

**ELEMENTS
OF
TELEVISION
SYSTEMS**

GEORGE E. ANNER

PRENTICE-HALL ELECTRICAL ENGINEERING SERIES

Elements of Television Systems

George E. Anner

ASSISTANT PROFESSOR OF ELECTRICAL ENGINEERING
COLLEGE OF ENGINEERING, NEW YORK UNIVERSITY

Prentice-Hall, Inc.
Englewood Cliffs, N. J.

PRENTICE-HALL ELECTRICAL ENGINEERING SERIES

W. L. EVERITT, *Editor*

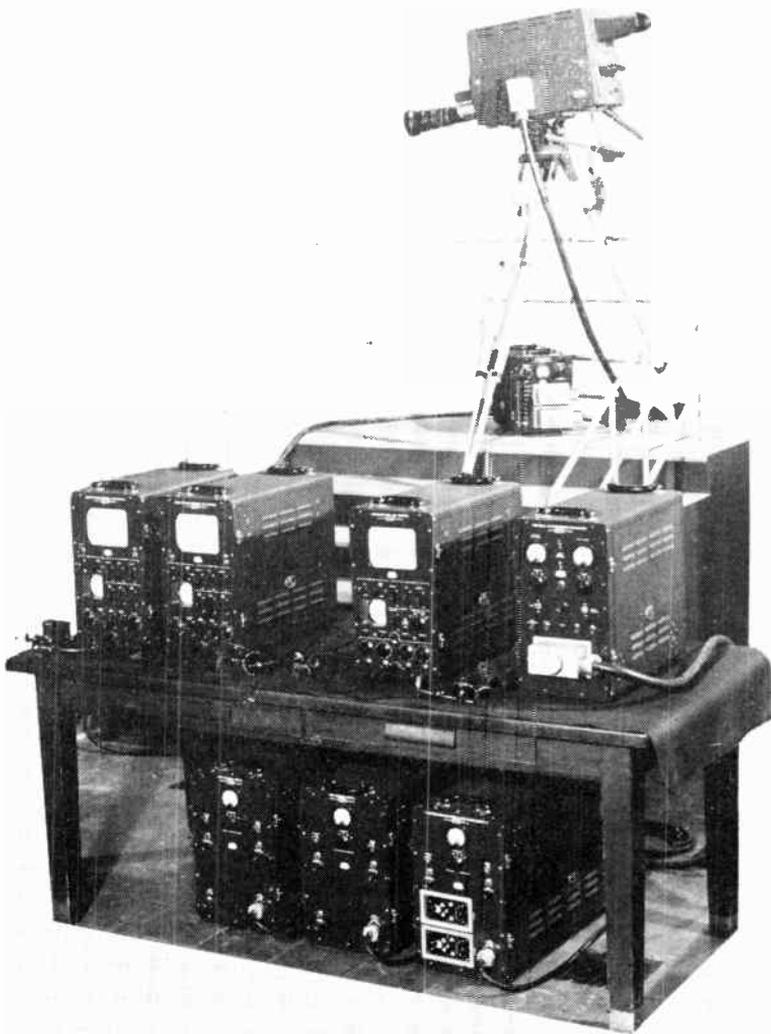
- SKRODER AND HELM *Circuit Analysis by Laboratory Methods*
PUMPHREY *Electrical Engineering*
PUMPHREY *Fundamentals of Electrical Engineering*
WARD *Introduction to Electrical Engineering*
RYDER *Electronic Engineering Principles*
RYDER *Networks, Lines, and Fields*
RYDER *Electronic Fundamentals and Applications*
MENZEL *Elementary Manual of Radio Propagation*
GOLDMAN *Transformation Calculus and Electrical Transients*
VAIL *Circuits in Electrical Engineering*
JORDAN *Electromagnetic Waves and Radiating Systems*
STOUT *Basic Electrical Measurements*
MARTIN *Ultrahigh Frequency Engineering*
THOMSON *Laplace Transformation*
BENEDICT *Introduction to Industrial Electronics*
MOSKOWITZ AND RACKER *Pulse Techniques*
FICH *Transient Analysis in Electrical Engineering*
ANNER *Elements of Television Systems*

COPYRIGHT 1951 BY PRENTICE-HALL, INC., ENGLEWOOD CLIFFS, N. J. NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY FORM, BY MIMEOGRAPH OR ANY OTHER MEANS, WITHOUT PERMISSION IN WRITING FROM THE PUBLISHERS.

PRINTED IN THE UNITED STATES OF AMERICA.

Third printing June, 1960

27326-C



An image orthicon camera chain. The camera is on the upper level. Monitors and control equipment are on the table. The units on the floor are power supplies. (Courtesy of Allen B. Du Mont Laboratories, Inc.)

Elements of Television Systems

To my wife,

BERENICE

PREFACE

Recent trends in the development of television systems for specialized industrial uses and for the transmission of images in full color have made classical methods obsolete for teaching the principles of those systems, if "classical" may be applied to a ten-year interval. These methods have concentrated on the television broadcasting system as standardized in the United States of America. It is believed that this limited point of view leads the student to misconceptions about transmission standards and, in some cases, to incorrect interpretations of design philosophy. For this reason the present work begins with a study of closed systems, those that rely upon cable connections between sending and receiving apparatus. In this manner those features of the commercial system that depend solely upon the use of radio waves as a carrier are eliminated, and the field of interest is restricted to the problems of converting an image to an electrical signal and of reversing that process.

In the second portion of the book the point of view is expanded to include the complications introduced by using a radio link in place of interconnecting cables. The last portion is concerned with methods of superimposing color-perception on a system which is inherently color blind. Since details of a new art change rapidly, an attempt is made to concentrate on the basic principles involved. Liberal use of footnote references to articles in the literature has been made so that the interested student may readily locate source material. Where there is need for an equation or a method of approach which may have been crowded out of the mind of the practicing engineer, reference has been made to only one of the many standard texts in electronics or radio engineering that cover such material.

In regard to notation an attempt has been made, in so far as possible, to use a consistent set of symbols throughout the book, even when this policy requires notation different from that used in the source material. Free use has been made of the symbols a-c and d-c as adjectives, and frequency is specified in cycles, kilocycles, or megacycles, the "per second" being omitted to conform with the current,

though inelegant, trend. In several sections, the angular frequency, ω , is referred to as frequency when no chance for ambiguity is present.

It is a pleasure to acknowledge the invaluable assistance of Mr. Leonard Mautner and Dean W. L. Everitt, who furnished constructive criticism of the manuscript, and also Dr. Irving Wladaver for his help in preparing much of the photographic material. My thanks are also extended to Mr. P. H. Gridley and his associates for their cooperation and excellent workmanship in the preparation of the illustrations.

A number of technical societies, publishing companies, and manufacturers have been generous in furnishing diagrams, technical information, and samples that were needed for the preparation of the book. While too numerous to list here, they are credited throughout the book for the material they have furnished.

GEORGE E. ANNER

Yonkers, N.Y.

CONTENTS

Part I: CLOSED TELEVISION SYSTEMS

| | | |
|----|--|-----|
| 1. | TRANSMISSION OF PICTURES | 1 |
| | Definitions. A matter of dimensions. Facsimile vs. television. A facsimile system. | |
| 2. | PICTURE STANDARDS | 13 |
| | Picture shape. An estimate of the system bandwidth. Speed and direction of scan. The frame frequency or picture repetition rate. | |
| 3. | SCANNING METHODS | 37 |
| | Mechanical scanning. Development of electronic scanning. ELECTRONIC SCANNING. Electrostatic focus and deflection. Magnetic focus and deflection. Comparison of electrostatic and magnetic deflection. SCANNING GEOMETRY. Progressive scan. The random scan. Interlaced scan. Even-line interlace. Summary. | |
| 4. | SCANNING GENERATORS | 78 |
| | Transients in the $R-C$ circuit. Generators of saw-tooth voltage. Trigger tubes. Multivibrator. Design procedure for the multivibrator. The order-of-magnitude equation. Wave forms of the plate-coupled multivibrator. Synchronizing the multivibrator. The blocking oscillator. MAGNETIC DEFLECTION SYSTEMS. Increasing or decreasing sweep current. Driver grid voltage. Trapezoidal generator. Flyback considerations. Shunt-resistance damping. Resistance damping. Advantages and disadvantages of resistance damping. Diode damping. Reaction scanning. The flyback power supply. | |
| 5. | SCANNING AND PICTURE REPRODUCTION. | 150 |
| | The one-dimensional series. The two-dimensional series. Interpretation of the double Fourier components. The electrical signal and its spectrum. Confusion in the signal. The effect of interlaced scanning. Aperture distortion. A digression on exponential series. One-dimensional case. Reconstruction of the image at the reproducer. The two-dimensional case. The normal component. Extraneous components. Scanning with electron beams. | |

| | |
|--|------------|
| 6. CAMERA TUBES | 185 |
| Static-image generators. Some elements of photometry. Photo-emission. Noise considerations. The flying spot pickup device. THE IMAGE DISSECTOR. Method of operation. The output current. Electron multiplier image dissector. Multiplier noise. Magnetic focusing. THE ICONOSCOPE. The storage principle. Electron bombardment of an insulated surface. Shading. The coupling circuit. Bias lighting. D-C insertion. Keystone correction. An appraisal. The image iconoscope. THE ORTHICON. Low-velocity-electron scanning. The deflection system. Characteristics. THE IMAGE ORTHICON. The image intensifier section. The target and scanning section. The electron multiplier. Noise. Characteristics. The image orthicon preamplifier. | |
| 7. VIDEO AMPLIFICATION | 265 |
| Steady-state requirements. The resistance-coupled amplifier. Transient response of a video amplifier. Unit function. Transient response of the <i>R-C</i> amplifier. The Fourier integral. Cascaded video amplifiers. HIGH-FREQUENCY COMPENSATION. Shunt compensation. Simple shunt compensation. Freeman-Schantz compensation. The vector diagrams. Transient response of the shunt-compensated amplifier. Series compensation. Series-shunt compensation. Summary of high-frequency compensation. Square-wave testing. Calculation of square-wave response. LOW-FREQUENCY COMPENSATION. Coupling-network effect. Cathode and screen degenerative effects. Compensation network. Bias circuit compensation. Coupling circuit compensation. Summary. Low-frequency transient response. Clamping. FEEDBACK VIDEO AMPLIFIERS. Mid-band. Low-band. High-band. The cathode follower. | |
| 8. CLOSED TELEVISION SYSTEMS | 342 |
| Type I—closed system. Type II—closed system. Type III—closed system. The supersync signal. The composite video signal. Supersync and picture separation. | |
| Part II: THE COMMERCIAL TELECASTING SYSTEM | |
| 9. SPECIAL PROBLEMS OF TELECASTING | 365 |
| The standardization committees. The commercial television system, block diagram. | |
| 10. THE OPTIMUM NUMBER OF LINES | 374 |
| Some factors governing resolution. Overlap, line structure, and resolution. Vertical resolution. Horizontal resolution. Over-all resolution. Circuit factors. Test pattern and resolution. | |

| | | |
|-----|--|-----|
| 11. | SYNCHRONIZATION | 395 |
| | Synchronization requirements. Amplitude components of the composite signal. Wave-form components of the composite sync signal. Differentiating circuit. Integrating circuit. DEVELOPMENT OF THE COMPOSITE SYNC SIGNAL. Horizontal sync pulse. Vertical sync pulse. Interlace requirements. Serration. Equalization of vertical fields. Blanking. THE SYNCHRONIZING SIGNAL GENERATOR. General requirements. Counter circuit. Reactance tube. Comparator circuit. Voltage-adding circuit. Delay networks. Familiar components. Operation. SYNC SIGNAL TESTING. Saw-tooth sweep. Sinusoidal sweep. Phase shifter. Dot generator technique. Pulse cross testing. | |
| 12. | VESTIGIAL-SIDEBAND TRANSMISSION | 446 |
| | Amplitude modulation. Single-sideband transmission. Vestigial-sideband transmission. RECEIVER AND TRANSMITTER ATTENUATION. Transmitter attenuation. Receiver attenuation. Comparison. VESTIGIAL-SIDEBAND FILTERING. Off-center tuning. Type A or crossover filter. Calculation of the distributed constants. Type B or notching filter. The reactive vestigial-sideband filter. The Balun. | |
| 13. | THE VIDEO TRANSMITTER | 484 |
| | Positive or negative transmission. Transmitter block diagrams. PLATE VS. GRID MODULATION. Plate modulation. Grid modulation. THE R-F SECTION. Oscillator. R-F amplifiers. Grounded-grid amplifier. THE VIDEO SECTION. Block diagram. Line amplifier. The video amplifiers. The constant-resistance network. The d-c component. THE MODULATING SECTION. The modulated stage. The plate load. TRANSMISSION LINE AND ANTENNA SECTION. The coupling loop and Balun. Diplexer. TRANSMITTING ANTENNAS. General requirements. The dipole. Crossed dipoles. Vertical directivity. The turnstile antenna. Broadbanding. The super turnstile. Testing. | |
| 14. | RECEIVERS | 546 |
| | THE CONVENTIONAL RECEIVER. General discussion. Available power and noise figure. Minimum noise figure of an ideal receiver. R-F stage. The frequency mixer. The local oscillator. I-F system. The video system. The deflection circuits. Automatic gain control. Summary. THE INTERCARRIER RECEIVER. Principle of operation. The intercarrier circuit. Recommended changes in transmission standards. PROJECTION SYSTEMS. Refractive projection systems. Reflective projection. The Skiatron system. | |

15. STAGGER-TUNED AMPLIFIERS 622
Single-tuned amplifier. The staggered pair. The staggered "n-uple."
Asymptotic forms.
16. THE RECEIVING ANTENNA 648
Carrier frequency requirement. Elements of V.H.F. propagation.
General requirements for the receiving antenna. BASIC ANTENNA
STRUCTURES. The dipole. The folded dipole. The parasitic ele-
ment. Fans and vertical V's. Spikes. Horizontal V. Vertical
directivity. Rotatable antennas. MASTER ANTENNA SYSTEMS. Video
distribution. R-F distribution.
17. TELEVISION MOTION PICTURES 682
The basic problem. The film scanning interval. 24 frames to 60
fields. The 35-millimeter projector. The 16-millimeter projector.
The shutterless 16-millimeter projector.

Part III: COLOR TELEVISION SYSTEMS

18. COLOR TELEVISION 693
ELEMENTS OF COLOR. Response of the eye. Color matching. The
color triangle. The color range of any set of primaries. Analysis
and synthesis. The color filter. Color synthesis by addition.
Color synthesis by subtraction. COLOR TELEVISION SYSTEMS. Basic
color analyzing systems. Basic color transmission systems. Basic
color reproducing systems. Bandwidth and compatibility. THE
R.C.A. SIMULTANEOUS COLOR TELEVISION SYSTEM. Color-channel
bandwidth. Methods of transmission. The receiver. Bandwidth
reduction by the mixed highs technique. THE C.B.S. SEQUENTIAL
COLOR TELEVISION SYSTEM. The color-switching interval. Frame
and field standards. Color disk phasing and speed control. The
supersync signal. Pickup and receiving equipment. THE C.T.I.
LINE-SEQUENTIAL COLOR SYSTEM. Color interlace. THE R.C.A. DOT-
SEQUENTIAL COLOR SYSTEM. The sampling process. The receiver.
Compatibility. THE C.T.I. SEGMENTAL-SEQUENTIAL SYSTEM. Prin-
ciple of operation. Color synchronization. The color receiver.
Compatibility.
- APPENDIX: Dot Systems of Television Transmission 761
The sampling process. Pulse transmission bandwidth. Dot inter-
lace. Dot interlace and multiplex.
- PROBLEMS 773
- INDEX 791

Part I

CLOSED TELEVISION SYSTEMS

CHAPTER 1

TRANSMISSION OF PICTURES

1-1. Definitions

Television has been defined as the electrical transmission of a succession of images and their reception in such a way as to give a substantially continuous reproduction of the subject or scene before the eye of a distant observer. Any complete electrical network that meets these requirements is termed a television system and its function is to generate, transmit, and reproduce video or picture information. Since the information to be conveyed is not electrical in nature, the network must have at least three basic elements: first, a transducer that serves to convert the optical data into some sort of equivalent electrical signal; second, an electrical link to carry these signals to a distant point; and third, another transducer that reconverts the electrical signals into an optical image which is acceptable to an observer at the distant point. These three elements are directly related to the threefold function of the television system stated above: a transducer to generate, a link to transmit, and a transducer to reproduce the desired information.

What additional components are required over and above these three depends largely upon the character of the information, the speed of transmission, and the characteristics of the transmission medium. Thus, if an intelligent engineering approach to the subject of television systems is to be taken, we must first look to the nature of the information which the system must handle.

The foregoing statements are more or less applicable to any electrical communications network which is to transmit information that is nonelectrical in itself. For example, in the transmission of speech by telephone, microphone and receiver serve as transducers with the connecting link furnished by wires. The additional elements of the telephone system—such as repeating amplifiers, modulators, filters and detectors in a carrier system—depend on the wire transmission characteristics, the time allowed for transmitting a given amount of

information, and the very nature of speech itself. In order to design transducers and other components that will give a satisfactory reproduction of the speech input, the engineer must first know the characteristics of the signals with which he is dealing. Hence, in our approach to the elements of television systems we start first with a brief analysis of a picture and the problems of transmitting such a picture over an electrical communication channel. From this we may proceed to a study of transducers and their associated equipment which will provide an electrical signal acceptable to the channel. The various components may then be joined into typical television systems.

Throughout this entire procedure we must remember that practical picture transmission involves the production of the illusion of a complete picture for the distant observer and that the function of the television engineer is to design, build, and maintain a commercially feasible system which will accomplish this illusion.

1-2. A Matter of Dimensions

Since the transmission of picture information is to take place over an electrical link of some sort, it seems desirable to see if such a link imposes any broad limitations on the type of signals which it can convey.

By definition, the response of an electrical circuit is a single-valued function of time. By this we mean that if signals from a number of sources are applied to the input terminals of a four-terminal network, they will add in such a manner that the voltage at the output terminals will have one, and only one, value at any instant of time. Whereas it is true that the output signal over a period of time may be resolved into separate frequency components, still it has only one amplitude instantaneously. We might say, then, that an electrical link is limited in that it may accommodate only a single-valued function of time. As an illustration of this principle we may consider the network shown in Fig. 1-1. An external noise source introduces a component of current $i_N(f,t)$, which adds to the signal current $i_S(f,t)$. A current $[i_S(f,t) + i_N(f,t)]$ which has only one amplitude at any instant will be delivered to the receiver. It would seem, therefore, that a communication channel is limited in the type of signal it can handle; all components will add to give a single-amplitude function of time.

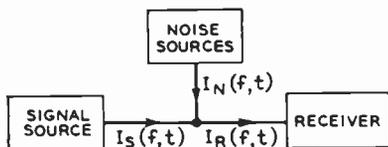


Fig. 1-1. Signal and noise combine to give a single-valued function of time.

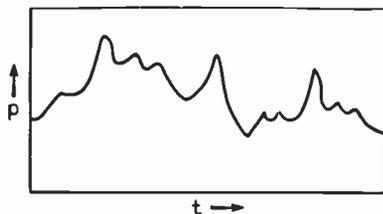


Fig. 1-2. Sound pressure is a single-valued function of time.

Fortunately for the communication art most forms of intelligence that are transmitted by electrical means are either inherently single-valued or may be forced into a single-valued form. Speech is an example of the first type because the spoken message may be represented at any point in space by a plot of sound pressure against time. This is illustrated in Fig. 1-2. If a suitable microphone is placed at that point, the variations in pressure would be converted into corresponding variations in voltage or current, and, in the absence of distortion, the electrical signal would also appear as in Fig. 1-2 but with the pressure scale replaced by voltage or current. Thus in the transmission of speech by wire—and the discussion may be extended to include radio communication—the entire message may be represented by a single-valued function of time.

Turning to the problem at hand, we have an example of the second type of intelligence. Let us consider that a picture or photograph of some sort is to be transmitted and, further, let us try to handle picture transmission in the same manner as speech. In place of a microphone a phototube is used as a transducer. If the light from the photograph is focused onto the phototube, an electrical current will be developed which once again is a single-valued function of time, but notice that all detail of the image is lost in the electrical signal. Since the phototube “sees” the entire image, the best it can do is to develop a signal that is proportional to the average brightness of the image. It has no way of recognizing that the picture is in fact a bounded, two-dimensional continuum of brightness. We might say that there is a complete incompatibility between the information to be transmitted, which is inherently two-dimensional, and the electrical system, which is only capable of handling a one-dimensional signal.

This basic incompatibility may be resolved, however, by a compromise. Experience with photographs and half-tone printing methods has shown that as a result of the limited acuity of the eye a perfectly satisfactory picture results if the continuum of brightness is replaced by a field of elemental areas, the brightness of each being the average brightness of the corresponding area in the original image. This concept is illustrated in Fig. 1-3, where the actual variation in

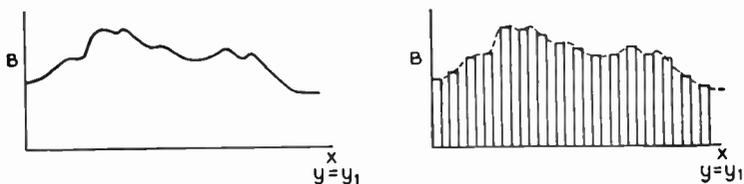


Fig. 1-3. The actual distribution of brightness is replaced by a number of samples.

brightness across the picture is replaced by a large number of samples of average brightness. If this process be extended in both dimensions over the picture, it may be seen that the original picture will be replaced by M smaller pictures, each of which is of constant brightness. M may be termed the "figure of merit" of the approximate picture, and the greater the number of samples, the closer will be the approximation of the original scene.

How this procedure is accomplished in an actual system will be described later. For the moment a simple system may be proposed. Let the picture be focused on a field of infinitesimally small phototubes, spaced closely together. Then the output of each phototube is proportional to the average brightness of an elemental area. Now

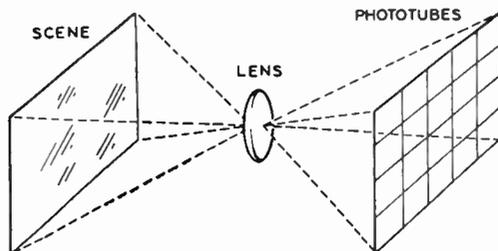


Fig. 1-4. The scene is focused onto an array of phototubes.

if the information so derived is to be transmitted to some point over a wire system, for example, a reproducing system which will convert voltage to brightness is required, which further has a one-to-one correspondence to the phototubes in regard to both number and position.

Since the phototube for each picture element delivers the proper type of electrical signal for a communication channel, the entire picture may be transmitted by connecting the M phototubes to M reproducing elements so that signals from all the elements are transmitted simultaneously in parallel. Care must be taken that each pair of pickup and reproducing elements so connected occupy corresponding positions in the over-all picture field so that the space co-ordinates are restored. Whereas such a parallel transmission system as shown in Fig. 1-5 appears feasible on paper, its accomplish-

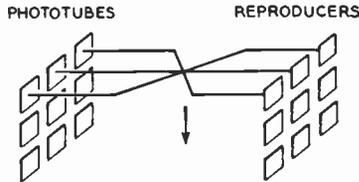


Fig. 1-5. Parallel transmission.

ment is practically out of question when it is realized that for a satisfactory picture some 100,000 elements and channels would be required. This number is determined by such factors as picture size and viewing distance, but 100,000 is a good round number to represent the order of magnitude of the problem posed by this parallel transmission system in which information from all M elements is transmitted simultaneously and over parallel communication channels.

One practical value of the parallel transmission system just described is that it illustrates the concept of sampling which is of great importance in picture transmission: one large area whose brightness varies from point to point may be replaced by M small areas or samples, the brightness of each being a sample of the brightness in the corresponding region of the original picture. The purpose of the sampling process is to render the bidimensional information amenable to transmission over M electrical systems.

We must now find a means of reducing the number of channels. We have seen that the picture may be sampled in space. If now it be sampled in time as well, the problem is solved. Suppose that two rotating, sampling switches are introduced into the system as shown in Fig. 1-6. With this arrangement the phototube outputs are

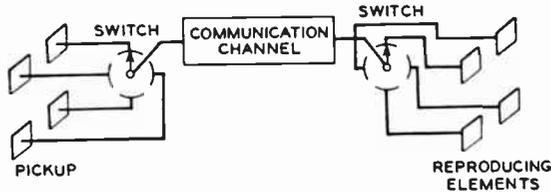


Fig. 1-6. Sequential transmission. Picture information from the several pickup elements is sampled in sequence.

sampled in an orderly fashion and a complete compatibility exists: the two-dimensional information has been reduced to a single-valued function of time. Nevertheless some sacrifice is made for the compromise. The original image has been resolved or broken up into M elements with a loss in resolution or detail, and the entire picture is never displayed in its entirety at any instant at the reproducer. This second fact requires that the sampling process occur at a rate sufficiently high that the eye may reassemble the picture and see its motion reproduced in a satisfactory manner.

This second or sequential system of picture transmission, in which the picture elements are sampled in time, forms the basis of modern television systems and, as such, needs closer examination. Consider the requirements on the three main components of the system.

Pickup System:

- (1) Is a photoelectric transducer.
- (2) Breaks the picture up into M samples or elements in space.
- (3) Has a selecting mechanism that samples these elements in some orderly manner in time and delivers these samples in sequence to the communication channel.

Reproducing System:

- (1) Is an electrophoto transducer.
- (2) Has a selecting mechanism which, operating in exact synchronism with the pickup selector, delivers the time samples from the communication system to the proper playback elements so that each sample is restored to its proper position in space.

Communication Channel:

(1) Must have the capabilities, i.e., bandwidth and delay characteristics, to deliver all signal and switch-synchronizing information furnished by the pickup system over the required distance to the reproducing system, and all this with a minimum of distortion.

Over a period of years considerable ingenuity has been used to develop devices which at the pickup end combine the functions of breaking up the picture into elements, selective switching, and converting light to voltage or current. Typical of these are the image dissector, the iconoscope, and the image orthicon. At the reproducing end the same functions have been combined into the cathode-ray tube, or CRT, which is used exclusively for direct-view television service and will be assumed as the reproducing transducer through the rest of this book.

The combined functions of breaking the scene up into elemental areas and switching them is termed "scanning." In the several devices mentioned above, scanning is accomplished by causing a beam of electrons to be deflected over either a suitable photoemissive surface for pickup or a fluorescent surface for reproducing.

1-3. Facsimile v. Television

Generally speaking there are two types of service in which picture information is transmitted: facsimile and television. In the former, a photographic image is delivered at the playback end, whereas in television an optical image must be presented which is suitable to the human eye. Because of this difference in display at the receiving end the actual details of the two systems differ considerably. Consider the problem in television. It has been stated before that the entire scanning process must proceed at a fast enough rate so that the eye can successfully integrate the samples which are displayed in sequence into a complete picture. The eye never sees a complete image of the picture, but only a small part of it at any given instant. The question arises, then, how fast must the entire picture area be covered by the scanning device? A number of factors are cogent, two of them being retina retentivity in the eye and, where phosphor-coated CRT's are used for display, the glow time of the phosphor. The former is a psychological phenomenon in which the eye "sees" an image for a short period after the stimulus is removed. It is this property of the eye which makes possible effective integration of all the information which has been presented in sequence.

The second factor is a characteristic of all phosphors or fluorescent materials which after being bombarded by an electron stream, continue to glow with the intensity decreasing exponentially in time. The actual decay time from the removal of excitation until the intensity reaches the threshold of visibility may be controlled by proper composition of the fluorescent material being excited. In P4 screens of the type used for television service this decay time is in the order of a millisecond.¹ This short but finite glow period is of considerable aid to the eye in reassembling the entire picture from its many parts.

Another factor to be considered is that motion in the original televised scene must also be reproduced. This problem is shared in "moving" pictures in which a series of "still" photographs must be presented in rapid enough succession so that apparent motion results. The cinema industry has found that if the eye is presented at least 16 complete pictures per second, then the illusion of motion is satisfactory. This figure of 16 pictures per second also presents the sequential data to the eye rapidly enough so that not only motion but the entire television picture itself appears satisfactory to the eye. It will be seen later that flicker difficulties place further demands on the speed of picture presentation.

Summarizing, it is seen that the eye demands at least one picture in $\frac{1}{16}$ second. In facsimile service on the other hand, a photographic image is constructed and the demands on speed of presentation are much less stringent. In fact, 20 or more minutes may be utilized in the transmission, reception, and reproduction of a single still picture. Translated in terms of scanning requirements, this means that the scanning in facsimile work is extremely slow and mechanical means for moving some scanning element may be used. With this slight background, a facsimile system may be described which will illustrate the principal components required and the interrelationship between them.

1-4. A Facsimile System

It will be assumed that a high-quality photograph of a normal size of 7 in. by $8\frac{1}{2}$ in. is to be transmitted. If this is illuminated, the light reflected from its surface will contain information as to the

¹I. G. Maloff and D. W. Epstein, "Luminescent Screens for Cathode Ray Tubes." *Electronics*, 10 (11), 31 (November 1937).

brightness distribution over the picture area. This must be converted to electrical energy so that the first component required in the system is a photoelectric transducer. A phototube or photocell will serve in this capacity.

Secondly, means must be provided for the reduction of the picture into a number of elemental areas and for switching them to the communication channel in orderly sequence. In the previous discussion of sequential transmission an array of infinitesimally small phototubes plus a switch was proposed. In the interest of simplicity, however, assume that this multiplicity of stationary phototubes is replaced by a single phototube which scans across the image. By this simple expedient the two functions of breaking into elements and switching may be combined. In a typical type of facsimile equipment² the scanning phototube is fixed, and the copy is made to move in front of it, as shown in Fig. 1-7. In order that the photocell see only a small elemental area of the picture a special optical system is used. As may be seen from the diagram, light from an exciter lamp

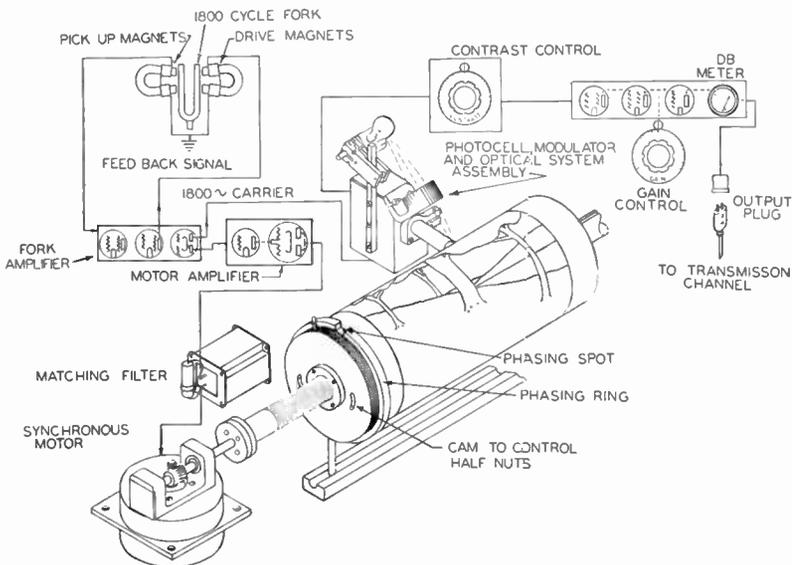


Fig. 1-7. The pickup equipment of a facsimile system. (Courtesy of Times Facsimile Corporation.)

² *Elements of Facsimile Communications*, The Times Facsimile Corporation.

of constant intensity is focused onto the moving copy. The reflected light passes through a small defining aperture in the light-tight housing to the photocell. Ideally this aperture should be infinitesimally small since it determines the size of the picture elements and hence the detail in the final image. Practical considerations rule out this possibility, however, and the actual aperture size is about 0.01 in. square. We shall see in Chapter 5 that the finite size of this scanning aperture introduces distortion in the electrical signal and that the nature of the distortion is such that it may be corrected electrically.

Since the photocell provides the required conversion from light to voltage, we next consider the scanning means which move the picture area across the aperture. It may be seen in Fig. 1-7 that the copy is fixed on a drum that is keyed to a revolving lead screw. Thus when the synchronous motor operates, the copy drum revolves and moves along the lead screw. In this manner the entire copy is scanned. Speed control of the entire process is maintained by driving the synchronous motor from a stable-frequency tuning-fork oscillator. The output of the fork oscillator is also used as a carrier which is amplified and fed to the communication channel.

The inverse process must be carried out at the reproducing end of the system; that is, a similar scanning drum is required to restore the sampled information to its proper location in the image, and the electrical signal must be converted back to light so that it may expose the photosensitive paper which is wrapped around the receiving drum. The major portions of the reproducing system are shown in Fig. 1-8. The drum is driven at proper speed by a fork-controlled synchronous motor and the optical system consists of a focusing lens and a neon light whose luminous output is proportional to the incoming electrical signal. The final image is produced by developing the photographic image, which has been exposed on the revolving drum.

Now it can be seen that if the image is to be reproduced properly, both drums must rotate in synchronism. Since both motors will rotate at the same speed as a result of the action of the fork oscillators, it is only necessary that the two drums start scanning in phase. This is accomplished in the following manner: At the beginning of a transmission the reproducing drum is prevented from rotating by the clutch, stop arm, and trip magnet shown in Fig. 1-8. As the recording drum of Fig. 1-7 starts to rotate, the photocell scans the phasing spot and transmits a phasing pulse which is fed to the communication

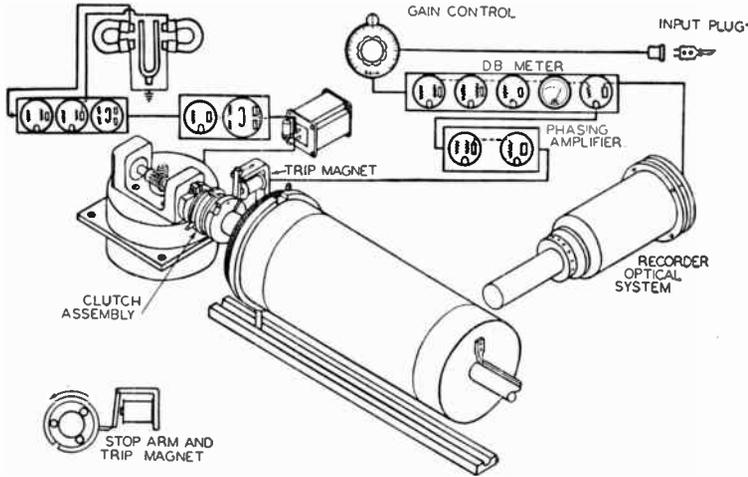


Fig. 1-8. The reproducing equipment of a facsimile system. (Courtesy of Times Facsimile Corporation.)

channel. When this pulse arrives at the receiving end, the trip magnet releases and the reproducing drum begins to revolve. As we have seen, the control afforded by the fork oscillators will hold the two drums in synchronism for the duration of the transmission.

Some idea of the performance of this facsimile equipment may be had from the following figures: With the two drums rotating at 90 revolutions per minute an image of 7-in. by $8\frac{1}{2}$ -in. size may be transmitted in seven minutes within a 2000-cycle bandwidth. In more familiar terms, average typewriter copy may be transmitted at a rate of approximately 130 words per minute.

As we extend the ideas of picture transmission we shall see that, for a given amount of transmitted detail, bandwidth and time of transmission may be interchanged. We shall also see that in the television system a more complex system of synchronizing the pickup and reproducing scanning motions is required.

It would be desirable at this point to set up an analogous simple television system so that the requirements of the two types of service, facsimile and television, could be compared. It has been pointed out, however, that the scanning speeds required in television are so high that the use of some mechanical scanning system such as described

above is precluded.³ Recourse must be made to electronic scanning systems and a description of the all-electronic television system will be postponed until certain of its important components have been discussed.

These chief components are the pickup and display devices plus the units required to produce scanning in them. The over-all approach in the remainder of this book will be to describe, and where feasible, to set up design procedures for the various components required for a closed system comprising a single pickup assembly and a single reproducing assembly interconnected by wire. This simple type of system has many fields of application.⁴ With the simple system completed, the problems of a commercial television broadcasting system will be considered. Here the problems are more complex, first because a large number of reproducing assemblies or receivers are associated with any given transmitter, and second because the wire link between sending and receiving ends is replaced by a radio communication system. These added complexities require considerable modifications in the simple closed system. The over-all treatment, then, is a progression from the specific to the more general.

In developing the closed system we shall follow a definite method of approach: In Chapter 2 we shall set up a number of standards which specify the pattern in which the image is scanned. In Chapter 3 we study the various means of producing the required scanning pattern and set up the wave forms of voltage required to cause an electron beam to follow the prescribed scanning pattern. With the scanning voltage specified we consider circuits for generating these voltages in Chapter 4. Next a study is made of the distortion introduced by the scanning process. The three remaining chapters take up the television camera tubes, the video amplifier circuits used for amplifying their output signals, and assemble these components into three basic types of closed television systems.

³ A notable exception to this statement is the Scophony system of scanning which has enjoyed some popularity in Great Britain. The post war trends in the United States, however, have been toward all-electronic scanning. See H. W. Lee, "The Scophony Television Receiver." *Nature*, **142**, 59 (July 9, 1938). See also Fig. 14-35.

⁴ R. W. Sanders, "Industrial Television." *Radio and Television News* (Radio and Electronic Engineering Edition), **12**, No. 2, 3 (February 1949).

CHAPTER 2

PICTURE STANDARDS

A successful communication system must have agreement between transmitter and receiver concerning how certain functions are to be performed. For example, in speech it is understood that sound is to be the common carrier and, except for the rare case of the lip reader, it is assumed by the message source that an ear, rather than an eye or nose, is to be the receiver. Once this agreement between transmitter and receiver has been reached, standards may be set up specifying how these various functions shall be performed. In the case of the various commercial broadcasting services where individual agreements between the tremendous number of transmitters and receivers would be impractical, the Government in the form of the Federal Communications Commission establishes and enforces the standards. The need for this standardization is particularly strong in telecasting services because of the so-called "lock and key" relationship between the two ends of the system; all receivers within the service area of a given transmitter must be able to receive and display its program material satisfactorily. On the other hand, in a simple closed wire-connected system involving only one transmitter and one receiver these standards may be held to the barest minimum.

In the discussion of standards, then, the elements requiring standardization fall into two broad categories: those pertaining to any system, closed or broadcasting, and those relating more specifically to commercial telecasting by radio. Those of the first group which have to do with the picture proper and the method by which it is scanned are the subject of this chapter.

In a closed system what factors must find agreement between the pickup and playback ends? Probably the most obvious answer is picture shape. Furthermore, if reference is made to the general requirements of scanning outlined in the last chapter, it will be seen that speed, geometry, and direction of scan must also be included.

Other factors are the figure of merit or number of elemental areas used and the rate of picture transmission. These will now be considered.

2-1. Picture Shape

At first glance the choice of one picture shape from the infinite number of possible shapes is a difficult one. Practically speaking, however, the choice is quite limited because scanning problems require that the shape of the scanned area be some regular and simple geometric pattern. Thus the field of choice is narrowed on this basis to three shapes: the circle, the square, and the rectangle.

It has been stated previously that the most common display device in modern television is the cathode-ray tube. Since for constructional reasons these tubes are often built with circular viewing screens,¹ it seems reasonable to choose a circular picture shape, because then the entire fluorescent surface of the tube may be used for display of the picture. The following question must now be answered: Given a circular picture shape, can it be scanned in some relatively simple pattern? Two reasonable possibilities exist which are shown in Fig. 2-1. A set of closely spaced horizontal (or vertical) lines may be used. The solid lines indicate a "forward" trace; the dotted lines, a "reverse," retrace, or flyback path over which the electron beam returns to its pattern-starting point. For ease of scanning, picture data is presented on the forward trace only. During flyback, which is made as short as possible, the electron beam is blanked out to prevent contamination of the image display. Despite the circular tube shape all the scanning lines are shown of equal length in order to ease the requirements on the scanning generator.

From the viewpoint of utilizing as much as possible of the scanning time for picture presentation the horizontal linear scan of Fig. 2-1a is poor. On the other hand, it provides maximum utilization of the tube face. Notice that an alternative system would shrink the

¹ Noncircular cathode-ray tubes have been built but they were not available commercially until the latter part of 1949. Tubes of this type were used by the Columbia Broadcasting System in receivers developed for their sequential color television system. These tubes had an end shape approximately rectangular with slightly rounded corners. See, for example, P. C. Goldmark, J. N. Dyer, E. R. Piore, and J. M. Hollywood, "Color Television" Part I. *Proc. IRE*, **30**, 4 (April 1942). See also, P. C. Goldmark and R. Serrell, "Color and Ultra-High-Frequency Television." *Proc. 1st National Electronics Conference*, 182 (October 1944).

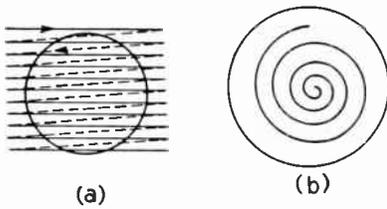


Fig. 2-1. Scanning a circular area. (a) Linear horizontal scan. (b) Spiral scan.

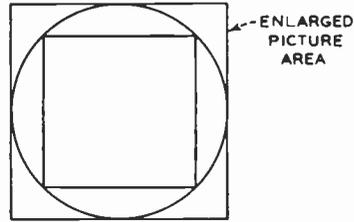


Fig. 2-2. An enlarged image may be obtained by expanding the sweep, but only with the loss of a large portion of the picture.

scanning pattern so that it lies entirely within the limits of the tube face, in which case all of the sweep time is utilized, but part of the fluorescent screen is lost. In effect, then, we may trade time utilization for space utilization. These two alternative schemes are illustrated in Fig. 2-2. It may be seen that the smaller scan results in a square or rectangular picture shape, which will be discussed later in this section.

We have just seen that the linear horizontal scan of a circular image is wasteful of large intervals of scanning time. Consider a second alternative, the spiral scan. Here the entire scan may be used for display. Hence it appears that the spiral scan is satisfactory on this basis. But other questions arise, such as: Can this scanning pattern be generated without too much difficulty, and what, if any, are its disadvantages?

In answer to the first question assume that a CRT employing electrostatic deflection is used. Figure 2-3 shows the circuit diagram of one type of circuit which will provide the necessary voltages to generate the spiral. In the diagram the center-tapped input transformer with R_1 and C_1 constitutes a phase shift network which feeds sinusoidal voltages in quadrature to the grids of V_1 and V_2 . If, for the moment, the screen grids are held at a constant voltage, the amplified grid voltages will cause the electron beam to follow a circular path on the cathode-ray tube screen. If the screen voltages now be varied in a saw-tooth manner as shown in the diagram, the amplification of each stage will change with the screen voltage. This means that the electron beam follows a circular path of increasing radius or a spiral.

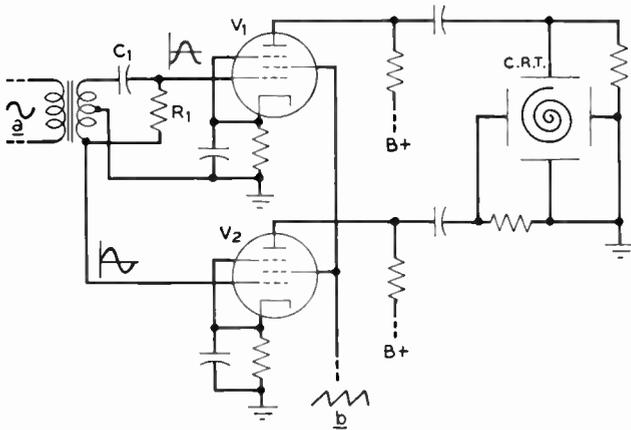


Fig. 2-3. Circuit for generating a spiral scan on an electrostatic deflection cathode-ray tube.

The number of revolutions in the complete spiral is determined by the relative periods of the two wave forms, the sinusoid *a* and the saw-tooth *b*. For example, if the rise time of the saw-tooth is 1 millisecond and the period of the sinusoid is 10 microseconds, the electron beam will complete 100 revolutions as it moves radially from the center to the outer edge of the screen.

The circuit just described, although practical from the point of view of circuitry, has a prime disadvantage, which is the result of the nonconstant linear velocity of the scanning beam. The same sort of problem exists in the making of phonograph records. Since the angular speed (or speed of rotation) remains constant, the linear speed increases as the spot moves away from the center of the screen. On the cathode-ray tube this means that in the regions near the tube center where the linear velocity is low, the fluorescent screen is excited for a longer interval than are the outer portions of the screen where the electron beam is moving faster. This results in a spurious brightness modulation of the screen, which shows up as maximum brightness in the center and tapers off linearly to the edges. Compensation of the center brightness may be accomplished by applying a saw-tooth voltage to the control grid of the cathode-ray tube, which gradually raises the grid voltage as the scanning beam moves outward on the screen. Compensation of this type is difficult to accomplish, however, because of difficulties in properly combining the correcting

voltage and the incoming picture signal voltage, which also must be applied to the control grid of the CRT.

This disadvantage alone is sufficient to rule out the circular picture and spiral scan for most practical purposes. Some further objection has been raised that the circular shape is not pleasing on aesthetic grounds. Probably the greatest single factor which caused elimination of a circular shape is that the cinema industry has standardized on a horizontal rectangle. Thus any television system which has hopes of using motion pictures as a source of program material is forced to eliminate the circular shape on the grounds that approximately 40 per cent of the picture information would be eliminated.

Two other picture shapes remain for discussion—the square and the rectangle. If a rectangle is inscribed in a circle it may be shown quite easily that its area will be maximum if it is a square; hence, of the two shapes the square gives better utilization of the circular tube face. Furthermore, scanning problems are minimized in a square picture and the simple horizontal linear scan shown in Fig. 2-1a may be used. Although this form of picture may be satisfactory for certain industrial or military uses, once again if moving pictures are to be used as program material, it must be ruled out because of loss of program material. It is of interest to note, however, that in several systems, designed specifically for naval and military use, the square picture shape has been rejected for the more common shape of the telecasting standards to be described.

The square may also be excluded on aesthetic grounds for, as Van Dyck² has pointed out, a rectangle of width-to-height ratio equal to unity is not “powerful” or “pleasing.” Whereas this last reason for excluding the square is open to criticism on subjective grounds, at least the demands on a commercial telecasting system to transmit moving pictures makes the decision final for that type of service.

The remaining picture shape to be discussed is the rectangle, and although there are an infinite number of possible width and height combinations, it seems reasonable to choose those of the motion picture industry. Considerable argument has been put forth for this choice for commercial telecasting,³ but the fact still remains that

² A. Van Dyck, “Dynamic Symmetry in Radio Design,” *Proc. IRE*, **20** (9), 1481 (September 1932).

³ D. G. Fink, Ed., *Television Standards and Practice*. New York: McGraw-Hill Book Company, Inc., 1943.

agreement with the cinema industry is advantageous to a new industry which must rely to a large degree on the relatively inexpensive and good-quality program material available on film.

It shall be assumed, then, in the rest of this book that the horizontal rectangle of width-to-height, or "aspect," ratio 4 to 3 is the standard of picture shape which gives a pleasing picture and is capable of handling all normal programming requirements.

2-2. An Estimate of the System Bandwidth

In order to discuss intelligently certain other standards we shall digress at this point to develop an approximate equation for the bandwidth required to transmit a television signal. It must be stressed that the following derivation is approximate and is based on rather crude assumptions. Its use is justified, however, because it does show exactly how bandwidth is related to the number of scanning

lines and the rate at which the pictures are transmitted.

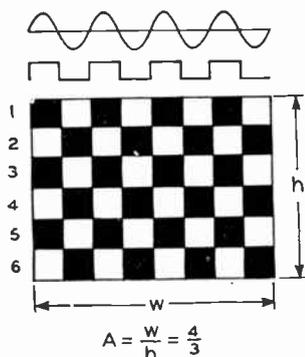


Fig. 2-4. The checkerboard pattern.

In Fig. 2-4 is shown a checkerboard pattern composed of alternate squares of black and white. Assume that this pattern is scanned by an aperture which is infinitesimally narrow but of the same height as the individual squares in the checkerboard. Assume further for the moment that the aperture is in perfect vertical alignment with the rows of squares, moves with constant velocity from left to right on row 1, returns in zero time from right to left, and then continues the scan on row 2 and so on.

Under these conditions, the generated signal would be a square wave.

Now let N_v = number of elements along a vertical line,
 N_h = number of elements along a horizontal line, and
 f_p = frame frequency or number pictures scanned per second.

Then the number of cycles of square wave generated per picture is

$$\frac{1}{2}N_vN_h$$

If it be further assumed that this square wave may be represented by its fundamental component, the frequency f_v of this component becomes

$$f_v = \frac{1}{2}N_v N_h f_p \quad (2-1)$$

and since $N_h = AN_v$ for the checkerboard, A being the aspect ratio,

$$f_v = \frac{1}{2}AN_v^2 f_p \quad (2-2)$$

Under the assumed condition that the aperture height is the same as that of the squares in the pattern, N_v is equal to n , the total number of lines in the scanning pattern, and we may rewrite (2-2)

$$f_v = \frac{1}{2}An^2 f_p \quad (2-3)$$

Now it will be realized that eq. (2-3) has been derived under an ideal condition because we assumed the scanning aperture to be in perfect vertical alignment with the rows of squares in the checkerboard pattern. Under this condition the reproduced image would be identical to the original image as shown at a' in Fig. 2-5; in general however,

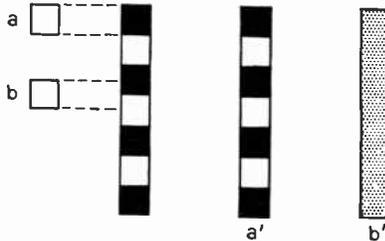


Fig. 2-5. Effect of aperture alignment on reproduced detail.

this ideal alignment will not occur and detail will be lost in the image because of it. For example, let us assume that the aperture is so much out of alignment that it just straddles two adjacent horizontal rows in the checkerboard, as shown at b in the figure. Then, since the aperture responds to the average brightness of the area which it covers, it will interpret the entire checkerboard as 50 per cent gray, and no detail would be reproduced as illustrated at b' . Between these two extremes of perfect and zero reproduction of detail lie any number of possibilities. It is evident, then, that eq. (2-3) obtains only in the ideal case. It has been customary, therefore, to introduce an empirical factor K , the utilization coefficient, into the equation to allow for

the inevitable misalignment between the aperture and scanned material. For the general case, then, the expression for the maximum video frequency should be rewritten

$$f_v = \frac{1}{2}KAN^2f_p \quad (2-4)$$

Kell, Bedford, and Trainer⁴ have recommended a value of 0.64 for the coefficient on the basis of resolution tests and this value has been widely accepted throughout the industry.

It must be realized that the checkerboard described above, where square and aperture heights are equal, represents the top limit of the system. Thus f_v in (2-4) is the maximum video frequency. For larger elements in the picture, f_v will be lower, and in the limit the average brightness of the entire field of view shows up as a d-c component. Thus the bandwidth of the system extends from zero frequency to f_v . Hence (2-4) is also an approximate relationship for the bandwidth of the system.

It must be admitted that eq. (2-4) is an approximation only. We shall see in Chapter 10 that the checkerboard pattern is a poor choice to test the resolution of a television system. Its use here is justified on the grounds that it is a pattern which has equal vertical and horizontal resolution: the ratio of elements on a horizontal line to those on a vertical line is identical to the aspect ratio. Furthermore, we assumed an infinitesimally narrow scanning aperture, a condition which cannot be fulfilled in practice. Nevertheless, eq. (2-4) is of value because it shows how the bandwidth required for transmitting a television signal is related to the number of scanning lines and to the frame frequency.

With eq. (2-4) the digression is complete: an equation has been derived giving an approximate expression for the bandwidth of a video system which is proportional to n^2 and to f_p , the picture repetition frequency. Some feeling for the order of magnitude of the quantities involved in television practice may be had by direct substitution into eq. (2-4). Thus a 500-line system sending 30 pictures per second requires an approximate video bandwidth of

$$f_v = \frac{0.64}{2} \left(\frac{4}{3} \right) (500)^2(30) = 3.2 \text{ megacycles.} \quad (2-5)$$

⁴ R. D. Kell, A. V. Bedford, and M. A. Trainer, "An Experimental Television System," *Proc. IRE*, **22**, 11, 1247 (November 1934).

Notice also that a 4:1 reduction in bandwidth may be obtained by using only 250 scanning lines, but the resolution or detail in either the horizontal or vertical direction would be reduced by two.

2-3. Speed and Direction of Scan

In choosing the horizontal rectangle as the standard picture shape it is assumed that the picture will be scanned in straight parallel lines. Four additional requirements on the scan are of concern here. First, regardless of its direction the scan must proceed with a constant linear velocity. This is necessary if the resolution is to remain constant over the entire picture area. Stated in other terms, constant linear velocity of scan permits equal fineness of detail in all parts of the picture. That this is a desirable condition may be seen from the following consideration: It is well known that the light-sensitive area of the eye, the retina, is made up of bundles of rods and cones. Since these are of small, yet finite, size the ability of the eye to resolve or separate two dots under view is limited. Experiments tend to indicate that on the average the eye can distinguish the two dots as being separate if they subtend an arc of one minute⁵ or more. In the reproduced television picture, the information is presented in the form of small areas, each of constant brightness. Hence the viewer has a tendency to adjust his viewing distance until adjacent areas blend into a continuous picture. This distance will depend, of course, upon the element separation in the picture or upon its resolution. Thus if different areas in the picture have different resolutions, the viewer tends to increase his viewing distance to render satisfactory the poorer parts of the image. This results in loss of detail in the higher quality regions and so the "best" parts of the picture are wasted in a sense. This sort of reasoning justifies the demand for constant linear velocity of scan.

Consider the second additional requirement on linear scan, namely, that it shall be unidirectional. More specifically it is meant here that picture data shall be presented only as the aperture moves in one direction and not in the reverse direction also. The reason for this is primarily one of circuitry. For example, if a device employing electrostatic deflection of the electron beam is used, a more complex deflection voltage is required for a bidirectional scan. Figure 2-6

⁵ E. W. Engstrom, "A Study of Television Image Characteristics," *Proc. IRE*, 21, 12, 1631 (December 1933).

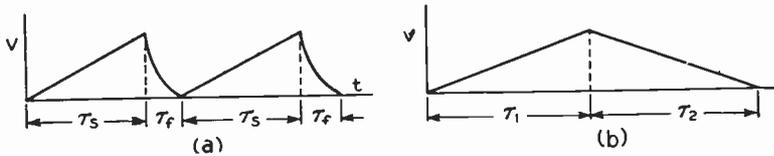


Fig. 2-6. (a) Deflection voltage for unidirectional scan. (b) Deflection voltage for bidirectional scan.

shows the deflection voltages required for unidirectional and bidirectional types of scan. At *a* the former is shown. Picture data is presented only during the scan intervals τ_s , τ_f , the flyback interval, being used to return the aperture to its starting position. During τ_s , the voltage must build up linearly with time to provide constant resolution. During τ_f , on the other hand, the voltage may follow *any* mathematical function of time just so long as it reaches its initial value at the end of τ_f . At *b* the bidirectional scan requirement is shown. During both τ_1 and τ_2 the deflection voltage must change linearly with time, once increasing and then decreasing, i.e., the deflection voltage must be an even function of time. It may be seen that the requirements in the former case are less severe. Furthermore considerable difficulty is encountered in building circuits to meet the requirements of linearity during τ_2 in curve *b*. Generation of the voltage shown at *a* may be accomplished by the use of a series *R* and *C* circuit and is discussed in Chapter 4. A further disadvantage is present with the bidirectional scan in that an overlapping of adjacent lines, which have finite width, occurs at the edges of the scanned pattern.

The third consideration relative to the linear scan is its direction. Clarification is required here because actually two directions of scan are involved. A moment's reflection will show that if a moving aperture is to cover the entire area of the rectangular image, it must move in two directions. The assumption has been made in the discussion thus far that the aperture moves with constant velocity along a "line," returns to the picture edge corresponding to its initial position and scans another "line" parallel to the first, separated from it by the dimension of the aperture normal to the "line." The question here, then, is shall these scanning lines be horizontal or vertical as shown in Fig. 2-7*a* and *b*, respectively? Actually the horizontal direction has been chosen as standard, first, because it conforms to the direction of writing English, and, second, because there is a feeling

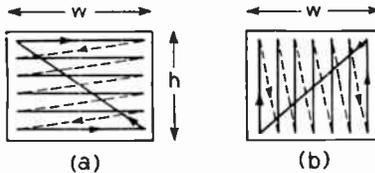


Fig. 2-7. Types of scan. (a) Horizontal. (b) Vertical.

in some quarters that the horizontal scan is better able to reproduce motion, which is predominantly horizontal rather than vertical, in typical televised scenes.

The fourth consideration pertaining to the linear scan is its sense, i.e., shall the aperture during the scan interval move from left to right or vice versa. No technical advantage is at stake in the selection to be made here, and again the "natural" choice seems to be to conform to the direction of writing English. Hence the standardized scan senses are from left to right and from top to bottom of an upright reproduction of the image televised.⁶ The last clause may seem to be superfluous but it is necessary because in the pickup end of the system an optical image of the scene is thrown on a photosensitive surface by a lens. This image is inverted relative to the physical world and, hence, must be scanned from right to left and from bottom to top in order that scanning conform to the standard.

Two other questions inevitably arise relative to scan at this point. The first is: What shall be the flyback ratio defined by

$$p = \text{flyback ratio} = \frac{\tau_f}{\tau_s} = \frac{\text{retrace or flyback time}}{\text{trace or scan time}} \quad (2-6)$$

where τ_f and τ_s are the intervals shown in Fig. 2-6a? The second is: What shall be the number of lines, which in the final analysis determines the resolution of the entire system? Both of these questions more properly relate to the requirements of the particular television system under consideration and as such are treated in another section. For the moment we may at least say that p should be as small as practicable in order that a minimum of time be lost from the scanning process.

⁶ It is interesting to note that Fink reporting on the deliberations of the National Television Systems Committee, which set up the proposed standards for commercial television service for the Federal Communications Commission, states that the sense of scan, in contrast to several other standards, was adopted unanimously. D. G. Fink, *op. cit.*, p. 195.

These scanning standards may be summarized as follows: The active picture area shall be scanned in parallel lines, horizontal (or nearly so) and from left to right. The sequence of lines will be from top to bottom of an upright picture image. The scan in both horizontal and vertical directions shall be at a constant speed.

2-4. The Frame Frequency or Picture Repetition Rate^{7,8}

Thus far in the discussion it has been shown that the picture information of the television picture is built up element by element until the entire field of view has been supplied with brightness information. Such a single scan of the entire picture area may be called a "picture" or a "frame." The latter notation, which is more commonly used, is derived from moving picture terminology, wherein a single complete picture in the series of pictures on the film is termed a frame. It has been shown further that these pictures or frames are then presented in sequence at a rate sufficiently high, such that the resulting picture satisfies the requirement for fusion of all the elements into a smooth picture, for reconstituting motion in the original scene, and for unnoticeable flicker. It remains at this point to investigate the factors that control these considerations.

At the outset it must be stressed that the disagreement between standards for industrial or military applications on the one hand, and for commercial telecasting on the other may be considerable. In setting up the requirements for the commercial system⁹ the criterion for "satisfactory operation" is the performance of the 16-millimeter sound moving picture system. In the former applications performance below this par may be tolerated and, in fact, may be mandatory in order that other requirements such as system simplicity, light-weight components, and minimum cost (where equipment life may be relatively short)¹⁰ may be met.

⁷ E. W. Engstrom, *op. cit.*

⁸ E. W. Engstrom, "A Study of Television Image Characteristics," Part II. *Proc. IRE*, **23**, 295 (April 1935).

⁹ It should be mentioned that the remarks of this section apply to the standard commercial black-and-white television system. Developments have been made in color television systems and in the so-called "dot-interlace" system of transmission which permit frame frequencies considerably lower than 24 per second to be used without adverse effect on the reproduced picture.

¹⁰ In the "Roc Bird" a miniature pickup and transmitting system is mounted in a bomb. The entire airborne system weighs but 50 pounds and is blown up at the end of the flight. See R. D. Kell, and G. C. Sziklai, "Miniature Airborne Television Equipment," *R.C.A. Rev.*, **VIII**, 3 (September 1946).

Thus, since the 16-mm movie uses 24 frames per second as the frame, or picture repetition frequency, f_p , this value sets the lower limit for the commercial telecasting system. Hence, $f_p \geq 24$ frames per second and it need only be determined what the actual value shall be. For the more general case, however, some investigation must be made as to the factors which place a lower limit on f_p . From the viewpoint of reproduction of motion, it has been determined that the eye sees no discontinuity of motion if the frame frequency exceeds 16 pictures per second. This low value of f_p is generally unsatisfactory because of the accompanying flicker. Hence we now consider flicker which is the first of these factors to be discussed.

If a beam of light of constant intensity be interrupted at a variable rate, it may be shown experimentally that above a certain minimum interruption frequency the eye responds as if it were receiving the light stimulus continuously. Under these conditions the separate stimuli arrive at a rate which is high enough for the eye to integrate the intensity over a complete cycle, and the effect is just as if a source of constant intensity were shining continuously. At frequencies below this critical value the eye is unable to so integrate the information; it sees the instantaneous variations. In this case flicker is said to be present and the frequency at which crossover between the two effects occurs is known as the critical flicker frequency.

If, now, the same experiment be performed at different levels of source intensity, it is found that the critical flicker frequency varies with the logarithm of the illumination intensity. This is a statement of the Ferry-Porter law and is illustrated in Fig. 2-8. Data for these curves were taken by Engstrom by having several observers view the image of a light source focused on a screen, the light beam being chopped by the rotating disk shown in the diagram.

If these results are to be related to television the immediate question is this: To what angle of disk opening, b , does the television system correspond? In the absence of a complete set of standards this question cannot be answered rigorously at this time. Some order-of-magnitude conclusions may be drawn, however. Consideration of the scanning pattern previously proposed shows that after a complete frame has been scanned, the scanning aperture must return to its starting point at the top left-hand corner of the picture. A finite time is required to accomplish this retrace, or flyback, in the vertical direction. From considerations that will become apparent

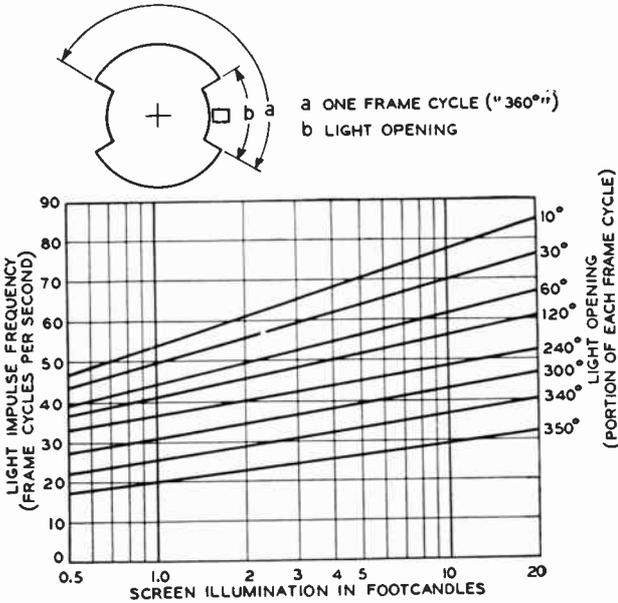


Fig. 2-8. Critical flicker frequency measured with a slotted disk.
(Courtesy of *Proc. IRE.*)

later in our work, this vertical flyback ratio, p_v , which is defined in (2-6) is generally of the order

$$p_v = 0.05$$

That is to say that the ratio of retrace to trace time is 0.05. Since the electron beam is "blacked out" during retrace to prevent contamination of the picture, this means that the eye sees light for approximately 95 per cent of the frame interval which corresponds to a disk opening of approximately 340° in Fig. 2-8. It would appear, then, that a 340° opening curve could be used for determining the critical flicker frequency. As a practical matter this guess does not correspond to the facts because the type of flicker generated in Engstrom's test does not correspond to that present in the television system. Despite this fact, which will be expanded presently, the curves of Fig. 2-8 do show, first, that the Ferry-Porter law obtains and, second, that if the screen viewed during the test were located in an ambient light level other than zero, the same results

would hold true, provided the abscissas were interpreted as apparent illumination, i.e., the difference between the illumination of the test screen and the ambient level of illumination.

In the preceding paragraph it was stated that the flicker present in a television image and in the test previously described are different. This is easy to see when the two systems are compared even in the most elementary respects. First, in the test for almost the entire slot width b , the whole screen is illuminated uniformly. On a CRT screen, on the other hand, only a few elements are illuminated at a time as the spot traces out its pattern of straight, parallel lines. Note the difference then: Entire screen illumination during the unblanked scan time as compared to partial screen illumination—with the illuminated area moving so as to cover the entire field of view in the unblanked scan time.

As might be expected, in the latter case where the conditions are more severe, a higher value of critical flicker frequency obtains. Thus a curve corresponding to a light opening of considerably less than 340° in Fig. 2-8 more closely approximates the actual conditions that exist in a television system.

An additional factor aids the retina in integrating the image over the scan period and that is that the illumination of an excited phosphor on a CRT tends to persist; it decays exponentially after the excitation is removed rather than dropping to zero instantaneously. Realizing this, Engstrom has conducted additional tests to measure the flicker threshold on an actual cathode-ray tube. The results of this test are given in Fig. 2-9, which shows critical flicker frequency vs.

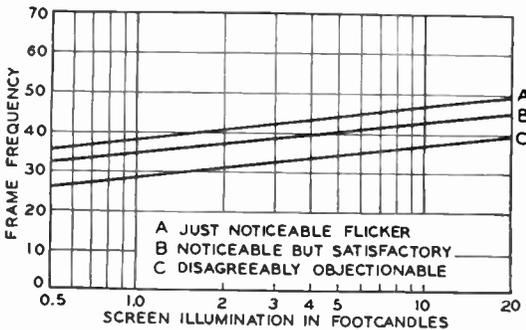


Fig. 2-9. Critical flicker frequency for a cathode-ray tube. (Courtesy of *Proc. IRE.*)

apparent illumination on the kinescope. The persistence characteristic of the cathode-ray tube used for the test is shown in Fig. 2-10.

In applying Fig. 2-9 to any television system, it must be realized that any curve is not absolute; it may be displaced vertically, depending on the persistence characteristic of the particular fluorescent material used on the viewing screen. The curves of Fig. 2-9 and 2-10 are typical, however, for the P4 screens commonly used for television service.

We are in a position now to decide on the value of flicker frequency for the television system. The choice is inevitably dependent upon the level of apparent illumination required. In the lack of a better criterion we may turn to the 16-millimeter sound moving picture system for a guide. Goldmark¹¹ has presented the following data:

| Levels of Illumination for 16-millimeter Movies | |
|---|--|
| Condition | Apparent illumination on 3-ft wide screen |
| No film in gate | 10 foot-candles |
| Clear film in gate | 8 foot-candles |
| High lights in color film | 4 foot-candles |

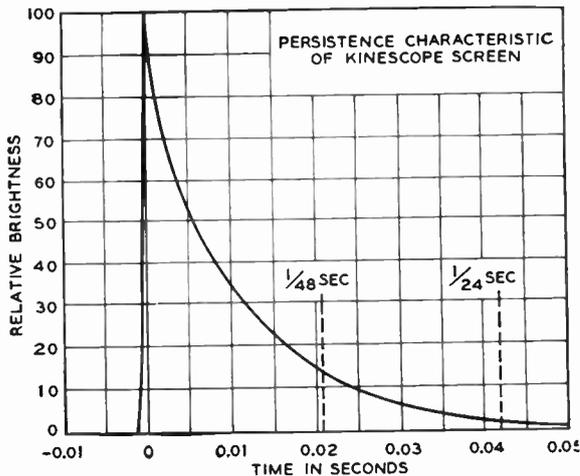


Fig. 2-10. Persistence characteristic of the cathode-ray tube screen used in obtaining the data for Fig. 2-9. (Courtesy of *Proc. IRE.*)

¹¹ P. C. Goldmark, First Lecture in the Series Modern Television, *New York Section IRE* (Fall 1948).

Using these figures as representative, reference to Fig. 2-9 shows that a flicker frequency of some 45 "flicks" per second are needed. Yet the 16-millimeter movie system displays only 24 frames per second. The discrepancy between these two figures of a factor of about 2 to 1 may be explained by a simple device that is used in the movie projectors.

In the first place, movie flicker is of the type given by the slotted-disk test described previously. Thus a lower value than that for the television system is permissible, but at any rate, it will be greater than the 24 frames per second displayed by the 16-millimeter sound movie system. As we have seen, this figure is perfectly satisfactory for reconstituting motion in the scene and presents enough footage of film for satisfactory sound reproduction. Thus raising the frame frequency to meet flicker requirements would result in the use of greater lengths of film with little improvement of the picture (except, of course, with regard to flicker). To obviate the need for this waste of film, a rotating slotted disk, not unlike that of Fig. 2-8 is incorporated in the projection system. The bislotted disk revolves once in the time that each frame is in the film gate. Thus, light is projected twice onto the screen through each film frame, and the 24 frames per second on the film appear as 48 "flicks" per second on the screen. In this manner the flicker frequency for the cinema is raised well above the threshold value.

But the objection might well be raised that all this talk is about a moving picture system; how does it tie in with television? Following this line of thought, it seems that it is only necessary to refer to Fig. 2-9 to determine the critical flicker frequency corresponding to the apparent illumination required by the given application, and to choose the next higher, convenient value of frequency as the frame frequency for the television system. The objection to this procedure is furnished by reference to equation (2-4) which shows that the bandwidth of the video system is directly proportional to f_p , the frame frequency. It is in the interest of bandwidth economy, then, to choose a frame frequency as low as possible. A compromise is indicated: f_p must be chosen to effect a balance between video bandwidth on the one hand, and unnoticeable flicker on the other. The compromise depends upon the requirements of the particular television system being designed.

At this point the 2 to 1 flicker-frame relationship described for the

moving picture projector seems intriguing. If a similar device could be adapted to the television system, a 2 to 1 saving in video bandwidth could be effected with no degradation of the flicker characteristic. Because of the sequential manner in which the picture information is presented on the CRT it is impossible to incorporate a rotating slotted disk as described above. It remains, then, to see if some alternate mechanism can be proposed which will achieve the same result.

In 1931 U. A. Sanabria¹² proposed a system of interlaced scanning that effectively gives two vertical flyback and blanking periods during the presentation of a single frame which accomplishes the same overall result. Figure 2-11 shows a number of lines covering the picture

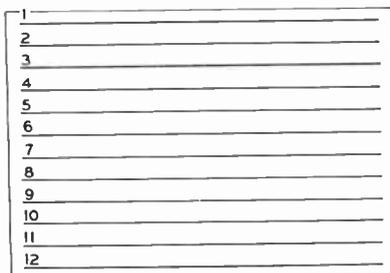


Fig. 2-11. Position of an arbitrary number of scanning lines. Progressive scan results if the lines are scanned in order from top to bottom. The sequence 1, 3, 5, 7, 9, 11, 2, 4, 6, 8, 10, 12 is used in interlaced scanning.

area. These lines are scanned in a "progressive" scan of the type already described, in the order of the numbers shown, i.e., 1, 2, 3, . . . , from top to bottom. At the end of line 12, the spot retraces vertically during the vertical blanking interval back to its starting point on line 1. As the name implies, in the interlaced system of scanning, two sets of lines, which interlace each other, are scanned alternately. Thus, for example, in Fig. 2-11 the scanning sequence would be 1, 3, 5, 7, 9, 11 retrace to the beginning of line 2 with the beam blanked out, then lines 2, 4, 6, 8, 10, 12 and retrace with the beam blanked out to the beginning of line 1.

If this interlaced scanning sequence is carried out, each picture is scanned in two interlaced halves composed of alternate lines with two

¹² U. A. Sanabria, U. S. Patent #1805848, March 1931.

blanking intervals, and the desired flicker-to-frame frequency ratio of 2 to 1 is accomplished. Each half-picture which is composed of one set of alternate lines is termed a "field" and the field frequency, f_f , is seen to be twice the frame frequency, where a 2 to 1 interlace of 2 fields to a frame is used.

Use of this interlaced scanning pattern imposes more severe restrictions on the scanning systems and, as will be seen in Chapter 3, demands that the sum of the active scanning lines plus the inactive lines lost during the vertical retrace intervals be an odd number.

A critical examination of the flicker present in this interlaced scanning system shows that actually two flicker effects are present. The first or "over-all effect" is the result of each line flickering at the rate of once per frame. The second is the "interline effect," caused by adjacent lines flickering in a time relationship differing by one field period. These two effects were also the subject of a series of tests performed by Engstrom. Again resorting to tests employing a rotating disk slotted in a different manner (Fig. 2-12), he was able to

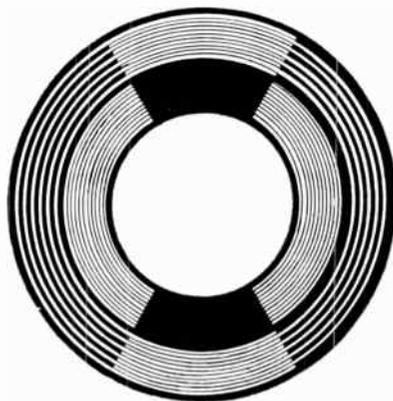


Fig. 2-12. Special disk used by Engstrom for interlaced scanning flicker tests. (Courtesy of *Proc. IRE.*)

draw some general conclusions. The interlaced scan pattern does minimize flicker, but the viewing distance must be increased over that required with progressive scan in order to minimize the interline flicker. Once again, the hangover or persistence characteristic of the CRT phosphor aids in creating the illusion of a fused picture and the interlaced system has been adopted as the commercial standard in

the United States of America. It must be stressed, however, that it *need* not be used; the interlaced scan pattern may be used where the 2 to 1 saving in bandwidth for a given flicker and illumination requirement is desirable at the expense of a more complex system.¹³

Summarizing, it appears that the choice of frame frequency is contingent upon a number of factors, all peculiar to the particular television system under consideration; screen illumination, flicker requirements, bandwidth, and permissible degree of system complexity. We digress on additional requirements peculiar to commercial telecasting.

It has been pointed out repeatedly that agreement between cinema and commercial television standards wherever possible is highly desirable from the programming viewpoint. It seems reasonable, then, to standardize the frame frequency at 24 frames per second which is common to both 16-mm and 35-millimeter moving picture systems. This with a 2 to 1 interlaced scan would place the flicker frequency at 48 flicks per second. Reference to Fig. 2-9 shows that this choice would permit operation at levels of apparent illumination up to some 10 foot-candles. Unfortunately other factors must be taken into consideration. Foremost among these is the effect of power supply ripple frequency on the scanning pattern.¹⁴

Probably one of the most objectionable defects in a televised image is unsteadiness. On this basis, we may investigate some of the undesirable effects caused by spurious modulation of the scanning pattern produced by power supply ripple voltages. Since commercial standards are based on an interlaced scanning geometry, let us determine what effect, if any, these ripple voltages will have on an electrostatic deflection system operating at 24 frames, 48 fields per second. As a specific example, we shall consider what happens to the horizontal scanning lines.

It must be realized at the outset that if ripple components appear on the horizontal deflecting plates of the cathode-ray tube, spurious sidewise displacements of the scanning lines will occur. Assume for

¹³ The military systems previously mentioned illustrate this point. In the Ring system where the pickup equipment is carried in a bomber, a 2 to 1 interlace is used. The Bloek systems which enjoy a comparatively short life in a guided missile use the simpler progressive scan. See V. K. Zworykin, "Flying Torpedo with an Electric Eye," *R.C.A. Rev.*, VII, 3 (September 1946).

¹⁴ R. D. Kell, A. V. Bedford, and M. A. Trainer, "Scanning Sequence and Repetition Rate of Television Images," *Proc. IRE*, 24 (4), 559 (April 1936).

illustration that a half-wave rectifier operating from a 60-cycle line is the voltage source for the circuit. Then the fundamental ripple component will modulate the lateral positions of the scanning lines at a frequency of 60 cycles per second. Since in the 24/48 frame-field system proposed each field has a duration V of $1/48$ second, the number of spurious ripple cycles per field will be

$$\text{No. of ripples/field} = \frac{60}{48} = 1.25 \text{ cycles/field} \quad (2-7)$$

The result of this spurious deflection of the scan pattern is shown in Fig. 2-13, where the lines of field 1' are displaced vertically so that

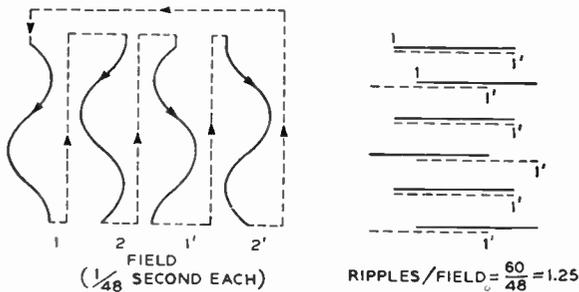


Fig. 2-13. Spurious line deflection due to 60-cycle ripple in a 24/48 interlaced scanning system. Any line moves horizontally between alternate fields. (After Kell, Bedford, and Trainer.)

their positions may be seen. Since the field and ripple frequencies are not integrally related, the envelope of the line positions for field 1 and 1' (and for fields 2 and 2' for that matter) are not the same. This means that a given line in alternate fields will occupy different positions causing an objectionable unsteadiness in the final image. Just how objectionable this effect will be depends, of course, upon the magnitude of the spurious deflection. A simple calculation will show that if the envelope amplitude is 25 per cent of the interline spacing, the maximum change of position of a single line is approximately 100 mils on a 10-in. CRT. This value may cause considerable degradation of the picture. A quick estimate of the situation shows that remedial measures may follow two avenues. First, the ripple voltage may be reduced at the expense of adding more filtering to the

power supply, or second, the frame and field frequencies may be increased so that they are related integrally to the power line frequency. An increase rather than decrease in these frequencies is necessary in order that flicker requirements may be met, but the price to be paid is an increase in video bandwidth. Other factors tend to favor a choice of the latter alternative and in telecasting the 30 frame-60 field per second values are chosen as standards.

If ripple voltage is present in the display equipment, it will probably also appear on the control grid of the kinescope. Its effect here will be to produce dark horizontal bars across the face of the televised picture. Furthermore in the 24/48 frequency system, these bars will move downwards across the face of the picture. A change to a 30 frame-60 field system will not eliminate the bars but at least they will remain stationary.

Since line position modulation and the bar pattern described above may be minimized by additional filtering and shielding, let us next consider an effect completely independent of ripple within the television equipment. The tendency over the post-war period has been to a larger number of remote pickups at points other than the television studio. At these remote locations direct current is virtually never encountered. Since all light sources operating from a-c sources have a stroboscopic effect, hum signals are picked up by the television camera tube. If of sufficient amplitude, these may produce dark bar modulation which moves vertically on the screen when the light source and field frequencies differ. With the 30/60 frequency standards the dark bars remain stationary.

It must be realized that only a few of these ripple effects have been mentioned. A more complete analysis such as that presented by Kell, Bedford, and Trainer must take additional account of spurious *vertical* deflection of the scan lines or "pairing," spurious effects as a result of higher ripple frequencies, and the like. All these considerations point to the adoption of a 30 frame-60 field system as opposed to the 24/48 system of the cinema industry. The disagreement between these two systems is unfortunate for it requires a conversion of the 24 movies frames per second into 30 frames per second for the television system when film is televised. This conversion may not be accomplished by simply increasing the film speed in the ratio of 5 to 4, for motion would be speeded up with the humorous effect

of films of the 1920's run in modern projectors.¹⁵ Since the sound track is on the same film as are the pictures, the increased speed would cause improper sound reproduction. How the conversion of film to television standards is accomplished without speed-up of the film is covered in Chapter 17.

Care should be taken to realize that adoption of the 30 frames and 60 fields per second does not eliminate the various ripple effects discussed, but it does minimize them. Because of the 1 to 1 relationship between supply and field frequencies, the envelope of the spurious line deflection has only 1 cycle per frame and the result is shown in Fig. 2-14. Notice that here the field-to-field motion of any given

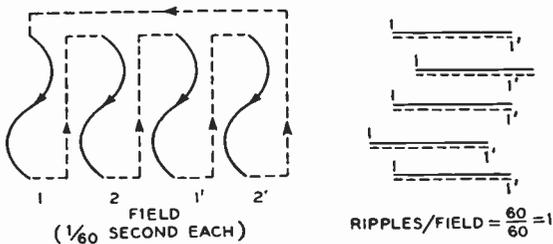


Fig. 2-14. Spurious line deflection due to 60-cycle ripple in a 30/60 interlaced scanning system. Field-to-field motion of the lines is eliminated. (After Kell, Bedford, and Trainer.)

line in the scan pattern is eliminated. Furthermore, the maximum peak-to-peak motion of a line is replaced by a fixed peak displacement from the normal of about 50 mils for the same conditions assumed in the previous example.

Whereas the discussion above was carried out for the commercial system which employs interlaced scan, it must be emphasized that a similar treatment will show that an integral relationship between power line and frame frequencies in a progressive scan results in the same sort of improvement.

In summary a balance sheet may be prepared comparing the two systems of standards.

¹⁵ The old silent films were taken at 16 frames per second. The playback speed was from 16 to 30 frames per second depending upon the whim of the projectionist. The advent of the sound track made rigid adherence to a constant speed of 24 frames per second (90 feet per minute) mandatory.

| | <i>24/48 system</i> | <i>30/60 system</i> |
|---------------------------------|---------------------|-----------------------|
| Power line frequency | 60 cps | 60 cps |
| Spurious line deflection | Yes | Yes |
| Spurious horizontal line motion | Yes | No |
| Spurious vertical deflection | Yes | Yes |
| Spurious vertical line motion | Yes | No |
| Black bar pattern | Yes | Yes |
| Motion of bar pattern | Yes | No |
| Flicker | Satisfactory | Satisfactory |
| Bandwidth | Δf | $\frac{5}{4}\Delta f$ |

In retrospect it is interesting to consider the verisimilitude of the assumption of a 60-cycle power line frequency which was made in the foregoing discussion. This is standard for most of the United States with two principal exceptions. Generation of power in the Niagara Falls section is at 25 cycles per second. St. Louis and certain parts of Metropolitan New York enjoy this distinction also, but in the latter region the service is being converted to 60 cycles per second, except for industrial loads. In large areas of California where the frequency is locked in with adjacent 60 cycle areas 50 cycles has been standard. In none of these regions would adoption of the 24/48 standards give improved service over the 30/60 system.

CHAPTER 3

SCANNING METHODS

We have seen in the previous sections that the scanning process is basic in television, that picture transmission requires that some sort of sampling aperture be made to move across the entire picture area in some orderly, predetermined fashion. Furthermore, in setting up means of determining the picture standards we confined the discussion to receiving-end terms where a cathode-ray tube was assumed to be the display device. This led to the general notion that scanning involves the deflection of an electron beam. The immediate question, then, is whether or not both ends of the system scan by the same method, thus making the previous discussion of picture standards general. Now it must be realized that the present day all-electronic system evolved in more or less definite steps. Let us briefly review some of the major scanning methods that have been used for television transmission over a period of the last twenty years.

3-1. Mechanical Scanning

In 1884 P. Nipkow proposed a mechanical means of scanning an image which involved the use of a rotating disk. The development of this scanning device may be described on the following basis. If an area wh is to be scanned in its entirety by a single aperture, that aperture must be made to move in two directions. This bidirectional motion of a single aperture may be replaced, however, by motion in a single direction of an array of apertures which are displaced from each other in two dimensions. To make this more clear, consider Fig. 3-1. If the array of apertures, n_a in number, is moved downward at a constant velocity, the entire area wh will be scanned in a series of vertical lines. Certain restrictions on the dimensions of the scanning array must be made so that only one of the several apertures is on the picture area at any given instant. From the diagram it can be seen that the vertical inter-aperture distance must be equal to or greater than h , the picture height, and if adjacent lines are to touch each

other, the width, a , of the aperture must be w/n_a , where w is the picture width and n_a is the number of identical apertures in the scanning array.

It is apparent that if the picture area is to be scanned continuously,

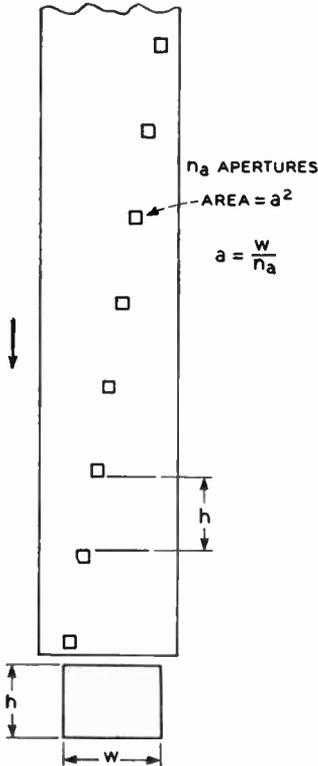


Fig. 3-1. A single aperture that moves in two directions may be replaced by an array of apertures that move in one direction.

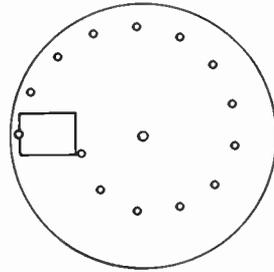


Fig. 3-2. The Nipkow disk.

an infinite number of such scanning arrays would be required. Clearly, a compromise is indicated. One such compromise is this: let the array of n_a apertures be closed back on itself to form a spiral and let the vertical motion of the array be replaced by a rotational one. This substitution results in the Nipkow disk, with n_a apertures arranged in a spiral, and this disk is to be rotated at a constant angular speed. Notice that the only change in scanning different from that of the original system is that the lines are sections of circles whose centers are located at the shaft.

Given the Nipkow disk as a scanner a simple mechanical television system may be set up. Let a phototube be placed behind the disk. Then, because of the scanning action of the pickup disk, the phototube "sees" only one element of the object at a time. The amplified output of the phototube is delivered by the link to a neon tube opera-

ting in the region of abnormal glow, causing the total light intensity to be proportional to the object brightness. The brightness information is restored to its proper position in space by a second Nipkow disk operating in exact synchronism with the first. The simple circuit which is shown in Fig. 3-3 represents one of the earliest steps in the development of television systems.

