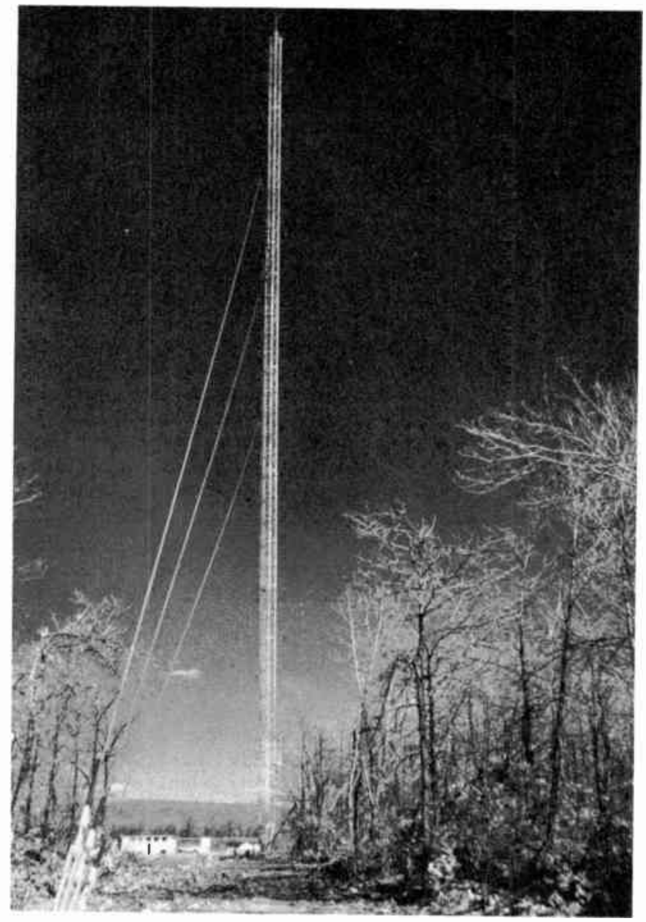


Fig. 12
Five-bay helical antenna at top of 1000-foot tower, WHUM-TV.

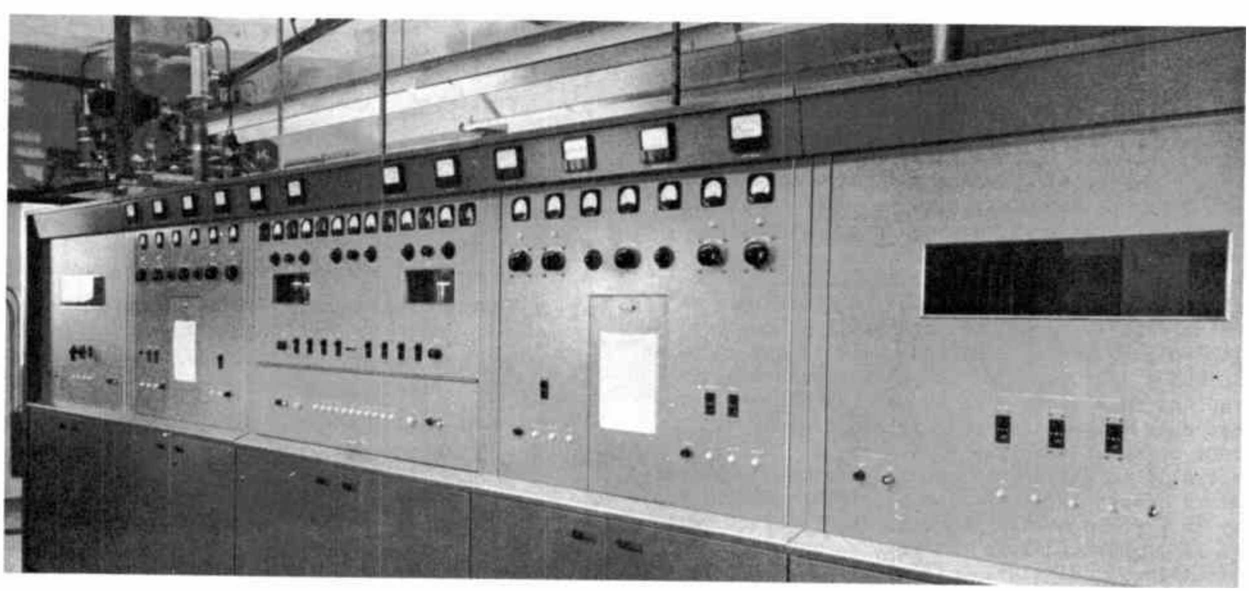


Fig. 11 - General Electric 12-kw transmitter installed at WHUM-TV.

GAIN STABLE MIXERS AND AMPLIFIERS WITH CURRENT FEEDBACK

Gail E. Boggs
National Bureau of Standards
Washington 25, D. C.

Abstract

Narrow band radio frequency amplifiers and mixers may be stabilized by negative feedback without increasing the bandwidth excessively. A couple using current feedback is described. This couple requires only a simple resistive beta circuit and may be designed such that the band-pass characteristic is largely independent of the feedback. Consideration is given to the problem of input impedance and a design procedure is outlined. A discussion of the experimental results concludes the paper.

Introduction

Inverse feedback may be used to substantially improve the gain stability of radio frequency amplifiers and mixers.[1,2,3] The use of voltage feedback results in a reduced output impedance and an increased bandwidth which is often desirable for wide band applications. Substantially improved gain stability requires considerable feedback. If a high degree of voltage feedback is applied over two stages with a resistive beta circuit, it can be shown that the gain will rise considerably at either extreme of the pass band. Montgomery[2] has presented a system using a reactive beta circuit in order to obtain an essentially uniform band pass response. This arrangement also has the advantage of yielding a greater gain-bandwidth product. When the bandwidth is increased by the application of feedback, as is the case with voltage feedback, the bandwidth becomes dependent upon the zero feedback voltage gain of the amplifier; the greater the increase in bandwidth the greater is this dependence. This is obviously an undesirable effect since the bandwidth will decrease as the amplifier tubes age even though the center frequency voltage gain remains substantially constant. Narrow band amplifiers using voltage feedback require an unusually high Q for the tank circuits in order to obtain a relatively narrow band pass characteristic.

The subject of feedback amplifiers has received extensive treatment in the literature and a considerable amount of this material has dealt with voltage feedback systems. It is not uncommon to find a discussion of the characteristics of feedback amplifiers with no mention of the fact that only voltage feedback is considered. In order to avoid possible confusion, it is felt desirable to define voltage and current feedback before proceeding. A voltage feedback amplifier may be defined as one in which the feedback voltage is proportional to the voltage across the output load. Similarly, in a current feedback amplifier, the feedback voltage is proportional to the current through the output load impedance and is independent of the load. Obviously only degenerative feedback is considered in this paper.

It can be readily shown that current feedback results in an increase in the output impedance of an amplifier and hence in many cases a reduced bandwidth. In amplifiers using pentode tubes the output impedance of the amplifier will generally be sufficiently high without feedback to present negligible loading of the tuned output circuit. Hence, if the bandwidth of a current feedback amplifier is determined by the output circuit, the bandpass characteristic will be essentially independent of both the feedback and the zero feedback voltage gain.

This paper will be concerned with radio frequency amplifiers and mixers arranged in a feedback couple with a simple resistive beta circuit. The narrow band amplifier and mixer couples to be discussed use current feedback and the bandpass characteristic may, by proper design, be largely independent of the feedback.

Amplifier Theory

Referring to Figure 1, it is apparent that the feedback voltage for this system is derived from the plate current of the second stage. As a result, the system may be considered as a current feedback couple.

The nomenclature to be used in the following brief analysis is given below. It is recognized that many of the quantities listed are properly complex but since we are interested primarily in the center frequency considerations where phase shifts are generally very nearly integral multiples of pi, these quantities are listed as reals.

A = voltage gain of couple without feedback

A_1 = voltage gain of first stage

A_2 = voltage gain of second stage

A_f = voltage gain of couple with feedback

A_n = normalized voltage gain of couple

a_1 = Fourier series coefficient proportional to conversion transconductance

b_o = Fourier series coefficient proportional to the average transconductance

B = return difference

C_1 = grid to cathode capacitance of first stage

C_2 = cathode to ground capacitance of first stage

e_{g1} = grid to cathode variational voltage of first stage

e_{g2} = grid to cathode variational voltage of second stage

e_i = variational input voltage

e_k = cathode variational voltage of first stage

e_o = variational output voltage

E_i = maximum value of input voltage

E_o = maximum value of output voltage

f_o = center frequency of couple in cycles per second

Δf = bandwidth in cycles per second

g_m = tube transconductance in mhos

g_c = conversion transconductance in mhos

k_o = constant proportional to the average transconductance

k_1 = constant proportional to the conversion transconductance

Q_k = Q of cathode circuit

Q_L = Q of plate load impedance

R = cathode resistance in ohms

R_L = effective plate load resistance in ohms

$u = \frac{f}{f_o} - \frac{f_o}{f} \cong \frac{\Delta f}{f_o}$

$x = Qu$

y_i = input admittance in mhos

Z_L = plate load impedance

ω_1 = angular velocity of signal voltage

ω_2 = angular velocity of oscillator voltage

It will be assumed that identical pentodes are used in the couple to be analyzed and that the load impedance is the same for each stage.

Figure 2 gives the equivalent circuit of the couple. The grid to cathode and cathode to ground capacitors are shown since input admittance considerations will be considered later. For gain determinations, these capacities will be neglected and the equivalent circuit reduced to three node pairs.

Writing the three nodal equations and placing them in matrix form gives

$$\begin{vmatrix} g_m e_i \\ -g_m e_i \\ 0 \end{vmatrix} = \begin{vmatrix} (g_m + \frac{1}{R} + \frac{1}{Z_L}) & 0 & -\frac{1}{Z_L} \\ -g_m & \frac{1}{Z_L} & 0 \\ -\frac{1}{Z_L} & g_m & \frac{1}{Z_L} \end{vmatrix} \begin{vmatrix} e_k \\ e_{g2} \\ e_o \end{vmatrix} \qquad (1)$$

where the input voltage may be expressed as

$$e_i = e_{g1} + e_k. \qquad (2)$$

Solving (1) for e_o and simplifying yields

$$e_o = \frac{\left[R(1 + g_m Z_L) + g_m Z_L^2\right] g_m e_i}{1 + g_m R + g_m^2 Z_L R}. \qquad (3)$$

For high gain, narrow band amplifiers $R \ll Z_L$, hence

$$e_o \cong \frac{g_m^2 Z_L^2 e_i}{1 + g_m^2 Z_L R}. \qquad (4)$$

Now by definition, the gain with feedback is, $A_f = \frac{e_o}{e_i}$ and the zero feedback gain (A)

$$A \cong g_m^2 Z_L^2.$$

Substituting in Equation (4)

$$A_f \cong \frac{A}{1 + \frac{AR}{Z_L}}. \qquad (5)$$

Now by definition,

$$B = \frac{AR}{Z_L} \qquad (6)$$

and

$$A_f \cong \frac{A}{1 + B}, \qquad (7)$$

which is a conventional form of feedback equation. It is worthy of note that Z_L appearing in the defining equation for B represents the load impedance of the second stage. A change in this impedance will result in a linear change in A_f for large values of B.

Bandwidth Considerations

In many applications of this type of amplifier it is desirable to have the bandwidth independent of the feedback. Therefore, the effect of the interstage tank circuit is considered briefly. For simplicity consider a resistive load impedance R_L on the second stage. Equation (5) is modified by this change to

$$\underline{A}_f = \frac{\underline{A}}{1 + \frac{\underline{A}R}{R_L}}. \qquad (8)$$

The gain, $\underline{A}_f$ and $\underline{A}$ must be considered to be complex for the purpose of this discussion and the subscript 0 denotes center frequency gain. The complex gain without feedback can be shown to be

$$\underline{A} = \frac{A_o}{1 + jQu} \qquad (9)$$

since only one anti-resonant impedance network is present. Substituting (9) in (8) yields

$$\underline{A}_f = \frac{A_{fo}}{1 + \frac{jQu}{1 + \frac{A_o R}{R_L}}}. \qquad (10)$$

From inspection of Equation (10) it is apparent that the effective Q of the interstage circuit is reduced by the feedback factor $1 + B$, as is the case with a single stage, voltage feedback amplifier. If the resistive load R_L is now placed on the first stage and the complex impedance Z_L on the second stage, Equation (5) is still valid but $\mathbf{A}$ is now $g_m^2 R_L Z_L$. Following a similar procedure yields

$$A_f = \frac{A_{fo}}{1 + jQu}. \qquad (11)$$

Obviously the bandpass characteristic is unaffected by the application of feedback provided pentode tubes are used. For amplifiers using similar tank circuits and feedback factors of ten or greater, the bandpass characteristic is essentially that of the output circuit. The bandwidth of such a system will be largely independent of the zero feedback voltage gain until the gain and hence the feedback factor, have fallen to a level such that the interstage network begins to make an appreciable contribution to the overall selectivity. In addition, this system permits the use of the feedback couple, with many of the attendant advantages, without the necessity of using a critical reactive beta circuit. Obviously, if the interstage Q is less than that of the output circuit, the bandwidth will remain independent of voltage gain for lower values of $1 + B$.

Input Impedance

It is well known[4] that with a cathode follower, the real part of the input admittance may be negative. Since the first stage of this couple operates with an essentially unbypassed cathode, this stage may have a negative input conductance. Fortunately, for high gain amplifiers operating at frequencies of 10 Mc or less and with nominal values of feedback, the magnitude of the negative input conductance is sufficiently small that its effect may be neglected. With couples requiring a large value of cathode resistance or operating at high frequencies, the negative input admittance may result in oscillation at or near the signal frequency when the couple is operated with a high impedance input. This discussion will be limited to input admittance effects due to the unbypassed cathode of the first stage.

The admittance equation is given in order to allow the engineer to check a particular design. Referring again to Figure 2, the nodal equations may be written and the matrix solved for the real part of the input admittance y_i. Substituting for the plate load impedance

$$Z_L = \frac{R_L}{1 + jx}$$

and assuming that $g_m Z_L \gg 1$, gives

$$R_e\, y_i \cong \frac{\omega^2 C_1{}^2 R x^2 - \omega C_1 B x + \omega^2 C_1 R \left[C_1 - C_2 B\right]}{x^2 - 2\omega(C_1 + C_2)\, RBx + (1 + B)^2}. \qquad (12)$$

Equation (12) may be differentiated with respect to x for a particular case, to find the value of x at which $R_e\, y_i$ is a maximum or minimum. A negative value of conductance may occur with a positive value of x. It is this condition which may lead to oscillation in certain cases with a tuned input circuit. For center frequency ($x = 0$) operation the conductance will be negative when $B > \frac{C_1}{C_2}$. At high frequencies or large values of R, the input resistance may become sufficiently negative to substantially increase the Q of a tuned circuit driving the couple. If this occurs, the system bandwidth will be a function of the feedback factor. Fortunately, the cathode to ground capacitance may be reduced at center frequency by shunt resonating the cathode. Then the input conductance will always be slightly positive at the center frequency. If $Q_K \ll Q_L$, the cathode circuit may be considered essentially resistive at the value of x at which the input conductance is a maximum or minimum. Hence the resonant cathode has very little effect upon the admittance occuring at the values of x mentioned above. For this condition C_2 may be taken as zero in Equation (12). From the foregoing it is apparent a negative input conductance may occur at two regions, one at the operating frequency and the other at a slightly higher frequency under the proper conditions.

In cases where the desired feedback is greater than 20 db or the operating frequency is higher than 10 Mc the problem of input admittance may assume serious proportions. Hence it is necessary for the design engineer to recognize these conditions. For example, at 30 Mc a couple may be designed with a resonant cathode circuit such that the input conductance is positive at the operating frequency but the couple could conceivably oscillate at a higher frequency if the driving impedance is of the correct form.

Mixer Couple

If the first tube of the couple is replaced by a mixer such as the 6SA7, the difference frequency voltage may be used as negative feedback to stabilize the mixer and amplifier operation.[3] As a result of the oscillator excitation, the mixer transconductance is driven from maximum on one-half cycle to zero on the other. Hence the instantaneous transconductance is a function of the oscillator voltage and may be represented by the Fourier series.

$$g_{m1} = b_o + a_1 \sin \omega_2 t + a_2 \sin 2\omega_2 t + \ldots \qquad (13)$$

For the brief analysis which follows only the first two terms of Equation (13) need be considered.

Since the plate circuit of the mixer is tuned to the difference frequency $(\omega_2 - \omega_1)$, only this component will be considered in the output voltage e_o, thus

$$e_o = E_o \cos (\omega_2 - \omega_1)t. \qquad (14)$$

Let the signal input voltage at angular frequency ω_1 be

$$e_i = E_i \sin \omega_1 t. \tag{15}$$

Assuming identical load impedances Z_L for each stage, Equation (4) may be modified to read

$$e_o = \frac{g_{m1}\, g_{m2}\, Z_L^2\, e_i}{1 + g_{m1}\, g_{m2}\, RZ_L}. \tag{16}$$

Substituting the expressions given in Equations (13), (14) and (15) in the above, expanding, and neglecting all terms which do not contain the difference frequency, yields

$$E_o = \frac{a_1\, g_{m2}\, E_i\, Z_L^2}{2(1 + b_o\, g_{m2}\, RZ_L)}. \tag{17}$$

For simplification let

$$b_o = k_o \frac{\hat{g}_{m1}}{\pi} \quad \text{and} \quad a_1 = k_1 \frac{\hat{g}_{m1}}{\pi}$$

where k_o and k_1 are constants determined by the switching function of the mixer tube selected.

Writing Equation (17) in terms of gain and substituting for b_o and a_1

$$A_f = \frac{A}{1 + \dfrac{2k_o\, RA}{k_1 Z_L}}. \tag{18}$$

The return difference may be defined by

$$B = \frac{2k_o\, RA}{k_1 Z_L} \tag{19}$$

and Equation (15) reduces to

$$A_f = \frac{A}{1 + B}.$$

Obviously, for large values of B, the gain is relatively independent of the transconductance of either tube. It should be noted that k_o and k_1 are subject to variation due to changes in the shape of the switching function which may result from a change in oscillator voltage as well as other causes. While this may be considered as a limiting factor for stability improvement, experimental results indicate that for pentagrid tubes, these constants tend to change together. With a high degree of feedback, inspection of Equation (18) indicates little change in A_f with changes in the value of the constants, provided that the changes of k_o and k_1 are in the same direction and of like percentage. Typical values of the constants for a 6SB7-Y are $k_o = 1.2$ and $k_1 = 1.75$. The author has observed that larger variations occur in the values of the constants of pentode mixers.

Since the conversion transconductance is less than the amplifier transconductance, a larger cathode resistor will be required to provide the degree of feedback obtained with an amplifier couple. At the signal frequency, the input admittance may be negative due to the unbypassed cathode. In fact, in some instances, Colpitts oscillation may take place in the signal circuit. The deleterious effects of a negative input impedance at the signal frequency can usually be minimized by a suitable filter on the mixer cathode or by a low impedance signal grid circuit.

In the case of amplifier couples it was observed that the input impedance was negative at a frequency slightly higher than that at which the amplifier was designed to operate. Similarly, with mixer couples, particularly those with an i-f in the order of 10 Mc, oscillation may conceivably occur when the input circuit is tuned slightly above the i-f. While these conditions would rarely be realized, the danger of oscillation can be minimized by the use of low impedance signal circuits.

Design Procedure

For the procedure which follows it will be assumed that the load impedances are identical. The design engineer may use this procedure with suitable modification for both the mixer and amplifier couples. Other variations may be readily obtained to fit a particular case. It will be further assumed that the feedback factor 1 + B, is large enough to permit the assumption that the bandwidth is determined only by the output circuit. No consideration is given to the problem of input admittance.

Given: Δf, f_o, A_f, $1 + B$, g_{m2} and g_{m1} or g_c.

(a) Calculate $A = (1 + B)A_f$.

(b) For mixer operation determine values of k_o and k_1 for the tube under consideration. These constants may be determined by graphical integration of the switching function.[3] If this is not expedient, assume $k_1 = 1.5$ and $k_o = 1.0$ and continue solution; suitable corrections must then be made experimentally when the unit is under test.

(c) Calculate the gain of each stage;

for amplifiers $A_1 = A_2 = \sqrt{A}$,

for mixers let $k = \dfrac{g_{m2}}{g_c}$,

$$A_2 = \sqrt{kA}$$

and $A_1 = \dfrac{A}{A_2}$.

(d) Calculate $R_L = \dfrac{A_2}{g_{m2}}$.

(e) Calculate $Q = \dfrac{f_o}{\Delta f}$.

(f) Calculate $C = \frac{Q}{\omega R_L}$.

(g) Calculate R;

for amplifiers $R \cong \frac{BR_L}{A}$;

for mixers $R = \frac{k_1 \; BR_L}{2k_0 A}$.

Alignment Procedure

To remove feedback for alignment and the measurement of gain, the feedback line should be broken and the cathode of the first tube adequately bypassed. The output circuit must now be returned to ground through a value of resistance approximately equal to R. The plate circuits may now be aligned to the correct frequency and the gain measured. When feedback is reapplied to the amplifier, the output circuit only should be readjusted for maximum at the same alignment frequency. If the peak of the bandpass response does not occur at the alignment frequency, the output circuit should be adjusted for this condition. A slight response shift may occur in some units due to extraneous phase shift of the feedback voltage.

It is worthy of mention that steps must be taken to minimize regeneration since an amount which normally would be tolerated may result in violent performance deviations in a feedback couple.

Experimental Results

For amplifiers operating at low frequencies and with relatively small values of 1 + B it is often possible to replace the interstage tank circuit with a resistor without encountering excessive phase shift. An experimental amplifier of this type was designed for operation at 50 kc with a gain of 400. The bandwidth of the amplifier was 0.5 kc. A supply voltage stability curve with and without feedback is shown in Figure 3 with normalized gain A_n plotted against plate supply voltage. The substantial stability improvement is immediately apparent. A 450 kc amplifier, with each plate circuit tuned, was also tested experimentally with the results shown in Figure 4. In this case higher gain was required resulting in a feedback factor of only 23 db. The bandwidth under each condition of operation is noted on the curve. The type of amplifier described in this paper has been operated successfully at 15 Mc and can probably be used at higher frequencies if suitable precautions are taken to restrict extraneous phase shift of the feedback voltage. Obviously, the input admittance problems discussed earlier must be given full consideration at frequencies in the order of 50 Mc.

The use of this system for mixer couples results in operation very similar to that obtained with the amplifiers just described. The stability improvement of mixer couples is, in general, not as great as that obtained in amplifiers due to the variation in the constants k_0 and k_1 described earlier. Satisfactory operation has been obtained with a 6AK5 pentode mixer operating at a 50 Mc signal frequency and a 5 Mc i-f frequency. The i-f stage of this couple was also a 6AK5.

Conclusion

It has been shown that current feedback may be used to stabilize radio frequency amplifier and mixer couples without the use of a reactive beta circuit. The use of this system will permit the design of a narrow band amplifier whose bandwidth is essentially independent of the feedback.

Since the output circuit is essentially unaffected by the application of feedback to an amplifier using pentode tubes, a complex filter designed to provide the desired band pass characteristic may be readily employed. It should be noted that the impedance of any paths from the output circuit to ground must be much larger than the beta circuit impedance.

References

1. D. G. Tucker, "Frequency Changers and Amplifiers with Constant Gain", Proc. I.R.E., 37, pp. 1324-1327, November 1949.

2. G. F. Montgomery, "Intermediate Frequency Gain Stabilization with Inverse Feedback", Proc. I.R.E., 38, pp. 662-667, June 1950.

3. G. E. Boggs, "Improvement in Gain Stability of the Superheterodyne Mixer through the Application of Negative Feedback", Proc. I.R.E. 40, pp. 202-207, February 1952.

4. F. D. Clapp, "Some Aspects of Cathode-Follower Design at Radio Frequencies", Proc. I.R.E., 37, pp. 932-937, August 1949.

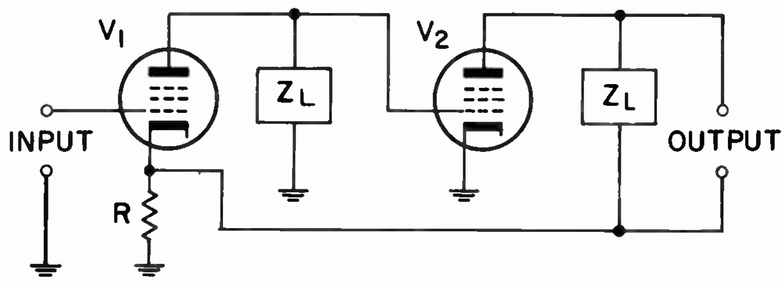


Fig. 1
Schematic of current feed feedback couple.

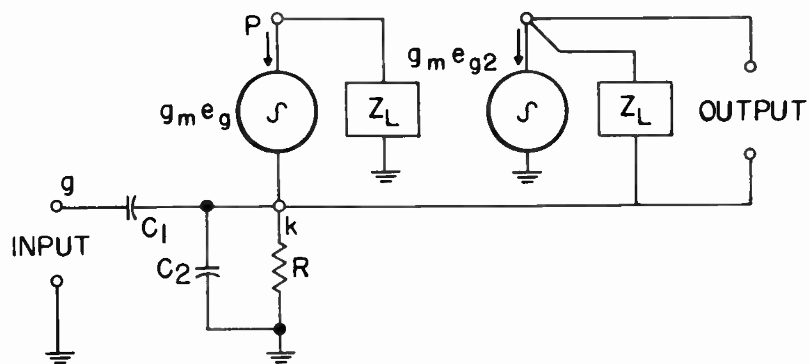


Fig. 2
Equivalent circuit of current feedback couple.

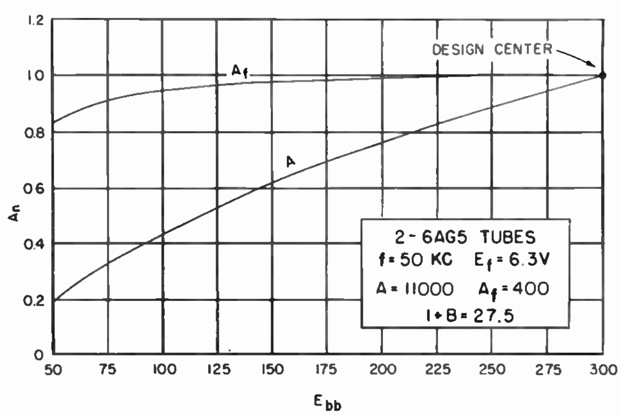


Fig. 3
Gain stability of 50-kc amplifier couple.

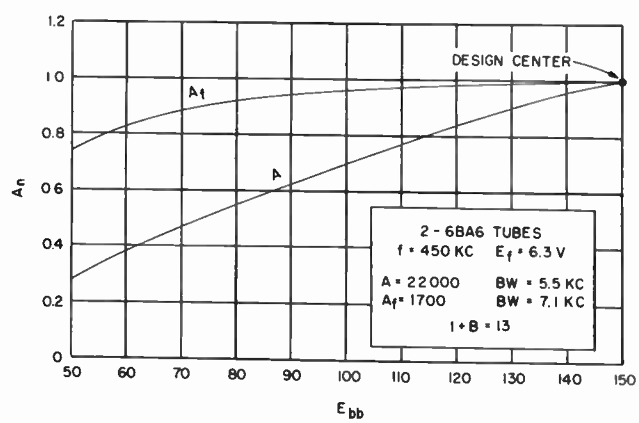


Fig. 4
Gain stability of 450-kc amplifier couple.

VIDEO AMPLIFIERS WITH INSTANTANEOUS AUTOMATIC GAIN CONTROL

William E. Ayer
Electronics Research Laboratory
Stanford University

ABSTRACT

Circuits are described which allow essentially complete control of the output/input amplitude characteristic of a multi-stage video amplifier handling both positive and negative input signals. The incremental gain of each stage is determined instantaneously by the signal current through the tube. Germanium diodes are employed in the cathode circuits to introduce degeneration for signals above a predescribed amplitude. The gain reduction achieved in this manner lasts only as long as the signal, so that recovery time is not adversely affected.

Design relations are given for single stages, and for paired stages employing negative voltage feedback. This latter configuration, while more complicated from the design standpoint than individual stages, permits a wide choice of gain characteristics (including zero or negative gain for large signals), and is much less susceptible to power supply ripple.

The process of cascading several stages or pairs of stages to realize a desired output/input characteristic is discussed. Some of the problems that may be encountered with several cascaded stages are indicated, and solutions outlined.

Using the techniques described, amplifiers of 100 decibel dynamic range are readily constructed. As a typical example, amplifiers of this dynamic range and 100 decibel small-signal gain have been built using three dual triodes. The over-all bandwidth for this case was approximately 500 kc. With the gain characteristic adjusted so that the output signal is proportional to the logarithm of the input signal, such an amplifier may be used to provide a crystal-video receiver with cathode-ray tube display without resorting to a manual gain control.

The ability of these amplifiers to handle signals of both polarities simultaneously greatly reduces the overshoot problem usually encountered in pulse amplifiers. When pulse-width fidelity is required, an overshoot may be introduced near the front end to prevent the pulse stretching which otherwise occurs for strong signals.

An important feature of the circuit arrangements used is the stability with respect to changes of diode and tube characteristics with age and temperature.

1. INTRODUCTION

In many microwave receiving systems, the incoming signal is detected directly, and then amplified to provide a usable output signal. When the received signals are of a pulsed nature, a video amplifier must be used following the detector to reproduce the envelope of the input signal. Arrangements of this sort are often called crystal-video systems.

Work at this laboratory with crystal-video systems has prompted an investigation of video amplifiers possessing special characteristics. The principal results of this investigation are reported in this paper.

Prior to the current investigation, a considerable amount of work was done by others along the same lines of attack. Mr. J. R. Wilkerson of Airborne Instruments Laboratory was the first, to the author's knowledge, to work with the diode degeneration scheme, a form of which is employed here. Mr. W. R. Rambo (formerly of Airborne Instruments Laboratory and now at Stanford) made major contributions to the development. The chief work of the author has been in extending the circuitry to allow bipolar operation, and to develop design relations.

2. BASIC CIRCUIT ARRANGEMENT

The circuit of a single-stage amplifier is shown in Fig. 2.1. The quiescent currents through the diodes are established by the resistors R_{t1} and R_{t2}, which are large in comparison with R_k. The condenser C_k is large, so that it performs essentially as a battery for pulse amplification. The grid bias for the tube is obtained from the divider R_{b1} and R_{b2}.

For small signals the cathode-to-ground impedance is typically of the order of a few hundred ohms, as determined by the sum of the dynamic resistances of the diodes. R_k is always several thousand ohms, so that it has little effect upon the impedance to ground when the diodes are conducting. The equivalent circuit is shown in Fig. 2.2. Under these conditions the stage exhibits maximum gain. As the size of the incoming signal is increased, the current through the tube changes. For positive input signals, the current through D_2 increases, while the current through D_1 decreases. When the grid signal is sufficiently large, conduction through D_1 will cease, and for all larger signals the equivalent circuit will be that shown in Fig. 2.3. For large negative input signals, the diode D_2 will be out of conduction, and the operation may again be understood from Fig. 2.3.

For small signals, then, the cathode degeneration is small, and the incremental gain is large. As the input level is raised, conduction through one or the other of the diodes stops, the cathode degeneration increases to that provided by the cathode resistor R_k, and the incremental gain of the stage drops. The shape of the output/input characteristic of a single stage is shown in Fig. 2.4. The upper curve illustrates the behavior that would be expected on the basis of "ideal" diodes (constant resistance for all forward voltages and infinite resistance for all back voltages). The incremental gain would be large and constant up to the transition point, and then drop abruptly to a lower constant value for larger signals.

In the practical case of high-conduction germanium diodes, the dynamic resistance of the series diodes increases as the transition point is approached, so that the incremental gain falls off before the actual transition occurs. The shape of the output/input curve under these conditions is illustrated by the lower curve of Fig. 2.4.

In designing stages for specific applications, considerable freedom is available. Thus, the small-signal gain may be varied over wide limits. Similarly, the output voltage at which the gain transition occurs, and the incremental gain for large signals, may be adjusted to meet various requirements. Stages of different characteristics may be cascaded to provide a desired over-all characteristic. For example, it is frequently desired that the output voltage of the amplifier be approximately proportional to the logarithm of the input voltage. This characteristic may be closely approximated by using a fairly large number of identical stages, so that there are many transitions or "knees" in the output/input curve. In view of the gradual nature of the transitions realized with available germanium diodes, the approximation is better than one based upon straight line segments. The logarithmic characteristic is obviously one of many which may be achieved. Several others will be discussed later.

The control of gain characteristic for both positive and negative input signals which is provided by the circuits described is important in many practical applications wherein pulses of only one polarity are to be amplified. First, it is often very convenient to introduce a short time constant somewhere in the amplifier to reduce stray low-frequency signals from power supplies, etc. If this is done in an amplifier which has large dynamic range for signals of only one polarity, the overshoot produced by such a short time constant can easily paralyze the system, resulting in a long recovery time. If the gain is controlled for signals of both polarities, this problem is greatly simplified.

In some applications it is necessary to preserve pulse width. When large signals are applied to an amplifier with IAGC, the weaker trailing portion of a received signal will be amplified more than the main pulse, so that very large amounts of stretching may occur. This may be circumvented with the bipolar amplifier by purposely introducing an overshoot early in the amplifier. The end of the pulse is thus defined by the base-line crossover point, and will be faithfully reproduced at the output.

Design Procedure

The design of a stage to provide a given output/input variation may be carried out in a fairly direct manner. Some cut-and-try is generally necessary, but convergence to the proper component values is rapid, once a reasonably good approximation has been made.

Basic specifications are the small-signal gain, the output voltage at transition, and the large signal gain. It is assumed that the operating point of the tube has been chosen, and that the gain characteristic specified is within the capabilities of the tube.

First, a reasonable value for the plate load resistor should be assumed. A satisfactory value for the first trial may be found by computing the plate load required to realize the desired small-signal gain without cathode degeneration, and then doubling this answer to allow for the dynamic resistance of the cathode diodes under quiescent conditions. In the typical case of triode amplifiers, this estimated plate load is given by[1]

$$R_L = 2\,\frac{A_1 R_p}{(\mu - A_1)} \qquad (2.1)$$

where A_1 is the desired small-signal gain for the stage. R_p is the dynamic plate resistance, and μ the amplification factor, evaluated at the operating point.

Next, the transition point may be approximated. This is done by finding the change in tube current required through R_L to give the desired output transition voltage. Since the cathode resistor R_k will generally be large in comparison with the dynamic resistance of the diodes, most of the tube signal current will come from the diode circuit. Thus, the output voltage at the time one of the diodes goes out of conduction will be given to a good approximation by

$$V_t \approx i_o R_L \qquad (2.2)$$

where i_o is the quiescent diode current. When greater accuracy is desired in determining the output transition voltage, the expression

$$V_t = i_o R_L + v_o R_L / R_k \qquad (2.3)$$

may be used, where v_o is the quiescent diode voltage drop. In typical applications involving high-mu tubes, the large-signal incremental gain is given approximately by

$$A_2 = R_L / R_k \qquad (2.4)$$

so that (2.3) and (2.4) may be used to complete the first approximate evaluation of R_L, R_k, and i_o.

When these tentative values have been determined, the final design may be started. First, the dynamic resistance of the series diodes for the quiescent current i_o should be found. This is simply twice the resistance for a single diode. Fig. 2.5 shows the measured dynamic resistance and voltage drop v_o as functions of i_o for 1N56 germanium diodes.

At this point the plate load resistor may be recalculated. An equivalent plate resistance of the tube with cathode degeneration may be found from

$$R_p' = R_p + (1 + \mu)\, R \qquad (2.5)$$

where R_p is the actual plate resistance, and R

[1]The derivations of the equations of this section are indicated in Appendix A.

designates the resistance between cathode and ground. In this case R would be the total dynamic resistance of the diodes with the quiescent current found from (2.2) or (2.3). The corrected plate load resistance value may now be found from

$$R_L = \frac{A_1 R'_p}{(\mu - A_1)} \tag{2.6}$$

where R'_p is the effective plate resistance obtained from (2.5) above.

With the new plate load a slightly different i_o will be required to re-establish the proper transition point, which, in turn will result in a different dynamic diode resistance. This "closing in" on the proper values for R_L and i_o generally goes very rapidly, however.

When R_L and i_o have been established, the small-signal gain and the transition point are fixed, and it remains only to obtain the correct value for the incremental gain A_2. The circuit element used to control A_2 is the cathode resistor R_k. Its value may be found from

$$R_k = \frac{R_L(\mu - A_2) - R_p A_2}{(1 + \mu) A_2} \tag{2.7}$$

To complete the design, values must be found for R_{t1} and R_{t2}, and R_{b2}, R_g, C_g, and C_k. Since the quiescent diode current i_o has been determined and the operating tube current i_p has been chosen, the zero-signal potential of the cathode is given by

$$V_{ko} = (i_p - i_o) R_k \tag{2.8}$$

The potential at the junction of the two diodes will be V_{ko} less the drop across D_2 for the current i_o. This latter voltage is small for high-conduction germanium diodes--typically of the order of 0.3 volts for 1N56's. The voltage drop across D_1 will be equal to that across D_2, but of opposite polarity, so that the voltage at the lower end of R_{t1} will be equal to V_{ko}. Since R_{t1} is to carry a current i_o, it may be found from

$$R_{t1} = \frac{(V_b - V_{ko})}{i_o} \tag{2.9}$$

The size of R_{t2} may be determined in a similar way. The voltage at its upper end will be $V_{ko} - v_o$, where v_o is the quiescent drop across diode D_2. Since R_{t2} is to carry the quiescent currents of both diodes, the proper value is given by

$$R_{t2} = \frac{(V_c + V_{ko} - v_o)}{2i_o} \tag{2.10}$$

The grid-circuit components R_{b1}, R_{b2}, R_g, and C_g are readily determined. First, it is usually convenient to make the sum of R_{b1} and R_{b2} equal to or greater than R_{t1}. Since the voltage across C_k is generally small in comparison with the positive supply voltage, this choice simplifies design in that no appreciable fraction of the current through R_{t1} will be diverted from D_1. It also helps to maximize the discharge time-constant of C_k. The division of resistance between R_{b1} and R_{b2} is determined by the operating grid bias chosen for the tube, and the potential V_{ko} existing at the junction of R_{b1}, R_{t1}, D_1, and C_k. If the desired grid bias is represented by V_g, the values of the bias resistors may be found from

$$\frac{R_{b1}}{(R_{b1} + R_{b2})} = \frac{V_g}{V_{ko}} \tag{2.11}$$

Since the sum of the two resistances is assumed to have been established, the individual values may be found from (2.11).

The values for R_g and C_g are determined by the usual considerations involved in pulse amplifier design, and need not be discussed here.

The choice of the condenser C_k is governed by the consideration that the voltage at its upper end should not vary appreciably in the presence of strong signals. When one of the diodes is out of conduction, a current i_o flows into or out of the condenser, tending to change its voltage. The magnitude of this voltage change under specified conditions is easily found from

$$\Delta V = \frac{i_o t}{C_k} \tag{2.12}$$

where t is the length of the pulse. The manner in which a voltage variation across C_k affects the operation of the amplifier is rather complex, but the primary result is a change in cathode potential. The quiescent cathode voltage following a strong positive input signal, for example, is increased by a somewhat smaller amount than the condenser voltage change given by (2.12). The result at the output of the stage is thus an overshoot. In practice, C_k is generally made as large as convenient on the basis of available condensers, typically of the order of tens of microfarads.

3. A TYPICAL STAGE DESIGN

The design procedure just described may be illustrated by a practical example. Two stages of an amplifier are shown in Fig. 2.6. The plate load resistor R_L was chosen to provide an overall bandwidth of a five-stage amplifier of approximately 500 kc, corresponding to a stage bandwidth of 1.3 Mc. A small-signal gain of about 85 db for the five stages was required. On this basis, the voltage gain per stage (A_1) would be 7.1. The output voltage at the gain transition was to be approximately 3.5 volts, and the large-signal gain (A_2) approximately 1.

Using (2.3), (2.4), (2.5), and (2.6), and the diode dynamic resistance curve of Fig. 2.5, it was found that the desired performance could be obtained with a plate load resistor of 4.7k and a quiescent diode current of 0.7 ma. A tube current of 4 ma was used, giving a plate resistance of 18k and an amplification factor of 60. The required grid bias for this operating point was 2.8 volts. Substituting 2(150) = 300 ohms from Fig. 2.5 in (2.5) and solving (2.6), the required plate load was found to be 4.9k. The effective interstage capacitance was estimated to be 25 micromicrofarads, giving an upper 3 db point of 1.35 Mc.

The cathode resistor R_k was next calculated using (2.7).

The values of the remaining circuit elements may be found easily. The quiescent cathode voltage from (2.8) is

$$(4.0 - 0.7)\text{ma}(4.7)\text{k} = 15.5 \text{ volts.}$$

R_{t1} will then be, from (2.9)

$$(250 - 15.5)\text{v}/0.7\text{ma} = 335\text{k.}$$

The other diode bias resistor R_{t2} will be, from (2.10)

$$(250 - 15.5 - .3)\text{v}/1.4\text{ma} = 189\text{k.}$$

The grid bias divider composed of R_{b1} and R_{b2} may be determined using (2.11). This relation may be rewritten as

$$R_{b1} = \frac{V_g/V_{ko}}{R_{b2} - V_g/V_{ko}} \quad \text{megohms}$$

with R_{b2} also in megohms. This form is sometimes more convenient, as it allows R_{b2} to be chosen as a standard value. In the example, R_{b2} was assumed to be 1 megohm. For a grid bias voltage of 2.8 volts and a quiescent cathode voltage of 15.5 volts, R_{b1} is found to be 220k.

The cathode condenser C_k was made 30 microfarads. This value was chosen partly because dual units of convenient size were available. From (2.12) the change in condenser voltage during a 5-microsecond pulse (of sufficient amplitude to drive one of the diodes out of conduction) would be

$$\frac{0.7 \text{ ma } 5\cdot10^{-6}\text{sec}}{30\cdot10^{-6} \text{ farads}} = 116 \text{ microvolts.}$$

In practice it has been found that voltage changes up to about 500 microvolts produced no appreciable degradation of the performance of high-gain multistage amplifiers, so that the 30 microfarad value for C_k is adequate in most cases. Trouble may be encountered in some applications when an overshoot is purposely introduced early in the amplifier. Under such conditions, the overshoot from a strong signal would be stretched in passing through the amplifier, so that the "off-time" of some of the later diodes might be many times greater than the length of the pulses being amplified. The important time in determining voltage change on C_k would then be the duration of the overshoot, and this could be very large. In such cases, the overshoot should be made as small as possible, consistent with the requirements to be met, and C_k should be made as large as possible. Additional improvement can sometimes be realized by limiting all signals above a prescribed amplitude before entering the amplifier.

The choice of the interstage coupling components is based upon normal pulse amplifier considerations of desired low-frequency response, allowable grid-circuit resistance, and physical size of the coupling condenser. Values of 1 megohm for R_g and 0.01 microfarads for C_g have been found satisfactory for most applications which required pulse lengths of the order of 0.5 to 5 microseconds to be handled.

Fig. 2.7 shows a series of oscilloscope photographs of input and output signals for the two stages of Fig. 2.6. The input signal was a triangular wave with a fundamental frequency of 1200 cycles/second. With this sort of signal the diodes are out of conduction for long periods, so that appreciable change in cathode condenser voltage takes place. This causes some shift in the positions of the gain transitions. The lower picture shows the input and output signals for the condition of only the second stage going through a gain transition. The center picture is for a larger input signal, both stages being driven into their low-gain regions. The upper picture shows the change in output signal shape as the input signal is varied. The lower of the three curves shows linear performance, neither of the stages having passed a transition. The upper two waveforms are for successively larger input signals, and are identical to the output signals shown in the lower two photographs.

The measured performance of the two-stage amplifier with pulse signals is shown in Fig. 2.8. A pulse width of 5 microseconds and a repetition rate of 2000 cycles/second were used. The output has been plotted linearly in voltage, and the input linearly in decibels. The closeness to which a true logarithmic response is approximated is somewhat surprising in view of the fact that a 50-decibel input range is being covered with only two gain transitions (three segments). The maximum departures from the logarithmic characteristic are approximately 2 decibels.

4. MODIFICATIONS OF THE BASIC CIRCUIT

There are several variations of the basic circuit which may be used to advantage in various applications. First, there is the obvious possibility of using pentodes in place of triodes. The design procedure in the case of pentodes will be slightly different than that for triodes because of the presence of screen current in the cathode circuit. The quiescent diode current for a given output transition voltage will thus be somewhat larger, as will the quiescent cathode voltage. The gain-bandwidth considerations will, of course, be different. To provide a given gain at some specified bandwidth, a considerably smaller number of tube sections will be required than in the triode case. Typically this ratio approaches one to two, so that the number of pentodes would be only slightly greater than the number of dual triodes to do the same job. In applications requiring the pass-band to go up to several megacycles, the instability of triodes, due to their high grid-plate capacitance, would enter the picture, and pentodes would be the best choice. At the lower bandwidths (up to 1 Mc or so) triodes are quite satisfactory. For many applications, the use of triodes is advantageous because of the relatively large number of stages used. Since there is one gain transition per stage, a large number of stages results in an output characteristic made up of a large number of segments. A relatively good approximation to a desired over-all characteristic may thus be obtained.

There are several other minor modifications of the basic circuit of Figs. 2.1 and 2.6 which have frequently been employed. The negative high-voltage supply may be eliminated, and the common diode bias resistor R_{t2} returned to ground. The effect of this change is twofold: it reduces the value of R_{t2} to such an extent that the large-signal incremental gain is affected, and it causes a slight reduction in the stability of the circuit with respect to changes in components. For large positive signals R_{t2} will be effectively in parallel with R_k so that the large-signal gain will be increased if R_{t2} is not several times greater than R_k. (In the two-stage amplifier of Fig. 2.6, R_{t2} would have to be 10.5k to maintain the original diode currents.) Since R_k alone determines the incremental gain for large negative input signals, the result of this change is to make the large-signal gain less for positive than for negative signals. Since the polarity of the signal is reversed in going through each stage, this difference in gain tends to average out in a multistage amplifier. The over-all input/output characteristic will not be as smooth as that obtainable using a negative supply, but for many applications the operation is entirely satisfactory.

The reduction of R_{t2} also causes some loss of the circuit's ability to adjust to changes in tube and diode characteristics, although this has not been found at all serious in practical amplifiers where R_{t2} has been of the order of 10k to 20k.

A second type of circuit change which may be made is to keep the negative high-voltage supply, and to return both R_{t2} and R_k to it. This allows R_k to be made very large, which means that the large-signal incremental gain can be made very small. Thus the circuit of Fig. 2.6 could be modified in this way, raising R_k to 80k. The large signal gain would then be (using (2.7)) only 0.057. This corresponds to rather effective limiting action, and is easily obtained circuit-wise.

Another modification of the basic circuit which may be found useful is to bias the two diodes unequally so that the gain transitions occur at different voltages for positive and negative input signals. This brings out the fact that the over-all characteristic of an amplifier may be adjusted separately for positive and negative input signals. The principal circuit change is in the values of the diode bias resistors R_{t1} and R_{t2}. With unequal quiescent currents, the dynamic resistances of the two diodes will, of course, be different. The sum of these resistances would be used to determine the cathode degeneration in the small-signal case. Output signal voltage at transition would be found from (2.2) or (2.3), with the appropriate quiescent current and voltage substituted for i_o and v_o, respectively.

5. THE FEEDBACK PAIR

Up to this point the discussion has been limited to simple grounded-cathode stages. For two reasons which will be stated shortly, it is often advantageous to operate the stages in pairs with voltage feedback between grid and plate of the second tube. This sort of arrangement has been described in the literature as a feedback pair.[1] The particular arrangement which has been found useful in this work is shown in Fig. 5.1.

In this circuit the plate voltage for the first tube is supplied from the plate of the second tube through the feedback resistor R_F. Since the second stage has grid-to-plate voltage feedback, it will have a resistive input impedance determined by the amount of feedback. This input resistance is given by[2]

$$R_{in} = \frac{R_F R_2 + R_F R_L + R_2 R_L}{R_2 + (1 + \mu_2) R_L} \qquad (5.1)$$

where R_2 is used to designate the effective plate resistance of the second tube. In circuits with no cathode degeneration, R_2 would be just the dynamic plate resistance of the tube. With cathode degeneration such as would be introduced in using the IAGC circuits, R_2 would be given by Eq. (2.5).

Since the plate load of the first stage is provided by the input resistance of the second, the first stage gain is

$$A_1 = \frac{\mu_1 R_{in}}{R_{in} + R_1} \qquad (5.2)$$

where R_1 denotes the effective plate resistance, and μ_1 the amplification factor of the first tube. The first stage voltage gain given above is measured between grid and plate of the first tube.

The voltage gain of the second stage may be found from

$$A_2 = \frac{\mu_2 R_L(\mu_2 R_F - R_2)}{R_F R_2 + R_F R_L + R_2 R_L} \qquad (5.3)$$

where μ_2 is the amplification factor of the second tube, and the remaining notation is as previously defined.

The voltage gain of the pair is just the product of A_1 and A_2. This may be written as

$$A_T = \frac{\mu_1 \mu_2 R_L(\mu_2 R_F - R_2)}{R_F R_2 + R_F R_L + R_2 R_L + R_1\left[R_2 + (1+\mu_2) R_L\right]} \qquad (5.4)$$

Another quantity of particular interest is the impedance seen at the plate of the second tube. The output resistance due to the feedback is given by

$$R_{out} = \frac{R_1 + R_F}{R_1/R_2(1+\mu_2) + R_F/R_2 + 1} \qquad (5.5)$$

The actual impedance seen at the plate of the tube would be the parallel combination of R_{out} and R_L.

[1] Several forms of this circuit, though not the exact one used here, are discussed by Van Voorhis in "Microwave Receivers," vol. 23 of MIT Radiation Lab. Series, pp. 524-526.

[2] The derivations of the equations in this Section are indicated in Appendix B.

The advantages of the feedback pair for this type of amplifier may now be stated. First, note that the gain of the second stage is proportional to the difference between $\mu_2 R_F$ and the effective plate resistance of the second tube, R_2. If the IAGC diode arrangements are incorporated in the cathode circuits of the pair, it is seen that the large-signal gain of the second stage may be varied over wide limits by changing the second stage cathode resistor R_{k2}, which, in turn, controls R_2. It is readily possible to make the large-signal gain (which will be designated by A_{22}) zero, or even negative, if desired.

Since this same difference term appears in the over-all gain expression given by (5.4), the gain of the pair above the first transition point may be controlled to a large extent by R_{k2}. In the particular case of A_{22} equal to zero ($\mu_2 R_F = R_2$), the total gain A_T is also zero for all signals greater than that required to drive the second stage to its gain transition. Pairs arranged in this way thus operate as limiters, and as such are very useful in certain applications. For this special case the value of R_{k2} may be readily found using Eq. (2.5). Thus

$$R_F = \frac{R_{p2} + (1 + \mu_2)R_{k2}}{\mu_2} \qquad (5.6)$$

which, in the case of typical high-mu triodes may be reduced to

$$R_{k2} = R_F \qquad (5.7)$$

with good accuracy.

The second advantage of the feedback pair results from the low second-stage output resistance resulting from its grid-to-plate resistive feedback. As a result of this low resistance, the feedback pair is relatively insensitive to power supply ripple. This consideration is of importance at low signal levels (stages operating at maximum gain), and this is the condition under which maximum discrimination against supply voltage ripple is realized. As will be shown later, the ripple output of a feedback pair is typically something like 20 decibels less than that of a pair of stages without feedback, but of similar gain characteristics.

Design Considerations

The techniques for determining component values in the basic grounded-cathode circuit are applicable, in general, to the feedback pair, but several new considerations must be dealt with. The necessary departures in design procedure are best brought out by following through the action of a pair with IAGC circuits incorporated.

For small input signals, the output will be proportional to input, and the two stages will operate at maximum gain. As the input level is raised, a point will be reached at which one of the diodes in the second-stage cathode circuit will cease to conduct, so that a change in over-all incremental gain will occur. One point to note here is that all of the change in second tube current required to reach this first transition point does not flow through the load resistor R_L, some going into the feedback resistor R_F.

When the gain of the second tube drops, the input resistance of the stage increases. Accordingly, the gain of the first stage increases, so that the decrease in over-all gain is much less than the decrease in second stage gain. (This is to be expected because of the gain-stabilizing properties of negative voltage feedback.) As the input signal is increased further the first stage goes through its transition, and the over-all incremental gain decreases again. The over-all characteristic may be stated in terms of the individual stage parameters. The required stage gain characteristics to realize a desired over-all performance are illustrated in Fig. 5.2. The small-signal gain of the pair, A_1, is equal to the product of the small-signal gains of the first and second stages, $A_{11}A_{21}$. The intermediate pair gain A_2 is given by the product of the first-stage intermediate gain, and the second-stage large-signal gain, $A_{12}A_{22}$.. The large-signal gain of the pair is given by the product of the first-stage large-signal gain and the second-stage large-signal gain, $A_{13}A_{22}$. The first output transition voltage V_{T1} is equal to V_2; the second, V_{T2}, will occur at a voltage $V_{T1} + (V_{12} - V_{11})A_{22}$.

When the desired over-all characteristic is known, the design of the individual stages must be worked out in a cut-and-try fashion, due to the dependence of first-stage operation upon that of the second stage. Thus to obtain the proper intermediate gain, the large-signal gain A_{22} of the second stage must be less than it would be in an amplifier made up of separate grounded-cathode stages, to overcome the increase in first-stage gain from A_{11} to A_{12}. With this lower A_{22} gain, A_{13} must then be increased to provide the desired final gain A_3. As a rough working rule, the cathode-to-ground resistors in a feedback pair will be approximately twice the size used in separate stages to obtain a given response.

As mentioned earlier, the quiescent diode currents in the second stage must be somewhat greater than would be used with separate stages to realize the same output voltage at the first transition. The required quiescent current may be found from

$$i_{o2} = V_{T1}\left[1/R_L + (1 + 1/A_{21})/R_F\right] - v_{o2}/R_{k2} \qquad (5.8)$$

The first term represents the change in current through the load resistor to provide the desired transition. The second term accounts for the additional current which must be supplied to the feedback resistor. The last term represents the increment of tube current which is supplied from the cathode resistor R_{k2}, as discussed in Section 2 and Appendix A.

The diode currents required in the first stage may be specified in a similar way by the relation

$$i_{o1} = \frac{V_{T1}}{R_{in1}A_{21}} + \frac{V_{T2} - V_{T1}}{R_{in2}A_{22}} - \frac{v_{o1}}{R_{k1}} \qquad (5.9)$$

Here R_{in1} designates the input resistance of the second stage below its gain transition, and R_{in2}

the input resistance above the gain transition, as calculated from Eq. (5.1).

From the preceding discussion of the feedback pair, it is apparent that the design procedure is considerably more complicated than for the basic circuit without feedback. Perhaps the most difficult phase of the design is to achieve a rough approximation to the desired performance. Once this has been obtained, the directions and approximate amounts of component changes may be estimated, and a suitable final arrangement achieved without particular difficulty. To facilitate the process of obtaining a first approximate circuit, a general procedure is given below.

Once reasonable parameters (m and plate resistance_ have been chosen for the tube to be used, an approximate value for the feedback resistor R_F may be found from (5.4). To do this values must be assumed for the dynamic resistances of the cathode diodes so that the effective plate resistances of the tubes (R_1 and R_2) may be found. Also, a value must be assumed for R_L. (The effect of a poor value for R_L is not very serious, as it has relatively little effect upon the small-signal over-all gain A_1.)

When a first value for R_F is determined, the small-signal gain of the second stage may be found from (5.3). An approximate value for the output voltage at the first gain transition may now be found from Eq. (5.6) by neglecting the last term. The value of R_2 above the first transition may be found using (5.4). The second cathode resistor R_{k2} may now be found using Eq. (2.5). The actual value of V_{T1} may now be determined using (5.6). The required value for R_1 may now be found from (5.4). R_{k1} may now be found from (2.5). The output voltage at the second transition can now be determined directly, using (5.1), (5.3), and (5.7).

Using this procedure, the desired over-all gains (A_1, A_2, and A_3) will be realized, but the output transition voltages (V_{T1} and V_{T2}) will not generally be correct. The process of correcting the transition points involves changing the quiescent diode currents and recalculating the over-all response. If the initial assumptions are not too far out, the final values can be found with one or two such recalculations.

The values for the diode bias resistors, grid bias and coupling resistors, etc., may be determined directly from the appropriate relations given in Section 2.

The circuit of a typical feedback pair designed for a logarithmic output/input characteristic is shown in Fig. 5.3. The measured performance with positive input pulses is as follows: Small-signal voltage gain, $A_1 = 36$. Output voltage at first transition, $V_{T1} = 3$ volts. Gain above first transition, $A_2 = 5.7$. Output voltage at second transition, $V_{T2} = 6.5$ volts. Gain above second transition, $A_3 = 1.0$. The performance above the first transition is slightly different with negative input pulses because no negative supply is used for diode bias. The relatively small resistors used between the diode junctions and ground thus cause the gains to be different for positive and negative signals, as discussed in Section 4.

In this circuit the first tube operates with a plate current of 3.5 ma, and the second with 4.5 ma. The quiescent diode currents are approximately 0.75 ma. Since the plate voltage for the first tube is supplied through the 22k feedback resistor, it operates at a relatively low voltage. In this design the first tube operates with a plate voltage of 135 volts, and a cathode voltage of 22.5 volts. The grid bias is about 0.75 volts. The plate voltage of the second stage is 212 volts, while the cathode voltage is 31 volts. The second tube bias is 2.0 volts. Under these conditions, each tube has an amplification factor of about 60, and a plate resistance of approximately 15k.

To illustrate the insensitivity of the feedback pair to power supply ripple, a comparison between the circuit of Fig. 5.3 and that of Fig. 2.6 will be made. For this purpose, the fraction of supply ripple appearing at the first plate of each type of amplifier will be calculated. For these computations the amplifiers will be assumed to be operating at their maximum (small-signal) gain, since this is the condition under which ripple is bothersome. First, for the basic circuit of Fig. 2.6, the effective plate resistance for the calculated operating conditions may be found using Fig. 2.5 and Eq. (2.5). This is found to be 36k. The fraction of available ripple which will appear at the plate will be determined by the voltage division between plate load and effective plate resistance. This is found to be 0.86. In the case of the feedback pair, the fraction of the total ripple voltage at the second plate is first found using Fig. 2.5 and Eqs. (2.5) and (5.5). The effective plate resistances of the two stages is found to be 32k. From (5.5) the output resistance of the second stage due to feedback is found to be 0.86k. The fractional ripple at the second plate is found as above to be 0.155. The ripple on the first plate will be less than this because of the dividing action of the 22k feedback resistor and the 32k effective output resistance of the first tube. The fractional ripple at the first plate is thus found to be 0.095, so that the feedback pair gives an improvement of 0.86/0.095 = 9.05, or 19.1 decibels.

The amplifier of Fig. 5.3 may be converted into a limiter by changing only two resistors. For this type of operation the feedback resistor is made 8.2k as indicated by (5.7). This changes the operating conditions of the first tube considerably, so that the resistor determining its grid bias must be changed. The original 3.5 ma plate current is obtained if the 33k bias resistor is raised to 68k.

6. CONSIDERATIONS RELATING TO COMPLETE AMPLIFIERS

Up to this point, attention has been focused upon the problem of designing stages or pairs of stages to operate in a certain manner. There are a few points pertaining to the cascading of stages which should be mentioned at this time.

When several identical stages are cascaded, the increments of output voltage between successive transitions is not constant. This may be understood by considering an example. Let the small-signal gains be A_1, the stage output transition voltages be V, and the large-signal gains be A_2. As the input signal to the amplifier is increased from zero, the last stage will go through its gain transition at an amplifier output voltage of V. The output of the preceding stage at this point will be V/A_1. When its output increases to V the second transition will occur. The output of the preceding stage will thus increase by an amount $V - V/A_1$ before the second transition occurs at the output, so that the voltage between first and second transitions will be $A_2V(1 - 1/A_1)$. At the time the second transition occurs at the output, the output of the second from last stage is again $V - V/A_1$. The next increment of output voltage is thus $A_2^2V(1 - 1/A_1)$, etc.

In the special case of $A_2 = 1$ which is frequently used in practice, the first increment at the output is V, and all remaining increments are of amount $V(1 - 1/A_1)$. When A_2 is significantly different from unity, the size of the output voltage increments will change progressively as the input signal is increased.

The same sort of increment variation will be present in the case of feedback pairs. Here the first two transitions are determined by the pair design. Assuming identical pairs, the output voltage increment between second and third transitions will be $\left[V_{T1} - V_{T1}/A_1 - (V_{T2}-V_{T1})/A_2\right] A_3$. The fourth increment will be $\left[V_{T2} - V_{T1}\right] A_3$. The fifth increment will be the same as the third, except that A_3 appears squared. Similarly, the seventh, ninth, etc., have A_3^3, A_3^4, etc. The sixth, eighth, etc., output voltage increments will be the same as the fourth, except for A_3^2, A_3^3, etc., terms.

In a typical case of identical pairs with equal increments between first and second transitions and unity large-signal gain, A_3, it is found that there is a repetitive variation in the size of the output increments. If, for example, the pair characteristics are $A_1 = 36$, $A_2 = 6$, $A_3 = 1$, $V_{T1} = 3$ volts, and $V_{T2} = 6$ volts, the output voltage increments are found to be 3, 3, 2.4, 3, 2.4, 3, etc. If all but the last pair are altered to have output transition voltages of 3.6 volts and 6.6 volts, the output increments will all be 3 volts.

Another consideration relating to the cascading of several stages or pairs is that of overdriving in the later tubes. Difficulty of this sort may be encountered when the grid of the last tube of a high-gain chain is driven negative. If the signal at this point exceeds the quiescent cathode voltage, clipping will occur. This situation will be worst when feedback pairs are used because of the higher gain of the first stage, as illustrated in Fig. 5.3.

There are several ways to get around this difficulty. First, other considerations permitting, the size of the output voltage increments of all stages may be reduced. This, of course, results in a smaller amplifier output voltage for a given large input signal. Second, it may be possible to reverse the signal polarity so that the larger signals at the last grid are positive. Since the stage operates with a large unbypassed cathode resistor, very large positive signals may be applied without drawing grid current. Another possible solution , when pairs are being used, is to convert the last two stages to the basic circuit (no feedback). In typical cases this will reduce the signal at the last grid by a factor of two.