The chief disadvantages of the double Nipkow disk apparatus are more or less self-evident and are listed below.

(1) Picture size is limited by the maximum allowable disk dimensions and the size limitations of the playback discharge tube. In the early stages of the art a 1-in. square image was typical.

(2) n_a , the number of lines, and hence the resolution, is limited by the scanning-disk size.

(3) The reproducing tube does not give off white light causing the gray scale to be replaced by shades of some monochrome.

(4) The system scanning speed is limited by mechanical considerations, e.g., by motor rotation.

(5) The efficiency of the optical systems is low. In illustration of this fifth point it can be seen that despite the high illumination of the object being televised, the instantaneous output of the phototube is proportional only to the illumination of the very small area framed by the aperture.

The next principal stage of development incorporated the "flying-spot" scan at the pickup end to increase the optical efficiency. It may be noted that in Fig. 3-3 the televised object is under constant

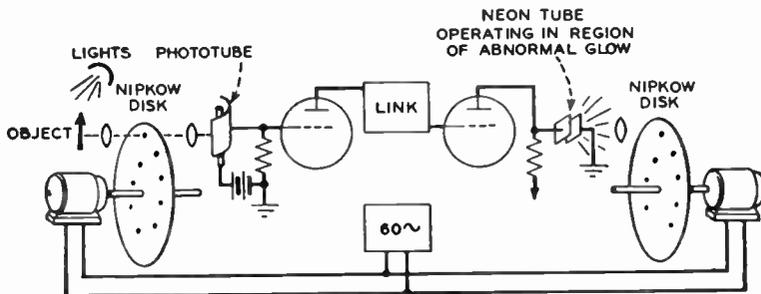


Fig. 3-3. A simple television system employing mechanical scanning.

high-level illumination, and the phototube effectively scans across the area. In the flying-spot scanner, on the other hand, the televised object is in darkness and is viewed by a battery of phototubes. Illumination in the form of a moving light spot of high intensity is then caused to scan across the object by a scanning disk. As a result of concentrating all available light on a single picture element and using several, rather than a single, phototubes for pickup, a greater output is obtained than in the previously described system. Reference to Fig. 3-4 shows that the prime difference between the two systems lies in the placement of the scanning disk relative to the lighting source, the object, and the phototube. In the earlier system, the scanning disk lies between the illuminated object and the transducer. In the flying-spot scanner, the disk is between the light source and the object. In the interests of even greater optical ef-

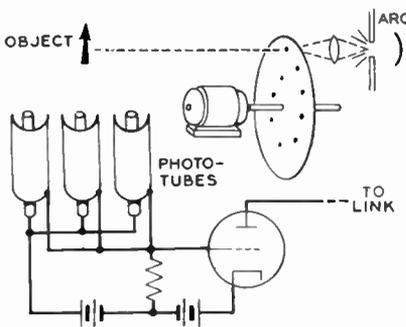


Fig. 3-4. Flying spot scanning.

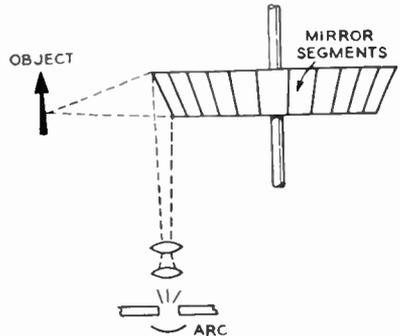


Fig. 3-5. Mirror scanner.

iciency, the Nipkow disk was later replaced by a revolving drum with lenses over the individual apertures.

It must be realized that in this same period of development several other mechanical systems of scanning at both pickup and reproducing ends were developed. As an example, a rotating mirror scanner is shown in Fig. 3-5. Here the spot of light is made to scan the image by rotating a drum of tilted mirrors, each having a tilt which gives a scanning line at a height different from all the others. For our discussion little advantage is to be gained by reviewing the details of each system, for all agree in the essentials. It should be mentioned,

however, that the Scophony^{1,2,3} system, involving double rotating mirrors whose axes are normal, was used—though primarily in Europe—for reproducing the image. Its chief advantage lay in its ability to project relatively large images at the receiving end. The principal components of the Scophony apparatus are diagrammed in Fig. 14-35.

3-2. Development of Electronic Scanning

In the early 1920's interest in electronic, as opposed to mechanical, scanning began to take form. Although it is beyond the scope of our work to trace the development of the all-electronic television system, it is of interest to note a few highlights in two major branches of the work which were finally combined in 1934.

At the beginning of that era, work was carried out at the Radio Corporation of America toward the development of a system that employed the cathode-ray tube as the reproducing element. 1929 marked one of the first receptions of a television image which was reproduced electronically. During the same interval research on an electronic pickup device was also under way but it lagged behind the work on the cathode-ray tube. While Zworykin had conceived of the principle of the iconoscope camera tube, it was not until 1932 that it was developed sufficiently for reporting to the industry. At the time of the 1929 test a mechanical galvanometer was used as the transmitter.⁴

During the same decade members of the staff at Purdue University had been doing work in the design and operation of cathode-ray tubes.⁵ Cognizance of this work caused the Grigsby-Grunow Company, which was manufacturing Majestic radio receivers at that time, to contact the University in 1929 with the result that a joint

¹ H. W. Lee, "The Scophony Television Receiver." *Nature*, **142**, 59 (July 9, 1938).

² Scophony Television, *Electronics*, **9** (3), 30 (March 1936).

³ D. M. Robinson, "The Supersonic Light Control and Its Application to Television with Special Reference to the Scophony Television Receiver." *Proc. IRE*, **27**, 483 (August 1939).

⁴ V. K. Zworykin, "Television with Cathode-Ray Tube for Receiver." *Radio Engineering* (December 1929).

⁵ C. F. Harding, R. H. George, and H. J. Heim, "The Purdue University Experimental Television System." *Engineering Bulletin*, R.S. No. 65, Engineering Experiment Station, Purdue University, 1939.

research project was established, whose purpose was the development of a television receiver employing a cathode-ray tube for scanning and display.

In retrospect it may seem odd that the use of the cathode-ray tube for display of television images was so late in arriving on the scene of system development because basically it is a comparatively old device in the field of electronics. Developed originally by Braun⁶ in 1897 the tube employed a cold cathode, and deflection of the ion beam was accomplished by the use of magnetic fields. Since modern high-vacuum techniques were virtually unknown at that time, the envelope contained air at low pressure. In its early form the cathode-ray tube fell far short of the requirements of the television system. It remained for subsequent improvements, such as the introduction of the coated, thermionic cathode developed by Wehnelt in 1905, better evacuation and manufacturing techniques, and improved phosphors, to make the tube of use in television.

Under the aegis of the Grigsby-Grunow-Purdue agreement, work went forward in the development of the receiver. In its early form, a cathode-ray tube employing electrostatic deflection of the beam was used and continual pumping of the envelope was required to

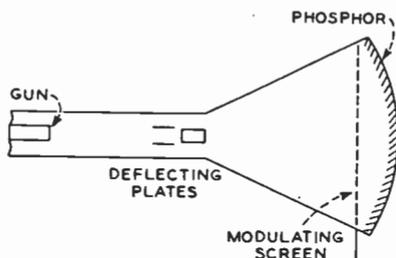


Fig. 3-6. Change-of-velocity modulated cathode-ray tube.

maintain a vacuum. Essentially the reproducing system was quite similar to that of the present day with the prime difference in the manner of applying the video information to the cathode-ray tube. In addition to the conventional elements of electron gun and fluorescent screen, the early Purdue tube incorporated a fine mesh of

⁶ See, for example, A. I. Albert, *Fundamental Electronics and Vacuum Tubes*. New York: The Macmillan Company, 1947.

nickel gauze just behind the phosphor. Application of the video voltage to this screen caused a change in velocity of the electrons in the beam at their point of impact on the phosphor. Since the excitation and intensity of glow of the phosphor are dependent upon the kinetic energy of the impinging particles, the video voltage could thus control the light output from the fluorescent screen. Deflection of the scanning beam was accomplished by the application of suitable deflection voltages, generated in grid glow tube oscillators, to the pairs of deflection plates.

At about the same time steps were taken to develop an all-electronic pickup device for test purposes which pointed the way to an entirely electronic system as contrasted to the prototype systems, which relied upon a rotating disk or drum to carry out the scanning process.

The pickup tube developed was a static image generator in that it was capable of generating only a single pattern of a silhouette form. Thus, although it was not a television "camera" in the usual sense, it did provide a satisfactory test signal and pointed the way to the later development of more modern static image generators, such as the monoscope. This image generator was simple in both construction and principle of operation. A metal silhouette cut from sheet metal was mounted in an evacuated envelope so that it could be scanned by an electron beam. Behind the silhouette was a collecting plate, which caught those electrons which passed by or through holes in the silhouette. The output was the current obtained from this collecting electrode. Pending the construction of a radio transmitter for test purposes, the system was operated as a closed one, using wires to connect the image generator and cathode-ray tube.

During the next three-year interval further research was conducted to improve the cathode-ray tube, which finally resulted in a glass envelope tube with an artificial phosphor screen and which employed the now standard method of controlling the screen brightness of applying the video signal to a control grid mounted in the electron-gun assembly.

A receiver incorporating this new playback tube and a Nipkow disk scanner pickup system served the basis of several "on the air" transmitter tests which, at that time, were considered satisfactory and which provided reception at distances up to 150 miles.

In 1934 support of and patent rights for the Purdue project were secured by the Radio Corporation of America. This brought together

strong forces for the development of a practical all-electronic system for, as we have seen, the research organization of RCA Victor had been doing parallel work on a system based on the cathode-ray tube.^{7,8} At the same time a number of systems for scanning moving picture film were developed.^{9,10}

The year 1934 also saw the development of the image dissector and the iconoscope, both of which incorporated electronic scanning of some sort and the electro-optical transducer into a single unit or camera tube. Subsequent developments along this line have resulted in the image iconoscope (1938), the orthiconoscope (1939), and the image orthiconoscope (1946). All these tubes use a moving electron beam as the scanning element.

Thus we see that the assumption of electronic scanning which was used in the last chapter is perfectly general: scanning is accomplished at both ends of the system by deflecting a beam of electrons. Since this concept is basic in modern television systems, we next investigate the means available for causing this deflection.

ELECTRONIC SCANNING

An electron beam such as used in the various cathode-ray devices common to television requires motion in three mutually perpendicular directions. For example, in the kinescope any beam electron must move first from gun to screen, and then must be capable of being deflected either up and down or horizontally across the screen to produce the scan pattern. Thus we must consider means of accomplishing these different types of motions. In general two fields are required, one to produce the toward-screen motion, and the second to produce the across-screen motion. Three types of combinations of fields are used for this:¹¹ (a) two normal electric fields, (b) two normal magnetic fields, and (c) two parallel fields, one magnetic and one electric. We now proceed to derive the relationships between beam deflection and fields for these combinations.

⁷ E. W. Engstrom, "An Experimental Television System." *Proc. IRE*, **21**, 12 (December 1933).

⁸ V. K. Zworykin, "Description of an Experimental Television System and the Kinescope." *Proc. IRE*, **21**, 12 (December 1933).

⁹ C. F. Hardy, R. H. George, and H. J. Heim, *op. cit.*

¹⁰ R. D. Kell, "Description of Experimental Television Apparatus." *Proc. IRE*, **21**, 12 (December 1933).

¹¹ I. G. Maloff and D. W. Epstein, *Electron Optics in Television*. New York: McGraw-Hill Book Company, Inc., 1938.

3-3. Electrostatic Focus and Deflection

In Fig. 3-7 an electron, free to move under the accelerating field due to E , is shown between two plates. Since the electron is free to move, it will eventually arrive at the plate of higher potential with potential energy $E\varepsilon$. By the law of the conservation of energy it must be true that on impact its potential energy is converted into kinetic energy. Thus,

$$\frac{1}{2}mv_0^2 = E\varepsilon \quad (\text{rationalized M.K.S. units})$$

$$\text{or} \quad v_0 = \sqrt{\frac{2E\varepsilon}{m}} \quad (3-1)$$

where $\varepsilon =$ electronic charge
 $= 16 \times 10^{-20}$ coulomb

and $m =$ electronic rest mass
 $= 9.03 \times 10^{-31}$ kilogram

Thus the final velocity, v_0 , of the electron is proportional to the square root of the accelerating voltage E .

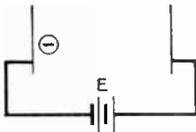


Fig. 3-7. An electron is free to move between two plates.

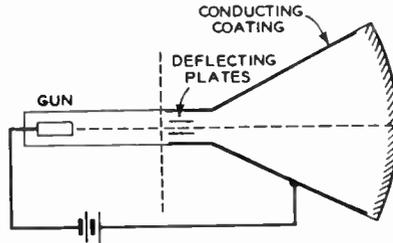


Fig. 3-8. The basic cathode-ray tube. The beam electrons maintain a constant axial velocity v_0 in the region to the right of the vertical dotted line.

Now in the cathode-ray tube the parallel plates of Fig. 3-7 are replaced by an electron gun and an anode consisting of a conducting coating inside the envelope at the end away from the gun (Fig. 3-8). Assuming the coating to have zero resistance it is an equipotential surface and, in fact, the whole region to the right of the dotted line and interior to the bulb is at constant potential and corresponds to the

“plate” of the previous figure. Thus an electron arrives in that region with an axial velocity v_0 given by (3-1) and further maintains that axial velocity throughout that entire region.

In its practical form the cathode-ray tube employs a steady beam of electrons in place of the single electron just discussed. In fact, it is this electron beam which causes the fluorescent screen on the tube face to give off light at the point of impact. We have already seen in the last chapter that the cross-sectional area of this beam should be as small as possible because it serves as the scanning aperture whose size determines to a large degree the reproduction of detail in the final image. Our immediate problem, then, is to consider by what means the electrons that are emitted from the thermionic cathode in the CRT may be focused so that they converge into a small spot at the fluorescent screen. In accomplishing this, use is made of the fact that the paths of electrons passing through electrostatic or magnetostatic fields are bent in a manner quite similar to that in which light rays are refracted when they pass through regions of different refracting index. As a matter of fact, Snell's law of optics may be carried over in modified form to give a corresponding equation in electron optics. The analogy between the two forms is illustrated in Fig. 3-9.¹² Comparison of the two examples in the figure shows

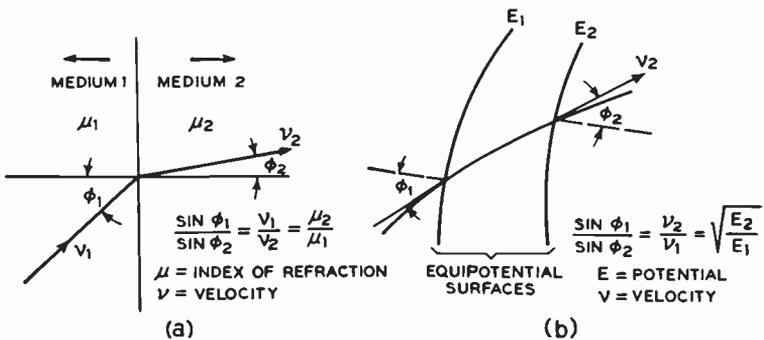


Fig. 3-9. A close analogy exists between optics and electron optics. (a) Optical path in optics. (b) Electron trajectory in electron optics.

¹² See, for example, M.I.T. Electrical Engineering Staff, *Applied Electronics*. New York: John Wiley and Sons, Inc., 1943; Chap. 1. Also I. G. Maloff and D. W. Epstein, *Electron Optics in Television*. New York: McGraw-Hill Book Company, Inc., 1938.

that the square root of voltage may be interpreted as the index of refraction in electron optics.

It follows more or less directly from Snell's law that the path of an electron in an electrostatic field may be traced point by point if the field configuration is known, say, in terms of the equipotential lines. Although the application of this technique is beyond the scope of our work, we may state the results, which are of importance here: By use of a suitable electrostatic field the electrons that leave the cathode in a cathode-ray tube may be focused into a small spot at their point of impact on the fluorescent screen. In this process, the electrostatic field may be made to serve as an electron lens. Let us see how this might be accomplished in a typical CRT.

The several parts of the electron gun are illustrated in Fig. 3-10, where the equipotential lines resulting from the applied voltages and electrode geometry are shown. Two lenses are provided, one by the control grid and first anode, and the second by the two anodes. The former serves to focus those electrons which leave the cathode in the direction of the defining aperture in the grid onto the crossover point, which is labeled x in the diagram. Notice, then, that a much greater constriction of the beam is obtained than if the defining aperture served only as a physical stop. The net result is that the crossover point, which is very small, serves as a virtual point source of electrons, which then enter the second lens. There the electric field bends the electron trajectories so that an image of the crossover point appears at the fluorescent screen; that is, the electron lens

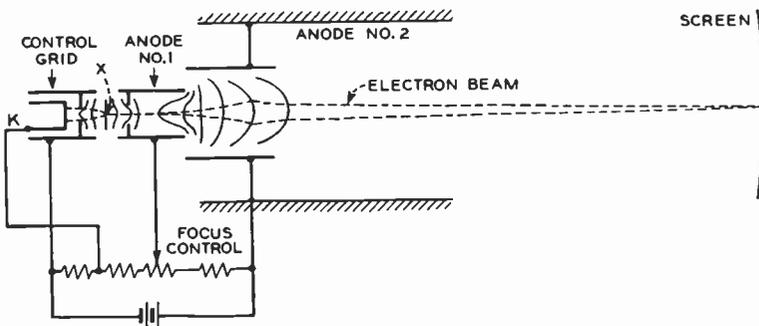


Fig. 3-10. Electrostatic focusing in a cathode-ray tube; x represents the crossover point.

focuses the crossover point onto the screen. In this manner the electrons that leave the cathode are made to arrive at the screen as a beam of very small cross-sectional area. In a typical case the diameter of the beam at the point of focus may be as small as 0.1 millimeter or less. Since the shape of the focusing field may be varied by controlling the voltage on the first anode, that voltage is generally made variable by returning it to a potentiometer, the focus control, in the high-voltage bleeder system.

The focusing system just described utilizes the bending action of an electrostatic field and hence is termed electrostatic focusing. Its advantages are that it draws no power (few, if any, electrons reach the first anode) from the high-voltage supply, and that the lens system itself is contained within the cathode-ray tube envelope. Its chief disadvantage is that the lens size is small and not capable of providing as fine a spot on the fluorescent screen as does the magnetic system of focusing, which is described later in the chapter. In current practice electrostatic focusing is confined to cathode-ray tubes up to 7-in., or occasionally a 10-in., size and is not used in the larger picture or television camera tubes.

Since we have seen how the electrons are formed into a narrow beam in the CRT, we must now consider how that beam is made to move across the face of the tube in the prescribed scanning pattern. It must be stated that after focusing in the electron gun the electrons enter the constant potential region which is provided by the aquadag coating on the tube envelope. In this region the electrons have a constant axial component of velocity v_0 . If, now, the beam be made to pass between a pair of deflecting plates, such as shown in Fig. 3-8,

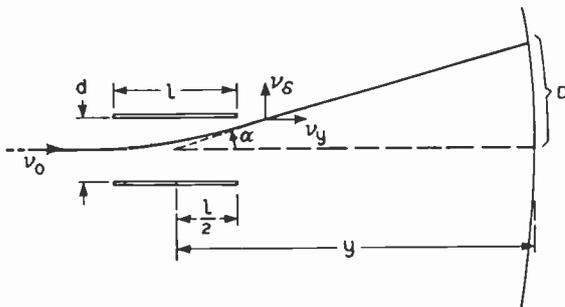


Fig. 3-11. Electrostatic deflection.

it may be deflected normal to the direction of v_0 by a deflecting voltage V across the plates. We shall now derive the equation which relates the beam deflection to the deflecting voltage. An enlarged view of the deflecting system is shown in Fig. 3-11.

Since v_0 remains constant, the time of flight, t , during which an electron remains between the plates is

$$t = \frac{l}{v_0} \quad (3-2)$$

Assuming a uniform field between the deflecting plates and neglecting fringing effects, the field intensity normal to the plates is

$$\varepsilon = \frac{V}{d} \quad (3-3)$$

and the normal force on an electron is given by

$$f = e\varepsilon = \frac{Ve}{d} = ma$$

The normal acceleration is

$$a = \frac{Ve}{md} \quad (3-4)$$

Then, integrating, $v_\delta =$ normal velocity at time t

$$= at = \frac{Ve}{md} \frac{l}{v_0} \quad (3-5)$$

and $\delta =$ vertical deflection from axis at time t

$$= \frac{1}{2} at^2 = \frac{Ve l^2}{2mdv_0^2} \quad (3-6)$$

Once the electron has passed beyond the edges of the deflecting plates, it is subject to no accelerating forces and, hence, moves with a constant velocity, which is the resultant of the two velocity components v_δ and v_0 both of which are constant. Thus,

$$\tan \alpha = \frac{v_\delta}{v_0} = \frac{Ve l}{mdv_0^2} \quad (3-7)$$

Comparison of (3-6) and (3-7) reveals that the angle of deflection is also given by

$$\tan \alpha = \frac{\delta}{l/2} \quad (3-8)$$

This last equation is important for it shows that the actual parabolic path of the deflected electron may be considered to be replaced by a path that continues in the direction of v_0 until the center of the plates is reached, and then undergoes an abrupt change of direction by an angle α . This equivalent electron path is shown by the dashed line in Fig. 3-11.

Then, continuing the derivation—for we wish to relate the total deflection D to the deflecting voltage V —we have:

$$\begin{aligned} D &= \text{deflection on screen corresponding to angular deflection of } \alpha \\ &= y \tan \alpha \end{aligned} \quad (3-9)$$

and substituting from (3-1) we get

$$D = \frac{ly}{2d} \frac{V}{E} \quad (3-10)$$

We may illustrate the use of eq. (3-10) with the following example: The dimensions of a certain cathode-ray tube are

$$\begin{aligned} l &= 2 \text{ cm} & y &= 20 \text{ cm} \\ d &= 1 \text{ cm} & E &= 1 \text{ Kv} \end{aligned}$$

Then the voltage required to produce a deflection of 5 centimeters will be

$$V = \frac{2dDE}{ly} = 2 \left(\frac{1}{2} \right) \left(\frac{5}{20} \right) 10^3 = 250 \text{ volts}$$

Further calculation will show that the maximum deflection which may be obtained in this tube without a change in accelerating voltage is 10 cm, which corresponds to a deflecting voltage of 500 volts. If the deflecting voltage is increased above this value, the deflection angle α will exceed $26.5^\circ = \text{arc tan} \left(\frac{d}{2} \left/ \frac{l}{2} \right. \right)$ and the electron beam will hit the deflecting plates. This illustrates one of the limitations on deflection angle in the electrostatic CRT: for a given deflecting voltage, V , the deflection is inversely proportional to the interplate spacing, d . It seems desirable, then, to make this spacing as small as possible but, of course, this is incompatible with large values of α , for if the electrons are to reach the phosphor, they must not hit the deflecting plates. This condition is remedied in practice by either putting a sharp bend in the plates or by giving them a gradual flare. In either case the deflection equations above must be modified by

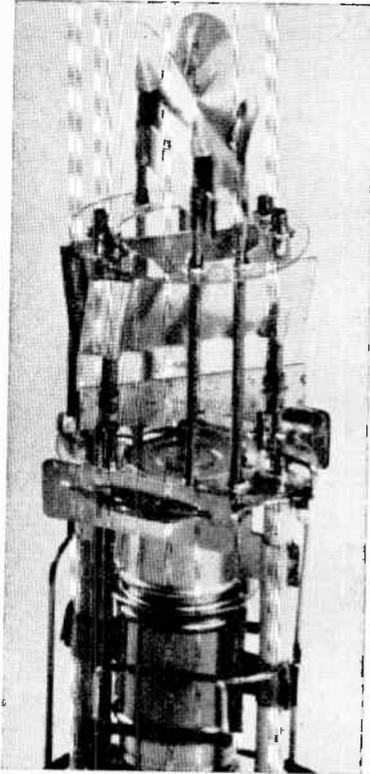


Fig. 3-12. Detail of the curved plates used in electrostatic deflection. (Courtesy of Allen B. Du Mont Laboratories, Inc.)

suitable factors,¹³ but in any case the deflection remains directly proportional to V , the deflecting voltage across the deflecting plates.

Since in any given CRT the physical dimensions remain constant, it is the usual practice to arrange eq. (3-10) in the following manner:

$$\frac{D}{V} = \left(\frac{l_y}{2d} \right) \frac{1}{E} \quad (3-11)$$

where D/V is termed the deflection sensitivity and which may be expressed in millimeters per volt. It may be seen from the last

¹³ An excellent treatment for deflection with curved deflecting plates is given by I. G. Maloff and D. W. Epstein, *op. cit.*

equation that the deflection sensitivity will be constant for a given tube operating at a fixed value of accelerating voltage, E . Furthermore, since the deflection sensitivity may be determined experimentally, it gives an easy way of handling the cases where bent or flared deflection plates are used. In recent years there has been an industry-wide trend to replace the deflection sensitivity of eq. (3-11) by its reciprocal V/D , the deflection factor. This factor is generally expressed in volts per inch. Values of these quantities for three typical television cathode-ray tubes of the electrostatic type are given in Table 3-1. Two sets of deflecting plates at right angles to each other are usually provided in cathode-ray tubes so that deflections in both the horizontal and vertical directions may be obtained. Since one set is generally nearer the tube face than the other, the dimension y is different for the two. In the table values are given for the set nearer the fluorescent screen.

TABLE 3-1
DEFLECTION SENSITIVITIES OF TYPICAL TELEVISION CRT'S

| Type | Approx. screen size, in. | Accelerating voltage, volts | D/V , mm/volt | V/D , volt/in. |
|-------|-----------------------------|--------------------------------|--------------------|---------------------|
| 5BP4 | 5 | 1500 | 0.45 | 57 |
| 7JP4 | 7 | 6000 | 0.94 | 27 |
| 10HP4 | 10 | 5000 | 0.251 | 100 |

We may summarize electrostatic deflection as follows: For a constant accelerating potential, the deflection of the electron beam is directly proportional to the deflecting voltage and is at right angles to the plane of the deflecting plates. Therefore, if the beam is to trace out a definite pattern on the fluorescent screen, the deflecting voltage must vary and, further, its shape must be the same as that of the desired deflection. Another fact which will be of importance in a later discussion of ion spot may be seen directly from eq. (3-10); the total deflection is independent of m and ϵ , the mass and charge, respectively, of the deflected particles.

3-4. Magnetic Focus and Deflection

In the last section we saw that the electron beam in a cathode-ray tube may be focused and deflected by means of electric fields. In

the present section we shall consider how the same results may be accomplished with magnetic fields. We shall first discuss magnetic focusing.

The need for magnetic focusing has already been stated: the electrostatic method gives a larger spot on the fluorescent screen of the CRT. The reason for this is that both types of lenses exhibit spherical aberration, that is, those electrons which approach the lens on trajectories that are inclined at a large angle relative to the axis of the tube are not bent sufficiently to return to a common focal point with electrons which remain near the tube axis. Thus there is a scattering of electrons about the focal point, and the resulting spot of impact is larger than it should be. One means of overcoming this form of aberration is to use a larger lens field so that all of the beam electrons remain in a relatively small portion of the field. This remedy is difficult to apply in the electrostatic case because, as we have seen, the lens electrodes lie within the envelope of the CRT neck. With magnetic focusing, on the other hand, the magnetostatic focusing field is produced by a coil or permanent magnet, which may be exterior to the tube envelope, and a similar limitation on size is not present. We may add parenthetically that the use of the external magnetic focus and deflection system also eliminates the need for the electrostatic deflection plates, and a CRT with a smaller tube neck diameter may be utilized. Then in the magnetic case superior focusing is obtained primarily because the lens size is increased.

In order to study how magnetostatic focusing takes place we shall investigate the system shown in Fig. 3-13a. That system uses a short focusing coil.¹⁴ The coil is formed of several turns of wire on a soft-iron ring which has an annular air gap. The upper part of the coil is surrounded by a magnetic shield so that opposite sides of the ring on either side of the air gap serve as pole faces. The general shape of the magnetic field which results when direct current flows through the coil is illustrated in the figure.

Whereas equations giving the focal length of such a lens may be

¹⁴ A second case where a long focusing coil extends over the entire length of the electron beam is described in Chapter 6. The short coil is the form used in cathode-ray tubes.

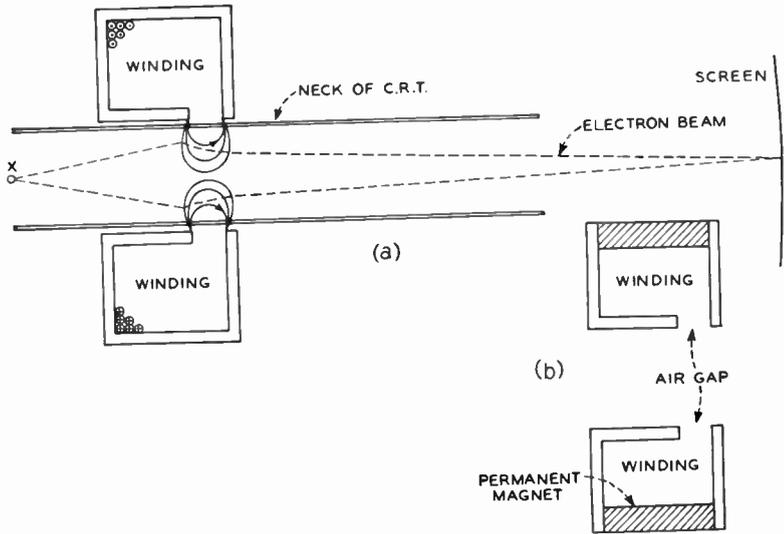


Fig. 3-13. (a) Magnetostatic focusing. (b) Detail of a PEM focus coil.

derived, subject to a number of simplifying assumptions,¹⁵ we shall merely describe the physical action which takes place in the focusing field. Any electron which enters the field in such a direction that it crosses the lines of flux is subject to a bending action whose direction may be determined by the left-hand motor rule. It may be shown that such an electron is subject to a force that causes it to follow a spiral path. This path will cause the electron to be directed toward the focal point as it leaves the magnetic field. It is by this means that the focusing action takes place.

A more rigorous analysis shows that, for a given accelerating voltage, the focal length is proportional to the square of the current in the focus coil and so is independent of the direction in which that current flows. It follows at once that focus control may be obtained by varying the current through the coil. It is common practice, however, to use this method for fine control only, coarse control being

¹⁵ See T. Soller, M. A. Starr, and G. E. Valley, Jr., *Cathode Ray Tube Displays*, Volume 22, M.I.T. Radiation Laboratory Series. New York: McGraw-Hill Book Company, Inc., 1948; also, I. G. Maloff and D. W. Epstein, *op. cit.*; and V. K. Zworykin and G. A. Morton, *Television*. New York: John Wiley and Sons, Inc., 1947.

obtained by adjusting the position of the focus coil along the neck of the CRT. In typical focus coils for 10- to 16-in. tubes the winding resistance is approximately 475 ohms and the direct current required ranges from 75 to 200 milliamperes, the exact value being dependent upon the accelerating voltage that is applied to the tube. It will be realized that the voltage drop across such a coil is small and that it may be energized from a tap on the low-voltage supply of equipment.

Some saving of focusing power may be had by making a composite focus coil which utilizes a permanent magnet as well as an electromagnet to produce the required field. The construction of a coil of this type, which is commonly designated as a PEM coil, is shown in Fig. 13-13*b*.¹⁶

The use of an external magnetostatic focusing field eliminates the need for the electrostatic focus system in the CRT and an electron-gun structure different from that previously described may be used. Present practice utilizes a tetrode structure, which consists of a cathode, control grid, screen grid, and anode, the latter three being axial cylinders of the same diameter. As in the electrostatic tube shown in Fig. 3-10, the anode is connected to the aquadag coating on the inner surface of the tube envelope.

For the sake of completeness it should be stated that it is not necessary to use magnetostatic focusing in the tubes which employ magnetic deflection of the electron beam. In fact the 9-in. cathode-ray tubes used in prewar television equipment generally used electrostatic focusing with magnetic deflection. Postwar practice, however, is to use magnetic focusing in the larger tubes, which employ magnetic deflection of the beam, because of the superior performance it affords.

Since we have discussed means for focusing the electron beam into a fine spot on the CRT screen, we are now in a position to consider the second method of deflecting the electron beam, the method in which a magnetic deflecting field is used. For the purposes of the present discussion we will assume this field to be uniform and restricted to a definite region of length l as shown in Fig. 3-14. We further assume that the direction of the field is out of the page as indicated. The placement of electron gun and anode is the same

¹⁶ For constructional details of focus coils, see Soller *et al.*, *op. cit.*

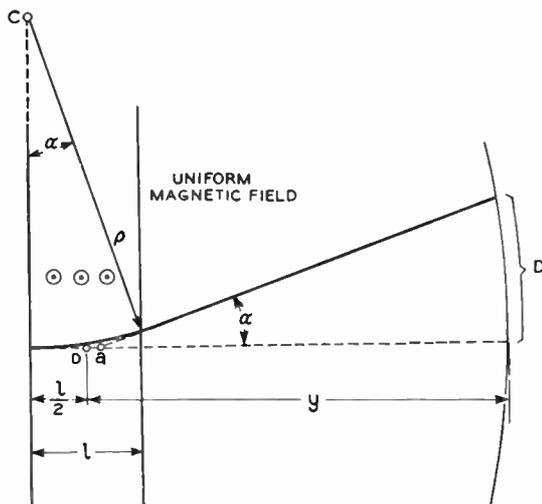


Fig. 3-14. Magnetic deflection.

as that shown in Fig. 3-8, only in this case the deflecting electric field applied across the deflecting plates is replaced by the uniform *magnetic* field. Application of Fleming's left-hand motor rule (remember that conventional current flow is in the direction opposite to that of the electron motion) shows that the deflection will be upward in the diagram. We may now derive the relationship for D , the screen deflection.

As in the previous case the electron maintains a constant axial velocity, v_0 , given by (3-1).

Let

H = field intensity, amperes/meter

B = flux density, webers/meter²

μ_0 = permeability of space = $4\pi \times 10^{-7}$ kilogram-meter/coulomb²

Now the force on the electron resulting from the magnetic field is normal to the direction of v_0 and causes the electron to move in a circular path with radius of curvature ρ . The force due to this field may be calculated by Ampere's law

$$f = B\varepsilon v_0 = \mu_0 H \varepsilon v_0 \quad (3-12)$$

and this must be equal to the centrifugal force

$$f = \frac{mv_0^2}{\rho} \quad (3-13)$$

Thus,

$$\rho = \frac{mv_0}{\mu_0 H \epsilon} \quad (3-14)$$

Now since the direction of electron motion at any point on the path is always normal to the radius vector from the center of curvature to that point, it follows that the radius vector swings through an angle α equal to the change in path directions of the electron on entering and leaving the magnetic field. Then from the diagram it may be seen that

$$\sin \alpha = \frac{l}{\rho} \quad (3-15)$$

It should be noticed that, in general, the point a does *not* lie midway between the boundaries of the magnetic field. Thus,

$$\begin{aligned} D &= (y - \overline{oa}) \tan \alpha \\ &= (y - \overline{oa}) \frac{l}{\rho} \\ &= \frac{(y - \overline{oa}) l}{\sqrt{1 - \left(\frac{l}{\rho}\right)^2}} \end{aligned} \quad (3-16)$$

If now the expression for ρ be substituted in (3-16), a rather formidable equation results which is difficult to interpret. If, however, the angle of deflection be sufficiently small, we may assume

$$\text{that} \quad \alpha = \sin \alpha = \tan \alpha$$

and further, since l is small compared to y

$$y - \overline{oa} \approx y$$

and the approximate equation results:

$$D \approx y l \mu_0 H \sqrt{\frac{\epsilon}{2Em}} \quad (3-17)$$

Substituting for ϵ and m for the electron in M.K.S. units we get

$$D \approx 4\pi(0.0297) \frac{y l H}{\sqrt{E}} \quad (3-18)$$

where the units of H are amperes/meter, the units of D , y , and l

are meters and the units of E are volts. Hence it may be seen that in magnetic deflection the total *screen deflection is proportional to H* , the field intensity, provided that the deflection angle is small. Again the second direction of deflection required by the scan may be obtained by using another magnetic field normal to that described above, and the same deflection equations obtain.

At this point it might be well to check the range of α , for which the approximation of eq. (3-18) is valid. The procedure using the exact forms would be to determine $\sin \alpha$ from (3-15) and then find $\tan \alpha$ from the tables, whereas in the approximate form the calculated value of the sine is assumed equal to the tangent. Hence, the per cent error involved in the approximation is

$$\left(\frac{\tan \alpha - \sin \alpha}{\tan \alpha} \right) 100\% = (1 - \cos \alpha) 100\%$$

The following data may be obtained from any set of mathematical tables:

| Sin α | Tan α | % Error | Approx. α° |
|--------------|--------------|---------|------------------------|
| 0.343 | 0.365 | 3.3 | 20 |
| 0.371 | 0.399 | 7.0 | 22 |
| 0.426 | 0.471 | 9.6 | 25 |
| 0.497 | 0.573 | 13.3 | 30 |

Thus the approximation is good within 10 per cent for a deflection of 25° . This is convenient, for several postwar wide-angle cathode-ray tubes use a peak-to-peak angular deflection of about 50° , which corresponds to an α of 25° . Since in all cases $\sin \alpha < \tan \alpha$, the deflection calculated from (3-18) will be conservative, i.e., it will be less than the actual deflection predicted by the exact equations.

In magnetic deflection the required magnetic field is established by causing current to flow through the coils of a deflection yoke, which is slipped over the neck of the cathode-ray tube and whose geometry is such that a uniform magnetic field is produced in the deflection region. Such a yoke of simple form is shown in Fig. 3-15. Assuming for a moment that the field produced in the cathode-ray tube by the yoke is perfectly homogeneous and square in cross section, we may derive a relationship between D , the total deflection, and I , the deflection current flowing through the yoke windings. Let: N be the total number of turns, and h the height of the field in meters. Then

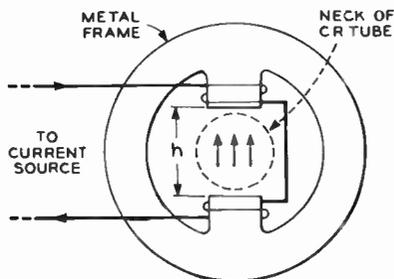


Fig. 3-15. Simple yoke for magnetic deflection.

$$H = \frac{NI}{h} \quad (3-19)$$

and substitution into (3-18) yields

$$\begin{aligned} D &= \frac{4\pi(0.0297)yl}{\sqrt{E}} \frac{NI}{h} \\ &= 0.373 \frac{yl}{h} \frac{NI}{\sqrt{E}} \end{aligned} \quad (3-20)$$

It can be seen, therefore, that under the conditions assumed, the total screen deflection is directly proportional to the deflecting current. Therefore, in *magnetic deflection the deflecting current must have the same shape as that of the required deflection*. Let us use eq. (3-20) to find the number of ampere turns required to produce the scanning raster in a 10BP4 magnetic deflection tube. The width of the scanning lines is to be 8 in. so that the complete raster of standard aspect ratio will fit on the face of the 10-in. CRT. The deflection yoke is 5 centimeters long and produces a field of height 5 centimeters. The yoke is centered on the neck of the tube at a point 9.5 in. from the fluorescent screen. The anode accelerating voltage is 8 kilovolts.

Our equation is derived on the basis of an angular deflection, α , away from the center of the screen. Thus a current which increases to a maximum and then reverses direction to the same value in the opposite direction is required. Thus, if we let D equal the line width of 8 in. NI will be the peak-to-peak value of the ampere turns. Then, by substituting into (3-20), we have

$$\begin{aligned} NI &= \frac{\sqrt{E}hD}{0.373ly} = \frac{\sqrt{8000}}{0.373} \left(\frac{5}{5}\right) \left(\frac{8}{9.5}\right) \\ &= 201 \text{ ampere turns, peak-to-peak} \end{aligned} \quad (3-21)$$

Notice that the units of h and l , and of D and y cancel; hence they need not be converted to meters.

Now as a practical matter the simple deflection yoke of Fig. 3-15 does not produce a homogeneous magnetic field as assumed. This condition may be remedied, however, by eliminating the iron core completely and by using instead an air-core coil whose windings are

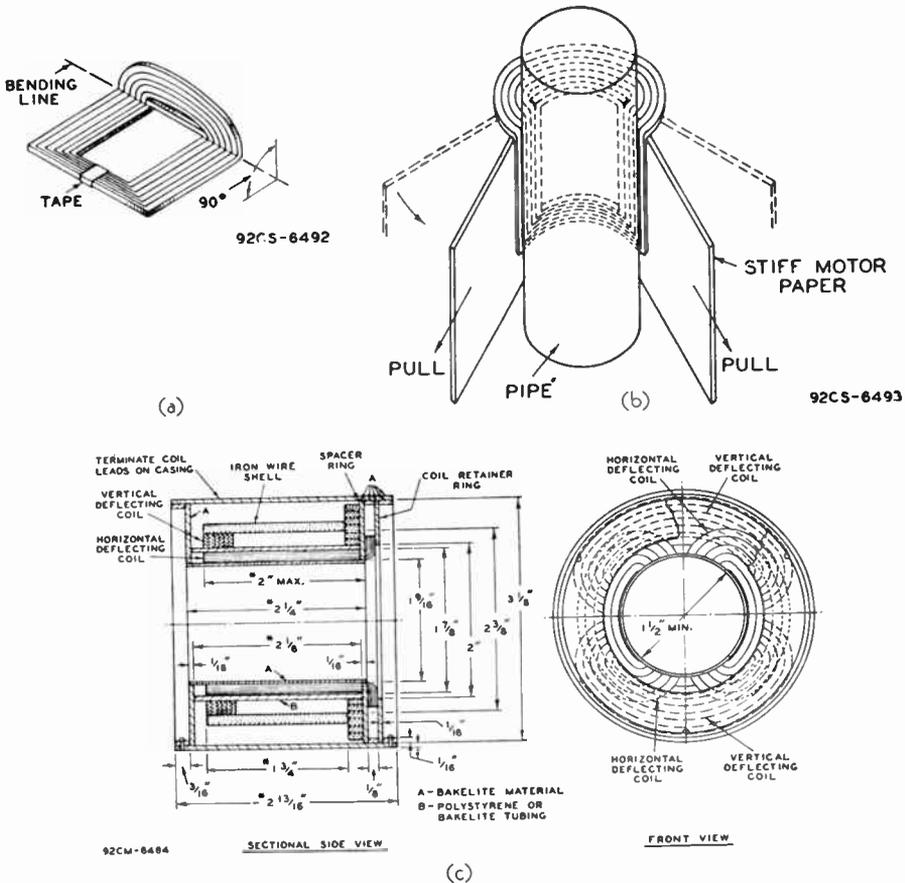


Fig. 3-16. A typical television deflection yoke. (a) The screen end of the flatwound coil is bent up 90° . (b) The flat coil is bent into shape around a mandrel. (c) The complete yoke, showing the horizontal and vertical deflection coils in place. (Courtesy of Radio Corporation of America.)

shaped in a particular fashion. In practice a pair of distributed winding coils are bent into a cylindrical shape and a good approximation to the uniform field may be obtained.^{17,18}

The manufacture of such a coil is not too difficult. Each coil of the pair is formed, in the flat, on a wiring board, and after taping, is bent around a mandrel to the required shape. Generally the coil ends toward the screen are bent up 90° from the axis of the tube to extend the field slightly in that direction. The chief steps in forming such a coil are shown in Fig. 3-16. The complete yoke, comprising two coil pairs at right angles (one pair for each deflection direction, vertical and horizontal), is mounted on an insulating form and surrounded by a sleeve of high permeability metal, which serves as a magnetic shield.¹⁹ The manufacturing procedure of a deflection yoke which has just been described is primarily used at the present time for fabricating experimental coils only. Once the final design of a yoke has been established, the separate coils are machine-wound on dies so that they are preformed to approximately the correct shape. The final forming is carried out in a press. This latter method is much more suitable for modern production methods and lowers the unit cost of a yoke by an appreciable factor.

3-5. Comparison of Electrostatic and Magnetic Deflection

Some interesting points of comparison between the two deflection systems described may be made by reference to eq. (3-10) and (3-20). To illustrate one point having to do with accelerating voltage effects, these may be written as

$$\text{Electrostatic Deflection} \quad D = K_1 \frac{V}{E} \quad (3-22)$$

$$\text{Magnetic Deflection} \quad D \approx K_2 \frac{I}{\sqrt{E}} \quad (3-23)$$

where E is the cathode-anode accelerating voltage, V the deflecting voltage, and I the deflecting current.

¹⁷ Even better results may be obtained if the winding density is distributed according to the cosine law. See K. Schlesinger, "Magnetic Deflection of Kinescopes." *Proc. IRE*, **35**, 8 (August 1947).

¹⁸ For an excellent treatment of the manufacturing techniques of air-core deflection yokes, see Soller *et al.*, *op. cit.*

¹⁹ *Preliminary Data on Television Scanning Circuits and Components*, Part I, RCA, RCA Victor Division, 1944.

Now in a cathode-ray tube both the screen brilliance and beam size are related to the accelerating voltage and both require a value of accelerating voltage ranging upwards from 5 to 15 kilovolts depending upon the tube size and type of service. Reference to the above equations indicates that as E is raised, V must increase in *direct* proportion for electrostatic deflection to maintain a given screen deflection. Evaluation of the constant K_1 for typical cathode-ray tubes shows, for tubes having screen diameters greater than about 7 in., that values of deflecting voltage are required that are higher than may be conveniently generated with receiving-type tubes.

In magnetic deflection tubes, however, the values of deflecting current required for tubes up to 20 in. in diameter may be obtained with relatively small beam power tubes. In consequence the 7-in. tubes have come to be a transition size for electromagnetic and electrostatic deflection. Thus we find that tubes smaller than 7 in. tend toward electrostatic deflection, whereas the larger tubes use magnetic deflection almost exclusively. In the 7-in. size both types are available, the 7EP4 being representative of the former type, and the 7DP4 of the latter.

In the postwar development of television receivers, emphasis has been on tubes having screen diameters of 10 in. or greater; hence, emphasis has been on the design of better and more efficient magnetic deflection systems and components. This subject is covered in some detail in the next chapter.

Another comparison may be made between the two types of deflection relative to the presence of the so-called "ion spot" or "negative ion blemish."²⁰ In this regard it is unfortunate that in the completed cathode-ray tube negative ions, other than electrons, are present. Various experiments tend to indicate that oxygen and chlorine ions, O_2^- and Cl^- , predominate in this category. The origin of these ions in the tube has been the subject of considerable research, and various theories have been postulated concerning this. Bowie suggests that the following three mechanisms are involved: (1) direct emission of the ions by the cathode, (2) secondary emission of the negative ions resulting from positive ion bombardment of the grid and cathode, and (3) formation of the ions by beam electrons attaching themselves to the appropriate molecules present within the tube envelope.

²⁰ R. M. Bowie, "The Negative-Ion Blemish in a Cathode Ray Tube and Its Elimination." *Proc. IRE*, **36**, 12 (December 1948).

Why the presence of these ions is a source of trouble in magnetic deflection may be seen from eq. (3-17). In this case the deflection of any charged particle depends on its charge-to-mass ratio, e/m . Hence any ion of charge-to-mass ratio different from that of the electron will be deflected by the magnetic field through a different angle than will be the electrons. Thus the scan pattern for the ions differs from that of the electrons. The final visible effect of the ions on the CRT screen will depend upon the type of focusing used in the tube. In electrostatic focusing as in electrostatic deflection, the focusing properties are independent of the e/m ratio of the beam particle; hence the ions as well as the electrons are focused onto a small spot on the fluorescent screen. Since the heavy ions suffer negligible deflection in the magnetic deflecting field, they continuously strike a small spot at the center of the tube face and actually burn or poison the screen at the point of impact so that it no longer gives off light. As viewed on the screen the ion blemish appears as a small,



Fig. 3-17. The appearance of ion spot on the face of a cathode-ray tube that employs electrostatic focus and magnetic deflection. The ion spot is the small, dark circular area over the subject's left eyebrow. (Courtesy of Sylvania Electric Products, Inc.)

stationary, brown spot. An example of this for the conditions just cited, that is electrostatic focus and magnetic deflection, is shown in Fig. 3-17.

The ion spot appears as a larger brown spot about 1 in. in diameter when magnetostatic focusing is used with magnetic deflection. The reason for the larger spot size under these conditions is that the heavy ions are relatively unaffected by the magnetostatic focusing field and so arrive at the screen in a larger, unfocused region.

The presence of either of these forms of ion blemish may be very annoying to the television viewer. Not only does it cause severe degradation of the picture detail near the screen center, but also if the camera is scanned horizontally across a scene, the ion spot exhibits the objectionable property of apparently moving in the opposite direction.

Apparently the formation of ion spot may be reduced by closer control in the manufacturing of the tubes. The resulting increase in cost because of more rejects has made more direct corrective methods attractive. Three methods of attack on the problem have been used. First, it has been found experimentally that blemish effects tend to decrease with increasing accelerating voltage and are practically negligible on sulfide-type screens at 12 to 15 kilovolts. The majority of receivers at the time of writing however, use high voltages limited to about 9 or 10 kilovolts which is below the threshold value. Thus other cures are indicated.

Secondly, attempts to reduce the blemish have been made by backing the fluorescent screen with a thin metal coating.²¹ While this type of screen backing was developed primarily to increase screen brilliance, it nevertheless is of aid in the present problem, for it has been demonstrated that the depth of penetration into a substance of a moving particle is inversely proportional to the particle mass. Thus the massy ions are impeded from reaching the phosphor by the metal backing coat. Obviously the coating must be limited in thickness so that it remains pervious to the electrons. Thus the metal-backed screen does not eliminate the problem but only ameliorates it.

The third approach involves the use of an ion trap, whose function is to introduce an auxiliary bend in the beam path to prevent the

²¹ D. W. Epstein and L. Pensak, "Improved Cathode Ray Tubes with Metal Backed Luminescent Screens." *RCA Rev.*, VII, 1, 5 (March 1946).

massy ions from reaching the fluorescent screen. One of the earliest forms of ion trapping, according to Bowie, employed the simple expedient of bending the envelope proper of the cathode-ray tube. This is illustrated in Fig. 3-18. A d-c biasing field is produced by a direct current of the proper magnitude in the deflection yoke. This trap field is of sufficient strength to bend the electrons around toward the screen, but the path of the massy ions remains unaffected and they hit the edge of the screen outside the picture boundaries.

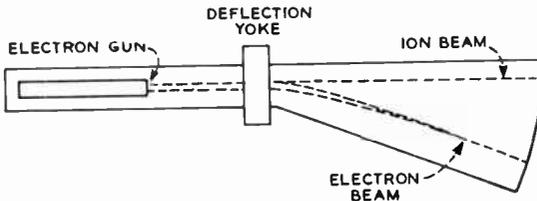


Fig. 3-18. An early ion-trap tube. The heavy ions are unaffected by the deflection field and hit the fluorescent screen outside of the picture area. (Courtesy of *Proc. IRE.*)

Although this simple construction performs the ion-trap function satisfactorily, it nevertheless results in a cumbersome tube. But once a principle is established, improvement needs only time.

In some of the later prewar Philco sets, the same principle was used, but the electron gun rather than the tube was bent. Postwar trends have been toward a design that bends the gun electrically rather than mechanically by proper shaping of the electrodes and electric fields. In either case a separate field provided by an ion-trap magnet is required. This may be furnished by either a permanent-magnet or an electromagnet. Two typical systems, employing these variations,

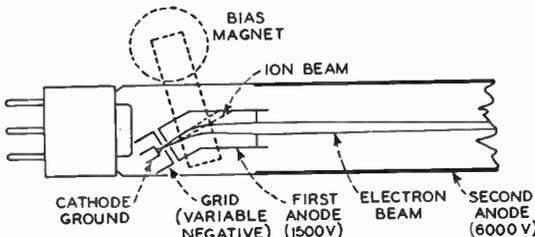


Fig. 3-19. A bent-gun ion trap. (Courtesy of *Proc. IRE.*)

are shown in Figs. 3-19 and 3-20. The latter diagram illustrates the gun and trap structure used in the popular 10BP4, which has proved to be the tube around which the majority of television receivers produced in the period 1946 to 1949 were designed.

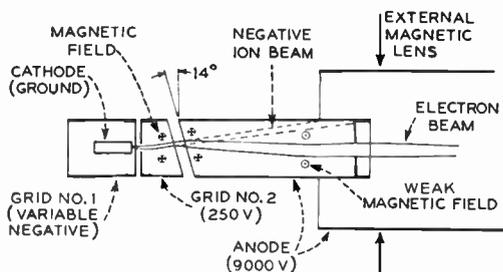


Fig. 3-20. The gun structure of the ion trap employed in the 10BP4 cathode-ray tube. (Courtesy of *Proc. IRE*.)

In the interests of completeness the two types of deflection should also be compared as to focusing defects and the like, but these are problems more closely related to electron optics and will not be discussed here. They are adequately covered in the literature.²²

SCANNING GEOMETRY

In the last section it was shown that in a cathode-ray tube the beam deflection is proportional to the deflecting voltage or current, depending upon the particular type of deflecting system being used. Furthermore, in Chapter 2 the scanning pattern was standardized. It remains, then, to determine what shape of deflecting voltage (or current) is required to cause the electron beam to trace out this prescribed pattern. Since two directions in the scan are involved, these may be obtained by two pairs of mutually perpendicular fields as previously described. Hence, as a matter of convenience we may resolve the motion of the deflected spot as it traces out the standard scan pattern into horizontal and vertical components. These, in turn, will determine the shapes of the deflecting voltages or currents. Proceeding from the simple to the complex we first consider the geometry of a progressive scan.

²² See, for example, I. G. Maloff and D. W. Epstein, *op. cit.*

3-6. Progressive Scan

The scanning raster (pattern of scanning lines) required for the progressive scan is shown at the upper left in Fig. 3-21. The picture is to be scanned from left to right and from top to bottom in a fixed number of lines, n . From the standards of the last chapter we know that these lines must be scanned at a constant speed. This requirement may be met if the horizontal deflection is of a saw-tooth form,

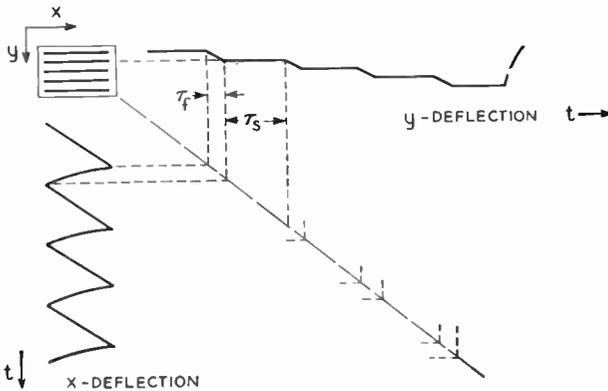


Fig. 3-21. Horizontal and vertical components required to produce a progressive scan pattern.

that is, if it increases linearly across the width of the picture and then decreases to its initial value. During this latter retrace interval no picture information is to be presented so the deflection may be of any convenient form. Now if only this horizontal deflection be imparted to the electron beam, the beam will continue to scan along a single line, retracing it over and over again. Clearly a vertical deflection component is also required in order to spread the several lines over the picture area in the vertical direction.

Now by definition in the progressive scan, the several lines are traced out in numerical order from top to bottom of the picture. Hence we require a vertical deflection which remains constant during the scan or trace portion of a horizontal line, then increases in the downward direction the distance of one line pitch during the line retrace interval, then remains constant during the next scan, and so

on. It may be seen at once that this vertical deflection will have to be locked in precisely with the horizontal sweep in order that the downward motion occur only during the horizontal retrace interval, and that it will have the general shape of a flight of stairs. This vertical deflection along with its horizontal saw-tooth deflection counterpart are necessary for the proposed scanning raster and are shown in Fig. 3-21.

We must now answer the question: whether these wave forms can be reproduced in current or voltage, for we have seen that magnetic or electric deflection require, respectively, current or voltage of the same shape as the desired deflection. The horizontal component, the saw tooth, is familiar in electronics and may be generated quite readily. The staircase wave, on the other hand, presents some problems. Its shape may be produced²³ but the requirement of locking its "treads" with the saw-tooth retrace is unduly cumbersome. Thus we seek some compromise which will simplify the vertical deflection component. Reference to Fig. 3-22 shows that the compromise

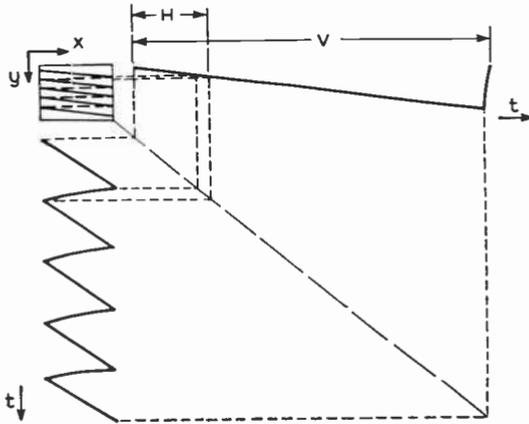


Fig. 3-22. The modified progressive scan pattern with its saw-tooth horizontal and vertical components.

is afforded by using a second saw tooth for the vertical deflection. This substitution eases the requirements on the deflection-generating circuits, simplifies the problem of synchronizing the two components, and affects the scanning raster only by causing a slight tilt downward

²³ See Linear Counter Circuit, chap. 11.

to the right. It is this compromise progressive scan, then, which utilizes two saw-tooth deflection components, that we shall consider.

Since the two scans are integrally related as shown in diagram, we may define a number of quantities and then set up the relationships between them.

Let

$$\left. \begin{aligned}
 (\tau_s)_r &= \text{vertical scan interval} \\
 (\tau_f)_r &= \text{vertical flyback interval} \\
 V_p &= \text{vertical (frame) period} \\
 f_p &= \text{frame frequency} \\
 p_v &= \text{vertical flyback ratio} \\
 &= \frac{(\tau_f)_r}{(\tau_s)_v}
 \end{aligned} \right\} \text{progressive scan} \quad (3-24)$$

then
$$V_p = \frac{1}{f_p} \quad (3-25)$$

It has been pointed out previously that the horizontal lines that occur during the vertical retrace cannot be used for picture data presentation and that they are blanked off from the screen by a negative pulse applied to the CRT control grid. It can be seen, then, that a number of lines will be lost during the vertical retrace and

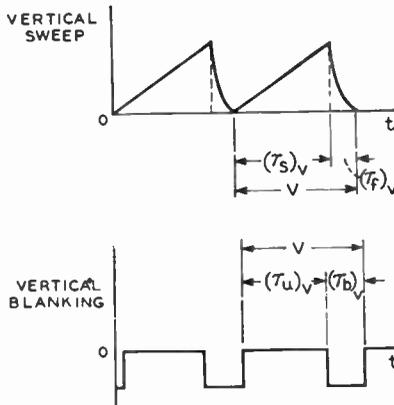


Fig. 3-23. The blanking signal is made wider than the retrace interval.

blanked interval. We shall call these "inactive lines" and designate their number by n_i . The actual number of lines which are used actively for the presentation of picture information, n_a , depends upon the length of time the CRT grid is unblanked. Thus in Fig. 3-23, if

$$\left. \begin{aligned} (\tau_u)_v &= \text{vertical unblank interval} \\ (\tau_b)_v &= \text{vertical blank interval} \\ b_v &= \text{vertical blanking ratio} \\ &= \frac{(\tau_b)_v}{(\tau_u)_v} \end{aligned} \right\} \quad (3-26)$$

$$\text{Then} \quad n_a = \frac{n(\tau_u)_v}{(\tau_u)_v + (\tau_b)_v} = \frac{n}{1 + b_v} \quad (3-27)$$

$$\text{and} \quad n_i = \frac{nb_v}{1 + b_v} \quad (3-28)$$

It may be seen that by having the blanking interval greater than the flyback interval a factor of safety is provided: picture data can be presented only on the linear portion of the scan where the requirements of constant beam velocity are met. The need for distinguishing between p_v and b_v will become apparent when we discuss scanning generators.

For the sake of completeness, we may also write down a similar set of definitions and equations for the horizontal scan. Thus,

$$\left. \begin{aligned} \text{Let} \quad (\tau_s)_h &= \text{horizontal scan interval} \\ (\tau_f)_h &= \text{horizontal flyback interval} \\ H &= \text{horizontal period} \\ &= (\tau_s)_h + (\tau_f)_h = (\tau_s)_h(1 + p_h) \\ \text{and} \quad p_h &= \frac{(\tau_f)_h}{(\tau_s)_h} \\ \text{then} \quad (\tau_b)_h &= \text{horizontal blanking interval} \\ (\tau_u)_h &= \text{horizontal unblank interval} \\ b_h &= \text{horizontal blanking ratio} \\ &= \frac{(\tau_b)_h}{(\tau_u)_h} \end{aligned} \right\} \quad (3-29)$$

and note that

$$H = \frac{1}{nf_p}$$

3-7. The Random Scan

In the progressive scanning pattern just discussed, the frequencies of the horizontal and vertical deflection components are integrally related through n , the number of scanning lines. Under this condition the lines of consecutive frames will superimpose on each other, provided that no spurious deflections resulting from hum and pickup are present. Such a system requires that the generators for the two deflection components be rigidly synchronized and means for accomplishing this are covered in later chapters.

Under certain conditions when the deflection at both pickup and reproducing ends of the television system are fed directly from a common deflection generator, a progressive scan may be used in which the vertical and horizontal deflection generators are free-running, no attempt being made to hold them in synchronism. This condition makes n a variable, and even though any given frame is scanned in a progressive pattern, lines of successive frames will not necessarily coincide with each other. This gives what might be termed a "random" or "helter-skelter" scan and is sometimes used in industrial television systems. The type I closed television system described in Chapter 8 is of this type.

3-8. Interlaced Scan

In the last chapter we saw that the interlaced system of scanning provided a means of decreasing the effect of flicker in the televised image without increasing the bandwidth of the transmitted signal. We next consider the geometry of the interlaced scanning raster in which alternate lines are scanned in sequence, each half set of lines being termed a field.

Since the horizontal deflection component is the same in both the progressive and interlaced scans, the x deflection will be a saw tooth once again. The immediate problem, then, is to find what type of y deflection will produce the desired interleaving of the two fields.

As a trial answer to the problem let us assume that the frequency of the vertical scanning saw tooth is double its value in progressive scanning. Then, during one field interval or the duration of a single vertical deflection cycle, only one-half the total number of lines will be scanned but they will be spread out over the entire picture height. During the second field interval the remaining $n/2$ lines of the frame

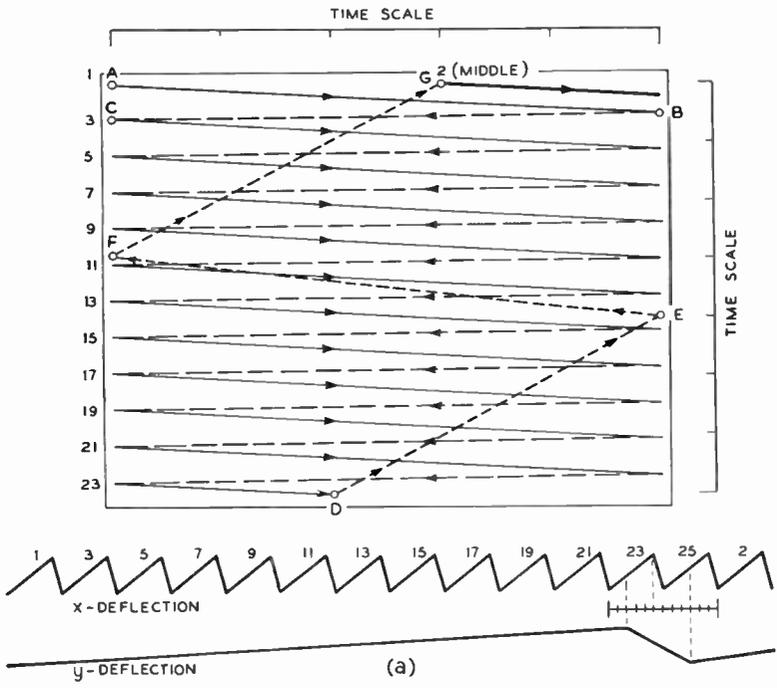
will be traced out. In order to interlace the lines of these two successive fields we must make certain that the lines of the second field fall midway between those of the first field. This condition will be met automatically if the second field starts at a point displaced one-half a line interval from the starting point of the first field. This requirement, in turn, may be satisfied if n is an odd number, for then the first field will end and the second field will begin in the middle of a line. These conditions may be understood with the help of a specific example which is shown in Fig. 3-24. In that example the total number of lines is 25, the horizontal flyback occupies one-sixth of the total line interval, and the vertical flyback has a duration of one and one-sixth line intervals. To simplify the example, the horizontal and vertical deflections are assumed to be linear during retrace.

Before studying through the figure we must make careful distinction between the terms line "width" and line "interval" in order to avoid confusion. By line width we mean the length of the line as it appears on the scanned raster. Line interval refers to the duration of one entire horizontal line period which includes both the trace and retrace. Thus, to be accurate we should say that in interlaced scanning where an odd number of lines is used, the first field ends in the middle of a line interval (rather than in the middle of a line width) and that the second field begins in the middle of a line interval. An extension of this idea will show that the term "line" which was used earlier in the discussion should be, more precisely, "line interval."

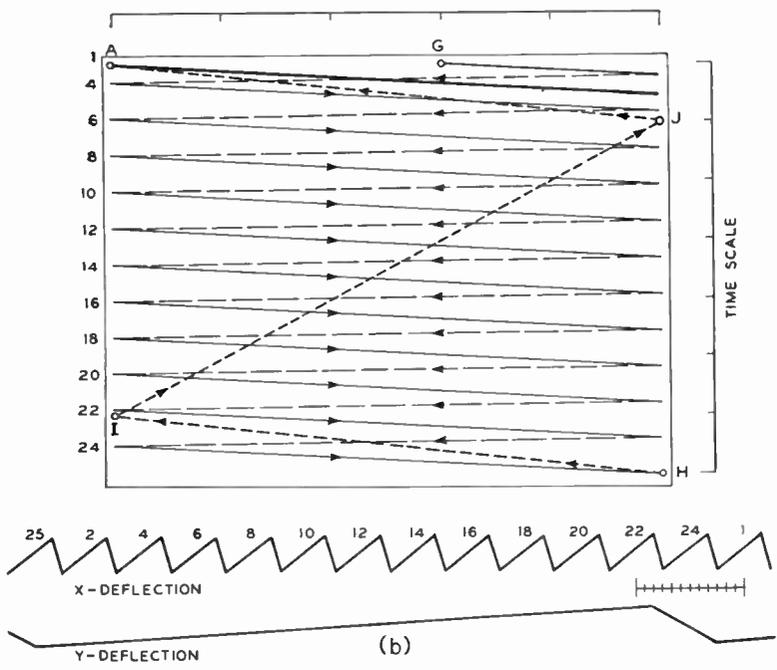
A second point which needs clarification is the system used for numbering the several horizontal lines and line intervals. It is conventional to number these in sequence from top to bottom as they appear in the complete raster. It follows that those lines which are associated with the first field are all odd-numbered and those with the second field are all even-numbered. It is convenient, therefore, to replace the terms "first field" and "second field" by "odd field" and "even field," respectively.

Let us now examine in detail Fig. 3-24*a*, which shows the odd field. Because of the combined effect of the x and y deflections the scan starts at point A . The horizontal trace lasts for five time units and carries the spot to point B . During the horizontal retrace of one

Fig. 3-24. The interlaced scan pattern with an odd number of lines. Successive fields begin at points separated in time by one-half a line interval. (a) The odd field. (b) The even field.



(a)



(b)

Fig. 3-24 (see caption on opposite page.)

time unit, the spot is carried back to the left-hand edge of the picture, point *C*, where the scan of line 3 begins. The deflection in the downward direction is the result of the vertical saw tooth, which increases linearly in time. This action continues until finally, at *D*, in the 23rd line interval the vertical retrace begins, and the scanning spot moves upward toward the top of the raster. Since the slope of the vertical saw tooth is greater during flyback than trace, the upward motion of the spot is faster than the downward motion.

By reference to the detail of the deflections in the diagram it may be seen that during the first three time units of the vertical flyback, the horizontal motion is in the trace direction, and the spot will be carried to *E*. During the next time unit a horizontal retrace occurs and the spot is carried to *F* at the left edge of the raster. At this point line intervals 1 through 23 are completed. The succeeding three time units, which belong to the first half of the line 25 interval (remember how the lines are numbered), occur during a horizontal trace and the spot is carried through to point *G*. At this point, which occurs at the middle of the line 25 *interval*, the even field begins, the downward motion starts again, and, as may be seen from the diagram, the even lines will interlace midway between the odd lines. Interlace is accomplished.

The scanning pattern for the even field is illustrated at *b* in the diagram. Particular note should be made of the manner in which this field ends. As may be seen from the detail of the deflections, the vertical trace ends at the end of line 22 trace, where the corresponding spot position in the raster is *H*. The vertical retrace interval is divided as follows: one time unit during a horizontal retrace to point *I*, 5 units to a horizontal trace to *J*, and the remaining unit to the horizontal retrace which carries the spot to *A*, where the entire raster begins to repeat itself for the second frame.

From this example it may be seen how the choice of an odd number of lines that are shared equally between two identical vertical saw-tooth deflections gives the desired 2 to 1 interlaced scanning raster. Since the duration of the odd line (25 in the example) is shared equally between the two fields, and the two vertical deflections are identical, the two fields have starting points which are displaced one-half line *interval* apart and proper interlace occurs. It must be stressed that the particular vertical flyback interval of six time units which was used in the example is not necessary to produce interlacing. The

vertical retrace may begin at any point during the duration of a line, the essential requirement being that the starting points of the two fields be spaced one-half line interval apart. This condition may be met if the number of lines is odd and if they are shared equally between two fields whose vertical deflections are identical.

An ingenious system has been developed for synchronizing the vertical and horizontal deflection generators so that these relationships are maintained. It is illustrated in block diagram in Fig. 3-25

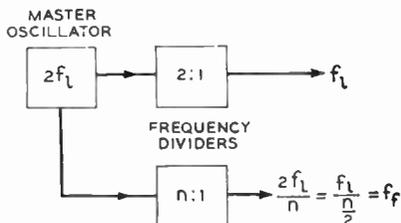


Fig. 3-25. Block diagram of the timing unit that provides 2 to 1 interlace. Both horizontal and vertical synchronizing pulses are derived from a common oscillator operating at twice the line frequency.

where f_l is the horizontal scanning, or line, frequency, and f_f is the vertical scanning, or field, frequency. It may be seen that the required one-half line interval relationship between fields is assured because the line and field frequencies are derived from a common source. Further details on the method of synchronization are covered in Chapter 11.

As was the case in progressive scanning it is necessary to blank out the scanning beam during the horizontal and vertical flyback intervals in order that the picture not be contaminated. Once again this blanking may be obtained by applying a negative square pulse of proper duration and frequency to the control grid of the CRT. In commercial practice it is usual to make the blanking pulse somewhat longer than the corresponding flyback interval to ensure that the sweep is under way before picture data are presented on the tube face. Fig. 3-26 shows the raster of our previous example as it would appear on a CRT when appropriate blanking signals are applied.

The equations relating the various quantities in interlaced scanning may now be written.

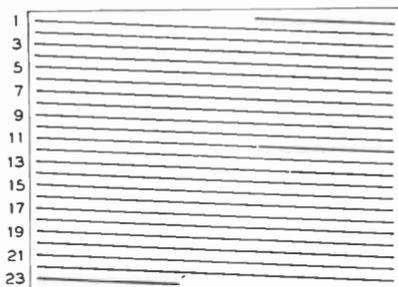


Fig. 3-26. The final raster produced by interlaced scanning. All flyback portions of the sweep in both the horizontal and vertical directions have been blanked out.

Let

$$\begin{aligned}
 V &= \text{field period} \\
 &= \frac{V_p}{2} = (\tau_u)_v + (\tau_f)_v \\
 &= \frac{1}{f_f} = \frac{1}{2f_p}
 \end{aligned}
 \quad \left. \vphantom{\begin{aligned} V &= \text{field period} \\ &= \frac{V_p}{2} = (\tau_u)_v + (\tau_f)_v \\ &= \frac{1}{f_f} = \frac{1}{2f_p} \end{aligned}} \right\} (3-30)$$

where f_f is the field frequency.

With reference to the blanking intervals, for interlaced scanning $(\tau_b)_v$ and $(\tau_u)_v$ are defined for one field and only $n/2$ lines occur in this interval; therefore, eq. (3-25) through (3-29) apply to interlaced geometry as well. One further equation may be stated which relates the line or horizontal scanning frequency to the frame frequency, namely,

$$f_l = nf_p \quad (3-31)$$

3-9. Even-line Interlace

We have just seen that interlace requirements may be met by using an odd number of scanning lines and a saw-tooth vertical deflection which repeats itself at twice the frame frequency. We shall now consider how the interlaced raster may be produced when an even number is used. It is apparent that the number of lines has no effect upon the horizontal deflection. The choice of an even n does affect the vertical deflection however. The effect is illustrated in Fig. 3-27 where, for the sake of simplicity, only eight lines are shown and zero flyback time is assumed for both deflections. The odd lines are scanned during the odd field, beginning at A and ending at B . Since the first field terminates at the end of a line interval, the new field

must start at the left edge of the raster. Thus, if line 2 is to fall mid-way between lines 1 and 3, it must begin at a lower level on the raster than point *A*. This may be accomplished only if the second vertical saw-tooth differs from the first; it must begin at level *c* rather than at

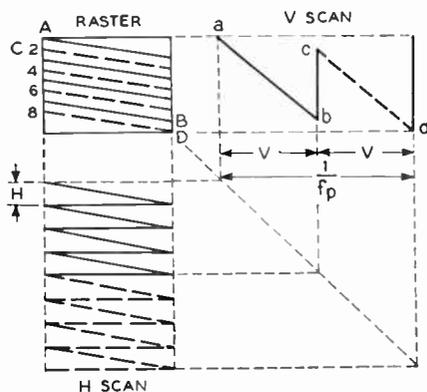


Fig. 3-27. Two-to-one interlace with an even number of lines. Notice that successive cycles of the vertical saw-tooth differ. Zero flyback in both directions is assumed.

level *a*. It is at once apparent, then, that interlace with an even number of lines requires a vertical saw-tooth in which alternate cycles differ, a condition which is severe on the vertical-deflection-generating circuits. It is for this reason that odd-line interlace is standard; by making *n* odd a vertical saw-tooth of twice the frame frequency produces the required scanning geometry.

3-10. Summary

We may summarize the results of this chapter in the following manner: The picture area may be scanned in either a progressive or interlaced raster. In either case both the horizontal and vertical deflections must have a saw-tooth shape. In the former system, the vertical saw-tooth must occur at frame frequency, whereas in interlaced scanning the vertical saw-tooth must be at twice the frame or field frequency. In magnetic deflection the deflection of the beam is proportional to the deflecting current which produces the magnetic field; in electric deflection, the deflection is proportional to the deflecting voltage. Therefore we must next consider means of generating currents and voltages of saw-tooth shape. This is the subject of the next chapter.

CHAPTER 4

SCANNING GENERATORS

We have seen in the last chapter that the scanning requirements are such that a saw-tooth wave form of voltage is required in electrostatic deflection and a saw-tooth wave form of current in magnetic deflection. In this chapter we shall discuss several of the various generator circuits which are commonly used to produce these waves. As a starting point we consider the simple series combination of resistance, capacitance, and e.m.f., which circuit forms the basis for all these generator circuits.

4-1. Transients in the R-C Circuit

Although the transient solutions for voltage and current in this simple circuit are well known, we shall review a few of the principal relationships, taking into account general boundary conditions.

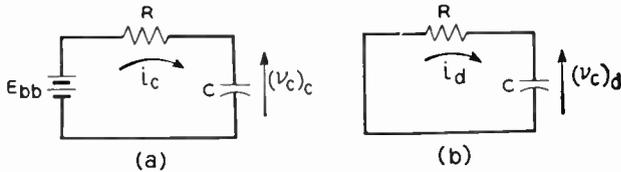


Fig. 4-1. The series R - C circuit is the basic circuit for saw-tooth generation. The condenser is assumed to have an initial voltage E_c .

Thus, in Fig. 4-1a, we have for a condenser uncharged at $t = 0$ and charging from the battery E_{bb}

$$(v_c)_c = E_{bb}(1 - e^{-t/T}) \quad (4-1)$$

$$i_c = \frac{E_{bb}}{R} e^{-t/T} \quad (4-2)$$

And similarly during discharge, where E_c is the initial voltage on C at $t = 0$,

$$(v_c)_d = E_c \epsilon^{-t/T} \tag{4-3}$$

$$i_d = -\frac{E_c}{R} \epsilon^{-t/T} \tag{4-4}$$

And in all four of the above equations

$$\left. \begin{aligned} \epsilon &= \text{Napierian base} \\ T &= \text{time constant of the circuit} \\ &= RC \text{ seconds} \end{aligned} \right\} \tag{4-5}$$

where R is given in ohms and C in farads.

Now, in the general case, at $t = 0$ and with a battery in the circuit the condenser will have an initial voltage E_c . Since the circuit is composed of linear, bilateral impedances, the solution for this general case may be obtained by use of the superposition theorem. Thus, by combining (4-1) and (4-3), when

$$v_c = E_c \quad \text{at} \quad t = 0$$

$$\text{then} \quad v_c = (v_c)_c + (v_c)_d = E_{bb} - (E_{bb} - E_c) \epsilon^{-t/T} \tag{4-6}$$

and, similarly, for the currents

$$i = i_c + i_d = \frac{E_{bb} - E_c}{R} \epsilon^{-t/T} \tag{4-7}$$

These equations show that the relative magnitudes of E_{bb} and E_c determine whether the condenser voltage increases or decreases. In

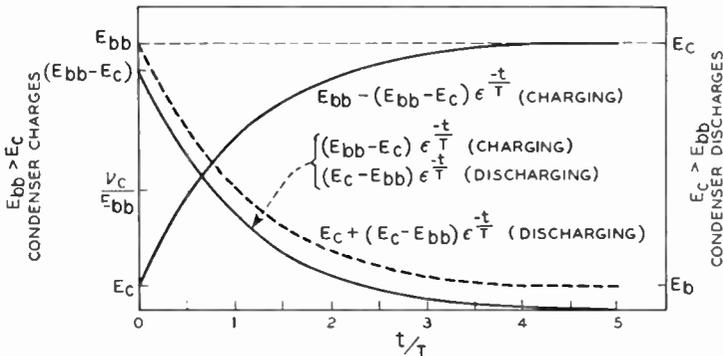


Fig. 4-2. Charge and discharge curves for the R - C circuit. E_c = initial condenser voltage, E_{bb} = battery voltage, T = circuit time constant.

either case, the voltage follows an exponential curve in time as does the current. Typical curves are shown in Fig. 4-2. From these it may also be seen that the change in condenser voltage or voltage swing is limited to the difference between battery and initial condenser voltages.

4-2. Generators of Saw-tooth Voltage

The principle of using a series $R-C$ circuit for the generation of a saw-tooth voltage may be stated qualitatively in the following terms: The initial build-up part of the charging curve of Fig. 4-2 looks "fairly linear." Thus, if the ratio of charging time, τ_c , to charging time constant, T_c , were sufficiently small, the build-up of voltage would be approximately linear. At $t = \tau_c$ the condenser must be discharged to furnish the flyback portion of the saw-tooth wave. In Fig. 4-3 is shown a simple switch-operated version of the saw-tooth voltage generator and the voltage developed across the condenser. At $t = 0$, S is opened and v_c builds up along the exponential curve a . At $t = t_1$, S is closed and v_c decreases exponentially

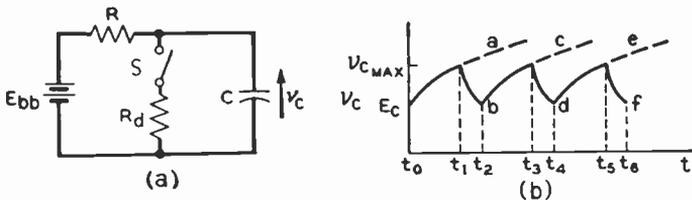


Fig. 4-3. The circuit is modified to produce saw-tooth waves. (a) A switch and a resistor, R_d , are shunted across the condenser. (b) An exponential saw-tooth voltage is developed across C when the switch is opened and closed periodically.

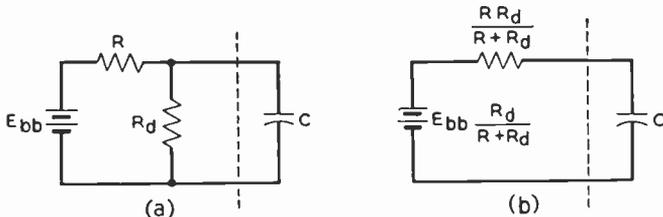


Fig. 4-4. The discharge circuit of Fig. 4-3 is simplified by application of Thevenin's theorem. (a) The equivalent discharge circuit. (b) The simplified discharge circuit.

along b until at $t = t_2$, S is opened again, and the cycle repeats itself. It is of importance to notice here that the charge and discharge intervals, τ_c and τ_d , respectively, are determined by the operation of the switch. It is of further importance for future work to notice the equivalent charge and discharge circuits. Since during charge S is open, the equivalent charge circuit is identical to that of Fig. 4-1. On discharge, however, S is closed, causing R_d to be shunted across the condenser. In order to simplify the analysis Thevenin's theorem may be applied to the battery circuit, with the result shown at b in Fig. 4-4. From this it may be seen that (1) E_c , the voltage at t_0, t_2, t_4 , etc., cannot be less than $E_{bb} \frac{R_d}{R + R_d}$, and (2) T_d , the discharge time constant, is given by

$$T_d = \left(\frac{RR_d}{R + R_d} \right) C \tag{4-8}$$

Since the charging time constant is given by

$$T_c = RC \tag{4-9}$$

it follows that the condenser discharges at a faster rate than it charges. Furthermore, the lower the value of R_d , the shorter will be the discharge-time constant. This makes for a greater loss of condenser voltage during the discharge interval.

$$\tau_d = t_2 - t_1 = t_4 - t_3 = \dots$$

Thus, while the circuit of Fig. 4-3 serves to illustrate the principle of saw-tooth voltage generation, two questions must be cleared up in order to make it a practical device: First, what limits are placed on the ratio of charge time to time constant, τ_c/T_c , to satisfy the linearity requirements of the television scan, and, second, what sort of automatic discharge device may be used to replace the switch S , shown in the diagram?

The question of linearity may be approached by expanding the exponential factor of eq. (4-6) in a power series. Thus the condenser voltage during the charge interval becomes

$$v_c = E_{bb} - (E_{bb} - E_c) \left[1 - \frac{t}{T_c} + \frac{1}{2} \left(\frac{t}{T_c} \right)^2 - \frac{1}{3} \left(\frac{t}{T_c} \right)^3 + \dots \right] \tag{4-10}$$

Then, removing the first term from the brackets and factoring out (-1) , we get

$$v_c = E_c + (E_{bb} - E_c) \left[\frac{t}{T_c} - \frac{1}{2} \left(\frac{t}{T_c} \right)^2 + \frac{1}{3} \left(\frac{t}{T_c} \right)^3 - \dots \right] \quad (4-11)$$

In (4-11) it may be seen that the first term in the brackets represents a linear build-up of voltage, whereas the remaining terms cause curvature. The problem then is to keep t/T_c sufficiently low so that these higher order terms will be negligible relative to the first.

It may also be shown that the first term gives a slope equal to the initial slope of the actual exponential curve at $t = 0$, for by differentiating (4-6) we get

$$\frac{dv_c}{dt} = \frac{1}{T} (E_{bb} - E_c) \epsilon^{-t/T} = \frac{E_{bb} - E_c}{T} \Big]_{t=0} \quad (4-12)$$

The first three terms of the power expansion are plotted in Fig. 4-5. Inspection of these curves shows that at least t/T_c must be restricted to values below 0.5. This restriction also simplifies calculations

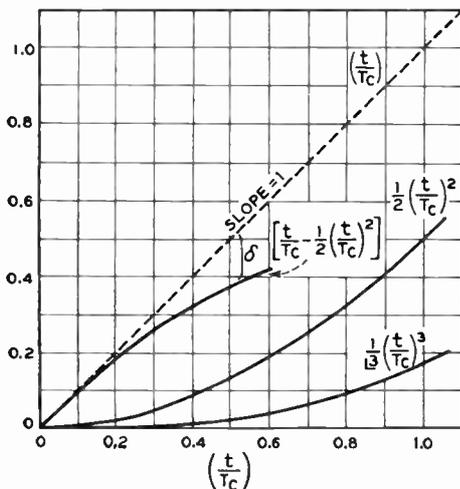


Fig. 4-5. The first three terms of eq. (4-11).

because lower values of the time ratio render the third and higher order terms negligibly small. Then, if we assume that the curvature of the voltage v_c time curve is the result of the second-order term alone, we may define the departure from linearity as the difference

between the ideal and actual curves expressed as a fraction of the available charging voltage ($E_{bb} - E_c$). This definition may be simplified as follows:

$$\delta = \frac{(E_{bb} - E_c) \left(\frac{t}{T_c} \right) - (E_{bb} - E_c) \left[\left(\frac{t}{T_c} \right) - \frac{1}{2} \left(\frac{t}{T_c} \right)^2 \right]}{(E_{bb} - E_c)} \\ = \frac{1}{2} \left(\frac{t}{T_c} \right)^2 100\% \quad (4-13)$$

By means of this equation we can determine the maximum permissible value of (t/T_c) for any given degree of linearity. For example, say that the actual curve is never to depart from a linear build-up by more than 2 per cent. Then, by (4-13), the limitation on (t/T_c) will be

$$\frac{1}{2} \left(\frac{t}{T_c} \right)^2 \leq 0.02 \quad \left(\frac{t}{T_c} \right) \leq 0.2$$

During this interval the corresponding change in condenser voltage which will be the sweep voltage will have the value

$$\Delta e_c \approx (E_{bb} - E_c) \left(\frac{t}{T_c} \right) \approx 0.2(E_{bb} - E_c)$$

This last equation shows one of the principal disadvantages of the R - C circuit as a sweep-voltage generator. The choice of a low value of (t/T_c) in the interests of linearity results in a poor conversion of available voltage ($E_{bb} - E_c$) to sweep voltage, Δe_c , the conversion ratio being approximately (t/T_c) .