7. CONCLUSIONS

One of the most important aspects of the operation of the type of amplifier described here is its stability with respect to variations in the characteristics of the tubes and germanium diodes employed. Considerable difficulty from these sources was experienced in earlier circuits which involved diodes connected directly between the cathode of the tubes and ground. A "floating" arrangement is used here in which the voltage across a large condenser adjusts to changes in diode and tube characteristics. In practice it has been found that tubes and germanium diodes may be selected at random with no appreciable change in circuit performance. The drop in the back-resistance of germanium diodes which occurs at high temperatures causes no appreciable change in circuit operation, since the diodes are effectively in parallel with the stage cathode resistors which are typically 10k or less.

APPENDIX A

DESIGN RELATIONS FOR THE BASIC IAGC CIRCUIT

The standard expression for the voltage gain of a grounded-cathode triode state is

$$A = \frac{\mu R_L}{R_L + R_p} \qquad (A.1)$$

When this is solved for R_L, the expression given in (2.1), without the factor of 2, results. Typically the degeneration due to the conducting diodes in the cathode circuit is such that the gain A is realized when the plate load resistor is doubled. The accuracy of (2.1) is dependent upon the quiescent diode current, the characteristics of the diodes, the tube parameters, etc., so that it is useful only for starting the design.

The simple relation given by Eq. (2.2) is usually of adequate accuracy for the practical determination of output voltage at the gain transition. The answer will always be somewhat low, since some change in the cathode potential is necessary to overcome the diode voltage drops. A correction term of

$$v_o R_L/R_k \qquad (A.2)$$

may be added to (2.2) to improve the accuracy. The actual change in cathode voltage is somewhat greater than v_o because of the increased current through the diode not being driven out of conduction. The slight under-correction realized with Eq. (2.3) compensates for the gradual decrease in gain before the actual transition voltage is reached, so that the output of the stage for large signals may be found with good accuracy if the output voltage at transition is determined from (2.3).

An equivalent circuit representation of a triode with cathode degeneration may be found using Thevenin's theorem. The circuit for signal frequencies is shown in Fig. A1, and the equivalent in Fig. A2.

The equivalent generator voltage is first found by assuming that the load (R_L) is infinite. The current through the circuit is then zero, so that the open-circuit voltage is just $-\mu V_s$. The impedance seen from the output terminals is then found with the internal generator V_s shorted out. For this determination R_L in Fig. A2 is replaced by a zero-impedance generator, and the V_s term in the equivalent generator is dropped. The net driving voltage is then $V - \mu iR$, so that the current that flows is given by

$$i = \frac{V - \mu iR}{R_p + R} \qquad (A.3)$$

Solving for V/i, the output impedance is found to be $R_p + (1 + \mu)R$, as given by Eq. (2.5).

Eq. (2.6) is obtained by solving the gain relation given by (A.1) for R_L with the plate resistance of the tube replaced by the equivalent output impedance obtained from (2.5).

The expression for R_k given by (2.7) is found by writing the expression for large-signal incremental gain

$$A_2 = \frac{\mu R_L}{R_L + R_p + (1 + \mu)R_k} \qquad (A.4)$$

and then solving for R_k. The approximate relation for A_2 given by Eq. (2.4) comes from the exact equation given above. It is based upon two assumptions which are usually met reasonably well in practice. The first is that the amplification factor of the tube is much greater than unity. This is a reasonable assumption for all tube types likely to be employed in pulse amplifiers of this sort. The second assumption is that μR_k is large in comparison with the sum of R_L and R_p. The error resulting from this latter approximation can become large if the large-signal incremental gain is appreciably greater than one. Considerable error then results because R_k must be fairly small to realize the desired gain. When the indicated assumptions are made, Eq. (A.4) reduces directly to (2.4).

Eqs. (2.8), (2.9), (2.10), and (2.11) require no explanation, as they result from straightforward application of Ohm's law.

Eq. (2.12), giving the change in potential of the cathode condenser during a strong signal, follows directly from well-known relations. The charge in a condenser is given by Q = CV, where C is the capacitance and V the voltage. If a constant current i_o now flows into or out of the condenser for a time t, the charge will change by an amount $\Delta Q = i_o t$, so that the incremental voltage change will be $\Delta V = i_o t/C$, as given in (2.12). (This expression will be very nearly exact even if ΔV should be several volts, because of the high-voltage, high-resistance sources which supply the diode currents.)

APPENDIX B

DESIGN RELATIONS FOR THE FEEDBACK PAIR CIRCUIT

The input resistance of the second stage of a feedback pair is readily found using the equivalent circuit shown in Fig. B1. The loop equations are

$$i_1(R_F + R_2) + i_2(R_2) = (1 + \mu_2)V \qquad (B.1)$$

and

$$i_1(R_2) + i_2(R_2 + R_L) = \mu_2 V \qquad (B.2)$$

The input resistance is found by solving for i_1 and then determining V/i_1. This is given in Eq. (5.1).

The voltage gain of the second stage is determined by solving (B.1) and (B.2) for i_2, multiplying by R_L to obtain the output voltage, and then dividing by the input voltage V. The result is given by Eq. (5.3).

The expression for first-stage voltage gain given by (5.2) is obtained from the standard triode gain relation of Eq. (A.1), the load being provided by the input resistance of the second stage.

The over-all gain of the pair may be calculated by multiplying the individual stage gains given by Eqs. (5.2) and (5.3). The result is given by (5.4).

The output impedance of the second stage may be calculated by replacing the plate load resistor by a zero-impedance generator. The equivalent circuit is shown in Fig. B2. The loop equations are found to be

$$i_1(R_F + R_1 + R_2) - i_2(R_2) = -\mu_2 e \qquad (B.3)$$

and

$$-i_1(R_2) + i_2(R_2) = \mu_2 e + V \qquad (B.4)$$

Since e is equal to i_1R_1, these may be rewritten as

$$i_1\left[R_F + R_2 + R_1(1 + \mu_2)\right] - i_2(R_2) = 0 \qquad (B.5)$$

and

$$-i_1\left[\mu_2 R_1 - R_2\right] + i_2(R_2) = V \qquad (B.6)$$

The output resistance is found by solving for i_2 and then determining V/i_2. The result is given by Eq. (5.5).

The relation for second-stage quiescent diode current given by Eq. (5.8) may be derived in a direct manner. The first term, V_{T1}/R_L might be termed the useful part of the total diode current, since it sets the output transition. The additional current that must be supplied to the feedback resistor is given by $V_{T1}(1 + 1/A_{21})/R_F$. This current is greater than V_{T1}/R_F by $1/A_{21}$ because of the action of the feedback. Thus, when there is unit signal voltage at the plate of the tube, there must be a signal of opposite polarity and magnitude $1/A_{21}$ at the grid, to which the feedback resistor is connected. The gain here is the small-signal gain of the second stage. The last term v_{o1}/R_{k2} accounts for the increased tube current which comes from the cathode resistor R_{k2} in pulling the diode out of conduction.

The quiescent diode current required in the first stage may be specified in terms of the output transition voltages of the pair. This relation is given by (5.9). The first term represents the change in the tube current of the first stage up to the first output transition. The output voltage at the first plate will be V_{T1}/A_{21} when V_{T1} is reached, so that this portion of the current will be $V_{T1}/(R_{in1}A_{21})$. The second term represents the additional change in first-tube current necessary to provide a second output transition voltage of V_{T2}. This term is similar to the first, the incremental change in output voltage appearing in the numerator, and the input resistance and gain of the second stage in the denominator. Since the second stage is above its gain transition, the appropriate values R_{in2} and A_{22} are used.

ACKNOWLEDGMENT

The work reported in this paper was made possible through support extended the Electronics Research Laboratory, Stanford University, jointly by the U. S. Army Signal Corps, the U. S. Air Force, and the U. S. Navy (Office of Naval Research and the Bureau of Ships) under Office of Naval Research Contract N6onr 251.

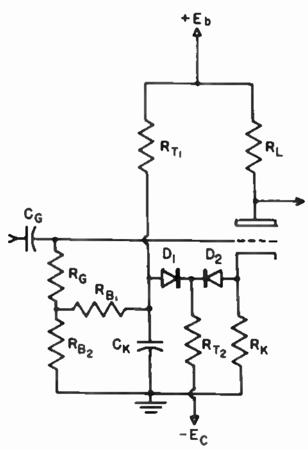


Fig. 2.1
Actual circuit.

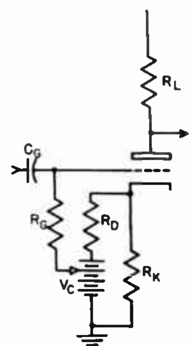


Fig. 2.2
Small-signal equivalent.

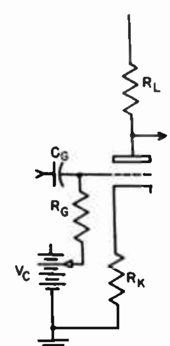


Fig. 2.3
Large-signal equivalent.

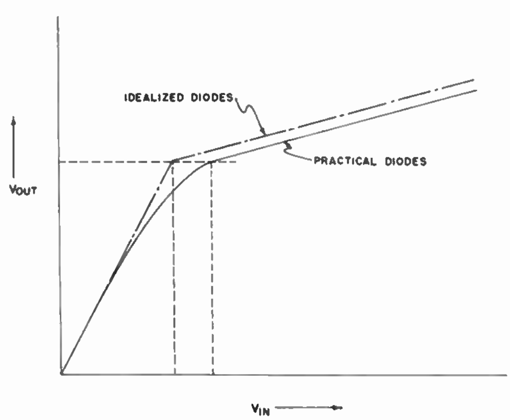


Fig. 2.4
Output/input characteristic
for one stage.

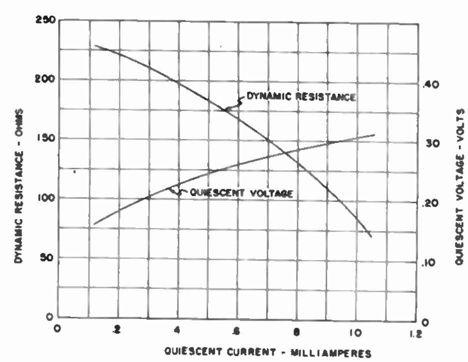


Fig. 2.5
Characteristics of IN56
germanium diodes.

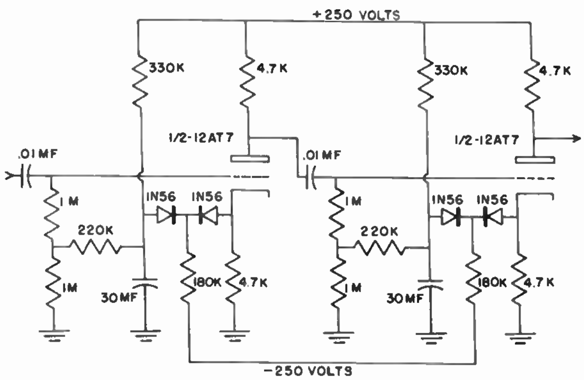


Fig. 2.6
Two stages of a typical amplifier.

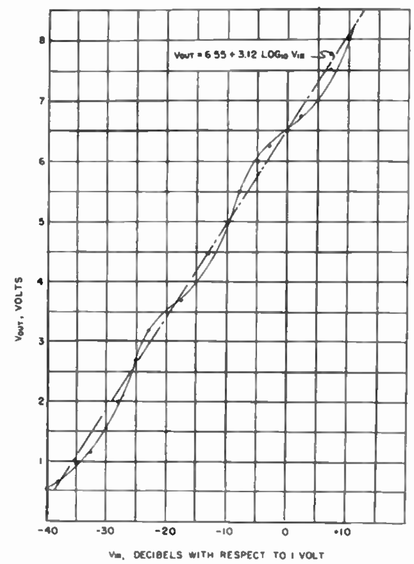


Fig. 2.8
Input/output characteristic of the two-stage amplifier of Fig. 2.6.

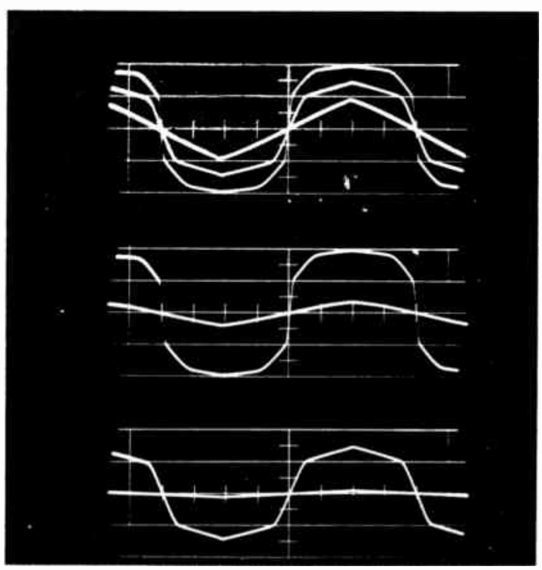


Fig. 2.7
Two-stage amplifier operation with triangular input signal.

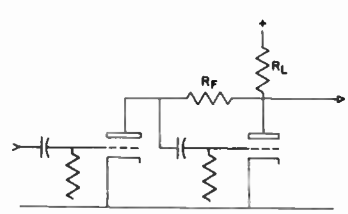


Fig. 5.1
The feedback pair.

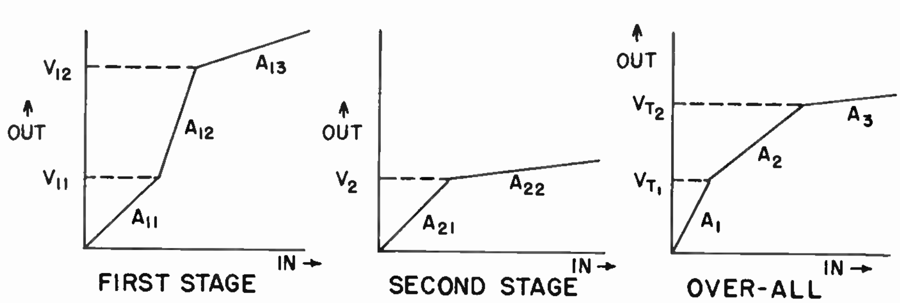


Fig. 5.2
Gain characteristics of pair with IAGC.

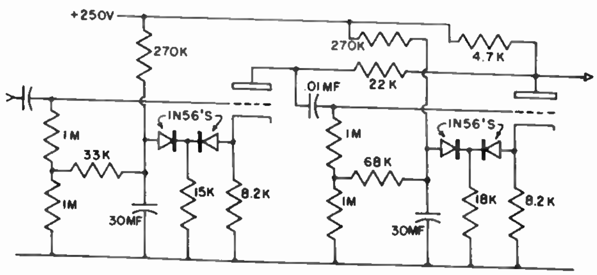


Fig. 5.3
Typical feedback pair with IAGC.

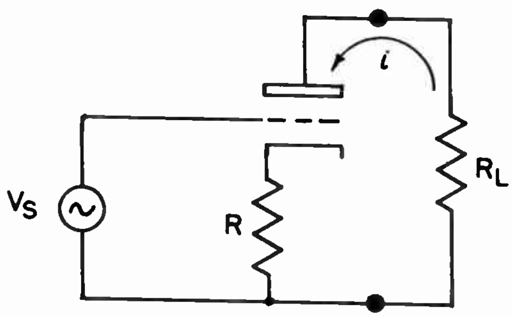


Fig. A.1
Triode with cathode degeneration - actual circuit.

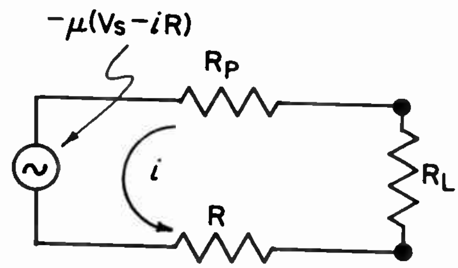


Fig. A.2
Triode with cathode degeneration - equivalent circuit.

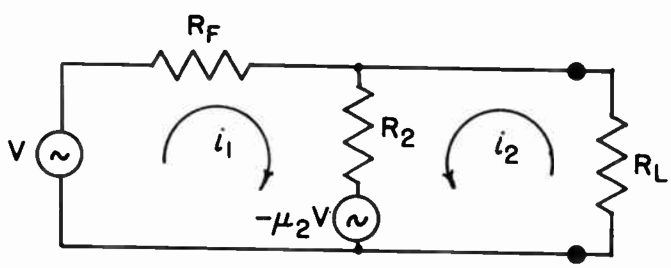


Fig. B.1
Equivalent circuit of second stage of feedback pair.

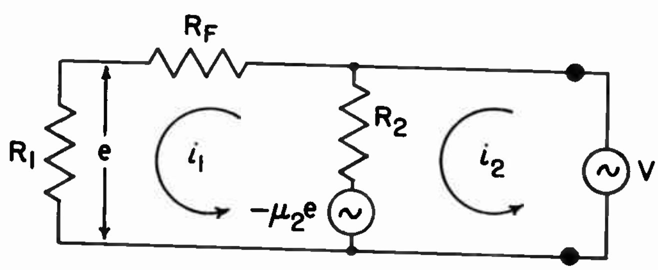


Fig. B.2
Circuit for computing output resistance of feedback pair.

AN AUTOMATIC LEVEL-SETTING SYNC AND AGC SYSTEM

E. O. Keizer
RCA Laboratories Division
Princeton, N. J.

and

M. G. Kroger
Motorola, Inc.
Chicago, Illinois

Introduction

When a modern television receiver is tuned to a good, clean signal--one without interference of any kind--the resulting picture is likely to be steady. However, if high energy impulse-type interference, such as ignition interference from automobiles, occurs along with the signal from the television station, the picture may jitter, may roll, or, in severe cases, may become completely unsynchronized and unusable. This is more likely to happen if the signal is weak, and may become a serious problem in some areas. In many cases the picture would still be usable, even with the interference, if it were steady.

The failure to synchronize in the presence of impulse noise is often the result of charging up of the sync separator and automatic-gain-control circuits by such noise pulses, causing the useful sync information to be suppressed or entirely lost in the sync separator output. It is the purpose of this paper to describe a sync separator and automatic-gain-control system wherein such charging up is reduced, resulting in steadier pictures.

Conventional System

Figure 1 is a block diagram of a conventional sync separator and automatic-gain-control system. The diagram has been simplified by lumping into one block all automatic-gain-controlled amplifiers, and into another block all other signal stages.

A portion of the video output is applied to the sync separator. In most receivers the sync separator is self-biased, the bias following the peaks of the composite video output so conduction normally occurs during the sync peak portion of the signal, but not during the picture portion of the signal. However, if there are long or numerous peaks of interference having greater amplitude than the sync portion of the desired signal, the self bias voltage will increase toward the level of these peaks. Thus the self bias voltage will tend to be too large and all or part of the sync portion of the signal will be biased beyond the sync separator operating range, resulting in loss of synchronization-picture jitter and roll.

The automatic-gain-control voltage may be derived from the video output also, or from before the video amplifier stage. In either case, it may also be susceptible to the same long or numerous noise pulses that affected the sync separator. As a result too much gain control bias may be developed, causing a reduction in the amount of desired signal present in the input to the sync separator; where the noise is especially bad, the desired signal may be almost completely suppressed.

Thus, if either the sync separator circuit or the automatic-gain-control circuit charges up on noise energy, poor synchronization is likely to result.

Fixed-Bias Sync Separator System

Assume that a given television receiver has high maximum gain and a noise-immune automatic-gain-control circuit with a flat video output amplitude characteristic. Then, under all usable signal conditions, the sync peaks would be at a certain dc level in the output. In such a receiver it would be possible to use a direct-coupled sync separator, externally biased from a stiff voltage source to operate at that sync peak level, and charging up would not occur; useful sync output would continue in the presence of impulse noise. The performance of such a system would depend upon the accuracy and reliability to which the sync peak level could be maintained with respect to the bias level under conditions of varying signal level and interference.

In practice, there are usually three difficulties to such an arrangement. First, the gain and speed of response of the automatic-gain-control loop may be inadequate to maintain the signal level sufficiently constant to keep the sync portion of the video amplifier output within the operating range of the sync separator, resulting either in loss of sync or in picture information being present in the sync output. Second, the automatic-gain-control voltage may be affected by interference, resulting in failure to maintain the proper signal level and in loss of sync. Third, the signal may be so weak that the maximum gain of the receiver may be insufficient to maintain full video output, again resulting in loss of sync.

Bias Level-Setting Sync and Automatic-Gain-Control System

The system to be described overcomes these three difficulties and provides good sync and automatic-gain-control performance, more immune to impulse noise than conventional circuits.

A simplified block diagram outlining this sync and automatic-gain-control system is shown in figure 2. The self-biasing sync separator of the first figure has been replaced by an externally biased dc-coupled amplifier stage. This serves the double purposes of sync separation and of additional automatic-gain-control loop amplification. The automatic-gain-control voltage itself is derived by another stage operating from the separated sync which appears at the output of this stage.

This double-purpose stage has a definite noise-limiting level and the automatic-gain-control stage which follows is biased so the sync peaks are automatically held very close to that level. The added gain in the loop is helpful in this respect. Also, because the automatic-gain-control stage operates on a signal from which the noise peaks have been clipped close to sync peaks, its noise immunity is considerably improved.

The external bias applied to the sync separator is a voltage, controlled by the automatic-gain-control voltage, which decreases when the video output decreases on very weak signals. Such a variable bias can be obtained from the low-voltage terminal of a resistor connected in series with the plate current of the gain-controlled i-f amplifier stages. At this point the voltage drops when a weak signal is being received, due to the increased current in the amplifiers at low automatic-gain-control voltages. This connection is indicated in figure 2.

Using the i-f amplifiers to bias the sync separator stage adds a parallel stage of gain in the control loop, one that is most effective for weak signals where the normal automatic-gain-control voltage is least effective. With this dc amplification added to the control circuit loop, high gain is maintained at all signal levels, insuring close automatic tracking between the level of sync peaks and the noise clipping bias. Since the additional amplification holds the video output to a level determined by the gain-controlled i-f amplifier plate voltage rather than to a fixed delay voltage, the resulting video output voltage versus r-f input voltage characteristic of the receiver is not as flat as it would be with a conventional high-gain delayed automatic-gain-control circuit.

Effects of the Variable External Bias

Figure 3 shows two output versus input curves illustrating the effect of the variable bias. Data for curve I was taken on a relatively low-gain receiver converted to the new system, while curve II is a composite curve made up in its sloping portion of the output versus input characteristic of the same receiver with the automatic-gain-control voltage shorted, and, in its level portion, of the constant output that would result if an ideal delayed automatic-gain-control circuit were employed.

In operation, this circuit always develops some automatic-gain-control bias. It is fortunate, therefore, for its operation that at very low biases the gain of the controlled i-f amplifiers does not change rapidly, and the maximum gain of the receiver is not greatly affected. However, delay of the application of automatic-gain-control voltage to the tuner is probably more important for best noise factor for this system than for systems developing no automatic-gain-control voltage of very weak signals.

Since the sync separator bias follows the signal level, the noise clipping level of the sync separator remains close to the top of the sync peaks for weak as well as for strong signals, thereby maintaining good noise immunity for signal levels that would be below the automatic-gain-control threshold in a conventional system.

Also, since the sync peak level of the video output is controlled by the external bias on the sync separator, that bias is useful in maintaining background level in the presence of rapid or large signal level changes.

Variations in Circuits Using This System

The foregoing description was intended to be an outline of the principles of the system, and therefore omitted circuit details. It was found that many different variations in the actual circuitry were workable and noise-immune. For example, in some cases the automatic-gain-control voltage was obtained using a keyed rectifier tube and in some cases using a non-keyed rectifier system. In some cases the video signal was applied to the cathode of the sync separator tube and the external bias to the grid, while in others the signal was applied to the grid and the bias to the cathode. In some cases the sync separator output was direct-coupled to the following sync amplifier stage and in others it was ac-connected. The keyed automatic-gain-control stage and direct-coupled sync amplifier stage were found to contribute to noise immunity.

In each case, it was necessary to be very careful about the initial design of the overall system, but once so designed, operation was reliable and non-critical.

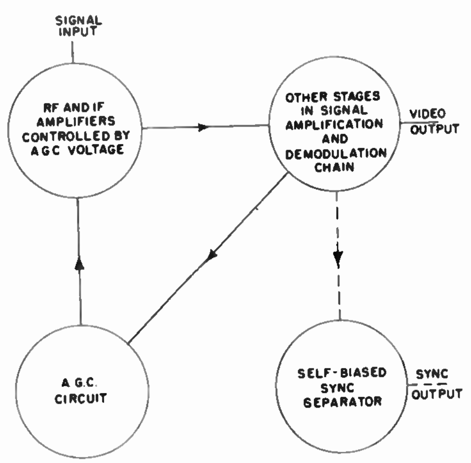


Fig. 1
A conventional sync and AGC system.

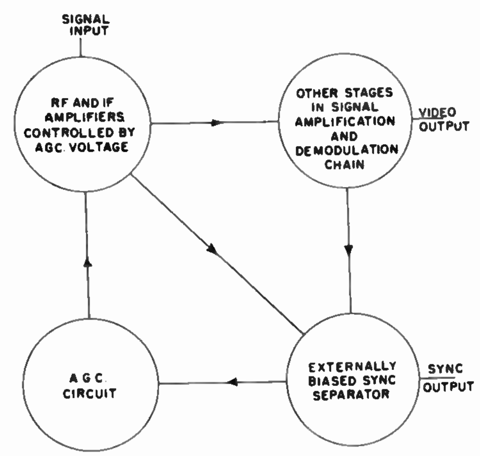


Fig. 2
Level-setting sync separator and AGC system.

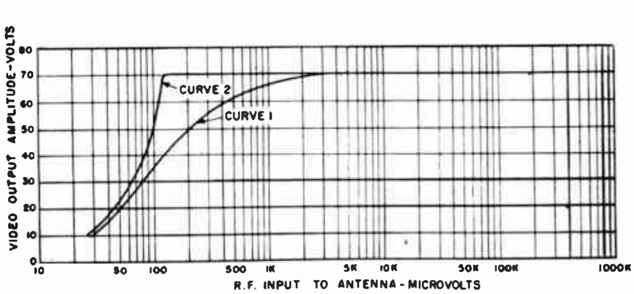


Fig. 3
Video output vs. signal input.

PACKAGED ADJACENT CHANNEL ATTENUATION
FOR TELEVISION RECEIVERS

John P. Van Duyne
Television Receiver Engineering Dept.
Allen B. Du Mont Laboratories, Inc.
East Paterson, New Jersey

The ability of a television receiver to perform adequately in the presence of adjacent channel interference has become, in spite of the best allocation work of the FCC, an increasingly important factor in receiver sales. Now that the long awaited thaw of the "freeze" on television transmitter Construction Permits has occurred, there will be more areas in the country in which adjacent channel interference will be a problem. These areas usually lie half way between centers of high population density in which adjacent television assignments exist and are usually at or beyond the secondary service areas of each. A typical area in which adjacent channel interference is a serious problem, is Trenton, New Jersey. Here, channels 2, 4, 5, 7, 9, 11, 13 from New York are received with varying success. Sandwiched in between are channels 3, 6, and 10 from Philadelphia. (See Figure 1) The directivity limitations of television antennas being what they are, a large portion of the burden of separating the desired from the undesired program material falls on the television receiver itself.

The designer of a Television receiver has to decide what degree of protection against adjacent channel interference he should design into his product. This decision must be made from the following factors:

1. The ultimate desired list price of the television receiver.

2. The relative importance of adjacent channel attenuation as weighed against simplicity of design, components, and alignment.

3. The percentage of sales of the product expected in areas where adjacent channel attenuation is a problem.

This decision which the designer must make, is bound to be a compromise in the face of the preceding conflicting requirements. Since it is usually undesirable to manufacture more than one chassis design, the manufacturer usually decides, in the interest of economy, to market a receiver with considerably less than adequate adjacent channel attenuation for those areas previously described.

The design problem posed by the relative importance of adjacent channel attenuation and a high, reproduceable, degree of electric fidelity is a severe one. A brief elaboration on the relation of the adjacent channel attenuation requirement to that of good electric fidelity is in order. As in most band pass system designs, the design criterion for maximum phase linearity is not necessarily compatible with the criterion for maximum adjacent channel attenuation. As in classical filter theory, the placement of the "poles" and "zeros" within the frequency range of interest determines the basic response characteristic of the television receiver. Due to physical and economic limitations, a high degree of adjacent channel attenuation is not achieved in the usual television tuner. Hence, the burden of adjacent channel rejection falls on the intermediate frequency amplifier. For similar reasons, the electric fidelity of the receiver is largely determined by the IF amplifier and secondarily so by the video amplifier. Since the IF amplifier must provide high gain, adjacent channel attenuation, and control the electric fidelity, it is readily seen that the major portion of the adjacent channel attenuation problem occurs in the I.F. amplifier design. To provide a high degree of electric fidelity, the IF amplifier system should have a nearly linear phase characteristic throughout the pass band and extending approximately 750 KC down the slope on the picture carrier side of the pass band. To achieve this high degree of phase linearity, sharp changes in amplitude response must be avoided, if expensive correcting networks are not to be used. At the same time however, the requirement for high adjacent channel attenuation implies the need for a steep slope in the amplitude response, beginning 750KC higher (at intermediate frequency) than the normal picture carrier frequency. The presence of the sharp corner in the amplitude response occasioned by this steep slope is not compatible with the requirement for good phase linearity in the region of $\pm$ 750KC about the normal picture carrier frequency. A compromise design between these two requirements is entirely possible, but involves critical, interacting adjustments which have not been practical in high speed, mass production of television receivers.

Another requirement on the television receiver IF amplifier is that of high gain at a low cost and with good stability. This desire for high gain with the fewest possible tubes limits the attenuation which may be achieved by absorption traps. These by their very nature reduce the available gain per stage. With the relatively low "Q's" possible at television intermediate frequencies and with economical construction, it is difficult to construct filter elements that will give high out of band attenuation and low pass band insertion loss. The major physical limitation is the need for small size. The impedance developed by the band pass circuits must be as high as possible in the interest of high gain. Calculation has indicated that one way to reduce the effects of these physical limitations would be to design a filter with a low characteristic impedance. This impedance would be low compared to that necessary if the filter were to be used as a transfer impedance between two high g_m pentode vacuum tubes. This implies the need for impedance transformation techniques ih order that high gain still be achieved, with the use of a low characteristic impedance for the filter. In view of the preceding considerations, it would be very desirable if a manufacturer could produce a television receiver in which the adjacent channel attenuation was adequate for a normal area. By "normal area" is meant a region in which the relative geography of the receiver and local television transmitters is such that the adjacent interfering signals are several times weaker than the desired local signals. Experience has indicated that adjacent channel attenuation of the order of 30 db is adequate protection in a reasonable percentage of most secondary service areas and in most of the primary service areas. An ideal situation would then result if a manufacturer could mass produce a television receiver with approximately 30 db of adjacent channel attenuation and then, by a simple field conversion, add 40 db of additional adjacent channel attenuation. It is the purpose of this paper to describe a packaged adjacent channel rejection filter suitable for this purpose.

A filter has been designed which is suitable for inclusion in the "link" between a television receiver tuner and an intermediate frequency amplifier. The receiver may be normally supplied without the filter. Upon receipt of the receiver by a distributor in an area where adjacent channel interference is known to be a problem, the distributor may insert the filter in the receiver with negligible deleterious effects upon the performance of the receiver. Thus, the cost of the additional attenuation is borne only by the customer who directly profits by the presence of this attenuation.

The filter design settled upon is a six element series "M Derived" band pass filter. (See Figure 3a and 3b) A study of the component values (calculated by the relations of Figure 3b) required for such a filter operating in the 44mc IF band indicated that mass produced components, readily available, were suitable. Typical values are shown in Figure 4. Such convenient values were not obtained with many other filter configurations, having similarly placed "poles" and "zeros", together with the more nearly constant image impedance which the six element series band pass filter exhibits. The constancy of image impedances is of importance if the filter is to be inserted in a length of transmission line. So called "Link coupled" circuits are frequently used between the RF tuner and the IF amplifier in television receivers. This link offers a convenient point to insert the filter. Since the length of cable used in this coupling circuit is an appreciable portion of a wave length, it is necessary that any network inserted in this line be reasonably well matched to the transmission line characteristic impedance to avoid serious distortion of the desired response. In the choice of the components which are to exhibit the values calculated with the aid of the equations in Figure 3b, extreme care must be exercised to be sure that the performance at 44mc may be accurately determined. Figure 5 shows the filter configuration together with those stray reactances which are likely to degrade the performance.

For example, C_{L1} of Figure 5, which represents the distributed capacity of L_1 will seriously degrade the maximum attenuation and will cause a tilted pass band response due to variation in its effective inductance across the pass band. These effects have been practically eliminated by splitting L_1 into two coils, separated by the capacitor $C_1/2$. The physical configuration of these parts may be seen in Figure 8.

Similarly, the lead inductance, L_{C2}, of the large capacitor C_2, is detrimental. This has been handled by minimizing and controlling the lead length through a mounting technique shown in Figure 8. The effective capacity is greater than the low frequency value as indicated in Figure 5b. The magnitude of this increase will be from 25% to 50% of the low frequency value in a typical case.

As a starting point for the design, the filter was considered to be operating between resistive generator and load impedances. A filter was constructed utilizing the component values of Figure 4. The measured performance of this filter is shown in Figure 6. It will be seen that 40 db of attenuation is provided with respect to the response at picture carrier. The insertion loss of 3db is small compared to similar attenuation achieved by absorption traps. At this stage of the design it was felt that this represented quite desirable attenuation performance. A check on the electric fidelity was made by comparing the transient response of a television receiver IF system driven from a 75 ohm resistive generator with the response of that system driven from the same generator with the filter of figures 4 and 8 inserted between the two. No additional transient distortion was introduced, thus indicating that the design was feasible for use in a television receiver under the conditions of resistive source and load.

It is desirable, for economic reasons, to insert this filter in the existing link circuit between the RF tuner and the IF amplifier. In our case, this circuit was an over-coupled double tuned circuit of rather narrow bandwidth. The use of this circuit was dictated by the selectivity requirements of the IF system when used without the filter. If an overcoupled condition is required due to gain - band-width product considerations, impedance matching of both source and load in the link circuit is impossible. Consequently it was decided to design the circuit such that the impedance viewed at the receiving end of the link was equal, at the center of the pass band, to the characteristic impedance of the coaxial cable used. It was then experimentally determined that the insertion of the filter of Figures 4 and 8 in the link circuit did seriously distort the resulting pass band response. This was felt to be due to the fact that the insertion of the filter changed the effective length of line in the link. This was proven to be the case by the fact that a length of cable could be chosen, which when used with the filter, permitted a pass band response in the link circuit similar to that of the double tuned circuit without the filter and with its normal length of cable. Figure 7 shows a plot of a typical television receiver amplitude response with and without the filter. Correction has been made for the "electrical length" of the filter. It will be noted that in the pass region, there is a negligible difference between the shapes of the two curves, but that extreme attenuation is achieved at the adjacent channel carrier frequencies.

Field results with the filter of figures 4 and 8 installed as mentioned above, have been quite gratifying. It has been found possible to receive interference free pictures on adjacent channels with interfering to desired signal ratios of +20db. Lest this seem like a small ratio, it should be pointed out that interference is annoying, even if the level of interfering signal is 40db down at the second detector. Figure 9 shows a photograph of a receiver with 30db of adjacent channel attenuation operating on channel 3 and being interfered with by channels 2 and 4. The interfering signals exceed the desired signal by approximately 20db. Negligible entertainment value is contained in this picture. Figure 10 shows the same receiver under the same signal conditions, but with the adjacent channel filter installed. Now, negligible interference is present. If the ratio of interfering to desired signal is increased still further, for example by an additional 20db, interference begins to appear in the picture. (See Figure 11) This interference is occasioned by cross modulation in the first RF amplifier. The particular interference shown is the result of the presence of channel 2 sound and picture carriers in the RF amplifier with sufficient amplitude to produce the difference frequency between twice the channel 2 sound carrier and the channel 2 picture carrier. This frequency is 64.25mc and is both amplitude modulated by the channel 2 video information and frequency modulated by channel 2 sound. This is evidenced by the appearance of <u>Channel 2</u> "blanking bars" and a 3mc beat which varies with channel 2 sound modulation. No amount of IF amplifier selectivity will eliminate or minimize this phenomenon. It would be necessary to introduce significant adjacent channel attenuation ahead of the first non-linear element, usually the RF amplifier tube, in order to reduce this effect.

It is also interesting to compare the transient response of the television receiver, with and without the filter installed in the link circuit. This is shown in Figure 12. There are some noticeable differences between the transient responses, but they represent relatively small effects compared to the distortions occasioned by variations in television transmitters, studio techniques, antenna characteristics, and customer manipulation.

In conclusion, it should be pointed out that this idea is simply the extension of a well recognized general principle, namely that of having the customer who benefits from increased performance be the one who bears the cost of that increased performance. The filter described in this paper, when used in conjunction with a suitably designed television receiver with 30db adjacent channel attenuation, results in a receiver in which the ability to reject adjacent channel interference is limited only by the non-linearity and selectivity of the RF tuner. Thus a simple and inexpensive IF filter, in conjunction with a very simple IF amplifier, can exceed the performance of multi-tube, multi-tuned circuit designs that were common a few years back.

The author wishes to extend his thanks to William Jacobus, whose efforts greatly facilitated this work.

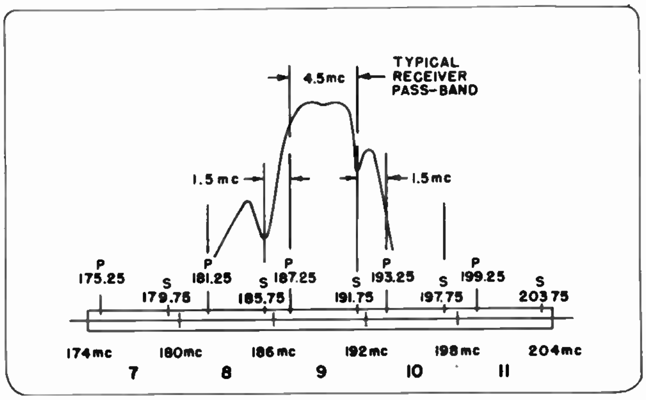


Fig. 1 - Television channel allocations.

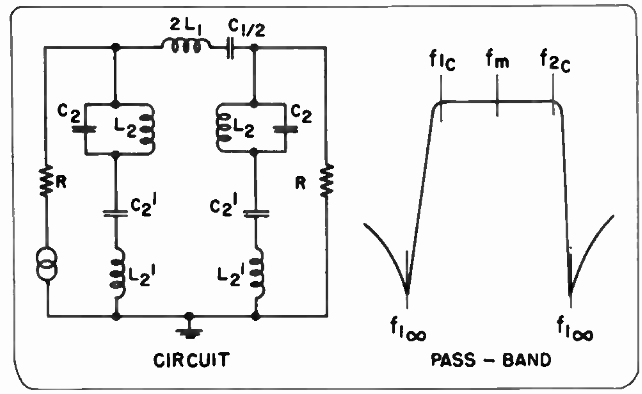


Fig. 2 - Six-element "M" derived band pass.

$$L_{1K}=\frac{R}{\omega_{2c}-\omega_{1c}},\quad C_{1K}=\frac{\omega_{2c}-\omega_{1c}}{R\,\omega_{1c}\,\omega_{2c}},\quad m_1=\frac{\dfrac{\omega_{1c}\,\omega_{2c}}{\omega_{2\infty}^2}\cdot g+h}{1-\dfrac{\omega_{1\infty}^2}{\omega_{2\infty}^2}}$$

$$g=\sqrt{\left(1-\frac{\omega_{1\infty}^2}{\omega_{1c}^2}\right)\left(1-\frac{\omega_{1\infty}^2}{\omega_{2c}^2}\right)},\quad h=\sqrt{\left(1-\frac{\omega_{1c}^2}{\omega_{2\infty}^2}\right)\left(1-\frac{\omega_{2c}^2}{\omega_{2\infty}^2}\right)},\quad m_2=\frac{g+h\,\dfrac{\omega_{1\infty}^2}{\omega_{1c}\,\omega_{2c}}}{1-\dfrac{\omega_{1\infty}^2}{\omega_{2\infty}^2}}$$

$$L_1=m_2L_{1K},\quad L_2=\frac{L_{1K}}{m_2}\left[\frac{(\omega_{2c}-\omega_{1c})^2}{\omega_{1c}\,\omega_{2c}}-\frac{(m_1-m_2)^2}{m_1\,m_2}\right],\quad L_2'=\frac{1-m_1^2}{m_1}\cdot L_{1K}$$

$$C_1=\frac{C_{1K}}{m_2},\quad C_2=\frac{m_1C_{1K}}{\dfrac{(\omega_{2c}-\omega_{1c})^2}{\omega_{1c}\,\omega_{2c}}-\dfrac{(m_1-m_2)^2}{m_1\,m_2}},\quad C_2'=\frac{m_2}{1-m_2^2}\,C_{1K}$$

Fig. 3 - Design relationships.

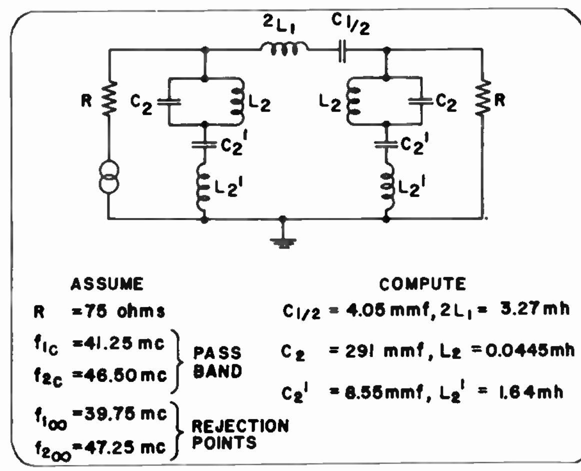


Fig. 4 - Computed values.

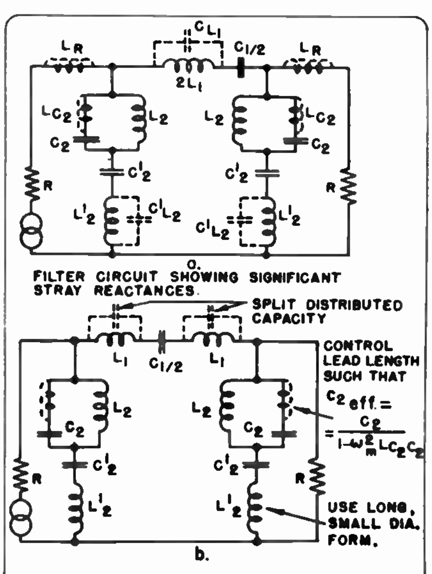


Fig. 5 - Modification to minimize stray reactance.

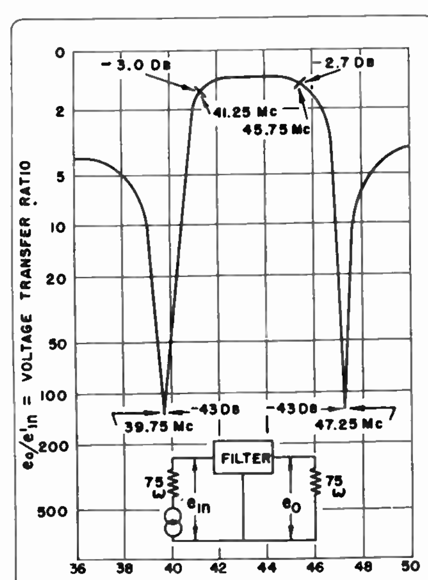


Fig. 6 - Voltage transfer ratio.

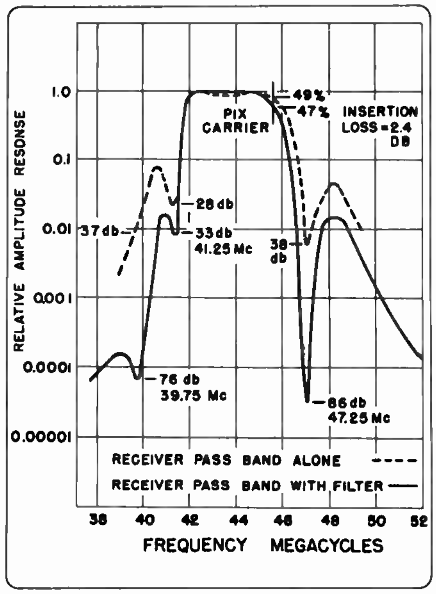


Fig. 7
Receiver pass with and without filter.

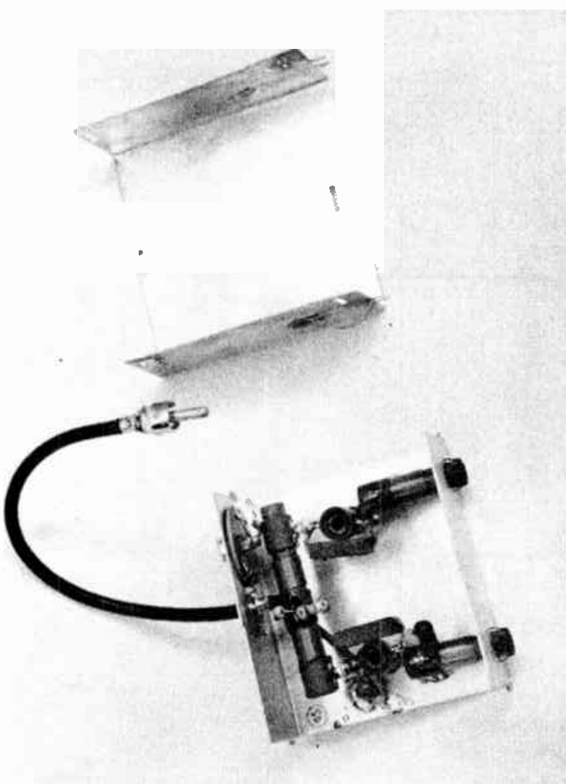


Fig. 8
Assembled filter.

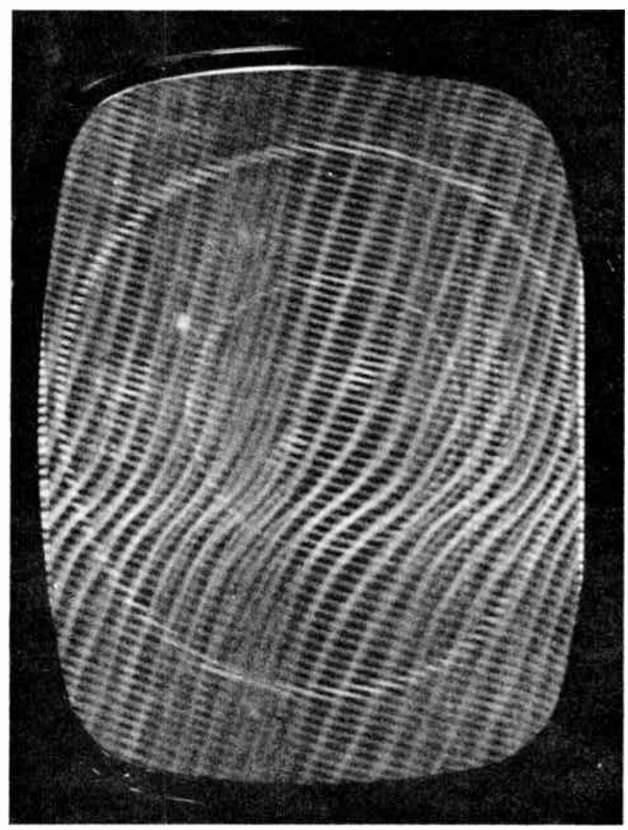


Fig. 9
Channel 3 interference without filter.

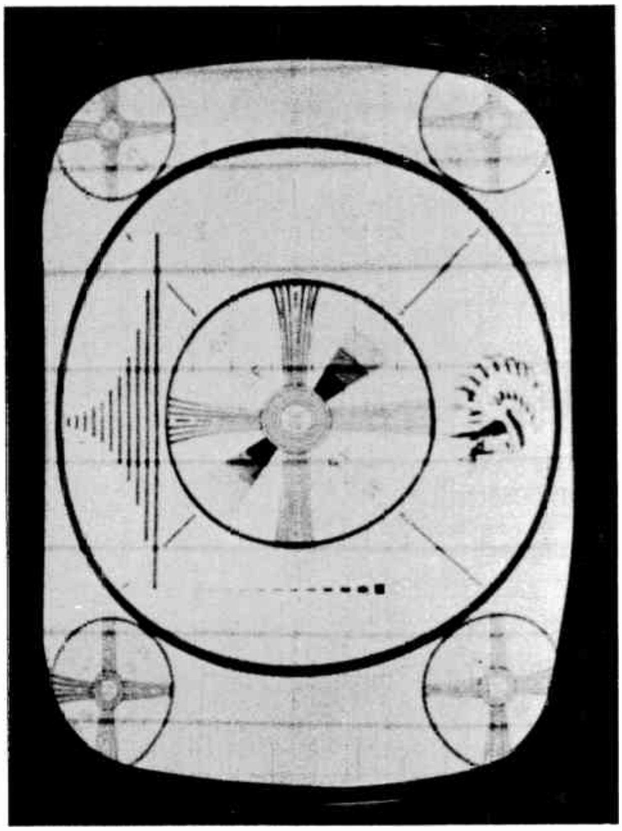


Fig. 10
Receiver on channel 3 with filter.

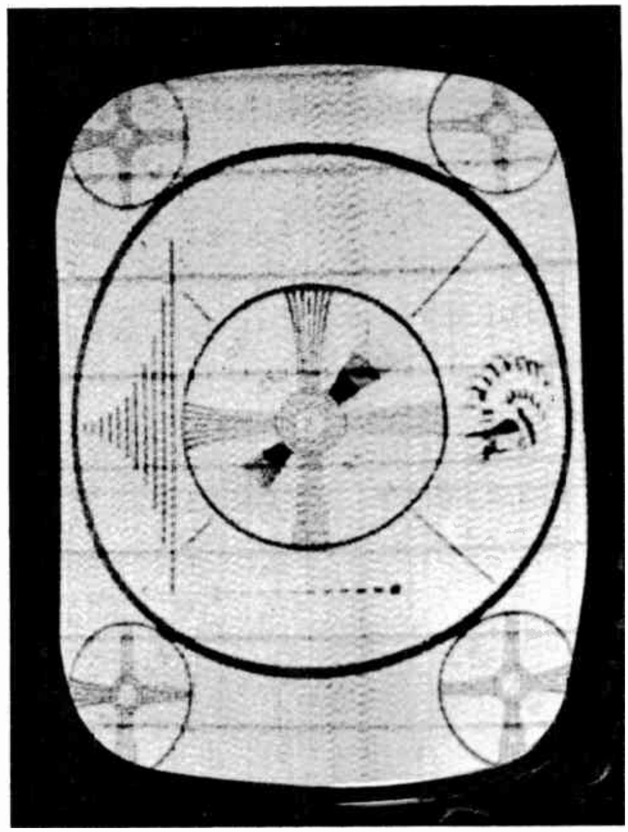


Fig. 11
Effects of RF cross modulation.

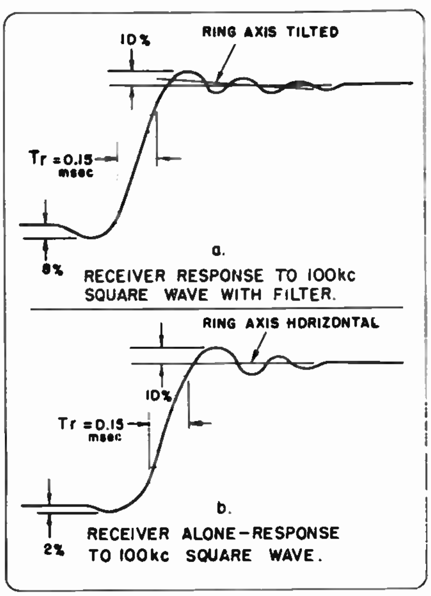


Fig. 12
Receiver transient response with and without filter.

METHODS OF MATRIXING IN AN NTSC COLOR TELEVISION RECEIVER

Will M. Quinn, Jr.
General Electric Company
Syracuse, New York

Introduction

An inherent problem in the NTSC color television receiver is that of combining the chrominance components with the luminance signal to obtain the red, blue, and green color signals for application to the control elements of a tri-color picture tube. Various well-known adder techniques have been employed in receivers to accomplish these operations. It is the purpose of this paper to present some of these techniques with a discussion of the relative merits of each.

Basic Principles of the NTSC Color Television System

For satisfactory color reproduction of the original scene being televised, three independent pieces of information are transmitted; luminance (or brightness), dominant wavelength (or hue) and purity (or saturation). The luminance information is transmitted in accordance with present monochrome standards as a full resolution black and white picture signal complete with horizontal and vertical synchronizing information. This is the compatible portion of the signal and will result in a high quality picture reproduction on existing monochrome receivers. The information corresponding to the dominant wavelength and purity is transmitted as phase and amplitude modulation respectively of a color subcarrier which is located 3.579545mc away from the picture carrier. Its frequency is so chosen that it is an odd multiple of half line-frequency and tends to cancel itself on alternate frames in existing monochrome receivers.[1]

Present NTSC field test specifications require that the signal have the following composition.

$$E_m = E_y + E_Q \sin(\omega t + 33^\circ) + E_I \cos(\omega t + 33^\circ) \qquad (1)$$

where $E_Q = 0.41\ (E_B - E_y) + 0.48\ (E_R - E_y)$ (1a)

$$E_I = -0.27\ (E_B - E_y) + 0.74\ (E_R - E_y) \qquad (1b)$$

$$E_y = 0.30E_R + 0.59E_G + 0.11E_B \qquad (1c)$$

and $\omega = 2\pi$ times the subcarrier frequency

For color difference frequencies below 500KC equation 1 may be represented by

$$E_m = E_y + \left[\frac{1}{1.14}\left(\frac{1}{1.78}(E_B - E_y)\sin\omega t + (E_R - E_y)\cos\omega t\right)\right] \qquad (2)$$

The first term of equations 1 and 2 supplies the luminance or y information which is usually obtained by combining certain proportions of the red, green and blue signal outputs of a tri-color camera or other pickup device. These proportions, specified by equation 1c, show that the red, green and blue reproducing primaries contribute 30, 59, and 11 percent respectively of the luminance of white. The remaining portions of equations 1 and 2 consist of the two chrominance components or color difference signals which supply the color information. This information may be recovered in a color television receiver through a process known as synchronous detection, that is, by heterodyning the color subcarrier and its sidebands with reference frequency of the same phase as the subcarrier component which carries the desired modulation. Subcarrier phase relationships of the NTSC signal are shown in Figure 1. A receiver may recover the color information by synchronous detection along either the E_I or E_Q (abbreviated I,Q) axes or along the $E_R - E_y$, $E_B - E_y$ (R-Y,B-Y) axes. There are certain advantages and disadvantages to detection along either set of axes which are beyond the scope of the material presented and will not be considered. However, regardless of whether the I,Q or R-Y, B-Y components are recovered in the color receiver the problem of matrices still exists.

The problem is somewhat simplified for the R-Y, B-Y receiver because these components need only be added to the luminance signal to produce the red and blue signals for application to the picture tube control element, that is,

$$\begin{aligned} R-Y + Y &= R \\ B-Y + Y &= B \end{aligned} \qquad (3)$$

The G-Y color difference signal may be directly obtained through synchronous detection at the proper phase angle or by the combination of certain negative proportions of the R-Y and B-Y components as may be proven by the substitution of equation 1c into equation 4.

$$G-Y = -0.51\ (R-Y) - 0.19\ (B-Y) \qquad (4)$$

The green color signal is then produced by the addition of the luminance signal to equation 4.

A block diagram of a typical NTSC receiver which recovers the R-Y, B-Y color components is shown in Figure 2.

Following the second detector in such a receiver, the composite signal is separated into the luminance channel and the chrominance channel.

In the luminance channel the color subcarrier and its lower frequency modulation components are removed by a low-pass filter or trap leaving only the brightness or Y signal.

Conversely the low frequency components of the luminance signal are removed from the chrominance channel by means of a band-pass filter. The remaining color subcarrier and its sidebands are then applied to the synchronous detectors where they are heterodyned with reference carrier of appropriate phases. Low-pass filters in the detector plate circuits recover the desired difference frequencies by excluding all other components. The G-Y color difference signal is obtained from the output of an adder which combines the R-Y and B-Y detector outputs in accordance with equation 4. The three-color difference signals are then individually combined with the luminance signal in adder circuits to produce the red, blue, and green signals necessary for color picture reproduction.

NTSC Color Bar Waveforms

Perhaps the operations described may best be understood by referring to Figure 3 in which a transmitted signal of saturated color bars has been assumed. Amplitudes shown are relative to a peak amplitude of unity for white. Figure 3a shows the signal after subcarrier filtering which leaves only the luminance signal. It may be recalled from equation 1c that $Y = 0.59G + 0.30R + 0.11B$. Brightness values for the color bars shown may be obtained by direct substitution into equation 1c. For example, yellow is composed of equal parts of red and green. Its brightness value is, therefore, $0.59 + 0.30 = 0.89$. Figures 3_b, 3_c, 3_d show the waveforms of the three color difference signals corresponding to the assumed color bar signal. It is the function of the receiver matrix unit to combine these signals individually with the luminance or Y signal to produce the three simultaneous red, green, and blue color signals. Considering the addition of the B-Y color difference signal to the brightness signal it is seen that for bars which contain no blue (e.g. red, yellow) the positive luminance values are exactly cancelled by negative values in the B-Y color difference signal. Those bars which contain blue (cyan, blue) produce positive components in the color difference signal which add to the positive luminance values to produce a peak value of unity (Figure 3_e). If the R-Y and G-Y color difference signals are added to the Y signal in the same manner the red and green color signals are produced (Figure 3_f, 3_g respectively). The blanking interval has been purposely omitted from Figures 3_e, 3_f, 3_g since it contributes nothing to the discussion.

Adder Requirements

Although the requirements of an adder for use in color television receivers are not as stringent as for many other applications, there are certain requirements that a good adder must meet for satisfactory performance. These are that it be simple, that it be capable of passing the highest frequencies present in the input signal, that it have negligible interaction between sources, and that it be free from distortion. Also desirable is that the adder be efficient or capable of performing the desired operations without severe attenuation of signal between input and output.

Picture Tube Adding

Perhaps the most frequently used and obvious method of combining the color difference signals and the luminance signal consists of applying the color difference signals to the three cathodes (or grids) of a tri-color picture tube and the luminance signal to the three grids (or cathodes), thereby utilizing the picture tube itself as an adder. This method possesses the advantages of simplicity and possible economy in that the three signals required at the three cathodes are R-Y, B-Y, and G-Y. For the R-Y, B-Y receiver the only operation needed after detection (neglecting gain) is to produce the G-Y signal which may be obtained by either of the methods described previously. However, a receiver which recovers I and Q must combine certain positive and negative amounts of these components in order to present the cathodes with the particular color difference signals required. Therefore, this method is not as advantageous for the latter type of receiver.

One disadvantage of using the picture tube as an adder is that the rather large tube and stray capacities encountered when feeding three parallel control elements place somewhat higher requirements upon the luminance amplifier in order to realize the required frequency response of 3.5mc.

Resistive Adders

Perhaps the most simple method of adding two or more signals is the linear passive network of the resistive adder shown in Figure 4.[2] In order to maintain the required isolation between sources in this type of network, the insertion resistors (R_1, R_2, R_3) must be large compared to the load resistor R_o. The output voltage becomes approximately

$$e_o = \frac{R_o}{R_1} e_1 + \frac{R_o}{R_2} e_2 + \frac{R_o}{R_3} e_3 \qquad (5)$$

If sources e_2 and e_3 are replaced by their respective source impedances Z_2 and Z_3 where $Z_2 = Z_3 = Z$ the coupling between sources may be shown to be

$$e_c = e \left(\frac{1}{1 + \frac{R}{R_o}} \right) \left(\frac{1}{1 + \frac{R}{Z}} \right) \qquad (6)$$

where e represents any source voltage and $R = R_1 = R_2 = R_3$ for convenience. For e_c to be small enough to be tolerated the ratio of $\frac{R}{R_o}$ must be large and the source impedance must be small compared to the insertion resistor. Results have shown that for the required accuracy the coupling factor $\frac{e_c}{e}$ must be five percent or less. At this figure the error introduced through cross coupling becomes small enough to be imperceptible.

This particular type of adder is very inefficient due to the attenuation needed to produce the required source isolation. Additional stages of gain must, therefore, be employed to compensate for this attenuation.

Other purely resistive configurations are possible but, in general, do not provide the necessary decoupling between sources and will not be considered. Purely inductive or purely capacitive configurations may also be used but are more expensive and tend to become troublesome at the higher frequencies.

Tube Adders

Other common methods used to perform the matrix functions utilize tube sections with common plate or common cathode impedances as shown in Figures 5 and 6.[2,3] For the two input voltage shown in Figure 5, the output voltage becomes

$$e_o = \frac{\mu}{2(\mu + 1) + \frac{r_p}{R_K}} (e_1 + e_2) \qquad (7)$$

When triodes of sufficient μ and low plate resistance are used, equation 7 reduces to its approximate form

$$e_o = \frac{1}{2} (e_1 + e_2) \qquad (8)$$

Common cathode adders are normally used whenever it is necessary to preserve phase without realizing gain.

Since some gain is usually desirable, common plate adders are generally used when a tube adder is desired, Figure 6.

$$e_o = K(e_1 + e_2) \qquad (9)$$

For pentodes

$$e_o = gm\, R_L (e_1 + e_2) \qquad (10)$$

assuming identical tubes.

Tube adders of this type provide excellent isolation between sources because of their high input impedance. The degree of isolation becomes primarily a function of tube and stray capacitances. A second advantage of importance is that impedance transformations are readily effected between input and output. Relative amplitudes of input signals desired in the output may be obtained by (1) adjusting the relative amplitudes of the input signals using equal tube section gains or (2) adjusting the gains of the respective tube sections using equal inputs.

Dual triodes are suitable for adders, particularly for those applications requiring the addition of only two signals. Adders employing tube type 12AT7 have been especially successful, providing linear addition for input signals of one volt or less with moderate gain in the case of plate adders. As is true with most wide band circuits, some form of peaking must be used in conjunction with tube adders to provide the required frequency response.