To a certain extent this low value of conversion ratio is caused by our choice of the unity-slope line as the reference of linearity. An obvious question, then, is whether some other reference of linearity can be chosen that will allow a greater output voltage for the same departure from the reference. In following this line of thought Fig. 4-5 may be of help, for the actual build-up curve lies below the reference curve for all values of time other than zero. Why not choose as a reference some straight line which cuts across the exponential so that for part of the time the exponential is slightly above the reference and for part, below? Inspection of the curves shows that a reference line of slope between 0.8 and 0.9 would meet this requirement. A good compromise of 0.85 has been suggested by

Fink.¹ Using this as the reference of linearity, we may define δ' as the difference between the ideal (slope = 0.85) and actual curves expressed as a fraction of the available charging voltage. It follows at once that δ' is

$$\begin{aligned}\delta' &= 0.85 \left(\frac{t}{T_c} \right) - \left[\left(\frac{t}{T_c} \right) - \frac{1}{2} \left(\frac{t}{T_c} \right)^2 \right] \\ &= -0.15 \left(\frac{t}{T_c} \right) + \frac{1}{2} \left(\frac{t}{T_c} \right)^2\end{aligned}\quad (4-14)$$

δ' and its components are plotted in Fig. 4-6.

The allowable time ratio for a given δ' may now be calculated. We choose a δ' of ± 1 per cent to correspond to the $+2$ per cent of the

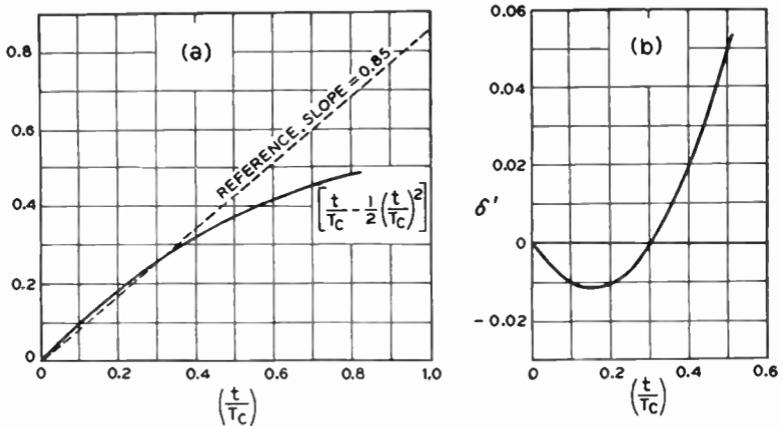


Fig. 4-6. (a) The first two terms of the charging curve, and the reference of linearity. (b) The departure from linearity plotted on an expanded scale.

previous example. This with (4-14) yields a value of (t/T_c) of approximately 0.4 and a corresponding output voltage swing of 0.33 ($E_{bb} - E_c$). Thus, by choosing a different reference line for the linearity check, we have obtained a gain of roughly 1.6 in the conversion ratio of available swing to sweep-voltage output.

A moment's reflection will show that even though δ' is within the

¹ D. G. Fink, *Principles of Television Engineering*. New York: McGraw-Hill Book Co., Inc., 1940.

limits of ± 1 per cent, the actual percentage error between the actual and reference curves may rise to over 5 per cent within the range $0 \leq \frac{t}{T_c} \leq 0.4$. Thus we shall consider the 0.4 figure to be the upper limit of time ratio permissible. Lowering this value purchases linearity at the expense of output voltage swing, but this is easily remedied by the addition of amplification.

We may now translate these results into a design condition. Since $\left(\frac{t}{T_c}\right)$ is not to exceed 0.4, for any given charging interval, τ_c , the circuit must have a time constant which satisfies the inequality

$$T_c = RC \geq \frac{\tau_c}{0.4} \quad (4-15)$$

We may now summarize the charge circuit by reviewing the design conditions. The v and h subscripts are omitted because the following equations apply to either of the two deflection systems, the horizontal or the vertical.

Given τ_b and τ_u for the system, τ_s and τ_f are chosen such that

$$\tau_s > \tau_u$$

and

$$\tau_s + \tau_f = \tau_b + \tau_u$$

Let

$$\tau_c = \tau_s \quad \text{and} \quad \tau_d = \tau_f \quad (4-16)$$

Then

$$T_c = RC \geq \frac{\tau_c}{0.4}$$

This specifies the R - C product.

It remains to consider the discharge part-cycle, which is easier to handle because no linearity requirement need be met. The chief concern is to make E_c as small as possible, and from Fig. 4-2 we see that if t/T_d is 5 or greater, this will be accomplished to all intents. Thus, if

$$\left. \begin{aligned} T_d &= \left(\frac{RR_d}{R + R_d} \right) C \leq \frac{\tau_d}{5} \\ \text{then } E_c &= E_{bb} \frac{R_d}{R + R_d} \end{aligned} \right\} \quad (4-17)$$

Judicious algebraic manipulation and the arbitrary choice of one of the variables permits solution for the other two. Typical values for R range from 100 kilohms to 1 megohm. A typical design problem

will be considered after we have found a means of replacing the manually operated discharge switch.

In the foregoing paragraphs the linear sweep is developed by using only a small portion of the build-up voltage across a charging condenser, the deviation from linearity being determined by the ratio of charge time to time constant. This system has the disadvantage that good linearity can be had only with loss of output voltage. An alternate approach to saw-tooth voltage generation overcomes this disadvantage but requires the addition of another vacuum tube in the circuit and may be termed a constant-current-charging circuit. In Fig. 4-7 the charging resistor of the previous circuit is replaced

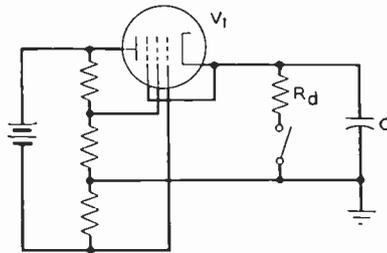


Fig. 4-7. Constant-current charging circuit that employs a voltage-saturated pentode.

by a voltage-saturated pentode that acts as a variable resistor or current limiter, which over a relatively large range of operating voltages passes a constant current I_c . Then, since the voltage across the condenser is the time integral of charging current,

$$v_c = \int_0^t I_c dt = I_c t \quad (4-18)$$

Thus the circuit delivers a linearly increasing voltage within the limits that V_1 can hold the charging current constant. The advantages of increased output voltage and linearity provided by this circuit do not sufficiently outweigh the simplicity of the series R - C circuit, however. Generally the output of Fig. 4-7 will not be sufficiently great to meet normal scan requirements and a stage or more of amplification is necessary. Because of this, the practice has been to use the series R - C generator with amplification to provide the saw-tooth sweep voltage required for electrostatic deflection.

4-3. Trigger Tubes

Since we have investigated the linearity requirements on the saw-tooth voltage generator, we must now answer the second question posed earlier in section 4-2, namely, what automatic circuit or device can be used to replace the discharge switch S shown in the last two circuit diagrams. Such a switching device must meet a minimum of three requirements:

- (1) It must have a relatively low internal resistance, R_a , so that the discharging time constant of the circuit will be low.
 - (2) It must be subject to control by an externally supplied synchronizing voltage.
 - (3) It must be positive in its "open" and "close" operations.
- These requirements are met in some degree by gas-filled triodes or thyratrons and by hard triodes. Consider first the thyatron trigger circuit.

In Fig. 4-8a the thyatron grid is biased low enough so that the maximum condenser voltage will not cause conduction. Under this condition, C will charge until the positive pulse of synchronizing

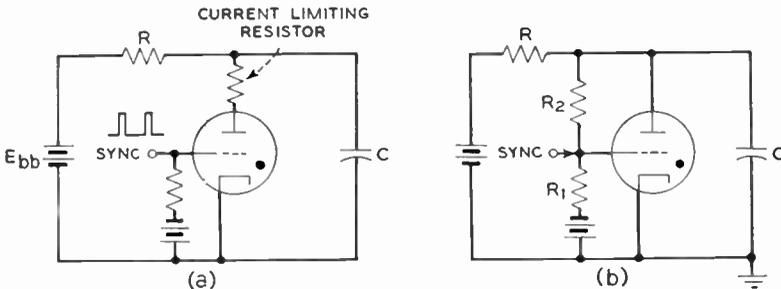


Fig. 4-8. (a) Thyatron trigger tube circuit. (b) Feedback-stabilized thyatron sweep circuit according to Koek. (Courtesy of *Electronics*.)

voltage allows the arc to be established. This conduction allows C to discharge until its voltage drops to the extinction voltage for the gas-triode, causing the arc to be extinguished, and the charge cycle is reinitiated. The current limiting resistor is required to protect the thyatron from excess current during the condenser discharge. Clearly, in this case the charging time is determined by the inter-

synchronizing pulse interval and the length of time required for the arc to extinguish.

It follows, therefore, that the thyatron meets the first two requirements for a satisfactory trigger tube but unfortunately it falls short of the third requirement. The crossover from conduction to non-conduction in such a gas-filled tube takes a finite length of time, which may be explained on the following basis. Once the arc has been established by the grid going above its critical value, the gas molecules in the tube envelope ionize and the arc is established. At the end of the synchronizing pulse the grid returns to a negative value determined by the bias and it becomes sheathed by a layer of positive ions. Thus the grid cannot regain control of the tube until these sheath ions have recombined with free electrons to form neutral gas molecules. These positive ions have relatively large mass and diffuse slowly to recombine with the electrons. The minimum time in which this recombination occurs after the reduction of anode voltage is termed the deionization time.²

It can be seen, therefore, that the off-position of the switch is not controlled by the termination of the synchronizing pulse, but depends also upon the time required for v_c to drop to the deionization voltage of the tube and upon the tube's deionization time. This latter also depends upon gas pressure, temperature, and aging. Thus, positive control of the switching action by the synchronizing voltage is not provided.

A further complication rises in that the deionization time of typical tubes used for this service, such as the 884 and 6Q5, ranges in the order of 10 microseconds or more and may become a considerable fraction of the total period for one scan. These disadvantages have rendered obsolete the use of gas-filled tubes in scanning generators for television service.

It should be mentioned, however, that considerable improvement in the stability of a gas-tube triggered sweep circuit may be obtained by the application of feedback. Kock³ has described such a circuit in which plate-to-grid feedback is provided by the resistor R_2 , shown in Fig. 4-8b. If R_1 and R_2 are approximately equal, the ignition

² J. Millman and S. Szeley, *Electronics*. New York: McGraw-Hill Book Co., Inc., 1941.

³ W. E. Kock, "A Stabilized Sweep Circuit Oscillator." *Electronics*, 12, 4 (April 1939).

potential of the tube is practically independent of such factors as electrode spacing and gas pressure in the tube. Koek also claims a decrease in deionization time caused by the negative throw applied to the grid by the feedback circuit at the instant of conduction.

Since the prime objections to the gas-filled trigger tube may be traced back to phenomena associated with the gas itself, it seems that a more desirable approach to the problem would be to replace the gas tube with a vacuum tube. This is the procedure which is normally followed, a typical circuit being shown in Fig. 4-9a. In essence the

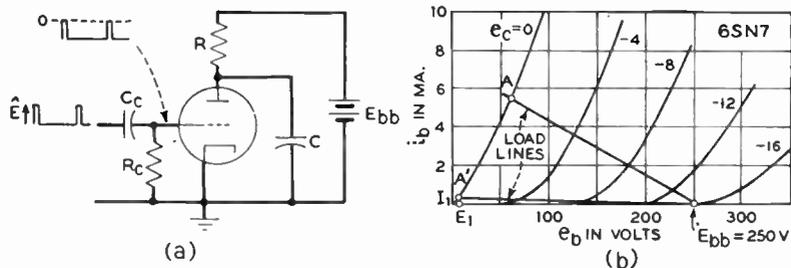


Fig. 4-9. (a) Saw-tooth generator employing a vacuum discharge tube. (b) Determination of the equivalent discharge resistance of the tube. The lower load line is for the illustrative example on page 90.

operation of the circuit is this: The discharge tube is normally biased beyond cutoff, and C charges from E_{bb} through R . The build-up of sweep voltage across the condenser is terminated by driving the tube into conduction, which allows the condenser to discharge through the tube itself. The duration of the condenser discharge is controlled wholly by the tube conduction and hence by the positive driving pulse on the grid of the tube. In order to analyze the action of the circuit carefully, we must consider what takes place first in the grid circuit and then in the plate circuit.

Notice that we have required in the above description that the grid be biased beyond cutoff and yet the circuit shows the grid return connected to the cathode. Actually the necessary bias is developed by a clamping action of the grid circuit. Consider the following action: A positive pulse is applied at the input side of the coupling condenser, C_c . Since no bias is present, the pulse drives the grid positive, causing grid current to flow. Also, during the positive

pulse, C_c charges through the parallel combination of R_c and the grid-to-cathode conduction resistance of the tube.

At the end of the positive pulse, grid current stops flowing and C_c discharges. Since the grid-to-cathode resistance is infinite in this condition, C_c discharges at a slower rate until the next positive pulse occurs and the entire cycle repeats itself. Notice that the discharge current flows in such a direction that the voltage drop across R_c makes the grid negative relative to the cathode. After a sufficient number of cycles have occurred, a steady state or repetitive condition exists, such that the charge lost by C_c between pulses is equal to the charge gained during a pulse. This steady state results in a negative bias across R_c of such a magnitude that the grid is held at nearly zero volts during the positive applied pulse. Thus the flow of grid current develops the required cutoff bias and the positive pulse is clamped at nearly zero volts on the grid. The action may be summarized this way: The applied positive, triggering pulse causes the grid to swing between cutoff and zero voltage, thereby controlling the charge and discharge of the R - C sweep-generating circuit. Then, if the pulse amplitude, E , is equal to or greater than the magnitude of the cutoff bias for the tube, the latter will be biased below cutoff during the interpulse periods.

Returning to the action in the plate circuit we require the value of the discharge resistance, R_d , which is the equivalent plate-to-cathode resistance of the tube when it is conducting. We must now consider how its value may be determined. Since the grid voltage remains at approximately zero during the entire conduction interval of the tube, and since E_{bb} and R are known, the point A , the intersection of the d-c load line and the $e_c = 0$ static plate characteristic curve, may be determined as shown in Fig. 4-9b. Then R_d is the ratio of E_1 and I_1 , the voltage and current at that point.

We may now consider the design of a typical sweep-generating circuit to illustrate our work this far:

Design a horizontal saw-tooth sweep generator where the line interval is 63.5 microseconds and the flyback ratio 1 to 19. The linearity is to remain within the limits ± 1 per cent. A supply voltage of 250 volts is available and a 6SN7 is used as the discharge tube.

We first calculate the values of the R - C charging circuit. Thus, by the specifications,

$$H = \tau_s + \tau_f = \tau_s(1 + p)$$

$$\text{or} \quad \tau_s = \frac{H}{1 + p} = 63.5 \left(\frac{19}{20} \right) = 60.4 \text{ microseconds}$$

$$\text{then} \quad T_c = RC = \frac{\tau_s}{0.4} = \frac{60.4}{0.4} = 151 \text{ microseconds}$$

As a starting point let us assume a value of C , thus,

$$C = 100 \text{ micromicrofarads}$$

$$\text{then} \quad R = \frac{151 \times 10^{-6}}{10^{-10}} = 1.51 \text{ megohms}$$

We must now check the discharge part-cycle. To do this we first calculate the static plate resistance of the discharge tube. This calculation is illustrated at b in Fig. 4-9, where a 1.51-megohm load line is drawn. Then, from the diagram,

$$R_d = \frac{E_1}{I_1} = \frac{8 \times 10^3}{0.2} = 40 \text{ kilohms}$$

The discharge time constant will be:

$$T_d = R_d C = (4 \times 10^4)(10^{-10}) = 4 \text{ microseconds}$$

It is at once apparent that the discharge interval, τ_f , is not five or more times greater than the discharge time constant, T_d ; therefore we cannot assume that the condenser discharges to $E_{bb}R_d/(R + R_d)$ as might be expected from eq. (4-17). On the contrary, the final voltage on the condenser at the end of each discharge part-cycle will increase until a repetitive steady state condition is reached. Once this repetitive condition is reached, the final voltage at the end of a charge part-cycle, namely

$$V_{bb} - (E_{bb} - E_c)\epsilon^{-\tau_s/T_c}$$

must be the initial condenser voltage at the beginning of the next discharge part-cycle. Thus, we may equate these two boundary values and obtain by means of eq. (4-6)

$$E_c = E_{bb} \frac{R_d}{R + R_d} - \left\{ E_{bb} \frac{R_d}{R + R_d} - [E_{bb} - (E_{bb} - E_c)\epsilon^{-\tau_s/T_c}] \right\} \epsilon^{-\tau_f/T_d}$$

whence

$$E_c = \frac{E_{bb}}{(1 - \epsilon^{-(\tau_s/T_c + \tau_f/T_d)}) \left[\frac{R_d}{R + R_d} (1 - \epsilon^{-\tau_f/T_d}) + \epsilon^{-\tau_f/T_d} - \epsilon^{-(\tau_s/T_c + \tau_f/T_d)} \right]}$$

where E_c is the minimum voltage to which the condenser discharges.

Substitution of the various quantities yields the value

$$E_c = 60 \text{ volts}$$

and the total sweep voltage swing will be

$$\Delta e_c \approx .33(E_{bb} - E_c) \approx 62.8 \text{ volts}$$

At this point a digression on the philosophy of synchronization is in order. In our preliminary discussions on picture transmission it was shown that some sort of synchronizing signal is required in the picture system to ensure lock-in of the scan at pickup and playback ends. Extending this idea we can anticipate that both horizontal-scan and vertical-scan synchronizing signals are required because these two components take place more or less independently. Let us assume for the discussion that a train of properly spaced square pulses of required duration is available to synchronize one of the sweep-voltage generators described above. Synchronization could then be accomplished by application of these pulses directly to the control grid of the trigger tube used.

Although such a proposal is entirely feasible it does not find general acceptance, particularly in television broadcasting systems where radio transmission is involved. In this type of system two probable difficulties are present. First, if because of fading or some other phenomenon synchronization signals are not received for an interval, the sweep circuit cannot discharge. This means an undeflected spot on the cathode-ray tube, resulting in burning of the screen.

Secondly, if randomly spaced noise pulses of the same polarity as the sync are picked up in transmission, the generator may be fired at incorrect intervals with a disruption of the correct scanning pattern.

In an attempt to minimize these difficulties, it is common practice to drive the grid of the trigger tube from the output of some form of impulse oscillator, which, in turn, is held to the correct frequency by the synchronizing signals. If this device be used, loss of sync will upset the scanning pattern but the spot will continue to be deflected

across the face of the cathode-ray tube at a slower rate determined by the unsynchronized oscillator and no screen burn will result.

In reference to the second difficulty, in the postwar period several forms of essentially long-time-constant oscillator circuits have been developed, which have greatly increased the immunity to noise impulses. These circuits are of a special nature and are discussed in connection with commercial telecasting receivers.

The impulse oscillators used for driving the sweep generator find several other uses in the television system. We shall, however, discuss two of the principal types in the present chapter so that the design of an entire sweep channel may be covered as a unit. The first of these impulse generators to be discussed is the multivibrator.

4-4. Multivibrator

The basic circuit of the multivibrator, shown in Fig. 4-10, consists of a two-stage resistance coupled amplifier employing regeneration, but the circuit constants are so chosen that a pulse output is derived.

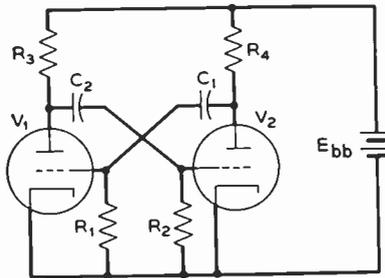


Fig. 4-10. The multivibrator.

Whereas the operation of the circuit may be described on the basis of a regenerative amplifier, a more powerful approach, developed by Shenk,⁴ which yields design equations, is that which considers the circuit as comprising two switches. On this basis assume that V_1 is conducting and that C_2 , which has been charged previously, is discharging through V_1 . The return circuit for this discharge current is through R_2 and a negative voltage is developed across R_2 , which appears on the grid of V_2 . This state, during which V_1 is conducting

⁴ E. R. Shenk, "The Multivibrator," Parts I, II, and III. *Electronics*, 17, 1, 2, and 3 (January, February, March 1944).

and V_2 is cut off, prevails until the decrease in current through R_2 causes the voltage developed on the grid of V_2 to reach cutoff, when V_2 conducts. The resulting drop in plate voltage of V_2 is coupled through C_1 to V_1 , driving it below cutoff. Then a similar state prevails, except that V_2 and V_1 have interchanged rôles. At the instant when V_2 begins to conduct, C_2 begins to charge through R_2 , causing the grid of V_2 to go positive. It is this positive pulse which may be used to trigger the discharge tube of the sweep-generating circuit. These various statements may be stated precisely in mathematical form and can be made to yield a set of equations that permit design of a multivibrator to give an output pulse of specified width at a specified frequency or pulse-repetition rate. The basis of the design, then, is that we shall assume two states of operation in the circuit which last for τ_1 and τ_2 , respectively, where

$$\left. \begin{aligned} \tau_1 &= \text{interval during which } V_1 \text{ is cut off and } V_2 \text{ conducts} \\ \text{and} \\ \tau_2 &= \text{interval during which } V_2 \text{ is cut off and } V_1 \text{ conducts} \end{aligned} \right\} \quad (4-19)$$

The change from one state to the other will be assumed to take place instantaneously. Thus each tube is in one of two *static* conditions, conduction at fixed plate voltage and essentially zero grid voltage, and nonconduction with the grid biased below cutoff.

As an initial condition we shall assume the circuit is in the τ_2 part-cycle, *i.e.*, V_2 is cut off and V_1 is conducting. Also assume that V_2 has been cut off for a sufficiently long interval that C_1 is charged up to the full supply voltage E_{bb} . Some transient causes the switching action to occur, and the plate voltage of V_2 drops. Since this causes the voltage on the V_2 side of C_1 to decrease, C_1 will start to discharge through the equivalent circuit, given in Fig. 4-11. The arrow direc-

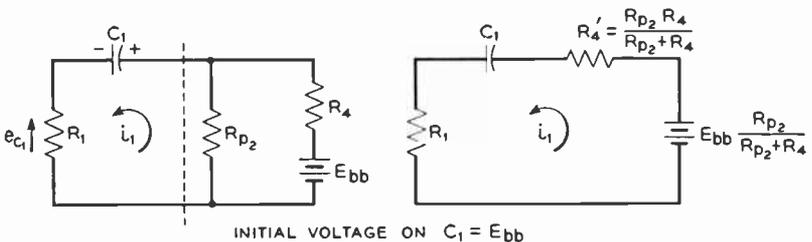


Fig. 4-11. Equivalent discharge circuit of C_1 during the interval τ_1 when V_1 is cut off.

tion of current is chosen to be the same as that defined for the series R - C circuit in section 4-1, and

$$R_{p2} = \text{equivalent static-plate resistance of } V_2 \quad \text{during conduction.} \quad (4-20)$$

At b in the diagram the battery and resistance network to the right of the dotted line has been replaced by its equivalent according to Thevenin's theorem. Then, since the circuit is a simple series R - C network, eq. (4-7) may be applied directly and we have

$$\begin{aligned} i_1 &= \left[E_{bb} \frac{R_{p2}}{R_{p2} + R_4} - E_{bb} \right] \frac{\epsilon^{-t/T_1}}{R_1 + R_4'} \\ &= - \frac{E_{bb}}{R_1 + R_4'} \left[\frac{R_4}{R_{p2} + R_4} \right] \epsilon^{-t/T_1} \end{aligned} \quad \left. \vphantom{\begin{aligned} i_1 &= \left[E_{bb} \frac{R_{p2}}{R_{p2} + R_4} - E_{bb} \right] \frac{\epsilon^{-t/T_1}}{R_1 + R_4'} \\ &= - \frac{E_{bb}}{R_1 + R_4'} \left[\frac{R_4}{R_{p2} + R_4} \right] \epsilon^{-t/T_1} \end{aligned}} \right\} (4-21)$$

where $T_1 = \text{discharge time constant for } C_1$
 $= (R_1 + R_4')C_1$

and $R_4' = \frac{R_{p2}R_4}{R_{p2} + R_4}$

By inspection of the circuit diagram the grid voltage on V_1 is the iR drop across R_1 . Thus,

$$\begin{aligned} e_{c1} &= \text{grid voltage on } V_1 \\ &= - \frac{E_{bb}R_1}{R_1 + R_4'} \cdot \frac{R_4}{R_{p2} + R_4} \epsilon^{-t/T_1} \\ &= - E_{bb}k_1 \epsilon^{-t/T_1} \end{aligned} \quad (4-22)$$

where

$$k_1 = \frac{R_1}{R_1 + R_4'} \cdot \frac{R_4}{R_{p2} + R_4} \approx \frac{1}{1 + \frac{R_{p2}}{R_4}} \quad \text{for generally } R_1 \gg R_4' \quad (4-23)$$

Equation (4-22) shows that the grid voltage on V_1 starts at $(-k_1E_{bb})$ and increases (*i.e.*, becomes less negative) exponentially in time.

If E_{c01} is the cutoff voltage of the tube, at

$$t = \tau_1 \quad e_{c1} = E_{c01} \quad (4-24)$$

and V_1 conducts, simultaneously terminating the τ_1 part-cycle and initiating the τ_2 part-cycle. Substitution of the terminal condition in (4-22) yields

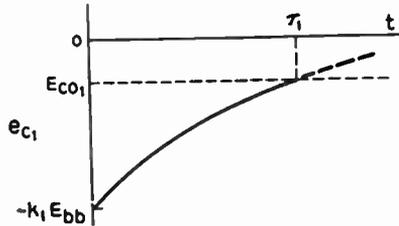


Fig. 4-12. The grid voltage on V_1 increases exponentially toward zero until cutoff voltage is reached at $t = \tau_1$.

$$E_{c01} = -k_1 E_{bb} e^{-\tau_1/T_1} \tag{4-25}$$

Simplification of the design equations results if we define a cutoff amplification factor

$$\mu_{c01} = -\frac{E_{bb}}{E_{c01}} \tag{4-26}$$

Now, if (4-25) and (4-26) be combined and the natural logarithm taken of both sides, there results

$$T_1 = \frac{\tau_1}{\ln(k_1 \mu_{c01})} \tag{4-27}$$

Now subject to the same assumption that C_2 starts to discharge from an initial voltage equal to the supply voltage E_{bb} , we may write down a similar set of equations which govern the discharge of C_2 during the τ_2 part-cycle. Thus we have

$$\begin{aligned} T_2 &= \text{discharge time constant for } C_2 & (4-28) \\ &= (R_2 + R_3')C_2 \end{aligned}$$

$$R_3' = \frac{R_3 R_{p1}}{R_3 + R_{p1}} \tag{4-29}$$

R_{p1} = equivalent static plate resistance of V_1 during conduction

$$\begin{aligned} k_2 &= \frac{R_2}{R_2 + R_3'} \cdot \frac{R_3}{R_{p1} + R_3} \\ &\approx \frac{1}{1 + \frac{R_{p1}}{R_3}} \quad \text{for generally } R_2 \gg R_3' \end{aligned} \tag{4-30}$$

μ_{co2} = cutoff amplification factor of V

$$= - \frac{E_{bb}}{E_{co2}} \tag{4-31}$$

and

$$T_2 = \frac{\tau_2}{\ln(k_2\mu_{co2})} \tag{4-32}$$

The frequency or pulse repetition rate of the output will be

$$f = \frac{1}{\tau_1 + \tau_2} \tag{4-33}$$

Notice that all of the quantities in the above equations are fixed circuit parameters or time intervals, which are part of the design specifications, except for the static plate resistances, R_{p1} and R_{p2} , and the cutoff amplification factors, μ_{co1} and μ_{co2} . These quantities may be determined to a sufficient degree of accuracy from the static plate characteristics of the tubes used in the manner indicated in Fig. 4-13.

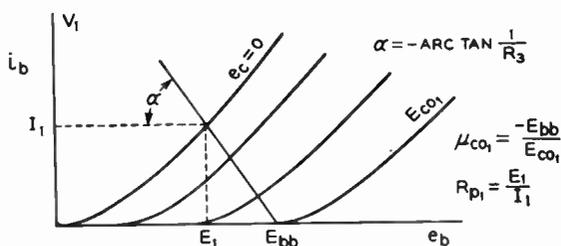


Fig. 4-13. Determination of R_{p1} and μ_{co1} from the static plate characteristics of V_1 .

μ_{co1} is the ratio of supply voltage to grid voltage which produces cutoff. Some philosophical arguments may be raised against this method of determining cutoff amplification factors because there is some doubt as to what value of plate current may be called cutoff. Stated more precisely, shall we say that the tube is conducting when 1 microampere flows or 500 microamperes or some other value? Since we have assumed that the change over from nonconduction to steady-state conduction condition is instantaneous, some sort of compromise is needed. Shenk has presented data to cover this, but for most design problems the method indicated in the figure for calculating μ_{co1} is sufficiently accurate.

R_{p1} is evaluated on the basis that V_1 conducts with approximately zero volts on its grid and with a plate voltage E_1 equal to the supply voltage minus the drop in the plate load resistance, or

$$E_1 = E_{bb} - I_1 R_3 \quad (4-34)$$

then
$$R_{p1} = \frac{E_1}{I_1} \quad (4-35)$$

where both E_1 and I_1 may be determined from the plate characteristics when E_{bb} and R_3 are given. It should be apparent that R_{p2} and μ_{co2} are evaluated in the same manner from the plate characteristics of V_2 . Generally V_1 and V_2 are identical halves of a twin triode such as the 6SN7 or the 6SL7, in which case μ_{co1} and μ_{co2} will be identical. The static plate resistances will also be equal, provided that R_3 and R_4 , the plate load resistances, are the same.

Now it might appear that eqs. (4-21) through (4-35) can be manipulated to permit design of the multivibrator for given values of τ_1 and τ_2 . If this were true, it would appear that some control could be had over the shape of the output waves because the ratio of τ_1/T_1 , for example, determines the shape of e_{c1} . This is illustrated in Fig. 4-14.

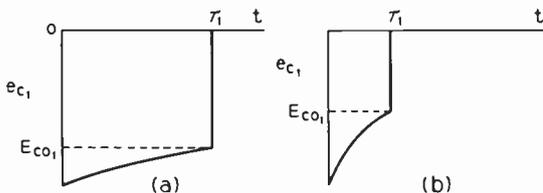


Fig. 4-14. Effect of τ_1/T_1 on the wave form of grid voltage.

- (a) τ_1/T_1 small;
 (b) τ_1/T_1 large.

Unfortunately the statements of the last paragraph are not true because all of the design equations presented above presume that C_1 is charged to E_{bb} at the beginning of the τ_1 part-cycle, and C_2 is charged to E_{bb} at the beginning of τ_2 . These two assumptions must be met or the foregoing equations are invalid. Hence we now investigate the charging of C_1 during the interval τ_2 . It will be seen that this will place a restriction on the upper limit of the product $(R_{p1} + R_4)C_1$.

First we note that the final charge on C_1 at the end of τ_1 must be equal to the initial charge on C_1 at the beginning of the τ_2 interval.

This transition value of condenser voltage V_{c1} may be obtained by the application of Kirchhoff's voltage law to Fig. 4-11. R_4' will generally be negligibly small compared to R_1 , thus at $t = \tau_1$

$$V_{c1} = E_{bb} \frac{R_{p2}}{R_{p2} + R_4} - E_{co1}$$

= initial voltage on C_1
at the beginning of the τ_2 part-cycle. (4-36)

During τ_2 the equivalent charge circuit for C_1 is that of Fig. 4-15. The direction of the charge current i_2 is such that e_{c1} goes slightly positive and grid current flows in the tube. R_{g1} is the equivalent

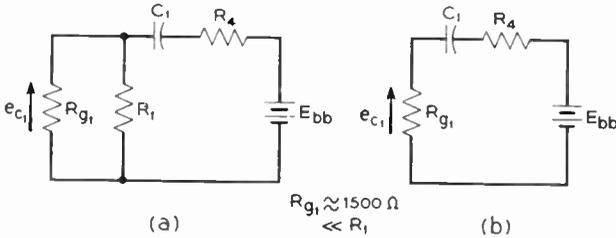


Fig. 4-15. Equivalent charge circuit of C_1 during the interval τ_2 when V_2 is cut off.

static grid resistance of the tube and generally is in the order of 1500 ohms. R_1 is usually very much greater and Fig. 4-15b shows the approximate circuit where the shunting effect of R_1 across R_{g1} is neglected.

Since the resulting circuit comprises resistance and capacitance in series with a battery, we may utilize eq. (4-6), which yields

$$v_c = E_{bb} - (E_{bb} - V_{c1})e^{-t/(R_{g1}+R_4)C_1}$$
(4-37)

But our design assumption is that at $t = \tau_2$, $v_c = E_{bb}$, hence

$$E_{bb} = E_{bb} - (E_{bb} - V_{c1})e^{-\tau_2/(R_{g1}+R_4)C_1}$$

or
$$\left[k_1 - \frac{1}{\mu_{co1}} \right] e^{-\tau_2/(R_{g1}+R_4)C_1} = 0$$
 (4-38)

Theoretically this condition cannot be satisfied unless the time constant for the circuit is zero or τ_2 is infinite. These conditions are impossible to satisfy but a good engineering compromise is possible

because the term in the brackets will always be less than unity. Thus, if $\frac{\tau_2}{(R_{\rho 1} + R_4)C_1} \geq 5$ the left-hand member of eq. (4-38) will not exceed 0.01. This will be assumed satisfactory. Hence, a maximum value of C_1 may be specified

$$C_{1 \max} = \frac{\tau_2}{5(R_{\rho 1} + R_4)} \quad (4-39)$$

and, similarly for the other half-circuit,

$$C_{2 \max} = \frac{\tau_1}{5(R_{\rho 2} + R_3)} \quad (4-40)$$

We see, then, that the intervals τ_1 and τ_2 are determined by the charging time constants T_1 and T_2 , respectively. But since C_1 must charge to supply voltage while C_2 is discharging during τ_1 and *vice versa*, maximum values for the capacitances are limited by the time available for them to recharge.

The equations which have been derived may now be collected into a design procedure.

4-5. Design Procedure for the Multivibrator

Given τ_1 , τ_2 , and E_{bb} .

(1) Pick a tube, generally a 6SN7 or 6SL7.

(2) From E_{bb} and the tube characteristics calculate μ_{c01} and μ_{c02} .

Where double triodes are used, these two values are equal.

(3) Choose R_4 , generally not to exceed 50 to 100 kilohms.

(4) Calculate
$$R_{p2} = \frac{E_2}{I_2} \quad (4-35)$$

(5) Calculate
$$R_4' : \frac{1}{R_4'} = \frac{1}{R_{p2}} + \frac{1}{R_4} \quad (4-21)$$

(6) Calculate
$$k_1 = \frac{1}{1 + \frac{R_{p2}}{R_4}} \quad (4-23)$$

(7) Calculate
$$T_1 = \frac{\tau_1}{\ln(k_1\mu_{c01})} \quad (4-27)$$

(8) Since R_1 must be at least 10 kilohms or more in order to validate the assumption that it has negligible shunting effect on $R_{\rho 1}$, choose $R_1 \geq 10$ kilohms. Then

$$C_1 = \frac{T_1}{(R_1 + R_4')} \tag{4-21}$$

(9) Check that C_1 is less than $C_{1 \max}$ which is given by

$$C_{1 \max} = \frac{\tau_2}{5(R_{\phi 1} + R_4)} \tag{4-39}$$

If this condition is not satisfied, assume a larger value for R_1 and repeat step (8). The right-hand section of the multivibrator is then completed. The procedure is then repeated for the left-hand section.

In regard to the latter it is sometimes possible to choose $R_3 = R_4$ and $R_2 = R_1$ and a simple set of equations for the left-hand section results. Thus if

$$R_3 = R_4 \quad \text{and} \quad R_2 = R_1$$

$$k_2 = k_1$$

and, from (4-27),

$$\frac{T_2}{T_1} = \frac{\tau_2}{\tau_1}$$

From (4-21)
$$C_2 = C_1 \frac{T_2}{T_1} = C_1 \frac{\tau_2}{\tau_1} \tag{4-41}$$

Since the same tube types are generally used in multivibrator circuits, some simplification may be realized by using design curves for the particular tube used. For example, Fig. 4-16a shows the

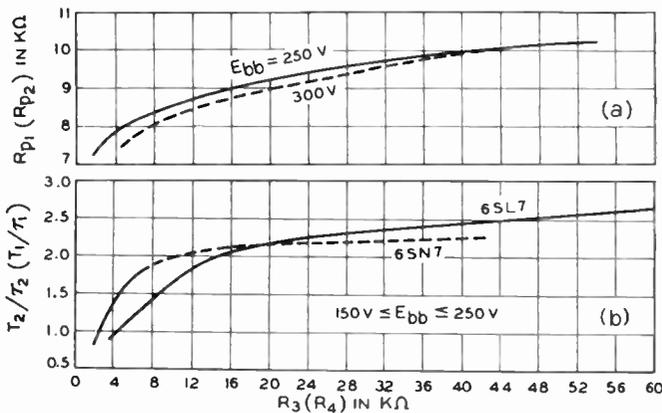


Fig. 4-16. Design curves for the multivibrator. (a) R_p for a 6SN7 (dashed line) and a 6SL7. (b) T/τ for the same tubes, after Shenk. (Courtesy of *Electronics*.)

variation of R_{p2} with R_4 for a 6SN7. For a given E_{bb} , μ_{co1} is constant and $k\mu_{co1}$ is a function of R_4 alone. Thus, from (4-27), T_1/τ_1 may be plotted against R_4 for a given tube. Curves of this type, developed by Shenk, are shown in Fig. 4-16*b*. In these curves Shenk assumes that k_1/R_4 remains constant over the normal range of E_{bb} , as indicated in the figure.

It must be stressed that the design procedure just outlined neglects the effects of shunt capacitance and hence takes no account of the rise time of the output wave forms. Furthermore the procedure does not allow designing for a given output pulse amplitude because it is based on a somewhat artificial and nonessential condition that both C_1 and C_2 charge up to the full battery voltage E_{bb} . This condition does simplify the design equations and certainly the method is useful in illustrating the manner in which design of the multivibrator may be handled.

Let us now apply our results to a typical problem: Design a free-running multivibrator for which τ_2 and τ_1 are 70 and 5 μsec , respectively. A 6SN7 is to be used and a plate supply voltage of 250 volts is available.

(1) Choose a 6SN7.

(2) From Fig. 4-9, $\mu_{co1} = \mu_{co2} = \mu_{co} = \frac{250}{16} = 15.6$.

(3) Since the time intervals involved are comparatively short, we choose a value of R_4 on the low side of the recommended limits; hence, let $R_4 = 40$ kilohms.

(4) For this example Fig. 4-16*a* may be used to determine R_{p2} .

$$R_{p2} = 9.9 \text{ kilohms}$$

$$(5) \frac{1}{R_4'} = \frac{1}{R_{p2}} + \frac{1}{R_4} = \left(\frac{1}{9.9} + \frac{1}{40} \right) 10^{-3}$$

$$= (0.101 + 0.025)10^{-3} = 0.126 \times 10^{-3}$$

$$R_4' = 7.94 \text{ kilohms}$$

$$(6) k_1 = \frac{1}{1 + \frac{R_{p2}}{R_4}} = \frac{1}{1 + \frac{9.9}{40}} = \frac{1}{1.247} = 0.802$$

$$k_2\mu_{co} = (0.802)(15.6) = 12.5$$

$$(7) T_1 = \frac{\tau_1}{\ln(k_1\mu_{co})} = \frac{5 \times 10^{-6}}{\ln 12.5} = \frac{5 \times 10^{-6}}{2.52} = 1.985 \mu\text{sec}$$

$$(8) \text{ Let } R_1 = 10 \text{ kilohms}$$

$$\text{Then } C_1 = \frac{T_1}{R_1 + R_4'} = \frac{1.985 \times 10^{-6}}{(10 + 7.94)10^3} = \frac{1.985 \times 10^{-6}}{1.794 \times 10^4} = 110 \mu\mu\text{f}$$

$$(9) C_{1 \text{ max}} = \frac{\tau_2}{5(R_{\rho 1} + R_4)} = \frac{70 \times 10^{-6}}{5(1.5 + 40)10^3}$$

$$= \frac{70 \times 10^{-6}}{5(4.15)10^4} = 337 \mu\mu\text{f}$$

Hence the calculated value of C_1 is satisfactory.

For the left-hand portion of the circuit,

$$(10) \text{ Let } R_3 = R_4$$

$$\text{and } R_2 = R_1$$

$$\text{Then } C_2 = C_1 \frac{\tau_2}{\tau_1} = 110 \left(\frac{70}{5} \right) = 0.00154 \mu\text{f}$$

Since $R_{\rho 1} = R_{\rho 2}$, it is clear that

$$C_{2 \text{ max}} = C_{1 \text{ max}} \frac{\tau_1}{\tau_2}$$

$$= 337 \left(\frac{5}{70} \right) = 24.5 \mu\mu\text{f}$$

Notice that in this case the simplified design procedure fails (as it usually will unless $\tau_1 \approx \tau_2$) and we must repeat the longer design.

(10a) Since $\tau_1 < \tau_2$ we must find some means of raising the value of $C_{2 \text{ max}}$. We see from eq. (4-40) that this may be accomplished by choosing R_3 less than its previous value. Hence let

$$R_3 = 20 \text{ kilohms}$$

$$(11) \text{ From Fig. 4-16a, } R_{\rho 1} = 9.25 \text{ kilohms}$$

$$(12) \frac{1}{R_3'} = \frac{1}{R_{\rho 1}} + \frac{1}{R_3} = \left(\frac{1}{9.25} + \frac{1}{20} \right) 10^{-3}$$

$$= (0.108 + 0.050)10^{-3} = 0.158 \times 10^{-3}$$

$$R_3' = 6.33 \text{ kilohms}$$

$$(13) k_2 = \frac{1}{1 + \frac{R_{\rho 1}}{R_3}} = \frac{1}{1 + \frac{9.25}{20}} = \frac{1}{1.452} = 0.688$$

$$k_{2\mu\text{co}} = (0.688)(15.6) = 10.72$$

$$(14) T_2 = \frac{\tau_2}{\ln(k_2\mu_{co})} = \frac{70 \times 10^{-6}}{2.37} = 29.5 \mu\text{sec}$$

$$(15) \text{ Let } R_2 = 10 \text{ kilohms}$$

then

$$C_2 = \frac{T_2}{R_2 + R_3'} = \frac{29.5 \times 10^{-6}}{(10 + 6.33)10^3} = \frac{29.5 \times 10^{-6}}{16.33 \times 10^3} = 1.81 \times 10^{-9} \text{ f}$$

$$(16) C_{2 \text{ max}} = \frac{\tau_1}{5(R_{\theta 2} + R_3)} = \frac{5 \times 10^{-6}}{5(1.5 + 20)10^3}$$

$$= \frac{5 \times 10^6}{5(2.15) \times 10^4} = 0.465 \times 10^{-10} \text{ f}$$

It is clear that our choice of R_2 is too low since $C_2 > C_{2 \text{ max}}$. To overcome this difficulty the time constant of the R_3C_2 charging circuit must be raised. Comparison of the last two equations indicates that 1 megohm is a good choice for R_2 . Then

$$(17) C_2 = \frac{29.5 \times 10^{-6}}{10^6 + 6.33 \times 10^3} = \frac{29.5 \times 10^{-6}}{1.006 \times 10^6} = 29.4 \mu\mu\text{f}$$

This completes the design of the multivibrator.

4-6. The Order of Magnitude Equation

Frequently various texts give the following equation for the order of magnitude of the multivibrator pulse-repetition rate:

$$f = \frac{1}{R_1C_1 + R_2C_2} \quad (4-42)$$

It may be shown quite readily that this equation is correct when the following conditions are met.

$$\text{If } k_1\mu_{co1} = \epsilon$$

$$\text{then } \tau_1 = T_1 \quad (4-43)$$

$$\text{and if } R_1 >> R_4'$$

$$\text{then } T_1 = R_1C_1 \quad (4-44)$$

and similarly for the other section.

4-7. Wave Forms of the Plate-coupled Multivibrator

It should be remembered that the multivibrator was introduced in this chapter because it may be used as the impulse generator required

in the saw-tooth voltage generator. In this particular application the exact wave-shape of the circuit's output is not of great importance. For the sake of completeness, however, we shall determine the shapes of the electrode voltages which are important in other uses of the multivibrator. Given the equivalent circuits and equations which have been derived in the preceding sections, the calculation of the wave forms is relatively easy. Consider first the ground to grid voltage, e_{c1} . During the τ_1 interval condenser C_1 is discharging in the equivalent circuit of Fig. 4-11 and the equation for the grid voltage in this interval is given by eq. (4-22).

During the τ_2 interval on the other hand, C_1 is charging as shown in Fig. 4-15. Application of our previous equations shows that during this part-cycle e_{c1} will be given by

$$\text{During } \tau_2 \quad e_{c1} = \frac{R_{g1}}{R_{g1} + R_4} E_{bb} \left(k_1 - \frac{1}{\mu_{c01}} \right) e^{-t/(R_{g1} + R_4)C_1} \quad (4-45)$$

The positive spike represented in the diagram as V_a will be

$$V_a = e_{c1} \Big]_{t=0} = \frac{R_{g1}}{R_{g1} + R_4} \left(k_1 - \frac{1}{\mu_{c01}} \right) E_{bb} \quad (4-46)$$

And the final voltage V_b will be

$$V_b = e_{c1} \Big]_{t=\tau_2} = V_a e^{-\tau_2/(R_{g1} + R_4)C_1} \quad (4-47)$$

Now it is well to take stock of our results at this point, for the values of V_a and V_b given above contradict an assumption used in the derivation of some of the multivibrator design equations. It will be remembered that the values of static plate resistance, R_{p1} and R_{p2} , are evaluated on the basis that during conduction (*i.e.*, during τ_2 for V_1) the grid voltage on the tube is zero, yet these equations show that actually the grid is positive during this interval. A consideration of the relative magnitudes of the relevant quantities shows, however, that the assumed and actual conditions are quite similar, for the first factor in the last two equations, namely, $\frac{R_{g1}}{R_{g1} + R_4}$ approaches zero in value.

It will be found however, that this slight positive peak of grid voltage does modify the calculated values of *plate* voltage slightly.

Proceeding to the calculation of plate wave forms, it would appear at first glance that e_{b1} is constant at supply voltage value during τ_1

when the tube is not conducting. An examination of the circuit diagram will show that this cannot be true because during that same interval C_2 is charging and the charge current which flows through R_3 produces a change in the plate voltage which is exponential; hence the actual shape of e_{b1} can be calculated by analyzing the condenser discharge circuit, which is similar to that of Fig. 4-15. We have from Fig. 4-17 that during τ_1

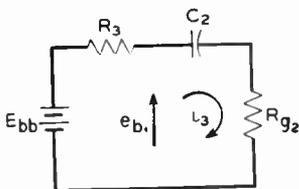


Fig. 4-17. e_{b1} may be calculated during τ_1 when C_2 is charging.

$$e_{b1} = E_{bb} - i_3 R_3$$

$$= E_{bb} \left[1 - \frac{R_3}{R_{g2} + R_3} \left(k_2 - \frac{1}{\mu_{co2}} \right) \epsilon^{-t/(R_{g2} + R_3)C_2} \right] \quad (4-48)$$

$$V_c = e_{b1} \Big|_{t=0} = E_{bb} \left[1 - \frac{R_3}{R_{g2} + R_3} \left(k_2 - \frac{1}{\mu_{co2}} \right) \right] \quad (4-49)$$

and

$$V_d = e_{b1} \Big|_{t=\tau_1} = E_{bb} \left[1 - \frac{R_3}{R_{g2} + R_3} \left(k_2 - \frac{1}{\mu_{co2}} \right) \epsilon^{-\tau_1/(R_{g2} + R_3)C_2} \right] \quad (4-50)$$

Actually the last equation may be simplified considerably, for from the assumption that C_2 becomes fully charged during τ_1 (which was made in the design procedure), it follows that the exponent had to be 5 or more. Furthermore, typical circuit values show the following inequalities to be true:

$$\left. \begin{aligned} \frac{R_3}{R_{g2} + R_3} &< 1 \\ \left(k_2 - \frac{1}{\mu_{co2}} \right) &< 1 \end{aligned} \right\} \quad (4-51)$$

Thus, for all practical purposes,

$$V_d = E_{bb} \quad (4-52)$$

From the above it would seem that e_{b1} during τ_2 could be calculated by setting up the equivalent discharge circuit for C_2 , but as a practical matter this is unnecessary for, during τ_2 , V_1 is conducting and the condenser discharge current in R_3 is negligibly small in comparison to the plate current of the tube. On this basis and assuming $e_c = 0$,

during τ_2

$$e_{b1} = E_{bb} \frac{R_{p1}}{R_{p1} + R_3} \tag{4-53}$$

and remains constant at this value throughout the entire part-cycle. This is represented by the solid line in Fig. 4-18. But we have previously seen that during τ_2 the grid of V_1 is positive and varying, and not constant at zero as assumed above. This means that R_{p2}

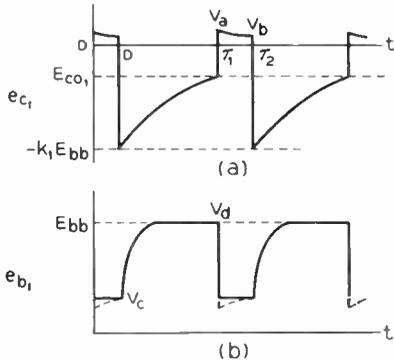


Fig. 4-18. Wave forms on V_1 . (a) Grid voltage. V_a and V_b are exaggerated. (b) Plate voltage. The dotted curve shows the effect of variation in R_{p2} during conduction.

varies slightly rather than remaining constant—or stated in more familiar terms, the grid voltage appears amplified and inverted on the plate. Thus the slight positive spike of grid voltage during τ_2 causes a slight dip in plate voltage below the value specified in (4-53) at the beginning of the τ_2 part-cycle.

Thus we see that the equivalent circuits provide a means of calculating the shape of the various tube-element voltages. It is apparent that similar methods may be applied to the second tube, V_2 , and the voltages obtained would be of the same general shape as those in Fig. 4-18, except that cutoff would occur during τ_2 and conduction during τ_1 . It should be emphasized that the plate wave forms can be made more steep than those in Fig. 4-18 by proper design of the appropriate τ/T ratios.

In all the preceding discussion, shunt capacitances and their effects have been neglected. At higher values of pulse-repetition rate, the shunt capacitances lower the gain of the two stages and discriminate against the higher frequency components of the waves which are

rich in harmonics. Where necessary these effects may be taken into account by adding the capacitances in the various equivalent circuits and deriving a new set of equations. Analysis becomes much more complex than that already given. In general, the over-all effect is to round off rising edges of the various waves.

4-8. Synchronizing the Multivibrator

The design procedure which was outlined in the last section results in a free-running multivibrator, *i.e.*, a multivibrator whose frequency of operation is dependent entirely upon its own circuit constants. In order to meet the requirements of synchronized scanning, however, we must now consider how the multivibrator may be locked in with some sort of synchronizing signal. Shenk⁵ has given a rather complete analysis of this problem which allows rigid control of both portions of the output wave form. Such a high degree of control is not necessary in the application of a scanning generator and we shall consider the design from the point of view of synchronizing the multivibrator to the correct frequency only. In order to make the situation facing us more explicit we shall consider the multivibrator with its associated parts which go to make up the entire saw-tooth voltage generator. One possible circuit configuration is given in Fig. 4-19, where the grids of V_3 , the discharge tube, and V_2 of the multivibrator, are tied together. Since the cathodes of both tubes are tied to ground, both tubes, if they are identical, will be cut off

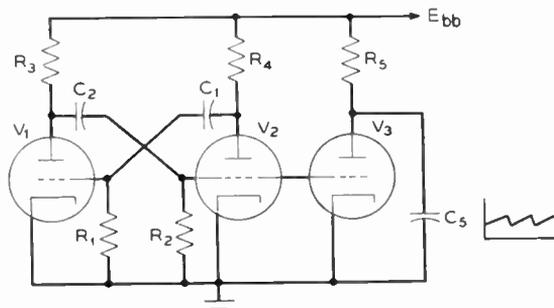


Fig. 4-19. A multivibrator-controlled sweep generator. The multivibrator feeds a vacuum discharge tube which controls the charge and discharge of C_5 .

⁵ E. R. Shenk, *op. cit.*

for the same interval of time. C_5 charges when V_3 is not conducting. It follows, therefore, that τ_2 of the multivibrator and τ_c , the charging time of C_5 , will be equal. If a free-running system were permissible, the design equations which have been given would permit calculation of all the circuit constants in the diagram. It should be noticed, however, that the parallel connection of the V_2 and V_3 grids would lower the effective grid resistance of V_2 , i.e., R_{g2} in the counterpart of eq. (4-39), to one-half its usual value, making it in the order of 750 ohms.

We now consider how this circuit may be synchronized by an external signal. Our previous analysis has shown that at the beginning of the τ_2 part-cycle, e_{c2} , the grid voltage on V_2 drops to $(-k_2 E_{bb})$ and then builds up exponentially, the part-cycle ending when e_{c2} reaches the cut-off voltage corresponding to the given value of E_{bb} . If, now, the multivibrator were designed for a τ_2 greater than τ_c , V_2 could be forced to conduct at any time less than τ_2 , say at τ_c , by the application to its grid of a positive-going, externally supplied voltage of proper amplitude, as shown in Fig. 4-20. The

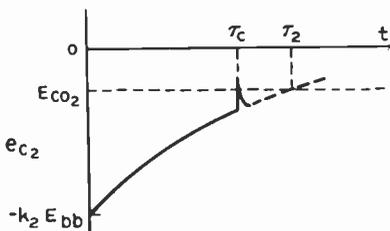


Fig. 4-20. Tube V_2 of free-running cutoff interval τ_2 may be forced to conduct at τ_c by an externally applied sync signal.

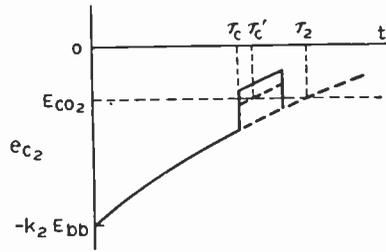


Fig. 4-21. A square synchronizing pulse provides poor synchronization. Any change in amplitude of the pulse causes a shift in time of the conducting point.

interval for which V_1 is cut off, τ_1 , would be unaffected. Hence the frequency of operation would become

$$f = \frac{1}{\tau_1 + \tau_c} = f_s \tag{4-54}$$

where f_s is the frequency of the synchronizing signal. It may be seen, then, that for satisfactory synchronization the free-running

frequency of the multivibrator must be less than f_s and the synchronizing signal must be of sufficient amplitude to satisfy the relationship

$$-k_2 E_{bb} e^{-\tau_c/T_2} + e_{\text{sync}} = E_{c02} \quad (4-55)$$

Two further aspects of synchronization must be considered; what shape of sync signal is best suited to our ends, and how shall that signal be injected into the multivibrator? In regard to the first, three types of signals may be considered, the sine and square waves and the pip, or short pulse of steep wave front. If the sine wave be chosen, (4-55) becomes

$$-k_2 E_{bb} e^{-\tau_c/T_2} + E_s \sin(\omega_s t + \phi) = E_{c02} \quad (4-56)$$

from which it may be seen that the chances for poor synchronization are large because any slight variation in E_s , ϕ , or in the multivibrator constants will result in a shift in phase of the output. Stated in other terms, we note that the time rate of change of a sinusoidal quantity is slow. Hence any small changes in grid voltage magnitude will result in uncertainty of the conduction time.

The use of a square wave as the synchronizing signal is open to the same sort of criticism because the resultant voltage of exponential plus square pulse is not flat-topped, and circuit variations will permit V_2 to conduct at different points along the top of the pulse as shown in Fig. 4-21.

When a pip is used as the sync voltage on the other hand, V_2 is only given a momentary chance to conduct. If its amplitude is below that required by (4-55), no synchronization will occur until the next cycle of operation. If its amplitude increases, no change in τ_c will occur because of its extremely short rise time. For these reasons the pip is most desirable as a sync signal and may be obtained by differentiating a square wave or some other pulse characterized by a relatively steep wave front.

The second aspect of synchronization which must be investigated is that of sync injection: how shall the synchronizing pip be injected into the multivibrator? At the outset it must be stated that any injection circuit will tend to load some portion of the multivibrator circuit, and this loading effect must be included in the multivibrator design. Generally, this effect will result in the lowering of the effective values of R_1 , R_2 , R_3 , and R_4 or the equivalent grid and plate resistances. Theoretically the injection may take place into grid,

plate or cathode circuits in the multivibrator but practical considerations generally rule out grid injection. The reason for this is that the grid leak resistances, R_1 and R_2 , generally have the highest values of all the circuit constants. Hence any injection circuit connected to the grid will have maximum shunting effect.

Several typical complete saw-tooth voltage generators are shown in Fig. 4-22, each of which employs a different scheme of sync injection. The circuit at *a* is desirable from a design point of view because the injection circuit has minimum effect on the multivibrator.

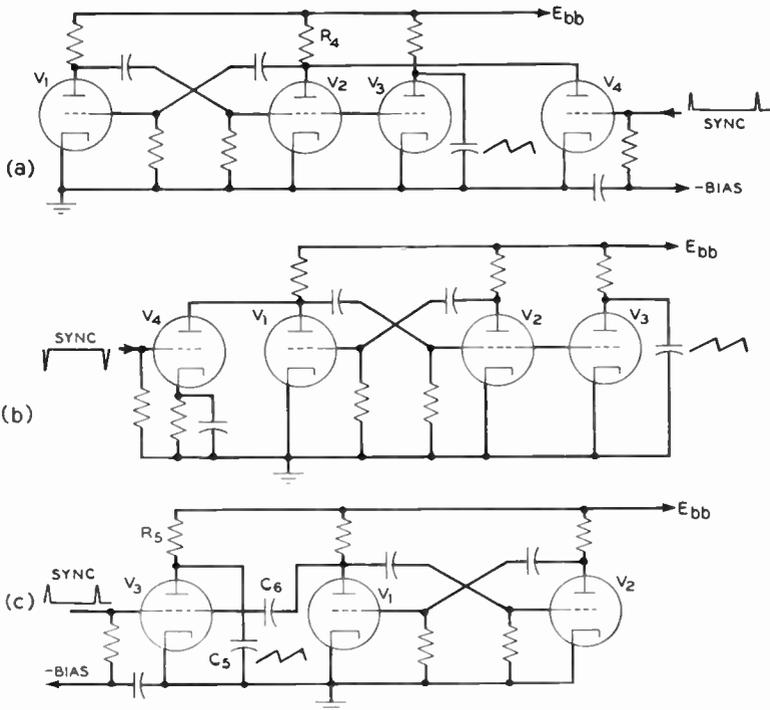


Fig. 4-22. Typical saw-tooth voltage generators employing the multivibrator. (a) Circuit employing positive-going sync. V_4 is normally cut off. Sync injection has negligible effect on the circuit constants. (b) Circuit employing negative-going sync. V_4 is normally conducting. The design is modified by R_{p4} , which shunts the plate of V_4 to ground throughout the entire cycle. The separate sync injection tube is eliminated. V_3 serves as sync injector and discharge tube.

V_4 is normally biased beyond cutoff and hence appears as an open circuit to V_1 and V_2 . During τ_2 the plate of V_2 is at battery potential. At τ_c the sync pulse causes V_4 to conduct and its plate current flowing through R_4 causes e_{b2} to drop momentarily. This drop coupled to the grid of V_1 causes the latter to cut off and the multivibrator circuit switches.

The circuit at c is interesting because the functions of injection and triggering the sweep circuit R_5C_5 are combined in the single tube V_3 . The operation of the circuit may be thought of in the following manner: In the presence of synchronizing pips, the discharge of C_5 is initiated by the pip, the duration of discharge being controlled by the multivibrator wave form which is also locked in with the pip. In the absence of the pip (and it was for this eventuality that the multivibrator was introduced into the sweep circuit) C_5 will continue to charge and discharge at a rate determined by the free-running operation of the multivibrator. Design of this circuit is complicated by the presence of the coupling condenser C_6 .

It is quite apparent from inspection of these circuits that saw-tooth voltage generators employing multivibrators as impulse generators involve three or more tubes plus several related components. In the construction of such a generator the cost of components and assembly runs high. For this reason, the trend has been toward simpler circuits which require fewer components. One form of these employs the blocking oscillator in place of the multivibrator as the impulse generator.

4-9. The Blocking Oscillator

The second form of impulse generator to be considered is the blocking oscillator, two forms of which are shown in Fig. 4-23. Inspection of these diagrams shows that they are identical in form to an audio oscillator of the feedback type. We might expect, therefore, that the wave form from grid to ground or from plate to ground would be a sinusoid of frequency determined by the inductance and distributed capacitance of the coupling transformer. Furthermore the bias would be produced by the flow of grid current charging C_1 during the positive half-cycles of grid voltage, and then by C_1 discharging through R_g for the remaining half-cycle. This action is basic in the oscillator and to see how a pulse output is developed, we consider the development of bias more closely. As the time constant

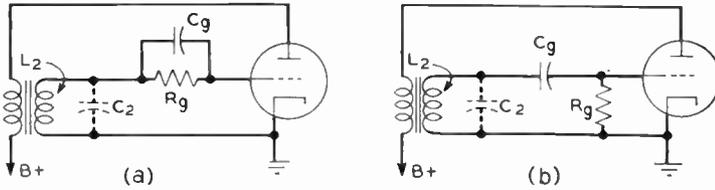


Fig. 4-23. Blocking oscillators are forms of tuned-grid oscillators.

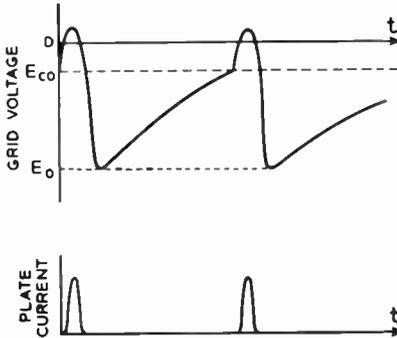


Fig. 4-24. Wave forms in the blocking oscillator.

$R_g C_1$ is increased, the voltage across R_g resulting from the discharge of C_1 becomes greater until the magnitude of discharge current through R_g is so large that the grid is driven beyond cutoff and the flow of plate current is blocked. The blocked condition prevails while the voltage across R_g decreases exponentially until cutoff is reached. The cycle then repeats itself. Typical wave forms for the blocking oscillator are illustrated in Fig. 4-24.

It will be observed that the cycle of operation consists of two distinct parts, one corresponding to positive voltage on the grid which we shall term the pulse interval, and a second when the grid is below cutoff, the interpulse interval. It would be desirable to present an analysis of the blocking oscillator at this point but generally the operation extends into regions of extreme nonlinearity in the tube characteristic and the use of the equivalent plate-circuit theorem as the basis for analysis is invalid. Any useful means of analysis must inevitably be graphical, one such method being that which uses isoclines or curves of constant slope.⁶ Such methods are usually

⁶ I. G. Maloff, and D. W. Epstein, *Electron Optics in Television*. New York: McGraw-Hill Book Company, Inc., 1938, chap. 13.

tedious to apply and generally an empirical approach is used. In general we may state that the pulse interval is governed by the constants of the coupling transformer, and the interpulse period by the time constant $R_g C_1$, cutoff voltage, and the charge accumulated by C_1 during the pulse. From the diagrams we may write

$$E_0 = \frac{1}{C_1} \int_0^{\tau_1} i_g dt \quad (4-57)$$

and

$$E_{e0} = E_0 e^{-\tau_2/R_g C_1} \quad (4-58)$$

In practice it is customary to use a transformer which provides the necessary value of τ_1 , and to adjust R_g and C_1 to give the proper value of τ_2 . Typical values of resistance and capacitance for television applications are given below:

| <i>Application</i> | R_g | C_1 |
|--------------------------------|------------|----------------|
| 60 cycle, vertical sweep | 1 megohm | 0.005 μ f |
| 15,750 cycle, horizontal sweep | 60 kilohms | 470 $\mu\mu$ f |

Since τ_1 is in the order of microseconds, it might be expected that typical transformers for blocking oscillator applications are characterized by low (as compared to audio interstage transformers) values of inductance. Generally this will be in the range of millihenrys. They are further characterized by tight coupling between the plate and grid windings. Typical step-down ratios from the plate to grid side are from 1:1 to 3:1.

In the foregoing discussion we have assumed that the τ_1 part-cycle pulse on the grid is sinusoidal, but under certain circumstances this shape may change because of action in the transformer.⁷ In any event, the τ_1 and τ_2 part-cycles are governed by the transformer and the $R_g C_1$ time constant, respectively.

The use to which a blocking oscillator is put determines the method of deriving its output pulse. Where the positive pulse itself is of importance, the output may be derived from either transformer winding or from a third winding added to the transformer expressly for that purpose. When derived in either of these methods the pulse may exhibit overshoot because of oscillations in the L - C circuit of the transformer winding in use. Overshoot of this type may be eliminated by deriving the output from a resistor in either the plate or

⁷ See Cruft Electronics Staff, *Electronic Circuits and Tubes*. New York: McGraw-Hill Book Company, Inc., 1947.

cathode returns. In this method the output voltage is directly proportional to the plate current of the tube and overshoot is completely eliminated since reversal of current cannot take place.⁸

In the present case, where we are concerned with the generation of a saw-tooth sweep voltage, the grids of the blocking oscillator and discharge tubes may be tied together directly. Then during τ_2 the discharge tube will be held below cutoff and will conduct during the pulse interval.

As a further simplification of the sweep circuit the discharge tube may be eliminated by use of the circuit shown in Fig. 4-25.⁹ The

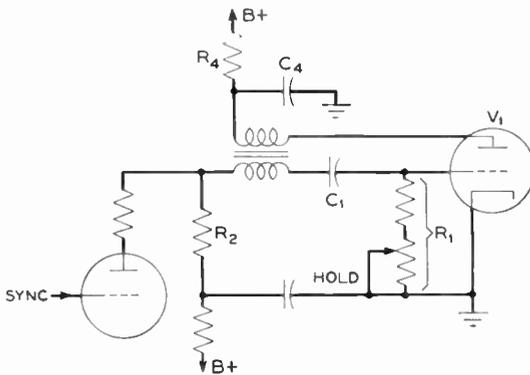


Fig. 4-25. Blocking oscillator sweep-voltage generator. The saw-tooth voltage is developed across C_4 when the tube is nonconducting.

addition of the R_4C_4 combination in the plate circuit of the blocking oscillator tube, V_1 , eliminates the need for a separate discharge tube because C_4 charges during τ_2 when V_1 is cut off, and discharges during τ_1 when the tube is conducting. Sync injection is obtained from R_2 , which is common to the plate circuit of V_2 and the grid circuit of V_1 . The variable portion of R_1 serves as a hold control because its adjustment allows proper setting of the free-running interpulse interval so that proper synchronization may be maintained. Comparison of the circuit with those of Fig. 4-22 shows the degree of simplification

⁸ T. Soller, M. A. Starr, and G. E. Valley, *Cathode Ray Tube Displays*, M.I.T. Radiation Laboratory Series. New York: McGraw-Hill Book Company, Inc., 1948.

⁹ A. Liebscher, "Learn as You Build Television." *Radio News*, 38, 3 (September 1947).

which results when the blocking oscillator replaces the multivibrator as impulse generator.

MAGNETIC DEFLECTION SYSTEMS

Our work in the last chapter showed that for magnetic deflection of an electron beam up to a half-angle deflection of 25° the required saw-tooth deflection would be produced by causing a saw-tooth current to flow in the deflection yoke. It might seem, then, that linear magnetic deflection could be produced by placing the deflection yoke in the plate circuit of a pentode driving tube which is driven by a saw-tooth grid voltage. Actually such a naïve approach to the problem which assumes the pentode to be a constant-current source is not admissible. In the first place, relatively high values of yoke current are required so that a power tube must be used. For such tubes the plate resistance is not sufficiently high to justify the assumption $i_p = g_m e_g$. In the second place linearity requirements demand that the yoke be isolated from the driver plate circuit. This may be seen from the following considerations. Let the yoke be placed in the plate circuit of the driver as shown in Fig. 4-26a. Then,

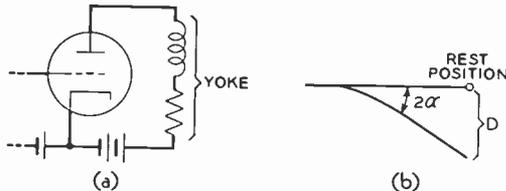


Fig. 4-26. Direct coupling of the deflection yoke to the driver tube.
(a) Circuit. (b) Required deflection is 2α .

since the d-c plate current flows through the yoke, a spurious d-c deflection of the electron beam will occur unless the no-signal plate current is zero. This, in turn, means that during the sweep cycle the total instantaneous plate current must start at zero and finally return to zero, a fact which requires that the tube operate in regions of extreme nonlinearity. It also means that the entire peak-to-peak angular deflection of 50° be provided by an increasing sweep current, as shown at *b* in Fig. 4-26. Thus an angular deflection of 50° is required for, say, a 10BP4 or 16AP4 CRT, and of 70° for the shorter 16CP4, instead of the corresponding half-angle deflections of 25° and

35° , respectively. Hence special design of the driving voltage would be required to overcome the nonlinearity introduced by the tube and the nonlinearity caused by the breakdown of eq. (3-18). To overcome these difficulties it is common practice in television design to isolate the yoke proper from the driver plate circuit by means of a transformer as shown in Fig. 4-27a. The introduction of the iso-

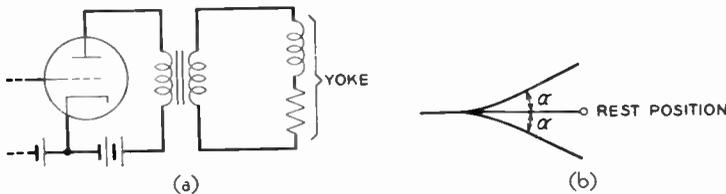


Fig. 4-27. Transformer coupling of the deflection yoke to the driver tube. (a) Circuit. (b) Required deflection is α on each side of the rest position.

lation transformer has a number of advantages. First, the yoke is isolated from the driver plate circuit; hence the driver quiescent point may be chosen so that the tube operates in the linear region of its characteristics. Second, since the yoke is isolated from the d-c plate current, a d-c positioning current may be introduced in the yoke to allow control of the no-signal position of the electron beam on the face of the CRT. Third, the actual yoke current may be alternating and the design may be based on a half-angle deflection, α , rather than the full deflection, 2α . Fourth, the transformer may be designed to step up the current in the yoke. This feature has two effects: the current demand on the tube is lowered, and a smaller region of the tube characteristic is utilized with an improvement in linearity.

We shall, therefore, assume in the rest of our treatment that the yoke is coupled to the driver through a transformer. Since the yoke is predominantly inductive, the usual concepts of impedance matching break down. The transformer may be designed for a current step-up, however, and to simplify our work we shall assume the transformer to be ideal, that it has negligible leakage reactance and infinite incremental primary inductance.¹⁰

¹⁰ For a more rigorous analysis which includes transformer leakage reactance, see A. W. Friend, "Television Deflection Circuits." *RCA Review*, VIII, 1 (March 1947).

4-10. Increasing or Decreasing Sweep Current

With the transformer isolating the yoke from the driver plate circuit we have complete freedom in choosing the direction of plate current during the trace portion of the scanning cycle. For example, we may have an increase in plate current during scan as shown at *a* in Fig. 4-28, or a decrease during scan as shown at *c*; both forms of

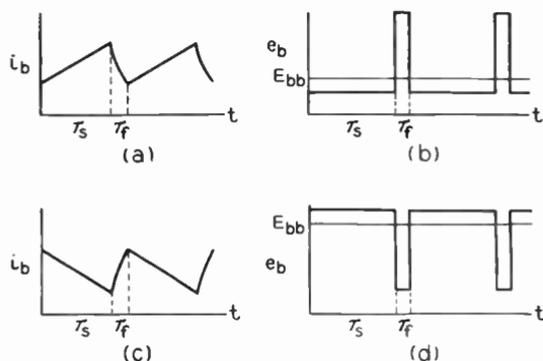


Fig. 4-28. Two alternative forms of driver plate current meet the requirements of linearity. (a) Driver current increases during scan. (b) Plate voltage corresponding to increasing sweep current. (c) Driver current decreases during scan. (d) Plate voltage corresponding to decreasing sweep current.

plate current will meet the requirements of linearity. Further investigation shows that each form has an advantage. In order to show this let us simplify the discussion by assuming that the yoke and transformer present a purely inductive load of magnitude L to the driver. Then, whenever the plate current is changing, the voltage drop across the load will be

$$e = L \frac{di_b}{dt} \quad (4-59)$$

and the corresponding instantaneous plate voltage will be

$$e_b = E_{bb} - L \frac{di_b}{dt} \quad (4-60)$$

From these equations we see that an increasing i_b (positive slope) causes the plate voltage to drop below the supply voltage and, con-

versely, a decreasing current produces a plate voltage greater than the supply voltage. In either case the magnitude of the drop across the load is determined by the slope of the driver current. Hence, during trace the load voltage is much smaller than during retrace when the current curve is changing rapidly. The actual shape of the load voltage *during flyback* will obviously depend on the shape of the current in that same interval. To a first approximation we may assume that the flyback current changes linearly and the resulting plate voltage for the two types of currents are plotted in Fig. 4-28 at *b* and *d*. From these it may be seen that the second system shown at *c* and *d* provides a means of effectively "boosting" the d-c supply voltage because the pulse across the load during flyback is large enough to raise the average value of plate voltage. A circuit which utilizes this boosting action is discussed in section 4-18.

Even though the idea of "getting something for nothing" by using the negative slope scan current is appealing, actually the positive slope scan current is generally used. The reason for this may be seen by considering the problem of current reversal during the flyback interval. During this interval the rate of current change is high and it is desirable in the interest of rapid flyback to have a high voltage across the load to help reverse the current.¹¹ Inspection of the wave forms of Fig. 4-28 shows that this condition is met by the plate current which increases during the scan part-cycle. A further advantage afforded by this form of current is that it permits use of a driving grid voltage which also increases during scan. This simplifies the circuitry of the entire sweep system.

4-11. Driver Grid Voltage

We well might inquire at this point what shape of driver grid voltage is required to produce the positively increasing saw-tooth current in the deflection yoke. To do this, consider the basic deflection circuit shown in Fig. 4-29.

It is assumed that the operating point of the driver is so chosen that its operation is linear and that the equivalent plate circuit shown at *c* is valid. The static plate resistance, R_p , rather than the dynamic plate resistance is used because the plate current has a form which is

¹¹ K. Schlesinger, "Magnetic Deflection of Kinescopes." *Proc. IRE*, **35**, 8 (August 1947).

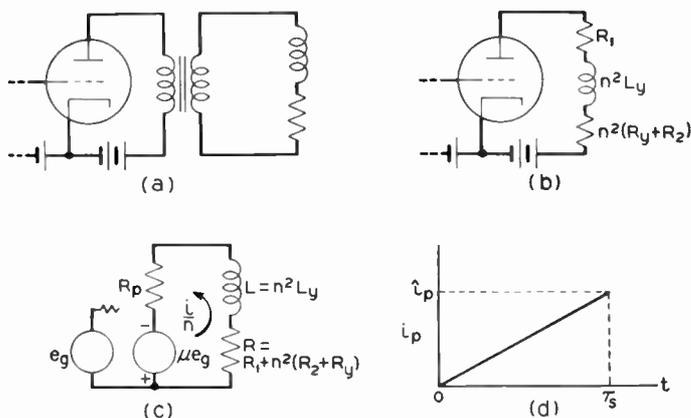


Fig. 4-29. Simplification of the transformer-coupled driver plate circuit. (a) The basic circuit. (b) Equivalent circuit referred to the primary side of the transformer. (c) Equivalent plate circuit. (d) Required a-c component of driver plate current.

matched better by the former condition. The notation of the diagram is defined below.

$$\left. \begin{aligned}
 R_p &= \text{static plate resistance of driver at the} \\
 &\quad \text{operating point} \\
 n &= \text{primary to secondary turns ratio} \\
 R_1 &= \text{transformer primary resistance} \\
 R_2 &= \text{transformer secondary resistance} \\
 R_y &= \text{yoke resistance} \\
 L_y &= \text{yoke inductance} \\
 \hat{i}_p &= \text{peak-to-peak plate deflection current} = \frac{\hat{i}}{n} \\
 \hat{i} &= \text{peak-to-peak yoke deflection current} \\
 R &= n^2(R_y + R_2) + R_1 \\
 L &= n^2 L_y
 \end{aligned} \right\} (4-61)$$

\hat{i}_p may be determined from eq. (3-19) or its equivalent in terms of the angle of deflection from rest position. Then, applying Kirchhoff's voltage law to the circuit of Fig. 4-29c, we get

$$\mu e_g = i_p(R_p + R) + L \frac{di_p}{dt}$$

or
$$e_g = \frac{\hat{i}_p}{\mu\tau_s} [L + (R_p + R)t] \quad \text{for } 0 \leq t \leq \tau_s \quad (4-62)$$

Equation (4-62) gives the driver grid voltage required to produce the deflection demanded by the system.¹² Inspection of the equation shows that the voltage at $t = 0$ jumps to a value $\hat{i}_p L / \mu\tau_s$ and then builds up at constant slope in time until $t = \tau_s$. Such a voltage, which may be said to be trapezoidal in form, is depicted in Fig. 4-30.

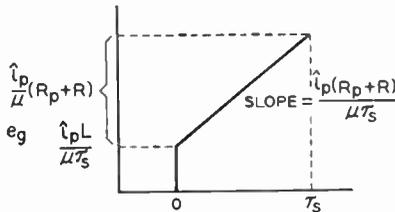


Fig. 4-30. The trapezoidal voltage required on the driver grid.

Given the grid voltage, we next consider circuits which may be used to produce it.

4-12. Trapezoidal Generator

The trapezoidal voltage just described may be derived from a modified version of the basic $R-C$ circuit described in the section 4-1. This modification is shown at Fig. 4-31 where the output voltage is

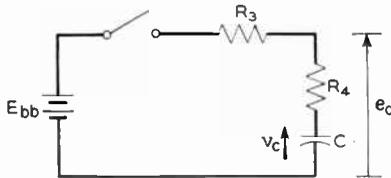


Fig. 4-31. The basic $R-C$ circuit modified for the generation of a trapezoidal voltage.

¹² M.I.T. Radar School Staff, *Principles of Radar*. New York: McGraw-Hill Book Company, Inc., 1946.

developed across the combination of R_4 and C_4 . As before, let E_c be the initial voltage on C_4 before the switch is closed, and T_c the charging time constant.

Then
$$e_o = v_c + iR_4$$

Substituting from eqs. (4-6) and (4-7) for v_c and i , respectively, we get

$$e_o = E_{bb} - (E_{bb} - E_c) \frac{R_3}{R_3 + R_4} e^{-t/T_c} \quad (4-63)$$

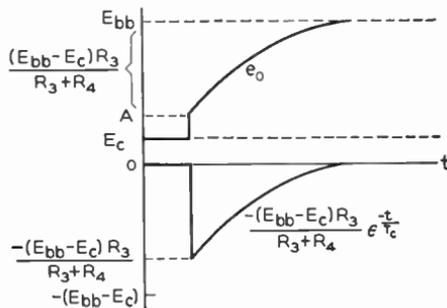


Fig. 4-32. Output of the trapezoidal generator shown in Fig. 4-31.

The components of e_o and e_o itself are plotted in Fig. 4-32. From the diagram it may be seen that the voltage at A , i.e., the value of e_o at the instant the switch is closed is:

$$\begin{aligned} (e_o)_0 &= E_{bb} - (E_{bb} - E_c) \frac{R_3}{R_3 + R_4} \\ &= E_c + (E_{bb} - E_c) \frac{R_4}{R_3 + R_4} \end{aligned} \quad (4-64)$$

Since E_c is the initial condenser voltage, it follows that the initial rise in voltage at $t = 0$ is

$$(E_{bb} - E_c) \frac{R_4}{R_3 + R_4} \quad (4-65)$$

Since the output of the trapezoidal generator will be coupled usually to the grid of the driver through an R - C network, the d-c component, E_c , will be removed. Thus for design purposes it is convenient to express e_o as a variable component added to E_c . Equation (4-65) gives the initial rise part of this varying component, and the ex-

ponential part could be found by algebraic manipulation of (4-64). It is instructive to use an alternative approach, however, which will give the varying components directly. To do this we replace the initially charged condenser by a series combination of a condenser of the same capacitance but with zero initial charge and a battery of terminal voltage E_c . The net effect in the circuit, then, is that the net battery voltage has been reduced from E_{bb} to $(E_{bb} - E_c)$. The value for T_c , of course, remains unchanged because the values of resistance and capacitance in the circuit are unaffected by the substitution. This equivalent circuit is given in Fig. 4-33. Then, applying again eqs. (4-6) and (4-7), we get

$$e_o = E_c + (E_{bb} - E_c) \left(1 - \frac{R_3}{R_3 + R_4} e^{-t/T_c} \right) \tag{4-66}$$

Equations (4-63) and (4-66) may be shown to be identical by suitable algebraic manipulation. If now we let \bar{e}_o be defined as the varying component of the output voltage, then

$$\begin{aligned} \bar{e}_o &= e_o - E_c \\ &= (E_{bb} - E_c) \left(1 - \frac{R_3}{R_3 + R_4} e^{-t/T_c} \right) \end{aligned} \tag{4-67}$$

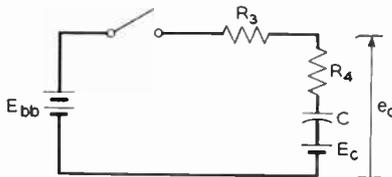


Fig. 4-33. Simplification of the trapezoidal generator circuit. The condenser with initial voltage E_c is replaced by a condenser with initial voltage zero in series with a battery whose voltage is E_c .

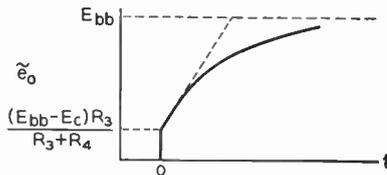


Fig. 4-34. The varying component of output voltage of the trapezoidal generator.

\bar{e}_0 is plotted in Fig. 4-34, which may be seen to be the same as Fig. 4-32, with the single exception that all the ordinates are displaced downward by the voltage E_c .

If we now apply the same technique that was used in saw-tooth voltage generation earlier in the chapter, we see that if the ratio t/T_c is restricted to sufficiently small values, the build-up after $t = 0$ will be linear in time and the required trapezoidal wave will result. Thus expanding the exponential wave and assuming that t is restricted so that

$$\left(\frac{t}{T_c}\right)^2 \ll \left(\frac{t}{T_c}\right) \quad (4-68)$$

$$\text{we get} \quad \bar{e}_0 = (E_{bb} - E_c) \left[\frac{R_4}{R_3 + R_4} + \frac{R_3}{R_3 + R_4} \left(\frac{t}{T_c}\right) \right] \quad (4-69)$$

Since \bar{e}_0 is of the same form as the required e_p , right-hand members of (4-62) and (4-69) may be equated. Then, equating coefficients of equal powers of t , we get

$$\left. \begin{array}{l} \text{For } t^0, \quad (E_{bb} - E_c) \frac{R_4}{R_3 + R_4} \equiv \frac{\hat{i}_p L}{\mu \tau_s} \\ \text{For } t^1, \quad \frac{(E_{bb} - E_c)}{T_c} \frac{R_3}{R_3 + R_4} \equiv \frac{\hat{i}_p}{\mu \tau_s} (R_p + R) \end{array} \right\} \quad (4-70)$$

Generally $R_3 \gg R_4$, so that these equations may be rewritten as

$$(E_{bb} - E_c) \frac{R_4}{R_3} = \frac{\hat{i}_p L}{\mu \tau_s} \quad (4-71)$$

and, substituting for $T_c = (R_3 + R_4)C_4 \approx R_3 C_4$,

$$\frac{(E_{bb} - E_c)}{R_3 C_4} = \frac{\hat{i}_p}{\mu \tau_s} (R_p + R) \quad (4-72)$$

Then, dividing (4-71) by (4-72), we get

$$R_4 C_4 = \frac{L}{R_p + R} \quad (4-73)$$

$= T_L =$ time constant of driver plate circuit

In order to meet the linearity requirements of (4-68)

$$\tau_s \leq 0.4 T_c \approx 0.4 R_3 C_4 \quad (4-74)$$

or

$$C_{4 \min} = \frac{\tau_s}{0.4 R_3}$$

With eq. (4-73) and (4-74) we have the time constants of the circuit specified in terms of known quantities, and if a value of R_3 be assumed (this will generally be 1 to 2 megohms), R_4 and C_4 may be calculated.