Disadvantages of tube adders which employ common plate or common cathode impedances are that one tube section is required for each signal source and that the matrix accuracy depends directly upon tube characteristics.

Feedback Summing Amplifiers

Perhaps the most successful method of matrixing employed is the feedback summing amplifier.[2,3,4] This circuit consists of a high gain amplifier with inverse feedback from plate to grid and which contains a resistive summing network in its grid circuit. This particular adder, therefore, combines the advantages of a passive adding network with those afforded by negative voltage feedback.

Basic requirements for correct operation of this adder are that there be no current drawn by the input (grid) circuit and that the open loop gain (K) of the stage be high in order to insure accuracy and provide sufficient source decoupling. Applying Kirchoff's current law to Figure 7

$$\frac{e_3 - e}{R_3} + \frac{e_2 - e}{R_2} + \frac{e_1 - e}{R_1} = \frac{e - e_o}{R_f} \qquad (11)$$

where

$$e = \frac{e_o}{K}$$

Substituting for e in equation 12 and solving for e_o

$$e_o = \frac{\frac{R_f}{R_1} e_1 + \frac{R_f}{R_2} e_2 + \frac{R_f}{R_3} e_3}{\frac{1}{K} - 1 + \frac{1}{K}\left(\frac{R_f}{R_1} + \frac{R_f}{R_2} + \frac{R_f}{R_3}\right)} \qquad (12)$$

If K is sufficiently large as is generally the case when pentodes are used

$$e_o = -\left[\frac{R_f}{R_1} e_1 + \frac{R_f}{R_2} e_2 + \frac{R_f}{R_3} e_3\right] \qquad (13)$$

The output voltage, is therefore, the negative sum of certain ratios of the input signals, these ratios being determined by the ratios of the feedback resistor to the series insertion resistors.

To compute the isolation between sources afforded by the feedback summing amplifier, refer to Figure 8 in which sources e_2 and e_3 have been replaced by their internal impedances, Z_2 and Z_3 respectively. Source e_1 becomes the only externally applied voltage and e_2' and e_3' refer to the voltages developed across Z_2 and Z_3 due to cross coupling of e_1.

When the feedback resistor is equal to the insertion resistor, the voltage at the adder grid is given approximately by

$$e = \frac{e_1}{K} \qquad (14)$$

where K is the open loop gain of the stage. The voltage developed across Z_2 due to a current from source e_1 is seen to be

$$e_2' = \frac{e_1}{K} \left(\frac{1}{1 + \frac{R_2}{Z_2}} \right) \qquad (15)$$

and similarly

$$e_3' = \frac{e_1}{K} \left(\frac{1}{1 + \frac{R_3}{Z_3}} \right) \qquad (16)$$

Since K can easily be made equal to 30 or more and Z_2 and Z_3 are normally much less than R_2 and R_3 the cross coupling is usually substantially less than one percent in practical circuits.

Application to an NTSC Receiver

The application of the feedback summing amplifier to a typical I, Q receiver is illustrated in Figure 9. The required matrices are:

$$\begin{aligned} R &= 0.956\,(I) + Y + 0.621(Q) \\ G &= 0.272\,(-I) + Y + 0.647(-Q) \\ B &= 1.106\,(-I) + Y + 1.703(-Q) \end{aligned} \qquad (17)$$

The ratios of $\frac{R_f}{R_1}$, $\frac{R_f}{R_2}$, ——— $\frac{R_f}{R_9}$ are predetermined by the coefficients of I, Y and Q of the above equations and must be adjusted to correspond to these values.

The example given considered only a receiver which recovered the I and Q chrominance components since this type of adder is particularly adaptable to receivers of this type. The principles involved apply equally well to a receiver which recovers the R-Y and B-Y components through detection. Such a receiver must perform the following operations after detection:

$$(R-Y) + Y = R$$

$$(B-Y) + Y = B$$

$$-0.51(R-Y) - 0.19(B-Y) + Y = G$$

Since the Y or luminance signal input contains frequency components extending to 3.5mc a response which is essentially flat to that frequency is desired. Results have shown that through minimizing the effects of shunt capacity to ground by careful selection of impedance levels in the summing network a flat response may be obtained without peaking and with more than sufficient isolation between sources. For example, when the circuit parameters below were used,

Tube Type 6AH6

$K = 50$

$R_1 = R_2 = R_3 = R_f = 20K$

$Z_s = 1K$ = Source Impedances

$R_L = 5K$

a frequency response of 4mc was obtained without peaking. Under these conditions, the coupling factor was computed to be $\frac{1}{1050}$.

Conclusion

Several adder techniques have been described, any of which will perform satisfactorily. Of these the feedback summing amplifier seems to possess certain advantages. These are improved frequency response, better linearity, increased stability, and better isolation between sources. This type of adder has been employed in developmental NTSC color television receivers very successfully, giving superior performance with improved stability over other types of adder circuits tried.

References

[1] C. J. Hirsch, W. F. Bailey, B. D. Loughlin, "Principles of NTSC Compatible Color Television," Electronics, February, 1952.

[2] Radiation Laboratory Series, Volume 19, "Waveforms," McGraw-Hill Book Company, Inc., New York, 1949.

[3] S. Seely, "Electron Tube Circuits," McGraw-Hill Book Company, Inc., New York, 1950.

[4] D. H. Dickens, "Electronic Analog Computer Fundamentals," Electronics, August, 1952.

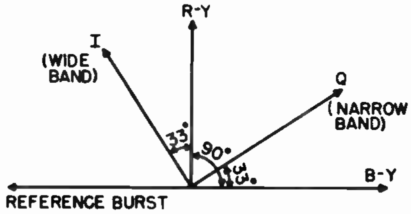


Fig. 1
Subcarrier phase relationships of the NTSC color television signal.

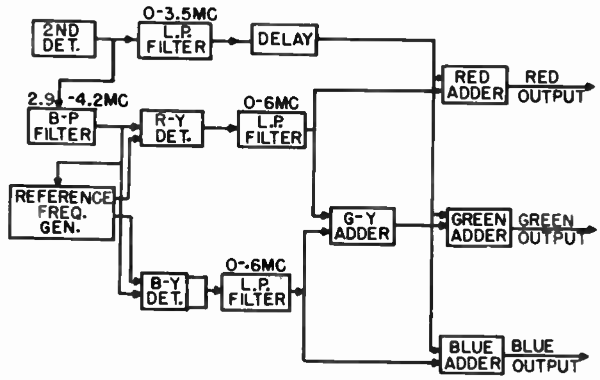


Fig. 2
Block diagram of a typical NTSC receiver.

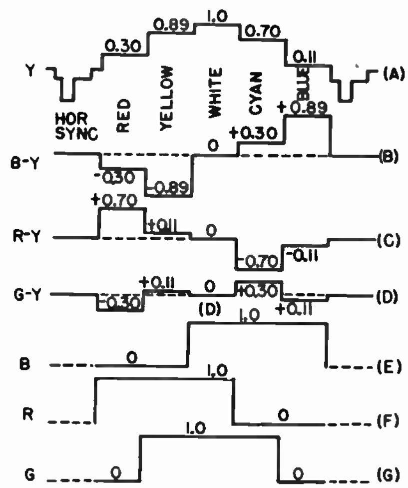


Fig. 3
NTSC color bar waveforms and amplitudes.

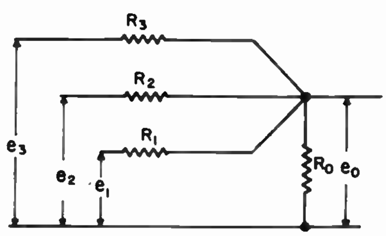


Fig. 4
Resistive adding network.

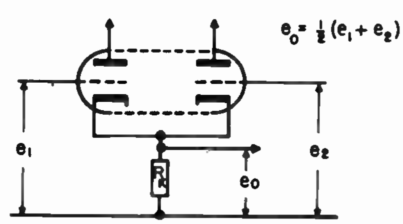


Fig. 5
Common cathode adder.

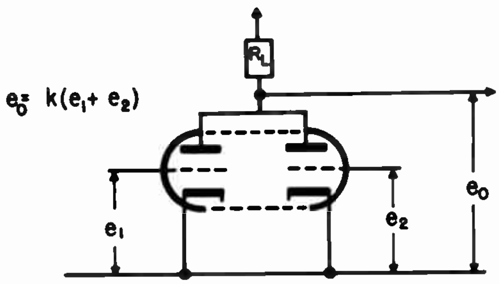


Fig. 6
Common plate adder.

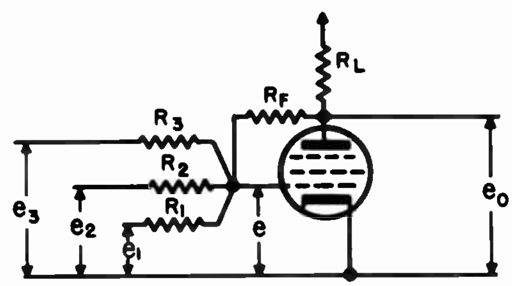


Fig. 7
Feedback summing amplifier.

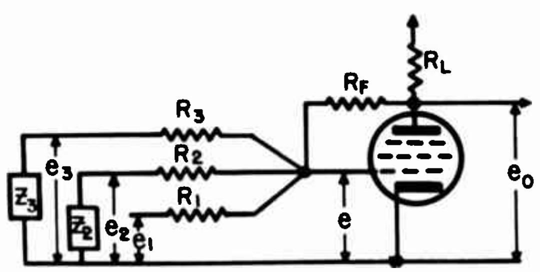


Fig. 8
Feedback summing amplifier circuit used in computing interaction between sources.

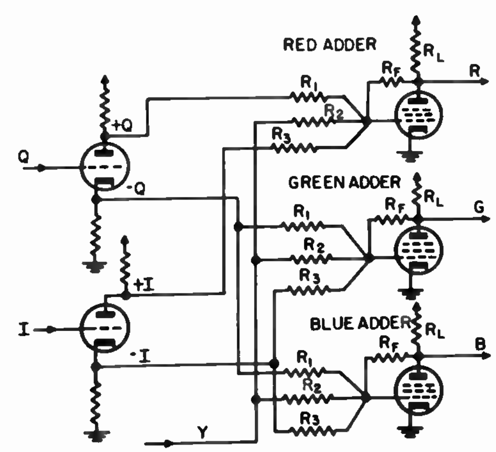


Fig. 9
Feedback summing amplifier used in and I-Q receiver.

$$R = 0.966I + Y + 0.621Q;$$
$$G = -0.272I + Y - 0.647Q;$$
$$B = -1.106I + Y - 1.703Q.$$

FACTORS AFFECTING THE DESIGN OF V H F-U H F TUNERS

By E. H. Boden - Advanced Application Engineer
Sylvania Electric Products Inc.
Emporium, Pennsylvania

INTRODUCTION

Since the addition of the seventy new u h f television channels, a variety of methods have been employed to add these channels to a television receiver. The methods employed range from external and internal converters, which cover the entire range, to single channel strips. Oscillator injection is obtained from fundamental oscillators, harmonic multipliers, or combinations thereof. Recently, some manufacturers have been installing both a v h f tuner and a u h f tuner. Each one of these methods has its advantages and disadvantages.

Prior to the addition of the u h f channels, the tuner, or head end, had the form shown in Figure 1. Here, an r f amplifier (one or two tubes), a mixer, and an oscillator were used. Until recently, the absence of suitable tubes for u h f channels had caused a deviation from this configuration. In order to return to the configuration in Figure 1, and include all the v h f and u h f channels in a single tuner, new tubes and techniques had to be developed.

The requirements placed on the circuit are much the same as for v h f with the addition that the tuning arrangement must add an approximate 2:1 frequency skip band and then tune another 2:1 frequency range of channels.

There are three important requirements placed on tubes for a combined u h f-v h f tuner.

1. They should provide equivalent performance on the v h f channels to that found with v h f tubes, and comparable performance on the u h f channels.

2. They should be designed for large volume production on existing machinery.

3. They should be priced competitively with v h f tubes for similar purposes.

Sylvania has developed tubes to meet the above requirements and is now manufacturing them in large quantities. These tubes are the type 6AN4, r f amplifier and mixer tube and the type 6T4, oscillator tube. Both the 6AN4 and 6T4 are seven-pin miniature tubes in a short bulb.

We shall now discuss some of the factors affecting the use of these tubes in a tuner capable of tuning both the v h f and u h f channels with no duplication of stages.

R F AMPLIFIER

The r f amplifier section has three equally important parts. They are, as shown in Figure 2, the input circuit, the tube, and the output circuit.

The design of the input circuit is of great importance to the overall performance of the tuner. If we think of the tuner as being of the form shown in Figure 3, where N_1 and L_1 are the noise figure and insertion loss of the input circuit, and N_2 is the noise figure of the remaining part of the receiver, we see that the overall noise figure is approximately N_2L_1. Therefore, the overall noise figure is increased by the insertion loss of the input circuit. This means that the input circuit must have a high no-load Q and be loaded to give minimum insertion loss. Since the input circuit will be followed by an amplifier stage which will attenuate the oscillator signal to the antenna, rejection at the oscillator frequency will not be required, thus permitting greater loading of the input circuit.

As was shown for the insertion loss, it may also be shown that losses due to mismatch and unbalance at the antenna input will cause the noise figure to be higher. Also, a loss in delivering the signal to the amplifier tube reduces the effectiveness of the amplifier to decrease the noise contributed by the stages that follow.

To prevent cross modulation at the input, the input circuit should present an insertion loss to signals other than the desired signal. The pass band of the input circuit should not need to be as narrow on u h f as on v h f due to the proposed frequency spacing of the u h f television stations.

The input circuit may use lumped or distributed constant elements. In some cases, it may be more feasible to use a combination of both so as to tune both the v h f and u h f channels and require a minimum of space. Some suggested input circuits are shown in Figure 4.

For maximum gain, the amplifier tube should have a high mu and g_m with low interelectrode capacitances and lead inductances. The Sylvania type 6AN4 has an average mu of 70, a transconductance of 10,000 micromhos when operated with 200 volts on the plate, and a 100-ohm cathode resistance. The grid to plate capacitance is 1.7 micromicrofarads while the grid to cathode capacitance is 2.2 micromicrofarads.

The input resistance of the 6AN4, including transit time and $1/g_m$, is approximately 100 ohms at v h f. At 470 mc, the input resistance is 50 ohms, increasing to 150 ohms at 900 mc. The increasing resistance is due to resonant effects of the cathode lead inductances and tube and socket capacities.

If a g c is used on the tube, it should be delayed or tapped down from the a g c bus. Otherwise, the noise figure may rise with signal strength. This could cause a snow-free signal to become snowy.

Performance measurements have been made at various points in the vh f and u h f television band. Conventional lumped constant circuits were employed on the vh f channels while u h f measurements were made in the amplifier shown in Figure 5. The amplifier shown in Figure 5 is tunable from 470-890 mc. In this circuit, the amplifier had an insertion power gain of 10 db at 900 mc with a bandwidth of 10 mc. The amplifier noise figure was 14 db at 900 mc.

Since the maximum gain of an amplifier for a given bandwidth is directly proportional to the ratio of transconductance to total plate-to-ground capacity, it is important that the circuit capacity be kept at a minimum. Plate-to-tube shield capacity is included in the total circuit capacity and may be reduced by use of a T-6 1/2 shield.

The tuning arrangement employed should not add to the total circuit capacity. The circuit shown in Figure 5 can be modified to include the v h f channels. This may be accomplished by fixing the line length for Channel 83 and capacitively loading the open end to tune to Channel 14. VH F channels may then be tuned by shunting lumped inductances across the tuning capacitor as shown in Figure 6. For the turret tuner and decade tuner, lumped constants have been used successfully to 900 mc.

MIXER CIRCUITS

The mixer circuit is quite simple. Having matched down from the amplifier plate impedance to the mixer cathode impedance, it remains only to add the necessary oscillator injection and couple the i f out at the plate.

The performance of the 6AN4 as a mixer is quite attractive because of its constant input impedance with oscillator injection variations and its conversion gain. When operated with 125 volts on the plate and 270 ohms in the cathode, a conversion transconductance of 2900 micromhos is obtained. This value of G_c is obtained with 1.4 volts rm s oscillator injection. The curve in Figure 7 shows the variation of G_c and plate current with oscillator injection. A careful study of this curve shows that the G_c is 10% down at 1 and 5 volts r m s and down to 0.7 at 0.6 and 6.4 volts injection. The input resistance change over this range of oscillator injection is negligible and of a value a little greater than the same tube as an amplifier.

The plate current curve serves as an indicator of the amount of oscillator injection to the tube. This is helpful when adjusting the tube in a tuner.

There is no oscillator injection shown in Figure 6. This is because of the variety of injection methods that may be employed. Injection into the cathode, heaters and grid have all been successfully tried. When injecting into the cathode, care must be taken that a portion of the received signal is not coupled out of the mixer. Also, when injecting into the grid, the impedance from grid to ground must be low for the received signal and the i f signal.

Although cathode bias was indicated above, other forms of biasing may and have been used. There seems to be an indication of some improvement in the noise figure when both cathode and grid biasing are used. If some grid biasing is used, then the I_b curve in Figure 7 will no longer be true. In this case, the oscillator injection is measured at the grid.

Noise figures for the mixer at 500 and 900 mc are 16 and 20 db respectively. Conversion gains of 3 db are realized when operated with between 1 and 5 volts of injection. With conversion gain and r f gain, special low-noise, pre-i f amplifier stages are unnecessary.

LOCAL OSCILLATOR

In the oscillator section there is the problem of maintaining a relatively constant output over the range of 100 to 950 mc, free from suckouts and with a minimum frequency drift.

The modified colpitts oscillator is the most common oscillator circuit at v h f and u h f. Tuning may be by a shorted line with a movable contact or capacitive short, capacitive tuning an open line, or capacitively tuning a lumped inductance. V H F and u h f channels may be tuned by the first method and by a combination of the second two.

Balanced tuning is preferred (plate and grid approximately the same impedance above ground) so that the large circulating currents between grid and plate are not in the chassis or tuner box. This is quite important for the reduction of oscillator radiation, frequency drift and dead spots.

The Sylvania 6T4, with 80 volts on the plate and a 10,000 ohm grid resistance, will develop 6 to 9 volts at the grid at 950 to 500 mc and 12 volts on v h f.

Common causes for dead spots are resonance in the unused portion of the tuning line, resonance between chassis and line, and resonance in heater and cathode chokes. It is recommended that both heater connections be kept at the cathode potential. This is best done by using two chokes. One choke is used for the cathode and one heater, the other for the second heater. Then, place a capacitor across the heaters. By so doing, the variations in heater-cathode capacity between tubes and variations due to temperature changes will be removed.

PERFORMANCE

Having now viewed the tube and circuit possibilities and techniques, it remains to examine the performance figures for various tuner configurations. Examining the several possibilities, there seem to be three probable arrangements. For purpose of discussion, we shall call these tuners A, B and C. These are shown in Figure 8. Tuner A has two grounded grid r f amplifiers in cascade on v h f and u h f. Tuner B has a cascode on v h f and a single stage of amplification on u h f. Tuner C has one stage of grounded grid amplification. All tuners have a 6AN4 grounded grid mixer and a 6T4 oscillator. Also, all three tuners tune the v h f and u h f channels.

Performance figures on these tuners include noise figure and insertion power gain at Channels 2, 13, 14 and 83. Comparison with v h f tuners is made with the more recent cascode v h f tuner.

Looking at noise figure first, in Figure 9 is a table of the noise figures of tuners A, B and C, on Channels 2, 13, 14 and 83. Tuner B with its cascode v h f has equivalent performance on v h f as a cascode v h f tuner, but does not have a comparable u h f performance. Tuner A, however, has nearly an equivalent noise figure on v h f as a standard cascode and a comparable performance on u h f. It would seem that tuner A would make a good compromise tuner. Tuner C could be said to have a fair noise figure on v h f and u h f. Such a tuner might serve well in a low cost set.

In the table shown in Figure 10 are the insertion power gains for these same tuners on the same v h f and u h f channels. Here, we see that tuners A and C have the most uniform gains in going from Channel 2 to Channel 83. However, it can be said that the loss in gain on the u h f channel on tuner B is accompanied by a nearly corresponding increase in noise figure and therefore would have the same sensitivity factor on u h f as on v h f.

CONCLUSION

It is concluded from this investigation, that a single v h f-u h f tuner is feasible and that the addition of the u h f channels may be done without degradation of the v h f performance, without greatly increasing the cost of the tuner, and without duplication of stages. It may further be concluded that low-cost, standard, miniature tube types are available for the purpose.

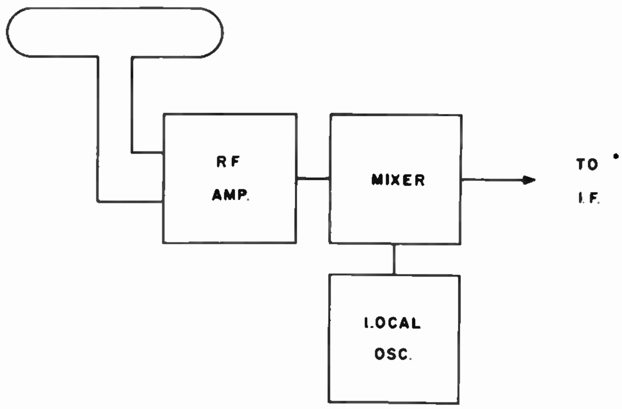


Fig. 1
Block diagram of vhf TV tuner.

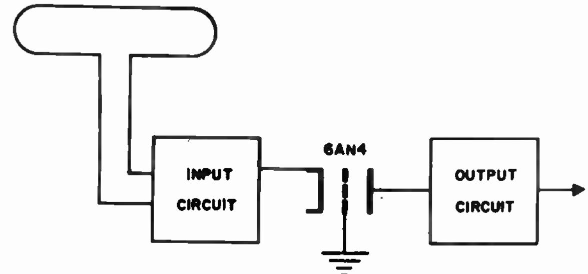


Fig. 2
RF amplifier.

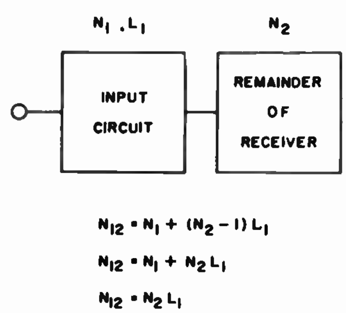


Fig. 3
Noise figure and the input.

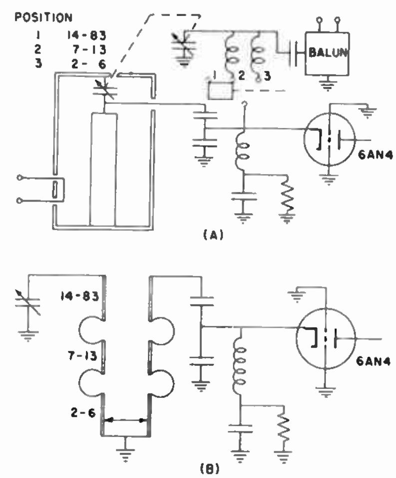


Fig. 4
Combined uhf-vhf circuits.

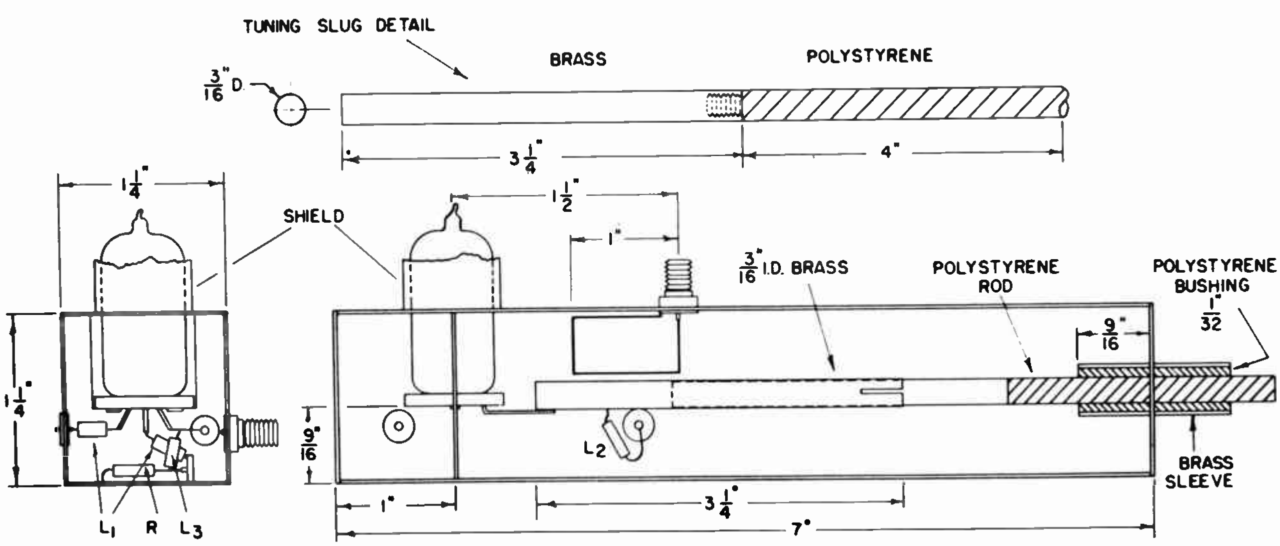


Fig. 5
Layout of coaxial line amplifier.

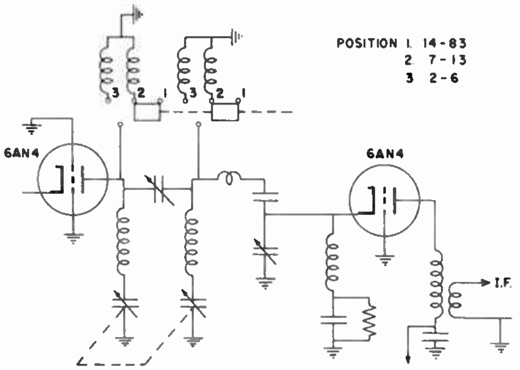


Fig. 6
VHF-UHF amplifier tuning.

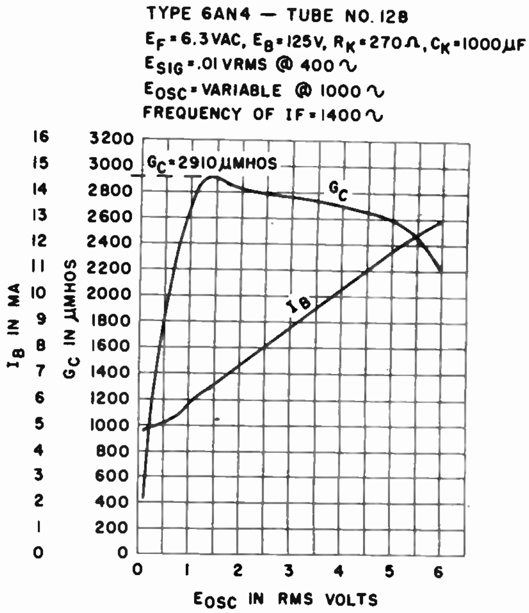
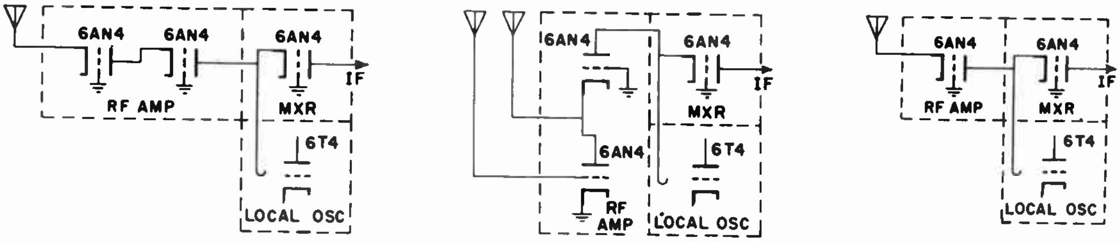


Fig. 7
Variation of plate current and
conversion conductance with
oscillator injection voltage.



(a) (b) (c)

Fig. 8
Three possible tuner arrangements.

CHANNEL	TUNER NOISEFIGURES (DB)		
	A	B	C
2	7	6	9
13	9.5	8	11
14	12	14	14
83	14	16	16

Fig. 9
Noise figure of tuners
(a), (b), and (c) of Fig. 8.

CHANNEL	TUNER GAIN IN DB		
	A	B	C
2	25	24	17
13	25	24	17
14	23	15	15
83	20	12	12

Fig. 10
Insertion power gain for tuners
(a), (b), and (c) of Fig. 8.

THEORY OF A.F.C. SYNCHRONIZATION

Wolf J. Gruen
General Electric Co.
Syracuse, N. Y.

Summary---The general solution for the important design parameters of an automatic frequency and phase control system is presented. These parameters include the transient response, frequency response and noise bandwidth of the system, as well as the hold-in range and pull-in range of synchronization.

I. Introduction

Automatic frequency and phase control systems have been used for a number of years for the horizontal sweep synchronization in television receivers, and more recently have found application for the synchronization of the color subcarrier in the proposed N.T.S.C. color television system. A block diagram of a general A.F.C. system is shown in Fig. 1. The phase of the transmitted synchronizing signal e_1 is compared to the phase of a local oscillator signal e_2 in a phase discriminator D. The resulting discriminator output voltage is proportional to the phase difference of the two signals, and is fed through a control network F to a frequency control stage C. This stage controls the frequency and phase of a local oscillator O in accordance with the synchronizing information, thereby keeping the two signals in perfect synchronism. Although in practice the transmitted reference signal is often pulsed and the oscillator comparison voltage non-sinusoidal, the analysis is carried out for sinusoidal signal voltages. The theory, however, can be extended for a particular problem by writing the applied voltages in terms of a Fourier series instead of the simple sine function. An A.F.C. system is essentially a servomechanism, and the notation that will be used is the one followed by many workers in this field. An attempt will be made to present the response characteristics in dimensionless form in order to obtain a universal plot of the response curves.

II. Derivation of the Basic Equation

If it is assumed that the discriminator is a balanced phase detector composed of peak-detecting diodes, the discriminator output voltage can be derived from the vector diagram in Fig. 2. For sinusoidal variation with time, the synchronizing signal e_1 and the reference signal e_2 can be written

$$e_1 = E_1 \cos \varphi_1 \tag{1}$$

and

$$e_2 = E_2 \sin \varphi_2 . \tag{2}$$

φ_1 and φ_2 are functions of time and, for reasons of simplicity in the later development, it is arbitrarily assumed that φ_1 and φ_2 are in quadrature when the system is perfectly synchronized, that is when $\varphi_1 = \varphi_2$.