We must now consider the magnitude of the trapezoidal voltage to make sure that the required driving current i_p is furnished. Clearly this will be determined by the quantity $(E_{bb} - E_c)$ in either (4-71) or (4-72). Let us evaluate this quantity. First, consider the charging circuit of Fig. 4-31. If the condenser charges from an initial voltage E_c , at any time, t , the voltage will be, by eq. (4-10),

$$v_c = E_c + (E_{bb} - E_c) \left[\frac{t}{T_c} - \frac{1}{2} \left(\frac{t}{T_c} \right)^2 + \frac{1}{3} \left(\frac{t}{T_c} \right)^3 + \dots \right]$$

Now we have assumed a τ_s/T_c ratio of 0.4 in order to ensure linearity; thus at $t = \tau_s$ the condenser voltage will be

$$(v_c)_{\tau_s} = E_c + (E_{bb} - E_c) \times 0.4$$

or
$$(v_c)_{\tau_s} = 0.6E_c + 0.4E_{bb} \tag{4-75}$$

At the end of the scan part-cycle the condenser must be discharged. Then, carrying over our knowledge of the saw-tooth generator, we see that this may be accomplished by placing a discharge tube in shunt with R_4 and C_4 as shown in Fig. 4-35a. After the circuit has

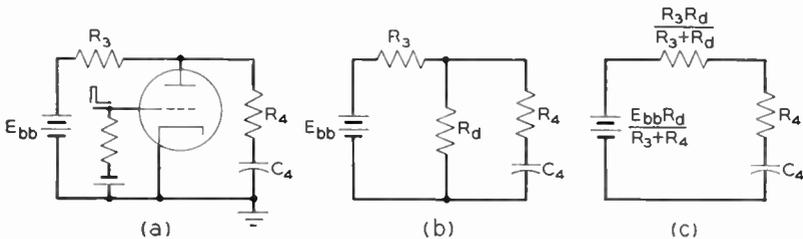


Fig. 4-35. Simplification of the trapezoidal generator during discharge.

been in operation for a long enough period so that an equilibrium condition prevails the condenser voltage must return to its initial value during the discharge interval. Thus from the equivalent circuit of Fig. 4-35c we may write

$$T_d = \left(\frac{R_3 R_d}{R_3 + R_d} + R_4 \right) C_4$$

but

$$R_3 \gg R_d$$

so

$$\begin{aligned} T_d &\approx (R_d + R_4)C_4 \\ &\approx R_d C_4 + T_L \end{aligned}$$

and

$$E_{bb} \frac{R_d}{R_3 + R_d} \approx 0$$

$$\left. \begin{array}{l} \\ \\ \end{array} \right\} (4-76)$$

Hence we have the result that during discharge

$$v_c = (0.6E_c + 0.4E_{bb})\epsilon^{-t/T_d}$$

and at $t = \tau_d$

$$v_c = E_c = (0.6E_c + 0.4E_{bb})\epsilon^{-\tau_d/T_d} \quad (4-77)$$

These equations may be arranged into a design procedure. With τ_d , τ_s , $\hat{i}_p L$, R_p , and R given

(1) Choose R_3 , nominally 1 to 2 megohms.

(2) Calculate R_d by the graphical method illustrated in Fig. 4-9.

An approximation is involved here because the exact value of E_{bb} is not known as yet.

(3) Calculate C_4 from eq. (4-74).

(4) Calculate T_L and R_4 from (4-73).

(5) Calculate $E_{bb}/E_c = k$.

$$\text{From (4-77)} \quad \epsilon^{\tau_d/T_d} = 0.6 + 0.4 \frac{E_{bb}}{E_c}$$

$$\text{or} \quad k = \frac{E_{bb}}{E_c} = 2.5(\epsilon^{\tau_d/T_d} - 0.6) \quad (4-78)$$

(6) Calculate E_{bb} .

$$\text{From (4-71)} \quad E_{bb} \left(1 - \frac{E_c}{E_{bb}}\right) \frac{R_4}{R_3} = \frac{\hat{i}_p L}{\mu \tau_s}$$

$$\text{or} \quad E_{bb} = \frac{\hat{i}_p L}{\left(1 - \frac{1}{k}\right) \mu \tau_s} \frac{R_3}{R_4} \quad (4-79)$$

Let us illustrate this design procedure with a typical problem. Design a vertical sweep generator for a progressive scan television system which operates with a frame frequency of 30 cycles per second. The flyback ratio is to be 1 to 19. The following constants are to be used:

Deflection yoke:

$$L_y = 50 \text{ millihenrys, } R_y = 65 \text{ ohms, } \hat{i} = 320 \text{ ma}$$

Transformer: $n = 10$, $R_1 = 600$ ohms, $R_2 = 10$ ohms

A triode-connected 6K6-GT is to be used as a driver so that $R_p = 6,250$ ohms and $\mu = 6.5$. A 6SN7 is to be used as the discharge tube, $R_d \approx 13$ kilohms,

$$\tau_s = \frac{19}{19 + 1} \left(\frac{1}{30} \right) = 31.6 \text{ msec}$$

$$\tau_d = \frac{1}{0.03} - 31.6 = 1.7 \text{ msec}$$

$$L = n^2 L_y = 100(50 \times 10^{-3}) = 5 \text{ h}$$

$$R = n^2(R_y + R_2) + R_1 = 100(65 + 10) + 600 = 8100 \text{ ohms}$$

Then, following the design procedure, we have

(1) Choose $R_3 = 1$ megohm

(2) Calculate R_d . For the present problem we shall assume this to be

$$R_d = 1.3 \times 10^4 \text{ ohms}$$

(3) $C_4 = \frac{\tau_s}{0.4 R_3} = \frac{31.6 \times 10^{-3}}{4 \times 10^6} = 7.7 \times 10^{-8} = 0.077 \mu\text{f}$

(4) $T_L = \frac{L}{R_p + R} = \frac{5}{6250 + 8100} = \frac{5}{1.435 \times 10^4}$
 $= 3.48 \times 10^{-4} \text{ sec}$

$$R_4 = \frac{T_L}{C_4} = \frac{3.48 \times 10^{-4}}{0.77 \times 10^{-7}} = 4.52 \times 10^3 \text{ ohms}$$

(5) $T_d = R_d C_4 + T_L = 1.3 \times 10^4(7.7 \times 10^{-8}) + 3.48 \times 10^{-4}$
 $= 10 \times 10^{-4} + 3.48 \times 10^{-4} = 1.348 \times 10^{-3} \text{ sec}$

$$\frac{\tau_d}{T_d} = \frac{1.7 \times 10^{-3}}{1.348 \times 10^{-3}} = 1.261$$

then $k = 2.5(\epsilon^{\tau_d/T_d} - 0.6) = 2.5(3.54 - 0.6) = 7.35$

(6) $1 - \frac{1}{k} = 1 - \frac{1}{7.35} = 1 - 0.136 = 0.864$

$$\hat{i}_p = \frac{\hat{i}}{n} = \frac{320 \times 10^{-3}}{10} = 3.2 \times 10^{-2} \text{ amp}$$

then

$$E_{bb} = \frac{\hat{i}_p L R_3}{(1 - 1/k)\mu\tau_s R_4} = \frac{(3.2 \times 10^{-2})(5)10^6}{(0.864)(6.5)(3.16 \times 10^{-2})(4.52 \times 10^3)} = 200 \text{ v}$$

Thus the design values for the circuit are

$$\begin{aligned} R_3 &= 1 \text{ megohm,} & C_4 &= 0.077 \mu\text{f} \\ R_4 &= 4,520 \text{ ohms,} & E_{bb} &= 200 \text{ v} \end{aligned}$$

and these values give a τ_s/T_c ratio of 0.4.

4-13. Flyback Considerations¹³

In our design equations of the last section we have neglected what happens in the circuit during the flyback interval, τ_f . Hence we must now examine the driver current wave form during flyback more critically. In this connection we may review a few points which have been made previously. First, we have seen that from the viewpoint of the scanning raster i_b may have any conceivable shape whatsoever during the retrace interval just so long as it reaches its trace starting value before the end of the blanking interval. This is true because no picture information is presented during blanking and hence during flyback. Then in section 4-10 we assumed i_b to be linear during retrace in order to permit a choice between the increasing or decreasing forms of driver current. We must now see if this second assumption is necessary, or even possible.

As a practical matter, such a linear retrace current cannot occur because any deflection yoke and transformer system has distributed capacitance associated with it. Consequently at the end of the sweep interval the field built up about the yoke and in the transformer coil begins to collapse, and if the driver plate resistance is high enough so that it has negligible damping effect, the L - C circuit comprising the total inductance and the shunt capacitance starts to oscillate. As may be seen from Fig. 4-36, the resulting current caused by the collapse of the magnetic field is a damped cosinusoid and, further, the shortest flyback time possible will be that corresponding to the half-period of the oscillation. The difficulty, however, is that unless means are provided to damp out the oscillations after τ_f they will contaminate the next trace. Our immediate problem, then, is to find some means of damping out these free oscillations. This may

¹³ K. Schlesinger, *op. cit.*

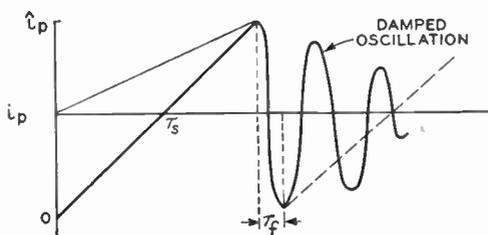


Fig. 4-36. A-C component of the driver plate current.

be done in a number of ways and the particular system chosen must be based on a compromise between economic factors, permissible flyback time, and the linearity of sweep required.

4-14. Shunt-resistance Damping

The most elementary form of damping utilizes the principle that a circuit of L and C may be rendered aperiodic or nonoscillatory by shunting the circuit with a resistance that is sufficiently small. This scheme of resistance damping is shown in Fig. 4-37, R_s being the damping resistor. Since, in this case, we are interested in suppressing oscillations in the yoke itself, it is convenient to construct the equiva-

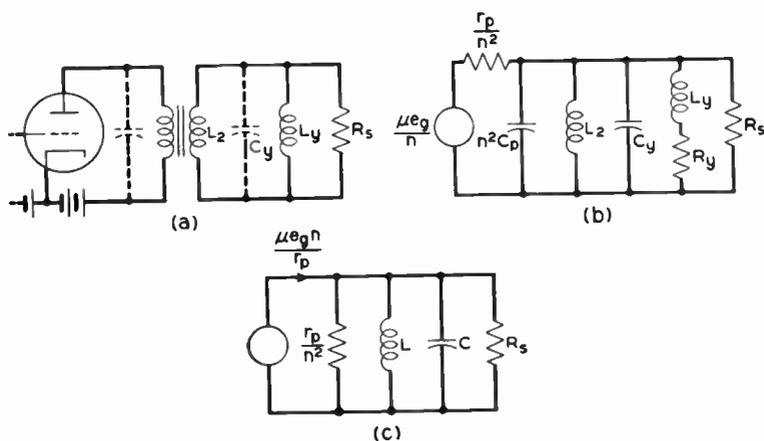


Fig. 4-37. Shunt-resistance damping. (a) The oscillations are damped out by R_s . (b) Equivalent circuit referred to the secondary side of the transformer. (c) The circuit reduced by Norton's theorem so that all the elements are in shunt.

lent secondary circuit given at *b*. Even further simplification yields the circuit at *c* where L_s , the transformer secondary inductance, and L_y have been combined into L , the yoke inductance, R_y , has been neglected, and n^2C_p has been added to C_y to give C . Since in a typical circuit which employs this form of damping n may range around 10,¹⁴ the effective applied voltage is assumed to be negligibly small. It follows then from a transient analysis that the parallel combination of R_s and r_p/n^2 , say R , should be

$$R \leq \frac{1}{2} \sqrt{\frac{L}{C}} \quad (4-80)$$

Thus R_s may be chosen to meet this design requirement.

It is of importance to note that the shunt resistance, R_s , changes the design equations for the trapezoidal voltage generator. That this is so may be seen from the following considerations. In place of Fig. 4-29c we now have Fig. 4-38. It may be seen that R_p , R_1 and

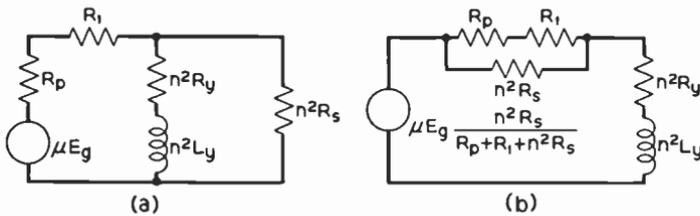


Fig. 4-38. Equivalent primary circuits for shunt-resistance damping.

$n^2 R_s$, comprise a voltage-dividing network and Thevenin's theorem may be applied to the equivalent generator circuit yielding the result shown at *b*. Thus we see that the effect of the damping resistor during the scan part-cycle is to lower the effective plate resistance, and the effective driving voltage in the plate circuit of the driver tube. This, in turn, raises the ratio of initial rise to slope in the trapezoidal generator. Once these facts are realized our previous generator design equations may be utilized with the various parameters redefined in terms of the circuit constants of Fig. 4-38.

It is instructive to note that the same results may be reached by

¹⁴ *Preliminary Data on Television Scanning Circuits and Components*, Part I. RCA Victor Division, 1944.

physical reasoning. Consider the simplified circuit in Fig. 4-39 where the yoke and transformer resistances have been neglected. The inductance current is required to be of saw-tooth shape during τ_s . Then, during this interval, the voltage drop across L is constant and equal to $L \frac{\hat{i}_p}{\tau_s}$. But this same voltage is across R , which shunts L .

Thus i_R is constant and equal to $\frac{L}{R} \frac{\hat{i}_p}{\tau_s}$. Then the total current is the sum of the two components shown at e and is trapezoidal. The applied voltage μe_0 is the sum of e and the iR_p drop. This required grid voltage μe_g is shown at f .

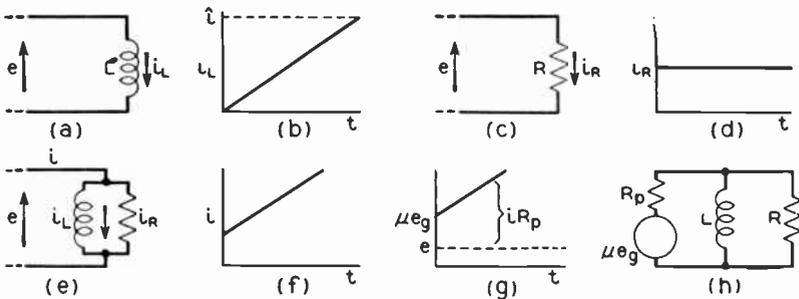


Fig. 4-39. Simplified derivation of the trapezoidal generator voltage. (a) The inductive part of the load. (b) The current in the inductance must increase linearly. (c) The resistive part of the load. (d) The constant voltage $e = L di/dt$ produces a constant current through the resistance. (e) The combined load. (f) The total current is the sum of (b) and (d). (g) The equivalent plate circuit voltage μe_g is the sum of e and the iR_p drop. (h) The complete equivalent circuit.

Whereas the damping system is excellent from the point of view of low-cost parts, it has the disadvantages of providing a relatively long flyback time and of being wasteful of input power. In regard to flyback time it is unsatisfactory for the horizontal deflection system which meets telecasting standards of a $63.5\text{-}\mu\text{sec}$ sweep interval. On the other hand it may be used in the vertical system which allows some 16 milliseconds for the entire vertical sweep and retrace cycle.

In order to examine the second disadvantage, we must consider the power requirements in magnetic deflection. It should be realized that in magnetic deflection only wattless power is required to build

up the deflecting field around the yoke; hence, it would be desirable to return this power to the driver plate supply during the flyback part-cycle when the field is collapsing. Actually with resistance damping all this field power is dissipated in the damping system and as a result can represent a power loss of considerable magnitude. We now calculate this field power:

During the scan interval, τ_s , the current build up in the net inductance is linear, or

$$\text{during } \tau_s \quad i = \frac{\hat{i}}{\tau_s} t \quad (4-81)$$

and the voltage drop across the inductance L of Fig. 4-37 is

$$e_L = \frac{L\hat{i}}{\tau_s} \quad (4-82)$$

Then the average power furnished to the yoke during scan is

$$P_y = \frac{1}{2} L \frac{\hat{i}^2}{\tau_s} \quad (4-83)$$

$$\text{but} \quad \tau_s = \frac{1}{(1+p)f_s} \quad (4-84)$$

where f_s = scanning frequency

$$\text{whence:} \quad P_y = \frac{1}{2} L \hat{i}^2 (1+p) f_s \quad (4-85)$$

Schlesinger has shown that for a 10BP4 cathode-ray tube sweeping at a horizontal frequency of 15,750 sweeps per second this yoke-power term represents 19 watts and, as we have seen, in simple resistance damping all of this power is lost. According to Spielman¹⁵ the average postwar television receiver using a 10-in. picture tube draws some 220 w from the supply line. Thus the power which resistance damping wastes represents an appreciable fraction of the total power requirement of the entire receiver. Furthermore, from the design point of view, the wasted power demands that a larger low-voltage power supply be incorporated in the receiver.

In the interests of completeness we should consider also the power requirements for the vertical deflection system, which may be calculated from (4-84). In going from the horizontal to the vertical

¹⁵ S. C. Spielman, "Input Power Requirements of Television Receivers," AIEE New York Convention, 1949.

system i , L , p , and f_s change, but for a first approximation we may concentrate on the ratio of frequencies, which is so large that it almost completely overshadows changes in the three other quantities for, as we have previously seen, in a 2 to 1 interlaced system the ratio of horizontal to vertical scan frequencies is one-half the number of scanning lines. For usual values ranging from, say, 300 to the 525 lines specified by commercial standards, it follows that the power lost in the vertical sweep system is negligible because of the relatively low vertical scanning frequency.

In summary it may be stated that simple shunt-resistance damping is not used in the horizontal deflection systems recently designed for television equipment. It may be used in the vertical deflection systems where a relatively long flyback interval is available and where the power lost in damping is small.

4-15. Resistance Damping

A second type of damping system that finds commercial application utilizes a series combination of resistance and capacitance shunted across the deflection yoke. Logically it would seem that such a system might be called "resistance-capacitance damping" but the popular term "resistance damping" will be used. The student should take care to note the differences between the shunt-resistance damping system described in the last section, and the one under consideration here. The basic circuit for the latter and its equivalent circuits are given in Fig. 4-40. One advantage of the circuit is immediately apparent, namely, that the presence of the condenser prevents the damping resistance, R_s , from offering a shunt path to the d-c positioning current (not shown in the diagram) which is introduced into the transformer secondary circuit to control the rest position of the electron beam. Other more important features of the system will become apparent as we proceed with the analysis of the circuit.

The equivalent circuits of Fig. 4-40 are derived on the assumption of an ideal transformer of primary to secondary turns ratio, n . From c in the figure it is seen that two oscillatory systems are present. These systems are indicated by the mesh currents i and i_1 . If, however, R_s is made *small* enough to suppress oscillations at a frequency determined by L and C , *i.e.*, at the natural frequency of the undamped deflection system, the equivalent circuit at d results. This, then, is the basic idea involved in resistance damping: the oscillatory circuit

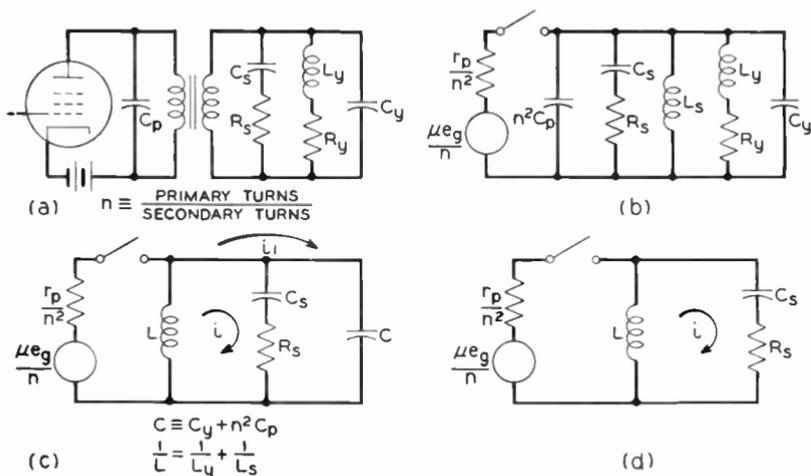


Fig. 4-40. Magnetic deflection circuit with resistance damping.
 (a) Basic circuit. (b) Equivalent circuit referred to the secondary.
 (c) Simplified equivalent circuit neglecting the yoke resistance R_y .
 (d) Simplified circuit during flyback. R_s is chosen to suppress oscillations in the $R_s C_s$ branch.

comprising L and C in parallel is replaced by a circuit of L , C_s , and R_s in series. Since two of these parameters may be chosen at will, considerable flexibility is provided in the design of the damping system.

It should be apparent that the change in the damping circuit requires no change in the basic requirement during the sweep part-cycle, namely, that the current in L builds up linearly in time. At the end of the sweep interval the current build-up ceases abruptly, and previous knowledge of the transient behavior of the series R , L , C circuit indicates that the shortest retrace obtainable is that given by critical damping or when R_s is adjusted so that the system is just rendered aperiodic. The required values of R_s and C_s to give this condition may now be derived. Thus, for the circuit of Fig. 4-40d, we have during the retrace part-cycle:

$$\left(Lp + R_s + \frac{1}{pC_s} \right) i = 0 \quad (4-86)$$

$$\text{or} \quad p = -\frac{R_s}{2L} \pm j \sqrt{\frac{1}{LC_s} - \frac{R_s^2}{4L^2}} = -\alpha \pm j\omega \quad (4-87)$$

where p is the differential operator.

Now for critical damping the imaginary term must be zero, hence

$$R_s = 2 \sqrt{\frac{L}{C_s}} \tag{4-88}$$

and under these conditions the roots of (4-86) are equal and the current in the circuit is¹⁶

$$i = (A + Bt)\epsilon^{-\alpha t} \tag{4-89}$$

A and B may be evaluated from the boundary conditions, namely, that at $t = 0$,

$$i = \hat{i} \quad \text{and} \quad A = \hat{i} \tag{4-90}$$

and further, at $t = 0$, the condenser is uncharged, hence

$$\frac{1}{C_p} i = 0$$

and
$$\left. \frac{di}{dt} \right|_0 = -\frac{R_s}{L} \hat{i} = -2\alpha \hat{i} \tag{4-91}$$

Then differentiating (4-89) and equating to (4-91)

$$\left. \frac{di}{dt} \right|_0 = -\alpha A + B = -2\alpha \hat{i} \tag{4-92}$$

Therefore,
$$B = -\alpha \hat{i} \tag{4-93}$$

and
$$i = \hat{i}(1 - \alpha t)\epsilon^{-\alpha t} \tag{4-94}$$

In order to interpret these results, (4-94) is plotted in Fig. 4-41. Looking at the resulting curve we must decide what is meant by the

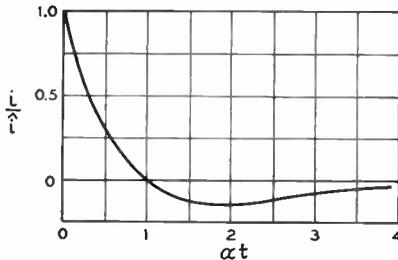


Fig. 4-41. Yoke current during flyback with R_s adjusted for critical damping. (Courtesy of *Proc. IRE.*)

¹⁶ See, for example, W. J. Seeley, *An Introduction to the Operational Calculus* Scranton, Pa.: International Textbook Co., 1941.

expression "flyback is completed." Since at $\alpha t = 2$ the current reverses direction we shall consider this value to define the end of the retrace part-cycle. It is of interest to note that at this value of αt the voltage across the inductance is zero. This latter fact may be verified intuitively because at $\alpha t = 2$ the slope of the current is zero.

We are now able to set up the design conditions for R_s and C_s because the flyback time, τ_f , will be

$$\alpha \tau_f = \frac{R_s}{2L} \tau_f = 2$$

or
$$R_s = \frac{4L}{\tau_f} \quad (4-95)$$

and from (4-88) and (4-95) we get

$$C_s = \frac{\tau_f^2}{4L} \quad (4-96)$$

A rather interesting interpretation may be attached to these results which is of importance in the design or selection of deflection yokes. Let $\omega_0 = 1/\sqrt{LC_s}$ = forced angular frequency of L and C_s .

Then

$$\alpha \tau_f = 2$$

or
$$\tau_f = \frac{2}{\alpha} = 2\sqrt{LC_s} \quad (4-97)$$

$$= \frac{2}{\omega_0} = \frac{\tau_0}{\pi} \quad (4-98)$$

From (4-98) it follows that where critical damping is used, the forced frequency of the L - C_s circuit must be high enough so that one cycle has roughly three times the duration of the flyback time. Since τ_0 is dependent upon L , (4-98) puts a limitation on the maximum value of L because of flyback requirements. Thus in a complete deflection yoke consisting of two sets of windings one of, say, 8-mh inductance, and the other of 50 mh, the second winding with the larger inductance would be used for vertical and not for horizontal deflection because it would not, in general, be able to satisfy the extremely short flyback-time requirement of the horizontal deflection system.

4-16. Advantages and Disadvantages of Resistance Damping

One of the principal disadvantages of resistance damping is that it does not provide a flyback time of the minimum value possible with a

given set of deflection yoke constants. As was previously pointed out this minimum would be one half-cycle of the oscillation in L and C (Fig. 4-40), and in resistance damping we concern ourselves with the circuit L , C_s , and R_s because we have more control over the parameters involved. In the commercial television system, at least, this is not too serious in the vertical deflection system because some 800 μsec are allowed for the vertical retrace. By the same token resistance damping in the horizontal system cannot adequately meet the short flyback-time requirements.

In order to explore this point further we may calculate the ratio of actual flyback time with resistance damping to the minimum possible flyback time with a given yoke and transformer combination. Our previous work (4-98) gives the actual flyback time with resistance damping. But without resistance damping the minimum possible flyback time is one half-cycle of oscillation in L and C , the notation being that of Fig. 4-40. Thus

$$\begin{aligned}(\tau_f)_{\min} &= \text{minimum possible flyback time} \\ &= \pi\sqrt{LC}\end{aligned}\quad (4-99)$$

$$\text{and} \quad \frac{\tau_f}{(\tau_f)_{\min}} = \frac{2}{\pi}\sqrt{\frac{C_s}{C}}\quad (4-100)$$

In order to evaluate this ratio we must calculate the ratio of capacitances which, at first glance, seems rather difficult. The key to the situation is that R_s was chosen small enough to suppress the high-frequency oscillations in the L - C circuit. Arguing backward from this proposition we may get an expression for the capacitance ratio.

In the last section we have the expression for the shunt-damping resistance required for critical damping. The same value may be verified by applying the principle of duality¹⁷ to the circuits of Fig. 4-42. From our previous results we know that in order for the circuit at b to be aperiodic R_s' must have a value

$$R_s' \geq 2\sqrt{\frac{L'}{C'}}\quad (4-101)$$

¹⁷ See, for example, E. A. Guillemin, *Communication Networks*. New York: John Wiley and Sons, Inc., 1935, vol. 2.

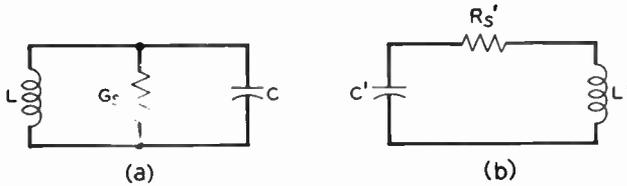


Fig. 4-42. Dual circuits. (a) The parallel circuit. (b) The series circuit is the dual of the circuit at (a).

Then by the principle of duality the network at *a* in the figure will be aperiodic if

$$G_s \geq 2 \sqrt{\frac{C}{L}} \quad (4-102)$$

or if

$$R_s \leq \frac{1}{2} \sqrt{\frac{L}{C}} \quad (4-103)$$

This checks eq. (4-80) of the last section. But R_s is a fixed resistor in the deflection system; hence (4-88) and (4-103) may be equated with the following result

$$\sqrt{\frac{C_s}{C}} \geq 4 \quad (4-104)$$

Substitution of which in (4-99) yields

$$\frac{\tau_f}{(\tau_f)_{\min}} \geq \frac{8}{\pi} \approx 2.54 \quad (4-105)$$

Therefore we see that, at best, resistance damping gives a flyback time which is two and a half times the minimum possible value.

As compared to the shunt-damping system, the present system requires the same ratio of initial rise to slope of the yoke current during the sweep portion of the cycle. This follows from the fact that the value of the damping resistor will be the same in both systems, assuming L and C are the same in both cases. These results may be verified from eq. (4-80), (4-88), and (4-104).

In regards to the dissipation of power supplied to the yoke during the scan intervals, it might seem at first glance that the present system is on a par with the system described in the last section because a resistance, R_s , provides the actual damping. An inspection of Fig. 4-41, however, shows that this is not so because during flyback the current drops below zero on the normalized current scale. In

physical terms this means that during flyback the yoke current reaches a value greater than the peak swing attained during trace. The effective current reversal indicates that the yoke acts as a source of power and delivers power back into the normal supply system. Thus the system of resistance damping wastes only some 77 per cent of the power wasted in the shunt-resistance damping system and hence is more desirable.

Another interesting by-product of the current overswing shown in Fig. 4-41 is that there is a gain in deflection sensitivity of about 13 per cent. Schlesinger has termed this as being caused by "flyback resonance." This is shown in Fig. 4-43.

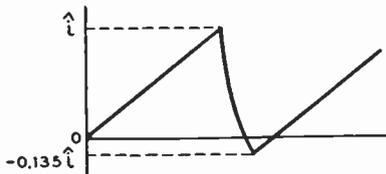


Fig. 4-43. Resistance damping increases the deflection sensitivity, for the sweep starts at $-0.135 \hat{i}$ rather than at zero.

In summary it may be said that the characteristics of resistance damping makes its use in low-frequency deflection systems quite satisfactory. Its comparatively slow flyback and power loss generally prevent its inclusion in the horizontal sweep system.

4-17. Diode Damping

We have just seen that one of the prime difficulties in the resistance-damping system is that it still does not provide a flyback time which is in the order of the minimum possible value which is determined by the total shunt capacitance in the circuit. This came about because the actual shunt oscillatory circuit was replaced by a series resonant circuit involving C_s which is larger than C . Since the problem revolves about some satisfactory means of placing the damping resistance in the circuit, the next step in development is to provide some means of switching this resistance in and out of the circuit. This is the basis for the so-called diode-damping system which employs a vacuum diode as a synchronous switching element. Shift is made so that during flyback the resistance is out of the circuit, and

the retrace speed is determined by L and C . After a complete reversal of current, the resistance is switched in to damp out further oscillations and the sweep generator voltage is applied to the driver grid. Since at the end of the current reversal the voltage across the inductance changes polarity, a diode, which is inherently sensitive to the direction of applied voltage, is quite feasible as a switching device. This idea will be expanded in the following paragraphs.

Two forms of the basic circuit are shown in Fig. 4-44. At a the diode and damping resistance are across the primary of the deflection

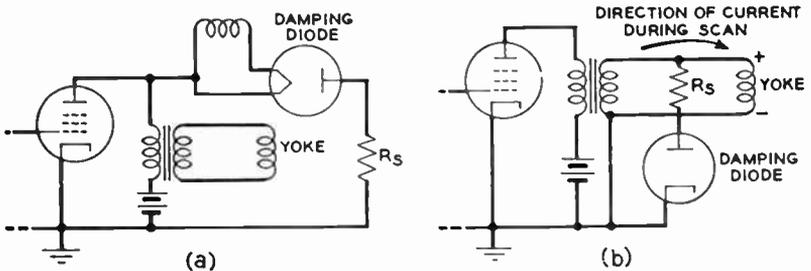


Fig. 4-44. Diode damping. (a) Primary damping. (b) Secondary damping. (Courtesy of *Proc. IRE*.)

transformer; at b they are on the secondary side. Whereas the action of both circuits is identical, the former requires a special filament winding for the damping diode which winding requires relatively high-voltage insulation and also adds to the total circuit shunt capacitance which in the final analysis sets the lower limit on the minimum flyback time. Hence, the circuit at b is preferred.

Consider the operation of the circuit in more detail. In Fig. 4-44b the direction of yoke current and polarity of voltage across the yoke during the trace interval is indicated. Notice that during this period the diode plate is positive relative to its cathode so that it conducts and the damping resistor, R_s , shunts the yoke; the voltage drop across the inductance is

$$(e_L)_{\tau_s} = I\hat{i}(1 + p)f_s \quad (4-106)$$

At the end of τ_s the current changes direction causing a reversal of polarity across the yoke and the damping diode. The latter ceases to conduct and R_s is removed from the circuit. Thus during flyback

the L - C circuit is free to oscillate at its "natural" angular frequency, which is, neglecting the secondary resistance,

$$\omega = \frac{1}{\sqrt{LC}} \tag{4-107}$$

where L and C are as defined in Fig. 4-37.

In this interval i is cosinusoidal and the yoke voltage is

$$(e_L)_{\tau_f} = -L\hat{i}\omega \sin \omega t \tag{4-108}$$

The current and voltage throughout an entire sweep cycle are shown in Fig. 4-45.

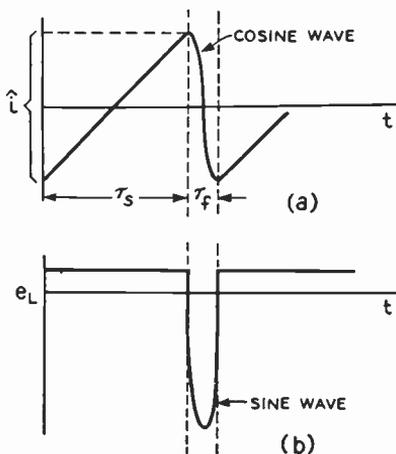


Fig. 4-45. Yoke current and voltage. Positive directions are those indicated in Fig. 4-44(b).

Then at the end of τ_f the voltage again reverses polarity and the damping resistor is switched in to damp out further oscillations in the L - C circuit as rapidly as possible. Since these cannot be suppressed instantaneously, a slight contamination at the beginning of the next sweep occurs but if ω is sufficiently high, this contamination will occur during the blanking interval and the picture proper will not be affected.

In summary, then, we see that with diode damping the flyback time approaches its theoretical minimum of one-half a cycle at ω defined by (4-107); R_s is in the circuit only during the *sweep* part-cycles. Two other important factors must be mentioned. First,

that since R_s is shunted across the yoke during scan, its effect must be taken into account when the driver and trapezoidal generator are designed. Secondly, since R_s is shunted across the yoke during the damping interval, its value may be estimated with the use of eq. (4-80). In this connection it must be remembered that the damping diode itself does not function as an ideal switch and its plate resistance enters into the effective value of damping resistance shunted across the deflection circuit.

Then, to finish the discussion, economically from the manufacturer's point of view diode damping is less desirable than the previously mentioned types because of the added cost of tube, socket, and increased demand on filament power supply. These are outweighed, however, by the short flyback attainable, and diode damping enjoyed extensive use in the horizontal systems of prewar television receivers. It still does not represent the best that may be reached in scan circuit design, for even though complete current reversal occurs during τ_f , still the energy stored in the yoke is damped out as rapidly as possible and the driver must furnish almost the total peak deflection current \hat{i} as shown in Fig. 4-45. This latter point is significant because it marks the difference between the diode damping just described and the more recent reaction scanning, some forms of which may use a diode as a control element.

4-18. Reaction Scanning¹⁸

We have just seen that in diode damping the energy which is stored in the deflection inductance at the end of τ_s is dissipated by switching the damping resistor, R_s , into the circuit. Since this represents a loss of power, the following idea occurs: Is it possible to control this energy in such a manner that it will produce a component of current which will contribute to the linear build-up required for scan? If such a system is possible, it will have an efficiency much greater than that of the previously described ones because the stored energy will be used rather than lost. To this end, the circuit of Fig. 4-46 may be used and the so-called system of reaction scanning will result.

We next consider the action of the circuit in steps, each of which corresponds to a distinct part of the operating cycle. Reference to Fig. 4-47, which shows each step, will aid in the discussion. At the

¹⁸ Instruction Book for Monoscope Camera Type TK-1A, RCA, Engineering Products Department, November 1946. See also A. W. Friend, *op. cit.*

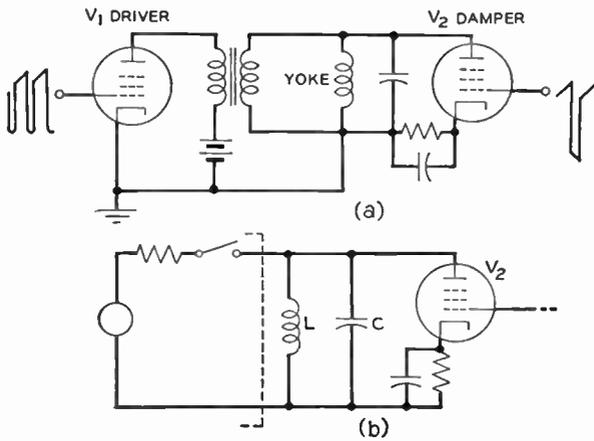


Fig. 4-46. Reaction scanning. (a) Basic circuit. (b) Simplified equivalent circuit.

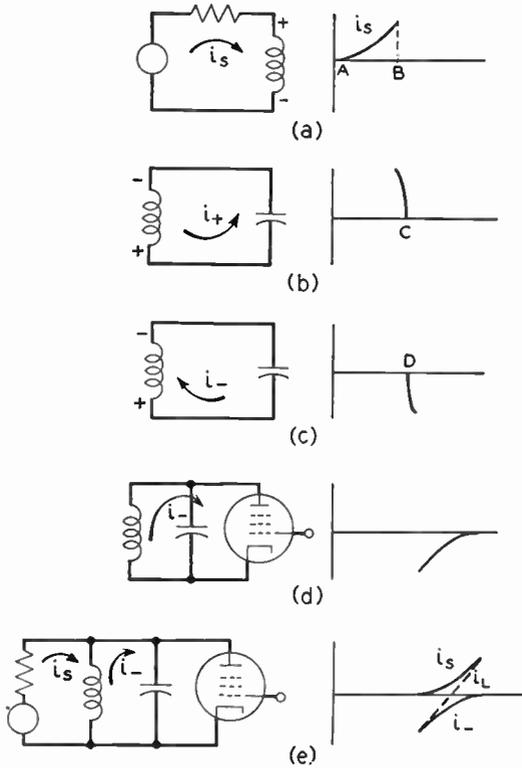


Fig. 4-47. Step-by-step operation of the reaction scanning circuit.

outset we note that the plate currents of both the driver and damper, V_1 and V_2 , respectively, will be determined by their respective grid voltages. For added clarity we shall assume V_2 cut off in the first τ_s interval from A to B . Thus from A to B the component of deflection current furnished by the driver, i_s , will build up linearly in time. At B , which defines the end of τ_s , V_1 and V_2 are cut off and the active circuit during τ_f consists wholly of the L - C combination which is free to oscillate. Hence as the field about L collapses, an oscillatory current i_+ flows as shown at b . This will be a damped cosinusoid, the damping being the result of the resistance inherently associated with the combined yoke and transformer inductance, L . Notice that at C , the current reverses direction and it is labeled i_- to avoid confusion, the negative sign indicating that the current is flowing upward through L as at c . Notice further that even though the current has reversed direction, its slope remains negative so that the voltage polarity across L remains unchanged.

At d the current has completed its reversal, that is, one half-cycle of the oscillation is completed, which incidentally defines τ_f . Then V_2 is made to conduct and i_- must follow the grid voltage on V_2 as far as shape is concerned instead of continuing as a cosinusoid. Hence, i_- may be made linear, and its slope is such that the voltage across L reverses polarity again as shown at d . Here is the essence of the whole system, for this component of current i_- , which is caused by the energy stored in the L - C network, is constrained by V_2 to follow a linear path, *which contributes to the total yoke current during the scan* instead of being damped out in a resistance.

But notice that at τ_f , V_1 also conducts, causing i_s to flow as previously described. Hence, the resultant current i_L which flows through L is the algebraic sum of i_s and i_- during the scan interval. This is shown at e . All of the wave forms involved are shown together in Fig. 4-48, where it may be seen that i_- contributes to the left-hand portion of the sweep and i_s to the right-hand portion.

The prime advantage of the circuit over those previously described is immediately apparent: the driver must furnish only one component, i_s , of the total current instead of the peak-to-peak value, \hat{i} , because the energy stored in the L - C network has been put to use rather than wasted. One might assume from Fig. 4-48 that i_s is only one-half of \hat{i} . Actually, however, since the current during τ_f is damped, the peak value of i_- is only some 60 to 80 per cent of the peak value of i_+ .

Thus the driver must furnish something over one-half of the sweep current, but the reduction in deflection power over that in the systems previously defined is quite significant. Still another advantage in the present system is that i_- may be used to iron out nonlinearity in i_s at the center of the sweep by suitable shaping of the grid voltage supplied to V_2 . This compensation can yield a net current i_L , which possesses excellent linearity.

On the other hand, it must be pointed out that the circuit is relatively complicated and expensive. Also, means must be provided for generating the V_2 control grid voltage. This voltage may be obtained from the square wave portion of e_L shown in Fig. 4-48d. The complete circuit is shown in Fig. 4-49. C_1 and R_1 comprise the circuit for developing the control grid voltage. P_1 and P_2 adjust the d-c component of damper grid voltage and so serve as linearity controls. P_3 controls the d-c current through the deflection coils and hence is the positioning control which determines the "rest" position of the electron beam in the horizontal direction.

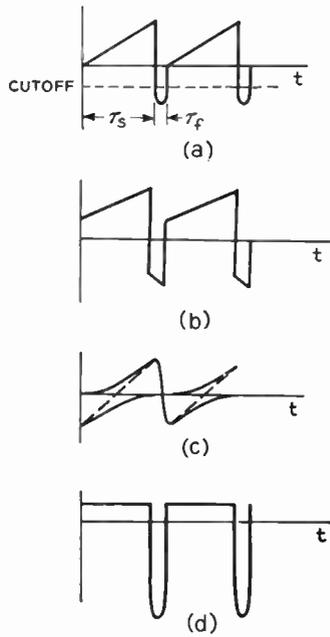


Fig. 4-48. Wave forms in the reaction scanning circuit. (a) Driver grid voltage. (b) Damper grid voltage. (c) Currents. (d) Voltage across the inductance.

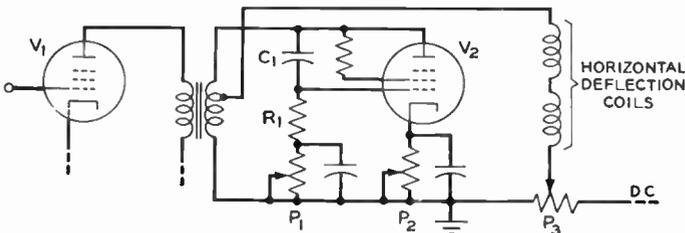


Fig. 4-49. Complete damping system for horizontal reaction scanning.

The added cost of the grid-controlled damper tube and its control circuit is quite justified in equipment used at the pickup end of the television system because it does provide a high degree of trace linearity. In the manufacture of television receivers which must sell in a highly competitive market, however, it is desirable to reach some compromise between the high cost and high efficiency. One such compromise is to replace the grid-controlled damper by a high-perveance diode.¹⁹ This allows a less expensive tube to be used, and the grid control circuit may be eliminated. The resulting circuit, which is shown in Fig. 4-50, closely resembles that previously described

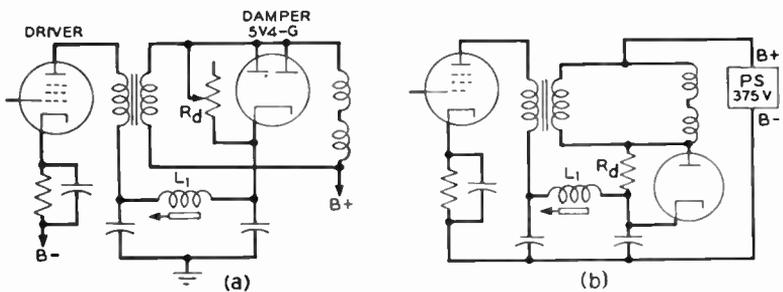


Fig. 4-50. Reaction scanning with a high-perveance diode damper.
(a) Basic circuit. (b) Circuit redrawn for convenience.

in the section on diode damping, but differs from it in that the damping of the flyback transient during τ_a is controlled so that once again it contributes to the magnitude and linearity of the final deflection current.²⁰

A disadvantage is incurred in the use of the diode in that the trace linearity is poor during the first 10 per cent of the sweep. The additional elements shown in the diagram serve to compensate for this, R_d providing a means of controlling the voltage applied across the damper tube by the sweep circuit and is set to the proper value at the factory. L_1 , which is variable, serves in the same capacity.

¹⁹ The term "perveance" is frequently neglected in electronics. In the empirical form of Child's equation for a space-charge-limited diode

$$i_b = K e_b^\alpha$$

K , a constant determined by the electrode geometry of the tube, is the perveance. See, for example, Y. Kusunose, "Calculation of Characteristics and the Design of Triodes," *Proc. IRE*, 17, 1706 (1929).

²⁰ A. W. Friend, *Television Deflection Circuits*, Part II, *op. cit.*

A final point to be mentioned about the circuit is that it utilizes the d-c voltage developed across the yoke during τ_s , the possibility of which was pointed out in the section 4-10. How this is effected may be seen by reference to Fig. 4-50*b*, where it is pointed out that the B supply voltage is furnished to the driver plate through the damper diode. Thus, since the damper and yoke are in series with the B supply, the voltage developed across the yoke will be added to $B+$ when the diode is conducting, *i.e.*, during τ_s . We have previously seen that the voltage developed in this interval is constant and in typical circuits has a value of some 50 volts. Thus this booster circuit furnishes an increase of 13 per cent over the power supply voltage to the driver plate circuit.

4-19. The Flyback Power Supply²¹

In those damping systems which permit one half-cycle of oscillation during the flyback interval a half-sine wave pulse of voltage is developed across the yoke in this same interval. This had been shown in Fig. 4-45 and the magnitude of the pulse is given by eq. (4-108). Since the transformer is wound step-down from primary to secondary with a turns ratio n , this voltage will also appear on the primary side of the transformer but with n times the amplitude. Its peak value, which occurs at the center of τ_f , will be

$$\hat{e} = nL\dot{i}\omega \quad (4-109)$$

and in typical circuits for, say, a 10-in. cathode-ray tube will range in the order of 3000 v for a sweep frequency of 15.75 kc. Furthermore its polarity is positive to ground and hence may be used as the voltage source for a pulse type d-c power supply. The current capabilities of such a supply would be small because of the low energy content of the individual pulses supplied, but since the beam current of typical cathode-ray devices lies below 100 μ a, such a power supply is ideal as the second anode supply for such a device. It is relatively inexpensive, light as compared to a 60-cycle iron-core transformer

²¹ The author is indebted to Drs. P. C. Goldmark and Kurt Schlesinger, both then of the Research and Development Division of the Columbia Broadcasting System, for making available to him early in 1946 preliminary data on the flyback power supply developed for the CBS Color Television receivers. Since that time a number of references have become available in the literature; for example, M. Cawwin, "High Voltage Power Supply for Television Receivers," *Radio News* (Radio Electronic Engineering Edition), **35**, No. 6, June 1946.

supply, and has the added highly desirable feature of being comparatively safe because of its poor regulation. A word of caution must be added, however. A high-voltage power supply is never completely safe; it should always be disconnected whenever work is done on associated equipment.

The basic circuit for the power supply is shown at Fig. 4-51a. Two prime difficulties are present in the basic circuit, the first of which is apparent from the diagram: The rectifier cathode is at high voltage relative to ground and hence requires a well-insulated filament supply,

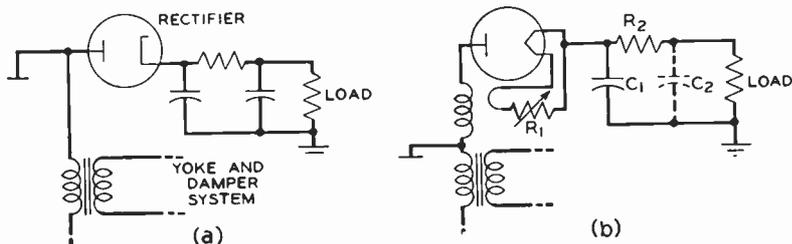


Fig. 4-51. The flyback power supply. (a) Basic circuit. (b) Typical circuit with an autotransformer to step up the pulse voltage applied to the rectifier.

separate from that furnishing the other tubes in the equipment. To this end special rectifier tubes which require only $\frac{1}{4}$ w of filament power have been developed, for example the 1Y3 by Chatham Electronics and the 8016 by RCA. This low power may be obtained from the pulse source itself by loosely coupling a single turn to the driver output transformer as shown in Fig. 4-51b. Proper adjustment is secured by varying R_1 until the filament exhibits a dull red color.

The second difficulty with the basic circuit is that its 3000-v capabilities are below the demands of the more recent cathode-ray tubes. This is overcome by adding an autotransformer winding with a 3 to 1 step-up ratio on the primary side of the driver transformer. With this device the d-c output from the rectifier is approximately 9000 v.

The high voltages developed require a driver tube with a high inverse peak rating, which may be had in the type 6BG6-G, a special form of the more familiar 807 beam power tube. Filtering problems in the power supply are negligible. Because of the low current drain imposed by the cathode ray tube R_2 may be used in place of an inductance as the series filter element. A resistance at this point is

also advantageous because it results in poor voltage regulation and added safety to personnel. The power supply is operated at line or horizontal sweep frequency so the fundamental ripple component of the voltage output is high, say above 15 kc; hence C_1 generally is in the order of 500 micromicrofarads. Special ceramic condensers have been developed for this type of service. The output capacitance, C_2 , is shown dotted in the diagram because it is furnished by either the capacitance-to-grounded-shield of the high voltage lead or by the capacitance between the aquadag coatings on the interior and exterior surfaces of the cathode-ray tube envelope.

In certain types of service, notably those employing the type 5TP4 projection cathode-ray tube, voltages far in excess of 9 kv are required. It might appear that the flyback power supply might be modified to deliver these higher voltages by the simple expedient of raising the step-up ratio of the autotransformer above the 3 to 1 value previously

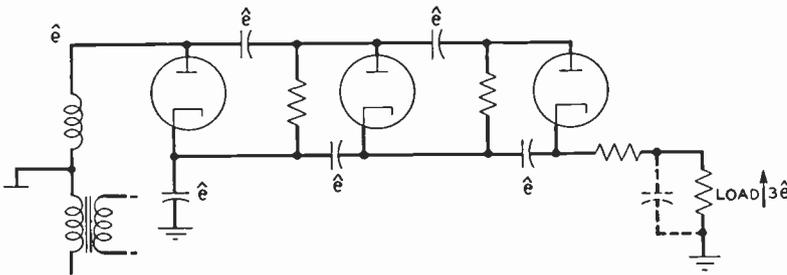


Fig. 4-52. Flyback voltage tripler.

specified. As a practical matter, this is not feasible because of the deleterious effects of shunt capacitance which would become prohibitively large with the larger step-up ratio. The solution to the problem lies in utilizing a voltage multiplier circuit in place of the simple rectifier shown. A typical voltage tripler of this type is shown in Fig. 4-52. Since the three lower condensers are each charged to approximately \hat{e} , the load voltage which is across all three in series will be three times as great. Practical values of output voltage for this circuit range up to some 27 kv, but this may be raised by cascading additional multiplier stages.

A highly desirable feature of all the flyback-type power supplies is that if the horizontal sweep should fail, the accelerating voltage is removed from the cathode-ray tube, thereby preventing the undeflected electron beam from burning the fluorescent screen.

CHAPTER 5

SCANNING AND PICTURE REPRODUCTION^{1,2}

We saw in Chapter 1 that the principle of scanning is basic in the transmission of television images and that in order to convert the picture image into an electrical image the picture must be scanned by some sort of a defining aperture backed by a photoelectric transducer. More precisely, if we consider the original picture to be a still photograph described by

$$B = f(x,y) \quad (5-1)$$

this must be transformed to give a corresponding electrical signal which is a function of time. Then since the scanning pattern is orderly and prearranged, every instant of time t is related to a specific space co-ordinate pair (x,y) , and there results a corresponding electrical signal

$$E = f(t) \quad (5-2)$$

which may be delivered to the communication channel. This signal for the picture assumed repeats itself after each frame interval and hence may be expanded in a Fourier series. As might be expected, the frequencies of the several Fourier components are independent of the picture content, being determined solely by the scanning frequency, and their amplitudes and phase are determined by the picture, being independent of the scanning process. In the present chapter we shall evaluate the frequency components to determine the basic picture structure, and then investigate the effect of a finite scanning aperture on the signal and on the reproduced television image.

To simplify the discussion we shall consider a simple facsimile

¹ P. Mertz and F. Gray, "A Theory of Scanning and Its Relation to the Characteristics of the Transmitted Signal in Telephotography and Television." *B.S.T.J.*, XIII, 3 (July 1934).

² P. Mertz, "Television—The Scanning Process." *Proc. IRE*, 29, 10 (October 1941).

system as proposed in Fig. 1-7 and 1-8. This allows a good physical interpretation of the concept of the scanning aperture. The concept will be extended to include the aperture of an electron beam in the latter part of the chapter. Since the scanning process is similar in facsimile and television systems, except for the speeds at which the scan is carried out, the results are general and in all respects directly applicable to the television system. As a starting point we wish to investigate only the frequency spectrum of the generated signal. Thus to eliminate other effects we shall assume a point aperture, *i.e.*, a scanning aperture of infinitesimally small size. Furthermore,

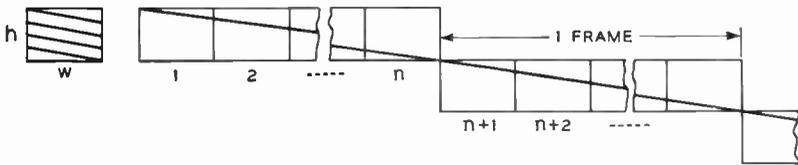


Fig. 5-1. The progressive scan pattern on a single image may be replaced by a single-line scan on an infinite array of identical images.

the retrace of the scanning spot, in both the horizontal and vertical directions, introduces mathematical difficulties without changing the frequency components. To overcome this difficulty we shall consider the actual picture to be replaced by an infinite array of identical images. By this device the actual *progressive* scan pattern on the single image may be replaced by a scan of constant speed and direction along a single line running across the infinite array of identical images. This is illustrated in Fig. 5-1.

5-1. The One-dimensional Series

If, for the moment, variations in intensity normal to the scan direction be neglected, the equation for the brightness along the scanning line may be represented by the Fourier series

$$B(x) = \sum_{k=0}^{\infty} b_k \cos \left(\frac{2\pi kx}{nw} + \theta_k \right) \tag{5-3}$$

Further, if the transducer now be made to move along the scanning line with a uniform velocity u , the voltage developed will be

$$E(t) = \sum_{k=0}^{\infty} a_k \cos(2\pi k f_p t + \theta_k) \quad (5-4)$$

where $f_p =$ frame frequency (5-5)

$$= \frac{1}{\text{time to scan } n \text{ images}} = \frac{1}{V_p}$$

Equation (5-4) results directly because x and t are directly related through the constant scanning velocity, u . Inspection of the equation shows that the frequencies present are all integral multiples of the frame frequency, f_p . This is to be expected because the frequency components in a Fourier expansion are always integral multiples of the lowest or fundamental repetition frequency, which in this case is the picture repetition rate, f_p .

Although the Fourier series of (5-4) is quite satisfactory from the point of view of the communication channel, which requires a single-valued function of time, neither it nor (5-3) contains information regarding the brightness variations in the direction normal to that of the scan. That is, neither expression contains information needed to reconstruct the two-dimensional image at the receiving end of the system. Thus, in place of the one-dimensional series, a two-dimensional one is required. Once this has been set up, the corresponding $E(t)$ may be written because, as we have seen, each value of t corresponds to a fixed point (x, y) on the image because of the orderly scanning pattern.

5-2. The Two-dimensional Series

To derive the bidimensional series, consider Fig. 5-2. Note that the positive direction of y is chosen downward to conform to our standard proposed in Chapter 2. Now along any single scanning line, say $y = y_1$, the variations in brightness may be represented by the trigonometric series

$$B(x, y_1) = \sum_{k=0}^{\infty} b_k \cos\left(\frac{2\pi k x}{w} + \theta_k\right) \quad (5-6)$$

in the interval $0 \leq x \leq w$

In a similar manner the variations along another scanning line, say $y = y_1$, may be represented by another series, similar in every respect

to (5-6) except that the amplitudes, b_k , will be different because the point aperture is moving along a different line. In other words, b_k is itself a function of y and hence may also be expanded in a Fourier series.

$$b_k = \sum_{l=0}^{\infty} b_{kl} \cos\left(\frac{2\pi ly}{h} + \theta_l\right) \quad (5-7)$$

This situation may become more clear by referring to Fig. 5-3, which shows a model of the brightness distribution in an arbitrary image composed of a number of concentric dark rings on a white background. In the model vertical height at each point is proportional to the brightness of the corresponding spot in the image. At c the model is split along $y = y_1$ and the brightness distribution along that line is given by (5-6). At d the block is split along a vertical line ($x = cst$) and we see that another distribution is apparent which is given by (5-7). It follows, then, that the distribution over the entire surface may be obtained by substituting the second of these equations into the first and there results

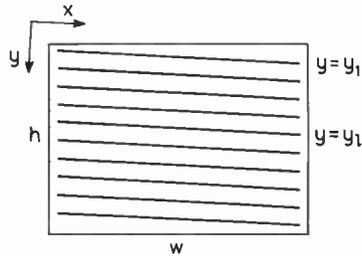


Fig. 5-2. The progressive scan raster. The positive directions of x and y are defined.

$$B(x,y) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} b_{kl} \cos\left(\frac{2\pi kx}{w} + \theta_k\right) \cos\left(\frac{2\pi ly}{h} + \theta_l\right) \quad (5-8)$$

and expansion of the cosine cosine product yields

$$B(x,y) = \sum_{k=0}^{\infty} \frac{b_{kl}}{2} \left\{ \sum_{l=0}^{\infty} \cos\left[2\pi\left(\frac{kx}{w} + \frac{ly}{h}\right) + (\theta_k + \theta_l)\right] + \sum_{l=0}^{\infty} \cos\left[2\pi\left(\frac{kx}{w} - \frac{ly}{h}\right) + (\theta_k - \theta_l)\right] \right\} \quad (5-9)$$

Further simplification results by changing the variable in the second summation within the braces from l to $-l$. This yields

$$B(x,y) = \sum_{k=0}^{\infty} \sum_{l=-\infty}^{\infty} \frac{b_{kl}}{2} \cos\left[2\pi\left(\frac{kx}{w} + \frac{ly}{h}\right) + \theta_{kl}\right] \quad (5-10)$$

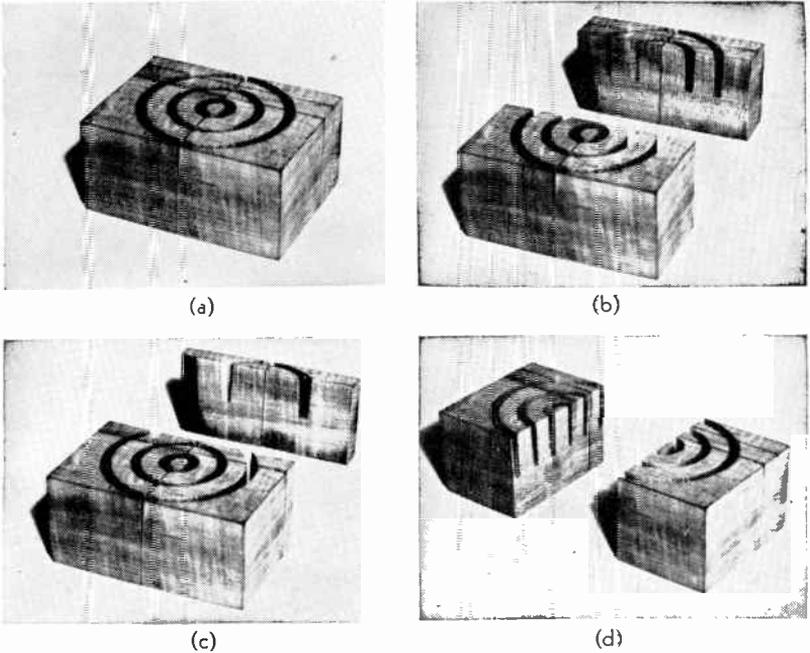


Fig. 5-3. A three-dimensional model illustrating the brightness variation across the image field. (a) The complete model. (b) The brightness distribution along an arbitrary line $y = y_2$. (c) The brightness distribution along another line $y = y_1$. (d) The brightness distribution along a vertical line $x = csl$.

where $\theta_{kl} = \theta_k + \theta_l \quad l > 0$
 and $\theta_{kl} = \theta_k - \theta_l \quad l < 0$

Equation (5-10) gives the brightness at any point (x,y) in the two-dimensional image in terms of a double Fourier series.

5-3. Interpretation of the Double Fourier Components

In the last equation b_{kl} is the amplitude of any single component (k,l) and may be evaluated in the same manner as is used for the one-dimensional series. We shall not attempt this actual evaluation because the work would be tedious and would contain no information about the frequency spectrum and the like. It would be desirable, however, to consider this question: What is the physical meaning of

each of the components of the double series? Now, in a one-dimensional Fourier series each component is a single sinusoidal wave. Extending this concept to cover the present case we see that each component here is a sinusoidal variation across an area. We can make this more clear by considering a specific component, say, for $k = 1$ and $l = 2$, and for further simplification we shall assume that $\theta_{12} = 0$. Then, if we consider the variation in x along the uppermost line where $y = 0$, we have from (5-10)

$$B_{12}(x,0) = \frac{b_{12}}{2} \cos\left(\frac{2\pi x}{w}\right) \tag{5-11}$$

which represents a single cycle of a cosine wave shown in Fig. 5-4a. Similarly, along another scanning line $y = y_1$ the equation is

$$B_{12}(x,y_1) = \frac{b_{12}}{2} \cos\left(\frac{2\pi x}{w} + \frac{2\pi y_1}{h}\right) \tag{5-12}$$

Since the second term is constant along the line, (5-12) again represents a cosine wave but shifted relative to the first by an angle $2\pi y_1/h$. This is shown at *b* in the diagram.

Along another line, $y = y_2$, we have still another cosine wave but shifted this time by an angle $2\pi y_2/h$ relative to the first as shown at *c*. Thus as we move downward across the image, the cosine wave suffers a progressive shift in phase toward the left. Now, remembering that these waves represent variations in brightness, we see that each component of the double series is a two-dimensional variation in brightness whose maxima and minima move toward the left as y increases, *i.e.*, downward in the diagram. For our specific example the brightness variation

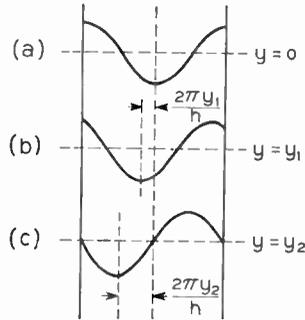


Fig. 5-4. For each value of y the cosine wave is shifted by an angle $2\pi y/h$ from the top wave.

along any vertical line, say $x = 0$, is, from (5-10),

$$B_{12}(0,y) = \frac{b_{12}}{2} \cos \frac{2\pi \times 2y}{h} \tag{5-13}$$

This, of course, represents 2 cycles of a cosine wave. The resulting component B_{12} is shown at *a* in Fig. 5-5. In the figure the relative

amplitude is represented by the closeness of the lines. The locus of a maximum lies where the lines are farthest apart.

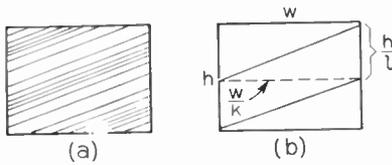


Fig. 5-5. (a) The Fourier component for the index pair $k = 1, l = 2$. (b) The slope of the striations may be calculated.

It is now a simple matter to extend this concept to the general component $k = k$ and $l = l$. From (5-11) and (5-13) we see that there will be k cycles of variation along a horizontal line and l variations along a vertical line. To aid in drawing these components we can calculate the slope of the maxima across the field.

In Fig. 5-5b the maxima of the striations are represented by the solid lines shown. Then, reading from the diagram,

$$\text{slope} = -\frac{kh}{lw} \quad (5-14)$$

The negative sign is introduced because the assumed positive direction of y is downward. With this information we can readily sketch the typical components of Fig. 5-6. The solid lines indicate the

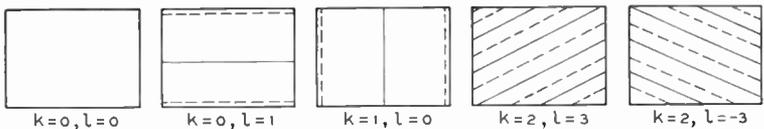


Fig. 5-6. Typical double Fourier components. θ_{kl} is assumed to be zero. Brightness maxima are represented by the solid lines and brightness minima by the dashed lines.

positions of the brightness maxima; the broken lines, the minima.

It is of interest to note that the concept of wave length, λ , of a picture component may be applied here since it is defined as the distance between two adjacent maxima in the direction normal to the line of the maxima. Thus

$$\lambda = \frac{1}{\sqrt{\left(\frac{k}{w}\right)^2 + \left(\frac{l}{h}\right)^2}} \quad (5-15)$$

In summary, then, we see that the concept of the Fourier series may be extended to cover a function of two variables. Thus any

picture, any distribution of brightness across a bounded area, may be considered to be composed of the superposition of an infinite number of field components of the type illustrated in Figs. 5-5a and 5-6. The wave length and slope of each such component is determined solely by the area of the picture. The amplitude of each component is determined by the picture content. As in one-dimensional Fourier analysis the amplitude and phase of every component in a still picture are fixed in time. A change in subject alters these component amplitudes and phases but has no effect on the wave length or slope.

5-4. The Electrical Signal and Its Spectrum

We have repeatedly stressed the point that the orderly scanning process relates each point (x,y) in the picture to a specific value of time, t . We can, therefore, write the equation for the output in time of the transducing element. Thus let

$$u = \frac{x}{t} = \text{horizontal scanning velocity} \quad (5-16)$$

$$v = \frac{y}{t} = \text{vertical scanning velocity}$$

Then, corresponding to (5-10), we have

$$E(t) = \sum_{k=0}^{\infty} \sum_{l=-\infty}^{\infty} \frac{a_{kl}}{2} \cos \left[2\pi \left(\frac{ku}{w} + \frac{lv}{h} \right) t + \theta_{kl} \right] \quad (5-17)$$

This is the electrical signal delivered to the communication channel and it does contain the information regarding brightness variations normal to the direction of scan.

In Chapter 1 we saw that the communication channel must be able to accommodate this signal. Consequently it is of importance to the engineer to investigate the frequency spectrum of this signal, which may be conveniently accomplished by introducing frequency into the last equation. Then, since in the progressive scan, which has been assumed, the vertical height of the picture is scanned in the time V_p ,

$$v = \frac{h}{V_p} = hf_p \quad (5-18)$$

and in the same interval n lines are scanned or the total horizontal distance covered is nw . Hence,

$$u = nwf_p \quad (5-19)$$

Substitution in (5-17) yields

$$E(t) = \sum_{k=0}^{\infty} \sum_{l=-\infty}^{\infty} \frac{a_{kl}}{2} \cos [2\pi(knf_p + lf_p)t + \theta_{kl}] \quad (5-20)$$

which has a frequency spectrum

$$k(nf_p) + lf_p \quad (5-21)$$

Notice that the first term gives multiples of the line, or horizontal, scanning frequency nf_p ; the second term gives multiples of the frame frequency. Hence the spectrum of the electrical signal consists of groups of frequencies, each group being centered on a multiple of the line frequency. In each group adjacent frequencies are separated by the frame frequency. A portion of a typical spectrum is sketched in Fig. 5-7. Amplitudes of the individual components are chosen

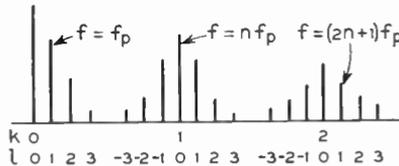


Fig. 5-7. Portion of the spectrum (5-21). Frequency groups are centered on multiples of the line frequency. Adjacent frequency components differ by f_p .

arbitrarily. It should be noticed, however, that the amplitudes decrease as either k or l increases, a necessary condition if the series is to converge.

Inspection of the figure would lead one to surmise that blank regions occur in the spectrum of the electrical system. The existence of such regions has been verified by actual measurement on the signal. It seems unusual, however, that such null regions are present between frequency groups; one might readily expect them to contain components of low amplitude. Actually the existence of such weak amplitudes may cause confusion in the reproduced image, and suitable

aperture choice may eliminate them as we shall see in the following sections.

5-5. Confusion in the Signal

In drawing the spectrum of Fig. 5-7 it was assumed that the picture content was such that l remained sufficiently small that adjacent frequency groups were separated by a null region in the spectrum. The assumption to the contrary, it might well be that two field components with indices (k,l) and (k',l') exist, such that

$$nk + l = nk' + l' \quad (5-22)$$

Under this condition both field components have frequencies in common, *i.e.*, the adjacent group side bands overlap. Then, since the equipment cannot distinguish between the coincident components, confusion results, a condition which is illustrated later in the chapter. A typical spectrum of this type is shown in Fig. 5-8 where

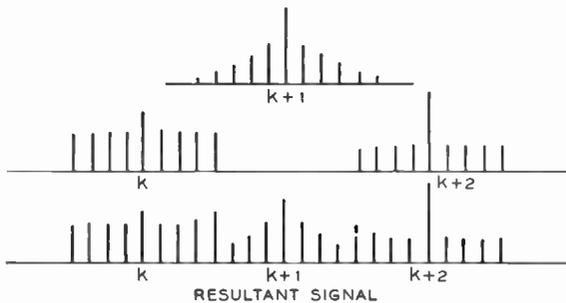


Fig. 5-8. Adjacent frequency groups may overlap, causing confusion.

adjacent frequency groups are shown on different levels for clarity. This overlapping of frequency components with its attendant confusion may be eliminated by insertion into the system at the proper point of a device which cuts off all frequencies corresponding to l in excess of $n/2$. This may be seen quite readily from the following considerations. The spacing between adjacent group center frequencies is nf_p . Then to prevent overlap the side bands of any group must be limited to $nf_p/2$. Hence,

$$lf_p < \frac{nf_p}{2} \quad (5-23)$$

which verifies the statement.

Notice that the suitable filtering device cannot be placed in the electrical portions of the over-all system because the transducer output contains the coincident confusion components, which—once generated—cannot be separated. It follows, therefore, that the filtering must be done between the original picture and the transducer, *i.e.*, by the aperture itself. It is fortunate that the aperture, which has been assumed to be a point thus far in the discussion, when properly designed eliminates these overlapping frequency components. Consideration will be given to this problem after we have investigated the effect of a finite rather than a point aperture on the picture signal.

It must be stressed that the various spectra discussed above are for a still picture. If the picture content varies in time, the amplitudes a_{kl} will also vary. Hence each frequency component will be amplitude-modulated with the result that additional, new side-band frequencies will be added in the spectrum. The analysis for such a varying picture becomes too involved to be of value.

5-6. The Effect of Interlaced Scanning

The work of the previous sections has been derived for a progressive scan pattern. We must now consider what effect, if any, a 2 to 1 interlaced scan would have on the frequency components of the

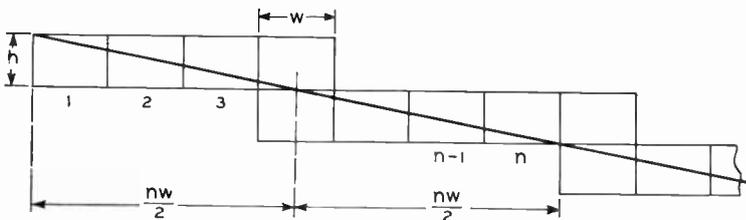


Fig. 5-9. The 2 to 1 interlaced scan pattern may be replaced by a single-line scan on an infinite array of identical images.

picture and the frequency spectrum of the corresponding electrical signal. In order to eliminate the mathematical difficulties associated with flyback, we shall assume once again an array of repeated

pictures. The particular configuration to be used is given in Fig. 5-9 which also shows a scanning path which meets the interlace requirements. Then, following the method used for the progressive scan, we observe that the brightness along the line over a width w is still given by eq. (5-6). Since, however, a vertical height of $2h$ is covered in scanning one complete picture, b_k now becomes

$$b_{k^*} = \sum_{l=0}^{\infty} b_{k^*l^*} \cos\left(\frac{2\pi l^*y}{2h} + \theta_{l^*}\right) \tag{5-24}$$

And, finally, the double series is

$$B(x,y) = \sum_{k^*=0}^{\infty} \sum_{l^*=-\infty}^{\infty} \frac{b_{k^*l^*}}{2} \cos\left[2\pi\left(\frac{k^*x}{w} + \frac{l^*y}{2h}\right) + \theta_{k^*l^*}\right] \tag{5-25}$$

where

$$\begin{aligned} \theta_{k^*l^*} &= \theta_{k^*} + \theta_{l^*}, & l^* > 0 \\ \theta_{k^*l^*} &= \theta_{k^*} - \theta_{l^*}, & l^* < 0 \end{aligned}$$

We use the notation k^* and l^* to indicate the set of indices for any Fourier component in the interlaced scanning system. The need for these additional symbols may be seen from the following considerations. The Fourier components for $k = 1, l = 1$ and $k^* = 1, l^* = 1$ are not the same. The first pair, which corresponds to the progres-

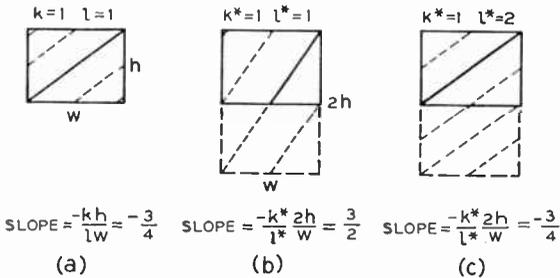


Fig. 5-10. Comparison of the Fourier components for progressive and interlaced scanning. For a given component in the picture $l^* = 2l$. (a) The Fourier component for $k = 1, l = 1$ in the progressive scan. Slope $= -kh/lw = -\frac{3}{4}$. (b) The component $k^* = 1, l^* = 1$ for interlaced scanning differs from (a). A total height of $2h$ is shown to simplify the construction. Slope $= -k^*2h/l^*w = -\frac{3}{2}$. (c) l^* must be doubled to produce the component shown at (a) in interlaced scanning. Slope $= -k^*2h/l^*w = -\frac{3}{4}$.

sive scan, represents a component with one cycle of brightness variation across each dimension of the picture area, while the second set corresponding to interlaced scanning represents a component with one cycle of variation across the picture width and one cycle of variation across *twice* the picture height. These facts may be verified from the arguments of eq. (5-17) and (5-25). The two Fourier components are illustrated at *a* and *b* in Fig. 5-10.

It follows at once, then, that if *k* and *l* be used to indicate a specified Fourier component in the picture, the corresponding index pair for that component in the interlaced scan will be

$$\left. \begin{aligned} k^* &= k \\ l^* &= 2l \end{aligned} \right\} \quad (5-26)$$

Then, to evaluate the frequency spectrum of the electrical signal corresponding to (5-25), we need only convert *x* and *y* into terms of frequency and time, thus,

$$\text{and} \quad \left. \begin{aligned} u &= nvf_p \\ v' &= 2hf_p \end{aligned} \right\} \quad (5-27)$$

$$\text{whence } E(t) = \sum_{k^*=0}^{\infty} \sum_{l^*=0}^{\infty} \frac{a_{k^*l^*}}{2} \cos [2\pi(k^*nf_p + l^*f_p)t + \theta_{k^*l^*}] \quad (5-28)$$

If, now, (5-26) be utilized, we have for the electrical signal corresponding to the (*k*,*l*) component in the picture

$$E(t) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{a_{kl}}{2} \cos [2\pi(knf_p + 2lf_p)t + \theta_{kl}] \quad (5-29)$$

Therefore we see that in interlaced scanning the spectrum of the electrical signal consists of frequency groups centered on multiples of the line frequency (knf_p), but adjacent frequencies in each group are separated by twice the frame frequency of field frequency. A comparison of the spectra for a given set of indices resulting from both types of scan are illustrated in Fig. 5-11. In the figure *n* is chosen to be 9 for simplicity; the progressive scan spectrum is shown at the top of the diagram and the interlaced scan spectrum at the bottom.

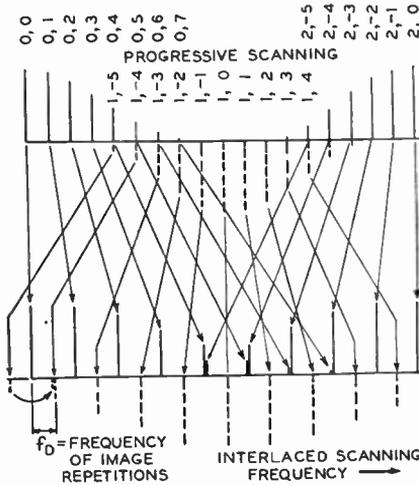


Fig. 5-11. Comparison of the spectra for nine-line progressive and interlaced scan systems. The arrows indicate components that correspond to a given index pair in the original picture. (Courtesy of *Proc. IRE.*)

5-7. Aperture Distortion

Our previous analysis has yielded an expression for the voltage that results from scanning a still picture with a photoelectric transducer through a point aperture. From this we found that the frequencies in the signal are functions of the scanning process alone. Further, the amplitude and phase of each component are determined solely by the picture itself. Now ideally the picture transmission system, whether television or facsimile, should reproduce at the receiving point a picture having the same components in frequency, amplitude, and phase that are present in the original picture. As a practical matter, of course, this ideal condition is not met. The response characteristics of the communication channel proper modify the relative amplitude and phase of these components. Furthermore at both ends of the system the scanning apertures are of finite size and also serve to modify the component amplitudes. This latter effect is called aperture distortion, and in section 5-12 we shall derive a quantitative expression for its effect. Our over-all method shall be to derive expressions for a_{kl} , θ_{kl} , $a_{k'l'}$, and $\theta_{k'l'}$, the primes referring to the

receiving end of the system. Since an ideal communication channel will be assumed, the differences between the original and reproduced components will be a measure of the distortion introduced by the finite aperture size.

5-8. Physical Concept of Aperture Distortion

Before proceeding to an analysis of the distortion it is well to get a physical picture of its effect. In Fig. 5-12 the original picture is assumed to be a square white spot on a black line, whose brightness distribution in x is shown at *b*. When scanned through an infinitesimally narrow aperture, the abrupt changes in brightness at the boundaries of the white spot are followed exactly by the transducer whose output is given at *c*. If, on the other hand, the transducer sees the picture through an aperture of finite width, it cannot reproduce the abrupt change at the boundaries. At each position of the aperture the transducer output is proportional to the average value of brightness which it sees. Hence its output reaches the maximum gradually, as shown at *d*. It will be observed that the effect of the aperture on the signal is similar to that of an ideal low-pass filter which has no delay or phase distortion: an applied square pulse has its corners rounded off.

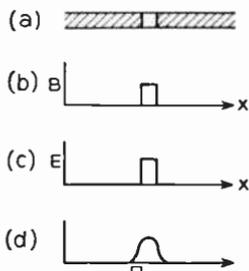


Fig. 5-12. Aperture distortion. (a) Original picture. (b) Brightness distribution in x . (c) Voltage after scanning by a point aperture. (d) Voltage after scanning by an aperture of finite width.

5-9. A Digression on Exponential Series

In the work to follow considerable simplification results if the Fourier series is expanded in terms of exponential rather than trigonometric components. The method of deriving the expression for the component amplitudes is similar to that used in the more familiar sine and cosine form of series. Thus say that a function $F(\phi)$, which is periodic in 2π , is to be expanded into an infinite sum of exponential terms which are harmonically related, that is, we assume the identity

$$F(\phi) = \sum_{k=-\infty}^{\infty} A_k \epsilon^{jk\phi} \quad (5-30)$$

where k is an integer.

In order for (5-30) to be an identity the complex amplitudes A_k must be evaluated correctly. To do this we multiply both sides of the equation by $\epsilon^{-jr\phi}d\phi$ and integrate between the limits 0 and 2π (or $-\pi$ and $+\pi$)

$$\int_0^{2\pi} F(\phi)\epsilon^{-jr\phi}d\phi = \sum_{k=-\infty}^{\infty} \int_0^{2\pi} A_k\epsilon^{j(k-r)\phi}d\phi \tag{5-31}$$

On expanding the right-hand member we find that two types of integrals are present

$$\left. \begin{array}{l} \text{(a) } k \neq r \\ \quad k = -r \end{array} \right\} \text{for which the integral is zero since the} \\ \left. \begin{array}{l} \text{(b) } k = r \end{array} \right\} \text{integrand is periodic in } 2\pi. \tag{5-32}$$

$$\text{for which } \int_0^{2\pi} A_k\epsilon^{j\phi}d\phi = A_k \times 2\pi$$

Therefore, (5-31) reduces to

$$A_k = \frac{1}{2\pi} \int_0^{2\pi} F(\phi)\epsilon^{-jk\phi}d\phi \tag{5-33}$$

and the amplitudes are evaluated.

The same result may be obtained by substituting Euler's identity into the trigonometric expansion of $F(\phi)$. This latter method does not require the intelligent guess of multiplying the equation through by $\epsilon^{-jr\phi}$ but it involves considerable algebraic manipulation. The series (5-30) is equivalent in every way to the more familiar trigonometric expansions. A_k will generally be complex, that is, it includes the phase-angle term of the sine or cosine series form, except where $F(\phi)$ is an odd or even function.

In the work to follow the exponents contain several terms. We shall, therefore, write the exponential in the notation of the complex variable

$$\exp \phi \equiv \epsilon^\phi$$

5-10. One-dimensional Case

For simplicity we shall begin the analysis of aperture distortion with the single-dimension case. This will then be extended to the more general one. Consider first the situation at the pickup end of the picture transmission system where the scanned picture has a brightness distribution along the scanning line as shown in Fig.

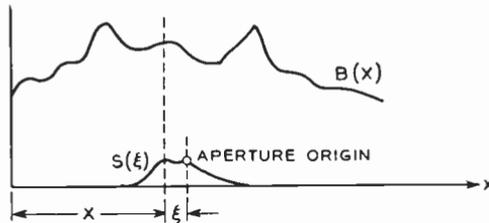


Fig. 5-13. Co-ordinate system for the finite aperture.

5-13. This brightness distribution may be represented by the trigonometric series (5-3) or by its exponential equivalent

$$B(x) = \sum_{k=-\infty}^{\infty} B_k \exp\left(\frac{2\pi j k x}{nw}\right) \quad (5-34)$$

If, now, this were scanned by a transducer through a point aperture, the corresponding electrical signal would have a form which we have previously derived. If, on the other hand, a finite aperture is used, a different condition obtains. Let the *response* of the aperture itself be that shown in Fig. 5-13 and further let it be represented mathematically by $S(\xi)$ where ξ is measured from the origin of the aperture which may be chosen arbitrarily. Then the brightness distribution seen by the transducer will no longer be $B(x)$ but, say, $B_1(x)$. The reason for this will be clear when it is realized that when the aperture is centered on x , the transducer sees the integrated brightness over the whole width of the aperture. Thus the apparent brightness $B_1(x)$ depends on both the original distribution, $B(x)$, and on the aperture response, $S(\xi)$. These ideas may be expressed analytically. Thus, if we let x be the co-ordinate of the aperture origin, the brightness of the original picture at the point $(x + \xi)$ will be by (5-34)

$$B(x + \xi) = \sum_{k=-\infty}^{\infty} B_k \exp\left[\frac{2\pi j k (x + \xi)}{nw}\right] \quad (5-35)$$

and the apparent brightness at x will be

$$B_1(x) = \int_{\text{aperture}} S(\xi) B(x + \xi) d\xi \quad (5-36)$$

Actually the integral should be taken over the entire domain of ξ but

the contribution to the integral for all ξ for which $S(\xi) = 0$ is zero. The \int_{aperture} is the only part of the total integral which is different from zero.

If, now, (5-35) be substituted into (5-36) and the factor in x , which is constant in the integration, be removed from within the integral sign, there results

$$B_1(x) = \sum_{k=-\infty}^{\infty} Y(k)B_k \exp\left(\frac{2\pi j k x}{nw}\right) \tag{5-37}$$

where
$$Y(k) = \int_{\text{aperture}} S(\xi) \exp\left(\frac{2\pi j k \xi}{nw}\right) d\xi \tag{5-38}$$

and the electrical signal at x corresponding to the apparent brightness is

$$E_1(x) = \sum_{k=-\infty}^{\infty} Y(k)A_k \exp\left(\frac{2\pi j k x}{nw}\right) \tag{5-39}$$

or, in time,
$$E_1(t) = \sum_{k=-\infty}^{\infty} Y(k)A_k \exp\left(\frac{2\pi j k u t}{nw}\right) \\ = \sum_{k=-\infty}^{\infty} Y(k)A_k \exp(2\pi j k f_p t) \tag{5-40}$$

In both of the last two equations the subscript 1 is used to indicate that the voltage corresponds to the apparent brightness $B_1(x)$ seen through the finite aperture.

Comparison of (5-40) to (5-4) shows that the effect of the finite aperture is to multiply A_k of each frequency component (*i.e.*, the complex amplitude corresponding to each component in the original image) by a quantity, $Y(k)$, which is a function of the aperture size and response. The same effect could be obtained electrically by cascading with the communication channel a linear network which has a transfer admittance $Y(k)$; hence, $Y(k)$ defined by (5-38) might properly be termed the "aperture admittance."

Frequently the aperture response $S(\xi)$ is symmetrical about its origin or, in other words, it is an even function in ξ . Where this is

true, the imaginary part of the expanded exponential in (5-38) is zero, and the aperture admittance reduces to

$$Y(k) = \int_{\text{ap}} S(\xi) \cos\left(\frac{2\pi k\xi}{nw}\right) d\xi \quad (5-41)$$

where $S(\xi)$ is symmetrical.

Since the imaginary part is zero under this condition $Y(k)$ is real and hence introduces zero phase shift and hence no phase distortion.