While one of the discriminator diodes is fed with the sum of e_1 and $e_2/2$, the other is fed with the difference of these two vectors as shown in Fig. 2. The resulting rectified voltages E_{d1} and E_{d2} can be established by simple trigonometric relations. Defining a difference phase

$$\varphi \equiv \varphi_1 - \varphi_2 , \tag{3}$$

one obtains

$$E_{d1}^2 = E_1^2 + \frac{E_2^2}{4} + E_1 E_2 \sin\varphi \tag{4}$$

and

$$E_{d2}^2 = E_1^2 + \frac{E_2^2}{4} - E_1 E_2 \sin\varphi . \tag{5}$$

The discriminator output voltage e_d is equal to the difference of the two rectified voltages, so that

$$e_d = E_{d1} - E_{d2} = \frac{2 E_1 E_2}{E_{d1} + E_{d2}} \sin\varphi . \tag{6}$$

If the amplitude E_1 of the synchronizing signal is larger than the amplitude E_2 of the reference signal, one obtains

$$E_{d1} + E_{d2} \cong 2 E_1 . \tag{7}$$

The discriminator output voltage then becomes

$$e_d = E_2 \sin\varphi \tag{8}$$

and is independent of the amplitude E_1 of the synchronizing signal. As φ_1 and φ_2 are time varying parameters, it should be kept in mind that the discriminator time constant ought to be shorter than the reciprocal of the highest difference frequency $d\varphi/dt$, which is of importance for the operation of the system.

Denoting the transfer function of the control network F as F(p), the oscillator control voltage becomes

$$e_c = F(p) E_2 \sin\varphi . \tag{9}$$

Assuming furthermore that the oscillator has a linear control characteristic of a slope S, and that the free-running oscillator frequency is ω_o, the actual oscillator frequency in operational notation becomes

$$p\varphi_2 = \omega_o + S e_c . \tag{10}$$

Substituting equations 3 and 9 into equation 10 then gives

$$p\varphi + S E_2 F(p) \sin\varphi = p\varphi_1 - \omega_o \tag{11}$$

The product SE_2 repeats itself throughout this paper and shall be defined as the gain constant

$$K \equiv SE_2 . \tag{12}$$

K represents the maximum frequency shift at the output of the system per radian phase shift at the input. It has the dimension of radians/second.

Equation 11 can be simplified further by measuring the phase angles in a coordinate system which moves at the free-running speed ω_o of the local oscillator. One obtains

$$\boxed{p\varphi + F(p)\,K\sin\varphi = p\varphi_1} . \tag{13}$$

This equation represents the general differential equation of the A.F.C. feedback loop. $p\varphi$ is the instantaneous difference frequency between the synchronizing signal and the controlled oscillator signal and $p\varphi_1$ is the instantaneous difference frequency between the synchronizing signal and the free-running oscillator signal.

Equation 13 shows that all A.F.C. systems with identical gain constants K and unity d.c. gain through the control network have the same steady state solution, provided that the difference frequency $p\varphi_1$ is constant. If this difference frequency is defined as

$$\Delta\omega \equiv p\varphi_1 = \omega_1 - \omega_o , \tag{14}$$

the steady state solution is

$$\sin\varphi = \frac{\Delta\omega}{K} . \tag{15}$$

This means the system has a steady state phase error which is proportional to the initial detuning $\Delta\omega$ and inversely proportional to the gain constant K. Since the maximum value of $\sin\varphi$ in equation 15 is ± 1, the system will hold synchronism over a frequency range

$$\left|\Delta\omega_{Hold-in}\right| \leq K . \tag{16}$$

Equations 15 and 16 thus define the static performance limit of the system.

III. Linear Analysis

An A.F.C. system, once it is synchronized, behaves like a low-pass filter. To study its performance it is permissible, for practical signal to noise ratios, to substitute the angle for the sine function in equation 13. Then, with the definition of equation 3, one obtains

$$p\varphi_2 + K\,F(p)\,\varphi_2 = K\,F(p)\,\varphi_1 . \tag{17}$$

This equation relates the output phase φ_2 of the synchronized system to the input phase φ_1. It permits an evaluation of the behavior of the system to small disturbances of the input phase, if the transfer function F(p) of the control network is specified.

A. F(p) = 1

This is the simplest possible A.F.C. system, and represents a direct connection between the discriminator output and the oscillator control stage. Equation 17 then becomes

$$p\varphi_2 + K\varphi_2 = K\varphi_1 . \tag{18}$$

If the initial detuning is zero, the transient response of the system to a sudden step of input phase $|\varphi_1|$ is

$$\frac{\varphi_2}{|\varphi_1|}(t) = 1 - e^{-Kt} . \tag{19}$$

Likewise, the frequency response of the system to a sine wave modulation of the input phase is

$$\frac{\varphi_2}{\varphi_1}(j\omega) = \frac{1}{1 + j\frac{\omega}{K}} . \tag{20}$$

The simple A.F.C. system thus behaves like an RC-filter and has a cut-off frequency of

$$\omega_c = K \quad \text{radians/sec.} \tag{21}$$

George[1] has shown that the r.m.s. phase error of the system under the influence of random interference is proportional to the noise bandwidth, which is defined as

$$B = \int_{-\infty}^{+\infty} \left|\frac{\varphi_2}{\varphi_1}(j\omega)\right|^2 d\omega . \tag{22}$$

The integration has to be carried out from $-\infty$ to $+\infty$ since the noise components on both sides of the carrier are demodulated. Inserting equation 20 into equation 22 then yields

$$B = \pi K \quad \text{radians/sec.} \tag{23}$$

It was shown in equation 15 that for small steady state phase errors due to average frequency drift, the gain constant K has to be made as large as possible, while now for good noise immunity, i.e., narrow bandwidth, the gain constant has to be made as small as possible. A proper compromise of gain then must be found to insure adequate performance of the system for all requirements. This difficulty, however, can be overcome by the use of a more elaborate control network.

B. $F(p) = \frac{1 + \tau_1 p}{1 + \tau_2 p}$

Networks of this type are called proportional plus integral control networks[2] and typical network configurations are shown in Fig. 3. Inserting the above transfer function into equation 17 yields

$$p^2\varphi_2 + \left(\frac{1}{\tau_2} + K\frac{\tau_1}{\tau_2}\right)p\varphi_2 + \frac{K}{\tau_2}\varphi_2 = K\frac{\tau_1}{\tau_2}p\varphi_1 + \frac{K}{\tau_2}\varphi_1 \, . \tag{24}$$

φ_1 and φ_2 are again relative phase angles, measured in a coordinate system which moves at the free-running speed of the local oscillator. To integrate equation 24, it is convenient to introduce the following parameters

$$\omega_n^2 \equiv \frac{K}{\tau_2} \tag{25}$$

and

$$2\zeta\omega_n \equiv \frac{1}{\tau_2} + K\frac{\tau_1}{\tau_2} \, . \tag{26}$$

ω_n is the resonance frequency of the system in the absence of any damping, and ζ is the ratio of actual to critical damping. In terms of the new parameters the time constants of the control network are

$$\tau_1 = \frac{2\zeta}{\omega_n} - \frac{1}{K} \tag{27}$$

and

$$\tau_2 = \frac{K}{\omega_n^2} \, . \tag{28}$$

With these definitions equation 24 becomes

$$p^2\varphi_2 + 2\zeta\omega_n p\varphi_2 + \omega_n^2\varphi_2 = \left(2\zeta\omega_n - \frac{\omega_n^2}{K}\right)p\varphi_1 + \omega_n^2\varphi_1 \, . \tag{29}$$

The transient response of the system to a sudden step of input phase $|\varphi_1|$ is found by integration of equation 29, and the initial condition for the oscillator frequency is obtained from equation 10. The transient response then is

$$\frac{\varphi_2}{|\varphi_1|}(t) = 1 - e^{-\zeta\omega_n t}\left[\cos\sqrt{1-\zeta^2}\,\omega_n t - \frac{\zeta - \frac{\omega_n}{K}}{\sqrt{1-\zeta^2}}\sin\sqrt{1-\zeta^2}\,\omega_n t\right] . \tag{30}$$

For $\zeta < 1$ the system is underdamped (oscillatory), for $\zeta = 1$ critically damped and for $\zeta > 1$ overdamped (no transient overshoots). In order to avoid sluggishness of the system, a rule of thumb may be followed making $.4 < \zeta < 1$ [2]. The transient response of equation 30 can be plotted in dimensionless form if certain specifications are made for the ratio ω_n/K. As the time constant τ_1 of the control network must be positive or can at most be equal to zero, the maximum value for ω_n/K is found from equation 27, yielding

$$\left.\frac{\omega_n}{K}\right|_{max} = 2\zeta \tag{31}$$

In this case the control network is reduced to a single time constant network ($\tau_1 = 0$). On the other hand, if for a fixed value of ω_n the gain of the system is increased towards infinity, the minimum value for ω_n/K becomes

$$\left.\frac{\omega_n}{K}\right|_{min} = 0 \, . \tag{32}$$

Fig. 4 shows the transient response of the system for these two limits and for a damping ratio of $\zeta = 0.5$.

The frequency response of the system is readily found from equation 24 and one obtains

$$\frac{\varphi_2}{\varphi_1}(j\omega) = \frac{1 + j2\zeta\frac{\omega}{\omega_n}\left(1 - \frac{\omega_n}{2\zeta K}\right)}{1 + j2\zeta\frac{\omega}{\omega_n} - \left(\frac{\omega}{\omega_n}\right)^2} \, . \tag{33}$$

Its magnitude is plotted in Fig. 5 for the two limit values of ω_n/K and for a damping ratio $\zeta = 0.5$. The curves show that the cut-off frequency of the system, for $\zeta = 0.5$, is approximately

$$\omega_c \cong \omega_n \quad \text{radians/sec.} \tag{34}$$

If φ_1 and φ_2 in equation 33 are assumed to be the input and output voltage of a four-terminal low-pass filter, the frequency response leads to the equivalent circuit of Fig. 6.

The noise bandwidth of the system is established by inserting equation 33 into equation 22 and one obtains

$$B = \omega_n\int_{-\infty}^{+\infty}\frac{1 + 4\zeta^2\left(\frac{\omega}{\omega_n}\right)^2\left[1 - \frac{\omega_n}{2\zeta K}\right]^2}{1 - (2 - 4\zeta^2)\left(\frac{\omega}{\omega_n}\right)^2 - \left(\frac{\omega}{\omega_n}\right)^4}\, d\left(\frac{\omega}{\omega_n}\right) . \tag{35}$$

The integration, which can be carried out by partial fractions with the help of tables, yields

$$B = \frac{4\zeta^2 - 4\zeta\frac{\omega_n}{K} + \left(\frac{\omega_n}{K}\right)^2 + 1}{2\zeta}\pi\omega_n \, . \tag{36}$$

For small values of ω_n/K, it is readily established that this expression has a minimum when $\zeta = 0.5$. Hence, the noise bandwidths for the limit values of ω_n/K and $\zeta = 0.5$ become

$$B\Big|_{\frac{\omega_n}{K}\to 1} = \pi\omega_n = \pi K \quad \text{radians/sec} \tag{37}$$

and

$$B\Big|_{\frac{\omega_n}{K}\to 0} = 2\pi\omega_n \quad \text{radians/sec.} \tag{38}$$

The above derivations, as well as the response curves of Figs. 4 and 5, show that the bandwidth and the gain constant of the system can be adjusted independently if a double time constant control network is employed.

C. Example

The theory is best illustrated by means of an example. Suppose an A.F.C. system is to be designed, having a steady state phase error of not more than 3 degrees and a noise bandwidth of 1000 c.p.s. The local oscillator drift shall be assumed to be 1500 c.p.s.

The required gain constant is obtained from equation 15, yielding

$$K = \frac{\Delta\omega}{\sin\varphi} = \frac{2\pi \cdot 1500}{\sin 3^\circ} = 180\,000 \text{ radians/sec.}$$

Since K is large in comparison to the required bandwidth, the resonance frequency of the system is established from equation 38

$$\omega_n = \frac{B}{2\pi} = \frac{2\pi \cdot 1000}{2\pi} = 1000 \text{ radians/sec.}$$

The two time constants of the control network, assuming a damping ratio of 0.5, are determined from equations 27 and 28 respectively

$$\tau_1 = \frac{2\zeta}{\omega_n} - \frac{1}{K} = \frac{1}{1000} - \frac{1}{180000} \cong 10^{-3} \text{ sec},$$

and

$$\tau_2 = \frac{K}{\omega_n^2} = \frac{180\,000}{1000^2} = .18 \text{ sec}.$$

These values of K, τ_1 and τ_2 completely define the A.F.C. system. A proper choice of gain distribution and control network impedance still has to be made to fit a particular design. For example, if the peak amplitude of the sinusoidal oscillator reference voltage is $E_2 = 6$ volts, the sensitivity of the oscillator control stage must be $S = 30\,000$ radians/sec/volt to provide the necessary gain constant of 180 000 radians/sec. Furthermore, if the capacitor C for the control network of Fig. 3a is assumed to be 0.22 uf, the resistors R_1 and R_2 become 4.7 kΩ and 820 kΩ respectively, to yield the desired time constants.

IV. Non-Linear Analysis

While it was permissible to assume small phase angles for the study of the synchronized system, thereby linearizing the differential equation 13, this simplification cannot be made for the evaluation of the pull-in performance of the system. The pull-in range of synchronization is defined as the range of difference frequencies, $p\varphi_1$, between the input signal and the free-running oscillator signal, over which the system can reach synchronism. Since the difference phase φ can vary over many radians during pull-in, it is necessary to integrate the non-linear equation to establish the limit of synchronization.

Assuming that the frequency of the input signal is constant as defined by equation 14, equation 13 can be written

$$p\varphi + F(p)\,K \sin\varphi = \Delta\omega. \tag{39}$$

Mathematically then, the pull-in range of synchronization is the maximum value of $\Delta\omega$ for which, irrespective of the initial condition of the system, the phase difference φ reaches a steady state value. To solve equation 39, the transfer function of the control network again must be defined.

A. $F(p) = 1$

The pull-in performance for this case has been treated in detail by Labin[3]. With $F(p) = 1$ equation 39 can be integrated by separation of the variables and it is readily found that the system synchronizes for all values of $|\Delta\omega| < K$. The condition for pull-in then is

$$|\Delta\omega|_{\text{Pull-in}} < K. \tag{40}$$

Large pull-in range and narrow noise bandwidth thus are incompatible requirements for this system.

B. $F(p) = \dfrac{1+\tau_1 p}{1+\tau_2 p}$

Inserting this transfer function into equation 39 and carrying out the differentiation yields

$$\frac{d^2\varphi}{dt^2} + \left[\frac{1}{\tau_2} + K\frac{\tau_1}{\tau_2}\cos\varphi\right]\frac{d\varphi}{dt} + \frac{K}{\tau_2}\sin\varphi = \frac{\Delta\omega}{\tau_2}. \tag{41}$$

This equation can be simplified by inserting the coefficients defined in equations 25 and 26, and by dividing the resulting equation by ω_n^2. This leads to the dimensionless equation

$$\frac{d^2\varphi}{\omega_n^2 dt^2} + \left[\frac{\omega_n}{K} + \left(2\zeta - \frac{\omega_n}{K}\right)\cos\varphi\right]\frac{d\varphi}{\omega_n dt} + \sin\varphi = \frac{\Delta\omega}{K}. \tag{42}$$

A further simplification is possible by defining a dimensionless difference frequency

$$y \equiv \frac{d\varphi}{\omega_n dt} \tag{43}$$

and one obtains a first order differential equation from which the dimensionless time $\omega_n t$ has been eliminated. It follows

$$\frac{dy}{d\varphi} = \frac{\frac{\Delta\omega}{K} - \sin\varphi}{y} - \frac{\omega_n}{K} - \left(2\zeta - \frac{\omega_n}{K}\right)\cos\varphi. \tag{44}$$

There is presently no analytical method available to solve this equation. However, the equation completely defines the slope of the solution curve $y(\varphi)$ at all points of a φ-y plane, except for the points of stable and unstable equilibrium, $y=0$; $\Delta\omega/K = \sin\varphi$. The limit of synchronization can thus be found graphically by starting the system with an infinitesimal velocity Δy at a point of unstable equilibrium, $y=0$; $\varphi = \pi - \sin^{-1}\Delta\omega/K$, and finding the value of $\Delta\omega/K$ for which the solution curve just reaches the next point of unstable equilibrium located at $y=0$; $\varphi = 3\pi - \sin^{-1}\Delta\omega/K$. The method is discussed by Stoker[4] and has been used by Tellier and Preston[5] to find the pull-in range for a single time constant A.F.C. system.

To establish the limit curve of synchronization for given values of ζ and ω_n/K, a number of solution curves have to be plotted with $\Delta\omega/K$ as parameter. The limit of pull-in range in terms of $\Delta\omega/K$ then can be interpolated to any desired degree of accuracy. The result, obtained in this manner, is shown in the dimensionless graph of Fig. 7, where $\Delta\omega/K$ is plotted as a function of ω_n/K for a damping ratio $\zeta = 0.5$. Since this curve represents the stability limit of synchronization for the system, the time required to reach synchronism is infinite when starting from any point on the limit curve. The same applies to any point on the $\Delta\omega/K$-axis, with exception of the point $\Delta\omega/K = 0$, since this axis describes a system having either infinite gain or zero bandwidth, and neither case has any real practical significance. The practical pull-in range of synchronization, therefore, lies inside the solid boundary. The individual points entered in Fig. 7 represent the measured pull-in curve of a particular system for which the damping ratio was maintained at $\zeta = 0.5$. For small values of ω_n/K this pull-in curve can be approximated by its circle of curvature which, as indicated by the dotted line, is tangent to the $\Delta\omega/K$-axis and whose center lies on the ω_n/K-axis. The pull-in range thus can be expressed analytically by the equation of the circle of curvature. If its radius is denoted by ϱ, the circle is given by

$$\left(\frac{\omega_n}{K} - \varrho\right)^2 + \left(\frac{\Delta\omega}{K}\right)^2 = \varrho^2 . \qquad (45)$$

Hence, for $\omega_n/K \to 0$, the pull-in range of synchronization is approximately

$$\left|\Delta\omega_{\text{Pull-in}}\right|_{\frac{\omega_n}{K}\to 0} \approx \sqrt{2\varrho\,\omega_n K} . \qquad (46)$$

ϱ can be interpreted as a constant of proportionality which depends on the particular design of the system, and which increases as the system gets closer to the theoretical limit of synchronization.

Equation 46 shows that the pull-in range for small values of ω_n/K is proportional to the square root of the product of the cut-off frequency ω_n and the gain constant K. Since the bandwidth of a double time constant A.F.C. system can be adjusted independently of the gain constant, the pull-in range of such a system can exceed the noise bandwidth by any desired amount.

V. Conclusions

The performance of an A.F.C. system can be described by three parameters. These are the gain constant K, the damping ratio ζ and the resonance or cut-off frequency ω_n. These parameters are specified by the requirements of a particular application and define the overall design of the system. It has been shown that among the systems with zero, single and double time constant control networks, only the latter fulfills the requirement for achieving good noise immunity, small steady state phase error and large pull-in range.

References

1 "Synchronizing Systems for Dot Interlaced Color TV," by T. S. George, Proc. I.R.E. February 1951.

2 "Principles of Servomechanisms," by G. S. Brown and D. P. Campbell, John Wiley & Sons, New York 1948.

3 "Théorie de la Synchronization par Contrôle de Phase," by Edouard Labin, Philips Research Reports, August 1949 (in French).

4 "Non-Linear Vibrations," by J. J. Stoker, Interscience, New York, 1950.

5 "The Lock-In Performance of an A.F.C. Circuit," by G. W. Preston and J. C. Tellier, Proc. I.R.E., February 1953.

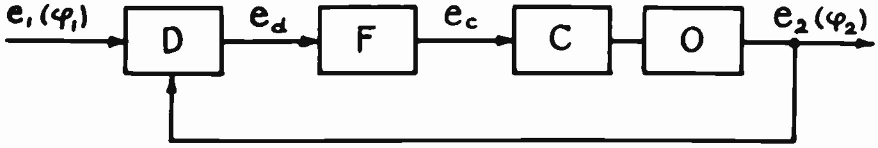


Fig. 1
Block diagram of AFC loop.

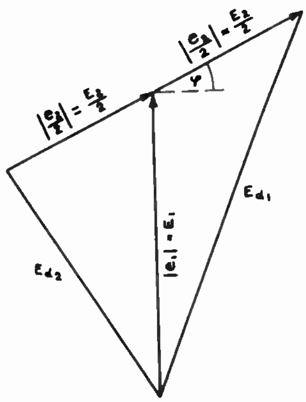


Fig. 2
Discriminator vector diagram.

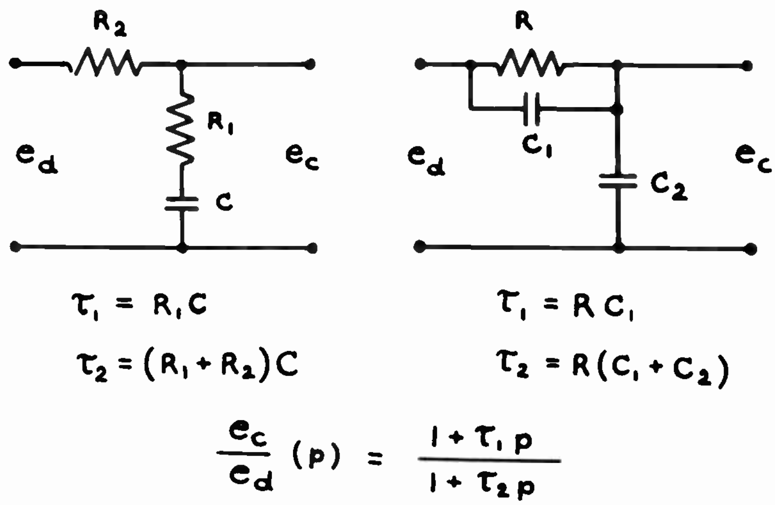


Fig. 3
Proportional plus integral control networks.

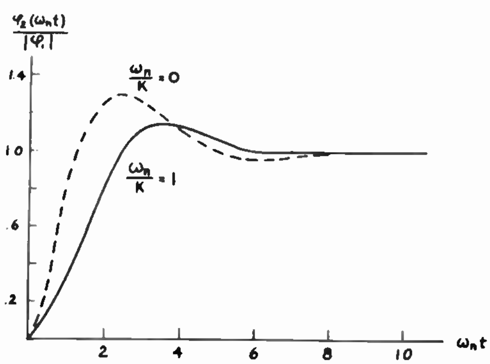


Fig. 4
Transient response for $\zeta = 0.5$.

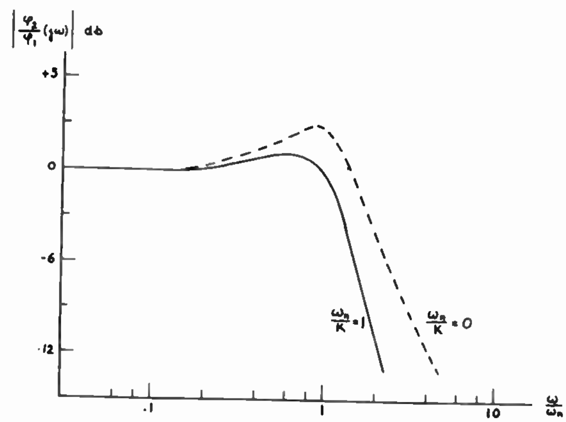


Fig. 5
Frequency response for $\zeta = 0.5$.

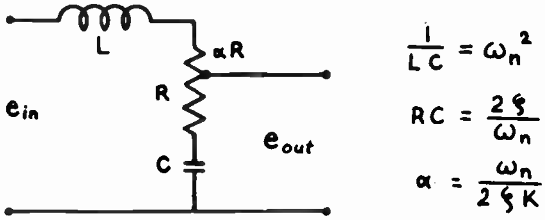


Fig. 6
Equivalent low-pass filter.

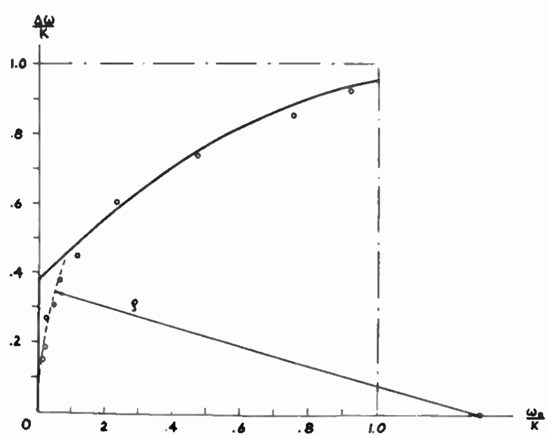


Fig. 7
Pull-in range of synchronization for $\zeta = 0.5$.

STANDARDIZATION OF PRINTED CIRCUIT
MATERIAL FOR MECHANIZED RADIO ASSEMBLY

W. Hannahs, J. Caffiaux, N. Stein
Sylvania Electric Products
Physics Laboratories
Bayside, L.I., N.Y.

Now that several brands of receivers incorporating some form of "printed" wiring are being released to the public for field evaluation, production and inspection departments are faced with the necessity of evaluating raw materials for which few standards have been established. Likewise the set designer, in replacing conventional point-to-point wiring by a more readily reproducible pattern of foil-like metal conductors on a plastic base, needs standards of good design practice and the means of evaluating "printed" constructions far in advance of Underwriters' inspection or public reaction.

The most common approach to improved assembling at this time is the etching of metal-clad plastic laminate and this important, but variable, material must be at the focus of standardization. However, in any consideration of standards and conditions of tests of printed circuits for radio use, the range of viewpoint must equally encompass the similar products of etching, stamping, plating or other processes. And it must likewise provide for judgment of flexible as well as rigidly supported foil conductors. The present viewpoint may reasonably be limited to connective harness joining all, or a major part, of radio or TV sets, as this promises to be of principle interest to the industry for the foreseeable future. Components employing "printed" parts and small silk-screened unit assemblies are generally amenable to evaluation under existing criteria.

In earlier papers by Danko[1] and by Swiggert[2] reference is made to tests of adhesion in clad-laminates and to tentatively utilized instantaneous measures of electrical properties. Now, however, engineers want a more perceptive look at these raw materials, one which will enable prediction of their endurance under field service conditions. Also needed is an understanding of the effects of dust, corrosion and other long term conditions upon variations of printed designs used in commercial radio and TV.

Samples of typical plated, etched, and stamped connective harness for radios are shown in Fig. 1. Both rigid and flexibly based types are represented.

Since shorting between conductors is obviously dependent on adhesion to the insulating member, delamination resistance is of primary concern. Details of two tests for this purpose - (1) a straight "pull" test and (2) a "peel test" were previously reported by the authors[3]. Recording of the stress required to peel a 1" wide conductor at room temperature has come into rather general use but agreement upon method is non-existent.

The essence of the peel test is well known in the adhesive and plating industries and several test procedures have been evolved[4-8, 14]. Variations in the stripping rates and angles of force application are dependent upon the specific material to be tested. Published stripping rates vary from one inch per minute with the force applied at 90°, to six inches per minute at 180°, to dead weights.

The rate of stripping for clad-laminates has been found uncritical up to six feet per minute, but gives results not correlative with static (dead weight) type tests. Values reported in this paper have been taken on the specially constructed extensometer shown in Fig. 2 at three feet per minute (exceptions noted) at which speed the dynamometer gave smoothest response.

A right angle pull has been found to give closer agreement between specimens having different foil thickness and probably also more closely simulates actually encountered conditions where the tag end of a conductor has begun to "lift". Tests performed under such conditions were in close correlation with similar tests performed on a low range commercial tensile testing machine.

As adhesives of varying thermal character may be used for foil bonding, cold adhesion strength becomes a questionable index to endurance at higher temperatures. In Fig. 2 it may be noted that provision has been made to heat the specimen under test by passing current from a welding transformer lengthwise through the foil strips. A thermocouple cemented to the foil provides a complete thermal history on the recorder at the right.

A first objective of this study has been to determine the thermal endurance of foil circuits under simulated conditions of manufacture. The short-term hot peel strength represents the resistance to delamination of conductors while heated by soldering and under the stress of handling and thermal deformation, without the presence of hardened solder to interfere with measurement. The graph of Fig. 3 is compiled from tests on many samples, all brands obtainable, and the several grades offered by some suppliers.

Fresh samples were used to obtain each temperature interval. All, however, were 2 oz. copper on 1/16" paper base phenolic and were raised in 10 seconds to the temperatures shown. The (initial) room temperature adhesion may be seen to range from about 2.5 lbs. to 19.5 lbs. per inch of width.

It has been found that four generalized types of response to heating may be recognized among specimens which are all within the same NEMA class, XXXP. Specimens having the highest cold adhesion declined in strength as the temperature was raised, in a gradual and easily predictable manner represented by Curve I. A second group of clad-phenolics, initially strong, were found to stand up well under mild heating but to suffer a sudden 50% loss of bond strength at a critical temperature level. This response has been labeled Type II.

Many of the poorer materials show improvement of strength upon heating, up to 50 - 70°C. Some then fail completely at 150° and others continue to show some adhesion even up near the limit of short-term thermal endurances of the best specimens. These are typified by Curves III and IV. A surmise of incomplete curing has been supported in some cases by examination of the bond layer. Notably, all materials have failed in the 10 seconds required to reach 240°C; faster heating, if obtainable, might enable differentiation at the upper level. This test, while yielding useful information about the behaviour of laminates under soldering and hot-punching conditions is too difficult for control use, because it is necessary to adjust experimentally the temperature rise program for each sample to a standard recorder diagram.

The method of selecting samples for adhesion testing was in accord with ASTM Specification D634-44 which essentially requires the rejection of material within three inches of any edge in order to minimize edge effects introduced in processing. Some data was taken to show the variation of strength in samples taken from the center versus samples taken from the edge of a sheet. This data may be found in the table of Fig. 4. Some sheets have also been found to have the best adhesion along the edges. Variations from sheet to sheet and from shipment to shipment of the same brand were in some cases quite pronounced.

Along with the averages for clad-XXXP which may be considered typical, there have been included a few values of more expensively based materials for comparison. The bond quality of clad-laminates has notably improved over the past three years so that minimum acceptance levels might well be raised several hundred percent over those mentioned in the references cited[1, 2]. This is warranted also by the further finding, reflected in the tabulated data, that those clad-phenolics having the best cold adhesion also best survived solder dipping.

The values in the second and third columns of Fig. 4 were taken on rapidly heated samples (as for the previous graph) and show a large, temporary, loss of bond strength during soldering. If a temperature critical for each type is not exceeded, a strong bond strength returns upon cooling.

Results of actual solder dipping tests are given in the last two columns and the temperatures shown are those of the solder on which the samples were floated. The solder temperatures were, on the average, 30°C above temperatures recorded by thermocouples placed within the samples at the metal-plastic interface. The 240°C level represents the lowest temperature considered practical for commercial solder dipping. These measurements were made on 1" strips in a bath of low melting point solder which permitted wiping the sample free of solder. Control checks with 50 - 50 solder baths and the use of narrower strips revealed no discernable differences.

In Fig. 5 are shown samples on which peel tests, cold and hot, have been made on three adjoining strips. Separation of three characteristic types have been encircled: (a) failure at the copper, (b) failure below the main interface, and (c) failure alternating between (a) and (b). The last mentioned type of separation was found in the strongest bonds. In Fig. 6 are blistered and exploded samples from elevated temperature tests.

The most severe test of the acceptability of materials and methods to be employed in the fabrication of "printed circuits" in radio and television receivers is their ability to withstand the rigors of service shop procedures. The operating temperatures of the most frequently used implements are tabulated in the upper left of Fig. 7. Measurements were taken by imbedding a fine-wire thermocouple at the tip of the "iron" and recording the stabilized temperature free standing in a 24°C ambient.

The duration of soldering tip contact to a typical rigid-backed copper foil conductor for several simple soldering operations was also determined, and the minimum range obtained for experienced service personnel is as follows:

	TIME	OPERATION
a)	1.5 to 4 sec.	Tinning a small cleaned area.
b)	2 to 4	Soldering a lead to tinned area.
c)	1.5 to 3	Unsoldering joint made in (b).
d)	2 to 4	Soldering lead in punched hole
e)	7 to 10	Unsoldering (d).
f)	4 to 8	Resoldering (d).

The temperature level reached by etched

conductors while being soldered may be judged by reference to the graph in the upper right of Fig. 7. The main curve in this figure depicts the impairment of bond under various typical hand soldering conditions, compiled from continuous recordings of samples peeled on a Tinnius-Olsen tensile machine. The plateaus of the curve reveal the normal resistance of the conductor to stripping and the valleys are the measured adhesion at spots along the conductors where typical soldering operations have been performed, the solder, of course, being wiped off before testing.

The second and third valleys are typical of carefully controlled assembly of phenolic based circuits with low wattage irons and indicate retention of a useful proportion of the original force retaining the conductor. The longer periods required for unsoldering in servicing, even though performed in minimum time with small irons is seen by the evidence of the first and fourth valleys to produce a severe deterioration. In most tests of the latter condition obvious blistering of the base or release of the conductor was observed.

The application of the stripping test to flexibly backed foil is illustrated in Fig. 8 with values from several specimens. The manner illustrated for holding the sample has been chosen purely as a matter of convenience and no direct logical agreement with the test of rigidly based foils is expected. However, measurements of rigid and flexibly-backed samples fabricated with identical adhesive give values of the same order. The last two values in the table raise the question of the condition of samples prior to analysis. Investigation at Stanford for WADC[9] has yielded evidence that etching generally effects bond strength adversely; moisture conditioning before testing adhesive strength is also a general practice in the ASTM procedures cited (4-8). Since peel tests may be performed on finished circuits as well as the raw material, these effects need to be taken into account.

In standardizing stripping tests, performance of the test on conductors less than 1" in width has been found to correlate within 1 or 2% except in the case of stamped circuits (see Fig.9). In narrower widths stamped conductors give increasing values of peel resistance per unit width, because of the turned and mechanically clinched edges produced in this process.

Other tests of the strength of clad-laminates such as shear (per ASTM D-816-46) and straight tension or "pull" (per ASTM D-429-47T) have been found of limited use. Both present the difficult problem of attaching a test member to the cladding (by a cold setting adhesive or a low melting solder) in a manner not deleterious to the bond under test, yet superior in strength to it. With any but the poorest 1/16" clad stocks simultaneous delamination of an area by a test plug extends deeply into or through the base laminate (Fig. 10). Such fractures often yield information as to the nature of the bond but the values resulting are widely dispersed and much dependent on the choice of plug diameter as may be seen in Fig. 11.

No completely satisfactory test for dielectric and loss factors has yet been proposed. The use of ASTM D150-47, wherein the average dielectric is measured between the two metal clad surfaces, is of limited control value because the adhesive layers, which are in closest proximity and have most influence upon fields generated by the conductor, are often a distinctly different dielectric from the main body of the non-conductor. Determining the "Q" of a specified coil etched on clad materials may be useful in control of a specific production situation but the standardizing of etching, washing and drying, so as to make the test of wide general use, does not look promising.

Determination of surface resistance by ASTM D257-49 is as feasible on clad as on non-clad laminates because the simple pattern of conductors required for this may be scribed and stripped, thus avoiding the variable effects of etching.

Surface resistance between closely printed lines is markedly lowered by accumulation of dust which absorbs moisture. In Fig. 12, pairs of etched conductors exposed for only 9 days in a dusty but dry location, show a ten-fold decrease of insulation resistance. The unprotected conductors show an appreciably greater sensitivity to dust plus humidity than do the pair protected by a single coat of a JAN approved fungus varnish, but both are seriously deteriorated. This is emphasized when the humidity is raised to a very common R.H. of 80, resulting in a drop of resistance of the unprotected lines from 2,500 to 25 megohms. This is illustrated at the right of Fig. 12. Adequate protection of printed circuits in radio design is strongly indicated.

Because of the lack of data concerning the useful life of etched foil circuits some hypothesis is warranted. One potential cause of failure is the deterioration of narrow conductors due to electrolytic corrosion effects. Such effects are apt to be noticeable when the piece is exposed to intermittent operation in moist environments, in which case electrolytic currents may be set up with the insulating material acting as an electrolyte. A semi-quantitative test to determine the extent of electrolytic corrosion caused by the combined effects of adhesive and backing material on fine wires has been reported by the Minnesota Mining and Mfg. Company[11] and with slight modification may be applied to the determination of these effects in foil-clad laminates.

Deterioration of insulation and conductors are quite generally accelerated by heat and circuits in the printed radios so far tested operate at surface temperatures of 75 to 102°C. Reference to existing standards[12, 13], concerned with the safe current capacities of conventional conductors reveals several possible ways of rating printed conductors. Generally, allowable current densities are derived from the permissible operating temperatures of the insulations in contact with the conductors. For rubber and thermoplastic hook-up wire (Type R & T insulation) this limit is 60°C. Varnished cambric (Type V) is rated at 85°C and asbestos insulations start at 90°C. Common insulating bases for printed circuits are most similar to Type V insulation and, as such, may be classed by Underwriters for operation at 85°. Ratings for phenolic laminates range from 93° to 177°C and copper conductors are generally not recommended above 150°C. Associated components and wires of lower rating (55° and 65°C) may be the dominating factor. This consideration for associated components is highly pertinent for "printed circuit" assemblies because of the usual practice of placing RC components in intimate contact with the connective circuit.

The permissible operating temperature of printed circuit conductors which suggested itself, with consideration of all of the above, is 85°C (surface). The temperature rise of typical etched conductors in a 60° ambient are shown in Fig. 13.

Limiting current densities derived from this data and an arbitrary 85° limit are quite ample, ranging from 85,000 amps. per square inch for isolated 1/32" conductors of 1 oz. copper, down to 24,000 amps. per square inch in groups of 1/8" x 2 oz. conductors. The lowest is ten to twenty times the density found in contemporary printed radio designs. The temperature in most printed decks, then, is the result of tube dissipation and the generally high ambient within cabinets. Better ventilation is indicated.

In addition to the suggestions made earlier in connection with specific tests and procedures, the results of the measurements here reported may well be embodied in certain minimal recommendations of good practice in the design of printed circuit receivers. In deference to the low hot strength of conductor bonds, components should be placed so as not to stress the bond. In Fig. 14a is shown how the component may be properly backed by lead binding, and 14b suggests that no soldering be done along a printed conductor except at points mechanically anchored.

It is believed that dust and moisture reduction of surface resistance call for adequate spacing of printed conductors (15c); probably a minimum of 1/32" on rigid bases and 1/16" on flexible. The life and voltage ratings of printed decks may both be extended by lowering operating temperatures, and a very simple method is illustrated in 15b.

Much remains to be done on the standardization of test procedures for clad-laminates and toward the establishment of performance minimums. The results of many aging tests will be needed to set accurately maximum continuous temperatures for printed decks, and to find better means of insulating and protecting them. It is hoped that by bringing some of these matters to attention that the development of an adequate body of information, concerning the application of printed circuits to receivers, will be stimulated.

References

1) Danko, S.F. and Lanzalotti, S.J. - "Auto-Sembly of Miniature Military Equipment"; Electronics, July 1951.

2) Swiggett, R.L. - "Printed Circuits on Foil-Clad Plastics"; Modern Plastics, August 1951.

3) Hannahs, W. and Caffiaux, J. - "Etching -- A General Fabrication Technique"; Radio and Television News, Electronic Engineering Section, March 1952.

4) "Adhesion of Vulcanized Rubber (D413-39)"; ASTM Standards 1949 (Book), Part 6, page 896.

5) "Peel or Stripping Strength of Adhesives (D903-49)"; ASTM Standards 1949 (Book), Part 4, page 798.

6) "Methods of Testing Rubber Adhesives (D816-46)"; ASTM Standards 1949 (Book), Part 6, page 1111.

7) "Methods of Test for Adhesion of Vulcanized Rubber to Metal (D429-47T)"; ibid, page 902.

8) "Method of Test for Bond Strength of Plastics and Electrical Insulating Materials (D952-48T)"; ibid, page 543.

9) "Development and Application of Automatic Assembly Techniques, Phase II: Interim Report #2, August 1952". Stanford Research Institute, Stanford, California, page 24.

10) "Electrical Resistance of Insulating Materials (D257-49T)"; ASTM Standards 1949 (Book), Part 6, page 396.

11) "Quantitative Measurement of Electrolytic Corrosion"; Electrical Insulation Section, Tape Research Laboratories, Minnesota Mining and Mfg. Co., St. Paul, Minnesota.

12) "Standard for Power Operated Radio Receiving Appliances"; Underwriters Laboratories, October 1952, page 21.

13) "Type R, T & V Insulation"; National Electrical Code, 1951.

14) A.L. Ferguson - "The Adhesion of Electrodeposits"; American Electroplaters Soc., Research Report Serial #1 (1945-46) Serial #2 (1946), Serial #8 (1948).

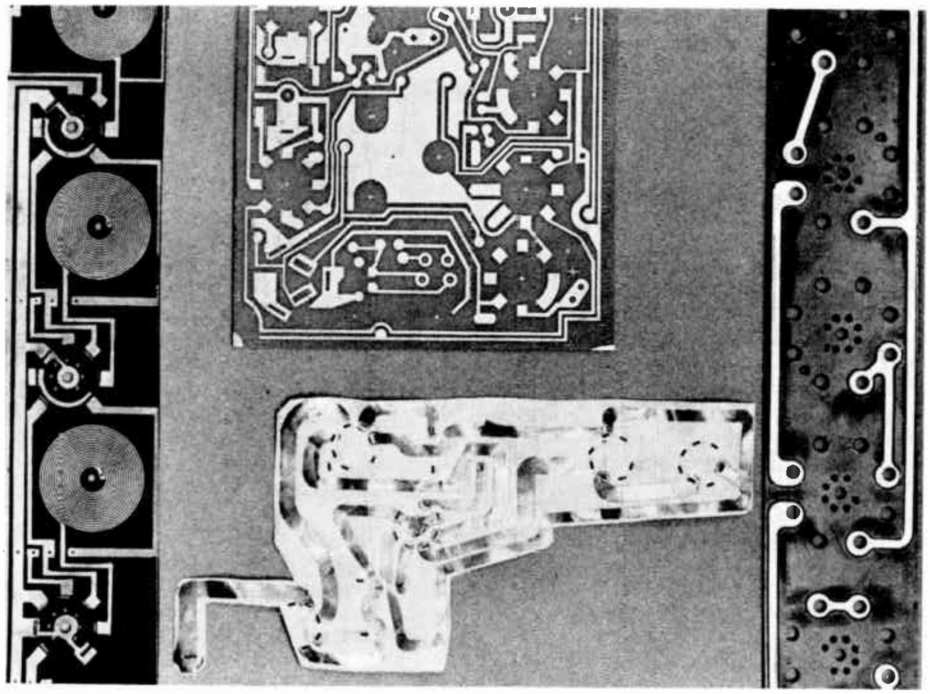


Fig. 1

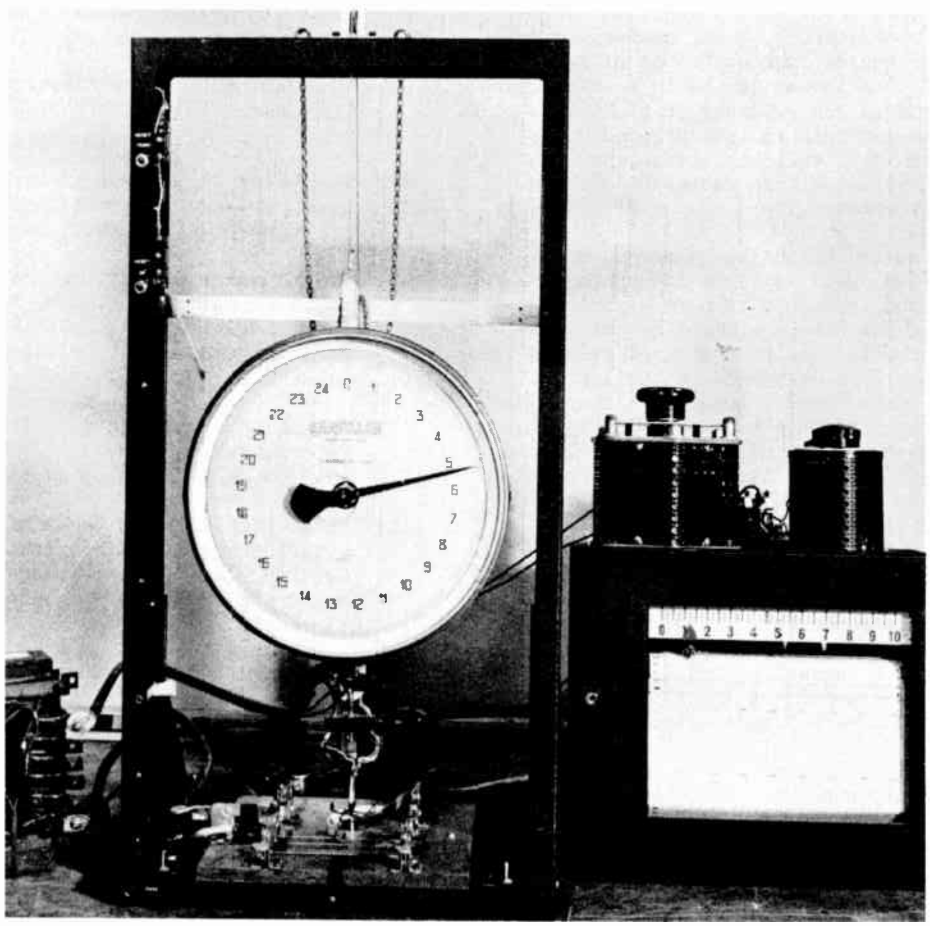


Fig. 2

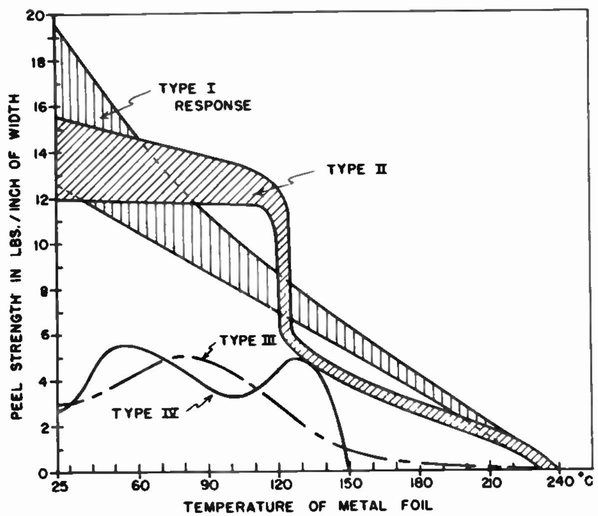


Fig. 3 - Hot peel strength of foil-clad laminates.

BASE LAMINATE (2 OZ. COPPER-CLAD EXCEPT WHERE NOTED)	ULTIMATE COLD AT 25°C	BOND STRENGTH *** IN LBS/INCH WIDTH — DURING SOLDERING: SHORT TERM HOT STRENGTH 210°C **	230°C **	AFTER SOLDER DIPPING 4 SEC. AT 240°C*	2 SEC. AT 275°C*
XXXP LAMINATE-AVERAGES					
BRAND "A" - BEST	16.2	3.0	1.4	16.2	15.4
BRAND "B" - MEDIAN	7.6	2.0	BLISTERED	8.4	EXPLODED
BRAND "C" - POOREST	2.4	0.2	BLISTERED	2.4 SL. BLISTERED	EXPLODED
EDGE OF A SHEET	1.6				
CENTER SAMPLE	3.3				
POLYTETRAFLUOROETHYLENE	6.8	13.6	9.2		
MEEAMINE FF-55	12.6	5.0 (125°)	BLISTERED (200°)		
SILICONE-GLASS G-7	1.1	6.2 (125°)	BLISTERED (200°)		
NYLON YN-25	15.5				
COTTON, FINE LE-41	8.7				
CANVAS CE	2.5				
ETHOXYLENE-GLASS	5.1				
STAMPED TO XXXP - AVERAGE	8.8 (1/4" WIDTH)			4.5 (1/4" WIDTH)	
ALUMINUM-CLAD XXXP-AVERAGE	8.3	2.0 (150°)		6.6	EXPLODED

* TEMPERATURE OF SOLDER IN POT.
** BOND INTERFACE TEMPERATURE (MOLTEN SOLDER MAY RUN 30° HOTTER).
*** ALL TESTS MADE ON 1" WIDTHS EXCEPT WHERE NOTED.

Fig. 4 - Peel strengths - significant values.

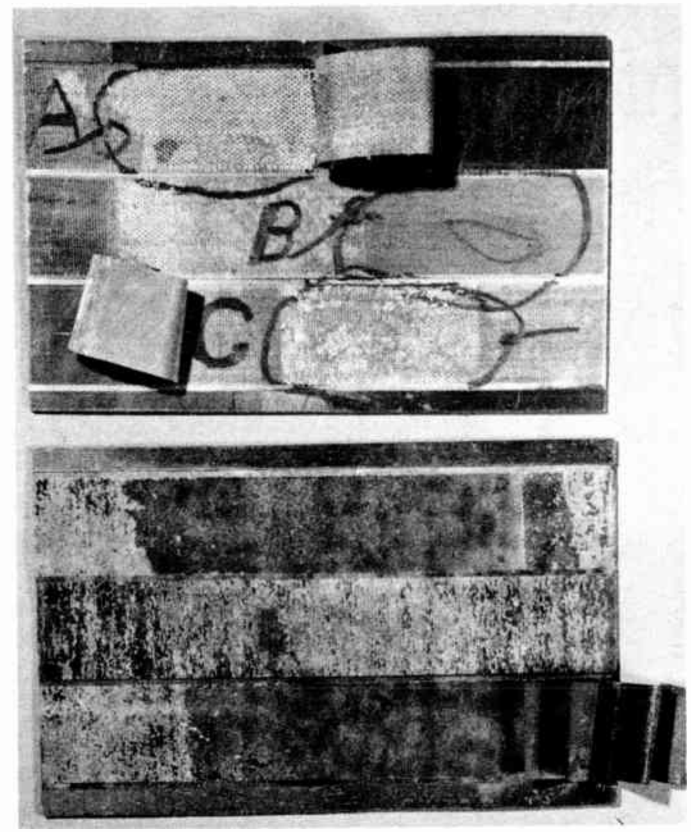


Fig. 5

Fig. 6

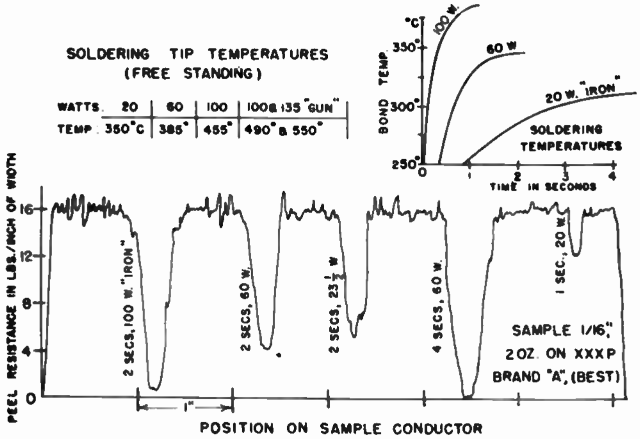


Fig. 7
Hand (iron) soldering:
impairment of peel strength.

TYPICAL TEST RESULTS

BACKING	LBS./INCH WIDTH
.003" FISH PAPER	8.7
.010" COTTON, 80 x 80	9.9
.004" NYLON, 109 MESH	1.85
.004" GLASS CLOTH ECC-60x56 (SINGLE COPPER FACE)	7.7
.004" GLASS CLOTH ECC-60x56 (DOUBLE *COPPER FACE)	5.9

* NOTE: ONE FACE REMOVED BY FERRIC CHLORIDE ETCHING

TEST METHOD

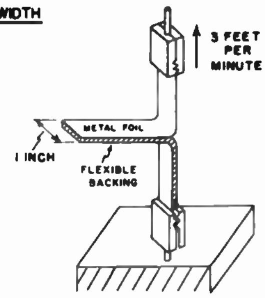


Fig. 8
Cold peel strength of flexibly
backed copper foils - 2 oz.

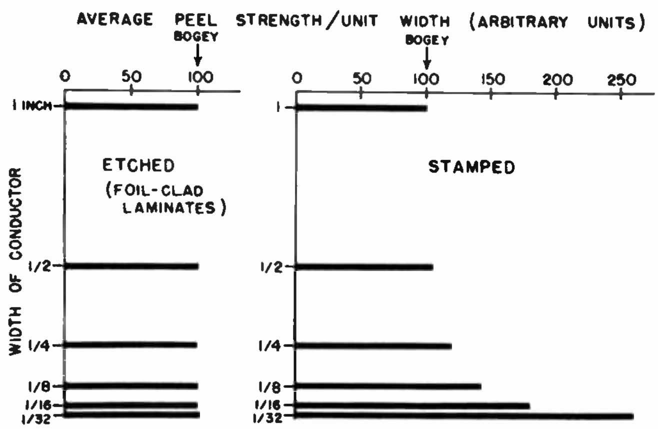


Fig. 9
Variation of ultimate bond strength with conductor width.

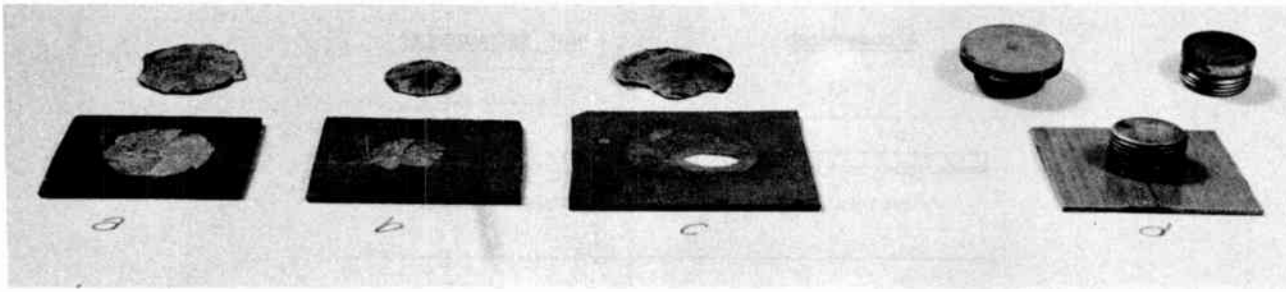


Fig. 10

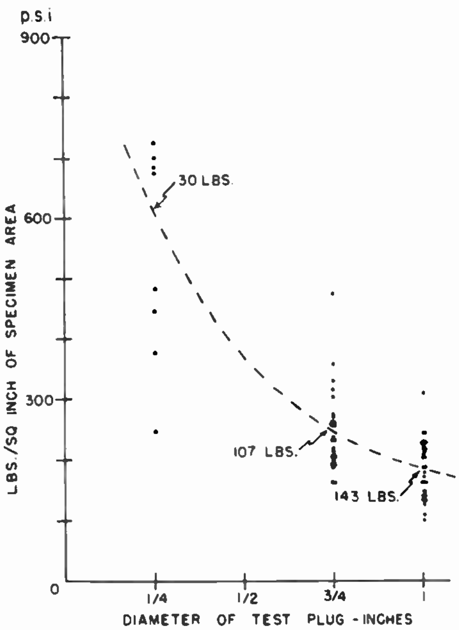


Fig. 11
Delamination (pull)
strength of clad-laminates.

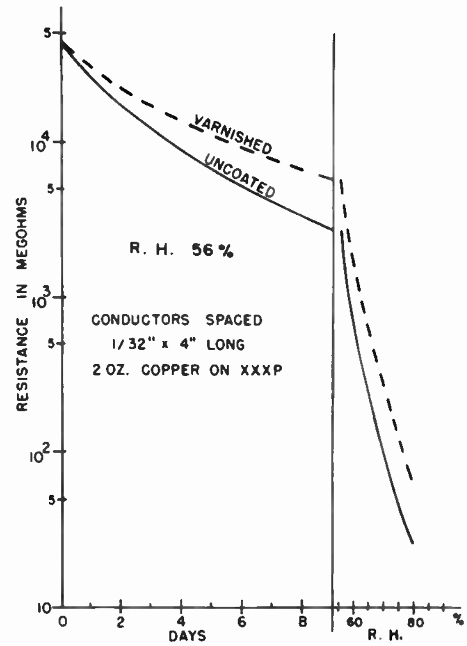


Fig. 12
Dust and humidity effect
on an etched circuit.

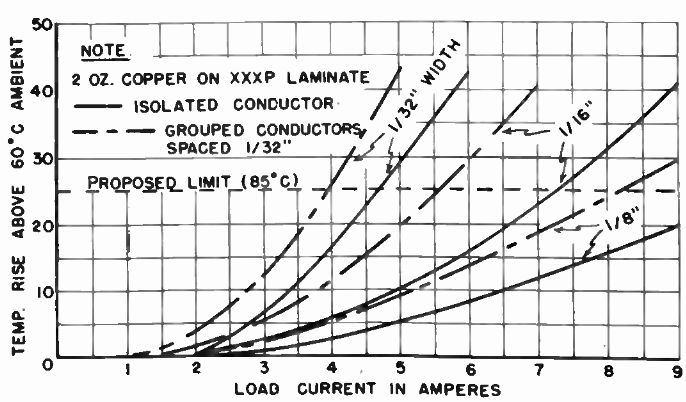


Fig. 13

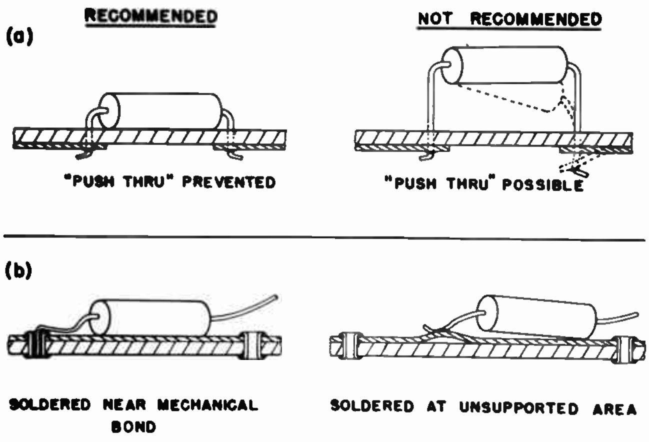


Fig. 14 - Design practices.

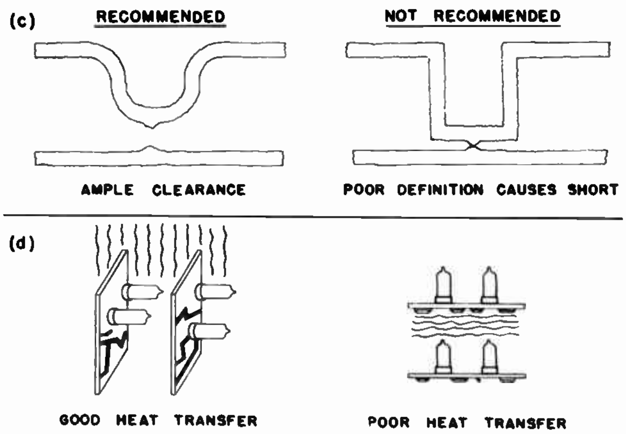


Fig. 15 - Design practices.

A COLOR TELEVISION RECEIVER FOR THE NTSC SYSTEM

by Kenneth E. Farr
Westinghouse Electric Corporation
Metuchen, N. J.

Introduction

This paper will describe a compatible color television receiver. Its basic circuits will be described and illustrated by block diagrams. Photo-Photographs of the receiver are also included. The receiver to be described may be considered a "medium-grade" type. It is capable of producing good commercial quality color pictures. Either more elaborate or less expensive designs can be produced within the framework of the NTSC system. This receiver is intended to be an approach toward a commercial production design. It uses an RCA tri-color kinescope tube producing a picture equivalent in size to a 12½" black-and-white picture. The receiver will operate on all VHF and UHF channels, and will receive standard monochrome transmissions as well as NTSC color signals.

In describing this receiver, the circuits will be broken down into four broad sub-divisions and each described in turn. These sub-divisions include the monochrome or brightness signal channel, the color decoder, the color sync circuits, and the deflection, convergence and power supply circuits. In conjunction with the color decoder section, the salient features of the NTSC system will be discussed.

Monochrome Video

Turning first to the monochrome channel, much of this portion of the receiver is patterned closely after present Westinghouse black-and-white receiver production. The tuner, IF strip, sync separation, AGC, and deflection oscillators are in this category and will not be described in this paper. The IF pass-band requirements, however, will be discussed in conjunction with the color circuitry.

The video section of the receiver differs considerably from black-and-white practice. It is shown in figure 1. A 6AS6 is used as the first video stage, with DC suppressor bias variation for master contrast control. Good linearity and uniform pass-band with changes in contrast control setting are required in this stage, since the color information passes through this stage in addition to the monochrome or brightness signal. The delay line here serves no function for monochrome reproduction. Its purpose will be described in conjunction with the color circuits. The terminals marked color difference inputs receive no signal for monochrome reproduction. Thus the red, green, and blue video amplifiers receive identical signals, which are applied to the three guns of the tri-color tube to produce a monochrome picture.

The resistors from the output of the first red and blue video stages to the input of the first green video are for color reproduction only, and will be explained later in that connection. As far as a monochrome signal is concerned, they operate merely to reduce the gain of the green video stage in a manner similar to negative feedback, since the voltages at the outputs of the first red, green and blue amplifiers should be identical. The gain reduction suffered by the green video amplifier as a result of this connection is made up elsewhere in the green video section.

To obtain a good monochrome picture as well as a good color picture, it is important that the video drive to the tri-color kinescope be adjusted with reference to the gains of the respective guns to provide the same color balance for all gradations in the picture. The red, green and blue video gains in this receiver are set at the ratio of 1 to .7 to .6 respectively, which is an average value for the RCA tri-color tube. Fine adjustments of color balance are made by varying the screen voltage and grid bias of each gun of the picture tube individually as required.

This video amplifier is AC coupled, and DC restorers are used as shown on the red, green and blue outputs. It is recognized that the noise performance of simple DC restorers as used here is poor. Direct coupling of the entire video system does not seem practical, however. Possibly some form of keyed clamp will be a better solution and will be tried as soon as time permits.

Triode video amplifiers are used to provide greater plate swing for a given power supply drain than could be obtained with pentodes. The two halves of the 6J6 are used in parallel for the red and green video output amplifiers but only one half is used for the blue. This is because the blue bandwidth need not be as wide as the red and green. The operation of this video amplifier when color pictures are being received will be described later.

Color Video Circuits

The NTSC Signal

Turning now to the color portion of the receiver, before describing the receiver circuits a brief resume of the technical features of the NTSC color television system may be in order. Red, green and blue video signals derived from a color camera are mixed to produce a brightness or monochrome signal. These red, green and blue signals are also mixed in different proportions to provide the chrominance information which is modulated onto a sub-carrier at the upper end of the video spectrum. The modulated sub-carrier is added to the brightness video signal and transmitted over the normal television channel. The frequency of the color sub-carrier is 3.579545 megacycles, which is an odd multiple (the 455th) of one-half the line scanning frequency. Selection of a sub-carrier in this manner effectively interlaces the chrominance components in frequency between those of the brightness video signal and renders them practically invisible on most black-and-white receivers.