Let us calculate the admittance of a typical aperture. Assume that a rectangular mechanical aperture of width δ is used. Its response will be that shown at Fig. 5-14a. Since the response is symmetrical in ξ , we have from (5-41)

$$\begin{aligned} Y(k) &= 2 \int_0^{\delta/2} \hat{S} \cos\left(\frac{2\pi k\xi}{nw}\right) d\xi \\ &= \hat{S}\delta \frac{\sin\left(\frac{\pi k\delta}{nw}\right)}{\left(\frac{\pi k\delta}{nw}\right)} \end{aligned} \quad (5-42)$$

which has the $(\sin x)/x$ form. The envelope of this function is plotted at *b* in the figure.

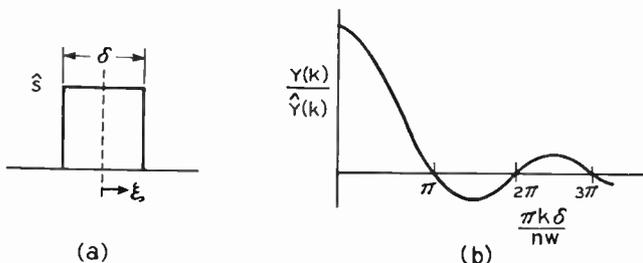


Fig. 5-14. (a) Rectangular aperture response characteristic. (b) Normalized aperture admittance $Y(k)$ for the response shown at (a).

It is interesting to note that $Y(k)$ for several common aperture shapes closely resembles that shown for the rectangular response, at least for values of k up to that of the first zero. It should be further noted that physically $Y(k)$ will not reverse sign as indicated in the figure.

The student might well wonder at this point what type of mechanical aperture shape would give a response different from that of Fig. 5-14a. In answer to this a special aperture and its response are shown at Fig. 5-15. This particular form has been used to synthesize

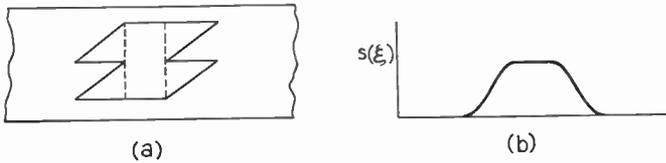


Fig. 5-15. A special scanning aperture. (a) Aperture shape. (b) Its response, $S(\xi)$.

mechanically the response of the electron beam effective aperture encountered in cathode-ray devices.³ This type of response for an electron beam is discussed at the end of the chapter.

5-11. Reconstruction of the Image at the Reproducer

Equation (5-40) gives the signal delivered to the communication channel by the transducer. If the assumed electrical response of this channel is taken to be ideal, the entire signal will be modified by a constant factor W and the signal applied to the reproducing transducer will be

$$E_1'(t) = WE_1(t) \quad (5-43)$$

and from this the transducer will produce an instantaneous brightness

$$b'(t) = PWE_1(t) \quad (5-44)$$

where P is a proportionality constant relating light to voltage. The primes indicate quantities at the receiving end of the over-all system.

Now this light is displayed to the final viewer through an aperture of finite size and of response $S'(\xi)$; hence the apparent brightness at any point x along the playback scanning line will be the integral of the product of transducer brightness and aperture response over that length of time from $-\tau_1$ to $+\tau_2$ required for the entire aperture to move past that point. This is illustrated in Fig. 5-16. A suitable change in the limits of integration results in simplification of the mathematical manipulation which is to follow. This change is

³P. C. Goldmark and J. N. Dyer, "Quality in Television Pictures" *J.S.M.P.E.*, XXXV (September 1940).

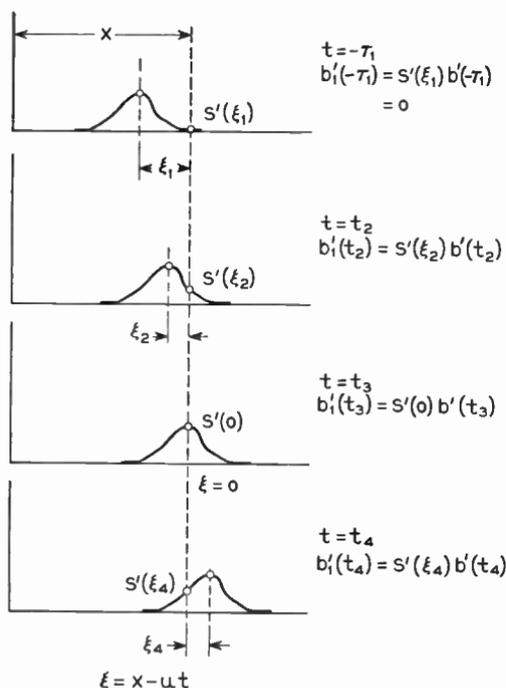


Fig. 5-16. Each point of $S'(\xi)$ contributes to the reproduced brightness at the point x .

brought about by noting that in assuming the single-line scan over an infinite array of identical image fields as shown in Fig. 5-1, the aperture moves past a given point x only once throughout all time from the remote past to the ever-so-distant future. It follows, therefore, that the contribution by the transducer to the apparent reproduced brightness at that point is zero for all times not in the closed interval $-\tau_1 \leq t \leq +\tau_2$, and the limits of integration may be changed to $-\infty$ and $+\infty$ with no change in the result. Then, integrating, we get for the reproduced brightness at the point x

$$B_1'(x) = \int_{-\infty}^{\infty} S'(\xi) PWE_1(t) dt \quad (5-45)$$

Equation (5-45) gives the reproduced brightness at x in terms of the incoming signal and hence, by suitable substitution, of the compo-

nents in the original picture. But this brightness may also be expressed in a Fourier series in terms of the components of complex amplitude $B_{k'}$ of the reproduced picture. Thus,

$$B_1'(x) = \sum_{k=-\infty}^{\infty} B_{k'} \exp\left(\frac{2\pi j k' x}{nw}\right) \tag{5-46}$$

From (5-33) we know that the amplitude of each component is

$$B_{k'} = \frac{1}{nw} \int_{-nw/2}^{+nw/2} B_1'(x) \exp\left(-\frac{2\pi j k' x}{nw}\right) dx \tag{5-47}$$

Once again mathematical simplification results from a change in limits. Since the exponential in (5-46) is periodic in intervals of nw , it reproduces itself identically at intervals of nw . Hence we may rewrite the equation as

$$B_{k'} = \frac{1}{nw} \int_{-\infty}^{\infty} B_1'(x) \exp\left(-\frac{2\pi j k' x}{nw}\right) dx \tag{5-48}$$

Now in order to observe the effect of the apertures on the reproduced image we wish to compare $B_{k'}$ to B_k for each value of k . It should be apparent that this may be effected by substituting for $B_1'(x)$ in terms of the original picture components. Hence, substituting (5-45) into (5-48) we get

$$B_{k'} = \frac{1}{nw} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S'(\xi) PW E_1(t) \exp\left(-\frac{2\pi j k' x}{nw}\right) dt dx \tag{5-49}$$

Consider now the amplitude of a reproduced image component (k') due to a single component (k) in the original picture. That is, we shall substitute in (5-49) for the k th component of $E_1(t)$ from (5-40). This yields

$$B_{k'} = \frac{1}{nw} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S'(\xi) PW Y(k) A_k \exp\left(\frac{2\pi j k u t}{nw}\right) \exp\left(\frac{-2\pi j k' x}{nw}\right) dt dx \tag{5-50}$$

This is the equation we wish to evaluate so that $B_{k'}$ and A_k (or B_k , its brightness equivalent) may be compared. The problem as far as integration is concerned is that x , u , t , and ξ are all interrelated.

Since identical scans are assumed at both ends of the system, u , the horizontal velocity, remains the same. Further, from Fig. 5-16 we have that

$$x = \xi + ut \quad (5-51)$$

and
$$dx = d\xi \quad \text{where } t = cst \quad (5-52)$$

Hence the second exponential may be reduced to terms of ξ and ut . If this substitution is performed and terms collected, there results

$$\mathbf{B}_{k'} = \frac{1}{nw} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S'(\xi) PWY(k) A_k \exp \left[\frac{2\pi j}{nw} (k - k') ut \right] \exp \left(- \frac{2\pi j k' \xi}{nw} \right) dt d\xi \quad (5-53)$$

Since each factor now contains only a single variable, t or ξ , we may write

$$\mathbf{B}_{k'} = \frac{PWY(k) A_k}{nw} \int_{-\infty}^{\infty} S'(\xi) \exp \left(- \frac{2\pi j k' \xi}{nw} \right) \int_{-\infty}^{\infty} \exp \left[\frac{2\pi j (k - k')}{nw} ut \right] dt d\xi \quad (5-54)$$

Consider the time integral. Since the integrand is periodic, one complete image being scanned in the interval V_p , the integral may have only two values depending upon the value of k' . (Remember k has already been chosen.) If $k' \neq k$, the integral is of the form

$$\int_0^{V_p} \exp (at) dt = 0 \quad (5-55)$$

or if $k' = k$, it is

$$\int_0^{V_p} \exp (0) dt = V_p \quad (5-56)$$

It follows, therefore, that $\mathbf{B}_{k'}$ is different from zero for only that value of k' which is equal to the k previously chosen, *i.e.*, there is a 1 to 1 correspondence between the original and reproduced field components. Thus we have for the k th component

$$\mathbf{B}_k = \frac{PWY(k) A_k V_p}{nw} \int_{-\infty}^{\infty} S'(\xi) \exp \left(- \frac{2\pi j k \xi}{nw} \right) d\xi \quad (5-57)$$

(5-57) may be simplified: Assume that the voltage-light proportionality constant, P , is the same at both ends of the system. Then

$$B_k = PA_k$$

And further, by similarity to equation (5-38), the integral is seen to be the reproducing aperture admittance

$$Y'(k) = \int_{-\infty}^{\infty} S'(\xi) \exp\left(-\frac{2\pi j k \xi}{nw}\right) d\xi \quad (5-58)$$

Hence
$$B_k' = \frac{V_p}{nw} W Y(k) Y'(k) B_k \quad (5-59)$$

We see, therefore, that each component in the original picture gives rise to a single component in the reproduced image. Also the equation shows that the complex amplitude of that component is modified by the channel response and the response of the two scanning apertures. B_k' is also proportional to V_p and hence to the speed of scanning, a result which is an entirely logical one.

5-12. The Two-dimensional Case

We have previously learned that the one-dimensional series of the type we have just considered fails to contain information for re-assembling the picture in the y direction. To overcome this deficiency, we must return to the bidimensional series and in so doing we shall find that the results stated in the last paragraph must be modified to a certain extent. Since the analysis of the bidimensional series is analogous to that which we have just completed, we need only write down the equations which correspond to the major steps in the simpler case. Thus, expanding the original image brightness in the double exponential series, we have

$$B(x,y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} B_{kl} \exp\left[2\pi j \left(\frac{kx}{w} + \frac{ly}{h}\right)\right] \quad (5-60)$$

As a further complication we must now admit a scanning aperture at the pickup end which has width as well as length. Let η be its ordinate measured from the aperture origin and its response be $S(\xi, \eta)$. This latter may be represented pictorially by a three-dimensional

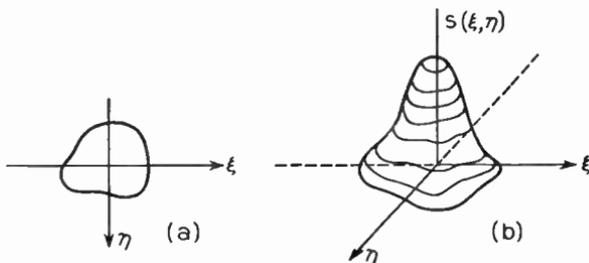


Fig. 5-17. (a) Co-ordinates of the bidimensional aperture. (b) The aperture response, $S(\xi, \eta)$.

diagram as in Fig. 5-17. By analogy to our previous work the admittance of this aperture will be

$$Y(k, l) = \int \int_{\text{ap}} S(\xi, \eta) \exp \left[2\pi j \left(\frac{k\xi}{w} + \frac{l\eta}{h} \right) \right] d\xi d\eta \quad (5-61)$$

and the apparent brightness seen by the transducer will be

$$B_1(x, y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} Y(k, l) B_{kl} \exp \left[2\pi j \left(\frac{kx}{w} + \frac{ly}{h} \right) \right] \quad (5-62)$$

And continuing the process we get as the analogue of equation (5-54)

$$B_{k'l'} = \frac{PWY(k, l)A_{kl}}{wh} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S'(\xi, \eta) \exp \left[-2\pi j \left(\frac{k'\xi}{w} + \frac{l'\eta}{h} \right) \right] \int_{-\infty}^{\infty} \exp \left\{ 2\pi j \left[\frac{(k - k')u}{w} + \frac{(l - l')v}{h} \right] t \right\} dt d\xi d\eta \quad (5-63)$$

where the primes once again indicate receiving-end quantities.

Consider the time integral which we have previously shown to be periodic with period V_p . Now in the one-dimensional case this integral was different from zero only for $k' = k$. In the present case, however, we see that the criterion is that

$$\frac{(k - k')u}{w} + \frac{(l - l')v}{h} = 0 \quad (5-64)$$

in which event the time integral becomes simply V_p . Equation (5-64) has implications not associated with (5-56), for it may be satisfied by several, not just one, pair of indices (k', l') for a chosen index pair (k, l) .

Let us consider this situation more carefully. If (5-64) is not satisfied,

$$B_{k'l'} = 0$$

or since $\frac{u}{w} = nf_p$ and $\frac{v}{h} = f_p$

we may write that if

$$(kn + l)f_p \neq (k'n + l')f_p \tag{5-65}$$

the reproduced component amplitude $B_{k'l'}$ is zero. The left-hand member of the inequality defines the frequency spectrum of the signal delivered to the reproducing transducer. Hence this is not a trivial result for it states that the amplitude of any frequency component (k', l') generated in reproduction is zero if that frequency is not equal to some frequency generated in pickup. In other words, no *new* frequencies are generated by the reproducing system even though it incorporates a scanning aperture of finite size. On the other hand, if

$$(k'n + l')f_p = (kn + l)f_p \tag{5-66}$$

or if $nk' + l' = nk + l$ (5-67)

the component amplitude is not zero. Ideally, (5-67) should be satisfied by one and only one index pair (k', l') such that

$$\left. \begin{aligned} k' &= k \\ l' &= l \end{aligned} \right\} \tag{5-68}$$

for then a 1 to 1 correspondence would exist between original and reproduced image components. Actually this ideal is not met as might be surmised from our discussion of the confusion components. The reproduced component specified by (5-68) we shall term the normal component. All other components whose k', l' values satisfy (5-67) we shall call the extraneous components. The amplitudes of the normal and extraneous components are not identical and we shall discuss them separately.

5-13. The Normal Component

When k' and l' satisfy (5-68) they define the normal component for which eq. (5-63) becomes

$$B_{kl'} = B_{k'l'} \Big]_{\substack{k'=k \\ l'=l}} = \frac{PWY(k,l)A_{kl}V_p Y'(k,l)}{wh} \tag{5-69}$$

where

$$Y'(k,l) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S'(\xi,\eta) \exp \left[-2\pi j \left(\frac{k\xi}{w} + \frac{l\eta}{h} \right) \right] d\xi d\eta \quad (5-70)$$

= admittance of reproducing aperture to the normal component.

Then, from (5-58), we may reduce the complex amplitude of the normal reproduced component to

$$B_{kl}' = \frac{V_p}{wh} W Y(k,l) Y'(k,l) B_{kl} \quad (5-71)$$

We may now compare B_{kl}' to B_{kl} . Equation (5-71) shows that the original image is reproduced at the receiver by the normal components, but this reproduced image will be distorted because each amplitude has been modified by the channel response, W , and the aperture admittances, $Y(k,l)$ and $Y'(k,l)$. Since the channel response has been assumed ideal, we may concentrate on the effects of the apertures. We see from the equation, then, that the apertures do introduce distortion which may be termed "simple loss of detail." This notation follows from the fact that typical apertures discriminate against the short wave length brightness components in the picture. To illustrate this Fig. 5-18 shows the square of the admittance of a circular aperture of uniform response plotted against wave length⁴ normalized with respect to the aperture diameter. It should be apparent that identical pickup and reproducing apertures would

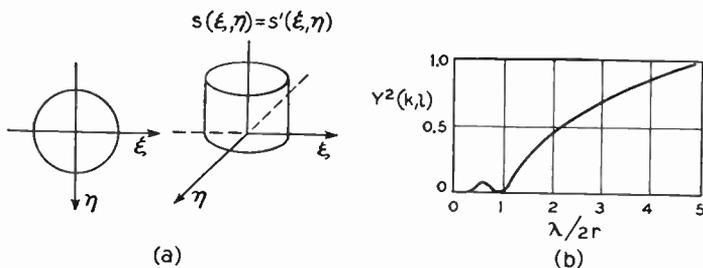


Fig. 5-18. The circular aperture of constant response exhibits a cutoff characteristic. (a) Aperture response. (b) Aperture admittance squared. (After Mertz and Gray.)

⁴ Wave length is used here in the sense of the normal distance between two adjacent maxima in a Fourier component, that is, as defined in eq. (5-15).

have the same admittance. Hence, the figure gives the product $Y(k,l)Y'(k,l)$ of eq. (5-71) for such an identical aperture system.

It will be observed that the very short wave length components for which $\lambda \leq 2r$ are virtually eliminated, which means loss of fine-grain detail in the reproduced image. This is a situation quite similar to that in the more familiar half-tone printing process where no detail smaller than a reproducing spot can be seen.

For those wave lengths for which $\lambda > 2r$ the loss of amplitude may be corrected for, provided that the admittance function is known for the apertures in use in any given picture transmission system. Since the aperture admittance terms appear in (5-71) in the same manner as the channel response, they may be equalized by a suitable electrical network. It is unfortunate that the use of such equalizers has to a large extent been neglected in television practice.

Equation (5-71) shows that both the one- and two-dimensional series give a reproduced brightness component amplitude proportional to V_p , the frame interval. This is an obvious result, for the longer the transducer light is seen in the reproduced image, the greater will be its apparent brightness.

5-14. Extraneous Components

We have previously predicted that values of k' and l' other than those defined by eq. (5-68) will also satisfy (5-64) or (5-67). Typical of such integer pairs are the following:

$$\begin{aligned} \text{For} \quad & k = 2 \quad \text{and} \quad l = 3 \\ & k' = 1, l' = 3 + n \quad \text{or} \quad k' = 3, l' = 3 - n \end{aligned}$$

These results simply confirm our previous discussion of the existence of the confusion components.

For our immediate purpose these confusion components resulting from a given set of values (k,l) may be defined in terms of a new variable μ such that

$$k' - k = \mu$$

Then these confusion components are given by

$$\text{and} \quad \left. \begin{aligned} l' &= l - \mu n \\ k' &= k + \mu \end{aligned} \right\} \text{Extraneous components} \quad (5-72)$$

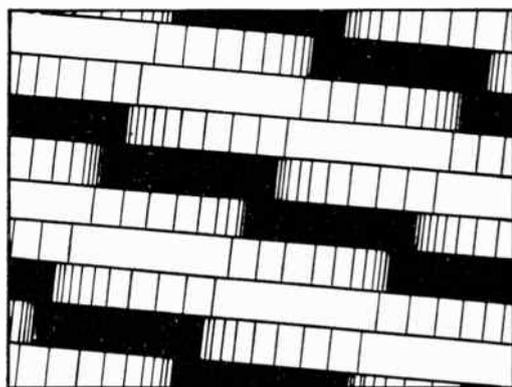
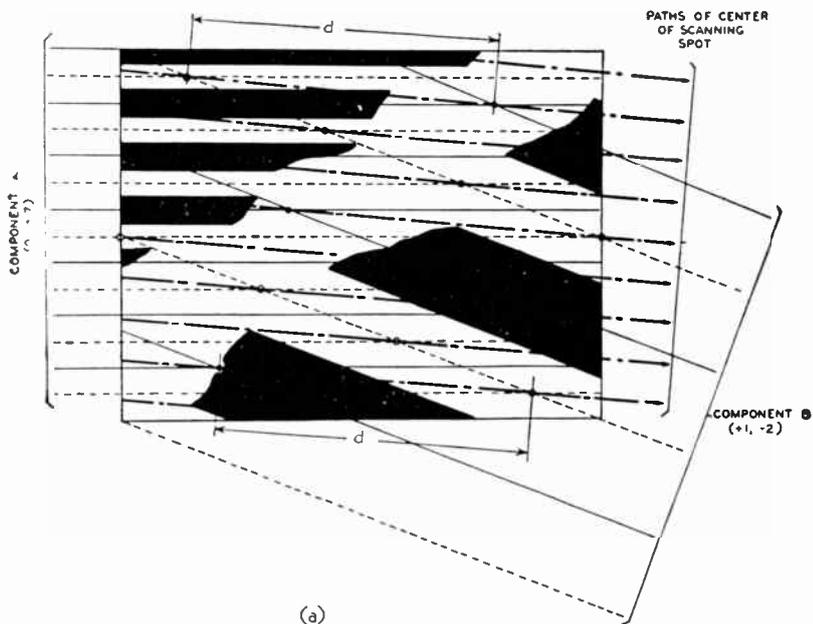


Fig. 5-19. Confusion in the image reproduced with a progressive 9-line scan. The Fourier components $(0, +7)$ and $(+1, -2)$ correspond to the same frequency in the incoming signal. (a) The two possible components are illustrated. (b) The reproduced picture contains both components even though only one of them may be transmitted. (Courtesy of *Proc. IRE.*)

and from eq. (5-69) each of these components will have a complex amplitude

$$B_{k'l'} = \frac{V_p}{wh} W Y(k,l) Y'(k,l) B_{kl} \quad (5-73)$$

where

$$Y'(k',l') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S'(\xi,\eta) \exp \left[-2\pi j \left(\frac{k'\xi}{w} + \frac{l'\eta}{h} \right) \right] d\xi d\eta \quad (5-74)$$

But since in general $Y'(k,l) \neq Y'(k',l')$

we find that $B_{k'l'} \neq B_{kl}$

i.e., the extraneous and normal components differ in complex amplitude.

Thus the single pickup component (k,l) of amplitude B_{kl} gives rise to a number of reproduced components (k',l') where $k' = k + \mu$ and $l' = l - \mu n$. Therefore, not only is the original picture reproduced in the normal component, $\mu = 0$, but also in a number of additional spurious images of different amplitude and phase. These latter images represent a form of noise in the picture and produce a loss of detail resulting from masking.

The situation in regard to the extraneous or confusion components may be stated in another manner. In the early part of the chapter we saw that two Fourier components in the original picture can give rise to the same frequency in the electrical signal. When the reproducing apparatus receives this signal it has no way of determining to which Fourier component it belongs; hence it reproduces both of them. This condition is illustrated for a progressive 9-line scanning system in Fig. 5-19. At *a* two components $(0,+7)$ and $(+1,-2)$ are shown, both of which give rise to the same frequency in the electrical signal. If, then, only the $(0,+7)$ component be transmitted, the reproducing system, being unable to determine this fact, will reproduce both components with the result shown at *b*. A somewhat similar condition exists for interlaced scanning.⁵

Another example of the effect of the extraneous components is shown in Fig. 5-20. The spurious pattern shown is the result of finite aperture effects only and should not be confused with ghost images which are caused by reception over multiple paths of different

⁵ See P. Mertz, *op. cit.*, Figs. 8 and 12.

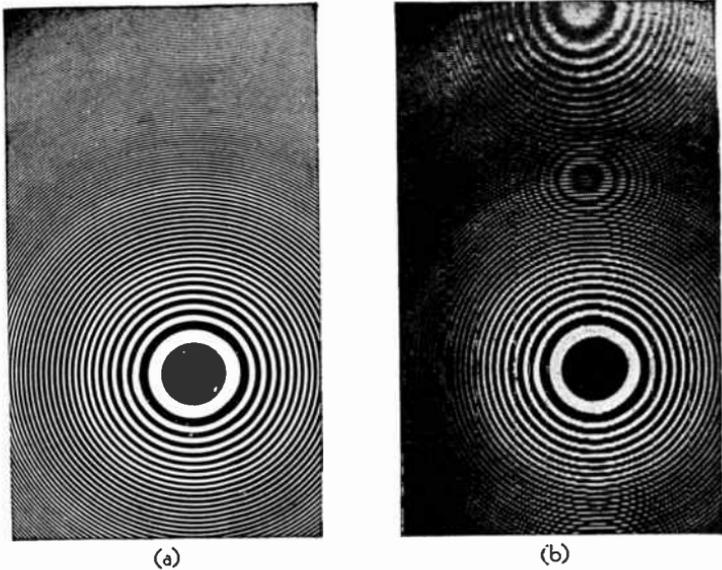


Fig. 5-20. Distortion in a reproduced image caused by a finite aperture. (a) The original subject. (b) The reproduced image. (Courtesy of *Bell System Technical Journal*.)

lengths. Fortunately, most of the masking components are usually of negligible amplitude.

Now it may be shown from a consideration of the cutoff characteristic of the aperture that the principal masking components, *i.e.*, those which produce the most objectionable effect in the final image, are those for which $|\mu| = 1$. This may also be seen on the basis that as l' approaches n , the field components become more horizontal and can be identified with the line structure itself. A good example of this effect is given when a flat white field is transmitted, *i.e.*, $k = 0$ and $l = 0$. Then the corresponding "most objectionable" components for which $|\mu| = 1$ are

$$k' = 1, \quad l' = \pm n$$

These components cause the final picture to be broken up into n lines, *i.e.*, the line structure is present. Since every scene has a d-c component corresponding to the average scene brightness, these lines will always be present to some degree, at least in the final image.

In connection with these "most objectionable" components, it is

interesting to note that it is theoretically possible to reduce their amplitudes to zero by proper design of the system apertures because, from (5-73), if

$$Y'(k', l) = 0$$

then

$$B_{k'l'} = 0$$

Hence the aperture should be so designed that

$$Y'(1, \pm n) = 0$$

This condition may be met by the use of a rectangular aperture of uniform response and of width equal to that of the spacing between adjacent scanning lines. Such an aperture may be at least approximated in a facsimile system where the size and shape of the aperture may be controlled by a physical mask. The very nature of the electron distribution in an electron beam, however, immediately precludes utilizing this principle and special techniques are required.

5-15. Scanning with Electron Beams

That this last condition cannot be attained easily with a scanning device which uses an electron beam may be seen from a consideration of the corresponding aperture response, $S(\xi, \eta)$. It has been experimentally determined that the electron density, ρ , in a beam of accelerated electrons has circular symmetry and decays toward the edges of the beam exponentially with the square of the distance from the beam center, *i.e.*, the distribution of ρ is Maxwellian.^{6,7} Then since $S(\xi, \eta) = K\rho(\xi, \eta)$, we see that some form of compromise must be made if we are to approximate the conditions for a flat field. Since the probability distribution shown in Fig. 5-21 implies

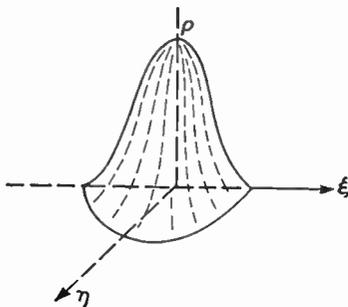


Fig. 5-21. The distribution of electron density in the scanning beam is Maxwellian.

⁶ V. K. Zworykin and G. A. Morton, *Television*. New York: John Wiley and Sons, Inc., 1940.

⁷ V. K. Zworykin, "Description of an Experimental Television System and the Kinescope." *Proc. IRE*, 21, 12 (December 1933).

an infinite cross-sectional radius, we shall first try to determine an alternate approximate distribution function for the brightness of a line which has been scanned by the beam, that will fit the physical limitation of a finite, in fact, small beam radius.

Since the beam is circular in cross section and has a nonuniform distribution of ρ , the electron density, the brightness of the scanned line normal to the direction of scan will vary. The distribution of ρ along y , $\rho(y)$, will now be investigated.

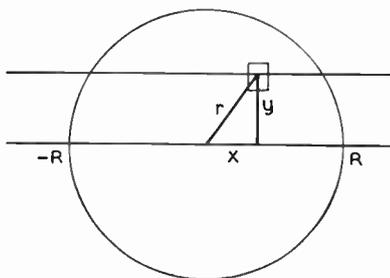


Fig. 5-22. Co-ordinate system for investigating the variation of brightness across a scanned line.

Let $\Sigma\rho(y)$ be the sum of the electron densities on a spot P as the beam passes *once* over P . It will be observed that this is the same as the sum of the ρ 's along the line $y = cst$, which is shown. Then assuming the exponential distribution stated above, we have

$$\rho = C\epsilon^{-Dr^2} \quad (5-75)$$

and

$$\Sigma\rho(y) = \int_{-R}^R \rho dx$$

—but the figure is symmetrical about the y axis and is of infinite radius. Therefore

$$\begin{aligned} \Sigma\rho(y) &= 2 \int_0^{\infty} C\epsilon^{-D(x^2+y^2)} dx \\ &= 2C\epsilon^{-Dy^2} \int_0^{\infty} \epsilon^{-Dx^2} dx \\ &= C\sqrt{\frac{\pi}{D}} \epsilon^{-Dy^2} \end{aligned} \quad (5-76)$$

Thus the distribution of light intensity—which is proportional to $\Sigma\rho(y)$ —across the scanned line is

$$I(y) = KC \sqrt{\frac{\pi}{D}} \epsilon^{-Dy^2} \tag{5-77}$$

with the maximum value

$$\hat{I}(y) = KC \sqrt{\frac{\pi}{D}} \tag{5-78}$$

occurring along the center where $y = 0$.

The nature of the exponential function is such that for $x < \pi$, $I(y) < 0.1\hat{I}(y)$ and the exponential distribution of infinite width may for a good engineering approximation be replaced by a \cos^2 function. The extent to which this is so may be seen from Fig. 5-23.

Thus if the line structure resulting from the “most objectionable” extraneous components is to be eliminated to give a “flat” field, the spot radius must be increased so that adjacent lines overlap. If this be done, the analysis by Mertz and Grey no longer obtains for it assumes no such overlap and other means of analysis must be used.⁸

It must be noted, however, that attempts to obtain a flat field by widening the spot must result in a loss of resolution. The one may be had but only at the expense of the other. That this is so may be seen from the following considerations: Overlap results in contamination of each line by information from the two adjacent lines; this decreases vertical resolution. Overlap also means larger spot radius, and since no detail smaller than the spot can be reproduced, horizontal resolution is lowered.

Clearly then, a compromise on spot size must be made in which flatness of field is balanced against resolution. This is a subjective matter and generally the consensus is that visible line structure and

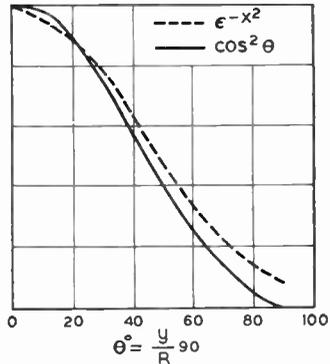


Fig. 5-23. Comparison of the exponential and cosine-squared functions.

⁸ See Chapter 10.

high resolution are more desirable than a flat field with its attendant image fuzziness. The compromise at the receiving end of the system is placed in the hands of the set owner in the focus control. Current practice in commercial television studios is to focus for best detail, that is, for minimum spot size.

CHAPTER 6

CAMERA TUBES

It has been convenient for the larger part of our discussion in the past chapters to think of the pickup end of the picture transmission system in terms of facsimile equipment. By so doing we confined the discussion to relatively familiar devices, such as the phototube as a transducer, a mechanical defining aperture, and a revolving screw mechanism to provide the motion along the prescribed scanning pattern. These elements were illustrated in Figs. 1-7 and 1-8. We have also learned, however, that modern television practice relies on all-electronic devices, the camera tubes, in which are combined the three basic functions of the pickup assembly, to wit: sampling, scanning, and the conversion of light to some form of corresponding electrical signal. In the present chapter we discuss the principal types of these camera or image-generating tubes.

6-1. Static-image Generators

Although in the strictest sense of the word the static-image generator is not a camera tube capable of converting light to electrical energy, it does deliver to the communication channel a picture signal identical in all ways to that from the camera tubes of the usual sense. Its sole difference is that it can generate signals corresponding to one and only one picture which is chosen during the tube manufacturing process. While such an image-signal generator is lacking in flexibility, it does provide a readily reproducible picture, an important advantage for test purposes. One such type of tube, the monoscope, is used largely by commercial television facilities to transmit "test pattern," an image which combines station identification and certain geometrical patterns which aid in the proper adjustment of television receivers. Typical test patterns are shown in Figs. 10-9 and 10-10.

The reader is referred to section 3-2 for a description of an early, developmental form of static-image generator used as part of the previously described Purdue Project. The modern commercial form

of this type of tube is embodied in the monoscope,¹ shown in Fig. 6-1. The electrodes K , G , A_1 , and A_2 make up the electron gun which is similar to that of the conventional cathode-ray tube. Particular care is taken in manufacture to build a gun structure capable of providing an electron beam of extremely small diameter. This beam is made to scan across the signal plate by the application of saw-tooth currents to the magnetic deflection yoke.

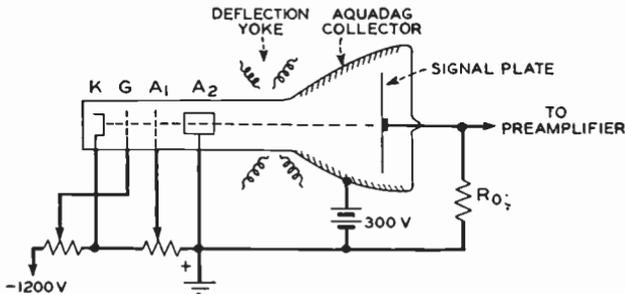


Fig. 6-1. The monoscope. Only a single image which is printed on the signal plate may be televised with this tube.

The signal plate, or pattern electrode, is an aluminum sheet, some $2\frac{1}{2}$ in. by $3\frac{1}{4}$ in. in size. Early in the manufacturing process the desired image, such as that of Fig. 10-9, is printed onto the aluminum in printer's ink. This plate and printed pattern are heated to reduce the ink to carbon. The result is a carbon image on an aluminum background.

The operation of the monoscope is based on the difference in secondary emission ratios for carbon and aluminum which are, for the voltages shown, 3 to 1 and 7 to 1, respectively. Inspection of the circuit diagram shows that the beam electrons hit the pattern electrode with a velocity of 1200 electron volts. These high-speed primary electrons produce secondary emission from the pattern electrode, the number of secondary electrons produced at any instant being proportional to the secondary emission ratio of that portion of the pattern under the scanning beam at that instant. These secondaries flow to the collector and constitute a current through the output resistance, R_o . Thus the output voltage developed across R_o follows

¹ C. E. Burnett, "The Monoscope." *RCA Review*, II, 4, 414 (April 1938).

the picture information on the signal plate as the latter is scanned by the electron beam.

Notice that a "black-negative" signal is generated. This follows from the fact that the carbon, or black, portions of the image produce a smaller output current than do the white, or aluminum, regions; thus a white-to-black transition in the scanned picture produces a negative-going voltage output.

If any portion of the pattern is in half-tone, care is taken to use a half-tone screen, giving smaller dots than the beam diameter. Under this condition the resolution of the monoscope is limited by the beam diameter rather than by the processing of the pattern plate in manufacture.

6-2. Some Elements of Photometry

Underlying the operation of the several types of camera tubes which we shall discuss is the principle of photoemission. Since in photoemission the energy required to remove free photoelectrons from the emitting surface is furnished by the incident light itself, we shall review some of the terms used to measure the energy contained in a given beam of light. By definition, light is radiant energy in the form of electromagnetic waves of such a wave length that they are visible to the human eye. Light sources, which of themselves generate light, *e.g.*, incandescent lamps, neon signs, the sun, or the fluorescent screen of a cathode-ray tube, are said to be self-luminous. One of the principal problems in the field of photometry, the science of measuring light intensity, has been to devise a system of units which relates the response of an observer to a standard unit system that is based on mass, length, time, and charge. Such a relationship is desirable because light is inevitably related to the human eye, which of course introduces a degree of arbitrariness into the photometric units. In the past the problem was handled by arbitrarily defining the candle, the unit of luminous intensity, as a fixed percentage of the luminous intensity of a bank of standard lamps, operating under specified conditions, maintained by the National Bureau of Standards.

In 1924 the International Commission on Illumination adopted a standardized luminosity curve, which has allowed a complete relationship to be established between the units of photometry and the M.K.S. system of units. Shown in Fig. 6-2, this luminosity curve of ordinates $\bar{v}(\lambda)$ is an arbitrarily standardized response curve of the

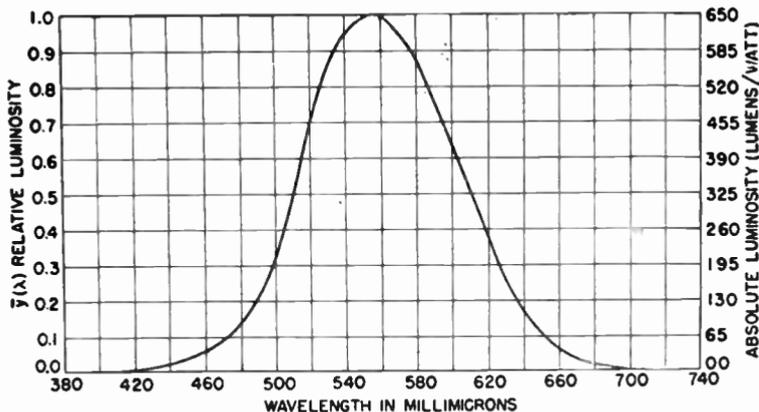


Fig. 6-2. The standard luminosity curve, which is used to relate photometric units to the M.K.S. system. It shows the frequency response of a standardized human observer. (Courtesy of *Electronics*.)

human eye. Its use in relating radiant quantities to their photometric equivalents is illustrated by the following equation:²

$$F = 650 \int_{\lambda_1}^{\lambda_2} P(\lambda) \bar{y}(\lambda) d\lambda \quad (6-1)$$

where P is the radiant flux in watts (a M.K.S. unit) and F is the luminous flux in lumens (a photometric unit). λ_1 and λ_2 denote the end points of the spectrum in which the radiant energy lies and are expressed in millimicrons. Once this basic relationship between the two systems of units has been established, all other photometric units may be derived from the lumen. In the work which follows we shall adopt the notation recommended by the Committee on Colorimetry of the Optical Society of America.

Thus far, then, we have defined the luminous flux, F , in lumens of a light source. A second property of such a source is its luminous intensity, I , which is given by

$$I = \frac{\Delta F}{\Delta \omega} \left[\frac{\text{lumens}}{\text{steradian}} = \text{candle} \right] \quad (6-2)$$

² D. W. Epstein, "Photometry in Television Engineering." *Electronics*, **21**, 7 (July 1948). See also, P. Moon, *The Scientific Basis of Illuminating Engineering*. New York: McGraw-Hill Book Company, Inc., 1936.

For a point source of light or a uniform spherical source, I is independent of the direction in which $\Delta\omega$ is taken. If, on the other hand, the source is extended, the flux intercepted in the solid angle $\Delta\omega$ will depend upon the direction α . We may illustrate this with Fig. 6-3a,

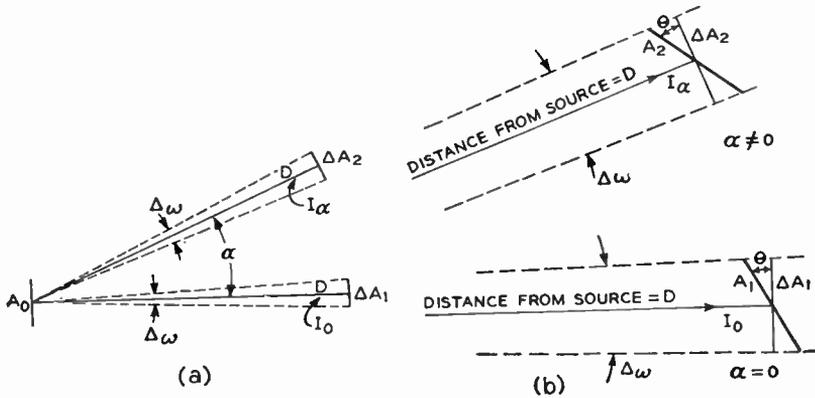


Fig. 6-3. (a) The luminous intensity, I , of an extended light source is a function of the angle α . (b) The illuminance, E , of a surface not normal to the direction from the light source depends on θ .

where the source has an area A_0 . Then, in general, two areas, ΔA_1 and ΔA_2 , which subtend equal solid angles, $\Delta\omega$, will not intercept equal values of luminous flux. Thus we must write

$$I_\alpha = \left(\frac{\Delta F}{\Delta\omega} \right)_\alpha \tag{6-3}$$

Frequently the directional characteristic of such an extended source follows a cosine variation

$$I_\alpha = I_0 \cos \alpha \tag{6-4}$$

where $I_0 =$ luminous intensity along the normal to A_0

which is a mathematical statement of Lambert's law.

If, now, a surface of area ΔA normal to the direction from the source intercepts ΔF lumens of flux from the light source, that surface has an illuminance³ E

$$E = \frac{\Delta F}{\Delta A} \left[\frac{\text{lumens}}{\text{square meter}} = \text{meter candle} \right] \tag{6-5}$$

³ Some texts replace the term illuminance by illumination.

Another frequently used unit of illuminance is the lumen per square foot or foot-candle which is equal to 10.76 lumens per square meter. Thus in Fig. 6-3*a* the normal surface ΔA_1 has an illuminance

$$E = \frac{\Delta F}{\Delta A_1} = \frac{I_0}{\Delta A_1} \Delta \omega$$

but from solid geometry

$$\Delta \omega = \frac{\Delta A_1}{D^2}$$

whence
$$E = \frac{I_0}{D_1^2} \quad \text{for} \quad \alpha = 0 \quad (6-6)$$

This is the well-known inverse square law of photometry, which is applicable to extended light sources, provided that D is 5 or more times greater than the largest dimension (the diagonal for a rectangle) of the source.

In a similar manner we calculate the illuminance of the normal area ΔA_2 at the angle α to be

$$E = \frac{I_\alpha}{D_2^2} = \frac{I_0 \cos \alpha}{D_2^2} \quad \text{for} \quad \alpha \neq 0 \quad (6-7)$$

Frequently we must calculate the illuminance of surfaces such as A_1 and A_2 of Fig. 6-3*b*, which are inclined at an angle θ with respect to the direction from the source. In such a case we have for A_1

$$E = \frac{\Delta F}{A_1} = \frac{I_0 \Delta \omega}{A_1}$$

But here
$$\Delta \omega = \frac{\Delta A_1}{D^2} = \frac{A_1 \cos \theta}{D^2} \quad (6-8)$$

whence
$$E = \frac{I_0 \cos \theta}{D^2} \quad \alpha = 0 \quad \theta \neq 0$$

Similarly, the illuminance of the surface A_2 is
$$(6-9)$$

$$E = \frac{I_0 \cos \theta \cos \alpha}{D^2} \quad \alpha \neq 0 \quad \theta \neq 0$$

Another property of a light source is its luminance,⁴ B . Thus, for example, the luminance of the surface A_0 in any direction α is defined as the ratio of intensity in that direction to the projected area of A_0 in that direction, or

⁴ Luminance is frequently termed brightness.

$$B_\alpha = \frac{I_\alpha}{A_0 \cos \alpha} \left[\frac{\text{candle}}{\text{square meter}} \right] \quad (6-10)$$

In the special case where the intensity from A_0 obeys Lambert's law, B is independent of α for

$$B_\alpha = \frac{I_0 \cos \alpha}{A_0 \cos \alpha} = \frac{I_0}{A_0} \quad (6-11)$$

In this case the human eye would observe the surface A_0 as being equally bright from all directions.

A fourth property of a light source is its luminous emittance, L , which is defined as the ratio of total luminous flux from the surface to its area, or

$$L = \frac{F_0}{A_0} \left[\frac{\text{lumens}}{\text{square meter}} \right] \quad (6-12)$$

Again if the intensity from A_0 obeys Lambert's law the luminous emittance may be related to the luminance, for

$$L = \frac{\int I_\alpha d\omega}{A_0} \quad (6-13)$$

But $I_\alpha = I_0 \cos \alpha$ and $d\omega = \frac{\Delta A}{D^2}$

and on a hemisphere, the element of area is

$$\Delta A = D^2 \sin \alpha d\alpha d\phi$$

whence

$$\begin{aligned} L &= \int_0^{2\pi} \int_0^{\pi/2} \frac{I_0 \cos \alpha D^2 \sin \alpha d\alpha d\phi}{A_0 D^2} = \frac{I_0}{A_0} \int_0^{2\pi} \int_0^{\pi/2} \cos \alpha \sin \alpha d\alpha d\phi \\ &= B 2\pi \int_0^{\pi/2} \cos \alpha \sin \alpha d\alpha = 2\pi B \left[\frac{\sin^2 \alpha}{2} \right]_0^{\pi/2} = \pi B \end{aligned} \quad (6-14)$$

Equation (6-14) serves as the basis for another unit of luminance, B , the meter-lambert which is equal to $1/\pi$ candles per square meter. This is the luminance of a perfectly diffusing surface which transmits or reflects one lumen per square meter. This meter-lambert unit is of particular use when one is dealing with perfectly diffusing reflecting surfaces. If these latter absorb no light, then the number of lumens intercepted and reflected remains constant and their luminance, B ,

in meter-lamberts is equal to their illuminance, E , in lumens per square meter, that is

$$B \quad [\text{meter-lamberts}] = E \quad \left[\frac{\text{lumens}}{\text{square meter}} \right] \quad (6-15)$$

If, on the other hand, these perfectly diffusing surfaces absorb a certain portion of the incident flux, the reflected flux is R , the reflection coefficient, times the incident flux and (6-15) becomes

$$B \quad [\text{meter-lamberts}] = RE \quad \left[\frac{\text{lumens}}{\text{square meter}} \right] \quad (6-16)$$

Two other common units of luminance are the millilambert and the candle per square foot. These are related to the meter candle as indicated below.

$$1 \text{ meter-lambert} = 0.1 \text{ millilambert} = 0.02957 \frac{\text{candle}}{\text{square foot}} \quad (6-17)$$

The several quantities which have been defined are collected in Table 6-1.

TABLE 6-1

| | | |
|--------------------|-----|---|
| Luminous flux | F | lumens |
| Luminous emittance | L | lumen/m ² |
| Luminous intensity | I | lumen/ω = candle |
| Luminance | B | lumen/ωm ² = candle/m ² |
| Illuminance | E | lumen/m ² = π meter-lambert |

In a typical television camera the light reflected from an illuminated object is focused by a lens system onto the photoemissive surface of the camera tube proper. Since, as we shall see, the current resulting from this light is proportional to the illuminance of the photoemissive surface, we need a relationship between the luminance of the televised object and the photocathode illuminance. Such a relationship will necessarily involve certain factors associated with the lens system. A number of forms of this relationship have been derived and the one recommended by De Vore and Iams⁵ is

$$E = \frac{\pi BT}{4f^2} \cos^4 \theta \quad (6-18)$$

⁵ H. B. De Vore, and H. Iams, "Some Factors Affecting the Choice of Lenses for Television Cameras," *Proc. IRE*, **28**, 8 (August 1940).

where

E = photocathode illuminance in lumens/ft²

B = televised object luminance in candles/ft²

θ = angle between the light ray striking the area under consideration and the system axis.

f = numerical aperture of the lens

$$= \frac{\text{lens focal length}}{\text{lens diameter}}.$$

T = transmission coefficient of the lens.

T is primarily a measure of the light losses which occur at the air-glass surfaces in the lens. Typical values for lens types used in television range from 0.4 to 0.7. We shall assume an average value of 0.64.

It may be realized that the application of (6-18) to any given televised image would be unnecessarily cumbersome because the surface luminance, B , varies from point to point on the image. Furthermore, because of the $\cos^4 \theta$ factor, the illuminance is less at the edges than at the center of the picture. Since the main region of interest is generally centered in the field of view, to simplify calculations we shall assume that the focused image is everywhere of illuminance corresponding to that at the center. Subject to this simplifying assumption, the $\cos^4 \theta$ becomes unity and for the assumed value of the transmission coefficient the equation becomes

$$E = \frac{B}{2f^2} \left[\frac{\text{lumens}}{\text{square foot}} \right] \quad (6-19)$$

The corresponding incident flux may be found by multiplying by the area of the photoemissive surface.

6-3. Photoemission

Whereas it is beyond the scope of the present work to include a lengthy description of photoemission, a few of its salient features will be set down for purposes of review. Four empirical laws relating to photoemission may be stated:⁶

When a photoemissive surface is illuminated by light of frequency ν , it is observed that

⁶ See, for example, J. D. Ryder, *Electronic Engineering Principles*. New York: Prentice-Hall, Inc., 1947.

- (1) The maximum velocity of the emitted electrons is proportional to ν .
- (2) The maximum velocity of the emitted electrons is independent of the illuminance.
- (3) The number of electrons emitted per unit time is proportional to the illuminance.
- (4) The time lag between emission and illumination is at least less than 10^{-9} sec.

As an explanation of the observed photoemission phenomena Einstein has proposed the photoelectric equation

$$h\nu = w + \frac{1}{2}mv^2 \quad (6-20)$$

where

- h = Planck's constant,
- w = work function of the photoemissive surface,
- m = electron mass,
- v = velocity of the emitted electrons.

The quantity $h\nu$ is the energy associated with a single quantum of the incident light.

The first rule follows directly from Einstein's equation. It also follows that if $h\nu$ is less than the work function of the material then no emission will take place. As a result we may define a threshold frequency for a given substance of work function w as

$$\nu_0 = \frac{w}{h} \quad (6-21)$$

Incident light of frequency less than ν_0 cannot release photoelectrons.

The second law also follows indirectly from (6-20), for the number of quanta which determines the number of emitted electrons is a measure of the energy supplied by the incident light. Hence, the number of electrons emitted per unit time is proportional to the intercepted power or the illuminance. It should also be true that for constant illuminance the number of electrons emitted per unit time would vary linearly with ν , which determines the energy per quantum. Actually photoemissive surfaces do not exhibit this property. Instead we find that typical surfaces have a color response which may take the form shown in Fig. 6-4. In television work we desire the camera tube to respond to visible light; hence, we must choose an

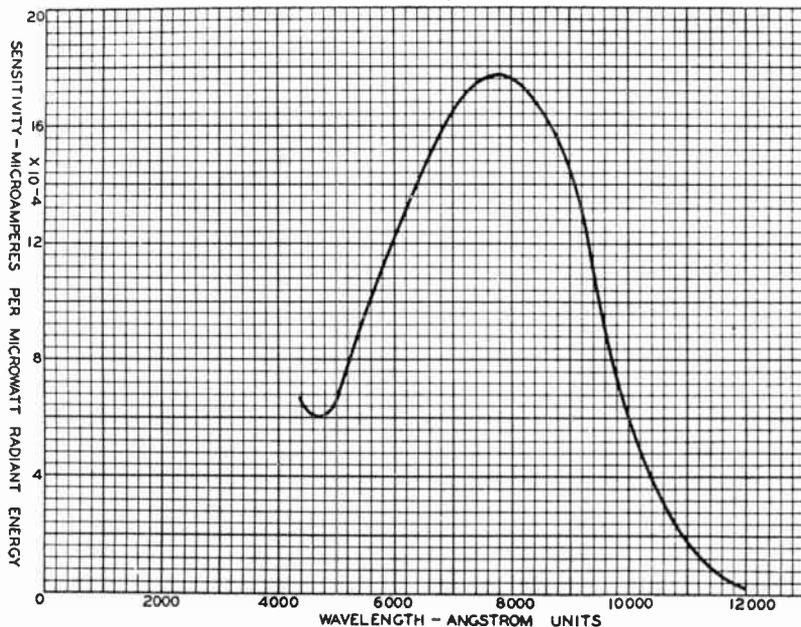


Fig. 6-4. Color response of a typical photoemissive surface. (Courtesy of Radio Corporation of America.)

emissive surface whose threshold frequency is below that of waves lying in the visible spectrum. Cesium has a low work function which is slightly under 2 ev.⁷ As a result the emitting surface in the camera tube will generally be composed of cesium in combination with one or more other substances. A common combination is cesiated silver.

⁷ It is common practice to express energy or electron velocity in terms of electron volts. The basic relationship between the quantities derives from the law of conservation of energy. If an electron of charge e coulombs falls through a potential difference of V volts, it has a kinetic energy in joules given by

$$\text{K.E.} = \frac{1}{2}mv^2 = V_e = 16 \times 10^{-20}V$$

m being in kilograms and v in meters per second. Thus an energy of V electron volts is that kinetic energy gained by an electron falling through V volts; one electron volt is equivalent to 16×10^{-20} joule or 16×10^{-13} erg.

Similarly V electron volts corresponds to an electron velocity in meters per second:

$$v = \sqrt{\frac{2V_e}{m}} = 5.95 \times 10^6 \sqrt{V} \quad \text{meters per second}$$

The mechanism behind this spectral response of the photoemissive surface is not understood but experience over a period of years has led to reliable working rules which provide responses satisfactory for television work. Zworykin and Morton⁸ have described the procedure for preparing the photosensitive surfaces of the iconoscope camera tube, in which a silver layer is activated with cesium. Extreme care is required during the activation schedule in order to produce a surface which at the same time has good emission and proper color response.

The importance of the color response of the camera tube cannot be overemphasized, for whatever is visible to the pickup system is displayed as light at the receiving-end cathode-ray tube. In particular, response in the infrared region causes difficulty because the camera tube sees, and the kinescope reproduces, information which the human eye would not identify in the original image. Also colors are not converted to proper tones on the gray scale. To illustrate this difficulty let us consider that a person's face is being televised by a camera with high infrared response. Upon translating the various color values to corresponding levels on the gray scale—for the television system, like black and white photography, is inherently color blind—we should expect the lips to appear almost black in the reproduced image. Actually an entirely different result is reproduced because the camera, being peaked to the reds, produces a high-level lip signal that reproduces as a white or light gray. Lips and face will appear in approximately the same tone in the final image, a result with a "washed-out" appearance. To correct for this the actor can use a deep red-brown lip covering which, though unsightly to the studio audience, produces the proper effect in the televised image. The use of color-compensating make-up and scenery paint is covered in the literature.^{9,10}

Heavy red response may also be corrected by the use of lighting which is rich in the blues and low in the red end of the spectrum. This particular approach to the problem of color response has been used in the studios at WRGB in Schenectady. Figure 6-5 shows the gray-scale equivalents of some of the principal colors used in studio

⁸ V. K. Zworykin and G. A. Morton, *Television*. New York: John Wiley and Sons, Inc., 1940.

⁹ W. C. Eddy, *Television, the Eyes of Tomorrow*. New York: Prentice-Hall Inc., 1945.

¹⁰ J. Dupuy, *Television Show Business*, General Electric Co., 1945.

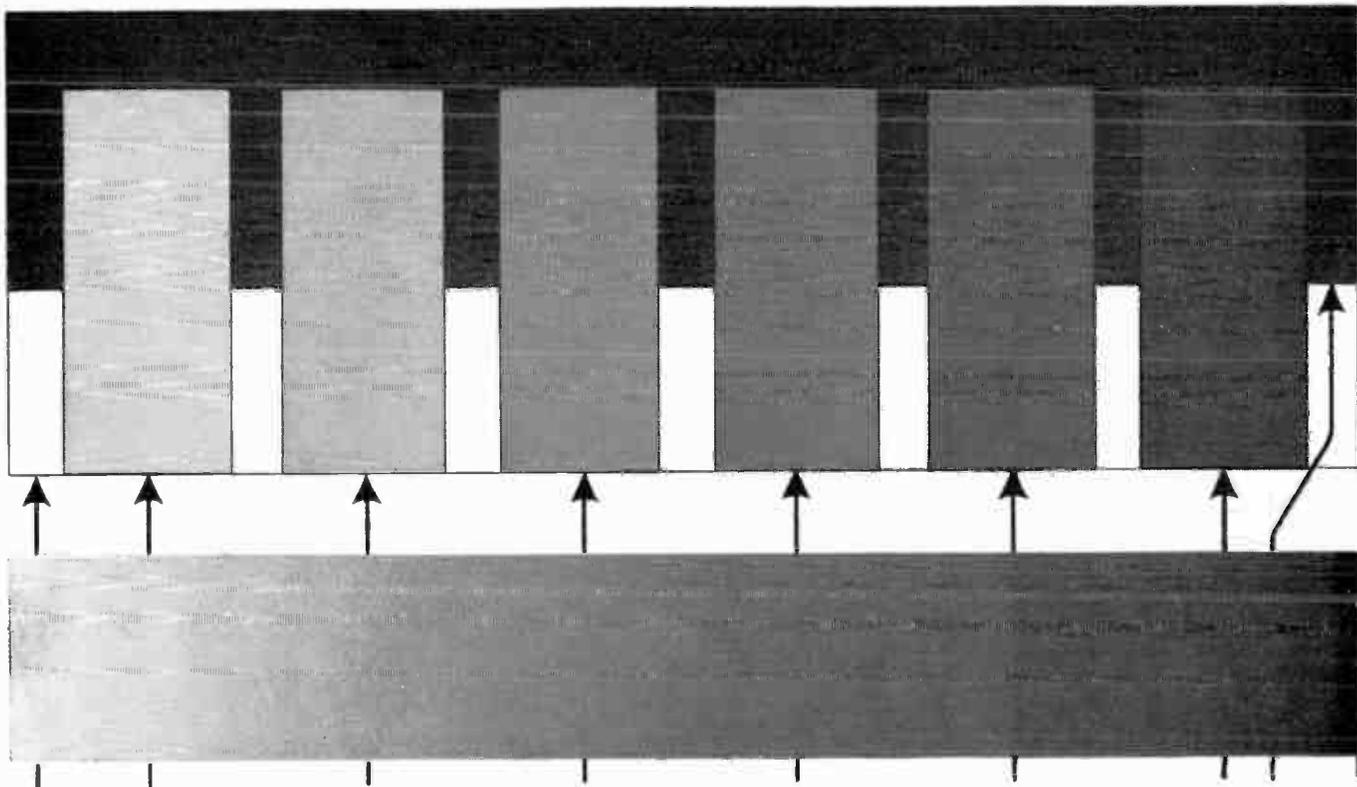


Fig. 6-5. Gray-scale equivalents of several colors as reproduced by an iconoscope working under mercury-vapor lights. The six blocks from left to right represent deep blue, light yellow, orange, violet, deep green, and bright red. A complete gray scale is shown at the bottom for comparison. Reproduced from Judy Dupuy, *Television Show Business*, General Electric Company, 1945.

setting design, with the studio lighting furnished by mercury-vapor water-cooled lamps. Since a large percentage of television pickup occurs at sites remote from the controlled light conditions of the studio, tube manufacturers continue to work on the improvement of the color response of camera tubes. It might appear that color filters in front on the camera tube could be used to correct the camera response. This expedient is not in general use because the low transmission coefficient of the filters lowers the photocathode illuminance.

Returning to the phototube we must realize that circuit components that will provide a closed electrical path must be provided if electrons emitted by virtue of the photoelectric effect are to constitute a useful output current. A basic circuit is given in Fig. 6-6, where

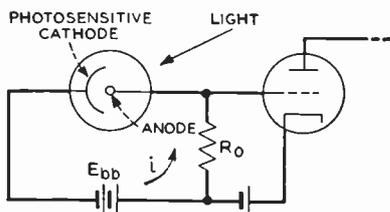


Fig. 6-6. The basic phototube circuit.

the photocathode and anode are placed in an evacuated envelope. If E_{bb} is sufficiently large so that voltage saturation obtains, the phototube current becomes directly proportional to the incident flux, F , and we have the relationship

$$i = sF = sAE \quad (6-22)$$

where s is the luminous sensitivity of the photocathode and A the projected area of the photocathode. Under the condition of voltage saturation, the phototube has an internal resistance of several megohms; hence, it behaves as a constant current source and the output voltage delivered to the preamplifier grid is

$$e_o = iR_o = sAER_o \quad (6-23)$$

We shall find that modern camera tubes differ considerably in form from the simple phototube just discussed, but the general mode of operation is similar and the output voltage has the form of (6-23).

6-4. Noise Considerations

Equation (6-22) shows that the average phototube current is proportional to the illuminance of the cathode, E . This question immediately arises: Is there any minimum value of illuminance below which the resultant phototube current would not be useful to produce an electrical signal? The answer lies in the value of the noise generated in the coupling network between the phototube and the preamplifier, and in the phototube and preamplifier tubes themselves. Since the maximum value of signal-to-noise ratio, S/N , in the entire system is set at this point, we must investigate it with considerable care. Since any noise generated before and in the preamplifier proper is amplified along with the signal, certainly S/N can never be improved over the value determined in the first coupling network. Whether or not the following stages in the amplifier chain deteriorate this initial value depends upon the available power gain of the first amplifier stage, a subject discussed in Chapter 14. Generally, if the voltage gain of the first stage is 3 or 4, the noise contribution of the succeeding stages may be neglected.¹¹

The situation here is typical of that encountered in the amplification of low-level signals. Unless the signal-to-noise ratio is sufficiently great, the signal is lost in the noise. In television systems this noise appears as tiny specks of "snow" over the entire reproduced image. Typical allowable values range from a peak picture signal to r.m.s. noise ratio of 10 to 1 for an acceptable picture; to 30 to 1, which results in an excellent picture. A 3 to 1 ratio is considered to be entirely unsatisfactory.¹² We see then that we need know not only a current-illuminance relationship but also one relating generated noise to the several noise-contributing circuit components. The chief causes of noise are shot, partition, and flicker effects in tubes as well as thermal agitation effects in resistors. We next consider these effects.

First consider the shot effect which is present in the phototube and, to a lesser degree, in the preamplifier. This phenomenon results from the random nature of thermionic or photoelectric emission. When any given electron in the emitting material receives energy from a

¹¹ V. K. Zworykin and G. A. Morton, *op. cit.*, p. 432.

¹² V. K. Zworykin, G. A. Morton, and L. E. Flory, "Theory and Performance of the Iconoscope." *Proc. IRE*, 25, 8 (August 1937).

photon or from thermal agitation in sufficient amount to overcome the work function of the material, it will not necessarily have a velocity component in the direction required for escape. Since the electron velocity distribution is random, the number of electrons emitted in unit time undergoes statistical variations. Thus the phototube current predicted by eq. (6-22) is an average value. The instantaneous value of current varies about this value in a purely random fashion. The total current, then, may be considered to consist of an average value, given by (6-22), upon which is superimposed the random component or noise. The variations in this noise component are completely irregular, so the noise is "white" in the sense that its energy is distributed uniformly throughout the frequency spectrum. Schottky¹³ has shown that the mean squared value of this shot noise current within the frequency band Δf is given by

$$\overline{i_s^2} = 2\varepsilon i \Delta f \quad \text{amperes squared} \quad (6-24)$$

where

ε = electronic charge, in coulombs,

i = average emission current, in amperes, and

Δf = noise bandwidth of the system, in cycles per second.

The quantity Δf is defined more precisely in Chapter 14, but at this point may be assumed to be equal to the steady-state half-power bandwidth of the amplifying system.

Equation (6-24) holds for tubes in which the current is emission-limited and hence applies to the phototube proper. For voltage- or space-charge-limited operation as in the preamplifier tube eq. (6-24) gives too large a value of mean squared noise current. A probable reason for this is that with the condition of space-charge limitation the dense cloud of electrons between the cathode and plate acts as a cushioning reservoir which irons out the random variations in the number of emitted electrons. Stated differently, the space charge serves as a virtual cathode whose emission is subject to less randomness than the actual cathode. Equation (6-24) must consequently be modified by some factor of magnitude less than unity when it is applied to the preamplifier tube. We shall presently see an alternative method of handling the effect of preamplifier shot noise.

Were the current output of the phototube utilized directly, we

¹³ W. Schottky, *Ann. Phys.*, **57**, 541 (1918); **68**, 157 (1922).

could immediately calculate the ratio of signal-to-noise components from the last two equations. Unfortunately, however, the output resistance R_o in Fig. 6-6 also contributes noise to the camera-tube output. This resistance, or Johnson, noise¹⁴ is a result of the thermal agitation of the electrons in any conductor. Because of the kinetic nature of matter, the electrons in the conductor are in a state of random motion, the motion being related to the average temperature of the conductor. Again, if an average current i flows through the conductor, the randomness of the electron motion causes an excess of negative charge at one end of the conductor at a certain instant. At some later instant the electrons will be bunched so that the excess of negative charge is at the opposite end of the conductor. Thus over an interval of time a noise voltage is developed across the ends of the conductor. Johnson and Nyquist have shown that the mean squared value of this noise voltage is related to the conductor resistance, R , and the conductor temperature, T , in the following manner:

$$\overline{e_j^2} = 4kT \int_{f_1}^{f_2} R df \quad \text{volts squared} \quad (6-25)$$

where

k = Boltzmann's constant = 1.374×10^{-23} joule per degree Kelvin

T = absolute temperature of the conductor, degrees Kelvin

$f_2 - f_1 = \Delta f$ = noise bandwidth, cps,

Under the special condition that R remains constant within the bandwidth Δf , (6-25) reduces to

$$\overline{e_j^2} = 4kTR\Delta f \quad (6-26)$$

In connection with this Johnson, or thermal agitation, noise it is important to notice that if the resistance is shunted by a capacitance, the simplified equation (6-26) may not obtain if Δf extends over a range of several megacycles as it generally does in television systems. The reason for this, of course, is that the high-frequency components of the white noise voltage appearing across the resistor tend to be shunted out by the capacitance. As an example of this, consider the network of Fig. 6-6 and assume a temperature of 300°K, a half-power bandwidth from 0 to 4 megacycles, and a resistance of 100,000 ohms. Application of (6-26) yields an r.m.s. noise voltage of 81.4 μ v.

¹⁴ J. B. Johnson, *Phys. Rev.*, **32**, 97 (July 1928).

If now we take into account the total capacitance in shunt across R_o , we have for the equivalent series impedance of R_o and C

$$Z = \frac{R_o}{1 + j\omega R_o C} \quad (6-27)$$

The resistive or real component of this series impedance which contributes noise is

$$R = \text{Re}(Z) = \frac{R_o}{1 + (\omega R_o C)^2} \quad (6-28)$$

To find the mean squared noise voltage we substitute (6-28) into (6-25) and there results

$$\overline{e_j^2} = \frac{4kT}{2\pi C} \text{arc tan } (\omega R_o C) \Big]_f^h \quad (6-29)$$

Evaluating this for a shunt capacitance of 20 micromicrofarads and the same values given above, we get

$$e_j = \sqrt{\overline{e_j^2}} = 13.75 \mu\text{v}$$

We must realize, however, that whereas shunt capacitance reduces the noise, it also attenuates the high-frequency signal components; Δf is no longer 4 mc. If the bandwidth is restored by compensation in the amplifier chain, both signal and noise increase. The student may verify that as Δf or R_o increases, the r.m.s. noise voltage becomes less dependent on the value of R_o .

The third source of noise in the circuit of Fig. 6-6 is the preamplifier tube. We have already seen that it contributes shot noise to the circuit, of magnitude given by eq. (6-24), modified by some factor less than unity which compensates for the cushioning effect of the space charge present in the tube. For purposes of calculation, however, it is convenient to replace the actual noisy tube by a noiseless tube whose grid circuit incorporates a resistance R_i . R_i has a value such that its Johnson noise causes the same mean square noise current $\overline{i_s^2} = 2\epsilon i \Delta f P$ in the noiseless tube plate circuit as is present in the actual tube. The value of R_i may be derived in the following manner: In a vacuum tube, the plate current is the product of the tube grid-plate transconductance and the grid voltage. The noise current $\overline{i_s^2}$ could be produced in a noiseless tube by a grid voltage

$$\overline{e_s^2} = \frac{\overline{i_s^2}}{g_m^2} \quad (6-30)$$

which is the mean squared voltage that, when applied to the grid of a noiseless tube, produces a mean squared noise current $\overline{i_s^2}$ in the plate circuit. But the Johnson noise of a constant resistance is given by (6-26). Then the value of R_t , the equivalent noise resistance, may be obtained by equating (6-26) and (6-30). There results

$$R_t = \frac{2\epsilon i P}{4kTg_m^2} \quad (6-31)$$

where P = a factor less than unity. Difficulty arises in evaluating P for a given tube. North¹⁵ has evaluated (6-31) for triodes and gives the approximate equation

$$R_t = \frac{2.5}{g_m} \quad \text{ohms} \quad \text{TRIODES} \quad (6-32)$$

where g_m is in mhos. The advantage of using the equivalent tube noise resistance, R_t , is that it expresses tube noise in the same form as thermal agitation noise, a form which is independent of the pre-amplifier gain and bandwidth. Equation (6-32) points out the desirability of having a high-transconductance preamplifier tube as far as low noise is concerned.

A second noise source in the preamplifier tube is the flicker effect, which is chiefly associated with oxide-coated cathodes. In such tubes the active area of emission moves about the cathode surface, introducing a further component of randomness in the plate current. Apparently the movement of the active emission area takes place slowly because the effect is significant only over a bandwidth extending up to 1 kc.¹⁶ Since this represents only a small fraction of the preamplifier bandwidth, the flicker effect may be assumed to contribute negligible noise in television circuits.

If a multigrid tube is used in the preamplifier, the partition effect must be reckoned with. In such tubes an electron must choose between the screen grid or plate as its destination, causing an additional degree of uncertainty in the instantaneous value of plate current. North¹⁷ has derived additional relationships which combine partition and shot noise into a single equivalent tube noise resistance.

¹⁵ D. O. North, "Fluctuations in Space-Charge-Limited Currents at Moderately High Frequencies," Parts II and III. *RCA Review*, IV, 4 (April 1940) and V, 2 (October 1940).

¹⁶ V. K. Zworykin and G. A. Morton, *op. cit.*, p. 431.

¹⁷ D. O. North, *op. cit.*

Thus for tetrodes and pentodes where G_1 is used as the control grid we have

$$R_t = \frac{I_b}{I_b + I_{c2}} \left(\frac{2.5}{g_m} + \frac{20I_{c2}}{g_m^2} \right) \quad \text{ohms} \quad \begin{cases} \text{TETRODES} \\ \text{PENTODES} \end{cases} \quad (6-33)$$

where

I_b = average plate current, in amperes,

I_{c2} = average screen grid current, in amperes, and

g_m = grid-plate transconductance, in mhos.

In general, an R_t of from 500 to 1000 ohms indicates a good tube from the standpoint of noise.

Summarizing these noise effects we see that the maximum signal-to-noise ratio is set by the ratio of signal voltage $e_o = |Z|sAE$ to the total Johnson noise of R_t , R_o , and the equivalent resistance of (6-24), all in series. Consequently given a minimum value of S/N , we may work backward to determine the lens constants for a given system. The student is referred to De Vore and Iams¹⁸ for a summary of these equations. A simplified calculation is outlined in the next section.

6-5. The Flying Spot Pickup Device¹⁹

We next consider the electronic flying spot scanner which is a comparatively recent version of the flying spot scanner of the Nipkow disk days. The system to be described is limited in application for it may only be used to pick up transparent program material of the type available on film and slides. For this sort of application it affords a relatively inexpensive form of camera-tube system and admirably suits our purposes for an illustration of typical signal-to-noise ratio calculations.

The basic principle of flying spot scanning has been described and diagrammed in Chapter 3. The present system which has found use in color television systems of the simultaneous color transmission type (cf. Chapter 18) and in the Multifax system of transmission differs from the mechanical scanners in that the flying spot of illumination is generated on the phosphor of a special cathode-ray tube and is caused to mark out the scanning raster by magnetic deflection. This moving light spot is focused onto the transparency through

¹⁸ H. B. De Vore and H. Iams, *op. cit.*

¹⁹ Vin Zeluff, "Television Flying Spot Generator." "Tubes at Work," in *Electronics*, 21, 6 (June 1948).

a lens system of large diameter, a numerical aperture of $f/1.9$ being typical. The light whose intensity has been modified by the transmission coefficient of the various portions of the slide or film is then spread over the photocathode of a phototube which generates a corresponding electrical signal. The physical arrangement of the apparatus is shown in Fig. 6-7.

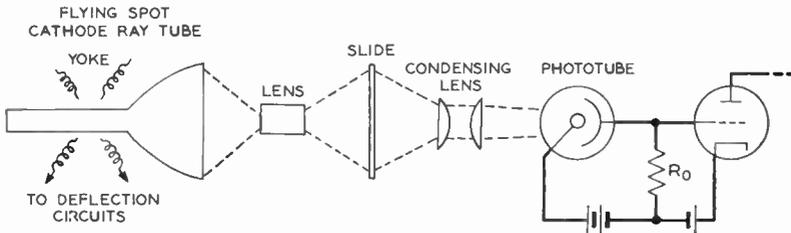


Fig. 6-7. (a) Components of an electronic flying spot pickup system. The flying spot originates on the face of the cathode-ray tube shown at the left.

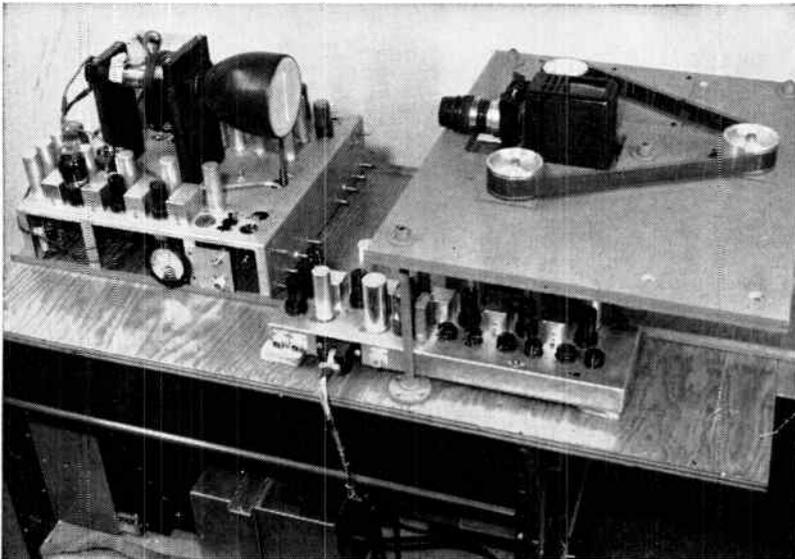


Fig. 6-7. (b) Photograph of the electronic flying spot scanner equipment. The phototube is located in the housing at the right. (Courtesy of National Broadcasting Company.)

Let us set up a hypothetical problem in which we assume negligible shunt capacitance across the output resistance, R_o . This assumption allows us to illustrate the type of calculations involved but simplifies the arithmetic to a considerable degree. The method may be extended to include shunt capacitance quite readily. For the specific problem we shall assume a type 917 vacuum phototube, which has a luminous sensitivity $s = 20 \mu\text{a/lumen}$. $R_o = 10,000$ ohms. The preamplifier is a 6AC7, operating as a conventional pentode for which $I_b = 10$ ma and $I_{c2} = 2.5$ ma. By (6-33) the equivalent preamplifier tube noise resistance is

$$R_t = 716 \text{ ohms}$$

which is an acceptable figure.

Let us further assume that we wish to calculate the photocathode illuminance that will give a combined signal to noise ratio of 30 to 1. To do this it will be convenient to convert all noise components into their equivalent mean squared current values. Thus for the tube noise

$$\overline{i_t^2} = \frac{\overline{e_t^2}}{R_o^2}$$

which is the equivalent mean squared current which, flowing through a noiseless resistance R_o , produces the same noise voltage as does R_t and equals

$$\frac{4kTR_t\Delta f}{R_o^2} \quad (6-34)$$

Similarly, for the Johnson noise developed in R_o we have

$$\overline{i_j^2} = \frac{4kT\Delta f}{R_o} \quad (6-35)$$

There is also a noise component as a result of shot effect in the 917 which is given by eq. (6-24). The total mean squared noise current is, then,

$$\overline{i_n^2} = \overline{i_t^2} + \overline{i_j^2} + \overline{i_s^2} = \frac{4kT\Delta f}{R_o} \left(\frac{R_t}{R_o} + 1 \right) + 2\epsilon i \Delta f \quad (6-36)$$

where i is the desired signal current in the highlights of the scanned image. Inspection of (6-36) shows that for the circuit components chosen the contribution of the preamplifier tube to the total noise is negligible because $R_t/R_o \ll 1$.

For the specified S/N ratio of 30 to 1 we may write

$$\frac{i_o}{\sqrt{i_n^2}} = 30 \quad (6-37)$$

Squaring and substituting from (6-36) we get a quadratic in i_o , the signal current, whose value on solution turns out to be $i_o = 0.07 \mu\text{a}$, for an assumed temperature of 300°K and a 4-mc bandwidth.

We may now work backward from eq. (6-22) to calculate the required illuminance of the photocathode which has a projected area of $\frac{1}{144}$ square feet. The required value of E is approximately 0.5 lumen/ft². It should be stressed once again that the results calculated in the example above are artificial in that the shunt capacitance across R_o has been neglected, an assumption which is generally not valid for a bandwidth of about 4 megacycles.

It might seem at first glance that the calculated value of E is low enough that no large demands are made on screen intensity of the flying spot cathode-ray tube. More careful consideration shows the converse to be true because the resolution of the whole system is determined by the size of the spot on the scanner tube; in fact, the spot size is the size of the pickup scanning aperture. Ideally, only one spot at a time on the surface of the tube is glowing; hence all the flux which produces the required photocathode illumination must come from a single spot whose diameter should be in the order of 0.001 in. This requires extreme screen intensity, a condition which has been met by the use of a screen phosphor of zinc oxide.²⁰ This type of screen has the additional advantage of providing a rapid decay of intensity, which falls to 5 per cent of its initial value in 1 μsec . A longer decay time would effectively increase the width of the scanning aperture in the direction along the scanning line. In the notation of the last chapter the ξ dimension would increase.

The flying spot scanner just discussed provides an excellent system of televising transparent subject matter. The basic equipment is relatively inexpensive in comparison to the more common types which employ camera tubes of the type to be described. The development in recent years of electron-multiplier-type phototubes which inherently have large outputs would indicate that the flying

²⁰ G. C. Sziklai, R. C. Ballard, and A. C. Schroeder, "An Experimental Simultaneous Color Television System. Part II—Pickup Equipment." *Proc. IRE*, **35**, 9 (September 1947).

spot scanner will enjoy increasing popularity for film pickup in black and white as well as in color television systems.

THE IMAGE DISSECTOR

We now turn our attention to the true camera tubes, those which are not limited to any particular type of subject matter and which combine the scanning, sampling, and transducing functions all in a single envelope. The major steps in the development of these true camera tubes have been listed in Chapter 3 and, in general, they have followed along two basic types, those that are storage devices or those that are nonstorage devices. Although tubes of both types were announced almost simultaneously in 1934, we shall consider first the nonstoring image dissector because the bulk of recent development has been along the storage line.

6-6. Method of Operation

Described originally in 1934 by Farnsworth,²¹ the image dissector was one of the first practical camera tubes. Its early form was that shown in Fig. 6-8. Physically the tube comprises an evacuated en-

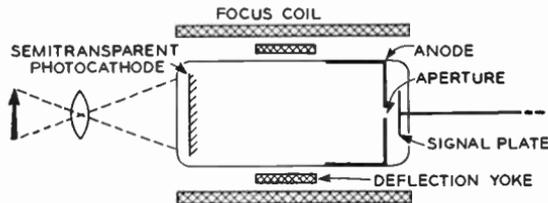


Fig. 6-8. Early form of the image dissector tube. The deflection signals cause an electron image which is produced at the photocathode to scan across the aperture in the anode.

velope which contains a photocathode, an anode in which is centered a small scanning hole (aperture), and a signal plate or collector. As shown in the diagram, the collector is directly behind the aperture and collects electrons which flow through the aperture. The student should notice the absence of the electron gun, which is common to all of the cathode-ray devices discussed so far. The focusing action is obtained magnetically rather than electrostatically with the help of a uniform axial magnetic field due to the focus coil shown.

²¹ P. T. Farnsworth, "Television by Electron Image Scanning." *J. Franklin Inst.*, **218**, 411 (October 1934).

Consider the operation of the dissector tube. A light image of the televised scene is projected onto the translucent photocathode by a lens system. Since the entire scene covers the photocathode, electron emission takes place over the whole cathode simultaneously, the number of electrons emitted at any instant from an elemental area being proportional to the illuminance of that elemental area. These emitted electrons are accelerated toward the anode by the anode voltage. In effect then, the light image on the photocathode has been transduced into a corresponding electron density image which moves toward the anode under the influence of the accelerating voltage and the focusing field. With proper adjustment of the focus coil current, this electron density image is in focus at the plane of the aperture. Obviously those electrons that fall in the area of the hole continue on to the signal plate or collector and constitute the output current.

Consider now that saw-tooth currents of the proper amplitude and frequency are applied to the horizontal and vertical deflection coils. The resulting magnetic field causes the entire electron density image to scan across the aperture. Thus the output current follows the illuminance across and down the picture area in a pattern determined by the scanning raster.

This scanning action may be stated in a different manner. Instead of having a moving aperture scan across a stationary electron density image, the image dissector causes the electron image to scan across a stationary aperture. In either case the result is the same: the aperture samples the electron image. Since the electron image conveys the luminance information of the original scene, the resulting output current is an $I(t)$ corresponding to a $B(x,y)$, t and (x,y) being interrelated by the scan pattern.

6-7. The Output Current

It is of extreme importance to note that at any instant the output current is proportional to the brightness of some area in the original scene as it appears on the photocathode, the area being equal to that of the aperture hole in the anode. Thus the resolution of the dissector is determined by the physical size of the aperture.

Let us call the aperture area one picture element. It follows from the previous discussion that *at any instant the output current is proportional to the instantaneous illuminance of that element on the photo-*

cathode which is focused on the aperture. For this reason, the image dissector is said to be of the instantaneous type of camera tube.

Let a be the element area. Then if w and h are, respectively, the photocathode width and height, and M is the figure of merit defined in Chapter 1, we have that

$$a = \frac{wh}{M} \quad (6-38)$$

Then for a photocathode of uniform luminous sensitivity, s , the output current will be

$$i = saE = \frac{sEwh}{M} \quad (6-39)$$

Equation (6-39) is basic for all camera tubes of the instantaneous type. Since M is about 100,000 for good resolution, the equation shows an inherent shortcoming of such devices: extremely high photocathode illuminance levels are required to produce a useful output, that is, one above the noise level.

It will be observed that the output current is limited by the photocathode area but this may not be increased without limitation because of the corresponding increase in tube size and lens cost. One distinct advantage of the form of dissector shown in Fig. 6-8 is that the translucent photocathode permits a short focal length lens to be used. This is economically desirable for if an f number—defined in (6-18)—is given, a lens of smaller diameter may be used. To counter-balance this, the light transmission coefficient of the translucent photocathode is low, which requires compensation in the form of higher scene lighting levels.

6-8. Electron Multiplier Image Dissector

A more recent form of the image dissector,²² which employs electron multiplication to boost the output current, is diagrammed in Fig. 6-9a. This form of the dissector differs from that just described in two ways. First, the semitransparent photocathode is replaced by an opaque one which requires illumination from the opposite end of the tube through a lens of long focal length. Secondly, the output current is increased by several stages of secondary emission multiplication. The basic form of the electron multiplier is shown in Fig.

²² R. W. Sanders, "Industrial Television." *Radio-Electronic Engineering Edition of Radio and Television News*, 12, No. 2, 3 (February 1949).

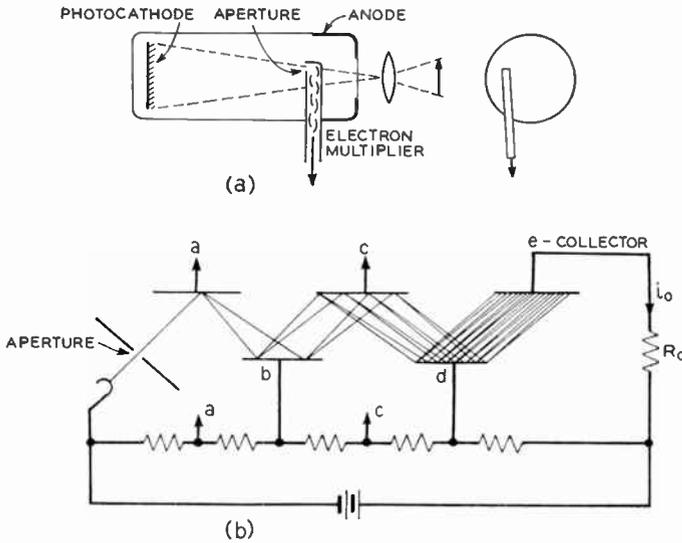


Fig. 6-9. (a) The output of an image dissector may be increased by means of an electron multiplier which is located off center in the tube in order to provide a clear optical path to the photocathode. (b) Operation of the electron multiplier. Current amplification occurs at each dynode, *a*, *b*, *c*, and *d*, because the secondary emission ratio is greater than unity.

6-9b. A number of accelerating electrodes or “dynodes,” *a* through *d* in the diagram, are arranged physically and electrically so that any electrons emitted from *a* go to *b*, any emitted from *b* go to *c*, and so on until the final output emission from *d* is collected by *e* and goes to make up the output current, i_o . Electron or current multiplication takes place because of a secondary emission ratio, r , greater than unity at each dynode. If i be the primary current passing from the photocathode through the aperture to the first dynode, *a*, the output current will be

$$i_o = i r_a r_b r_c r_d \tag{6-40}$$

or, in the general case, if there be n dynodes exclusive of the collector, each having the same secondary emission ratio, the output will be

$$i_o = i r^n \tag{6-41}$$

which indicates a current gain of r^n . In a typical image dissector with a voltage per dynode of 200 v, the secondary emission ratio is

in the order of 3 or 4. Consequently current gains in the order of one million are attainable.

We have seen from eq. (6-39) that the primary or aperture current is directly proportional to the aperture area, a . For an image dissector with an a of 40 square mils and a luminous sensitivity of 20 $\mu\text{a/lumen}$, we have for the primary current

$$\begin{aligned} i &= sEa \\ &= \frac{(2 \times 10^{-5})(4 \times 10^1)(4 \times 10^{-5})}{(1.44 \times 10^2)} = 2.22 \times 10^{-10} \text{ amp} \quad (6-42) \end{aligned}$$

for an illuminance of 40 foot-candles. Hence a typical output current is in the order of 200 $\mu\mu\text{a}$. This is a relatively high output, but the 40 square mil aperture by the same token would give poor resolution because of its comparatively large size. It still remains that in the image dissector high resolution at high output can be bought only at the expense of high illuminance levels. As a consequence its use at the present time is largely restricted to low-resolution systems having about 200 active scanning lines or to applications where extremely high light levels are possible. The low resolution system for telemetering is discussed in Chapter 8. In certain color television systems the image dissector is used for film pickup, because of its desirable color response. In this case illumination is furnished by a carbon arc, and high intensity is possible with no particular difficulty.

6-9. Multiplier Noise

It is almost axiomatic in electronics and communication work that the addition of a circuit element to improve one characteristic of a system deteriorates some other characteristic. For example, a transmission line may have its response equalized at the expense of gain, or an amplifier gain may be increased with a corresponding loss in bandwidth. The question might well be asked, then, as to what system characteristic has suffered because an electron multiplier has been added to increase the dissector output current.

In this particular it is fortunate that secondary emission takes place with little of the randomness which accompanies thermionic or photoelectric emission. The process is fairly definite: One incident electron literally dislodges r electrons from the secondary emission surface. For this reason the electron multiplier imparts equal gain to all components of the primary current; hence, the shot noise de-

veloped at the photocathode is multiplied by the same ratio as the signal current. As a result, eq. (6-24) applies to the electron multiplier image dissector with the exception that the noise and signal components must be interpreted in terms of the output values at the multiplier collector electrode.

Let m be a proportionality constant. Then, from (6-39) and (6-41), the output signal is

$$i_o = ma \quad (6-43)$$

Similarly, from (6-24) the r.m.s. output noise current resulting from shot effect is

$$i_{no} = \sqrt{i_s^{-2}} \propto \sqrt{ma} \quad (6-44)$$

It follows, therefore, that

$$\frac{S}{N} \propto \sqrt{a} \quad (6-45)$$

We may also reason that the resolution, being inversely proportional to aperture size, is proportional to $1/a$. We can see, then, that given an illuminance, (S/N) increases slower than the resolution decreases as the aperture size is increased. Nevertheless, the one may be traded for the other but the odds are not even. This confirms the results of the last section in a more precise fashion.

6-10. Magnetic Focusing

In the preliminary discussion of the image dissector it was noted that no electron gun is present in the tube and that focusing is obtained with the use of an axial magnetic field. Since this system of focusing is used in several of the camera tubes that will be discussed, we shall consider the action which takes place analytically. We assume that a uniform axial electric field, \mathcal{E} , is produced by the cathode-anode voltage. Further, a similar uniform magnetic field of intensity H is produced by a long coil wound around the circumference of the tube and extending over its entire length. An electron is released from the cathode with an initial velocity v_o inclined at an angle θ from the system axis. This initial velocity may be resolved into two components, one normal to the axis and of magnitude $v_o \sin \theta$, and the other parallel to the axis and of magnitude $v_o \cos \theta$. These two components may be considered independently.

Consider, first, the normal component which is unaffected by the electric field. Since it causes the electron to move normal to the H lines, the electron will be subjected to a force

$$f = \mu_o H \epsilon v_o \sin \theta \quad (6-46)$$

f is a constant force normal to H and to the velocity component $v_o \sin \theta$ and consequently causes the electron to rotate in a circular

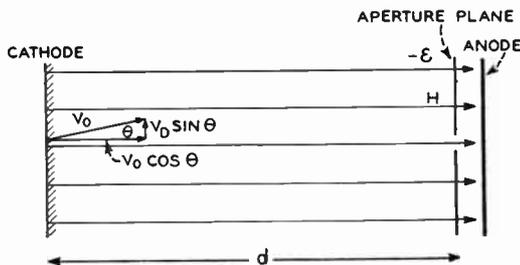


Fig. 6-10. Magnetic focusing with a long, uniform axial magnetic field.

path with peripheral speed $v_o \sin \theta$ in the plane normal to the magnetic field. Since the path is circular, f is balanced by a centrifugal force

$$f = \frac{m(v_o \sin \theta)^2}{\rho} \quad (6-47)$$

Equating (6-46) and (6-47) we get for the radius of the circular path

$$\rho = \frac{m v_o \sin \theta}{\mu_o H \epsilon} \quad (6-48)$$

Since the electron moves with constant speed around the circle of radius ρ , the time, τ , required for one trip around the circle is

$$\tau = \frac{2\pi\rho}{v_o \sin \theta} = \frac{2\pi m}{\mu_o H \epsilon} \quad (6-49)$$

This last equation has interesting implications because τ is independent of θ , v_o , and ρ . Thus all electrons emitted into the fields with components of velocity normal to the fields follow circular paths and they all complete one revolution in the same length of time, τ .

Simultaneously each electron is being attracted to the anode by the electric field which produces a constant axial acceleration

$$a = \frac{\varepsilon \varepsilon}{m} \quad (6-50)$$

and the resulting axial velocity component of the electron is

$$v_x = v_o \cos \theta + \frac{\varepsilon \varepsilon}{m} t \quad (6-51)$$

We see, then, that the electron has two components of motion, one circular in a plane normal to the fields, and the other linear and parallel to the fields. Therefore the actual electron path as it moves from cathode to anode is a helix, and in time τ the electron will move an axial distance

$$l = \int_0^\tau v_x dt = v_o \cos \theta \tau + \frac{\varepsilon \varepsilon}{m} \frac{\tau^2}{2} \quad (6-52)$$

and substituting from (6-51) for τ we get

$$l = \frac{2\pi m}{\mu_o H \varepsilon} \left(v_o \cos \theta + \frac{\pi \varepsilon}{\mu_o H} \right) \quad (6-53)$$

If now H is adjusted so that l is equal to d , the intercathode-aperture spacing, all those electrons leaving an area a on the cathode, and for which θ is small enough that $\cos \theta \approx 1$, will arrive in an equal area a in the plane of the aperture. Since this statement is true for any area a on the surface of the photocathode, it follows that an electron density image produced at the cathode reproduces itself in the plane of the aperture. Where l is made equal to d , the reproduced image is erect and of the same size as the density image at the cathode.

In more advanced treatments²³ of the problem it is shown that other ratios of l to d may produce amplification, and that improper adjustment of H will produce a rotation of the image at the aperture plane. It is sufficient for our purposes to note that it is possible to produce an erect image of magnification one.

In the analysis of magnetic focusing it is assumed that the magnetic field intensity H is uniform through the intercathode-anode space. The problem of producing such a uniform field with a coil is not without difficulties. One method that has been used employs a coil with a tapered winding, that is, the long focusing coil of Fig. 6-8 and 6-9 is wound with a varying winding density over its length.

²³ I. G. Maloff and D. W. Epstein, *Electron Optics in Television*. New York: McGraw-Hill Book Company, Inc., 1938.

Proper control of the winding density will give the desired field. Small irregularities in the field which may be seen as irregularities in the raster of the televised picture may frequently be removed by placing small chips of high-permeability metal between windings or on top of them. This procedure alters the permeability of the magnetic circuit in the region where the chip is inserted and though tedious may lead to excellent results.

An alternate approach for producing a magnetic field which is uniform over the length of the dissector tube employs a uniformly wound focus coil surrounded by a tubular magnetic shield wound of iron wire. This solution is used in connection with the image orthicon tube which, in common with the dissector, uses long-coil, magnetic focusing.²⁴

THE ICONOSCOPE^{25,26,27}

We have seen that the chief limitation on the use of an image dissector as the camera tube in high resolution television systems is that it requires extremely high levels of picture brightness and cathode illuminance. This comes about because the dissector is inherently an instantaneous device; at any instant the output is proportional to the instantaneous illuminance of the element being scanned. In 1934 Dr. V. K. Zworykin announced the iconoscope tube which represented the culmination of some ten or more years' work on the development of a camera tube which could utilize a new principle of operation. This principle required a single element of the photo-emissive surface to store up charge for the entire interval between successive scans. This revolutionary device at once was a camera tube of high output which made the electronic televising of studio scenes a practical matter. We shall consider in order a general description of the iconoscope, the storage principle, and operation of the device. Subsequent sections will discuss some of the equipment normally associated with the iconoscope pickup chain and some of its principal characteristics.

²⁴ R. D. Kell and G. C. Sziklai, "Image Orthicon Camera." *R.C.A. Review*, VII, 1 (March 1946).

²⁵ V. K. Zworykin, "The Iconoscope—A Modern Version of the Electric Eye," *Proc. IRE*, 22, 1 (January 1934).

²⁶ V. K. Zworykin, "Television," *J. Franklin Inst.*, 217, 1 (January 1933).

²⁷ V. K. Zworykin and G. A. Morton, *op. cit.*

Shown in Fig. 6-11, the iconoscope (*icon*—image, *scope*—observation) comprises an electron gun and a photosensitive mosaic mounted in an evacuated envelope. Of importance, too, is an aquadag collector electrode on the inner surface of the envelope, which is normally operated at some 1000 v positive with respect to the electron gun cathode. The mosaic that is the heart of the device consists of a thin uniform sheet of mica backed by a conducting metallic coating referred to as the signal plate. The front side of the mosaic consists of a very large number of small-sized insulated islands of cesiated

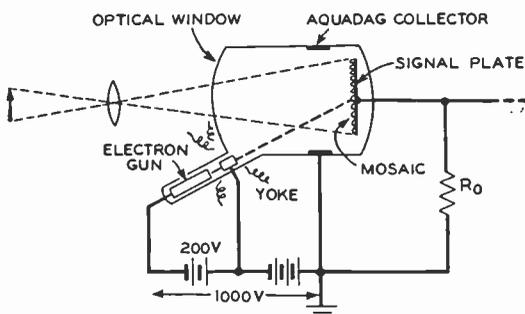


Fig. 6-11. Basic circuit of the iconoscope.

silver which are photoemissive. Each of these islands is capacitively coupled to the signal plate by a condenser consisting of the island, the signal plate, and the mica between them. Some idea of the minuteness of the photosensitive islands may be derived from the idea that the gun scanning beam of diameter between 0.01 and 0.02 in. covers a large number of the islands. In the commercial types of iconoscopes, such as the 1849 and 1850, the electron gun is inclined at an angle relative to the mosaic, a convenient mechanical arrangement to provide an unobstructed optical path between the window and mosaic. Magnetic deflection of the beam is used and means must be provided for correcting for the eccentric gun position.²⁸

The type 1847 experimental iconoscope overcomes the necessity of the out-of-line gun by using a semitransparent or translucent mosaic-signal plate system. The resulting need for increased illumination is not serious in experimental systems for which the 1847 is intended. In fact, a single 200-w lamp, silvered on the inside in conjunction

²⁸ See section 6-17 on keystone correction.

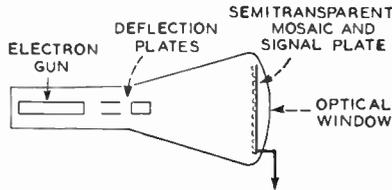


Fig. 6-12. The type 1847 iconoscope which employs a semitransparent mosaic and electrostatic deflection.

with an $f/2.3$ pickup lens, provides sufficient illumination for still scenes.²⁹

In brief, the operation of the iconoscope is as follows: Under the influence of the incoming light image the mosaic islands emit photoelectrons. By virtue of the island-to-signal plate capacitance charge is stored up in proportion to the number of electrons emitted, with the result that the illumination image is stored in the form of charge on the multitude of subelementary condensers. The beam of electrons emitted from the electron gun is caused to scan across the photosensitive face of the mosaic by the deflection yoke and suitable deflection currents. This *stored* information is then released in proper sequence to the output circuit by the electron beam which effectively restores the lost charge to each of the condensers in order. Thus the output current from each element is, theoretically at least, proportional to the illumination of that element *for the entire interval between two successive scans of the element*. The resulting increase in output over that of a corresponding instantaneous type of pickup tube is theoretically in the order of the system figure of merit, M . We show this in the next section. A more careful analysis of the iconoscope operation will follow and shows where the results in practice do not give the theoretical gain over the other type of system.

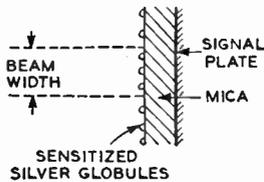


Fig. 6-13. Enlarged section of the iconoscope mosaic.

6-11. The Storage Principle

We have seen that the mosaic consists of a multiplicity of photoemissive globules of activated silver, each of which has capacitance to the signal plate. An enlarged diagram of a portion of the mosaic is shown in

²⁹ 1847 Iconoscope Data Sheet, RCA Manufacturing Co., Inc., 1940.

Fig. 6-13. As long as light falls on the photoemissive surface, electrons leave that surface, causing a charge to be built up in the subelementary condensers. If, now, each condenser be discharged at regular intervals, τ , the output current at discharge will be proportional to the total photoemission between successive discharges. Let

a = element area = area of scanning beam

Let i_p be the instantaneous value of photoemission current from an element. [This may be calculated from (6-39).] Then the charge stored by the element between successive scans is

$$q = i_p(\tau - \tau_e) \quad (6-54)$$

where τ_e is the time for the beam to scan over and discharge one element. Since there are M such elements in the mosaic, τ_e is the ratio of τ to M . Thus the output current delivered by a single element when it is scanned is

$$i_o = \frac{q}{\tau_e} = i_p(M - 1) \approx i_p M \quad (6-55)$$

M generally exceeds at least 10^4 ; hence, the term unity is negligible and we see that, in theory, the storage device gives an output M times greater than a nonstorage tube, other things being equal. Practically, (6-54) must be modified because the stored image is contaminated by emission from other elements and because the potentials present at the mosaic prevent saturated photoemission to take place. These effects may be handled by an efficiency factor, η , to which Zworykin assigns the approximate value of 5 per cent. Even with this low efficiency, for an M of 100,000, the iconoscope will have 5000 times the output of the image dissector. Translate this gain into terms of required scene illuminance and the revolutionary effect of the iconoscope on the development of practical television is at once apparent.

If the expression for i_p , eq. (6-39), be substituted into eq. (6-55), it will be seen that at least in theory the iconoscope output is proportional to the mosaic area. This result may be reached intuitively because for a given figure of merit, the element size, a , increases with the mosaic area. The larger a is, the larger the photoemissive surface and output current. In the commercial iconoscope, a compro-

mise mosaic size of 9×12 cm is used. Zworykin has given the following constants as typical of a commercial iconoscope.

$$wh \approx 100 \text{ cm}^2$$

$$s = 15 \text{ } \mu\text{a/lumen}$$

Then, assuming an efficiency of 5 per cent and a 10,000-ohm coupling resistance, we may calculate the output voltage of the iconoscope to be

$$e_o = 0.85E \quad (6-56)$$

where E is in lumens/cm².

The measured response of some typical iconoscopes shows that at low levels of illumination the output-voltage illuminance characteristic is linear as predicted by our equations. The slope of this initial rise in the measured characteristics is 1 v/lumen/cm² rather than the 0.75 value calculated above. It may also be seen from Fig. 6-14

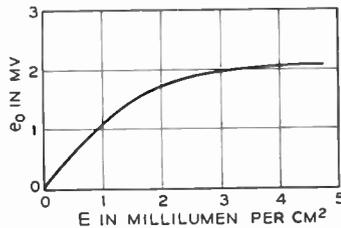


Fig. 6-14. Response of a typical iconoscope. $s = 15$ microamperes/lumen; $wh = 100$ cm²; $R_o = 10$ kilohms. (From V. K. Zworykin and G. A. Morton, *Television*. New York: John Wiley & Sons, Inc., 1940.)

that the linear relationship of the measured curves breaks down at higher level of illuminance. In fact, over a wide range of illuminance the e_o versus E characteristic is logarithmic, *i.e.*, the relationship may be expressed as $e_o = K \log E$.

6-12. Electron Bombardment of an Insulated Surface

It is of passing interest to note that three types of electron emission take place in the iconoscope: (a) thermionic—at the electron gun cathode, (b) photo—at the mosaic, and (c) secondary—at the mosaic. This secondary emission occurs at the surface of an insulated target, the mosaic, when it is bombarded by the high-speed beam electrons

which have sufficient energy to release the secondaries. We next consider this mechanism in detail. Maloff³⁰ has described some experiments which give an excellent insight into the mechanism of bombardment. The circuit used is shown in Fig. 6-15. Under the

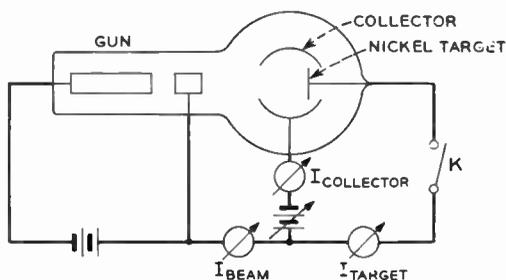


Fig. 6-15. Tube for investigating electron bombardment. (Courtesy of *Electronics*.)

influence of the accelerating voltage the beam of electrons emitted from the electron gun hits the nickel target, releasing secondary electrons. The number of secondary electrons emitted by the target is determined by the beam current (number of primary electrons) and the target secondary-emission ratio. It is not necessarily true that all the secondary electrons will go to the collector to form I_c , the target-collector voltage being a determining factor. The emitted secondary electrons not collected by the collector electrode must fall back onto the target. As a result the ratio of I_c to I_b is not identical to the secondary emission ratio, and depends upon the collector voltage.

The object of the experiment is to measure the ratio I_c/I_b as the collector-target voltage is varied; the results are plotted in Fig. 6-16. As might be expected, as the collector becomes more negative the collector current decreases, the excess emitted secondaries returning to the target.

We note that with a retarding voltage of 3 v the collector and beam currents are equal; hence the target current is zero, and the switch, K , may be opened without disturbing the circuit. Since nothing is changed electrically by opening the switch, we note that the target

³⁰ I. G. Maloff, "Electron Bombardment in Television Tubes." *Electronics* 17, 1 (January 1944).

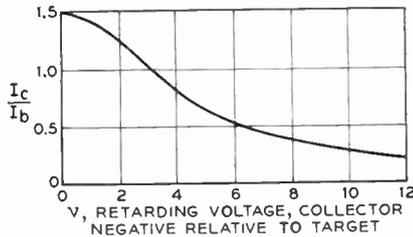


Fig. 6-16. The current-voltage characteristic of a nickel target.
(Courtesy of *Electronics*.)

is at beam potential, but the collector is negative by 3 v with respect to the target. This is tantamount to saying that an *insulated* nickel target under bombardment by a 500-v electron beam will become 3 v positive with respect to the collector electrode. Any change in target-collector potential will cause a corresponding change in collector current until this equilibrium value of 3 v is reached.

These results may be generalized for any insulated target having a secondary emission ratio greater than unity and under bombardment by a high-velocity electron beam: it will assume a potential of a few volts positive with respect to the electrode which collects the emitted secondary electrons. This positive voltage is known as the equilibrium potential of the bombarded surface. The velocity of the primary beam electrons is determined by the collector voltage plus the target voltage.

We may now carry over these ideas to the iconoscope, where the target is the front or photoemissive side of the mosaic. Strictly speaking, since each of the photoemissive islands is insulated completely from all other islands, we have an array of insulated targets. We shall still consider an element to be composed of the sum of all such islands under the electron beam at any given instant.

Let us first consider the action of the bombarding beam when the mosaic is in darkness, *i.e.*, when no photoemission occurs. From the results stated above we know that, directly as an element is scanned, its potential reaches the equilibrium value relative to the collector. This process applies repeatedly as the beam passes from element to element. The secondary emission ratio is greater than unity and yet, when each element is driven to the equilibrium potential, the collector and beam currents are equal. This means that we must account for the large number of secondary electrons released by the

beam, which represents the difference between the secondary emission ratio and collector-current-beam-current ratio. This idea may be made clearer by a crude example. Assume a secondary emission ratio of 5 to 1. This means that one beam electron releases five secondary electrons from the mosaic element. Since the element is insulated it reaches equilibrium potential which, in turn, is such that only one of the secondaries goes to the collector. Thus four electrons are left over; they are the difference between the five emitted electrons determined by the secondary emission ratio and the one electron comprising the collector current.

Since these "excess" electrons cannot go to the collector, they must fall back onto the mosaic itself and will naturally tend to return to the more positive elements—to those elements which have already been scanned. Notice that when the whole mosaic is viewed in darkness as it is scanned, we see the following process take place: As a given element is scanned by the beam it is driven to an equilibrium potential which is positive with respect to the collector. Then, as the beam moves on, the same element picks up some of the excess secondaries emitted from succeeding elements until its potential is between 0.5 to 1.5 v *negative* with respect to the collector. At this voltage the excess secondaries from other elements are repelled. To summarize this process: A given element of an unilluminated mosaic reaches a low value of, say, -1.5 v. Then, while traversed by the beam, it is driven to an equilibrium value of roughly $+3$ v. The *difference* between the two values is the operating range of the mosaic potential.

If we now add to this mechanism the effect of photoemission when the mosaic is illuminated by the light image of the televised object, we will find that a specific element gives off photoelectrons in proportion to its illuminance, and the element voltage increases positively from its negative value up to some value, say v , when the scanning beam arrives. Then the voltage will almost instantaneously jump from v to the equilibrium value. The resulting change in charge on the element scanned causes a flow of charging current through the capacitance to the signal plate and produces a corresponding voltage across R_s in the external circuit of Fig. 6-11.

Since a dark mosaic element produces an output voltage proportional to the difference between, say, -1.5 and $+3$ v, an illuminated element gives an output proportional to the difference between

-1.5 and $(+3 - v)$. v is determined by the photoemission. As a result the output always depends on $(+3 - v)$. For high mosaic illuminance v is high (near $+3 v$), and the output is low. Conversely, for low illumination, only a small number of photoelectrons are emitted; v remains close to $-1.5 v$; and the output on scanning is large. The iconoscope gives a black positive output.

Maloff has suggested an equivalent electrical circuit for the iconoscope, based on the charging characteristic of the bombarded mosaic surface. The use of such a circuit is necessarily limited because the

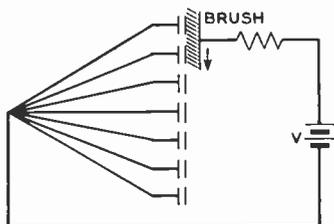


Fig. 6-17. Maloff's equivalent circuit for an iconoscope. (Courtesy of *Electronics*.)

entire process taking place in the iconoscope is not completely understood; the equations which result from the circuit do not permit the ready calculation of output voltage from the device. However, the circuit does serve to give a physical picture of some effects occurring in the tube. Figure 6-17 shows the scanning action of the electron beam to be equivalent to a brush commutating across the sub-elementary capacitances on the mosaic.

Each element is charged to some voltage E_o which depends upon the photoemission from that element and upon the number of excess electrons which were released from other elements, failed to reach the collector, and returned to that element. As the brush commutates across the elements a charging current proportional to the difference between v and E_o flows and restores each condenser to its equilibrium potential. Since E_o varies from element to element, the charging current varies in a like manner. It is this *difference* in charging current which is the output current of the iconoscope.

The shower of excess secondaries tends to neutralize some of the charge stored by photoemission. In fact, it has been estimated that only about one-quarter of the picture charge is left on an element when it is scanned. Moreover, the electric fields at the mosaic are such that they prevent saturated photoemission, and only about one-third of the predicted number of photoelectrons leave an element. These two effects combine to give the storage efficiency factor, η , previously defined in connection with eq. (6-55). From the data previously given its approximate, average value is

$$\eta \approx \left(\frac{1}{4}\right)\left(\frac{1}{3}\right) \approx 8 \text{ per cent}$$

We have just seen that the excess secondary electrons emitted at the mosaic tend to cancel stored charge. They also cause two other deleterious effects: They prevent a fixed output level corresponding to a black signal, and they cause the output current to contain a spurious component which causes a dark spot in the final reproduced picture. These effects will be considered in order.

In the absence of secondary emission an unilluminated region on the mosaic would remain at the equilibrium potential. When scanned by the electron beam the instantaneous charging current would be zero—and furthermore, it would be zero for *every* unilluminated mosaic region. Were these conditions to obtain in the iconoscope, the black regions of the picture would always deliver the same fixed level of output voltage or, stated in other words, the device would have a fixed black level.

As we have seen, however, this ideal condition does not obtain in the tube. The charge on an unilluminated mosaic area will not be fixed but will depend upon the number of excess secondary electrons which have arrived on that region. To further complicate matters the distribution of these excess secondaries is not uniform; it tends to be almost random, being affected to some extent by the picture content. The apparent reason for this quasi-randomness is that these secondaries are attracted to the more positive regions of the mosaic, *i.e.*, to those regions which have the highest illuminance. From these facts it follows that the unilluminated regions are not at equilibrium potential, and worse, their potential is not fixed but depends upon the distribution of the secondaries. When scanned, these black regions require a charging current and deliver an output which is not fixed. There is no fixed black level in the output.

This condition is further aggravated because the picture information is coupled to R_o through the mosaic-to-signal plate capacitance which makes the average value of the output voltage zero. Since the average value of the generated signal should be proportional to the average scene brightness or background level, the significance of this fact is that background level information is absent in the output voltage developed across R_o . These conditions are shown in Fig. 6-18. Fortunately they may both be corrected with the use of the d-c insert circuit described in a subsequent section.

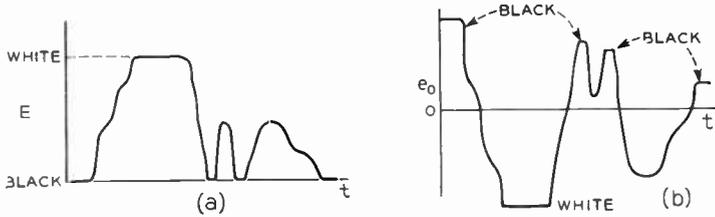


Fig. 6-18. The output of an iconoscope lacks a fixed black level and background level information. (a) Variation of illuminance along a scanning line. (b) The corresponding black-positive output signal. Notice that the black level is not fixed and that the average value of the signal is zero.

The so-called dark spot is also the result of the quasi-random distribution of the excess secondary electrons over the mosaic surface. The net potential distribution on the mosaic may be considered to consist of two components, one due to the picture and one due to the shower of excess secondaries. When scanned the mosaic delivers an output which again contains these same two components, the latter



Fig. 6-19. A televised image showing the effect of iconoscope dark spot. (Courtesy of American Broadcasting Company.)

of which is spurious and unwanted. The regions which receive the largest number of the redistributed electrons are most negative and will show up dark in the output; thus the spurious component of signal produces a dark region or dark spot in the final picture. A typical example of this is shown in Fig. 6-19.

We may sum up the situation this way: regardless of the mosaic illumination the iconoscope will deliver an output signal which is caused solely by the quasi-random distribution of the excess secondary electrons. When a picture is present this dark spot or spurious signal is superimposed on the picture signal and, like noise, cannot be separated from it. It is desirable, of course, to eliminate this spurious component of the iconoscope output. Generally speaking, there are two avenues of approach to a problem of this sort. The trouble may be eliminated at its source or some sort of compensating device may be used to cancel out its effect. Let us examine these possibilities.

If we are to eliminate the dark spot at its source, we must in some manner eliminate secondary emission at the mosaic for it is the secondaries that cause the difficulty. This, in turn, may be accomplished by reducing the accelerating voltage in the tube. Actually this is no solution at all, for if no secondaries are available to establish a conduction path between the mosaic and collector, there can be no output current from the iconoscope; the remedy is worse than the initial condition.

Other forms of camera tube, however, such as the orthicon, do not depend on secondary emission for operation and as a result do not have dark spot to the same extent as does the iconoscope. The alternate approach is to introduce into the signal a dark spot-canceling component. The method for accomplishing this is described in the next section.

6-13. Shading

In the preceding section we have seen that the output signal of the iconoscope has a spurious component, commonly called the dark spot, which results from a quasi-random distribution on the mosaic of excess secondary electrons. The process of compensating for this spurious component is known as shading and is the subject of the present section.

It is axiomatic that if we are to compensate for the effects of some

quantity, we must know something of its characteristics. In the case of shading we must know the common forms in which a dark spot manifests itself, and it is fortunate that these are comparatively regular and well known. To a first approximation, at least, the dark spot shows up in the following forms: a gradual shading across the picture, a shading from edge to center to edge, a gradual shading vertically on the picture, or a combination of these. Some of these are illustrated in Fig. 6-20.

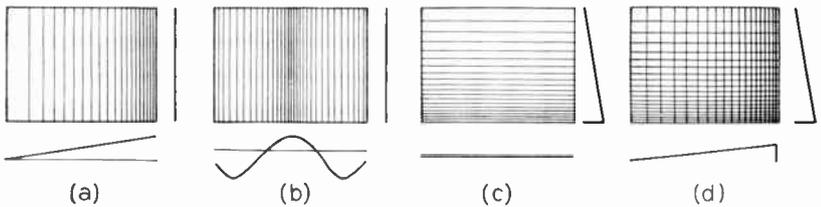


Fig. 6-20. Common forms of dark spot. Each may be canceled by addition of its inverse, which is furnished by the shading voltage generator. (Photos courtesy of American Broadcasting Company.)

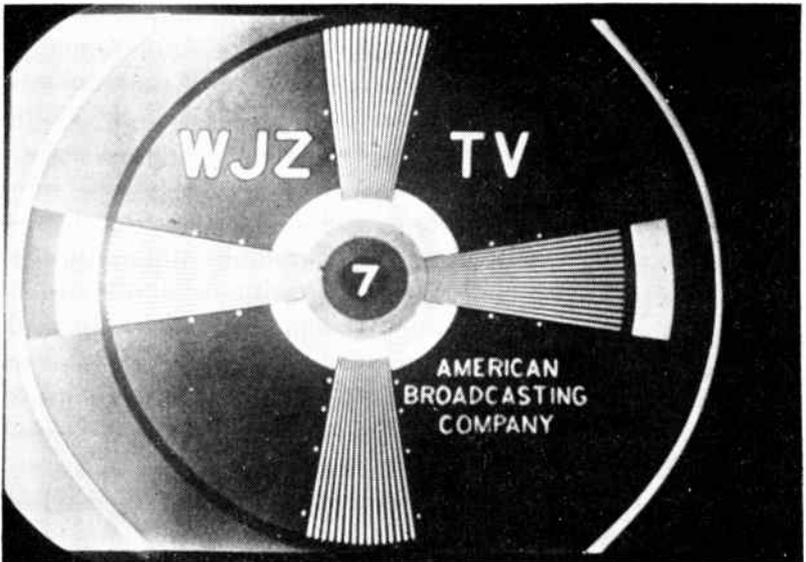


Fig. 6-20e. Actual appearance of the horizontal saw-tooth shade of Fig. 6-20 (a).

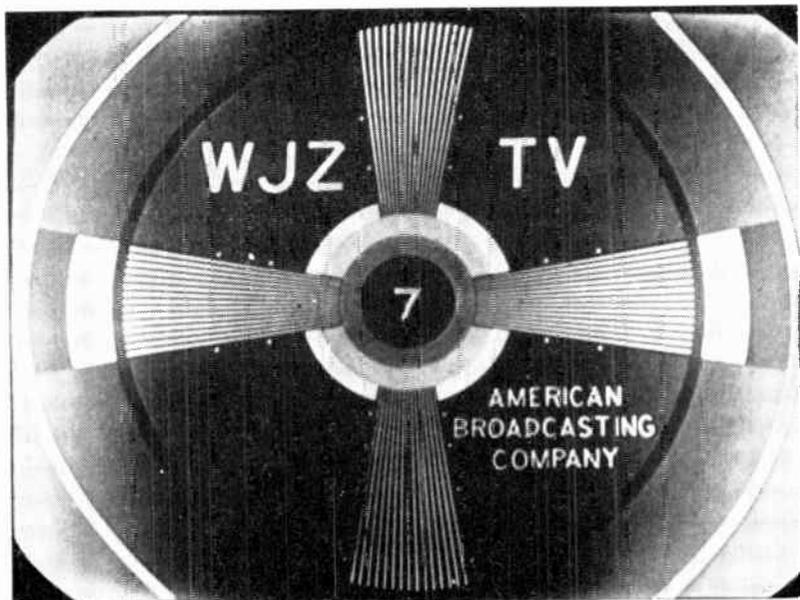


Fig. 6-20f. Actual appearance of the horizontal parabola shade of Fig. 6-20 (b).



Fig. 6-20g. Actual appearance of the vertical saw-tooth shade of Fig. 6-20 (c).

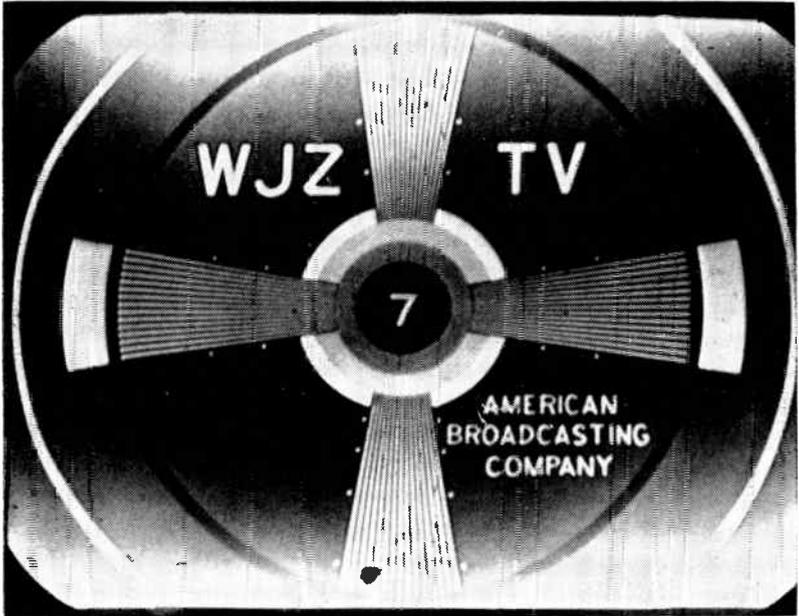


Fig. 6-20h. Actual appearance of a vertical parabola shade

Once the horizontal and vertical components of the dark spot are known, their inverses may be added to the signal. If this is done properly the spurious component will be canceled out of the picture signal. It must be realized that in the presentation of a television program there is insufficient time to analyze the dark spot, consequently the procedure used is to make available a number of suitable correcting voltages to a trained operator. Then, watching the final picture on a monitoring cathode-ray tube, he can adjust these voltages until the picture is free from the spurious signal. In practice the procedure is less difficult than it sounds; the controls require little adjustment except where the over-all picture level changes abruptly, such as on a change of scene. A typical shading generator diagram is given in Fig. 6-21. The controls for the horizontal saw-tooth shade voltage are indicated by P_1 and S_1 . P_1 is a conventional potentiometer that allows adjustment of the amplitude of the correcting saw-tooth wave. The polarity of the correction voltage may be reversed by means of S_1 , which changes the number of stages of amplification from an odd to an even number or *vice versa*.

The parabolic wave form, or "center push" as it is frequently called,

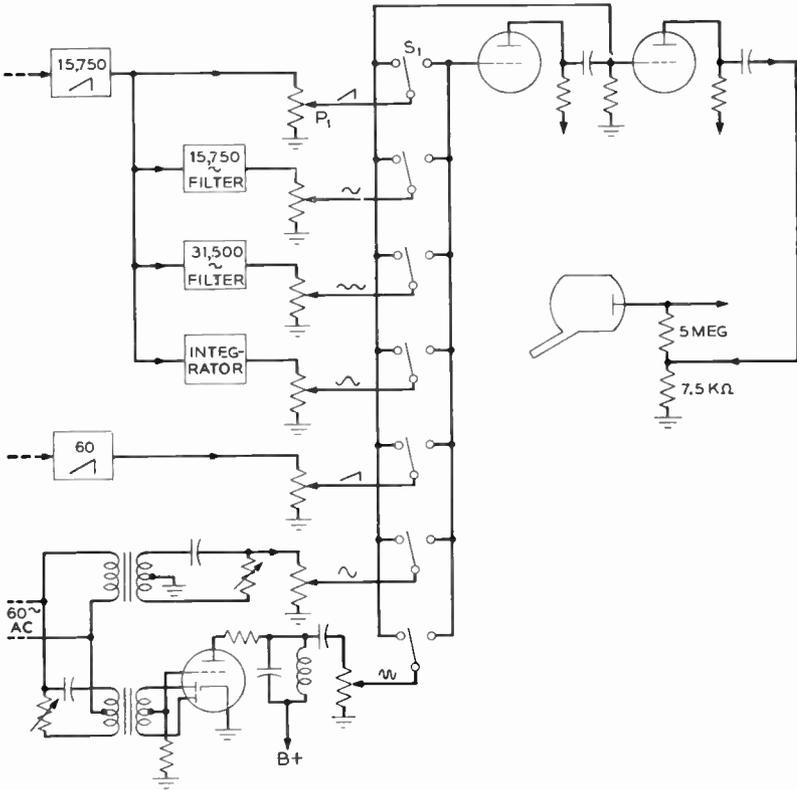


Fig. 6-21. A typical shading voltage generator. Each of the shading signals may be added to the iconoscope output signal to cancel the dark spot.

is obtained by electrically integrating a saw-tooth voltage with an R - C circuit.

In certain television studios it is felt that other wave forms in addition to the saw tooth and parabola are necessary for proper shading. The diagram of Fig. 6-21 shows provisions for correcting with sine waves at line and frame frequency and at twice these values. It might be well to discuss how these wave forms are produced. It is assumed that the saw-tooth voltages at horizontal and vertical sweep frequency are available. By means of Fourier analysis it may be shown that a saw-tooth wave contains both odd and even harmonics

of the fundamental repetition frequency; consequently there are present in the wave the fundamental and second harmonic. These components may be separated out by filters as shown in the diagram and used for shading. Under commercial telecasting standards the frame and power line frequencies are identical, so it is rather foolish to filter out a 60-cycle component from the vertical saw tooth. It may be obtained directly from the power line through a step-down transformer. The 120-cycle component is obtained by rectifying the 60-cycle component with a full-wave rectifier. The rectified output is rich in second harmonic, which may be filtered out by some circuit of the form shown in the diagram. The student should realize that numerous variations of Fig. 6-21 are possible; the diagram only suggests a few sources of the various shading voltages. It is felt in some quarters that a maximum of six voltages (saw tooth, parabolic, and fundamental sine wave each at line and field frequency) are adequate, and that the monitoring operator cannot handle more than this number.

The method of combining iconoscope output and shading signals is worthy of note. As may be seen from the diagram the shade signals are applied across a low-resistance tap on a 5-megohm resistance. By this device the iconoscope is made to see an essentially constant load regardless of adjustments in the shading circuit. It will be shown in the next section that the 5-megohm resistance is not the iconoscope load resistance, R_o , shown in Fig. 6-11.

6-14. The Coupling Circuit

We have already seen that the maximum possible signal-to-noise ratio of the whole television system is set at the coupling circuit between the camera tube and the first preamplifier stage. As a result this coupling circuit must be designed with some care to provide a compromise between high signal, maximum signal-to-noise ratio, and adequate bandwidth or high-frequency response. Such a compromise design has been described by Barco.³¹ From our previous work we know that the output signal voltage from the iconoscope will be the product of the output current and R_o , or

$$E_o \propto R_o \quad (6-57)$$

³¹ A. A. Barco, "An Iconoscope Pre-Amplifier." *RCA Review*, IV, 1 (July 1939).

On the other hand, the Johnson noise in the resistor—and we shall consider this to be the predominant noise source in the circuit—is proportional to the square root of R_o , or

$$E_n \propto \sqrt{R_o} \quad (6-58)$$

From these two relationships we see that the signal-to-noise ratio tends toward an optimum as R_o is increased for

$$\frac{S}{N} = \frac{E_o}{E_n} \propto \sqrt{R_o} \quad (6-59)$$

Viewed from the bandwidth point of view, however, R_o should be as small as possible to minimize the shunting effect of capacitance on the high-frequency signal components. This consideration imposes a severe limit on the value of resistance. For example, for a shunt capacitance of 25 μmf a value of 1270 ohms is required for R_o to give a 5-mc half-power point.

By way of compromise two paths are available: (1) reduce the shunt capacitance and (2) tolerate some frequency distortion and compensate for it in later stages of the video amplifier. The methods of compensating video amplifiers are covered in Chapter 7. In the circuit described by Barco both methods are used.

Consider first the shunt capacitance present. This capacitance is the sum of the circuit strays, the preamplifier input capacitance, and the iconoscope output capacitance. The first may be minimized by careful wiring techniques. The last two may be reduced by using degeneration in the associated circuits, for example, the first preamplifier stage is made a cathode follower. While the operation of this circuit is well known, we shall review it to show how a similar circuit may be applied to the iconoscope proper.

The cathode follower is a stage operated with its plate at a-c ground potential, the output being developed across a load resistor between cathode and ground. The basic circuit is shown in Fig. 6-22a. We may solve for the input admittance of the stage. Thus

$$Y_{in} = \frac{I_{in}}{E} \quad (6-60)$$

For the input circuit

$$\frac{I_{in}}{j\omega C_{pk}} + (I_{in} + I_p)R_K = E \quad (6-61)$$

Generally,

$$I_{in} \ll I_p$$

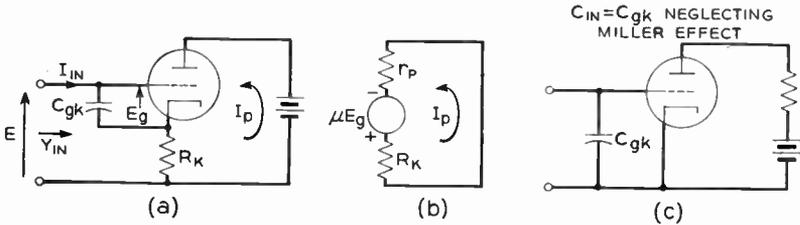


Fig. 6-22. Cathode degeneration may be used to reduce the input capacitance of the stage.

hence (6-61) may be simplified to

$$\frac{I_{in}}{j\omega C_{gk}} + I_p R_K = E \tag{6-62}$$

By inspection,

$$E_g = E - I_p R_K \tag{6-63}$$

and from the equivalent plate circuit of Fig. 6-22b

$$I_p = \frac{\mu E_g}{r_p + R_K} = \frac{\mu(E - I_p R_K)}{r_p + R_K} \tag{6-64}$$

whence

$$I_p = \frac{\mu E}{r_p + (1 + \mu)R_K} \approx \frac{\mu E}{r_p + \mu R_K} \tag{6-65}$$

where $\mu \gg 1$. Combination of eqs. (6-60), (6-61), and (6-65) yields

$$Y_{in} = \frac{j\omega C_{gk}}{1 + g_m R_K} \tag{6-66}$$

A conventional grounded cathode stage is diagrammed in Fig. 6-22c. If the Miller effect be neglected, the input capacitance will be simply C_{gk} . It can therefore be seen that the degeneration provided by the cathode resistor R_K in the cathode follower circuit effectively reduces the input capacitance of the stage by the factor $1/(1 + g_m R_K)$.

Let us consider the physical significance of this reduction in capacity. C_{gk} is the actual interelectrode capacitance between the grid and cathode. In the conventional stage the cathode is grounded and C_{gk} appears directly across the input terminals; hence C_{gk} and the input capacitance are one and the same. In the cathode follower circuit the cathode is above ground by the $I_p R_K$ drop. C_{gk} is no longer directly across the input terminals and the effective input capacitance

is determined by the magnitude of I_{in} , as shown in eq. (6-60). In the degenerative circuit, when a voltage is applied to the input, $E_o = E - I_p R_K$ is the voltage appearing across the circuit capacitance C_{ok} . The smaller condenser voltage gives a smaller input current with a corresponding reduction in effective capacitance appearing between the input terminals.

This idea may be restated in a slightly different manner. In Fig. 6-22c an applied voltage E causes a current $j\omega C_{ok}E$ to flow. In

Fig. 6-22a the same applied voltage produces the current $\frac{j\omega C_{ok}E}{1 + g_m R_K}$.

The lower current in the second case is the result of E being bucked by $I_p R_K$ and gives a lower effective input capacitance. We shall see presently that this same idea may be applied to the iconoscope itself.

Two other points must be mentioned in connection with the cathode follower input stage. First, in order to hold the tube noise to a minimum a triode-connected tube is used. It has previously been pointed out that the triode has no partition noise and consequently has less noise than an equivalent pentode. Since a high g_m tends to reduce tube noise, it is customary to use a high- g_m tube, such as the 6AC7, but triode-connected, *i.e.*, with both the screen and plate at a-c ground potential.

The second point has to do with the biasing of the cathode follower stage. Generally R_K will be of such a value that the d-c drop across it is greater than the rated bias for the tube in use. This is overcome by connecting the grid return resistor to a tap on R_K rather than to ground. By proper adjustment of the tap the d-c cathode-to-tap voltage may be set equal to the required bias. Tapping of R_K for bias adjustment is shown in Fig. 6-23a. Notice that this tap has no effect on the input capacitance; it will, however, change the conductive component of the input admittance.

We shall now determine how the same sort of treatment may be used to reduce the iconoscope output capacitance. In general, this capacitance will consist of two components, one within the tube between the signal plate and collector and another one which is the capacitance between the signal plate and the grounded shield which surrounds the camera tube and the entire video preamplifier chain. This latter external component may be reduced in a manner similar to that just described. As shown in Fig. 6-23b, a shield composed

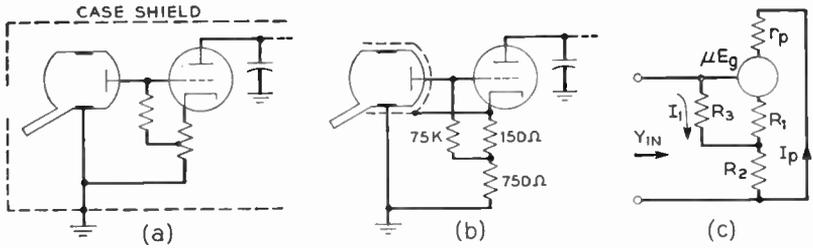


Fig. 6-23. Reduction of the signal plate-to-ground capacitance. (a) There is capacitance between the signal plate and case shield. (b) A second shield is placed between the iconoscope and the case shield and is connected to the triode cathode. (c) The equivalent circuit of the triode input.

of fine wires is placed on the iconoscope. Since this shield lies between the grounded case shield and the signal plate, it now becomes the determining factor in the output capacitance, *i.e.*, the signal plate to case-shield capacitance has been replaced by that between signal plate and iconoscope shield. If this iconoscope shield is connected to the cathode as shown instead of to the ground, the effective capacitance between signal plate and ground is reduced. This follows directly since the signal-plate to shield capacity is in parallel with C_{pk} and will be reduced by the same factor $1/(1 + g_m R_K)$.

At the beginning of this section it was pointed out that two approaches to the compromise design could be used. We have considered means for reducing the shunt capacitance; now we must determine how much loss in bandwidth can be tolerated at this point in the circuit. This will determine the maximum permissible value of R_o . Barco has recommended a value in the vicinity of 300 kilohms. This gives a reasonable signal-to-noise ratio and at the same time maintains sufficient high-frequency response that subsequent video compensation is able to give the required half-power bandwidth. A good question at this point is: What is R_o in Fig. 6-23b? Recall that R_o is the output load resistance for the iconoscope. The output current from the tube flows from signal plate to collector through the external circuit. Thus R_o is the total resistance between these two points. Since the plate current of the first preamplifier tube flows through R_2 , which is part of R_o , R_o does not comprise a completely passive network. Its value may best be checked by solving for the real part of the input admittance of the first preamplifier

stage. The equivalent circuit, neglecting the shunt capacitance, is given in Fig. 6-23c. Recalling that no grid current will flow because the tube is biased negatively and assuming that $I_1 \ll I_p$ we may write

$$E = I_1 R_3 + I_p R_2 \tag{6-67}$$

and
$$E_g = E - I_p(R_1 + R_2) \tag{6-68}$$

For the plate circuit

$$I_p(R_1 + R_2 + r_p) = \mu E_g = \mu[E - I_p(R_1 + R_2)] \tag{6-69}$$

whence
$$I_p \approx \frac{\mu E}{\mu(R_1 + R_2) + r_p} \quad \text{for } \mu \gg 1 \tag{6-70}$$

Substitution of (6-70) into (6-67) gives

$$E = I_1 R_3 + \frac{\mu E R_2}{\mu(R_1 + R_2) + r_p} \tag{6-71}$$

and, finally,
$$G_{in} = \frac{1}{R_3} \left[\frac{1 + g_m R_1}{1 + g_m(R_1 + R_2)} \right] = \frac{1}{R_o} \tag{6-72}$$

Evaluation of this equation gives an effective R_o lying between 200,000 and 300,000 ohms. It is interesting to note how the flow of plate current through R_2 raises the effective contribution of this resistance to R_o .

Actually (6-72) gives a value of coupling resistance which is slightly high because some shunt resistive network is required between signal plate and ground to permit injection of the shading signals. As described in the last section, this shunt network has a

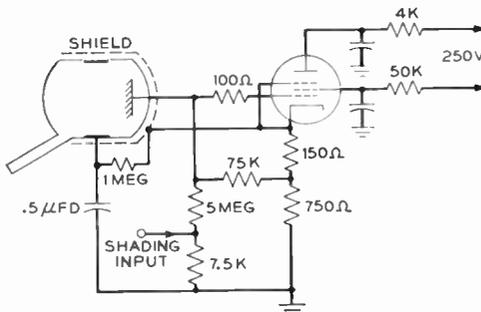


Fig. 6-24. The complete iconoscope output coupling network. The output voltage is between cathode and ground.

value in excess of 5 megohms and percentagewise will lower the calculated value of R_0 only slightly. This network is included in the complete coupling network of Fig. 6-24.

Reference to this figure will show that the iconoscope collector is at the same d-c potential as the first preamplifier cathode. It has been found experimentally that the shading is more satisfactory if the collector is at a slightly positive d-c potential with respect to the signal plate. It may be seen from the diagram that this potential difference is provided by the d-c IR drop across R_2 . The 1-megohm resistor and 0.5- μ f condenser comprise a filter for this voltage and ensure that the collector is at a-c ground potential. The 100-ohm resistor in series with the first preamplifier grid is used to suppress high-frequency parasitic oscillations.

6-15. Bias Lighting

In the actual use of an iconoscope as a television pickup tube a number of techniques are used which are the result of experience and not of theoretical considerations. One such technique involves the use of "bias-lighting" or "back-lighting," which affords an appreciable increase in the sensitivity of the tube. In its usual form back-lighting is provided by a number, say 4 to 6, of small incandescent lamps of the flashlight type driven from a d-c source. Means are provided to control the exciting current through them and they are physically arranged to illuminate the signal plate and walls of the iconoscope envelope in back of the mosaic. Care must be taken so that they do not contaminate the picture by shining on the photosensitive surface of the mosaic itself.

Although the mechanism by which this off-mosaic lighting increases sensitivity is not clearly understood, a probable explanation presumes that some of the cesium is deposited on the interior walls of the iconoscope envelope during the mosaic activation process. Apparently back-lighting prevents the building up of negative charge on these walls by causing electrons to be released from them. This, in turn, permits a better transfer of mosaic secondary electrons from the mosaic to the collector with a corresponding increase in the utilization of the stored information.

A second practical consideration in the use of the iconoscope requires that the entire surface of the mosaic, rather than only a large portion of it, be scanned by the electron beam. Aside from the

obvious fact that under-scanning produces an undersized picture, it can also introduce electrical difficulties. The unscanned portions of the mosaic tend to build up a negative charge resulting from the excess secondaries arriving there. This charge tends to leak over into the active mosaic area and cause a bloom which contaminates the picture information.

6-16. D-C Insertion

We have seen in a previous section that the iconoscope output lacks a fixed black level and background information. The importance of this fact is apparent from the following example, illustrated in Fig. 6-25, where the given object produces the same output voltage *swing* regardless of whether it is televised in broad daylight or in the shadow. In the diagram the contrast range between the spot and background is constant. At *a* the lighting is such that the background is black, whereas at *b* a 50 per cent gray background is present. The swing or contrast range in each case is the same but the average value locates the over-all position on the gray scale. At *c* the iconoscope signal has a zero average value and the position on the gray scale is lost.

It should be apparent that the signal at *c* could be converted to that of either *a* or *b* by adding the appropriate d-c or average value. If this is done, the black level will be reestablished at a certain fixed value. In actual practice the process is worked in reverse order; the black level is fixed at a definite d-c value, thus causing the average component to be reinserted.

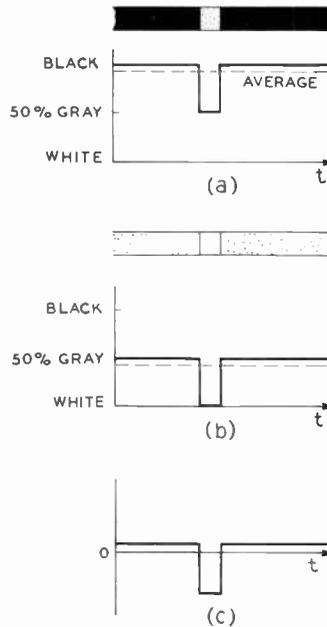


Fig. 6-25. Contrast level is maintained but background level is lost in the iconoscope output. (a) The output signal corresponding to a 50 per cent gray bar on a black background. (b) The output signal corresponding to a white bar on a 50 per cent gray background. (c) Both signals produce the same a-c output signal. The average value is reduced to zero.

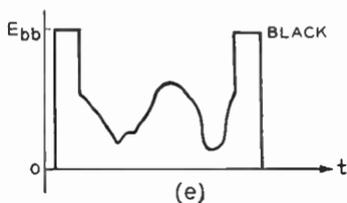
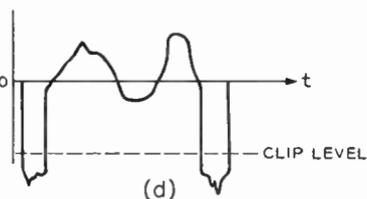
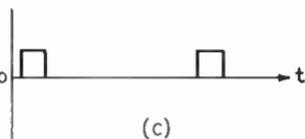
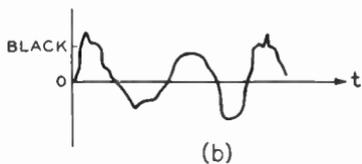
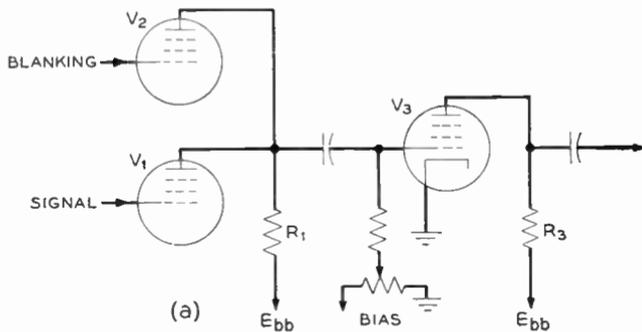


Fig. 6-26. The d-c insert circuit. Black level is set by adjusting the bias on V_3 . (a) Basic circuit. (b) Signal input to V_1 . The average value is zero. (c) Blanking input to V_2 . (d) Input to V_3 . (e) Output voltage across R_3 .

We have seen previously that no picture information is presented to the ultimate cathode-ray tube during the flyback or retrace intervals of the scan. To meet this condition we must make shift to ensure that the signal voltage is at black level during these intervals. This may be accomplished with the circuit of Figure 6-26. The amplified black-positive iconoscope signal is fed to the control grid of V_1 and will appear amplified and inverted across R_1 . Simultaneously a positive-going blanking signal is applied to the grid of V_2 . Since R_1 is the common load resistance for both tubes, the combined output across it will appear as at d in the diagram. Notice that the combined signal is forced in the negative direction during the blanking intervals. Any portion of the blanking interval signal may be forced below the cutoff level of V_3 by adjustment of the latter's bias. This cutoff level then becomes the black level for the output signal. Control of the bias sets the clip level, which becomes the established black level of the signal that must always correspond to the blanking intervals. "Pedestal" is the name given to this blanking interval black level.

It is immediately apparent that whatever agency sets the bias of V_3 must know the background level of the original scene. Hence the bias may be set by an operator who has the original scene under view. Alternatively a phototube may be used to view the scene and to automatically control the bias. Since it functions without scanning, this phototube responds to the average illumination; its control will be proportional to the average scene brightness.

Figure 6-26a shows the output of V_3 capacitively coupled to the following amplifier chain. It appears, therefore, that the d-c insert is probably of no avail since the coupling condenser will remove the d-c component just inserted. This subsequent loss of average value is not serious. Throughout the remaining portions of the signal channel the pedestal is recognized as black level, and clamping circuits of the type described in Chapter 7 and section 13-9 may be used to keep successive pedestals at the same voltage level, thereby reinserting the requisite average value. The function of the circuit of Fig. 6-26 is to establish for once and all the position of the pedestal relative to, say, a maximum white signal.

6-17. Keystone Correction

Little has been said about the sweep circuits associated with the iconoscope, except that magnetic deflection is used. From the work

of Chapters 3 and 4 we know that a saw-tooth deflection current is required and we have seen the type of circuit required to produce this current. An interesting variation of the usual sweep current is required for the commercial types of the iconoscope, however, because of the eccentric position of the electron gun with respect to the mosaic. As shown in Fig. 6-11 the gun is mounted in an off-axis position in order to provide a clear optical path between the window and the mosaic. Consider the effect which this has on the raster if the usual deflection currents are applied to the yoke. As the raster is scanned vertically, the amplitude of horizontal *angular* deflection remains constant. Since the gun is nearer the bottom than the top

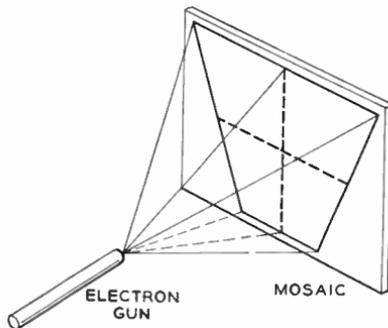


Fig. 6-27. Constant angular deflection produces a keystone-shaped raster on the mosaic because of the eccentric position of the electron gun.

of the mosaic, the horizontal lines will be narrower at the bottom than at the top, and the raster will be of keystone, rather than of the required rectangular, shape as shown in Fig. 6-27. This results in an intolerable situation for the over-all television system. Only the keystone portion of the mosaic is scanned and at the final kinescope this region is stretched into the conventional rectangle causing distortion of the viewed image.

To remedy this fault means must be devised so that the scanning signals applied to the iconoscope yoke produce a scan pattern of constant width on the mosaic, rather than of constant angular deflection. By way of review it should be recalled that the standard direction of scan requires that the mosaic be scanned from bottom to top and from right to left because the optical lens system inverts the image

on the mosaic. Thus the correction must be such that the *angular* horizontal scan *decreases* as the scan progresses along the field. This may be accomplished by causing the amplitude of the horizontal deflection current to decrease as indicated in Fig. 6-28a.

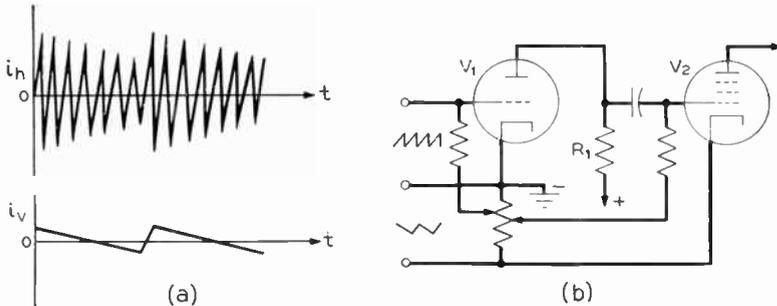


Fig. 6-28. Keystone correction. (a) Line and field deflection currents for keystone correction. (b) Simple form of keystone correction circuit.

A simple form of a keystone correction circuit³² is shown at b in the diagram. Since the required change in amplitude in i_h is in phase with i_v , the latter is used to develop a variable bias on the grid of V_1 . The resulting output across R_1 consists of the horizontal wave modulated at field frequency plus a spurious field frequency component resulting from the change in bias. This latter component is canceled out by returning the grid of V_2 to the same source of variable bias. The two field frequency components appearing at V_2 are in phase opposition and may be made to cancel out. The current output of V_2 , then, is of the required form to correct for the keystone pattern. Similar results may be obtained with a balanced modulator.

It may be seen from Fig. 6-27 that the electron gun of the iconoscope should be aimed at a point below the center of the mosaic if equiangular vertical deflection above and below the mosaic center is to be used. Proper gun location is handled during the manufacture of the tube.

If more than one iconoscope is used in a television studio, it is common practice to generate the sweep signals for all the cameras at

³² R. D. Kell, A. V. Bedford, and M. A. Trainer, "An Experimental Television System—The Transmitter." *Proc. IRE*, **22**, 11 (November 1934).

a common point. This master sweep generator provides the necessary keystone correction circuits, and the corrected sweep signals are fed by cable to the several cameras. This use of a common sweep generator may give rise to an odd situation where the cameras are used for televising different media. Say, for example, that camera 1 is used for televising live talent on the studio floor and camera 2 is used for film pickup. In this case, inspection of the cameras would show the iconoscope of camera 1 mounted with the electron gun down, while the second iconoscope would be mounted upside down, *i.e.*, with the electron gun up. The reason for this anomalous situation is apparent when it is remembered that a film projector throws an *upright* rather than an inverted image on the mosaic. If, then, camera 2 used a "right-side up" iconoscope, it would have to be scanned in the reverse direction from camera 1 and a separate keystone correction generator would be required. The need for this additional correction circuit is eliminated by the simple expedient of mounting one of the iconoscopes in an inverted position.

6-18. An Appraisal

We shall briefly summarize the characteristics of the iconoscope in order to evaluate its position in the television industry. Foremost of its characteristics is its ability to store or integrate picture information over an entire frame interval. Undesirable effects within the tube reduce the actual output down to only 5 to 10 per cent of its theoretical value, but this actual output represents a tremendous gain over that obtained with the instantaneous type pickup tubes.

A second undesirable effect due to a quasi-random distribution of the excess secondary electrons on the mosaic is the generation of a spurious signal component, the dark spot. This may be canceled out with suitable voltages. Since the position and shape of the dark spot change in time a monitoring operator is required to make corresponding changes in the compensating shading voltages.

The color response of the iconoscope does not match that of the average eye. Considerable control of the color response is available during manufacture when the mosaic is being activated. To counterbalance the color response, a technique of lighting and make-up has evolved which give satisfactory results for studio work.

Over a period of years operating experience has shown that in televising the usual range of subjects a scene illuminance of 1000 to 2000

foot-candles is required. The general feeling among live talent at least is that this relatively high level of illuminance when provided by arc or incandescent lamps makes the television studio uncomfortably hot. Our previous work has shown that the mosaic illuminance may be increased by using a bigger aperture or f number on the pickup lens. This will generally permit lowering the scene illuminance to well below the 1000 foot-candle level, but the depth of field suffers as a consequence.³³ Some idea of the typical lighting levels required for good operation of the iconoscope may be had from the fact that direct illuminance by sunlight on a summer day is in the order of some 9000 foot-candles.

The Illuminating Engineering Society has recommended the following light levels for gymnasiums:

Games with spectators—30 foot-candles³⁴

General assemblies and dancing—5 foot-candles

Taking these figures as typical of the illumination available at remote indoor pickup sites we can see that the iconoscope will give relatively poor results and has been largely superseded by the image orthicon for this type of service.

Probably the greatest single feature of the iconoscope in comparison with the other camera tubes to be described is that it has excellent resolution and delivers pictures of high quality. This is possible because the 1-kilovolt second anode potential permits extremely sharp focusing of the scanning beam. We have already seen that the resolution is determined by the beam diameter rather than by the smaller subelementary globules which make up the mosaic. As a result of the need for high illuminance, the iconoscope has been largely relegated to film and slide pickup but it still is capable of the highest resolution of all the camera tubes discussed in this chapter.

The last feature of the iconoscope which we wish to consider is the effect of its almost logarithmic output-illumination characteristic. To do this we must examine both ends of the television system. Ideally the pickup end should have a linear voltage-illumination characteristic and the output end a linear intensity-voltage relationship. These ideal characteristics would provide a one-to-one rela-

³³ The subject of depth of field is handled analytically by H. Iams, G. A. Morton, and V. K. Zworykin, "The Image Iconoscope." *Proc. IRE*, 27, 9 (September 1939).

³⁴ *General Electric News Digest*, January-February 1949.

tionship between original scene brightness and reproduced image intensity.

Over a wide range of control grid voltage the typical kinescope is linear in light output. It follows, therefore, that the logarithmic response of the iconoscope prevents the desired one-to-one light relationship, giving some compression in the picture highlights. It is fortunate, however, that the eye responds logarithmically to the sensation of light. Thus, as long as the system provides the same *ratio* of extreme light levels that is present in the original scene the contrast range is satisfactory to the human eye and the highlight compression is not too serious.

Historically then, if one may refer to a decade as history, the iconoscope pointed the way to two main types of camera-tube development. Research has been directed to the development of tubes with greater sensitivity and better linearity. A tube which showed improvement in these features would at once provide greater flexibility than the iconoscope, which has remained an outstanding milestone in the development of television.

6-19. The Image Iconoscope³⁵

One of the means of increasing camera-tube sensitivity is to use secondary-emission image intensification, a principle used in the image iconoscope which was announced at the Annual Convention of the Institute of Radio Engineers in 1939. This tube found little use in commercial practice but did serve as an important stepping stone in the development of the later tube types. We shall consider it chiefly as an illustration of the principle of image intensification.

As shown in Fig. 6-29, the image iconoscope differs from the iconoscope in two main respects: the mosaic is not photosensitized, and a semitransparent photocathode is included. In operation an optical image is focused on the photocathode which emits a corresponding electron image in a manner similar to that previously described for the image dissector. Under the action of the axial magnetic field and the electric field present this electron image is accelerated toward the mosaic. The image electrons arrive at the mosaic with sufficient velocity to produce secondary emission from the subelementary islands. Thus picture information is stored in the form of charge in the mosaic-to-signal plate capacitance just as it is in

³⁵ H. Iams, G. A. Morton, and V. K. Zworykin, *op. cit.*

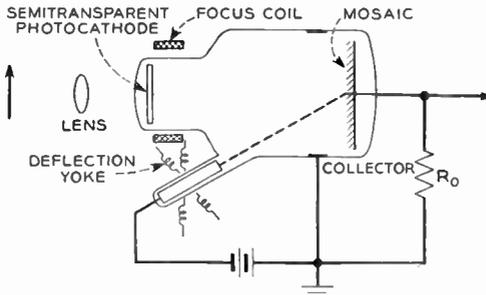


Fig. 6-29. The image iconoscope.

the iconoscope. The difference is that in the latter case the stored charge is the result of photoemission resulting from the incident optical image, whereas in the image iconoscope it is caused by the secondary emission due to the incident electron image. The output is developed when the stored charge is released by the action of the scanning beam moving across the mosaic.

The image iconoscope provides a sensitivity 6 to 10 times that of the iconoscope, the gain being the result of the following factors.

(1) The photocathode is more efficient as a photoemitter than a photosensitive mosaic. Typical values of luminous sensitivity which may be obtained are 20 to 50 $\mu\text{a}/\text{lumen}$, the higher values corresponding to those surfaces which are rich in infrared response.

(2) A high electric field intensity exists at the photocathode surface which permits saturated photoemission. This is in contrast to the condition in the iconoscope where only about 20 to 30 per cent of the photoelectrons are drawn away from the mosaic.

(3) The mosaic has a high secondary emission ratio, ranging from 3 to 11, thus the stored charge is r times as great as the charge leaving the photocathode, r being the secondary emission ratio. An added advantage is that the photocathode is close to the optical window in the tube envelope. This permits the use of a short focal length lens.

On the other side of the ledger, the image iconoscope is more expensive and requires more auxiliary equipment and more adjustments. Since the output is developed by scanning the mosaic with a high velocity electron beam the dark spot is still present, although to a lesser degree than in the iconoscope. Both tubes have approximately the same resolution.

THE ORTHICON³⁶

In both the iconoscope and image iconoscope the mosaic is scanned by high-velocity beam electrons which produce secondary emission at the mosaic surface. One deleterious effect of the secondary emission is the production of a spurious signal, the dark spot. If the scanning beam velocity be reduced sufficiently, the secondary emission may be eliminated and the dark spot with it. If this scheme is to be utilized in a camera tube, some mechanism other than that of the iconoscope must be found to generate the output current. We consider next the question of low-velocity scanning and how it may be used in a television camera tube.

6-20. Low-velocity-electron Scanning

In the tube shown in Fig. 6-30 a low value of accelerating voltage, 25 volts, is used. Further, the mosaic is considerably farther away from the electron gun than is the collector electrode. As a result of

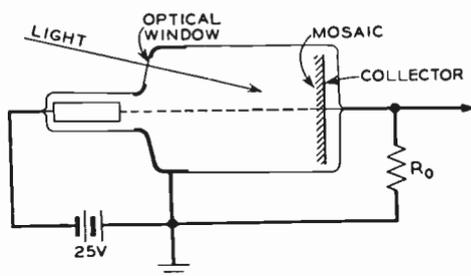


Fig. 6-30. The basic low-velocity electron-beam camera tube.

this configuration the beam electrons are accelerated as they move toward the collector and then decelerated as they move past it. Notice that an equilibrium condition will set in here which is quite different from that described for the iconoscope. When the device is first turned on, the mosaic will be at ground potential, *i.e.*, +25 volts relative to the electron gun cathode and the beam electrons will strike the mosaic. Since their velocity is too low to produce secondary emission, these primary electrons will remain on the mosaic which

³⁶ A. Rose, and H. Iams, "The Orthicon, a Television Pick-Up Tube." *RCA Review*, IV, 2 (October 1939).

gradually swings negatively relative to ground. When the mosaic and cathode potentials are approximately equal, the mosaic will repel the beam electrons, which will then reverse direction and return to the collector.

Let a light image be focused on the mosaic. Photoemission proportional to the illumination will take place. Since the mosaic is negative with respect to the collector, all the photoelectrons are pulled away to the collector and the photoemissive process is saturated. Furthermore there will be no dark-spot distribution on the mosaic and the image will be retained as stored charge in the mosaic-to-signal plate capacitance. Consider a small element of mosaic whose voltage is positive relative to its equilibrium value. Instead of repelling the scanning beam as the latter approaches that element position in the raster, it will attract just enough of the beam electrons to restore itself to equilibrium potential. The resulting change of charge produces an output current through R_o .

In a tube which is to utilize this principle of operation, it is of great importance to have the scanning beam approach the mosaic at right angles because the mosaic can only repel the normal component of the beam electron velocity. Hence if the beam is inclined to the mosaic, it tends to glance off the mosaic. This will cause the beam shape to vary with the point of impact. Furthermore, the repelling voltage necessary to prevent the beam from landing is determined only by the normal component of the beam velocity. Therefore the glancing beam would charge different points of the mosaic to different potentials, a condition which would result in the generation of a

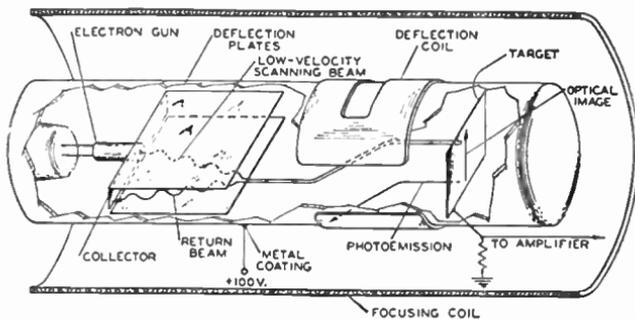


Fig. 6-31. Schematic diagram of the orthicon tube.
(Courtesy of *RCA Review*.)

spurious signal not unlike the dark spot. A second major problem in the proposed tube is to provide a sufficiently large beam current in the narrow scanning beam as it slows down near the mosaic. Obviously the first condition cannot be met by the tube of Fig. 6-30 if conventional scanning means are employed; it is met, however, in the orthicon tube shown in Fig. 6-31.

6-21. The Deflection System

The horizontal deflection system in the orthicon is composed of an axial magnetic field furnished by a long focus coil and an electric field normal to the tube axis. Neither field affects the axial component of the electron velocity, v_o , which will remain constant. Although the motion of an electron in these crossed fields may be derived analytically, let us approach the question from a physical point of view. Thus in Fig. 6-32a an enlarged view of the deflection plates

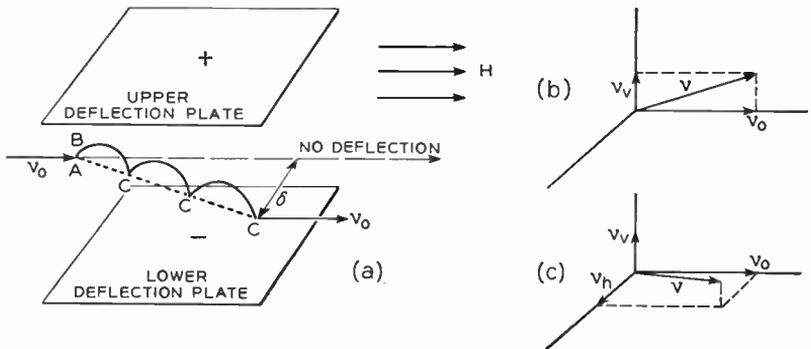


Fig. 6-32. Horizontal deflection in the orthicon. (a) Deflection with crossed electric and magnetic fields. H is due to the focus coil and is parallel to the direction of v_o . (b) Electron velocities at the point A. (c) Electron velocities at the point B.

may be seen. The upper plate is assumed to be positive with respect to the lower plate. Consider the forces on an electron as it enters the electrostatic field at point A with an axial velocity v_o . Since the electron is moving parallel to the magnetic field, the latter will have no effect on the electron motion. The electrostatic field, on the other hand, attracts the electron upward, giving it a vertical component of velocity v_v . The resultant velocity will be the vector sum, v ,

shown at b in the diagram. The electron now has a component of motion cutting across the lines of magnetic flux. Use of the left-hand rule (remember current and electron motion are in opposite directions) shows that at a nearby point B the electron will have a horizontal component of velocity v_h . This added to the corresponding v_v and v_o gives the resultant velocity shown at c in the figure. If this argument be carried out point by point, it will be seen that the electron follows a cycloidal motion in moving from left to right through the electric field deflection plates. At each cusp on the cycloidal path (the points labeled C in Fig. 6-32a) the horizontal velocity component is zero; hence if adjustments can be made so that the electron emerges from between the deflection plates just as it reaches a cusp, it will emerge with zero horizontal velocity, *i.e.*, it will once again be moving perpendicular to the mosaic. The net effect of the horizontal deflection system is to cause the electron to suffer a lateral displacement δ without affecting its velocity in magnitude or direction. δ is proportional to the electric field and the time of flight from one edge of the plates to the other, and inversely proportional to the magnetic field intensity, H .

It may be observed that as the amplitude of the cycloids is reduced, so is the horizontal velocity component at any point along the cycloid. If this amplitude were made very small, the location of a cusp at the point of emergence would be less critical, a necessary condition because the electric field varies in time. To this end the deflection

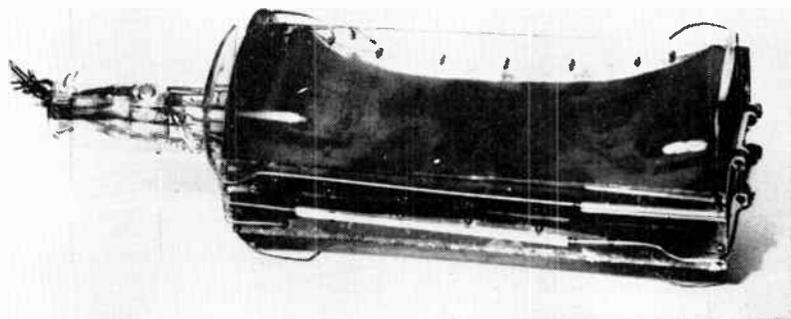


Fig. 6-33. The horizontal deflection plates of the type 1840 orthicon. The plates are flared to reduce the amplitude of the cycloidal motion of the electron as it leaves the plates.

plates actually used flare outward toward both edges. This produces an electric field which weakens as the plates flare out and the desired reduction in cycloid amplitude results. Figure 6-33 shows the horizontal deflection plates from a type 1840 orthicon tube.

We have already noted that δ , the lateral displacement, is proportional to the electric field intensity or deflection voltage; hence a horizontal saw-tooth deflection may be produced by applying the usual saw-tooth voltage associated with electrostatic deflection to the deflection plates. It is significant to notice, however, that in contrast to the simple electrostatic case the deflection is normal to the direction of the electric field, a result which is produced by the presence of the axial magnetic field. Also note that lateral motion of the beam can occur only when the beam is between the plates. Therefore the plates must be at least as wide as the desired raster width on the mosaic.

Figure 6-31 shows that the vertical deflection, a low-frequency phenomenon, is accomplished with a magnetic yoke. Once again the axial magnetic field modifies the deflection that would be obtained with the yoke alone as described in Chapter 3. Its net effect is to produce a deflection parallel, rather than perpendicular, to the deflecting magnetomotive force. Actually the path of the electrons when they are in the region of the crossed magnetic fields is helical. For small deflections the amplitude of the helix is small and the system may be considered to translate the electron beam vertically without adding any new components to its velocity. Again, the vertical translation only occurs within the crossed-field region; hence, the vertical deflection yoke must be at least as high as the mosaic. As shown in Fig. 6-31, it is convenient to place this yoke around the exterior of the orthicon envelope. Saw-tooth current in the vertical deflection yoke produces the required vertical deflection.

By way of summary we note that the horizontal and vertical deflection systems just described do meet the requirement of delivering the scanning beam always in a direction normal to the mosaic.

6-22. Characteristics

Once again consider the action that takes place within the orthicon. Saturated photoemission charges the regions of the mosaic positively. A low-velocity scanning beam is repelled by the dark, unilluminated regions of the mosaic. This means that black always corresponds to zero output volts, *i.e.*, the orthicon has a fixed black level.

As the low-velocity beam scans across a mosaic element which has lost photoelectrons, a number of beam electrons just equal to the number of photoelectrons lost during the storage interval are attracted to that element. Thus the difference between beam and collector currents is the charging current. This, in turn, is identical to the output current of the tube. The output current will be maximum from those elements which have emitted the greatest number of photoelectrons; hence the orthicon delivers a white-positive output.

Since no secondary emission occurs at the mosaic, no dark spot is present, and none of the stored charge is neutralized by a shower of excess secondary electrons. Furthermore saturated photoemission takes place; hence, compared to the iconoscope the orthicon has the same expression for output current, except that η , the storage efficiency factor, is very nearly 100 per cent. Thus the orthicon has roughly 20 times the sensitivity of the iconoscope.

In addition a linear relationship between electrical output and light input is obtained with the orthicon. It is this straight-line characteristic which gives the tube its name (*ortho*—straight, iconoscope—image viewer). Orthicon is the accepted abbreviation of the full name orthiconoscope.

In this connection it might be well to define the term “gamma,” which has been carried over into television practice from the photographic industry. Our remarks will be confined to the γ of the electrical system only. The output-input characteristics of several of the common transducers are power-law curves and have the general form

$$\text{Output} = K(\text{Input})^\gamma \quad (6-73)$$

Thus γ is defined as the exponent in (6-73). Taking logs of both sides of the equation we get

$$\log (\text{Out}) = \log K + \gamma \log (\text{In}) \quad (6-74)$$

which is the equation of a straight line of slope γ . γ for any device may be obtained, therefore, by plotting the output-input characteristic of the device on log-log paper.

The orthicon, being a linear device, has a γ of unity. Contrast this with the iconoscope, which has a γ of approximately unity for low levels of illumination but changes from power to logarithmic law over a wide range of illumination.

It is of interest to compare a typical orthicon, type 1840, and an

iconoscope, type 1850. The former is more sensitive but tends to give best operation at medium light levels. The iconoscope is more satisfactory at high levels and is capable of greater resolution. The 1840 is superior to the 1850 in its ability to reproduce the full length of the gray scale. The orthicon saw little use in television because of curtailed televising during the war years. It was superseded in 1946 by the more sensitive image orthicon.

THE IMAGE ORTHICON³⁷

At their Annual Convention in New York in 1946 the members of the Institute of Radio Engineers observed a demonstration of a new type of camera tube whose sensitivity was great enough to permit televising a subject illuminated only by the light from a single candle. This supersensitive tube, the image orthicon, derived its extreme sensitivity by combining the best features of the several

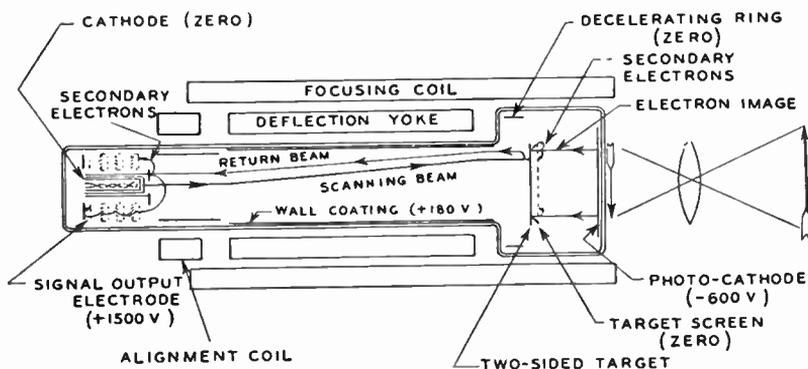


Fig. 6-34. Schematic diagram of the image orthicon tube. (Courtesy of *Proc. IRE*.)

camera tubes which we have studied. It incorporates the storage principle of the iconoscope and low-velocity beam scanning of the orthicon. It utilizes the electron image intensification of the image iconoscope, and derives further gain by adding an electron multiplier similar in principle to that described for the image dissector. In

³⁷ A. Rose, P. K. Weimer, and H. B. Law, "The Image Orthicon—A Sensitive Television Pickup Tube." *Proc. IRE*, 14, 7 (July 1946). See also, R. B. Janes, R. E. Johnson, and R. R. Handel, "Producing the 5820 Image Orthicon." *Electronics*, 23, 6 (June 1950).

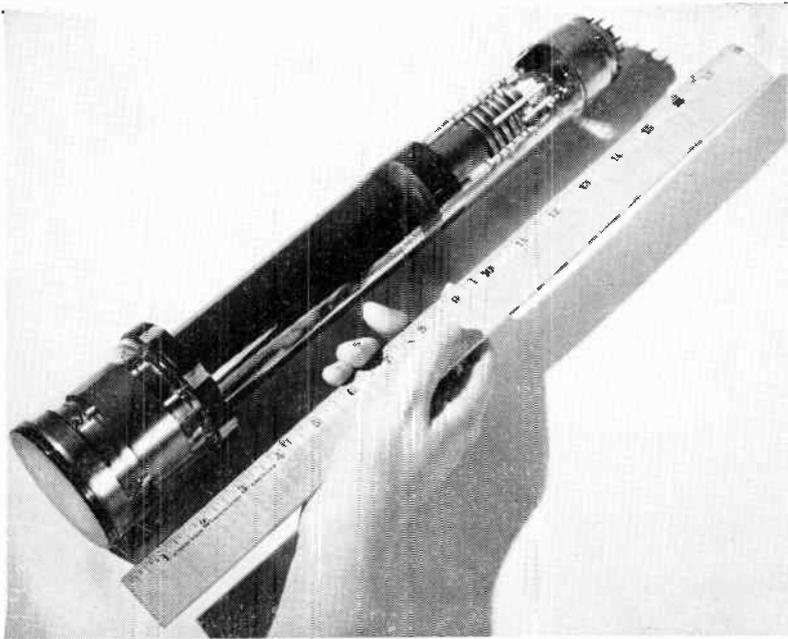


Fig. 6-35. The image orthicon tube. The five dynodes of the electron multiplier may be identified at the upper end of the tube. (Courtesy of Radio Corporation of America.)

physical form the image orthicon resembles an orthicon to which has been added an image intensifier section and an electron multiplier; it is shown in Figs. 6-34 and 6-35. Operation and construction of the tube may best be considered in three separate sections: the image intensifier, the target and scanning section, and the electron multiplier. We next consider these in order.

6-23. The Image Intensifier Section

The operation here is quite similar to that described for the image iconoscope. At the optical window end of the tube envelope is located a semitransparent photocathode, whose surface is held at -600 v relative to a mosaic target which is described in the next section. The electric field at the photocathode is favorable and saturated photoemission takes place from each element, the emission being proportional the illuminance. These photoelectrons, which

form a complete electron image of the scene are focused by the axial magnetic field onto the mosaic target. The image electrons arrive with sufficient velocity to produce secondary emission. The emitted secondaries are collected by a fine wire mesh which is held at one volt positive relative to the equilibrium potential of the target. With normal operating voltages, the secondary emission ratio, r , of the target is greater than unity. Thus a single incident image electron releases r secondary electrons. As will be shown, the target stores information in the form of charge over the interval between successive scans, τ . Thus per element, the photocurrent is given by eq. (6-39). This is intensified or increased by the factor r at the target and the charge stored during τ will be

$$q = rsaE\tau \quad (6-75)$$

where it is assumed that the time to release the stored charge, τ_e , is a negligibly small fraction of τ . This was shown in reference to (6-54). The factor r combined with the higher photoemission obtained gives roughly five times the stored charge obtained in the iconoscope.

Mention must be made of the screen which collects the secondary electrons released from the mosaic target. It will be realized that this mesh lies between the target and the photocathode, consequently if it is not to cast an electron image shadow on the target it must have a very fine structure which is largely open space. To this end special techniques have been developed which yield screens with 500 to 1000 mesh per linear inch and which are 50 to 75 per cent open space! These screens are so fine that they cast negligible shadow and permit the generation of high-resolution pictures.