In Figure 2 is shown a specification of the NTSC signal. The equation at the top is the complete expression of the color video signal for all video frequencies. It is seen to be made up of a brightness signal E_Y', and in the brackets a color sub-carrier consisting of two parts. The chrominance information which is modulated onto the sub-carrier is formed in two packages, E_Q' and E_I'. E_Q' lies along an axis of the CIE color diagram passing from purple through white to green while the E_I' component lies along an axis passing from orange through white to cyan. These two signals are modulated onto two sub-carriers of the same frequency but 90° different in phase. To conserve bandwidth and improve compatibility of the signal, this sub-carrier transmission is of the vestigial-side-band suppressed-carrier type.

The vector diagram depicts the two components of the modified sub-carrier, E_Q' and E_I', at 3.58 Mc. in their correct phase relative to the burst. This is the color synchronizing signal, transmitted on the back porch of each horizontal blanking period. The make-up of the brightness signal, E_Y', as well as the chrominance signals E_Q' and E_I', is shown alongside the vector diagram. The primes refer to gamma correction of the signals in all cases. The compositions of E_Q' and E_I' are such that they can be combined in the proportions shown in the lower right corner of the figure to produce the color difference signals $(E_R'-E_Y')$ and $(E_B'-E_Y')$. Since the bandwidth of E_Q' is limited to 500 Kc. these relations hold true only up to 500 Kc. The equation second from the top, and the dotted vectors, are alternative ways of specifying the NTSC signal for color video frequencies below 500 Kc. There is a real significance to these alternative expressions as they indicate two possible types of receiver circuitry.

It should be realized that a single sub-carrier vector is produced which is the resultant sum of these two vectors. This resultant vector can have any phase from zero to 360 degrees around the diagram and any amplitude from zero (which is a monochrome picture), to unit value. While it is not immediately obvious from the way that this vector is synthesized, it can be shown that the phase of the transmitted sub-carrier is a function only of the hue of the picture,while its amplitude is a function only of the saturation. Synchronous demodulation is used at the receiver which resolves the components of the transmitted vector along any desired axes. Thus, either E_Q' and E_I' or $(E_R'-E_Y')$ and $(E_B'-E_Y')$ could be recovered. The medium-grade receiver under discussion here utilizes only about 600 Kc of color bandwidth, and recovers $(E_R'-E_Y')$ and $(E_B'-E_Y')$ directly.

A simplified diagram of a color transmitter encoder is shown in Figure 3 to illustrate the generation of the NTSC color signal. If a tri-color picture tube is substituted for the camera, and the signal paths are traced from right to left, this figure would then be a simplified illustration of a color receiver utilizing the entire transmitted chrominance signal.

Pass-band Considerations

The transmitted frequency spectrum of the sigmal, and to the same scale, the idealized RF-IF response, are shown in Figure 4. It should be noted that the E_Q' component is transmitted full double side-band, while the wider band E_I' is transmitted with vestigial side-band. The receiver being described here utilizes only the double-side-band portion of the signal. A somewhat more complex receiver can be designed incorporating matrixing to take advantage of the greater chrominance bandwidth available in the E_I' signal, along the orange-cyan axis. This "wide-band orange-cyan" type receiver could be considered a deluxe ceiling performance receiver.

The receiver response at the bottom of the figure is represented here as being 6 db down at the picture carrier in the conventiel manner, and 6 db down at approximately 4.1 Mc. at the upper end of the pass-band. It is desired, however, to have the effective response to the color demodulator inputs flat over a somewhat wider range to utilize as much as possible of the double-side-band portion of the sub-carrier.

Figure 5 illustrates how this is accomplished. The sub-carrier portion of the transmitted spectrum is reproduced at the top for reference. In the center is shown, in dotted lines, the IF pass-band from Figure 4. Also shown here, in solid line, is the chroma amplifier pass-band. This is the pass-band for the sub-carrier after the first demodulation. It covers the region of 2.5 to 4.5 megacycles. When the receiver is properly tuned, the IF response will be complementary to this chroma amplifier response. The effective overall pass-band to the color demodulator inputs will then be as shown at the bottom of the figure, being flat over about plus and minus 600 Kc from the color sub-carrier.

This idealized overall pass-band will result only when the receiver is properly tuned. The IF pass-band shifts relative to the signal, as the receiver is tuned, while the chroma amplifier pass-band does not. For this reason, as well as the practical difficulties encountered in matching the slopes of an IF and a video response, a 6 db boost in the chroma amplifier pass-band was felt to represent the maximum practicable amount of compensation. A further limitation on the amount of upper side-band which the receiver can use is the requirement that the chrominance pass-band be well down (about 30 db) at 4.5 megacycles so that the sound carrier will not produce a 920 Kc beat in the chrominance output.

Chrominance Channel

The chrominance channel, with a portion of the monochrome video channel duplicated, is illustrated in Figure 6. The chroma amplifier receives the color sub-carrier signal from the output of the first video stage. The saturation control, which is a panel control for the viewer's use, varies the gain of this stage. It can be set to zero for monochrome signals. The synchronous demodulators are 6AS6 tubes with suppressor injection. The outputs of these demodulators are ($E_R'-E_Y'$) and ($E_B'-E_Y'$). These outputs are combined with E_Y' in the resistive adders to form E_R' and E_B'. Since the band-pass to which the color difference signals have been subjected is much narrower than that through which the E_Y' signal has passed, the color difference signals suffer a delay relative to the E_Y' signal. The delay line shown in the E_Y' signal path introduces a delay of approximately .7 microseconds to the E_Y' signal to make the signals coincident for mixing.

When color difference signals are supplied to the resistive adders, the outputs at the red and blue video plates are no longer identical and equal to E_Y' as was the case for a monochrome signal. They are now E_R' and E_B', respectively. Connecting these plates to the green video input now functions to produce the green signal E_G', by subtracting the appropriate amounts of E_R' and E_B' from E_Y'. (It will be recalled that E_Y' is made up of a linear addition of E_G', E_R' and E_B'.) The signals E_R', E_B' and E_G' are applied to the three guns of the tri-color tube to control the beams for their respective colors, thus producing a simultaneous color picture.

Color Sync

The local sub-carrier regenerator, which furnishes injection to the demodulators, is illustrated in Figure 7. The output of the chroma amplifier contains, in addition to the color video information, the color synchronizing burst. In order to separate this burst from the color video information, a gated amplifier is used. The output of the chroma amplifier is applied to the burst amplifier, along with a gate pulse from a multivibrator triggered by the trailing edge of the horizontal retrace. The output of the burst amplifier then consists only of the color synchronizing burst, which is applied to a balanced phase detector. The output of the self-excited local oscillator is also applied to the phase detector. The DC error signal from the phase detector is integrated and used to control the frequency and phase of the local oscillator.

Auxiliary Circuits

The high voltage supply for the tri-color kinescope is a regulated fly-back supply delivering 20,000 volts at 700 microamps. This has been described in a recent RCA License Lab. Bulletin.[1] A 6V6 is used for the vertical output. Both horizontal and vertical oscillators are multivibrators. Dynamic convergence circuitry used for the tri-color tube is essentially the same as developed by RCA[2]. Two low voltage power supplies are used, one delivering 300 volts at 225 ma. and one delivering 450 volts at 225 ma.

Physical Contruction

A top view of the receiver chassis is illustrated in Figure 8. This chassis contains everything except the picture tube and speaker. The IF strip and tuner can be seen along one side. The high voltage supply is in a separate enclosure. Figure 9 shows the underside of the chassis.

A rear view of the receiver in the cabinet is shown in Figure 10. The picture tube is mounted in a sub-assembly which slides into the cabinet on rails fastened to the cabinet top. A front view of the receiver is shown in Figure 11. Controls directly exposed include on-off volume concentric with master contrast. Saturation and master brightness are thumb wheels on the side of the channel selector dial. Fine tuning is in the center of the channel selector. Service controls are accessible behind center cross-bar, which is actually a trap door. Additional service controls are on the rear of the chassis.

Summary

To summarize, a receiver has been described which may be considered to be an approach to a commercial product. It is a medium grade receiver. Both cheaper, as well as higher performance more costly type receivers are possible with the present NTSC system. Whether this medium-grade type of receiver will ultimately be the most practical, time will tell.

References

1. Deflection Systems with Regulated Kickback High-voltage Supplies for Tri-color Kinescopes - LB-877.

2. A. W. Friend, Deflection and Convergence in Color Kinescopes, Proc. IRE, October, 1951.

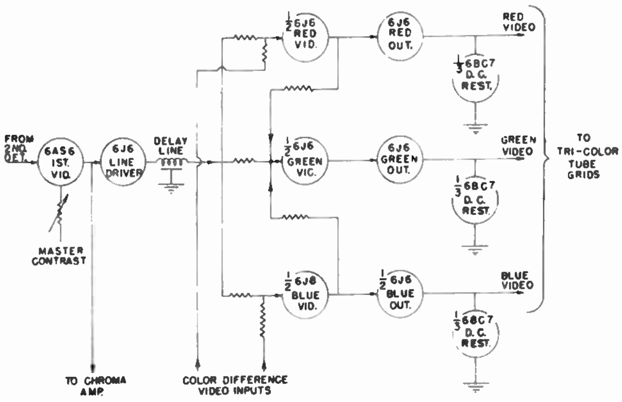


Fig. 1 - Video section.

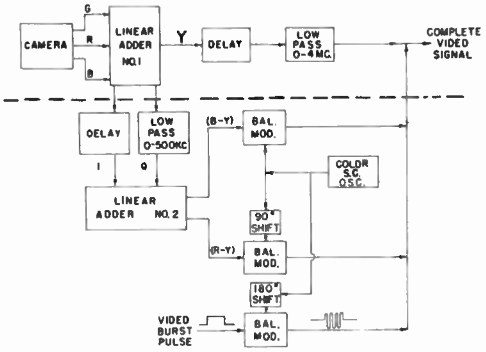


Fig. 3 - Color encoder (not used).

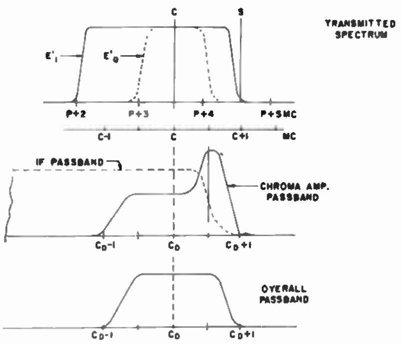


Fig. 5 - Chrominance pass-band.

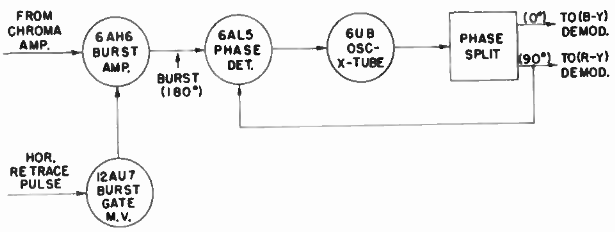
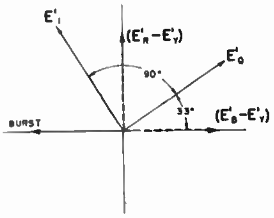


Fig. 7 - Subcarrier regeneration.

$$E'_M = E'_Y + \left[E'_Q \sin(\omega t + 33°) + E'_I \cos(\omega t + 33°)\right] \quad \text{(ALL FREQUENCIES)}$$

$$= E'_Y + \left[.493(E'_B - E'_Y)\sin \omega t + .877(E'_R - E'_Y)\cos \omega t\right] \quad \text{(BELOW 500K)}$$



$$E'_Y = .59E'_G + .30E'_R + .11E'_B$$

$$E'_Q = -.52E'_G + .22E'_R + .31E'_B$$

$$E'_I = -.28E'_G + .60E'_R - .32E'_B$$

$$(E'_R - E'_Y) = .63E'_Q + .95E'_I$$

$$(E'_B - E'_Y) = 1.71E'_Q + 1.11E'_I$$

Fig. 2 - The NTSC signal.

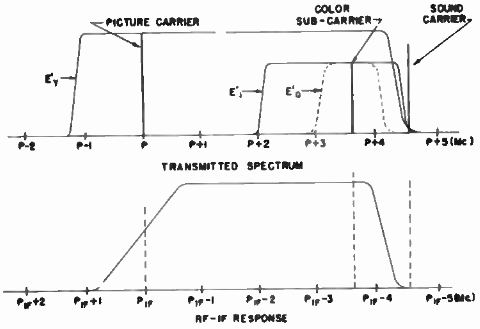


Fig. 4 - Pass-band requirements.

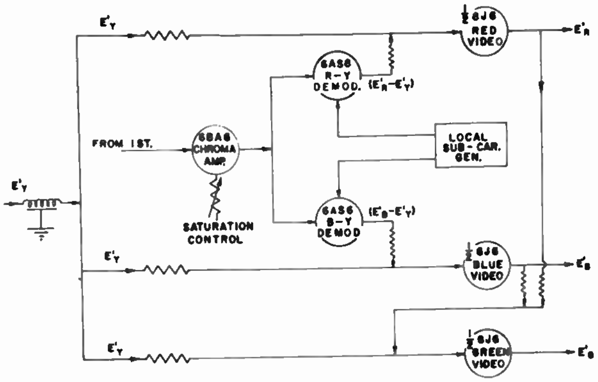


Fig. 6 - Chrominance channel.

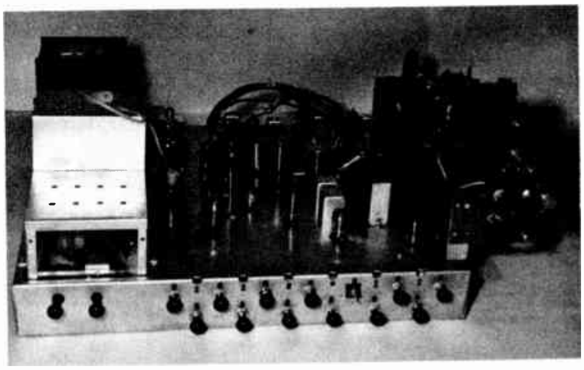


Fig. 8 - Chassis - top view.

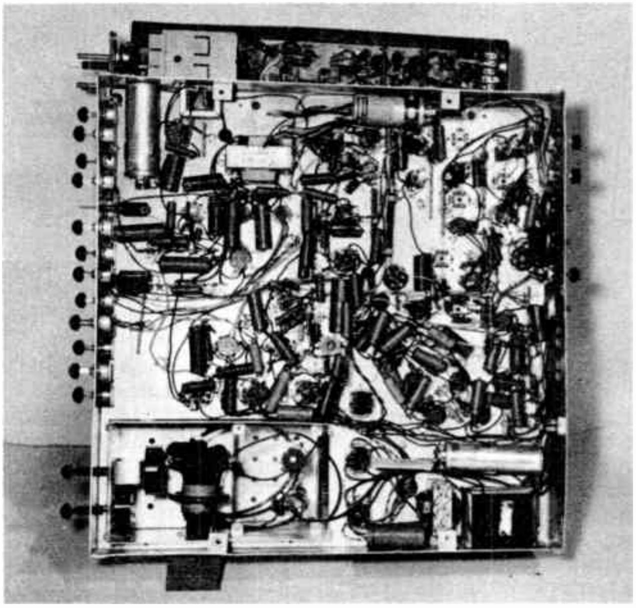


Fig. 9 - Chassis - bottom view.

Fig. 10 - Receiver in cabinet - rear view.

Fig. 11 - Receiver in cabinet - front view.

A SIMPLIFIED VIDICON TELEVISION CAMERA

V. K. Zworykin, L. E. Flory, W. S. Pike, G. W. Gray
RCA Laboratories Division
Radio Corporation of America
Princeton, N. J.

Within a relatively few years television has attained a pre-eminent position in the American home. It has become a primary medium of entertainment and an important source of instruction. The world of art, music, drama, current events, and sports has been opened up to the average American family as never before. In the political realm the voter has been given a familiarity with candidates, representatives, and government functions which should add new vitality to the democratic process.

All this has been accomplished by broadcast television. Up to the present it has far overshadowed all other television applications. Even so, closed-circuit, or industrial, television is finding increasing use in industry, research, and education. Great progress has been made in simplifying the construction, improving the performance, and reducing the cost of industrial television equipment. The introduction of the compact Vidicon camera has been a major step in this development. Nevertheless, costs of all types of industrial television equipment developed in the past have been too high to permit a distribution commensurate with its potential usefulness.

Hence it is not surprising that some of the greatest fields of application of closed-circuit television have not been touched. The utility of television for classroom teaching is unquestioned. Its use with the light microscope permits showing biological specimens, crystals and other materials to large groups with an ease heretofore unheard of. Televising equipment used for lecture experiments in chemistry, physics and other subjects has been amply demonstrated to be of tremendous value. Interoffice or interbuilding television for communication and transmission of written or other visual information could save large amounts of time. In small shops where the family living quarters and the shop are often in the same building a television camera can "watch" the store and transmit the information to the living quarters. Home use of closed-circuit television has never been seriously considered because of the complex and expensive equipment required.

We shall describe a device which is aimed at the fullest realization of these potential uses of television pickup equipment. This device is a camera adapter which makes every existing television receiver a potential closed-circuit television system. It takes the form of a simple Vidicon camera which transmits video signals over a cable to the receiver on carriers corresponding to unused television channels (Fig. 1). Its use does not impair the normal operation of the receiver in any way.

The camera attachment is so simple and easy to adapt to the receiver that the applications mentioned as well as many others become immediately feasible.

As before indicated, the school presents a fertile field for inexpensive closed-circuit television equipment. In an increasing number of schools the television receiver is coming to be looked upon as a standard piece of classroom equipment. The addition of even one camera attachment linked by cable to the several receivers can enhance their usefulness by a large factor. With its aid, the school principal can address all classes simultaneously, introduce distinguished visitors, and present visual demonstrations at close range to the students. All this can be accomplished with a minimum expenditure of time and without disruption of classroom schedules. If, in addition, camera units are placed in each classroom, this enables the principal to obtain an instantaneous check on classroom conditions from his office. Furthermore, teachers in training and visitors can study classroom procedures and student reactions without affecting them by their presence.

In the home a simple television system would be in great demand by hobbyists and experimenters as are movie cameras, tape recorders and the like. Further, utilitarian uses in the home such as watching children at play, monitoring the front door or checking on the baby asleep in the nursery would make the receiver in the living room not merely a means of entertainment but the real nerve center of the home (Fig. 2).

Two cameras and receivers tuned to complementary channels and linked by an inexpensive coaxial or in many cases twin lead line immediately become a two-way television communication system for inter-room use for transmitting visual information. Even a larger number of pickup points may be included in a network with a single cable connection, the observer at each receiver selecting the camera whose picture he wishes by the usual channel selector switch.

In addition to the home and the school, industry, commerce and institutions of many types can find extensive use for the equipment in question. Without elaborating upon this further, let us examine its construction and operation.

As in the standard RCA Industrial Television Equipment, the Vidicon forms the heart of the camera unit. Its small dimensions, high-quality performance, and simplicity of operation contribute materially to the success of the system. In it the photoconductive target on which the picture is projected is scanned directly by a magnetically deflected low-velocity beam. The video signal is obtained from the target electrode (Fig. 3).

In other respects considerations of economy have led to distinctly novel design features. It was realized that in many ways the circuitry required to operate a closed-circuit television system using the Vidicon pickup tube is quite similar to that already in use in television receivers. This circuitry has been highly developed for mass production and represents a very high value in terms of cost for the functions performed. Therefore, it seemed logical to investigate the possibilities of using these highly developed circuits to perform the additional functions necessary to complete a closed-circuit television system.

In general there are four functions necessary to operate such a system. First, there must be provided a viewing screen or monitor on which to view the picture. This, of course, is already available in the receiver. Second, the relatively weak signal generated by the pickup tube must be amplified to a level suitable for modulation of the kinescope. Amplification of a high order is accomplished in the receiver. This amplification is, however, performed at radio and intermediate frequencies while the signal generated by the pickup tube is usually at video frequencies extending from near dc to several megacycles. Third, voltages must be supplied to the pickup tube to generate and focus the scanning beam. In the case of the Vidicon the voltages required are in the range of those present in the receiver. Lastly, signals of proper waveform must be supplied to cause the pickup tube beam to scan the target in synchronism with the monitor kinescope. Since the Vidicon requires about one tenth the number of ampere turns of deflecting field as a 70 degree kinescope operating at 15 kilovolts the necessary deflecting power can be taken from the receiver with negligible effect upon its normal operation.

In designing the adapter the second operation mentioned, that of amplification of the weak signals from the Vidicon, required the greatest amount of attention. A typical signal current from the Vidicon is one tenth microampere. It has been found in practice that an input resistance of 50 thousand ohms is the maximum that can profitably be used and still allow adequate high frequency compensation. This means a maximum low frequency signal input of 5 millivolts. Two methods of handling the amplification are possible. First, the signal may be amplified directly as a video signal to a level to operate the video portion of the amplifier in the receiver. In most receivers this is a level of several volts. Alternatively, use can be made of the existing RF and IF amplifying circuits in the receiver by converting the video signal to a modulated carrier at some point in the camera. The latter method was chosen because the signal can be transmitted over a low impedance line to the receiver at a low power level requiring little power output from the camera and because of the convenience of coupling the signal to the receiver at RF frequency.

The signal from the Vidicon is amplified in a four stage amplifier consisting of two double tubes, to a level of about two volts. A third double tube in the camera is used as an oscillator, electron-coupled to a modulator section which in turn feeds a coaxial line to the receiver. A simplified schematic diagram of the camera is shown in Figure 4. The oscillator frequency may be adjusted to fall in any unused television channel. In this area it may, for example, be adjusted to make use of channel 8 which has a picture carrier frequency of 181.25 megacycles.

A cable connects the camera to the receiver through a control box to provide the necessary coupling circuits and to provide control of the Vidicon beam and focus from the viewing point. The cable may be up to several hundred feet in length.

The RF signal from the camera is coupled directly to the antenna circuit of the receiver through an attenuating pad to prevent any interference with normal operation of the receiver. Vertical deflection is provided by simply connecting the camera coils in series with the low side of the receiver coils. Horizontal deflection is provided through a transformer on the low side of the receiver horizontal coils as shown in the schematic of the control box in Figure 5. The transformer is used rather than a direct connection because in many receivers the coils do not return to an ac ground. Direct connection would hence require two wires in the cable carrying the signal. With the transformer one side of the secondary may be grounded and one wire is sufficient.

It is also necessary to supply a blanking signal to the Vidicon in the form of pulses occurring at the horizontal and vertical frequencies. These blank the Vidicon beam during the flyback time and prevent the generation of spurious signals due to beam return. Horizontal blanking is obtained from a pulse of the order of ten volts which appears across the camera horizontal coils. This pulse, which is made positive, is applied to the cathode of the Vidicon. Vertical blanking is generated in the control box by using a blocking oscillator transformer with a pulse stretching and clipping circuit shown on the schematic of Figure 5. The negative pulses so obtained are sent to the camera on the Vidicon grid lead and serve to bias off the beam during vertical return. A small amount of these blanking signals may also be mixed into the video signal to provide a means of synchronizing additional receivers.

The necessary dc voltages are drawn from appropriate points in the receiver circuits and are modified and controlled as necessary in the control box. As can be seen there are no tubes in the control box, the only tubes required in addition to those in the receiver being the three double tubes in the camera itself (Figure 6).

In the receiver shown in use with the camera attachment (Figure 7) and in many other standard receivers no modifications of any kind are necessary. All connections to the receiver are made by means of adapters placed under tube bases and deflection coil plugs (Figure 8). Actual adaptation of the camera to makes of current receivers representing over 50% of the total receivers made has been done. In most cases the only change to be made in the camera is in the wiring of the actual adaptors to match the particular line up of tubes. In some receivers there is no plug-in connection for the deflection coils. In these cases it is necessary to cut the wires to the coil and insert connectors in the leads. This is a simple operation which can be done without removing the chassis. A study of schematic diagrams of all of the other leading makes of receivers in current production has indicated no difficulties to be expected in designing a universal attachment which would need to differ only in minor connection detail from receiver to receiver. The camera unit can be connected to a receiver in a few minutes and thereafter can be used at will. Its presence does not interfere with normal use of the receiver. Operation can be changed from camera to off the air by a switch on the control box. The heaters in the camera may be left operating so that this changeover becomes instantaneous in either direction. Only minor modifications in the control box are required to permit the use of several camera attachments operating with a single receiver. The major change required is the provision of individual beam current and focus controls for each Vidicon camera and the provision of a selector switch to connect power and deflection to the desired camera.

A bandwidth of four megacycles is obtained so that picture definition is limited by the receiver itself. The sensitivity of the Vidicon is adequate to permit operation under room lighting conditions when standard 16-mm movie camera lenses are employed as objectives.

In summary, an extraordinarily simple and flexible closed television system is obtained by adding a suitably designed Vidicon camera attachment to a standard home receiver. The simplicity and economy of the system is the result of maximum utilization of receiver circuits and components. Its potential uses go far beyond those of the usual industrial television system. In particular, widespread use of television as a teaching aid in schools, two-way intercommunication systems, and even the use of closed-circuit television in the home become feasible, bringing us a step nearer the fullest utilization of television in its broadest sense as an extension of vision.

References

1. The Vidicon Photoconductive Camera Tube, P. K. Weimer, S. V. Forgue, R. R. Goodrich, Electronics, May 1950

2. Simplified Television for Industry, R. C. Webb, J. M. Morgan, Electronics June 1950

3. The Television Microscope, L. E. Flory, Cold Spring Harbor Symposium on Quantitative Biology, June 1951.

4. Television in Medicine and Biology, V. K. Zworykin, L. E. Flory, Electrical Engineering, Jan. 1952.

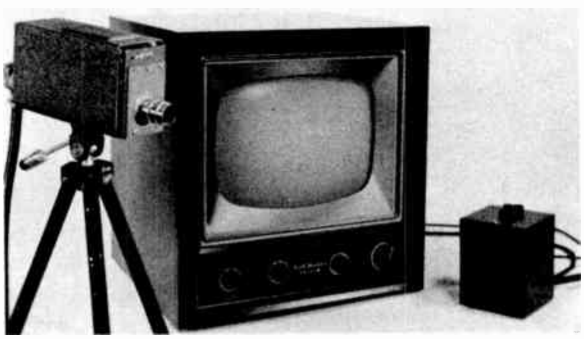


Fig. 1
Camera adapter with receiver and control box.

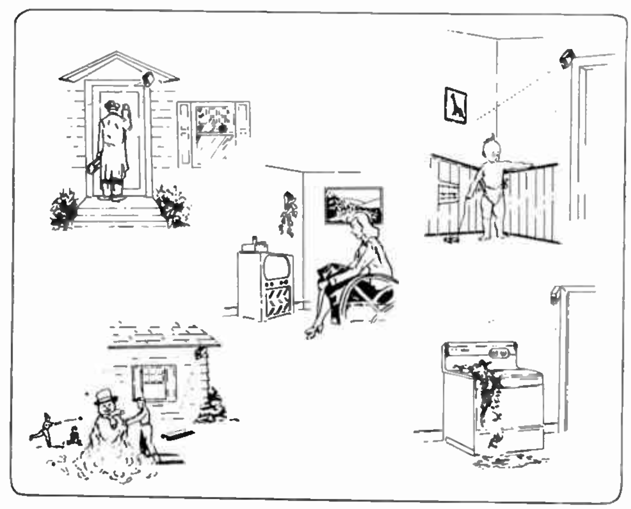


Fig. 2
Home uses of camera adapter.

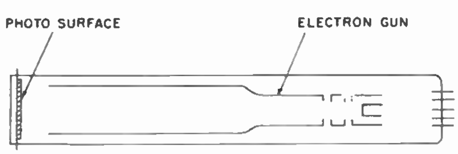


Fig. 3
RCA 6198 Vidicon.

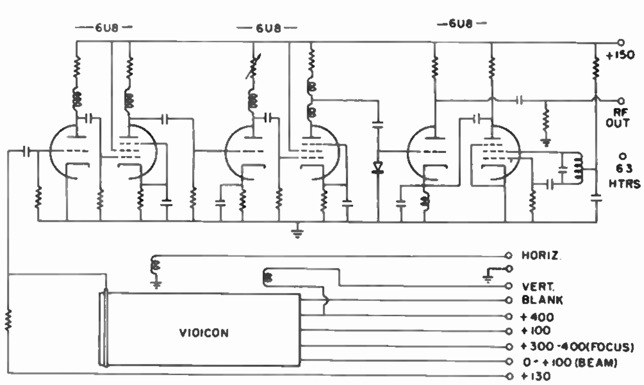


Fig. 4
Schematic diagram of camera.

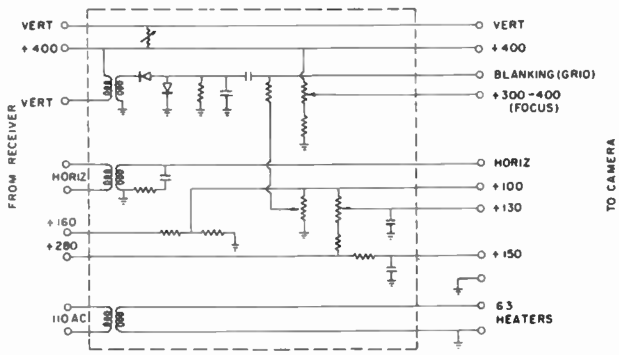


Fig. 5
Schematic diagram of control box.

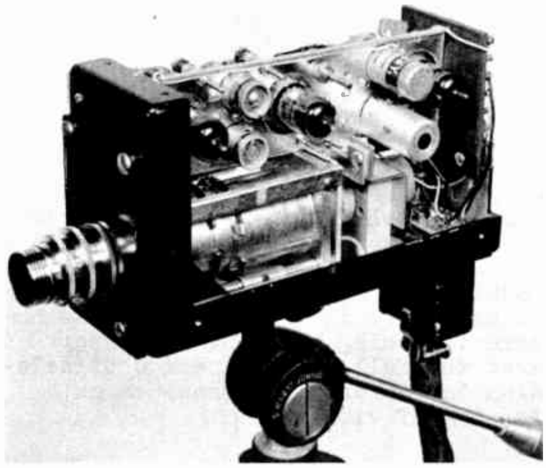


Fig. 6
Interior view of camera.

Fig. 7
Camera adapter in use with standard television receiver.

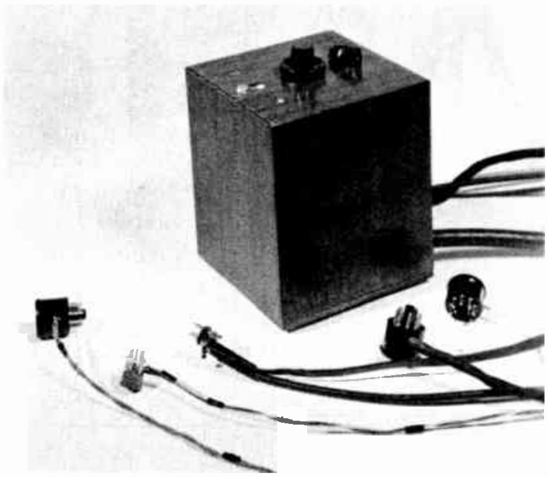


Fig. 8
Control box and adapter plugs.