6-24. The Target and Scanning Section

Reference to Fig. 6-34 shows that the electron gun and scanning system is similar to that of the orthicon, except that the horizontal deflection plates have been replaced by an additional winding in the deflection yoke. Thus both horizontal and vertical deflections are produced by magnetic deflection fields which are at right angles to the axial magnetic field.

It may also be seen from the diagram that as the scanning beam approaches the rear surface of the mosaic target, it is decelerated by an annular ring which is held at the same potential as the cathode of the electron gun (0 volts). From our experience with the orthicon,

we see that low-velocity beam scanning occurs and that the equilibrium potential of this rear target surface is zero volts.

Before considering the electrical action which occurs as the low-velocity beam scans the rear of the mosaic target we must investigate the structure of this latter element. The mosaic in the image orthicon differs in operation from those previously discussed in that the charge-storing emission does not occur from the surface which is scanned by the electron beam. What is used is a two-sided mosaic. This may be verified by reference to Fig. 6-34. Up to the present in our discussion we have thought of a mosaic as a large number of subelementary islands, each insulated from the others and each having capacitance through a dielectric to a rear signal plate. If, now, a sheet of fairly low-resistivity glass is blown thin enough, its resistance is extremely high laterally on the surface of the glass. If charge is stored between the surfaces on a minute area of this glass, that charge will not spread laterally over the surface but will in time neutralize itself between the two surfaces. These are the exact properties required for the two-sided mosaic target: the high across-surface resistivity of the glass effectively provides the myriad of insulated islands, the glass is a dielectric, and the comparatively low intersurface resistivity allows charge between the surfaces to be neutralized slowly, that is, during one frame interval. How these characteristics apply in the operation of the image orthicon is described below.

Actually three effects occur simultaneously at the two-sided target. We describe them separately as a matter of convenience. The cycle of operation in terms of voltage variation is shown in Fig. 6-36, where the reference is the cathode of the scanning beam gun, which is grounded. The secondary electron collecting screen or mesh is held constant at +1 v. At *a* the element of target, which is of the same area as the scanning beam, has just been scanned on its scanned or rear surface and has been restored to the equilibrium potential of 0 v. The front or secondary emission surface is also at ground potential for reasons to be described at the end of the cycle. At *b* secondary emission for an entire frame interval has driven the front surface of the element to +1 v. Note that this is the maximum positive value which it can attain because once its potential equals that of the collecting mesh, any released secondary electrons will return to the target rather than to the collector mesh. Notice further

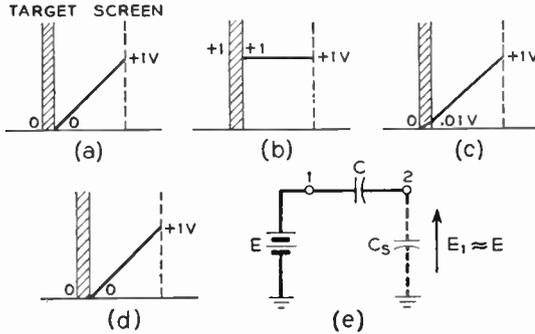


Fig. 6-36. Target potentials in the image orthicon. (a) Before scanning and exposure. (b) Before scanning and after exposure. (c) After scanning. (d) One frame later and before the next exposure. (e) Charging an insulated condenser. (After Rose, Weimer, and Law.)

that since the rear and front surfaces are coupled to each other through the interface capacitance, the rear surface will also rise to +1 volt. The mechanism behind this may become more clear if the reader refers to Fig. 6-36*e*. A battery of e.m.f. V is connected to one plate of an isolated condenser of capacitance C . The return path from terminal 2 to the grounded battery terminal is the stray capacitance to ground, C_s . Since the two condensers form a series voltage divider, the voltage drop from terminal 2 to ground will be

$$V_1 = V \frac{C}{C + C_s}$$

but in general $C_s \ll C$, so that V_1 and V are very nearly identical. This shows why both sides of the element are at the same voltage as shown at *b* in the diagram.

At *c* the rear surface of the element has just been scanned by the low-velocity beam. A sufficient number of electrons has been left by the beam to restore that surface to its equilibrium value of 0 v. At this instant the front surface of the element is charged positively and the rear surface negatively. The voltage drop between the two surfaces will be the stored picture charge divided by the interface capacitance of the element. This voltage drop is small compared to 1 v. and accounts for the small positive voltage shown on the front surface in the diagram. At *d*, this voltage has dropped to zero by conduction through the glass during the frame interval.

Remember, all these actions that have been described occur simultaneously. Since the illuminance of the photocathode varies over its surface, different elements of the target will be driven to different values of voltage between 0 and +1 v. Thus the number of charge-neutralizing electrons removed from the scanning beam varies in proportion to the illuminance and the change in *returning* beam current is the output.

We may state these results analytically in the following manner. As before, let τ_e be the time required for the beam to scan one element. Then the charging or output current is

$$I_{ch} = I_b - I_c = \frac{q}{\tau_e} = rsdE \frac{\tau}{\tau_e} \quad (6-76)$$

where I_b is the beam current and I_c is the collector current or current returning from the mosaic. But τ/τ_e is the figure of merit and Ma is the target area wh . Therefore

$$I_{ch} = rsEwh. \quad (6-77)$$

6-25. The Electron Multiplier

The electrons in the scanning beam have been described as having low velocity because they are almost at rest when they arrive at the rear target surface. They nevertheless have sufficient velocity in the body of the tube so that their transit time from gun to mosaic and back to gun again is a small fraction of the horizontal and vertical deflection cycles. Since this is true, on the return trip from the mosaic the unaccepted beam electrons will retrace very nearly their path toward the mosaic and will arrive very near to their original point of departure in the electron gun. This fact is utilized in the construction of the electron multiplier, which comprises a number of annular dynodes that are mounted concentrically to and along the length of the electron gun. The first, or collector, dynode, being near to the aperture which emits the beam from the gun, collects the returning, unaccepted beam electrons. As these are passed on from stage to stage, a current gain results by virtue of secondary emission at each dynode surface. Under normal operating conditions the five-stage multiplier affords a current gain of some 200 to 500 times. If m be the over-all multiplication ratio, the final output current of the image orthicon becomes from eq. (6-77)

$$I_o = mrsEwh \quad (6-78)$$

6-26. Noise

In its commercial form, such as the type 2P23, the image orthicon provides a sensitivity 100 to 1000 times that available with the iconoscope or orthicon tubes. Inasmuch as eq. (6-78) shows that the tube's output is proportional to m one might forget the story of Midas and ask why not increase the number of multiplier stages for an even more substantial increase in output. The answer lies in noise.

With the iconoscope the output current is so small that the shot noise from the scanning beam is largely masked by the Johnson noise in the coupling resistance, R_o . In the image orthicon on the other hand, the beam shot noise is multiplied along with the signal current, and shot noise may become a determining factor. Because of this the gain due to electron multiplication is of use only up to the point where the amplified shot noise is of the same order of magnitude as the resistance Johnson noise. Given the beam current of the tube one may calculate the maximum useful value of m by the use of eqs. (6-24) and (6-26) and Ohm's law.

6-27. Characteristics

From the equation for output current (6-78) we see that the relationship between output current and mosaic illuminance is linear just as it was in the orthicon tube. A review of the operating cycle of the two-sided target will show, however, that this linear relationship must break down when E reaches a sufficiently high value. This condition is reached when the voltage of a mosaic element becomes equal to the voltage of the mesh collector. For illumination levels in excess of this value linearity is not maintained but the tube can still deliver satisfactory pictures. An explanation has been proposed for this anomaly.

Consider the typical output-input characteristic shown in Fig. 6-37. In the range from A to B the output current is proportional to the illuminance. This is the "ortho" range described by eq. (6-78). At B the illuminance is great enough so that the whole target gets charged up to the potential of the collecting mesh in one frame interval. It would seem, therefore, that at or above B the whole image would saturate, but as a practical matter this saturation does not occur as far as the final image is concerned. In order to consider the mechanism behind this condition let us say that a white bar on a

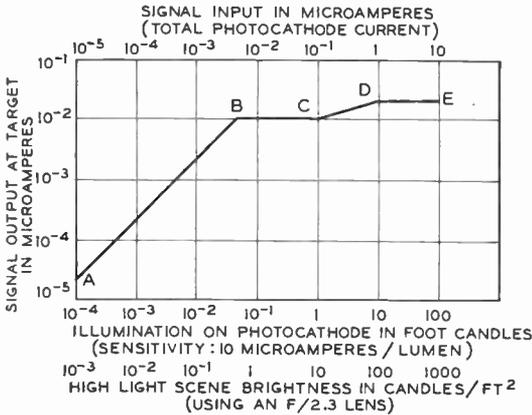


Fig. 6-37. Signal versus light characteristic of the image orthicon. (Courtesy of *Proc. IRE.*)

black background is being televised and that the average illumination is in the *BC* range of Fig. 6-37. The charge over the whole mosaic tends to saturate. Recall that the electron image from the photocathode releases secondary electrons from the front surface of the target. Some of these secondaries emitted from the region corresponding to the white bar will have sufficiently low velocity that they will return to adjacent areas on the target. The returning secondary electrons decrease the voltage on these adjacent areas and the latter deliver the required black signal when they are scanned. The remainder of the target which does not receive the returning secondaries

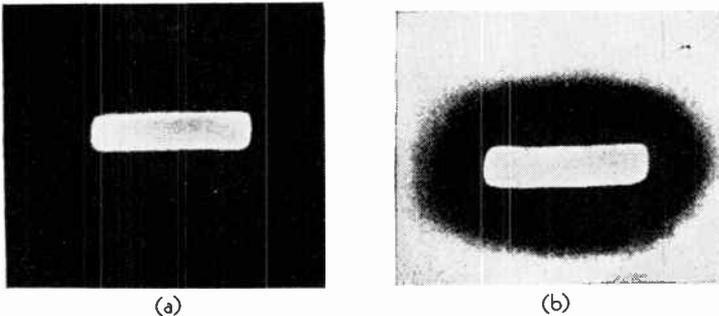


Fig. 6-38. Reproduction of a light spot at low and high spot brightness. (Courtesy of *Proc. IRE.*)

remains saturated and delivers a white signal. This effect is shown in Fig. 6-38, where the actual white spot on the black background is reproduced as a white spot surrounded by a black "halo." The surrounding background is white. The net effect is that contrast is preserved in the vicinity of the highlight. In a normal picture with fine detail the halo cannot be observed because of its small size; good contrast is maintained except in large dark areas.

If the mosaic illuminance is raised above C in Fig. 6-37 the characteristic becomes linear again. This change in mode of operation is the result of an effective increase in element capacitance that takes place when the illuminance is sufficient to give the entire target some charge during a small portion of the frame interval. This condition occurs at C where the capacitance per element exceeds the value given by the total target capacitance divided by the figure of merit. Once this larger capacitance comes into play, the output rises linearly until a saturation of charge determined by the new capacitance value sets in. This last condition occurs at D .

We see, therefore, that at low light levels the image orthicon is a linear device. At high levels of illuminance the contrast is maintained by the effect of secondary electrons released from the electron image side of the target and is substantially independent of the overall scene brightness. These characteristics combine to make the image orthicon satisfactory for operation over a wide range of illuminance and make it the most flexible of all the camera tubes covered in this chapter.

In contrast to the orthicon, the image orthicon requires shading. The need for this is caused by the low-velocity secondary electrons which, after being released from the target's front surface, return in a semirandom shower to that same surface. It has been determined that a single shading voltage, a saw tooth at line frequency, provides satisfactory compensation for the resulting dark spot. This compensating voltage is usually applied across a tap on the grid return resistance of one of the video amplifiers in the same manner as that described for the iconoscope.

Considerable care is required in operating the image orthicon tube to prevent damage to the target. Operating instructions for the tube are available from the manufacturer.³⁸

³⁸ For example, *Tips on the Use of Image Orthicons*, RCA Tube Department, 1949.

6-28. The Image Orthicon Preamplifier

It is interesting to contrast the design problems of the preamplifiers for the iconoscope and image orthicon camera tubes. In the former the output current from the tube is low; hence a relatively large coupling resistance is required to provide an output voltage amplitude which is adequate. To complicate the design problem, the shunt capacitance across R_o is inherently large, and the larger the value of R_o , the greater is the shunting effect of the capacitance on the high-frequency signal components. A compromise must therefore be made in the choice of R_o to balance the requirements of signal amplitude and high-frequency response. Since the limitation on the maximum value of R_o can be eased by decreasing the shunt capacitance, considerable effort is expended to reduce the capacitance by special techniques. At best, these reduce the capacitance to roughly $8 \mu\mu\text{f}$ and the compromise is effected by using an R_o of 300 kilohms. Even this compromise value leaves something to be desired in the output voltage amplitude, and high-frequency peaking is required in the video amplifier to make up for the loss in highs due to the compromise.

With the image orthicon the design problem is less severe, primarily because of two reasons: First, a much larger output current is available, and second, a comparatively low value of shunt capacitance, about $30 \mu\mu\text{f}$ ³⁹ is present without any reduction due to degeneration. The first factor allows a lower value of coupling resistance to be used for any given output voltage, which, in turn, eases the bandwidth requirements. The low value of capacitance, which includes the output capacitance of the last multiplier dynode, the capacitances of the shielded video cable, the circuit strays, and the input of the first preamplifier, is due primarily to the low output capacitance of the dynode as compared to that of the signal plate in the iconoscope.

A typical image orthicon preamplifier circuit is shown in Fig. 6-39.⁴⁰ The effective coupling resistance used is about one-tenth of that used with the iconoscope. The first two stages are designed to be flat out to about 8 megacycles and compensation for the loss in high-frequency components across the input circuit is afforded by the "high peaker,"

³⁹ Allen B. Du Mont Laboratories, Inc., Operating and Maintenance Manual—Du Mont Model TA-124-B Image Orthicon Chain, 1948.

⁴⁰ "Television Field Pick-up Equipment Instructions." RCA Engineering Products Department.

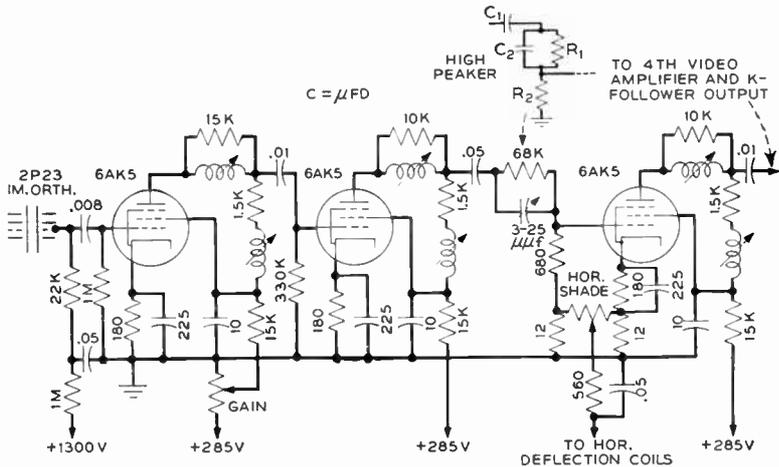


Fig. 6-39. Typical image orthicon preamplifier. The "high peaker" circuit is shown in the inset. (Courtesy of Radio Corporation of America.)

which is inserted between V_2 and V_3 . As may be seen from the inset, this network is a frequency-sensitive voltage divider. At high frequencies, C_2 shorts out R_1 and the full signal voltage appears across R_2 . At low frequencies, however, C_2 appears as an open circuit and only a small fraction of the signal is delivered to the grid of V_3 . It will be observed from the magnitudes of the circuit components that the peaker provides severe attenuation of the lows, a factor which provides the required high-frequency compensation and tends to eliminate low-frequency microphonic noise generated in the first two preamplifier stages.

The shading voltage, a saw tooth at line frequency, is derived from the horizontal deflection system and is fed to a potentiometer. The potentiometer and the two 12-ohm resistances form a bridge circuit such that the potentiometer may be used to control both the magnitude and polarity of the shading voltage which appears on the grid of the third tube.

CHAPTER 7

VIDEO AMPLIFICATION

Reference to Fig. 6-37 of the last chapter shows that an image orthicon camera tube operating under normal conditions will develop a voltage in the order of 0.1 mv across a 10,000-ohm load resistance. This voltage is much too small to drive the final kinescope, in fact, several of the common direct-view cathode-ray tubes require a voltage swing of some 60 or 70 v in order to reproduce the full gray scale from black to white. In order to make up for the discrepancy between the generated and required voltage levels, amplification must be introduced into the system. The need for amplification is further increased by the losses which are inherent in the communication link which interconnects the pickup and reproducing transducers; these tend to attenuate the signal below its value at the output of the camera tube.

The amplifiers required for raising the level of the video signal must be capable of handling signals extending over a frequency range of several megacycles. As compared to the audio amplifier, the video amplifier must be an extremely broad-band device. In the present chapter we shall first consider the analysis and design of video amplifiers based on their steady-state requirements. The relationship between the steady-state and transient responses will then be demonstrated.

7-1. Steady-state Requirements

The propagation characteristics of the general four-terminal network consisting of input terminals connected by a linear bilateral network to an output terminal pair and working between equal image impedances may be specified in terms of the complex propagation constant of the network γ , defined by

$$\epsilon^\gamma = \frac{E_i}{E_o} = \frac{E_i}{E_o} / \phi_i - \phi_o \quad (7-1)$$

where E_i and E_o are the complex input and output voltages, respectively. Where the network contains active, unilateral elements, such as in the video amplifier, the output voltage is of greater magnitude than the input voltage, and it becomes convenient to define a ratio which is the reciprocal of that used in (7-1). This ratio

$$A = \frac{E_o}{E_i} \quad (7-2)$$

is the complex amplification or gain of the network and may be used to describe the response of the network to a given input signal. Following the practice used in audio amplifiers we shall generally express A in polar form, thus,

$$A = A / \phi = \frac{E_o}{E_i} / \phi_o - \phi_i \quad (7-3)$$

$$\left. \begin{aligned} \text{hence} \quad A &= \frac{E_o}{E_i} \\ \text{and} \quad \phi &= \phi_o - \phi_i \end{aligned} \right\} (7-4)$$

The magnitude of A we shall call the amplification of the network; ϕ is the network phase shift and is defined as a negative quantity when the output voltage E_o lags the input, E_i .

If a number of stages of amplification are cascaded and each stage has a complex amplification

$$A_k = \frac{E_{ok}}{E_{ik}} \quad (7-5)$$

the over-all complex amplification will be

$$A_T = \frac{E_{on}}{E_{i1}} \quad (7-6)$$

In the system of cascaded amplifiers the output of any stage is the input of the following stage; hence we may expand the right-hand member of (7-6) as

$$\begin{aligned} A_T &= \frac{E_{o1}}{E_{i1}} \frac{E_{o2}}{E_{o1}} \frac{E_{o3}}{E_{i3}} \cdots \frac{E_{on}}{E_{in}} \\ &= A_1 A_2 A_3 \cdots A_n \end{aligned} \quad (7-7)$$

$$\left. \begin{aligned} \text{whence } A_T &= A_1 A_2 A_3 \cdots A_n \\ \text{and } \phi_T &= (\phi_{o1} - \phi_{i1}) + (\phi_{o2} - \phi_{i2}) + \cdots + (\phi_{on} - \phi_{in}) \\ &= \phi_{on} - \phi_{i1} \end{aligned} \right\} (7-8)$$

Let an input signal consisting of two or more components of different frequency be applied to the first amplifier in the cascaded chain. Then if the output signal is to be a distortionless reproduction of the input, the following conditions must be met:

1. The original frequency components and no additional components shall be present.
2. The relative amplitudes of the several components shall remain unchanged.
3. The time delay of the several components shall be constant.

From these three conditions we may set up the steady-state requirements of an ideal or distortionless amplifier. Condition 1 requires that the operation of the amplifier system be linear. *i.e.*, there shall be no amplitude or nonlinear distortion. In order to maintain relative amplitudes of the signal components unchanged, all components, regardless of their frequency, must be amplified by the same factor; therefore A must be independent of frequency over the entire range of signal frequencies. The last condition of constant delay time may be expressed in terms of ϕ , the total phase shift. The input voltage of a single component may be expressed as

$$e_i = E \sin(\omega t + \phi_i) \quad (7-9)$$

After passing through the amplifier, it will appear across the output as

$$e_o = AE \sin(\omega t + \phi_o) \quad (7-10)$$

or we may also write

$$e_o = AE \sin(\omega t + \phi_o + n\pi) \quad (7-10a)$$

where n is any positive or negative integer. That this is valid may be seen from the following considerations: if n is even, the value of the sine is unchanged; if n is odd, A becomes negative, indicating only a reversal of the output voltage. Equations (7-9) and (7-10a) are, of course, valid for any value of t ; hence we may define a new time variable t' , such that

$$\omega t' + \phi_i = \omega t + \phi_o + n\pi \quad (7-11)$$

and rewrite (7-10a) as

$$e_o = AE \sin(\omega t' + \phi_i) \quad (7-12)$$

We may also express the output voltage as the input voltage delayed by a time interval τ , thus

$$e_o = AE \sin [\omega(t + \tau) + \phi_i] \quad (7-13)$$

Then from (7-12) and (7-13) we see that

$$t' = t + \tau \quad (7-14)$$

Substitution of (7-14) into (7-11) finally yields

$$\phi = \phi_o - \phi_i = \omega\tau - n\pi = 2\pi\tau f - n\pi \quad (7-15)$$

The condition of constant delay time in the amplifier may therefore be expressed as follows: The phase shift of the amplifier plotted against frequency shall be a straight line passing through an integral multiple of π at zero frequency.

That these conditions give distortionless amplification may be verified from the following example. Let a signal voltage of two components, one of frequency f_a and phase ψ_a and the other of frequency f_b and phase ψ_b , be applied to the amplifier input. At any time t this voltage will be

$$e_i = E_a \sin (\omega_a t + \psi_a) + E_b \sin (\omega_b t + \psi_b) \quad (7-16)$$

and the corresponding output voltage will be

$$e_o = A_a E_a \sin (\omega_a t + \psi_a + \phi_a) + A_b E_b \sin (\omega_b t + \psi_b + \phi_b) \quad (7-17)$$

where $\left. \begin{array}{l} A_a = \text{amplification} \\ \phi_a = \text{phase shift} \end{array} \right\} \text{ at } f_a \quad \left. \begin{array}{l} A_b = \text{amplification} \\ \phi_b = \text{phase shift} \end{array} \right\} \text{ at } f_b$

Now if f_b is k times as great as f_a , or

$$f_b = k f_a \quad (7-18)$$

and if the amplifier satisfies the no-distortion criteria which have been stated,

$$A = A_a = A_b \quad (7-19)$$

and

$$\left. \begin{array}{l} \phi_a = 2\pi\tau f_a - n\pi \\ \phi_b = 2\pi\tau f_b - n\pi \\ = k2\pi\tau f_a - n\pi \end{array} \right\} \quad (7-20)$$

substitution of (7-19) and (7-20) shows that the output signal will be

$$e_o = AE_a \sin [(\omega_a t + 2\pi\tau f_a) + \psi_a - n\pi] \\ + AE_b \sin [k(\omega_a t + 2\pi\tau f_a) + \psi_b - n\pi] \quad (7-21)$$

Since this equation is true for any time t , we may define a new time variable t' such that

$$\omega_a t' = \omega_a t + 2\pi\tau f_a \tag{7-22}$$

and the output voltage becomes

$$e_o = \pm A [E_a \sin(\omega_a t' + \psi_a) + E_b \sin(\omega_b t' + \psi_b)] \tag{7-23}$$

the sign depending upon the value of n . With either sign the output voltage is an amplified version of the input; both components are increased by the same factor and their phases remain unchanged. We have shown, therefore, that the criteria do provide distortionless amplification of steady-state components.

Let us summarize the requirements for an ideal amplifier: A shall be independent of frequency in the pass band; ϕ shall be linear with frequency in the pass band with a phase intercept $n\pi$ or, alternatively, the time delay, τ , shall be constant in the pass band; and the amplifier shall be linear.

7-2. The Resistance-coupled Amplifier

Of all the common types of amplifiers used in audio work the resistance-coupled form may best be adapted to the broadband requirements of the video system. We shall briefly review the characteristics of this circuit which is shown in Fig. 7-1. The use of the equivalent plate circuit in analysis is justified for we have specified that the amplifier shall be linear in its operation.

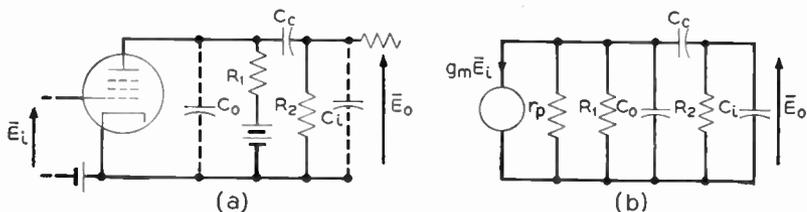


Fig. 7-1. The resistance-coupled amplifier. (a) The basic circuit. (b) Equivalent plate circuit.

The exact analysis of the circuit shows that some of the reactive circuit parameters have negligible effect over certain ranges of frequency and that the operation may best be considered in three separate ranges of frequency. Thus there is a certain mid-frequency band defined as that range of frequency for which C_c has negligible reactance and the shunting effect of C_i and C_o on R_1 and R_2 is negli-

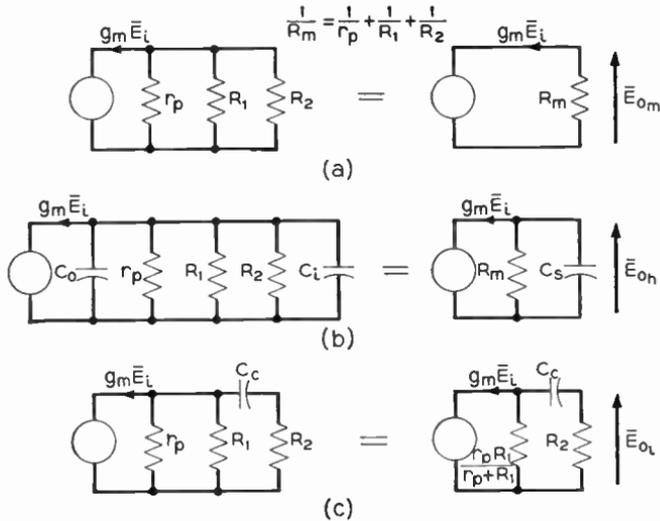


Fig. 7-2. The operation of the resistance-coupled amplifier may be considered in three separate frequency bands. (a) Equivalent mid-band circuit. (b) Equivalent high-band circuit. (c) Equivalent low-band circuit.

gible. In this mid-band we may use the equivalent circuit of Fig. 7-2a, and we have for the complex output voltage, which is *defined as a voltage drop from plate to ground*,¹

$$E_{om} = -g_m E_i R_m \quad (7-24)$$

where

$$\frac{1}{R_m} = \frac{1}{r_p} + \frac{1}{R_1} + \frac{1}{R_2}$$

$$A_m = \frac{E_{om}}{E_i} = -g_m R_m \quad (7-25)$$

Therefore

$$A = \frac{E_o}{E_i} = g_m R_m \quad \left. \vphantom{\frac{E_o}{E_i}} \right\} \text{MID-BAND} \quad (7-26)$$

and

$$\phi = 180^\circ \quad (7-27)$$

As the frequency increases above the mid-band range, the reactance of the three condensers decreases further, and the shunting effect of C_o and C_i is no longer negligible. Since C_c has negligible reactance,

¹ In the amplifier diagrams an arrow is used to indicate the positive direction of a voltage rise.

C_o and C_i are in parallel and may be lumped into an equivalent shunt capacitance C_s , equal to

$$C_s = C_o + C_i \tag{7-28}$$

The simplified equivalent circuit of Fig. 7-2b obtains and the output voltage is

$$E_{oh} = \frac{-g_m E_i}{\frac{1}{R_m} + j\omega C_s} = \frac{-E_i g_m R_m}{\sqrt{1 + (\omega C_s R_m)^2}} / -\text{arc tan } (\omega C_s R_m) \tag{7-29}$$

and the amplification in the high-band of frequencies, A_h , is

$$A_h = \frac{E_{oh}}{E_i} = \frac{-g_m R_m}{\sqrt{1 + (\omega C_s R_m)^2}} / -\text{arc tan } (\omega C_s R_m) \tag{7-30}$$

In the work which is to follow it is convenient to use the relative high-band amplification defined as the ratio of high- to mid-band amplification; thus from (7-25) and (7-30)

$$\frac{A_h}{A_m} = \frac{1}{\sqrt{1 + (\omega C_s R_m)^2}} / -\text{arc tan } (\omega C_s R_m) \tag{7-31}$$

In magnitude the relative amplification is

$$\frac{A_h}{A_m} = \frac{1}{\sqrt{1 + (\omega C_s R_m)^2}} \quad \text{HIGH-BAND} \tag{7-32}$$

and the *relative* high-band phase shift, θ_h , is

$$\theta_h = \phi_h - \phi_m = -\text{arc tan } (\omega C_s R_m) \quad \text{HIGH-BAND} \tag{7-33}$$

where the negative sign indicates that the high-band output lags the mid-band output.

(Generally in the calculation of response curves it is convenient to work in terms of a normalized frequency, the reference being the so-called half-power frequency. This may be forced into the last two equations in the following manner. If a constant-amplitude variable-frequency voltage be applied to the input of the resistance-coupled amplifier, (7-32) will also give the magnitude of the ratio of output voltage in the high-band to output voltage in the mid-band. Since power is proportional to the square of voltage we may write

$$\frac{P_h}{P_m} = \left(\frac{E_h}{E_m}\right)^2 = \left(\frac{A_h}{A_m}\right)^2 = \frac{1}{1 + (\omega C_s R_m)^2} \tag{7-34}$$

We now ask this question: Is there some frequency f_2 in the high-frequency band at which the amplifier delivers one-half of the power it delivers in the mid-band? To answer this, we make the power ratio $\frac{1}{2}$ in the last equation and there results

$$\frac{1}{2} = \frac{1}{1 + (\omega_2 C_s R_m)^2}$$

hence

$$\omega_2 C_s R_m = 1 \quad (7-35)$$

or

$$f_2 = \frac{1}{2\pi C_s R_m} \quad (7-36)$$

f_2 , defined as in eq. (7-36), is the upper half-power frequency. Substitution of (7-35) into (7-32) and (7-33) gives the relative response in terms of the frequency normalized with respect to f_2 .

$$\left. \begin{aligned} \frac{A_h}{A_m} &= \frac{1}{\sqrt{1 + \left(\frac{f}{f_2}\right)^2}} \end{aligned} \right\} \text{HIGH-BAND} \quad (7-37)$$

and

$$\theta_h = -\text{arc tan } \frac{f}{f_2} \quad (7-38)$$

It may be seen by inspection of (7-37) that the relative amplification at the half-power point is 0.707. Notice that the process of normal-

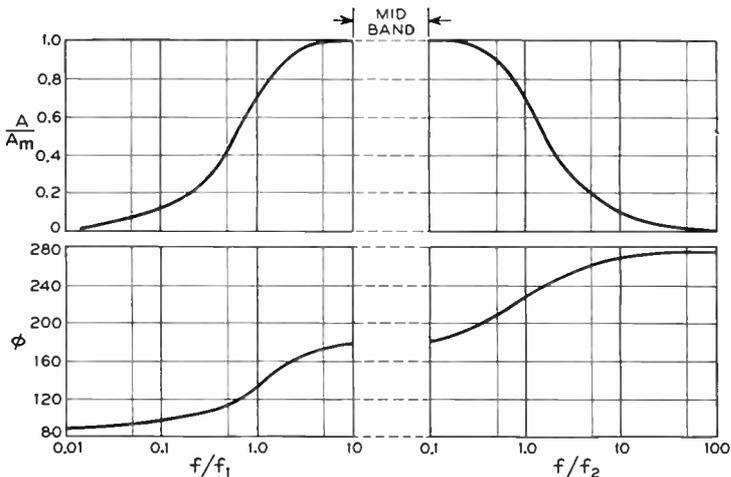


Fig. 7-3. Normalized response curves of the resistance-coupled amplifier. (a) Gain. (b) Phase shift.

ization makes the expressions for relative gain and relative phase shift independent of the specific circuit constants of the amplifier in use. These equations may be plotted to give universal response curves which apply to any resistance-coupled amplifier.² Inspection of these curves shows that to an excellent approximation *the upper limit of the mid-band of the amplifier lies at one-tenth of the upper half-power frequency.* The typical response in the high-band is shown in Fig. 7-3.

The low-frequency band is defined as that band of frequencies for which X_{C_c} is no longer negligible but the shunting effect of C_o and C_i on the plate load is unimportant. Subject to these restrictions the equivalent plate circuit of the amplifier for the low-frequency band is that shown in Fig. 7-2c. For the low-band output voltage we have

$$E_{ol} = -g_m E_i \frac{\frac{r_p R_1}{r_p + R_1}}{\left(R_2 + \frac{r_p R_1}{r_p + R_1}\right) - \frac{j}{\omega C_c}} R_2 \quad (7-39)$$

Then, factoring out the term in the parentheses from the denominator, we get

$$\begin{aligned} E_{ol} &= \frac{-g_m E_i \frac{r_p R_1 R_2}{r_p + R_1}}{\left(R_2 + \frac{r_p R_1}{r_p + R_1}\right) \left[1 - \frac{j}{\omega C_c \left(R_2 + \frac{r_p R_1}{r_p + R_1}\right)}\right]} \\ &= -E_i \frac{g_m R_m}{1 - \frac{j}{\omega C_c R_L}} \end{aligned} \quad (7-40)$$

where
$$R_L = R_2 + \frac{r_p R_1}{r_p + R_1} \quad (7-41)$$

It follows that the low-band complex amplification is

$$A_l = \frac{E_{ol}}{E_i} = - \frac{g_m R_m}{1 - \frac{j}{\omega C_c R_L}} \quad (7-42)$$

and for the relative amplification and relative phase shift we get

² See, for example, F. E. Terman, *Radio Engineers' Handbook*. New York: McGraw-Hill Book Company, Inc., 1943, p. 357.

$$\left. \frac{A_l}{A_m} = \frac{1}{\sqrt{1 + \left(\frac{1}{\omega C_c R_L}\right)^2}} \right\} \text{LOW-BAND} \quad (7-43)$$

$$\text{and} \quad \theta_l = \phi_l - \phi_m = +\text{arc tan} \left(\frac{1}{\omega C_c R_L} \right) \quad (7-44)$$

These expressions may also be normalized, this time with respect to the *lower* half-power frequency, f_1 . Defined in a manner analogous to f_2 , f_1 is that frequency at which the relative amplification falls to 0.707 and occurs when

$$\omega_1 C_c R_L = 1$$

$$\text{or} \quad f_1 = \frac{1}{2\pi C_c R_L} \quad (7-45)$$

Substitution of this identity into the previous low-band equations yields

$$\left. \frac{A_l}{A_m} = \frac{1}{\sqrt{1 + \left(\frac{f_1}{f}\right)^2}} \right\} \text{LOW-BAND} \quad (7-46)$$

$$\text{and} \quad \theta_l = \text{arc tan} \frac{f_1}{f} \quad (7-47)$$

Since θ_l is a positive quantity the low-band output leads the output in the mid-band. It may be shown from (7-46) that to an excellent approximation *the lower limit of the mid-band is 10 times the lower half-power frequency.*

A typical response curve for the resistance-coupled amplifier is shown in Fig. 7-3. It should be noticed that the mid-band is defined by the frequency range

$$10f_1 \leq f \leq 0.1f_2 \quad \text{MID-BAND} \quad (7-48)$$

Notice in the derivation above that the accent has been shifted from amplification and phase shift to their relative values, A/A_m and $\theta = \phi - \phi_m$. We therefore must convert the criteria for distortionless operation into terms of these relative quantities. Both conversions may be made by inspection. First, if A is to remain constant, A/A_m must remain at unity. Second, by definition,

$$\begin{aligned} \theta &= \phi - \pi \\ &= (2\pi\tau f - n\pi) - \pi = 2\pi\tau f - (n+1)\pi \end{aligned} \quad (7-49)$$

Then since n is any integer, positive or negative, the same criterion for no distortion applies to θ as to ϕ .

In summary we may say that the requirements for no distortion in a linear amplifier are that relative amplification shall remain constant at unity, and that relative phase shift shall vary linearly with frequency in the pass band of the amplifier. The latter condition does not preclude θ remaining zero.

In order to satisfy the criteria stated in the last paragraph a resistance-coupled amplifier must be designed in such a manner that the entire range of frequencies which comprise the video signal lie within the mid-band range of the amplifier. If such a design be carried out, the amplification will be so low that the stage is of little use in raising the signal level. We may demonstrate this by an example which is typical. Say the bandwidth of the signal is such that the amplifier must be flat to 4.5 mc. By (7-48) we see that the amplifier must be designed so that its upper half-power frequency is $10 \times 4.5 = 45$ mc. From eq. (7-35) we see that f_2 , the upper half-power frequency, is that frequency at which the reactance of C_s and R_m are equal. For a typical amplifier employing a 6AC7 tube C_s , the sum of input and output capacitances, will average around $20 \mu\mu\text{f}$, thus (7-35) requires a value of R_m

$$R_m = \frac{1}{2\pi(4.5 \times 10^7)(2 \times 10^{-11})} = 173 \text{ ohms} \quad (7-50)$$

With a transconductance of 9000 micromhos, the 6AC7 will have a corresponding mid-band amplification of

$$A_m = g_m R_m = (9 \times 10^{-3})(1.73 \times 10^2) = 1.56 \quad (7-51)$$

This value of amplification is too low to be of value.

To counteract the conflicting requirements of bandwidth and amplification, a compensating procedure is applied in the design of the amplifier. Since the decrease in high- and low-band amplification is the result of different circuit elements, we may consider the compensation of the two bands separately. This will be done after we have seen that the transient response of an amplifier is related to its steady-state characteristics.

7-3. Transient Response of a Video Amplifier

To this point in the present chapter we have been concerned with the steady-state response of the basic resistance-coupled amplifier

circuit. The general point of view in the steady-state case is this: if a sinusoidal signal of known amplitude, frequency, and phase is applied to the input terminals of the amplifier, what will be the amplitude and phase of the corresponding signal which appears across the output terminals? An investigation is then made of how these output quantities change as the frequency of the applied signal is varied. When this information is known, we have the steady-state response of the network, that is, we know how the amplifier responds to a *sinusoidal* signal of any frequency. Since Fourier's theorem gives us a means of relating any repetitive nonsinusoidal signal to its harmonically related sinusoidal components, we can then find the response of the amplifier to such a signal, by determining the output for each of the sinusoidal components. Then, in accordance with the superposition theorem, the total output signal will be the sum of the individual sinusoidal output components.

In television work, however, we deal almost exclusively with signals which are nonrepetitive in nature. The Fourier analysis does not obtain in this case and we are faced with the following problem: What sort of response characteristics can we use in place of the steady-state which will give us the required information for these non-repetitive input signals?

In the final analysis the measure of distortion in an amplifier is the answer to the question: To what extent is the shape of the output signal an exact, though amplified and delayed, reproduction of the shape of the input signal? Now if we speak of the shape or wave form of a signal, we must inevitably bring in a time variable, and we may describe the signal by plotting its instantaneous amplitude as a function of time. This concept is well known and in fact has been used several times in this book where a voltage or current has been plotted against time to illustrate its shape. It follows, then, that we must concern ourselves with the time response of an amplifier, that is, the shape of its output signal relative to a given input signal. We have two principal questions to answer: (1) What input signal shall be used to give the information we need about the reproduction of non-repetitive wave forms, and (2) How shall we calculate the amplifier response to that input signal? When these questions are answered, we shall discover that the time or transient response is related to the steady-state characteristics of the amplifier.

A number of different mathematical approaches to the calculation

of transient response are available^{3,4} in the literature and the recent trend has been toward almost exclusive use of the Laplace transform.⁵ Since our purpose is primarily the study of television systems, however, it is beyond the scope of our work to develop the mathematical tools needed for these methods; hence we shall approach the problem from a physical basis to gain insight of the circuit behavior, and state the mathematical results where required and reference them for the benefit of the interested student.

7-4. Unit Function

In order to find the transient response of a network we must first find a suitable test signal to apply at the input terminals. From an even elementary study of electrical transient phenomena we know that a common test voltage is that furnished by a battery of constant voltage which is suddenly applied to the network at $t = 0$. The resulting applied voltage has the familiar form plotted in Fig. 7-4

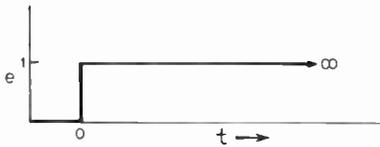


Fig. 7-4. Unit function, a basic test voltage.

and is known as a step function. If the amplitude of the step is unity, the form is known as “unit function” or “unit step.” The reason for the choice of this form of test signal for the calculation of transient response may be demonstrated mathematically by means of the superposition integral or Duhamel’s integral as it is sometimes known. The basic line of reasoning used is that any nonrepetitive wave form may be considered as the superposition of a number of unit functions suitably weighted in amplitude and time as illustrated

³ H. A. Wheeler, “The Interpretation of Amplitude and Phase Distortion in Terms of Paired Echoes.” *Proc. IRE*, **27**, 4 (April 1939). See also S. Goldman, *infra*, chap. 12.

⁴ M. J. DiToro, “Phase and Amplitude Distortion in Linear Networks.” *Proc. IRE*, **36**, 1 (January 1948).

⁵ S. Goldman, *Transformation Calculus and Electrical Transients*. New York: Prentice-Hall, Inc., 1949.

in Fig. 7-5. Then, if the network response to unit step is known, the response to the actual wave form may be obtained as the superposition of the response to each of the weighted steps which synthesize the wave.⁶ We should note that this is a direct analogue of the steady-state problem where the response to a sine wave is used to determine the response of a repetitive wave which may be synthesized from suitably weighted sinusoidal components. For our purpose it is

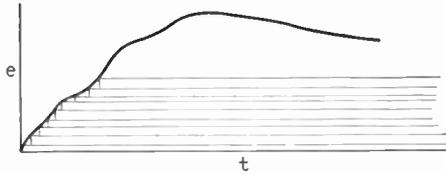


Fig. 7-5. A nonrepetitive wave form may be considered as the superposition of step functions properly weighted in amplitude and time. If the response of a network to unit step is known, the total response is the superposition of all the individual responses to each of the steps.

appropriate to justify the use of a step function as a test signal on physical grounds. To do this we consider two extremes of wave form which are encountered in television work. Assume that a large area of gray appears in the televised image. Then in this case the wave form of the signal is constant at some d-c value in the region of the gray signal. In the extreme case, where the whole picture is at the same gray level, the duration of that d-c level of the corresponding signal will last for at least one frame interval. Thus we may state that one requirement on a video amplifier is that it be able to maintain a constant d-c level over a relatively long interval, that is, long in terms of a line duration.

Further, let it be assumed that in one region of the televised image an abrupt transition from black to white occurs. This once again represents an extreme condition; if the amplifier can reproduce this abrupt transition, it certainly can reproduce any gradual one. It follows, then, that if a video amplifier can meet these two requirements—reproduction of an abrupt transition and maintenance of a constant level over a comparatively long interval—it should be able

⁶ See, for example, S. Goldman, *op. cit.*, chap. 4, or L. B. Arguimbau, *Vacuum Tube Circuits*. New York: John Wiley and Sons, Inc., 1948, chap. 4.

to reproduce any other conditions which occur in a video signal. A moment's reflection will show that these two conditions are met exactly by step function; hence step function is an ideal test wave form for video amplifiers and it is the basic test signal used for calculating transient response and for determining transient response experimentally.

7-5. Transient Response of the R-C Amplifier

With our test signal decided upon we may now calculate the transient response of the basic resistance-coupled amplifier whose circuit diagram is shown in Fig. 7-1. It is fortunate that the transient response of this circuit is amenable to solution by simple differential equations. To simplify our work, we shall assume the plate resistance of the amplifier tube to be much larger than the resistance R_1 . Actually further simplification of the circuit can be made. To illustrate this let us consider the physical behavior of the circuit when a step-function grid voltage is applied. At $t = 0$ the current $g_m e_i$ begins to flow. Now since the voltage across a condenser cannot change abruptly, initially $g_m e_i$ flows into the condensers, charging them. As the voltage across the condensers increases, some of the current begins to flow through the resistors. Furthermore, since $C_e \gg C_i$, most of the voltage drop across them will appear across C_i ; thus in the region of the step in the applied wave, we may neglect C_e and consider it to be shorted out. Further, in the typical amplifier $R_2 \gg R_1$, so that in the vicinity of the step, a much smaller current flows through R_2 than through R_1 . It follows at once, then, that as far as behavior near the jump in an applied step voltage is concerned, the equivalent circuit of the amplifier is identical to that shown for the high-frequency-band steady-state response in Fig. 7-2b, but with $R_m = R_1$. Remember that with C_e shorted out, C_o and C_i are in parallel and their sum is defined as C_s . We may now solve for the instantaneous output voltage, e_o , by using Kirehhoff's current law.

$$\left. \begin{array}{l} \text{Let} \quad i_1 = \text{instantaneous current in } R_1 \\ \quad \quad i_s = \text{instantaneous current in } C_s \end{array} \right\} \quad (7-52)$$

$$\text{Then} \quad i_1 + i_s = g_m e_i \quad (7-53)$$

$$\text{or} \quad \frac{e_o}{R_1} + C_s \frac{de_o}{dt} = g_m e_i \quad (7-54)$$

By collecting terms and separating the variables we get

$$\frac{-de_o}{g_m e_i R_1 - e_o} = -\frac{dt}{R_1 C_s} \quad (7-55)$$

and integrating

$$\ln (g_m e_i R_1 - e_o) = -\frac{t}{R_1 C_s} + \ln K \quad (7-56)$$

K being the constant of integration which may be evaluated for

at

$$t = 0, \quad e_o = 0$$

thus

$$\ln K = \ln g_m e_i R_1 \quad (7-57)$$

Then if (7-57) is substituted into (7-56) and antilogs are taken, there results

$$e_o = g_m e_i R_1 (1 - \epsilon^{-t/R_1 C_s}) \quad (7-58)$$

which is the response of the R - C amplifier to the discontinuity in an applied step function. We see, then, that the output voltage response to the discontinuity in the applied step voltage is of a familiar exponential form which is replotted in Fig. 7-6a for convenience. From this curve we observe that the time constant $R_1 C_s$ must be small if the rise from zero to full response is to occur in a short interval of time.

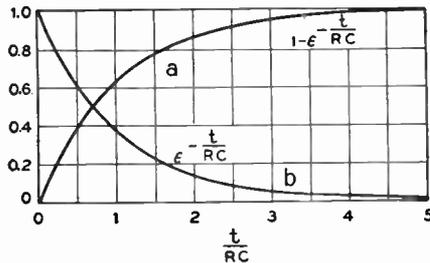


Fig. 7-6. The basic exponential curves. (a) $1 - \epsilon^{-t/RC}$. (b) $\epsilon^{-t/RC}$.

We shall investigate this condition more carefully later in the section. We may summarize our results so far by noting that: (1) The response to the discontinuity is exponential. (2) High gain demands a large value of R_1 whereas short rise time requires a small value of R_1 ; therefore high gain and short rise time are incompatible. (3) The response to the discontinuity depends upon R_1 and C_s , which are also

the determining factors in the high-band steady-state response; on this basis we might expect that the transient response to a discontinuity in applied voltage is related to the high-frequency steady-state characteristics of the amplifier.

Let us now consider the behavior of the circuit after some finite interval has elapsed after the application of the step voltage. When C_i has become almost fully charged, the voltage across the coupling condenser C_c must be taken into account. When this condition obtains we may effectively neglect C_s and assume that a constant voltage $g_m e_i R_1$ is applied to C_c and R_2 . Our previous studies have shown that the output voltage would be

$$\begin{aligned} e_o &= \text{voltage across } R_2 \\ &= g_m e_i R_1 \epsilon^{-t/R_2 C_c} \end{aligned} \quad (7-59)$$

which is the response of the R - C amplifier to the flat portion of an applied step voltage. From this equation which is replotted for convenience in Fig. 7-6b we see that the time constant $R_2 C_c$ must be very large, in fact infinite, if the output voltage is to remain constant at a fixed d-c level. In summary, we note that after a finite interval after application of the step voltage: (1) The transient response decays exponentially. (2) The time constant $R_2 C_c$ should be high. (3) The response depends upon $R_2 C_c$, which is also the determining factor in the low-band steady-state response; hence we might expect that the transient response to the flat-topped portion of step function is related to the low-frequency steady-state characteristics of the amplifier. Furthermore our analysis shows that we are justified in handling the response to the discontinuous and flat-topped portions of step function as two separate problems.

Let us apply these results to calculate the transient response of a resistance-coupled amplifier for the duration of one active scanning line interval, which under commercial telecasting standards is approximately 51 μ sec. Let the circuit constants of the amplifier be

$$\begin{aligned} g_m &= 9000 \text{ micromho} & R_2 &= 100 \text{ kilohms} \\ R_1 &= 5000 \text{ ohms} & C_c &= 0.1 \text{ } \mu\text{f} \\ C_s &= 20 \text{ } \mu\mu\text{f} \end{aligned}$$

Then the mid-band gain of the amplifier will be

$$A = g_m R_1 = (9 \times 10^{-3})(5 \times 10^3) = 45$$

$$R_1 C_s = (5 \times 10^3)(2 \times 10^{-11}) = 0.1 \mu\text{sec}$$

and

$$R_2 C_c = (10^5)(10^{-7}) = 10^{-2} \text{ sec}$$

The response to the discontinuity is shown in Fig. 7-7a. From this curve the 10 to 90 per cent rise time is seen to be $2.2R_1 C_s = 0.22 \mu\text{sec}$. Since the $R_2 C_c$ time constant is long in comparison to the active

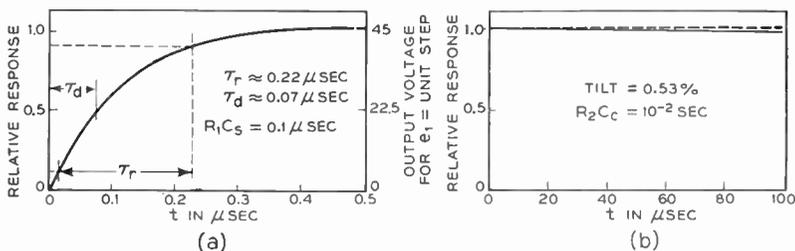


Fig. 7-7. Transient response of the resistance-coupled amplifier.
(a) In the vicinity of the discontinuity in an applied step function.
(b) After the full response has been reached.

line interval, we may assume that the exponential decay after the build-up has occurred may be replaced by a linear decay which has a slope equal to the initial slope of the exponential, that is,

$$\begin{aligned} \left. \frac{de_o}{dt} \right]_{\text{initial}} &= \left. \frac{d}{dt} \epsilon^{-t/R_2 C_c} \right]_{t=0} = -\frac{1}{R_2 C_c} \epsilon^{-t/R_2 C_c} \Big|_{t=0} \\ &= -\frac{1}{R_2 C_c} = 10^{-2} \text{ v/sec} \end{aligned}$$

This result may be stated in a different way. Assuming e_o to decay at a constant rate during the line interval we may write

$$e_o = g_m e_i R_1 \epsilon^{-t/R_2 C_c} \approx g_m e_i R_1 \left(1 - \frac{t}{R_2 C_c} \right)$$

and the "tilt" of the output wave will be

$$\text{tilt} = \frac{e_o]_{51} - e_o]_0}{e_o]_0} = \frac{t}{R_2 C_c} = \frac{53 \times 10^{-6}}{10^{-2}} = 0.53 \text{ per cent}$$

These results are plotted in Fig. 7-7b. Whereas they appear quite satisfactory for a single stage, if a number of these stages are cascaded, the rise time and decay will both increase and the specification

of the transient requirements becomes quite difficult. We shall consider this problem later in the discussion.

Thus far we have seen in a qualitative manner that the transient and steady-state responses are interrelated. We next consider an analytical method of expressing this interrelationship, the Fourier integral.

7-6. The Fourier Integral

It is beyond the scope of our work to derive the Fourier integral. We shall, however, briefly describe some of its properties and state the mathematical results which are of aid in our discussion of video amplifiers. The Fourier integral is a generalization of the Fourier series, which permits a nonrepetitive pulse, such as unit function, to be expressed as an integral (summation) of an infinite number of sinusoidal frequency components, each of infinitesimal amplitude and spread continuously over the frequency spectrum from zero to infinite frequency.⁷ If unit function is applied to a network having the steady-state characteristics $A(\omega)$ and $\phi(\omega)$, the method of the Fourier integral states that the response of the network will be

$$e_o(t) = \frac{A(0)}{2} + \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{A(\omega) \sin [\omega t + \phi(\omega)]}{\omega} d\omega \quad \begin{array}{l} \text{RESPONSE} \\ \text{TO UNIT} \\ \text{FUNCTION} \end{array} \quad (7-60)$$

Equation (7-60) shows the direct relationship between the transient and steady-state responses of a network which we anticipated in the last section. It further shows one of the chief difficulties of the Fourier integral approach: if mathematical evaluation is to be used, the steady-state response characteristics of the network must be stated analytically. This means, for example, that the curves of Fig. 7-3 would have to be stated in equation form and the resulting integration becomes quite difficult. Alternatively some form of graphical integration may be used.

The use of the Fourier integral may be demonstrated by calculating the transient response of an amplifier which has ideal characteristics from a steady-state point of view. We choose this example because the integration is straightforward and yields some interesting results

⁷ This should be contrasted to the Fourier series where the components have finite amplitude and discrete frequencies which are integral multiples of the fundamental repetition rate.

of general applicability. The ideal low-pass amplifier has a uniform amplitude response, A , up to a cutoff frequency, f_c , and zero for all frequencies higher than f_c . Its phase characteristic is linear in the pass band and given by $\phi = \omega\tau_d$, τ_d being a constant. These steady-state responses are plotted in Fig. 7-8a. The significance of negative

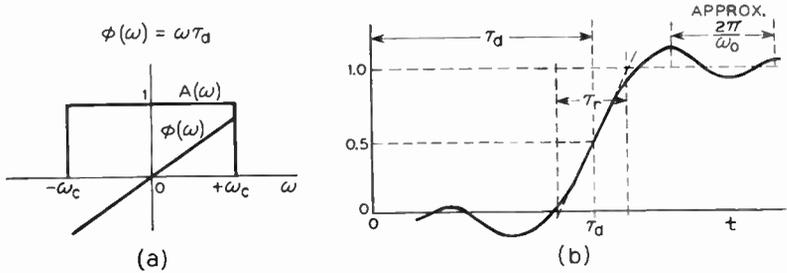


Fig. 7-8. The response characteristics of an "ideal" amplifier. (a) Steady-state response. (b) Transient response.

frequency is that the rotating vector, which represents the negative frequency component, rotates in a clockwise, or negative, direction.

We now substitute these responses into eq. (7-60). Notice that $A(\omega)$ is zero everywhere outside the pass band so that the contribution to the integral is zero at all frequencies $|\omega| > |\omega_c|$; hence we have

$$e_o(t) = \frac{A}{2} + \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} \frac{A \sin(\omega t - \omega\tau_d)}{\omega} d\omega \quad (7-61)$$

Then, multiplying numerator and denominator of the integrand by $(t - \tau_d)$, we get

$$\left. \begin{aligned} e_o(t) &= \frac{A}{2} + \frac{A}{2\pi} \int_{-\omega_c}^{\omega_c} \frac{\sin \omega(t - \tau_d)}{\omega(t - \tau_d)} d[\omega(t - \tau_d)] \\ &= \frac{A}{2} + \frac{A}{\pi} \int_0^{\omega_c(t - \tau_d)} \frac{\sin u}{u} du = \frac{A}{2} + \frac{A}{\pi} Si(x) \end{aligned} \right\} \quad (7-62)$$

where $u = \omega(t - \tau_d)$ and $x = \omega_c(t - \tau_d)$

The definite integral $Si(x)$ of (7-62) is well known and its evaluation as a function of x , the upper limit, is available in the literature.⁵ The result is plotted as a function of time in Fig. 7-8b.

⁵ See, for example, F. E. Terman, *op. cit.*, p. 16.

We may note a number of significant points from this response curve. First, the transient response of an "ideal" amplifier is not ideal at all; it exhibits "overshoot" and "ring," that is, it rises above and then oscillates about the 100 per cent response level, and displays a finite rise time, τ_r . This fact shows that extreme care is required in setting up the requirements for a video amplifier in terms of steady-state responses alone and further shows why a study of transient response is so necessary. Second, an amplifier having "ideal" steady-state characteristics is a mathematical fiction which cannot be realized physically. This follows because the response curve of Fig. 7-8b shows a response for $t < 0$, indicating that the network responds before the test signal is applied! These facts are of far-reaching consequence, in fact, it may be shown that $A(\omega)$ and $\phi(\omega)$ are interrelated⁹ so that only one but not both of these characteristics may be set up arbitrarily, as was done in Fig. 7-8a. This may be verified in a qualitative fashion for the resistance-coupled amplifier, for eqs. (7-32) and (7-33) show that high-band amplification and phase response are both functions of the same quantity, $\omega C_s R_1$, and hence must be interrelated. Third, τ_d , the delay time, or the interval by which the 50 per cent response point lags the time of application of the step test signal, is equal to the slope of the steady-state phase response curve. Fourth, if the linear region of the build-up portion of the curve be extended to intersect the 0 and 100 per cent response levels, τ_r , the rise or build-up time may be defined as the time interval between these two intersections.¹⁰ This interval may be related to the steady-state cutoff frequency because at any time t the slope of the output voltage may be obtained by differentiating (7-62), thus

$$\frac{de_o(t)}{dt} = \frac{A}{\pi} \frac{\sin x}{x} \omega_c \quad (7-63)$$

But the linear portion of the curve occurs at $t = \tau_d$, where $x = 0$ and $(\sin x)/x = 1$. Thus the slope of the linear portion is

$$\left. \frac{de_o(t)}{dt} \right]_{t=\tau_d} = \frac{A\omega_c}{\pi} = A2f_c \quad (7-64)$$

⁹ See, for example, S. Goldman, *op. cit.*, p. 128.

¹⁰ It should be noted that a more common definition of rise time is that interval required for the response to change from 10 to 90 per cent of its final value. This definition leads to more exact results in practice because the 10 and 90 per cent points are clearly defined. The definition used in this section is chosen for mathematical convenience.

and from the figure

$$\tau_r = \frac{A}{A2f_c} = \frac{1}{2f_c} \quad (7-65)$$

This last equation is a statement of Küpfmüller's rule that in a network having the ideal low-pass characteristics shown in Fig. 7-8a the rise time is one-half the reciprocal of the steady-state cutoff frequency.

Statements 3 and 4 show again that the transient and steady-state responses are interrelated, and that given f_c and $\phi(\omega)$ we may calculate the rise and delay times in the transient response, at least for this ideal case.

One other important general observation may be derived from the preceding results. Notice that if f_c is raised the rise time decreases. Furthermore for the ideal amplifier if f_c is raised the slope of the $\phi(\omega)$ curve will decrease with a corresponding decrease in delay time. It would seem that a good transient response would result if the steady-state bandwidth were infinite. Obviously this condition cannot be met in practice.

It must be stressed that the quantitative results which have been derived apply only to the "ideal" amplifier which was chosen because its $A(\omega)$ and $\phi(\omega)$ characteristics may be integrated easily. Hence some care must be exercised in extending the results to amplifiers which may be realized in practice. Moskowitz and Racker¹¹ have suggested a handy rule-of-thumb extension of statement 4, which relates rise time to cutoff frequency. It is a physical fact that pulses encountered in practice do not have abrupt transitions or discontinuities but actually have finite rise times. For example, the horizontal synchronizing pulses used in commercial television have a rise time of approximately 0.25 μ sec. The rule states that where the applied pulse has such a finite rise time, say τ_p , the transient response of an amplifier having a sharp cutoff characteristic will be satisfactory if τ_r for the amplifier is less than τ_p of the pulse. In other words, the cutoff frequency should be at least

$$f_c = \frac{1}{2\tau_p} \quad (7-66)$$

Where the amplifier has a gradual cutoff characteristic, a lower cutoff frequency may be tolerated, or

¹¹ S. Moskowitz and J. Racker, "Pulse Amplifier Design." *Radio-Electronic Engineering Edition of Radio and Television News*, 10, 2 (February 1948).

$$f_c = \frac{1}{3\tau_p} \quad (7-67)$$

7-7. Cascaded Video Amplifiers

In the interests of the work to follow we must consider what effect the cascading of n identical amplifiers has on the transient response. Palmer and Mautner¹² have shown that if the overshoot,¹³ γ , per stage is less than or equal to 2 per cent, the decrease in bandwidth of the n stages is the factor $1/\sqrt{n}$, or in other words, the rise time for the n stages is \sqrt{n} times that for a single stage. It may also be shown that γ increases with n , and that if the limit per stage is held to 2 per cent or less, the increase in γ due to cascading will not give unsatisfactory results.¹⁴ This will be demonstrated by curves later in the chapter.

These data allow us to specify roughly the requirements for each stage in a cascaded chain of video amplifiers: The overshoot, γ , shall be as small as possible and not exceed 2 per cent; and the rise time, τ_r , shall be $1/\sqrt{n}$ times, or less than, the maximum allowable rise time of the whole system. Since rise time is related to cutoff frequency, this is tantamount to putting a lower limit on the cutoff frequency. In more advanced texts it is shown that the tendency of a transient response to ring, that is, to overshoot and oscillate about the 100 per cent response value, is related to steepness in the $A(\omega)$ response characteristic. The steeper the cutoff, the greater is the ring characteristic; hence most video amplifier designs are based on a more or less gradual cutoff so that a less strict requirement is placed on the cutoff frequency.

From our study of the resistance-coupled amplifier the following facts are evident:

¹² R. C. Palmer and L. Mautner, "A New Figure of Merit for the Transient Response of Video Amplifiers." *Proc. IRE*, **37**, 9 (September 1949). See also S. Moskowitz and J. Racker, *op. cit.*

¹³ See Fig. 7-8b. γ is the amount by which the response goes above the 100 per cent level, expressed as a percentage of the full response.

¹⁴ G. E. Valley, Jr., and H. Wallman, *Vacuum Tube Amplifiers*, M.I.T. Radiation Laboratory Series No. 18. New York: McGraw-Hill Book Company, Inc., 1948. Where the per stage overshoot is from 5 to 10 per cent, γ increases approximately by the factor \sqrt{n} . The more exact expression for rise time is

$$\tau_r = \sqrt{\tau_{r1}^2 + \tau_{r2}^2 + \cdots + \tau_{rn}^2}$$

where τ_{rk} is the rise time of the k th stage.

Let $(\Delta f)_T$ = bandwidth of n cascaded identical stages,
 A_T = amplification of n cascaded identical stages,
 Δf = bandwidth per stage,
 A = amplification per stage.

Then if γ is 2 per cent or less

$$(\Delta f)_T = \frac{\Delta f}{\sqrt{n}} \quad (7-68)$$

and $A_T = A^n \quad (7-69)$

The gain-bandwidth product of the cascaded series will therefore be

$$A_T(\Delta f)_T = A^n \frac{\Delta f}{\sqrt{n}} \quad (7-70)$$

It is clear, then, that gain may be traded for bandwidth or vice versa, with n a controlling factor. Ideally the gain-bandwidth product per stage should be as high as possible to allow a minimum number of stages to be used. For the single-stage resistance-coupled amplifier, this product is

$$\begin{aligned} \Delta f &= \frac{1}{2\pi C_s R_1} \\ A &= g_m R_1 \\ A\Delta f &= \frac{g_m}{2\pi C_s} \end{aligned} \quad (7-71)$$

a function primarily of the tube. The quantity $g_m/2\pi C_s$ is often specified as the figure of merit for a video amplifier tube. Recognition of the fact that this quantity is the gain-bandwidth product for an amplifier which has a two-terminal coupling network has led to the development of high- g_m low- C_s tubes, such as the 6AC7.

HIGH-FREQUENCY COMPENSATION¹⁵

Our study so far has shown that the resistance-coupled amplifier affords a poor compromise between high gain and bandwidth, and does not exhibit an ideal transient response. We shall now consider a number of means for compensating these effects. Our general procedure will be to analyze the various compensating circuits on a

¹⁵ S. W. Seeley and C. N. Kimball, "Analysis and Design of Video Amplifiers," Parts I and II. *RCA Review*, II, 2 (October 1937) and III, 3 (January 1939).

steady-state basis and then to observe their effects on the transient response of the amplifier.

The general basis for the several methods of high-frequency compensation to be described is this: In the high-frequency band the change in $A(\omega)$ and the departure from linearity of $\phi(\omega)$ are the result of the shunting effect of C_s on R_1 , that is, of the decrease in plate load impedance at higher frequencies. To compensate for this effect we may insert inductance into the plate load.¹⁶ A number of interstage coupling networks which use compensating inductance will be considered.

7-8. Shunt Compensation

The decrease in plate load impedance at high frequencies may be compensated for by adding an inductance of proper value in *series* with R_1 , the plate load resistance. The resulting circuit, which is shown in Fig. 7-9a, is known as the shunt-compensated circuit. For

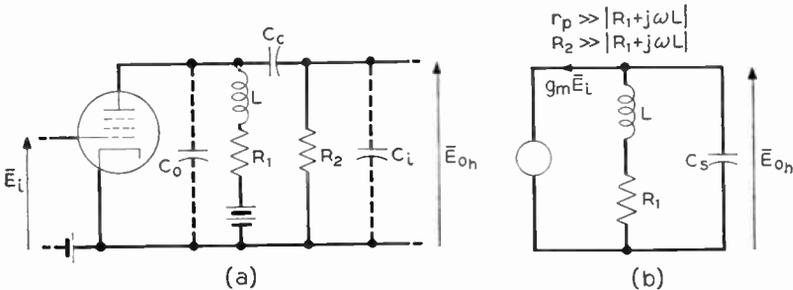


Fig. 7-9. The shunt-compensated video amplifier. (a) Basic circuit. (b) Simplified high-frequency-band equivalent circuit.

the analysis which is to follow we may make some simplifications in the basic circuit diagram. First, we shall be concerned with the response of the amplifier in the high-frequency band; hence C_c will have negligible reactance and may be omitted from the diagram. Second, omission of C_c places C_o and C_i in parallel; these may be lumped into the single capacitance, C_s , defined by eq. (7-28). Third, the values of plate resistance and grid leak resistance, R_2 , will be

¹⁶ High-frequency compensation may also be obtained by adding inductance in the cathode return. See A. B. Bereskin, "Cathode-Compensated Video Amplification," Parts I and II. *Electronics*, **22**, 6 (June 1948) and **22**, 7 (July 1948).

much larger than that of R_1 in typical circuits. The shunting effect of r_p and R_2 on the R_1 - L branch of the plate circuit will be negligible and r_p and R_2 may be omitted from the diagram. Under these simplifying conditions the equivalent high-band plate circuit is that shown in Fig. 7-9b.

Reading directly from the diagram we see that in the high-band

$$\mathbf{E}_{oh} = -g_m \mathbf{E}_i \mathbf{Z}_h \quad (7-72)$$

or

$$\mathbf{A}_h = -g_m \mathbf{Z}_h \quad (7-73)$$

where \mathbf{Z}_h is the complex load impedance in the high-frequency band, which may be expressed by

$$\mathbf{Z}_h = \frac{-jX_c(R_1 + jX_L)}{R_1 + j(X_L - X_c)} = \frac{R_1 \left(1 + \frac{jX_L}{R_1}\right)}{\left(1 - \frac{X_L}{X_c}\right) + j \frac{R_1}{X_c}} \quad (7-74)$$

Then, substituting for X_L and X_c in terms of ω , L , and C_s , we get

$$\mathbf{Z}_h = \frac{R_1 \left(1 + j \frac{\omega L}{R_1}\right)}{(1 - \omega^2 L C_s) + j \omega C_s R_1} \quad (7-75)$$

It is convenient to express \mathbf{Z}_h in terms of the upper half-power frequency, f_2 , of the amplifier *in the absence of the compensating inductance*, L . Inspection of the equivalent circuit of Fig. 7-9b and eq. (7-35) shows this uncompensated upper half-power frequency to be

$$\omega_2 = \frac{1}{C_s R_1} \quad (7-76)$$

It is of further convenience to express L in terms of a design parameter, K , and f_2 . This may be done by defining

$$K = \frac{\omega_2 L}{R_1} \quad (7-77)$$

whence

$$\frac{L}{R_1} = \frac{K}{\omega_2} \quad (7-78)$$

Substitution of (7-76) and (7-78) into (7-75) yields

$$\mathbf{Z}_h = \frac{R_1 \left[1 + jK \left(\frac{\omega}{\omega_2}\right)\right]}{\left[1 - K \left(\frac{\omega}{\omega_2}\right)^2\right] + j \left(\frac{\omega}{\omega_2}\right)} \quad (7-79)$$

The complex, high-band amplification then becomes

$$A_h = - \frac{g_m R_1 \left[1 + jK \left(\frac{f}{f_2} \right) \right]}{\left[1 - K \left(\frac{f}{f_2} \right)^2 \right] + j \left(\frac{f}{f_2} \right)} \quad (7-80)$$

This expression gives A_h in terms of the normalized frequency, f/f_2 . For simplicity we may use the identity

$$y = \frac{f}{f_2} \quad (7-81)$$

In common with the resistance-coupled amplifier, the compensated amplifier may have its complex amplification expressed on a relative basis, the mid-band response being the reference. We shall define the mid-band as that range of frequencies for which the reactance of C_c is negligible, the reactance of L is negligible in comparison to R_1 , and C_s has negligible shunting effect. The student should notice that this definition is consistent with the definition given in eq. (7-25). Then, from the diagram,

$$A_m = -g_m R_1 \quad (7-82)$$

and the relative complex high-band amplification becomes

$$\frac{A_h}{A_m} = \frac{1 + jKy}{(1 - Ky^2) + jy} \quad (7-83)$$

It is convenient to reduce this complex amplification ratio into its more useful polar form, thus

$$\frac{A_h}{A_m} = \frac{|1 + jKy|}{|(1 - Ky^2) + jy|} = \sqrt{\frac{1 + K^2y^2}{1 + (1 - 2K)y^2 + K^2y^4}} \quad \text{HIGH-BAND} \quad (7-84)$$

where
$$y = \frac{f}{f_2}$$

The expression for θ_h , the relative phase shift, may also be derived in the same manner, but the result which will appear as the difference between two angles is very difficult to manipulate. An equivalent expression which has a simpler form may be obtained by first rationalizing equation (7-83), thus

$$\frac{A_h}{A_m} = \frac{(1 + jKy)[(1 - Ky^2) - jy]}{(1 - Ky^2)^2 + y^2} = \frac{1 - jy(1 - K + K^2y^2)}{(1 - Ky^2)^2 + y^2} \quad (7-85)$$

and $\theta_h = -\arctan y(1 - K + K^2y^2)$ HIGH-BAND (7-86)

Equations (7-77), (7-84), and (7-86) are basic in the analysis and design of shunt-compensated video amplifiers on a steady-state basis. They show that shunt compensation can never give an amplifier which has the ideal amplification and phase response. This fact may be seen from the following: Consider first the amplification. If A_h/A_m is to remain constant, the right-hand member of eq. (7-84) must be independent of y , the normalized frequency. Since a y^4 term is present in the denominator of the expression but not in the numerator, no possible value of K can give the ideal condition.

A similar argument may be used to show that θ_h cannot vary linearly with y . This is demonstrated under Case III which is covered later in this section. Since the ideal conditions cannot be obtained exactly, the problem in design is to effect the best compromise between constant amplification on the one hand and linear phase shift on the other. Two general philosophies may be used for this compromise, the first, for the lack of a generally accepted term, we shall call *simple shunt compensation*. The second approach leads to the *Freeman-Schultz design condition*. We consider these in order.

7-9. Simple Shunt Compensation

The basis of the simple shunt-compensating design is that the upper half-power frequency, f_2 , is made identical to f_c , the top video frequency or highest frequency to be amplified. It is apparent, then, that all those frequencies lying between $0.1f_2$ and f_2 will suffer both phase and frequency distortion. The compensating inductance, L , must be chosen to minimize this distortion. Generally speaking, L (or its design parameter K) may be chosen to optimize either the amplification response or phase response, or to give some compromise between the two. Some typical cases are considered in the following paragraphs.

CASE I. Design Condition: The relative amplification at the top video frequency shall be 1.

The required value of K to meet this condition may be determined from (7-84).

At $f = f_2 = f_c$, $y = 1$ and $\frac{A_h}{A_m} = 1$. Then

$$1 = \sqrt{\frac{1 + K^2}{1 + (1 - 2K) + K^2}}$$

whence $K = \frac{1}{2}$ (7-87)

Equations (7-84) and (7-86) are plotted for $K = \frac{1}{2}$ in Fig. 7-10. The delay characteristic is also plotted because in general it is easier to judge constancy of delay than linearity of phase. It will be observed that whereas the response is unity at $f = f_c$, at lower frequencies the amplification response exhibits a hump. If several such stages are cascaded, this hump becomes extremely pronounced since

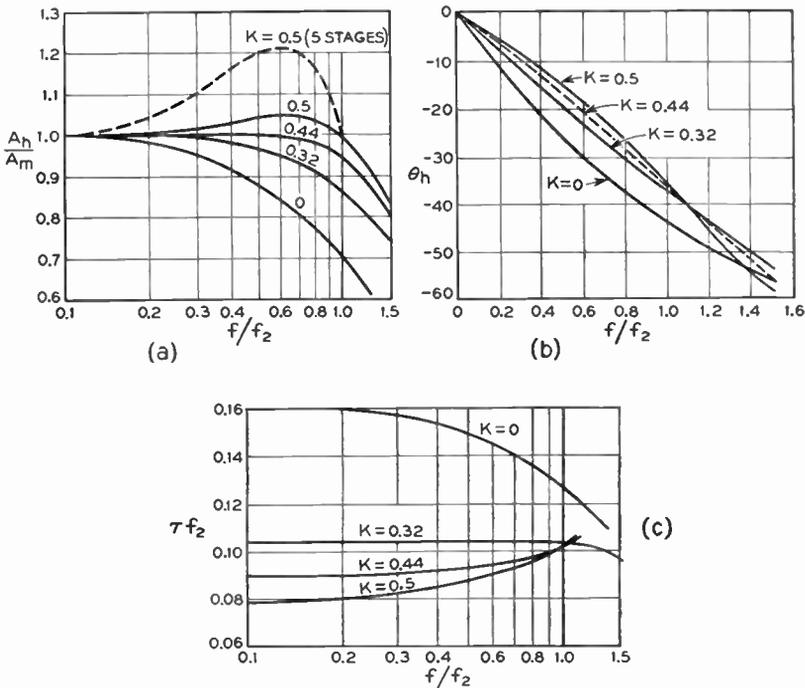


Fig. 7-10. The effect of shunt compensation on the steady-state response of an amplifier. $K = \omega_2 L / R_1$. (a) Gain. (b) Phase shift.

the total response is the product of the individual responses, and the amplifier becomes of little use. The response for five cascaded stages is shown on the graph. It may also be seen that $K = \frac{1}{2}$ does not meet the requirements of linear relative phase response.

CASE II. Design Condition: The amplification response shall have maximum flatness. Here we try to minimize the hump which Case I gave. We have already observed that it is impossible to have A_h/A_m remain perfectly constant. Subject to certain limitations, however, we can approximate this ideal condition. Since we have chosen f_c to be identical with f_2 , the operating range of the amplifier is limited to values of y equal to or less than unity. K is generally a quantity less than unity. It follows that for most of the operating range the K^2y^4 term in the denominator of (7-84) will make only a small contribution to the value of A_h/A_m . Then, subject to the condition K^2y^4 is small, we may write for (7-84)

$$\frac{A_h}{A_m} \approx \sqrt{\frac{1 + K^2y^2}{1 + (1 - 2K)y^2}} \quad (7-88)$$

This expression will remain constant if the coefficients of y^2 in the numerator and denominator are equal; hence we write

$$K^2 = 1 - 2K$$

or
$$K = 0.415 \quad (7-89)$$

If this value of K is substituted into (7-84) and (7-86) the two responses may be calculated. The departure from linearity of the relative phase characteristic and departure from constant delay are still noticeable; the decrease in A_h/A_m for values of y near unity is effected by the K^2y^4 term which is no longer negligible as assumed in the derivation. The relative importance of this term may be seen if the student calculates the curves.

CASE III. Design Condition: The relative phase characteristic shall be linear with frequency. The problem for this condition is to determine a value of K such that the slope of the θ_h vs. y curve is constant. Thus we differentiate (7-86):

$$\begin{aligned} \frac{d\theta_h}{dy} &= - \frac{1 - K + 3K^2y^2}{1 + [(1 - K)y + K^2y^3]^2} \\ &= - \frac{(1 - K) + 3K^2y}{1 + (1 - K)^2y^2 + 2(1 - K)K^2y^4 + K^4y^6} = \text{constant} \quad (7-90) \end{aligned}$$

Let us factor out the first term of the numerator, thus

$$\frac{d\theta_h}{dy} = -(1-K) \left[\frac{1 + \frac{3K^2}{(1-K)} y^2}{1 + (1-K)^2 y^2 + 2(1-K)K^2 y^4 + K^4 y^6} \right] \\ = \text{constant} \quad (7-91)$$

We can see immediately from this equation that $d\theta_h/dy$ can never be independent of y because the fourth- and sixth-power terms in the denominator are not matched by corresponding terms in the numerator. Once again, however, in the range where these two higher order terms are negligible the equation reduces to the approximation

$$\frac{d\theta_h}{dy} \approx -(1-K) \left[\frac{1 + \frac{3K^2}{(1-K)} y^2}{1 + (1-K)^2 y^2} \right] = \text{constant} \quad (7-92)$$

The condition expressed by the last equation may be satisfied if the coefficients of y^2 in both numerator and denominator are equal. Thus we write

$$\frac{3K^2}{1-K} = (1-K)^2 \quad (7-93)$$

whence

$$K \approx 0.32 \quad (7-94)$$

Inspection of the curves for this value of design parameter which are plotted in Figure 7-10 shows that while the θ_h and τ responses are excellent, the amplitude response is inferior to that provided by Case II.

It is evident from these three cases that a value of K should be chosen which will give a reasonable compromise between the two characteristics. One recommended value¹⁷ which gives such a compromise is

$$K = 0.44 \quad \text{COMPROMISE VALUE} \quad (7-95)$$

It is interesting to notice that the value of K for condition III may be derived in an alternative manner. If θ_h is to vary linearly with frequency, we have from (7-49) that θ_h at f_2 must be 10 times the value of θ_h at $0.1f_2$, that is,

$$\theta_{f_2} = 10 \theta_{0.1f_2} \quad (n = 0) \quad (7-96)$$

¹⁷ Cruft Electronics Staff, *Electronic Circuits and Tubes*. New York: McGraw-Hill Book Company, Inc., 1947.

Then, from (7-86)

$$-\text{arc tan } (1 - K + K^2) = -10 \text{ arc tan } [0.1(1 - K + 0.01K^2)] \quad (7-97)$$

A transcendental equation of this type may be solved graphically by letting

$$\left. \begin{aligned} c &= \text{arc tan } (1 - K + K^2) \\ \text{and } d &= 10 \text{ arc tan } [0.1(1 - K + 0.01K^2)] \end{aligned} \right\} \quad (7-98)$$

If these two quantities c and d are plotted against K , the intersection of the curves will give the value of K which satisfies (7-97). This method yields a value of K close to 0.32.¹⁸

The results of this section may be summarized:

| <i>Condition</i> | K |
|----------------------------------|-------|
| Relative response = 1 at f_c | 0.5 |
| Flattest response | 0.415 |
| Most linear phase characteristic | 0.32 |
| Optimum compromise | 0.44 |

The student should take note that all the designs specified above yield the same value of mid-band amplification for a given top video frequency. The reason for this is that in each case f_2 is set equal to f_c by proper choice of R_1 , thus a fixed f_c gives a fixed value of R_1 and corresponding fixed value of A_m . To illustrate how these various design equations may be used, let us calculate the circuit constants for a typical video stage. A 6AC7 is to be used and the design is to give an optimum compromise between flat amplitude and linear phase response up to 4.5 mc. The output feeds a second 6AC7. For the 6AC7

$$g_m = 9000 \text{ micromho}$$

$$C_i = 11 \text{ } \mu\mu\text{f}$$

$$C_o = 5 \text{ } \mu\mu\text{f}$$

The estimated stray capacitances are 5 $\mu\mu\text{f}$. The total shunt capacitance of the plate circuit will consist of the output capacitance of the stage plus the input capacitance of the following stage plus the stray capacity; therefore

$$C_s = 11 + 5 + 5 = 21 \text{ } \mu\mu\text{f}$$

¹⁸ The discrepancy in the results of the two methods is the result of the omission of the higher order terms in (7-92).

By the design condition, f_2 must be the top video frequency; then from (7-76)

$$R_1 = \frac{1}{\omega_2 C_s} = \frac{1}{2\pi(4.5 \times 10^6)(2.1 \times 10^{-11})} = 1,682 \text{ ohms}$$

The required value of compensating inductance is from (7-77)

$$L = \frac{KR_1}{\omega_2} = \frac{(0.44)(1,682 \times 10^3)}{2\pi(4.5 \times 10^6)} = 26.2 \text{ } \mu\text{henry}$$

The mid-band gain of the amplifier is

$$A_m = g_m R_1 = (9 \times 10^{-3})(1,682 \times 10^3) = 15.2$$

Comparison of these results with those obtained for the similar problem in section 7-2 shows that compensation has raised the gain of the circuit by a factor of 10. The remaining components of the coupling network, namely C_c and R_2 , must be chosen to satisfy the low-frequency requirements on the amplifier, a subject discussed later in the chapter.

7-10. Freeman-Schantz Compensation¹⁹

We have previously stated that the problem of high-frequency compensation may be approached in a manner alternative to that just described. This alternative manner provides that the amplifier be

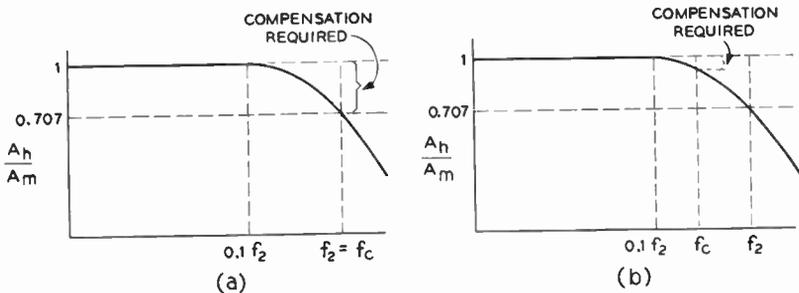


Fig. 7-11. Comparison of compensation by the simple shunt and Freeman-Schantz methods. (a) Simple shunt compensation. The amplifier is designed so that its uncompensated upper half-power frequency is equal to f_c , the highest frequency to be amplified. (b) Freeman-Schantz compensation. f_2 is made higher than f_c . Less compensation is required.

¹⁹ R. L. Freeman and J. D. Schantz, "Video Amplifier Design." *Electronics*, 10, 8 (August 1937).

designed so that its uncompensated half-power frequency is *greater* than, rather than equal to, the top video or cutoff frequency. The choice of the higher value of f_2 of necessity requires a lower value of R_1 , which will give a reduced mid-band amplification. The advantage gained is that less compensation is required. This concept is shown in Fig. 7-11. These diagrams illustrate the two alternate approaches to the shunt compensation of the high-band response.

Notice that even though the philosophy of design has changed, the circuit of the amplifier remains that shown in Fig. 7-9, consequently eq. (7-84) and (7-86) may still be used to calculate the response. Since f_2 and f_c are no longer identical, however, it is convenient to define a new design parameter

$$M = \frac{f_c}{f_2} \tag{7-99}$$

Then once the response has been calculated, the $y = f/f_2$ scale may be converted to, say, $y' = f/f_c$ by dividing y by M .

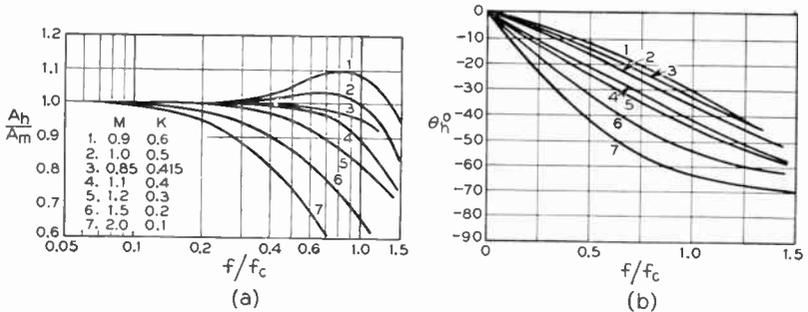


Fig. 7-12. Steady-state response of an amplifier employing Freeman-Schantz compensation. $K = \omega_2 L/R_1$, $M = f_2/f_c$. (a) Gain. (b) Phase shift.

The amplification and phase characteristics for several values of M and K are plotted in Fig. 7-12. Of these values, the set

$$M = 0.85 \quad \text{and} \quad K = 0.415 \quad \text{FREEMAN-SCHANTZ CONDITION} \tag{7-100}$$

are recommended by Freeman and Schantz to give a good compromise design.

7-11. The Vector Diagrams

Some physical picture of how the compensating inductance affects the relative phase characteristic may be had by comparing the vector diagrams for the compensated and uncompensated amplifiers. Reference to Fig. 7-3 shows that the uncompensated amplifier has too large a value of relative phase shift at each frequency in the high-band. Compensation should decrease the value of θ_h at each frequency so that the linear characteristic is approached. We now show that L does decrease θ_h by means of the vector diagrams of Fig. 7-13. To

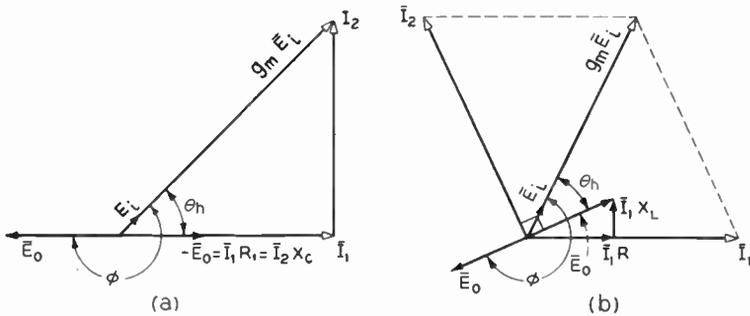


Fig. 7-13. The effect of shunt compensation on phase shift at f_2 , the uncompensated upper half-power frequency. (a) Uncompensated. (b) Compensated.

simplify the drawing of the sketches we have arbitrarily chosen I_1 , the current through the plate load resistor R_1 , to be the reference vector and the condition shown is that at f_2 , the upper half-power frequency. In the uncompensated amplifier X_c and R_1 are equal. Since the same voltage appears across the resistor and the condenser, I_1 and I_2 are equal in magnitude and in phase quadrature. θ_h , the angle between E_i and $-E_{oh}$, is 45° .

The diagram for the compensated case assumes $K = \frac{1}{2}$; hence at f_2 the impedance of the R_1 branch is

$$|Z_{L,R_1}| = \sqrt{R_1^2 + (KR_1)^2} = 1.12R_1 = 1.12X_c \quad (7-101)$$

Again the same voltage appears across the two parallel branches so that

$$|I_2| = 1.12 |I_1| \quad (7-102)$$

Where a number of video amplifiers are cascaded more care must be exercised in choosing K , for as we have previously seen, the effect of any overshoot in the transient response becomes quite pronounced as the number of stages in cascade is increased. This effect is illustrated by the curves in Fig. 7-15 which were calculated by Bedford and Fredendall.^{22,23} These curves show that a K of 0.5 is quite unsatisfactory from the viewpoint of overshoot and ring, and that a value between 0.384 and 0.438 gives a better compromise transient response. Palmer and Mautner²⁴ have calculated curves of rise time and overshoot for a single shunt-compensated stage and recommend $K = 0.388$, which gives a 2 per cent overshoot per stage, as the optimum value.

A word of caution in interpreting the curves of Fig. 7-14 for the Freeman-Schantz design condition should be mentioned. Since their design is based on the cutoff frequency, f_c , rather than on f_2 , the half-power frequency, the abscissas should be interpreted in terms of f_c .

A practical consideration arises when the foregoing results are applied in practice. In all of our analytical work on shunt compensation we have assumed L to be a pure inductance, the shunt capacitance across the coil having been neglected. This assumption cannot be realized physically and where a bandwidth of one or more megacycles is used, the shunt capacitance upsets the design assumptions to a considerable extent. It is often the practice to provide the compensating inductance with a movable core in order that the inductance may be adjusted to the correct value experimentally. Alternatively, the compensating network may be analyzed to take shunt capacitance across the inductance into account, a procedure which results in the so-called m -derived shunt-peaking network. It is demonstrated in the literature²⁵ that proper design of this network gives an increase in gain of 1.6 to 1.8 times that obtained with the shunt-peaking network. The shunt capacitance required across L ,

²² A. V. Bedford and G. L. Fredendall, "Transient Response of Multistage Video Amplifiers." *Proc. IRE*, **27**, 4 (April 1939). The reader should note that the design parameter K used by the authors is not the same as that defined by eq. (7-77). If K' be the value defined by Bedford and Fredendall, the relationship is $K = 1/(K')^2$.

²³ The calculation of these curves by the method of the Laplace transform is illustrated in G. E. Valley, Jr., and H. Wallman, *op. cit.*, chap. 1.

²⁴ R. C. Palmer and L. Mautner, *op. cit.*

²⁵ See, for example, S. Moskowitz and J. Racker, *op. cit.*

however, is greater than the stray value alone so that an additional circuit element is required to realize this increase in gain.

7-13. Series Compensation

In the shunt-compensated circuits which have been described the equivalent high-band coupling system between the plate of the stage under consideration and the grid of the following stage is a two-terminal network. This fact may be verified by reference to the equivalent circuit diagram of Fig. 7-9b where the reactance of the coupling condenser, C_c , has been assumed to be negligible. The use of a four-terminal coupling network will give a higher mid-band gain and a more linear phase characteristic than does the simpler two-terminal configuration. The price which must be paid for these advantages is the increased difficulty in proper adjustment of the compensating elements and a poorer transient response.

A number of forms of four-terminal compensating networks may be used. One of these, which is shown in Fig. 7-16a, results in what is

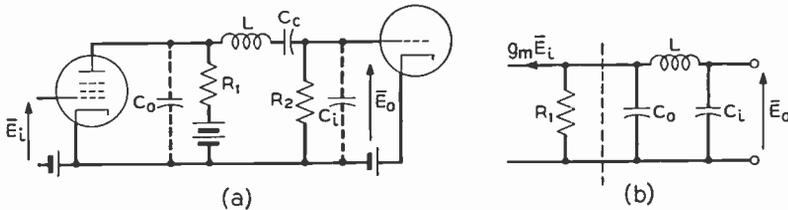


Fig. 7-16. Series compensation utilizes a four-terminal coupling network. (a) Series-compensated amplifier. (b) Equivalent high-frequency-band circuit.

known as series peaking or compensation because the compensating inductance is a series rather than a shunt element in the coupling network. Subject to the same simplifying assumptions which were used in the shunt-compensated circuit, the equivalent high-band circuit is that of Fig. 7-16b. These assumptions are that r_p and R_2 are large enough to have negligible shunting effect on the network, and that the high-band reactance of C_c is negligible. Inspection of the circuit shows why series compensation provides a higher mid-band gain than does the previous circuit. In the shunt case C_o and C_i are in parallel and combine into C_s . Then for a given half-power frequency R_1 and hence A_m are determined by the sum of C_c and C_i .

In the present case, on the other hand, the series inductance effectively separates C_o from C_i , and only the former shunts R_1 . Obviously, then, for the same f_2 , R_1 will be larger, being determined by C_o alone rather than $C_o + C_i$. The larger value of R_1 , in turn, gives a higher mid-band gain. Direct application of Kirchhoff's laws to the circuit yields the result that the high-band amplification normalized with respect to the mid-band amplification is

$$\frac{A_h}{A_m} = \frac{1}{(1 - \omega^2 LC_i) + j[\omega R_1(C_i + C_o) - \omega^2 LC_o - \omega^2 LC_o \omega C_i R_1]} \quad (7-105)$$

Seeley and Kimball have recommended the following design values to give an optimum compromise between the steady-state amplitude and phase characteristics. Let f_c be the top video or cutoff frequency.

$$\left. \begin{array}{l} \text{Then} \quad L = \frac{1}{2\omega_c^2 C_o} \\ \text{and} \quad R_1 = \frac{1.5}{\omega_c(C_i + C_o)} \end{array} \right\} \text{SERIES COMPENSATION} \quad (7-106)$$

$$(7-107)$$

Subject to these two design conditions, the several terms of eq. (7-105) may be reduced to some power of a normalized frequency. The resulting expressions for relative gain and phase are simplified considerably by this procedure. We also define a parameter m , the ratio of C_i to C_o ,

$$m = \frac{C_i}{C_o} \quad (7-108)$$

It is convenient to normalize the frequency with respect to f_c , the upper cutoff frequency; thus

$$y = \frac{f}{f_c} = \frac{\omega}{\omega_c} \quad (7-109)$$

$$\text{Then, from (7-107)} \quad \omega R_1(C_i + C_o) = 1.5 \frac{\omega}{\omega_c} = 1.5y \quad * \quad (7-110)$$

$$\text{Similarly, from (7-106)} \quad \omega^2 LC_o = \frac{\omega^2}{2\omega_c^2} = \frac{y^2}{2} \quad (7-111)$$

$$\text{and} \quad \omega C_i R_1 = [\omega R_1(C_i + C_o)] \frac{C_i}{C_i + C_o} = \frac{1.5m}{1+m} y \quad (7-112)$$

Combining the last two equations we get

$$\omega^2 LC_o \omega C_i R_1 = \frac{1.5}{2} \frac{m}{1+m} y^3 \quad * \quad (7-113)$$

It also follows directly from the definition of m and (7-111) that

$$\omega^2 LC_i = \frac{my^2}{2} \quad * \quad (7-114)$$

The three quantities marked with asterisks may be identified with their counterparts in eq. (7-105). Making the necessary substitutions and reducing the result to polar form, we finally have

$$\frac{A_h}{A_m} = \frac{1}{\sqrt{1 + (2.25 - m)y^2 + \left(\frac{m^2}{4} - 2.25 \frac{m}{1+m}\right)y^4 + 0.5625 \left(\frac{m}{1+m}\right)^2 y^6}} \quad (7-115)$$

and

$$\theta_h = -\arctan \left[\frac{1.5y - 0.75 \left(\frac{m}{1+m}\right) y^3}{1 - \frac{m}{2} y^2} \right] \quad (7-116)$$

It has been shown that best operation of the circuit on a steady-state basis is obtained when $m \geq 2$ and the Q of the inductance coil is equal to or exceeds 20.

In reference to the last statement it is of interest to note that in typical tubes used for video work m is greater than 2, *i.e.*, the input capacitance exceeds the output capacitance by a factor of 2 or more. To verify this refer to the 6AC7 constants given in the example in section 7-9. Some small degree of flexibility is afforded in adjusting the C_i to C_o ratio in the placement of C_c , the coupling condenser in Figure 7-16a. This comes about by virtue of the shunt capacitance between the outer foil of the condenser and ground. If C_c is placed on the grid side of L this capacitance contributes to C_i , conversely it adds to C_o if placed on the plate side of the compensating inductance. Because of low-frequency-band considerations C_c will be in the order of a microfarad, and the shunt capacitance to ground may range up to 2 micromicrofarads, the exact value being dependent upon the placement of C_c relative to the chassis and other grounded components.

In certain rare instances C_o may be much greater than C_i . In such

a case it is inadvisable to increase C_i by adding shunt capacitance because this procedure must inevitably result in a lower mid-band gain for a given bandwidth, as may be seen from (7-107). A better approach is to take advantage of the principle of reciprocity by interchanging the positions of R_2 , the relatively high grid leak resistance, and R_1 , the relatively low plate load resistance. It should be noted, however, that the resulting increase in voltage drop across the new plate resistance, R_2 , will lower the d-c plate voltage of the tube, which in turn will lower the mutual conductance. Let us now examine these steady-state results in terms of transient response. We shall assume a value of $m = C_i/C_o = 2$ and introduce a parameter

$$k = \frac{L}{C_o R_1^2} \quad (7-117)$$

It may be shown from eqs. (7-106) and (7-107) that the recommended values of L and R_1 give a $k = 2$. If the transient response of the series-compensated amplifier is plotted for this value as shown in Fig. 7-17, it is observed that an overshoot of approximately 10 per

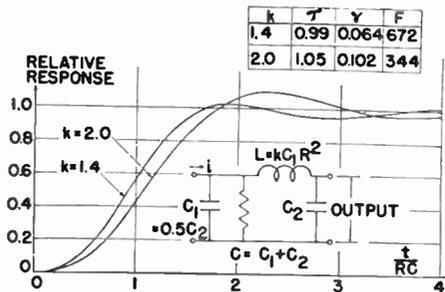


Fig. 7-17. Transient response of a series-compensated amplifier. Notice that $k = 1.4$ gives a 6 per cent undershoot at $t = 2.8RC$. (Courtesy of *Proc. IRE*.)

cent is obtained. Thus from our previous work we see that the recommended values are not satisfactory where a number of stages are to be cascaded. Some improvement in the response may be obtained by lowering the value of k to 1.4, as shown in the figure. It should be noticed, however, that an undershoot of 6.4 per cent occurs in the ring which, in turn, will give poor results in a cascaded amplifier. We must conclude therefore that in spite of the superior steady-state response of the series-compensated amplifier, when several

stages are to be connected in cascade, shunt compensation with $K = 0.388$ gives superior performance.

7-14. Series-shunt Compensation

A second form of four-terminal interstage coupling network may be had by combining the two types of compensating circuits which have been described. The advantage of using both the series- and shunt-compensating inductors (Fig. 7-18a) is that both series and shunt compensation occur simultaneously. L_2 acts to separate C_i and C_o ,

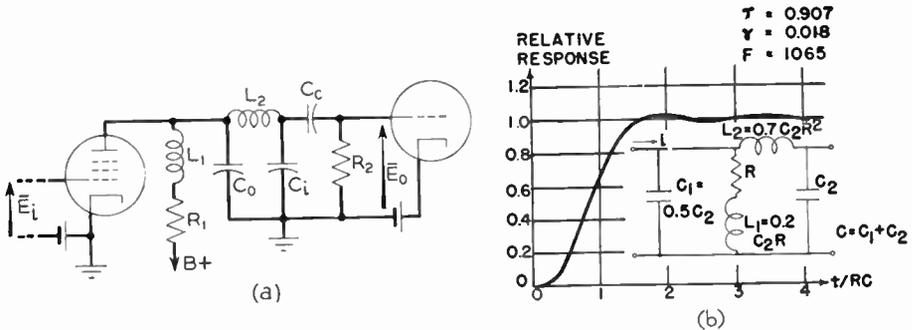


Fig. 7-18. Series-shunt compensation. (a) Basic circuit. (b) Transient response. (Courtesy of *Proc. IRE.*)

and L_1 tends to counteract the shunting effect of C_o on R_1 . The net result is that for a given bandwidth a larger value of R_1 may be used with a corresponding increase in the mid-band gain. Seeley and Kimball have recommended the following design values:

$$\left. \begin{aligned}
 \text{(a)} \quad m &= \frac{C_i}{C_o} = 2 \\
 \text{(b)} \quad R_1 &= \frac{1.8}{\omega_c(C_i + C_o)} \\
 \text{(c)} \quad L_1 &= 0.12(C_i + C_o)R_1^2 \\
 \text{(d)} \quad L_2 &= 0.52(C_i + C_o)R_1^2
 \end{aligned} \right\} \text{SERIES-SHUNT COMPENSATION (7-118)}$$

where ω_c is 2π (top video frequency).

The transient response for these values is shown in Fig. 7-18b. It will be observed that the results obtained are excellent. Let us now illustrate the use of these design equations by designing a typical

series-shunt-compensated amplifier. In order to show how compensation improves the performance of an amplifier, we shall design to meet the 10 to 90 per cent rise time of the resistance-coupled amplifier discussed in section 7-5. Thus

$$g_m = 9000 \mu\text{mho}$$

$$\tau_r = 0.22 \mu\text{sec}$$

$$C_s = 20 \mu\mu\text{f}$$

We shall further assume that C_s is divided between C_i and C_o to give an m of 2. Thus

$$C_i = 13.3 \mu\mu\text{f}$$

$$C_o = 6.7 \mu\mu\text{f}$$

Then, from Fig. 7-18*b*, the rise time is approximately

$$\tau_r = 1.65R_1C_s$$

or

$$R_1 = \frac{\tau_r}{1.65C_s} = \frac{2.2 \times 10^{-7}}{1.65(2 \times 10^{-11})} = 1.333 \times 10^4 = 13.33 \text{ kilohms}$$

$$L_1 = 0.12C_sR_1^2 = 0.12(2 \times 10^{-11})(1.333)^2(10)^8 = 0.426 \text{ mh}$$

$$L_2 = 0.52C_sR_1^2 = 0.52(2 \times 10^{-11})(1.333)^2(10)^8 = 1.85 \text{ mh}$$

and

$$A_m = g_mR_1 = (9 \times 10^{-3})(1.333 \times 10^4) = 120$$

It is interesting to notice from this example that the compensating elements give a mid-band gain of 2.7 times that obtained with the uncompensated amplifier without any appreciable change in the transient response. It is of further interest to note that if the two amplifiers were designed for the same bandwidth or cutoff frequency, the increase in gain resulting from compensation would be a factor of only 1.8. This fact may be verified by the use of eq. (7-118*b*).

The student should notice that the four-terminal compensating networks which have been described have the general form of low-pass filters. By the addition of other elements these basic forms may be converted to m -derived and dead-end filters. In general, the more compensating elements used, the better will be the response and the higher will be the midband gain. It is beyond the scope of this text to analyze these more complicated circuits but they are described in

the literature.²⁶ Needless to say the adjustment of the several compensating elements in the filter networks is more difficult than for the simpler circuits which we have considered. Where a number of video stages are in cascade, a considerable increase in gain per stage may be had by "stagger peaking" the several stages. Easton²⁷ has proposed such a design for determining the peaking of each of the compensating circuits which gives twice the mid-band gain afforded by simple shunt compensation.

7-15. Summary of High-frequency Compensation

Our results in this chapter show that the transient behavior of an amplifier is not readily expressed in terms of its amplitude and phase characteristics; hence we shall summarize our work on the basis of the transient response of the several compensating networks. Comparisons of these networks on a steady-state basis are available in the literature.²⁸ The uncompensated resistance-coupled amplifier exhibits a long rise time, zero overshoot, and a low value of gain. Shunt, series, and series-shunt compensation provide means of counteracting the effect of shunt capacitance and so allow a larger value of R_1 to be used with a corresponding increase in gain. In each case, the design parameters must be chosen to minimize rise time and overshoot, the latter factor being of great importance when several stages are connected in cascade.

7-16. Square-wave Testing

We have stressed the importance of the transient response of a video amplifier in television applications and have discussed analytical means of calculating this response to an applied voltage of step-function form. It is highly desirable, therefore, to discover some means of determining whether the proper amount of compensation has been applied to an amplifier. Since the response consists of an output voltage as a function of time, the cathode-ray oscilloscope suggests itself as an ideal piece of test equipment, the general layout

²⁶ A summary of design equations for several compensating circuits is given in D. E. Foster and J. A. Rankin, "Video Output Systems." *RCA Review*, V, 4 (April 1941).

²⁷ A. Easton, "Stagger-Peaked Video Amplifiers." *Electronics*, 22, 2 (February 1949).

²⁸ See, for example, S. W. Seeley and C. N. Kimball, *op. cit.*, and D. E. Foster and J. A. Rankin, *op. cit.*

of the test unit consisting of the step-function generator feeding the amplifier under test, the output being viewed directly on the oscilloscope.

It is immediately apparent that such a test setup would be difficult to use because a step function, being a nonrepetitive wave, would produce only a single trace of the output wave form on the oscilloscope screen, which would be difficult to observe. For the testing technique, then, we desire to replace the step-function signal by a repetitive signal which retains the essential properties of the step as a test voltage. Such a substitute voltage is the square wave.

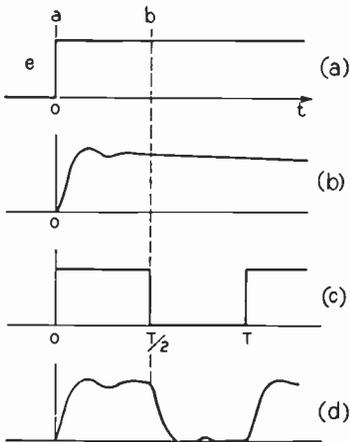


Fig. 7-19. If step function is replaced by a square wave of proper half-period, the response at the discontinuity remains unaffected. (a) Step function. (b) Response to step function. (c) Square wave of proper half-period. (d) Square-wave response.

The substitution of a square wave for step function for determining the transient response of a network may be justified on the following grounds: We have seen that the transient response may be conveniently divided into two separate regions, one at the discontinuity and the other one after the discontinuity when the step function remains at a constant d-c value. If, then, we confine our attention to the response at the discontinuity, we only require a signal which presents the discontinuity, that is, which duplicates the leading edge of the step function, over and over again at fixed intervals. A square wave does precisely this. We only need be careful in choosing the period of the square wave to be sufficiently long so that the response to the discontinuity is essentially

completed before a new cycle is applied. This concept is illustrated in Fig. 7-19. The response to the discontinuity in the applied step function is completed in the time interval \overline{ab} ; hence if the step function is replaced by a square wave of half-period $T/2 = \overline{ab}$, the response to the leading edge remains unaffected.

In general the interval \overline{ab} is not known until the test has been made; hence some independent means for determining the frequency

of the square-wave signal is needed. Since the steady-state response of the amplifier may be measured or calculated quite readily, it is convenient to use this information to determine the proper square-wave frequency. On this basis Bedford and Fredendall²⁹ have recommended that the square-wave frequency be chosen as that frequency at which the steady-state characteristics begin to vary with frequency.

In our analysis of the uncompensated amplifier at the beginning of the chapter we found that the transition between the mid- and high-frequency band occurs at approximately one-tenth of the upper half-power frequency. Thus a rule-of-thumb approximation to Bedford and Fredendall's criterion is to choose the square-wave frequency to be about one-tenth of the amplifier cutoff frequency. The final value may be determined during the test procedure by adjusting the frequency so that the discontinuity response is completed, as shown in Fig. 7-19*d*.

A typical test setup for experimentally determining the square-wave response of an amplifier is shown in Fig. 7-20*a*. The 20-mc

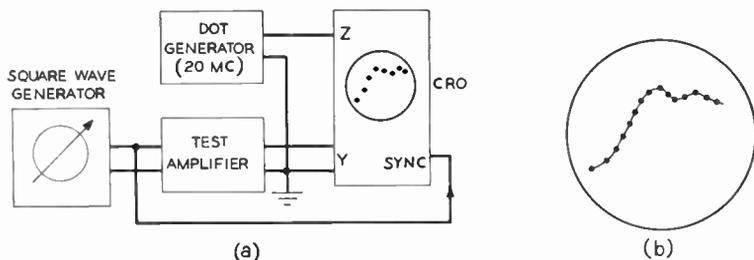


Fig. 7-20. Square wave testing. (a) Test setup for determining the square-wave response of an amplifier. (b) Appearance of the oscilloscope pattern for a typical shunt-compensated amplifier.

dot generator serves to provide a convenient time scale on the oscilloscope screen so that rise time and other pertinent intervals in the response curve may be determined. The operation of the unit is this: The intensity control of the oscilloscope is adjusted so that no trace is visible on the screen. The output of the dot generator, which consists of very narrow pulses which occur at a 20-mc repeti-

²⁹ A. V. Bedford and G. L. Fredendall, "Transient Response of Multistage Video-Frequency Amplifiers." *Proc. IRE*, **27**, 4 (April 1939).

tion rate, is fed to the z axis terminal (control grid) of the oscilloscope and serves to unblank the trace at $\frac{1}{20}$ -mc = 0.05- μ sec intervals. The appearance of the oscilloscope trace under these conditions is shown in Fig. 7-20b. Since the bright dots occur at known intervals, rise time, delay time, and the like may be read off the screen directly.

We see, then, that the square wave of proper frequency may be used to an excellent advantage in determining the response of an amplifier to the discontinuity in an applied step voltage. Later in the chapter we shall extend the method of square-wave testing to include the response to the d-c portion of a step voltage. The usefulness of the square-wave technique as a testing tool cannot be overemphasized.

7-17. Calculation of Square-wave Response

In the last section we saw that the transient response in the vicinity of a discontinuity in an applied step function may be determined experimentally by the use of an applied square wave. It would seem that this concept could be extended to include an analytical means of determining transient response, the advantage being that the square wave is subject to analysis by the more familiar Fourier series rather than by the Fourier integral. We shall merely outline the method which is covered at length in the literature.^{30,31}

Let a square wave of fundamental frequency ω be applied to the input terminals of the network under consideration. By Fourier's analysis the square wave may be reduced to the summation of an infinite number of harmonically related sinusoidal components, thus if the square wave of unit amplitude be expanded as an odd function of time we have

$$e_i(t) = \frac{1}{2} + \frac{2}{\pi} \left(\sin \omega t + \frac{1}{3} \sin 3\omega t + \dots + \frac{1}{n} \sin n\omega t \right) \quad (7-119)$$

where n is an odd integer. Since $e_i(t)$ consists of a number of discrete, sinusoidal components, we may calculate the output voltage result-

³⁰ A. V. Bedford and G. L. Fredendall, "Analysis, Synthesis, and Evaluation of the Transient Response of Television Apparatus." *Proc. IRE*, **30**, 10 (October 1942).

³¹ Philip M. Seal, "Square-Wave Analysis of Compensated Amplifiers." *Proc. IRE*, **37**, 1 (January 1949); also "Correction on 'Square-Wave Analysis of Compensated Amplifiers,'" **37**, 4 (April 1949).

ing from each component, using the steady-state response of the network at the frequency of that component, *i.e.*,

$$e_o(t)_{k\omega} = \frac{2}{\pi} A(k\omega) \sin [k\omega t + \phi(k\omega)] \quad (7-120)$$

Then by the superposition theorem the total output voltage caused by the applied square wave will be the sum of the output voltages due to each component, or

$$e_o(t) = \frac{A_o}{2} + \frac{2}{\pi} \left\{ A(\omega) \sin [\omega t + \phi(\omega)] + \frac{A(3\omega)}{3} \sin [3\omega t + \phi(3\omega)] \right. \\ \left. + \cdots + \frac{A(n\omega)}{n} \sin [n\omega t + \phi(n\omega)] \right\} \quad (7-121)$$

Notice that we now have an expression which involves no integration, and the analytical expressions of $A(\omega)$ and $\phi(\omega)$ are not required; we only need their values at certain discrete frequencies which are odd multiples of the fundamental frequency of the square wave.

By way of illustrating how the Fourier series method is applied, let us set up the equations for the response of the "ideal" amplifier, whose characteristics are given in Fig. 7-8*a*, to a square wave of frequency ω .

Now in the pass band

$$\left. \begin{aligned} A(\omega) &= A \\ \phi(\omega) &= \omega\tau_d \end{aligned} \right\} \quad (7-122)$$

and

Substitution of these values into eq. (7-121) shows that the expression for the output voltage is

$$e_o(t) = \frac{A}{2} + \frac{2A}{\pi} \left\{ \sin (\omega t - \omega\tau_d) + \frac{1}{3} \sin (3\omega t - 3\omega\tau_d) \right. \\ \left. + \cdots + \frac{1}{n} \sin (n\omega t - n\omega\tau_d) \right\} \\ = \frac{A}{2} + \frac{2A}{\pi} \left\{ \sin \omega(t - \tau_d) + \frac{1}{3} \sin 3\omega(t - \tau_d) \right. \\ \left. + \cdots + \frac{1}{n} \sin n\omega(t - \tau_d) \right\} \quad (7-123)$$

For specific values of ω and τ_d eq. (7-123) may be evaluated directly.

Alternatively we may calculate the response for any ω and τ_d by normalizing the time-frequency variable. Thus let

$$\left. \begin{aligned} & \text{then} \\ & x = \omega(t - \tau_d) \\ & e_o(t) = \frac{A}{2} + \frac{2A}{\pi} \left(\sin x + \frac{1}{3} \sin 3x + \dots + \frac{1}{n} \sin nx \right) \end{aligned} \right\} \quad (7-124)$$

which may be evaluated for different values of x . Notice that $t = 0$ corresponds to a value of $x = -\omega\tau_d$; the response on a normalized basis must include negative values of x .

If the student evaluates either of the last two equations, he will observe that a considerable amount of slide rule work is involved, particularly since at least 10 terms of the series, that is, up to $n = 19$, must be evaluated in order for a good approximation to the actual response to be obtained.³² For a smaller number of harmonics, the use of a finite number of terms rather than the whole infinite summation of the Fourier series contributes falsely to the overshoot and ring in the calculated response. The amount of work may be reduced appreciably by resorting to graphical charts used to evaluate the various sinusoidal components. Charts of this type have been published by Bedford and Fredendall.³³ The square-wave technique of analysis also affords a means of attacking the inverse problem of deriving the steady-state characteristics of a network when its transient response is known. This method which has been developed into a usable form by Bedford and Fredendall will now be outlined.

By direct application of Fourier's theorem we know that the function $e_o(t)$, given in Fig. 7-20b, may be expanded into an infinite series of harmonically related terms.³⁴ Since the function exhibits neither odd nor even symmetry, the full sine and cosine form of the series must be used, thus

³² For a discussion of this manifestation of the Gibb's phenomenon see E. A. Guillemin, *The Mathematics of Circuit Analysis*. New York: John Wiley and Sons, Inc., 1949, chap. 7.

³³ A. V. Bedford and G. L. Fredendall, "Analysis, Synthesis, and Evaluation of the Transient Response of Television Apparatus," *op. cit.* This paper also considers the inverse problem of constructing the steady-state characteristics from a given square-wave response.

³⁴ Remember that the square-wave response repeats itself in time. The action of the oscilloscope sweep circuit causes consecutive cycles to overlap, so that the response appears to be nonrepetitive.

$$e_o(t) = \frac{C_o}{2} + S_1 \sin \omega t + C_1 \cos \omega t + S_2 \sin 2\omega t + C_2 \cos 2\omega t + \dots + S_n \sin n\omega t + C_n \cos n\omega t \quad (7-125)$$

where

$$C_n = \frac{1}{T} \int_0^T e_o(t) \cos n\omega t \, dt$$

and

$$S_n = \frac{1}{T} \int_0^T e_o(t) \sin n\omega t \, dt \quad (7-126)$$

Each integral of (7-126) is clearly the area under the curve which is the product of $e_o(t)$ and $\sin n\omega t$ or $\cos n\omega t$, as the case may be; hence each of the coefficients in the series (7-125) may be evaluated graphically, if by no other means. Each pair of terms corresponding to a given value of n may be combined into a single term of amplitude E_n and of phase ϕ_n .

$$E_n = \sqrt{S_n^2 + C_n^2}$$

$$\phi_n = \text{arc tan } \frac{S_n}{C_n} \quad (7-127)$$

and the series reduces to

$$e_o(t) = \frac{E_o}{2} + E_1 \sin (\omega t + \phi_1) + E_2 \sin (2\omega t + \phi_2) + \dots + E_n \sin (n\omega t + \phi_n) \quad (7-128)$$

This is the term-by-term expression of the output voltage corresponding to the input voltage given by eq. (7-119) and hence must be equal to the expression (7-121). Then the termwise comparison of (7-121) and (7-128) shows that

$$\frac{2}{\pi} \frac{A(n\omega)}{n} \sin [n\omega t + \phi(n\omega)] = E_n \sin (n\omega t + \phi_n) \quad (7-129)$$

whence

$$A(n\omega) = \frac{n\pi}{2} E_n$$

and

$$\phi(n\omega) = \phi_n \quad (7-130)$$

Thus the steady-state responses may be evaluated at odd multiples of the fundamental frequency of the applied square wave.

It will be realized that in evaluating these equations an enormous amount of labor is involved, and the method described is not in com-

mon use. An alternative method which employs graphical aids has been developed by Bedford and Fredendall.³⁵

The difficulty in deriving the steady-state characteristics from the square-wave response must not be allowed to obscure the desirability of the square-wave testing technique; in itself it is a powerful tool for checking the response of video amplifiers experimentally.

LOW-FREQUENCY COMPENSATION

Our attention for several sections has been directed to the high-frequency response of a video amplifier and to its transient response in the vicinity of a discontinuity. We must now consider the second problem, that of compensation at the low-frequency end of the video spectrum. This will lead to a discussion of the transient response of the amplifier to the constant d-c portion of an applied step function. We have already noted for the uncompensated amplifier that these two problems are related. Consider first the steady-state low-frequency response.

We have derived the equations for the amplitude and phase response of the resistance-coupled amplifier in the low-band of frequencies. These results, given by eq. (7-46) and (7-47), may be simplified in the broad-band video amplifier case because of the relative

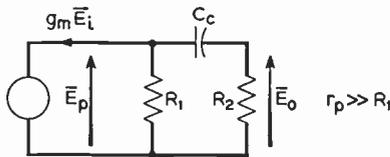


Fig. 7-21. Simplified equivalent low-frequency circuit of a video amplifier.

magnitudes of certain of the circuit parameters. For example, we have seen that high-band considerations require that R_1 be small, in the order of 1000 ohms; hence, R_1 is small compared to r_p and R_2 . Furthermore, in the low-band the reactance of C_c becomes important and the shunting

effect of C_c is negligible. These facts, combined with Fig. 7-2c, give the simplified low-frequency equivalent circuit for the video amplifier which is shown in Fig. 7-21. The expression for the output voltage subject to the conditions listed above may be derived by reducing R_L of (7-41) to its value here, namely R_2 . It is more desirable from the point of view of compensation, however, to derive expressions for gain and phase on a slightly different basis.

³⁵ *Op. cit.*

7-18. Coupling-network Effect

Under the conditions listed in the last paragraph $R_2 \gg R_1$. Regardless of frequency, then, the impedance of R_2 and C_c will have negligible shunting effect on R_1 and we may consider E_o as a fraction, determined by the voltage divider $C_c R_2$, of a constant voltage E_p , given by

$$E_p = -g_m E_i R_1 \quad (7-131)$$

In the low-band, we have for the output voltage

$$E_{ol} = E_p \frac{R_2}{R_2 - \frac{j}{\omega C_c}} = E_p \frac{1}{1 - \frac{j}{\omega C_c R_2}} \quad (7-132)$$

We may relate E_{ol} to the corresponding output voltage in the mid-band by noting that as the frequency of the applied voltage is raised, the reactance of C_c approaches zero, that is, the reactive term in the denominator of (7-132) drops out and we have

$$E_{om} = E_p = -g_m E_i R_1 \quad (7-133)$$

This result, of course, checks with the value of mid-band output voltage obtained in the analysis of the high-band response. It is convenient to combine the equations given above so that

$$E_{ol} = E_{om} \alpha \quad (7-134)$$

where

$$\alpha = \frac{1}{1 - \frac{j}{\omega C_c R_2}} \quad (7-135)$$

It is at once apparent that at a frequency f_1 , defined by

$$\omega_1 C_c R_2 = 1 \quad (7-136)$$

the magnitude of the low-band output voltage is 0.707 times its mid-band value, and f_1 is the lower half-power frequency, which was previously defined for the resistance-coupled amplifier. The variation of magnitude and phase of E_{ol} in the low-band will be identical with that plotted for A_l/A_m and θ_l in Fig. 7-3.

Notice that the factor α is similar in form to the corresponding factor which controls the response in the high-band, thus several of the high-band concepts may be carried over to the present case. For example, on a steady-state basis we immediately see that f_1 should be

made sufficiently low so that all the video frequencies to be amplified are passed with only a small amount of distortion. f_1 is lowered by raising the R_2C_c product. In practical amplifiers this product may not be raised without limit because of two primary factors: (1) In any vacuum tube the value of R_2 , the grid leak resistor, is limited by the flow of grid current. For most receiving tubes the manufacturers specify an upper limit in the vicinity of 500 kilohms. (2) As the capacitance of C_c is raised, the physical size of the condenser and its shunt capacitance to ground increase. We have already seen that this latter capacity to ground adds to C_s which is the controlling factor in the high-band response of the amplifier. Furthermore, as C_c increases in size, it will generally have a larger leakage current. Typical values of components are $C_c = 0.25 \mu\text{f}$ and $R_2 = 500$ kilohms, which give a lower half-power frequency of approximately 1.25 cycles per second. On a steady-state basis the response to, say, a 60-cycle video frequency seems excellent; we shall see later that on a transient basis an even lower value of f_1 should be used; some form of compensation is desirable.

7-19. Cathode and Screen Degenerative Effects

If the reader refers to the various circuit diagrams of video amplifiers in this chapter, he will notice in each case that the cathode of the amplifier tube is returned directly to ground, the tacit assumption being made that fixed bias rather than self-bias is used.

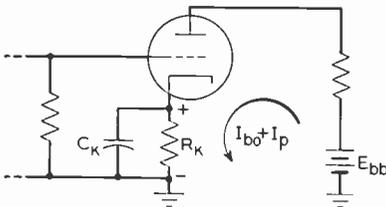


Fig. 7-22. Self-bias is developed by I_{bo} flowing through R_K .

As a practical matter, it is more desirable to develop bias across a shunt combination of a resistor R_K and a condenser C_K , located between cathode and ground as shown in Fig. 7-22. The d-c

component of total plate current, I_{bo} , produces a d-c voltage drop across R_K of proper polarity to make the grid negative relative to the cathode. The magnitude of this bias voltage is adjusted by proper selection of R_K . The function of the shunt condenser C_K is to make the impedance of the combination negligible to the a-c component of plate current, I_p . Use of this self-bias network eliminates the need for separate bias batteries or multiple bias taps on the power supply

bleeder and affords an additional advantage in that it permits a larger value of R_2 to be used as compared to the allowable value when fixed bias is utilized.³⁶ This reflects on circuit operation by providing a lower value of uncompensated half-power frequency, f_1 .

The use of the self-biasing network is not without disadvantages, for at the lower frequencies where the reactance of C_K increases, Z_K is no longer negligible. When this is true, the a-c voltage drop $I_p Z_K$ becomes significant and acts to oppose the applied voltage E_i . A condition of degeneration is present and the bias network serves to decrease the output voltage below its calculated value. Terman³⁷ has shown that the degenerative effect of the bias network may be represented by a factor γ

$$\gamma = \frac{1 + j\omega C_K R_K}{(1 + g_m R_K) + j\omega C_K R_K} \quad (7-137)$$

where the screen grid of the amplifier tube is adequately by-passed, a condition which is met quite readily in practice. The γ factor may be multiplied into (7-134) to give the low-band output voltage with the effects of both bias network degeneration and the coupling network $C_c R_2$ taken into account.

$$E_{ol} = E_{om} \alpha \gamma \quad (7-138)$$

If the α and γ factors are expressed in polar form it will be observed that both affect E_{ol} in the same direction, that is, as the signal frequency decreases, both act to lower E_{ol} and increase the positive or leading relative phase angle, θ_i . To compensate for these effects we need a network which introduces a lag and raises the output level as the frequency is lowered. We next try to find a network which exhibits these properties.

7-20. Compensation Network

In the construction of electronic equipment where several tubes are operated from a common power supply, it is the usual practice to isolate the several stages from feedback effects by the use of decoupling networks. The common form of this network is shown in Fig. 7-23 where R_3 and C_3 are the isolating elements. It is of interest to

³⁶ S. W. Seeley and C. N. Kimball, *op. cit.*

³⁷ Frederick E. Terman, *Radio Engineers' Handbook*. New York: McGraw-Hill Book Company, Inc., 1943.

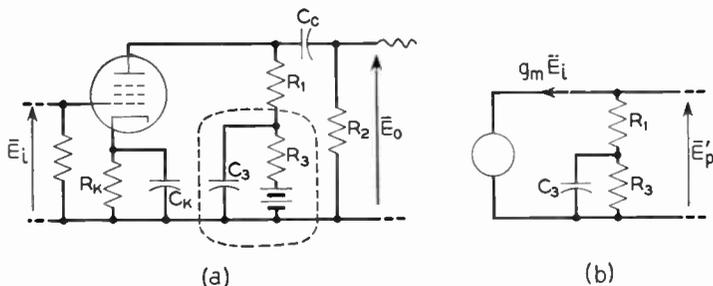


Fig. 7-23. Low-frequency compensation. (a) Basic circuit. (b) Equivalent circuit.

note that this decoupling combination has the very type of frequency response required for low-band compensation.

To show this we may consider that the $C_c R_2$ branch has negligible shunting effect on the circuit comprising R_1 , R_3 , and C_3 , which we shall term the plate load. In the mid-band, C_3 has negligible reactance and effectively shorts out R_3 ; hence the mid-band plate load is simply R_1 . At lower frequencies, however, the reactance of C_3 is no longer small enough to short out R_3 and the plate load becomes

$$Z_L = R_1 + \frac{R_3}{1 + j\omega C_3 R_3} \quad (7-139)$$

whose magnitude, of course, will increase as ω is lowered. Thus the plate decoupling network tends to compensate in magnitude for the loss in output voltage resulting from cathode degeneration and the network $C_c R_2$.

Inspection of (7-139) also shows that the $C_3 R_3$ network introduces a lag in phase angle which may be used to compensate the lead introduced by the α and γ factors. This fact may also be demonstrated with the aid of the vector diagrams shown in Fig. 7-24. At a we have the uncompensated case, and $-E_p$ is in phase with the reference vector $g_m E_i$. In the compensated case illustrated at b , the reactive component of the $C_3 R_3$ combination introduces a component of voltage normal to $g_m E_i$ and lagging. $-E_p$ also lags $g_m E_i$. We see, then, that the compensating network has a lagging angle labeled θ_δ in the diagram. As the frequency is lowered, the reactive component of Z_L increases and θ_δ increases in magnitude.

We have shown that at least the $C_3 R_3$ network operates in the right

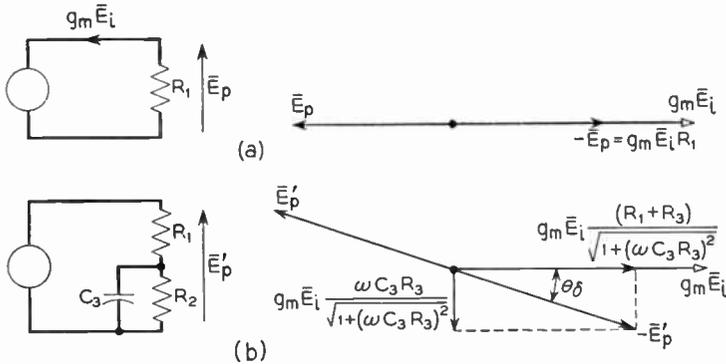


Fig. 7-24. Vector diagrams showing the effect of low-frequency compensation on phase shift. (a) Uncompensated case. (b) Compensated case.

direction to provide low-band compensation; we must now determine what specific values of C_3 and R_3 are required. To do this we must evaluate the factor δ by which the compensating network modifies the output voltage. Under the assumption of the constant current form of equivalent plate circuit, which is quite valid for voltage amplifier pentodes, the plate current remains unchanged regardless of the presence of C_3 and R_3 in the circuit. Then for the circuit of Fig. 7-23 we may write

$$E_p' = -g_m E_i Z_L = -g_m E_i R_1 \left(1 + \frac{R_3}{1 + j\omega C_3 R_3} \right) \quad (7-140)$$

$$= -g_m E_i R_1 \left[\frac{\left(1 + \frac{R_3}{R_1} \right) + j\omega C_3 R_3}{1 + j\omega C_3 R_3} \right] \quad (7-141)$$

$$= E_{om} \delta \quad (7-142)$$

We may now show how the circuit should be designed to compensate for either the α factor caused by the coupling network, or the γ factor which is introduced by the cathode bias system. The several equations given above may be combined to give the low-band output voltage of a low-band-compensated amplifier which uses self-bias.

$$E_{oi} = E_{om}\alpha\gamma\delta \quad (7-143)$$

$$E_{oi} = E_{om} \left(\frac{1}{1 - \frac{j}{\omega C_c R_2}} \right) \left[\frac{1 + j\omega C_K R_K}{(1 + g_m R_K) + j\omega C_K R_K} \right] \times \left[\frac{\left(1 + \frac{R_3}{R_1} \right) + j\omega C_3 R_3}{1 + j\omega C_3 R_3} \right] \quad (7-144)$$

7-21. Bias Circuit Compensation

The γ and δ factors in eq. (7-144) are reciprocal in form and may be canceled out completely if

$$1 + \frac{R_3}{R_1} = 1 + g_m R_K \quad (7-145)$$

and
$$\omega C_3 R_3 = \omega C_K R_K \quad (7-146)$$

Thus the compensating circuit can give *perfect* compensation of the degenerative effect of the self-bias network. The required design equations for C_3 and R_3 may be obtained from the last two equations

$$\text{and } \left. \begin{aligned} R_3 &= g_m R_1 R_K \\ C_3 &= \frac{C_K}{g_m R_1} \end{aligned} \right\} \begin{array}{l} \text{BIAS CIRCUIT} \\ \text{COMPENSATION} \end{array} \quad (7-147)$$

It must be stressed that the values of C_3 and R_3 specified by eq. (7-147) can only compensate for the self-bias network; the output voltage will still vary in accordance with the α factor. The condition is ameliorated for, as we have seen, self-biasing permits the use of higher values of R_2 , a determining factor in the variation of α with frequency.

With the effect of cathode degeneration compensated the low-band response is determined solely by the coupling network, and dividing both sides of eq. (7-144) by E_i we have

$$\frac{A_i}{A_m} = \frac{1}{1 - \frac{j}{\omega C_c R_2}} \quad (7-148)$$

and if f_1 be defined as in eq. (7-136) we have

$$\left. \begin{aligned} \frac{A_i}{A_m} &= \frac{1}{\sqrt{1 + \left(\frac{f_1}{f}\right)^2}} \\ \text{and } \theta_i &= \arctan \frac{f_1}{f} \end{aligned} \right\} \begin{array}{l} \text{LOW-BAND RESPONSE,} \\ \text{BIAS-CIRCUIT} \\ \text{COMPENSATION} \end{array} \quad (7-149)$$

These quantities are plotted in Fig. 7-3.

One final point must be stressed: the low-band compensating network has no effect on and is not affected by the presence of any of the several high-band compensating networks which may be used. In either frequency band, the circuit constants are such that operation in one band is unaffected by networks introduced to compensate in the other.

7-22. Coupling Circuit Compensation

The correction factor δ may alternatively be used to compensate for the effect of the coupling network factor, α . Inspection of these factors in eq. (7-144) shows that they are not of reciprocal form so that *perfect* compensation is not possible as it was in the last case. An excellent engineering approximation can be made to the ideal, however, for generally the circuit design will be such that

$$R_3 \geq \frac{10}{\omega C_3} \quad (7-150)$$

in the range of frequencies under consideration. The effect of this inequality on δ may be best seen by considering δ in its form shown in (7-140). From (7-150) $\omega C_3 R_3 \geq 10$, thus δ becomes

$$\delta \approx 1 + \frac{R_3}{j\omega C_3 R_3} \approx 1 - \frac{j}{\omega C_3 R_1} \quad (7-151)$$

In this form δ may be used to cancel α , provided that

$$C_3 R_1 = C_c R_2 \left. \vphantom{C_3 R_1} \right\} \begin{array}{l} \text{COUPLING-CIRCUIT} \\ \text{COMPENSATION} \end{array} \quad (7-152)$$

If the equality (7-152) is met in the design, the low-band response will be determined by the bias circuit factor. Then, taking the magnitudes of numerator and denominator of γ , we have

$$\gamma = \sqrt{\frac{1 + (\omega C_K R_K)^2}{(1 + g_m R_K)^2 + (\omega C_K R_K)^2}} \quad (7-153)$$

Again the phase angle of γ may best be obtained by rationalizing γ , thus

$$\begin{aligned}\gamma &= \frac{(1 + j\omega C_K R_K)[(1 + g_m R_K) - j\omega C_K R_K]}{(1 + g_m R_K)^2 + (\omega C_K R_K)^2} \\ &= \frac{[(1 + g_m R_K) + (\omega C_K R_K)^2] + jg_m R_K \omega C_K R_K}{(1 + g_m R_K)^2 + (\omega C_K R_K)^2} \quad (7-154)\end{aligned}$$

$$\text{and} \quad \theta_\gamma = \text{arc tan} \left[\frac{g_m R_K \omega C_K R_K}{(1 + g_m R_K) + (\omega C_K R_K)^2} \right] \quad (7-155)$$

If we define a frequency f_K at which

$$\omega_K C_K R_K = 1 \quad (7-156)$$

γ and θ_γ may be simplified and the low-band response equations become

$$\left. \begin{aligned} \frac{A_l}{A_m} &= \sqrt{\frac{1 + \left(\frac{f}{f_K}\right)^2}{(1 + g_m R_K)^2 + \left(\frac{f}{f_K}\right)^2}} \\ \text{and} \\ \theta_l &= \text{arc tan} \left[\frac{g_m R_K \frac{f}{f_K}}{(1 + g_m R_K) + \left(\frac{f}{f_K}\right)^2} \right] \end{aligned} \right\} \begin{array}{l} \text{LOW-BAND} \\ \text{RESPONSE,} \\ \text{COUPLING} \\ \text{CIRCUIT} \\ \text{COMPENSATION} \end{array} \quad (7-157)$$

It must be stressed that f_K is *not* the lower half-power frequency. Defined by (7-156) it is introduced merely to simplify calculation.

7-23. Summary

The decrease in gain and departure from linear phase shift in the low-band is the result of the coupling network and the self-biasing system. Either, not both, of these effects may be canceled by use and proper design of the compensating network shown in Fig. 7-23. Generally the bias network effect is canceled out, for the use of self-bias allows higher values of R_2 to be used which improves the low-band response of the coupling network. High- and low-band compensation networks act with mutual independence; each may be designed without consideration of the other, with the single exception that R_1 , which enters into the expressions for all three frequency bands, must be determined from the high-band considerations. A

typical amplifier compensated for both ends of the video band is shown in Fig. 7-25. Compensation is for the bias network in the lows. High response is for $K = \frac{1}{2}$ with an assumed shunt capacitance of $20 \mu\mu\text{fd.}$ and an upper cutoff frequency of approximately 4 megacycles.

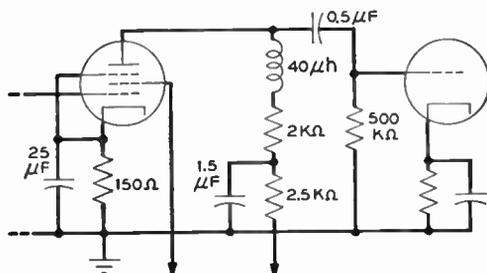


Fig. 7-25. A video amplifier employing high- and low-band compensation.

It should be stated that circuit configurations other than those described here may be used for bettering the low-band response of video amplifiers. One such circuit described by Zeidler and Noe³⁸ utilizes a circuit connection known as a pentriode amplifier, which completely cancels out the effects of screen grid and bias circuit degeneration down to zero frequency. The principal advantage of the pentriode connection is that condensers of small capacitance are used. At the time of writing the circuit has found little acceptance in commercial television equipment.

7-24. Low-frequency Transient Response

Having discussed the steady-state behavior in the low-frequency band of the compensated amplifier, we next consider its transient response. It is unfortunate that there are no simple means of handling this problem analytically, and the trend has been to use the square-wave technique. In contrast to square-wave testing for high-band response to determine rise time and overshoot where a high-frequency square wave is employed, here a square wave of the lowest video frequency, the frame frequency of 30 cycles per second under commercial standards, is used. Since the first 10 harmonic com-

³⁸ H. M. Zeidler and J. D. Noe, "Pentriode Amplifiers." *Proc. IRE*, **36**, 11 (November 1948).

ponents are the important ones, it is clear that in this case the output wave will give a measure of the low- and mid-band behavior of the amplifier under test. If viewed on an oscilloscope, the output voltage will exhibit zero rise time and overshoot, but the shape during the flat top portion of the test square wave will give an excellent indication of the low-frequency behavior. Typical types of distortion due to poor low-frequency amplifier characteristics are illustrated in Fig. 7-26. Of the several types listed, *b*, *d*, and *f* are those most

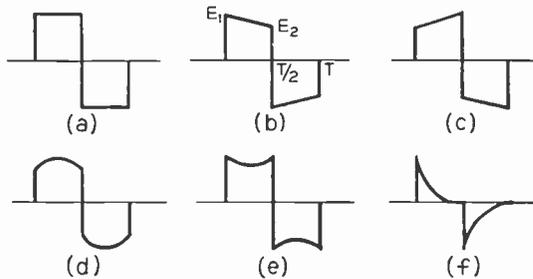


Fig. 7-26. Typical forms of low-frequency distortion. (a) Applied square wave. (b) Tilt due to leading phase shift of the fundamental signal component. (c) Tilt due to lagging phase shift of the fundamental signal component. (d) Convex rounding. The fundamental component has the proper phase but is too strong. (e) Concave rounding. The fundamental component has the proper phase but is too weak. (f) Differentiation caused by low R_2C_2 time constant.

commonly encountered. The cause of the sloping of the wave top shown at *b* may be considered in two ways, first, because of poor low frequency phase response, and second, as a result of insufficient time constant in the coupling network C_c and R_2 . The condition illustrated at *f* is an extreme case of that just mentioned.

The amount of distortion present in *b* may be conveniently expressed in terms of tilt or sag. The tilt or sag is defined as the ratio of the difference between the leading and trailing edge amplitudes to the leading edge amplitude in per cent. Thus from Fig. 7-26*b*

$$\text{tilt} = \frac{E_1 - E_2}{E_1} 100 \text{ per cent} \quad (7-158)$$

It is interesting to see once again how the steady-state criteria for

distortionless amplification fail to give a good indication of transient behavior. We have observed that a coupling circuit of $R_2 = 500$ kilohms and $C_c = 0.25 \mu\text{f}$ gives a steady-state half-power frequency of 1.25 cycles. It would seem, therefore, that a 30-cycle square wave would be passed with negligible distortion by this network because all the components of the square wave exceed ten times the half-power frequency. A simple calculation shows that distortion is present and in fact there is a tilt of 13.3 per cent in the amplified wave. This may be seen from the following calculations:

At time t , the voltage across R_2 will be

$$e_{R_2} = g_m e_i R_1 \epsilon^{-t/R_2 C_c} \approx g_m e_i R_1 \left(1 - \frac{t}{R_2 C_c} \right)$$

The half-period of the 30-cycle square wave is

$$t = \frac{1}{2(30)}$$

and from (7-158) the tilt will be

$$\text{tilt} = \frac{t}{R_2 C_c} = \frac{1}{2(30) (5 \times 10^5)(2.5 \times 10^{-7})} = 13.3 \text{ per cent}$$

Practically speaking, tilt may be reduced appreciably by the use of synchronized clamping, which is discussed in Chapter 13. In general an over-all tilt of 10 per cent may be tolerated in a cascaded chain of video amplifiers if small detail is also present in the picture signal. When no detail is present, a smaller tolerance of 2 to 3 per cent must be held if no variation in the background level is to be discernible.

The condition of rounding which is illustrated at d in Fig. 7-26 is among other factors the result of attenuation of the higher frequency components in the square waves which attenuation is caused by the shunting effect of the compensating condenser C_3 in Fig. 7-23. In other words, the rounding is caused by a fundamental component which has the proper phase but is of too great amplitude relative to the higher frequency components.³⁹ Rounding may be eliminated by shunting C_c by a resistor in series with a large blocking condenser.

In general the adjustment of low-frequency transient response may best be accomplished experimentally by using the square-wave tech-

³⁹ M. J. Larsen, "Low-Frequency Compensation of Video Frequency Amplifiers." *Proc. IRE*, **33**, 10 (October 1945).

nique. As in the high-band case, it is difficult to predict the transient behavior of the amplifier from its steady-state characteristics alone.

7-25. Clamping⁴⁰

Up to this point in our discussion of video amplifiers we have neglected the d-c, or average, component of the video signal. Clearly this component will not be passed by any of the amplifiers which have been described because they all utilize a blocking condenser C_c which serves to isolate the d-c plate voltage of one stage from the grid of the following stage. The need for preserving the average signal component was established in the last chapter; we must now consider how it may be restored at any point in the video chain.

We have previously observed that the loss of the average signal component in the video amplifiers is not serious⁴¹ once the black level pedestals have been established. We now seek a circuit which can recognize and clamp each pedestal to a fixed voltage level. This process will automatically reinsert the missing d-c component. Consider the clamping circuit shown in Fig. 7-27, where the output of a compensated video amplifier stage V_1 feeds the grid of an *unbiased* stage V_2 . The required clamping action is furnished by the diode V_3 whose polarity is that required for clamping a black-positive signal on V_2 .

Consider the action which takes place when the video signal is first applied to the circuit. Since V_2 has zero bias, the average value of the voltage applied to V_2 is zero. As each blanking pulse of duration τ_b arrives, the grid is driven positive and V_3 , the clamping diode, conducts. The conduction current allows C_c to charge, its left side in the diagram becoming positive. From the circuit diagram the time constant of the charging circuit is

$$\begin{aligned} T_c &= \text{charging time constant} & (7-159) \\ &= C_c(R_g + R_d) \end{aligned}$$

where

$$\begin{aligned} R_g &= \text{equivalent generator resistance} \\ R_d &= \text{equivalent static plate resistance} \\ &\quad \text{of } V_3 \text{ during conduction} \end{aligned}$$

and

$$R_2 \gg R_d$$

⁴⁰ Foster and Rankin, *op. cit.*

⁴¹ See Section 6-16.

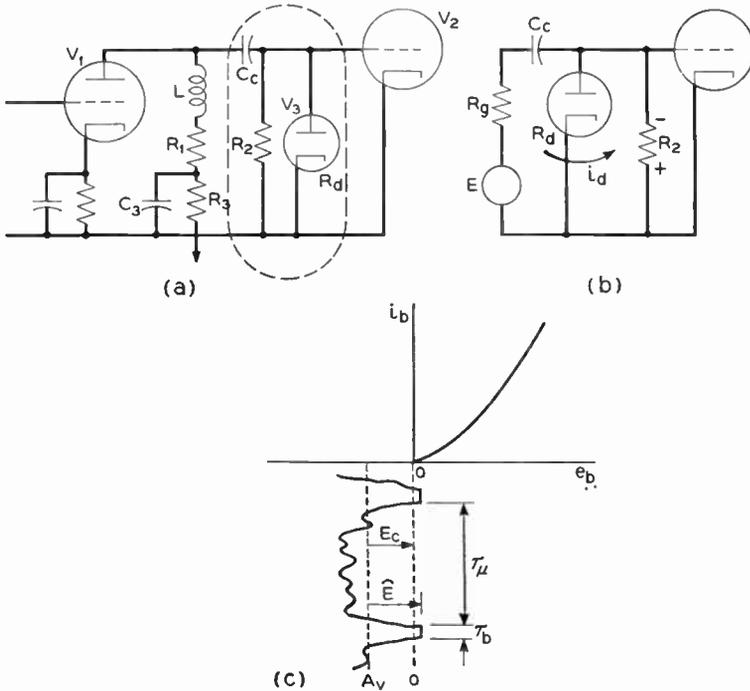


Fig. 7-27. Diode clamping. (a) Basic circuit. (b) Equivalent circuit. (c) Clamping diode characteristic.

The charging process will continue for the duration of the blanking pulse, τ_b .

Remembering that the average value of the grid signal on V_2 is zero, we see that at the end of τ_b , the applied voltage must swing negatively. V_3 ceases to conduct and C_c discharges as shown in the diagram. The discharge time constant is

$$T_d = \text{discharging time constant} = C_c(R_g + R_2) \quad (7-160)$$

The discharge current, i_d , flows in such a direction through R_2 that a negative bias is produced on the grid of V_2 .

At the end of τ_u another blanking pulse occurs, V_3 conducts C_c charges, and so on, the complete cycle repeating itself over and over again. We must note the following fact, however: $T_d \gg T_c$ and the condenser tends to build up a charge which, in turn, maintains

the bias across R_2 . Thus after several cycles of operation have taken place, sufficient bias is developed that each blanking pulse drives the plate of V_3 just enough positive to allow the charge lost by C_c during τ_u to be replaced during τ_b . In other words, a repetitive steady-state condition is reached which allows pedestal level to be a constant fraction of a volt positive on the plate of V_3 . Pedestal level has been clamped to a fixed level, and the proper bias is developed across R_2 . We shall demonstrate that the bias developed in this manner is proportional to the average signal level, black being the reference.

It may clarify matters to restate the operation in a different manner. The clamping circuit of Fig. 7-27a may be viewed as a half wave rectifier of the shunt type where the rectifying element is across, rather than in series with, the load resistance R_2 . In such a circuit the d-c voltage developed across the load is proportional to the positive peak value of the applied signal, which in this case is not sinusoidal. The action is also modified by the presence of C_c which permits the video signal to be superimposed on the developed bias as far as V_2 is concerned.

The value of bias voltage developed across R_2 may be calculated subject to a number of justifiable, simplifying assumptions.

(1) The average value of the original picture remains constant over several successive lines.

(2) $\tau_b \ll T_c$ so that during τ_b the charging current remains constant at its Ohm's law value.

(3) $\tau_u \ll T_d$ and i_d remains at its Ohm's law value.

(4) The change in condenser voltage, E_c , is negligibly small so that E_c may be assumed constant.

(5) The d-c plate voltage of V_1 remains constant and hence need not be considered in the calculations.

(6) $R_d \ll R_2$.

The notation used is that of Fig. 7-27b, thus \widehat{E} is the pedestal voltage measured from the average value of the signal.

Subject to these assumptions the condenser charging current is determined by the difference between \widehat{E} and E_c or

$$i_c = \frac{\widehat{E} - E_c}{R_d + R_2} = \text{constant} \quad (7-161)$$

The charge acquired by C_c during the charging interval τ_b is

$$Q_c = i_c \tau_b = \frac{\widehat{E} - E_c}{R_g + R_d} \tau_b \quad (7-162)$$

During discharge the source voltage is at its average value so that the discharge current is

$$i_d = \frac{E_c}{R_g + R_2} = \text{constant} \quad (7-163)$$

and the charge lost by C_c during τ_u is

$$Q_d = i_d \tau_u = \frac{E_c}{R_g + R_2} \tau_u \quad (7-164)$$

Q_c and Q_d are each small quantities, but if a steady-state condition prevails, it must be true that the charge gained by C_c during τ_b must be equal to the charge it loses during τ_u ; hence

$$Q_c = Q_d \quad (7-165)$$

Substituting (7-162) and (7-164) into this identity and clearing we get

$$E_c = \frac{\widehat{E}}{1 + \left(\frac{R_g + R_d}{R_g + R_2} \right) \frac{\tau_u}{\tau_b}} \quad (7-166)$$

which shows that the condenser voltage is proportional to the pedestal level as measured from the average value of the signal. Remember all quantities in the denominator are constant.

The steady-state bias voltage developed across R_2 is

$$E_{cc} = -i_d R_2 = -E_c \frac{R_2}{R_g + R_2} \quad (7-167)$$

and this bias is proportional to E_c . It follows, therefore, that the bias on V_2 is proportional to \widehat{E} which is the difference between average level and pedestal. The d-c component has been reinserted on the grid of V_2 by the clamping circuit.

For purposes of calculation it should be pointed out that R_g is approximately equal to R_1 , the plate load resistance of the compensated video stage. Typical circuit values for a clamping circuit at the grid of a cathode-ray tube have been listed by Foster and Rankin.

$$R_2 = 1 \text{ megohm} \quad R_d = 4 \text{ kilohms} \quad C_c = 0.1 \text{ } \mu\text{f}$$

Under certain circumstances when the magnitude of the applied black-positive signal is not too great, the clamping diode

V_3 of Fig. 7-27a may be eliminated and its action replaced by the grid-to-cathode circuit of V_2 . This is possible because the grid-current-grid-voltage characteristic of V_2 is similar to the plate-current-plate-voltage characteristic of the clamping diode. Where a black-negative signal is applied, a separate clamping rectifier must be used with its polarity the reverse of that shown in the figure. For low values of applied voltage the vacuum diode may be replaced by one of the newer types of germanium rectifier elements, such as the 1N34. Use of the crystal unit results in considerable saving in cost, space, and demand on filament power.

Having seen that the d-c component of the video signal may be restored by means of a clamping circuit, we must now consider a second question in regard to clamping: At what points in the video chain must the d-c component be present in the signal?

It should be apparent that if at no other place in the system the d-c signal component must be present on the control grid of the cathode-ray tube on which the final image is reproduced. Were this not so, the final picture would not contain the proper background level of the televised scene. At least, then, clamping must occur at the kinescope control grid, or that grid must be direct-coupled to some stage which is clamped. The second alternative also delivers the correct average signal component to the kinescope.

Reasoning in reverse direction from the last paragraph one might well arrive at the conclusion that clamping is required only at those points in the circuit where an image is to be reproduced, *i.e.*, at the cathode-ray tubes. As a practical matter the judicious use of the clamping circuit at certain amplifiers in the video chain can result in substantial reduction in the design requirements of these amplifiers. This is particularly true in the high-level video stages associated with a television broadcast transmitter.

Consider the following example. We shall arbitrarily assume that the peak white-to-black swing is 4 volts on the grid of a particular stage. At *a* and *b* of Fig. 7-28 are shown signals corresponding, respectively, to a white line on an almost black background, and a solid white background. It is further assumed that the ratio of blank to unblank time is such that the average values of the signals are as indicated in the diagram. If now the amplifier is to handle these extremes of signal, in the absence of clamping it must be so designed that it will operate linearly over a grid swing of $3.7 + 3 =$

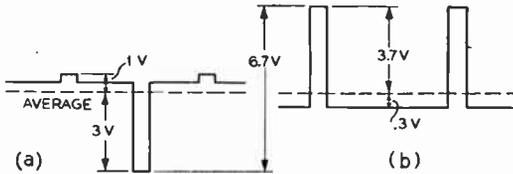


Fig. 7-28. Clamping may be used to reduce the grid swing for which an amplifier must be designed. (a) Signal for a black bar on a white background. (b) An all-white signal.

6.7 volts. If that same grid circuit is clamped, however, black is held at a constant level and the amplifier need handle a swing of only 4 volts. Notice, then, that for the figures used clamping reduces the grid swing requirements on the stage by roughly 40 per cent. In the high power stages of the video chain where linearity over a wide range of grid voltages is difficult to obtain, this grid swing reduction is extremely important. It will be seen in Chapter 13 that all the high-level video stages employ either clamping or direct coupling in order to effect this saving.

FEEDBACK VIDEO AMPLIFIERS

The problem of compensation in video amplifiers may be approached from an entirely different point of view from those which have been described in earlier sections. It is well known that negative voltage feedback may be used to decrease phase and frequency distortion in an audio amplifier. The same treatment may be extended to broad-band video amplifiers. This idea has been utilized by the Du Mont organization in the design of television video amplifiers for their image orthicon camera equipment.^{42,43} The chief advantage of the feedback video amplifier is that a single feedback network provides equalization over the entire video band. Controls are included so that the high- and low-band responses may be extended with little effect on the mid-band gain; adjustment of mid-band gain has no effect on the relative response in the highs and lows.

The basic circuit used comprises two tubes which function together and may be termed a "video doublet." Feedback is provided be-

⁴² Operating and Maintenance Manual, Du Mont Model TA-124-B Image Orthicon Chain, Allen B. Du Mont Laboratories, Inc., 1948.

⁴³ For a general reference on feedback video amplifiers see G. E. Valley, Jr., and H. Wallman, *Vacuum Tube Amplifiers*, M.I.T. Radiation Laboratory Series No. 18. New York: McGraw-Hill Book Company, Inc., 1948.

tween the tubes by a resistor, R_F , shown in Fig. 7-29a. Following the procedure set up in the first part of this chapter we shall consider the steady-state analysis of the video feedback doublet; the transient analysis of the circuit is available in the literature.⁴⁴ In the work which is to follow we shall assume that the plate resistance of each

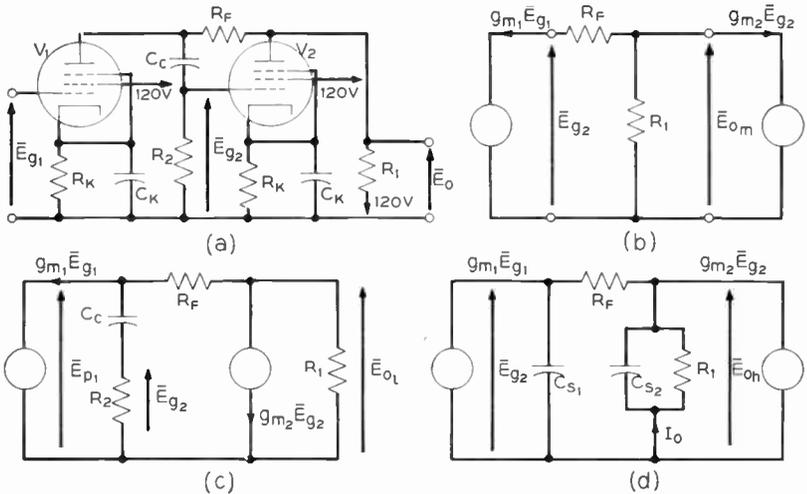


Fig. 7-29. The video doublet. Feedback is provided by R_F .
 (a) Basic circuit. (b) Equivalent mid-band circuit. (c) Equivalent low-band circuit. (d) Equivalent high-band circuit.

tube has negligible shunting effect on the plate load circuit. We shall follow the procedure used with the resistance coupled amplifier, that is, we shall consider three bands of operation, mid, low, and high. In order to observe the compensating effect of the feedback network we shall force the results into a form similar to those derived for the R - C amplifier.

7-26. Mid-band

We define the mid-frequency band as that band in which C_c and the shunt capacitances have negligible effect on the operation. The equivalent mid-band circuit subject to this definition is given in Fig. 7-29b where it is assumed that $R_2 \gg R_F + R_1$. From the diagram

⁴⁴J. H. Mulligan Jr. and L. Mautner, "The Steady State and Transient Analysis of a Feedback Video Amplifier," *Proc. IRE*, **36**, 5 (May 1948).

$$E_{om} = -(g_{m1}E_{\theta 1} + g_{m2}E_{\theta 2})R_1 \tag{7-168}$$

and

$$E_{\theta 2} = E_o - g_{m1}E_{\theta 1}R_F \tag{7-169}$$

whence $A_m = \frac{E_{om}}{E_{\theta 1}} = -g_{m1}R_1 \left[\frac{1 - g_{m2}R_F}{1 + g_{m2}R_1} \right]$ MID-BAND $\tag{7-170}$

Comparison of this equation with (7-26) shows that the effect of the second tube and the feedback network is to multiply the output of the first tube by the factor $(1 - g_{m2}R_F)/(1 + g_{m2}R_1)$ which is less than unity. It is characteristic of inverse feedback that the amplification is reduced.

Two interesting facts may be observed from (7-170). First, the mid-band gain may be controlled by the bias on V_1 which determines g_{m1} . Second, for typical values $R_1 = 1$ kilohm and $R_F = 4$ kilohms the mid-band gain is fairly well isolated from variations in g_{m2} .

7-27. Low-band

As the frequency of the applied signal drops below the mid-band limit, the reactance of C_c becomes significant and $E_{\theta 2}$ is determined by the voltage divider action of C_c and R_2 are shown in Fig. 7-29c. From the diagram

$$E_{\theta 2} = \frac{R_2}{R_2 - \frac{j}{\omega C_c}} E_{p1} = (E_o - g_{m1}R_F) \frac{R_2}{R_2 - \frac{j}{\omega C_c}} \tag{7-171}$$

Because of the high impedance of the R_2C_c branch E_o is given by (7-168) and may be written E_{o1} since we are analyzing the low-band circuit. Combination of these two equations gives:

$$E_{o1} \left[1 + g_{m2}R_1 \left(\frac{R_2}{R_2 - \frac{j}{\omega C_c}} \right) \right] = -g_{m1}E_{\theta 1}R_1 \left[1 - g_{m2}R_F \left(\frac{R_2}{R_2 - \frac{j}{\omega C_c}} \right) \right] \tag{7-172}$$

which may be reduced to the form

$$A_l = \frac{E_{o1}}{E_{\theta 1}} = -g_{m1}R_1 \left(\frac{1 - g_{m2}R_F}{1 + g_{m2}R_1} \right) \left[\frac{1 - \frac{j}{(1 - g_{m2}R_F)\omega C_c R_2}}{1 - \frac{j}{(1 + g_{m2}R_1)\omega C_c R_2}} \right] \tag{7-173}$$

This expression may be simplified if we note that the factors outside of the brackets are precisely the complex mid-band amplification. The quantity within the brackets may also be simplified subject to an approximation. For typical circuits constants and an assumed g_{m2} of 5000 micromhos the reactive term in the numerator drops out. Hence we have approximately that

$$\left. \begin{aligned} \frac{A_1}{A_m} &\approx \frac{1}{\sqrt{1 + \frac{1}{[\omega C_c R_2 (1 + g_{m2} R_1)]^2}}} \\ \text{and } \theta_1 &\approx \arctan \left[\frac{1}{\omega C_c R_2 (1 + g_{m2} R_1)} \right] \end{aligned} \right\} \text{LOW-BAND (7-174)}$$

Equations (7-174) are deceptive in the following respects: (1) Self-bias effects are neglected, and (2) C_c and R_L are *within* the doublet. If the output across R_1 is coupled to another similar pair through another R - C combination, the response will be affected by an α term defined in (7-135). The equation does show how the feedback provides compensation for R_2 and C_c within the doublet. In effect the time constant $C_c R_2$ is increased by a factor $(1 + g_{m2} R_1)$, which is subject to control by varying the bias on V_2 . We have already seen that the mid-band gain is fairly well isolated from changes in g_{m2} ; hence this low-band response control has little effect on the gain.

7-28. High-band

As the frequency of the applied signal is raised above the mid-band limits, the effect of the various shunt capacitances of the circuit must be reckoned with. We shall adopt the following notation:

$$\left. \begin{aligned} C_{o1} &= \text{output capacitance of } V_1 \\ C_{i2} &= \text{input capacitance of } V_2 \\ C_{s1} &= C_{o1} + C_{i2} \\ C_{s2} &= \text{total shunt capacitance across } R_1 \\ &= \text{sum of } V_2 \text{ output and input capacitance} \\ &\quad \text{of the following stage.} \\ X_1 &= \frac{1}{\omega C_{s1}} \quad X_2 = \frac{1}{\omega C_{s2}} \quad Z_1 = \frac{R_1}{1 + j\omega C_{s2} R_1} \end{aligned} \right\} (7-175)$$

The equivalent high-band circuit is given in Fig. 7-29*d*. By inspection of the diagram we see that the output current is no longer the sum of the constant currents $g_{m1}E_{g1}$ and $g_{m2}E_{g2}$ because of the shunt path offered by C_{s1} whose reactance must be reckoned with in the high-band of frequencies. The circuit may best be analyzed by use of the superposition theorem; thus I_o is the sum of the components of current flowing through Z_1 resulting from, respectively, $g_{m1}E_{g1}$ acting alone and $g_{m2}E_{g2}$ acting alone. Hence,

$$I_o = \frac{g_{m1}E_{g1}(-jX_1) + g_{m2}E_{g2}(R_F - jX_1)}{R_F + Z_1 - jX_1} \tag{7-176}$$

where

$$E_{g2} = E_{oh} - R_F \left[\frac{g_{m1}E_{g1}(-jX_1) - g_{m2}E_{g2}Z_1}{R_F + Z_1 - jX_1} \right] \tag{7-177}$$

and

$$E_{oh} = -I_o Z_1 \tag{7-178}$$

The last three equations may be combined and after considerable algebraic manipulation yield the result

$$A_h = \frac{E_{oh}}{E_{g1}} = -g_{m1}R_1 \left(\frac{1 - g_{m2}R_F}{1 + g_{m2}R_1} \right) \times \left\{ \frac{1}{\left[1 - \frac{R_F R_1}{(1 + g_{m2}R_1)X_1 X_2} \right] + \frac{j}{(1 + g_{m2}R_1)} \left(\frac{R_F + R_1}{X_1} + \frac{R_1}{X_2} \right)} \right\} \tag{7-179}$$

Again the last expression may be simplified by noting that the expression outside the braces is A_m . Furthermore the term

$$R_F R_o / (1 + g_{m2}R_1) X_1 X_2$$

is very small in comparison to unity, thus we have the approximate equations

$$\left. \begin{aligned} \frac{A_h}{A_m} &\approx \frac{1}{\sqrt{1 + \left\{ \left(\frac{\omega}{1 + g_{m2}R_1} \right) [(R_F + R_1)C_{s1} + R_1 C_{s2}] \right\}^2}} \\ \theta_h &\approx -\arctan \left\{ \left(\frac{\omega}{1 + g_{m2}R_1} \right) [(R_F + R_1)C_{s1} + R_1 C_{s2}] \right\} \end{aligned} \right\} \begin{array}{l} \text{HIGH-} \\ \text{BAND} \\ (7-180) \end{array}$$

From the last two equations we note that the degradation of high-band response in the doublet is the result of the shunting effect of C_{s1}

on $(R_F + R_1)$ and of C_{s2} on R_1 . The feedback compensating network acts to lower these R - C time constants by the factor $(1 + g_{m2}R_1)$, thereby producing a wider response. Notice that g_{m2} controls the high-band characteristics. The effect of bias on the circuit may be summarized as follows: V_1 bias sets mid-band gain with no effect on high and low response; V_2 bias controls response in the end bands with little effect on the mid-band gain.

7-29. The Cathode Follower

It is common practice in television work to monitor the operation of a picture system by checking voltage wave forms at critical points in the equipment. In order that the test equipment cause minimum disturbance of the channel under test it is imperative that the test voltage be sampled with a device of high input impedance. Consequently we seek some sort of circuit which has a high driving-point impedance and at the same time is capable of operating throughout the entire frequency range of the signal under test.

Another situation encountered in most television systems is that the video signal must be transmitted over relatively long lengths of cable. It is well known that optimum transmission over a line occurs when both the generator and receiving-end impedances are identical to the characteristic impedance, Z_o , of the line. The lines generally used in television work are of the coaxial type and have characteristic impedances in the range from 50 to 300 Ω and which, for practical purposes, may be considered as pure resistance. In order to feed such a cable properly we desire some device which has a low internal impedance of the same order of magnitude as the range given above.

Both of these desired characteristics, high input impedance and low output impedance, are available in the cathode follower circuit, in fact, the latter may be described as a broad-band impedance-matching device. In common with devices which provide a step-down of impedance level, the cathode follower has a gain of less than unity. These characteristics of the circuit will now be analyzed.

The basic circuit of the cathode follower is shown in Fig. 7-30. We have already seen in the last chapter that the input impedance is raised by the degeneration caused by the cathode-to-ground impedance. We consider next the gain and output impedance of the circuit.

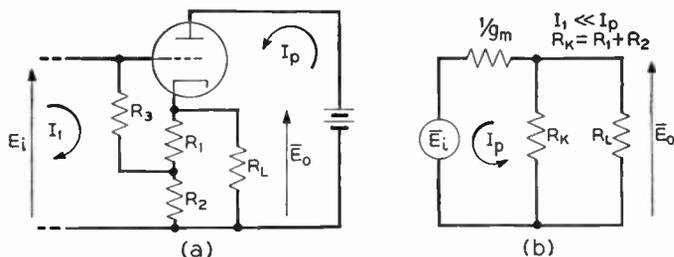


Fig. 7-30. The cathode follower. (a) Basic circuit. (b) Equivalent circuit.

Considerable work may be saved if we utilize the results of the last chapter. Let us define

$$R_K = R_1 + R_2 \tag{7-181}$$

Then replacing $(R_1 + R_2)$ of eq. (6-70) by the equivalent load $R_K R_L / (R_K + R_L)$ we have for the plate current

$$\begin{aligned} I_p &\approx \frac{\mu E_i}{r_p + \mu \frac{R_K R_L}{R_K + R_L}} \quad \text{for } \mu \gg 1 \\ &\approx \frac{E_i}{\frac{1}{g_m} + \frac{R_K R_L}{R_K + R_L}} \end{aligned} \tag{7-182}$$

The constant-voltage equivalent circuit for eq. (7-182) is given in Fig. 7-30b. The circuit amplification is

$$A = \frac{E_o}{E_i} = + \frac{I_p \frac{R_K R_L}{R_K + R_L}}{E_i} \approx \frac{R_K R_L}{R_K + R_L} \left(\frac{1}{g_m} + \frac{R_K R_L}{R_K + R_L} \right)^{-1} < 1 \tag{7-183}$$

By direct application of Thevenin's theorem to the equivalent circuit we see that the source impedance seen by R_L is

$$Z_{out} = \frac{R_K}{g_m \left(\frac{1}{g_m} + R_K \right)} = \frac{R_K}{1 + g_m R_K}, \tag{7-184}$$

the output impedance of the cathode follower. This result may be verified by the more conventional method of replacing R_L by a volt-

age E delivering a current I . With $E_i = 0$, Z_{out} will be the ratio of E to I .

We see from eq. (7-184) that the load, R_L , always sees an impedance equal to the parallel combination of $1/g_m$ and R_K , which of course will be less than $1/g_m$. A consideration of the range of transconductance available in commercial tube types shows that a large range of low output impedances is available. The constants for several typical tubes are given in Table 7-1.

TABLE 7-1

| Tube | g_m micromhos | $1/g_m$ ohms |
|--------------|-----------------|--------------|
| 6AC7 | 9000 | 110 |
| 6L6 | 6000 | 167 |
| 6A3 | 5250 | 190 |
| 6L6 (triode) | 4700 | 213 |
| 6SN7 | 3000 | 333 |
| 6SK7 | 2000 | 500 |
| 6SF5 | 1500 | 666 |

It can be seen from the table that a considerable choice in low output impedance may be had by the proper choice of R_K , which, it must be remembered, shunts $1/g_m$. The general design procedure calls for choice of a tube for which $1/g_m$ is larger than the Z_o of the cable being matched. R_K is then chosen to provide the proper impedance match, its components R_1 and R_2 being designed to furnish the correct bias for the stage. Consider the following illustrative example.

A cathode follower is to be designed to work into a 72-ohm coaxial cable. We choose a 6AC7 for the tube to be used because of its low $1/g_m$ value. The value of R_K required to match the tube to the cable is

$$\frac{1}{R_K} = \frac{1}{72} - \frac{1}{110}$$

$$R_K = 211 \text{ ohms}$$

Proper bias adjustment requires that we know the length and loop resistance per unit length of the cable. The reason for this is that to the a-c components of signal the properly terminated cable appears as a fairly constant resistance, Z_o , which is independent of the length of cable in use. The d-c component of current, however,

flows through the d-c cable resistance plus the terminating resistance Z_o . R_1 and R_2 must be adjusted with this in mind. The entire cable system may be isolated from the bias network by the insertion of a blocking condenser at x in Fig. 7-31. If the low-band response is to be maintained, an extremely large condenser will be required since Z_o is low.⁴⁵

In the analysis above we have neglected the effect of shunt capacitance on the output voltage. This is justified in most cases because of the low value of load resistance. For example, in the problem just cited the total a-c load resistance between cathode and ground will be R_K and Z_o in shunt, or 54 ohms. With a total shunt capacitance of, say, $20 \mu\mu\text{f}$ the upper half-power frequency will be roughly 150 megacycles; obviously there is no problem of poor high-frequency response.

On long runs of coaxial cable, the response of the cable is not uniform over the entire pass band. Compensation for this effect may be provided by a conventional lattice-type equalizer network.

We have seen that the cathode follower serves as a high-input low-output impedance device, capable of operation over a wide frequency range. In describing the over-all television system in subsequent chapters we shall make frequent use of the circuit. Since the output impedance of the cathode follower is low, its step function or transient response is usually quite satisfactory. Treatment of the transient response is available in the literature.⁴⁶

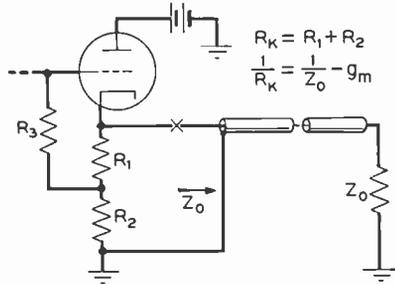


Fig. 7-31. A cathode follower is used to match a coaxial cable.

⁴⁵ A number of nomographs covering various aspects of cathode follower design are available in the literature. See, for example, M. B. Kline, "Cathode Follower Nomograph." *Electronics*, 22, 6 (June 1949).

⁴⁶ B. Y. Mills, "Transient Response of Cathode Followers in Video Circuits." *Proc. IRE*, 37, 6 (June 1949).

CHAPTER 8

CLOSED TELEVISION SYSTEMS

The television system is unique in electrical communication in that it extends the range of vision of the observer. This characteristic of the system leads to a wide range of applications in telemetering and monitoring of operations which take place in remote or dangerous locations. For example, television has provided a safe means for observers to watch nuclear reactions which of necessity must be confined within a protective enclosure. In the field of rocket-motor development the use of an insulated camera pickup assembly within the high-temperature test chamber has eliminated the need for observers to check motor operation through a small, laminated glass porthole in the chamber wall; the viewing may be done at a convenient location with the aid of 10-in. or larger cathode-ray tubes.¹

A suitable televising system permits the monitoring of several operations that take place at different points. Emerson Radio has developed such a video chain that enables a plant manager in his office to view any one of a number of key production points scattered throughout the factory.² Choice of the location viewed is provided by a master switching system which connects the monitoring receiver to any of several camera equipments.

In the field of telemetering, television, in contrast to the several servomechanism or follow-up devices, provides a remote meter indication with zero error. With the camera tube viewing the actual instruments, the observer literally "sees" these instruments whether they indicate voltage, current, or pressure, temperature, or other nonelectrical quantities. A system of this type built by Farnsworth

¹ "Television Monitors Dangerous Operations," *Electronics* (Tubes at Work, Vin Zeluff, ed.) **21**, 4 (April 1948).

² "Television System for Industrial Applications," *Electronics* (Tubes at Work, Vin Zeluff, ed.), **20**, 5 (May 1947). See also R. C. Webb and J. M. Morgan, "Simplified Television for Industry," *Electronics*, **23**, 6 (June 1950).

Research Corporation has been used to monitor boiler water gauges located up to 400 feet from a power station control room.³

These and many other similar applications show that there is an increasing demand for television in fields other than entertainment. In each of these applications the subject matter is of interest to one or at the most only a few observers; hence they are generally cable-connected systems, no radio transmitter is used, and the cameras are connected directly to the viewing equipment by wire facilities. We shall, then, call such a cable-connected system a closed television system.

The closed television system just defined is not necessarily required to produce a picture of high entertainment value. We shall find that in its development many compromises in picture quality may be made in the interests of economy and simplicity of operation. We shall also see that such systems tend to follow two trends in design: signal and synchronization information are kept separate and fed over different cables, or they are combined into a "composite video signal" which requires a single signal pair. The chief purpose of the present chapter is to show how the various components which have been studied may be combined into a closed television system.

8-1. Type I—Closed System

We propose to set up a closed television system in broad outline. We must first determine some of the standards around which the system is to be built. In the interests of simplicity we shall try to use cathode-ray devices which employ electrostatic deflection. Our first choice is the camera tube and a logical one is the type 1847 iconoscope or its more recent commercial counterpart, type 5527, both of which are of the electrostatic type and require no keystone correction. The 5527 is capable of 250 line resolution, can operate with illumination from roughly 600 watts of incandescent lighting, and has the added advantage of requiring voltages of less than 1 kilovolt for satisfactory operation. An additional feature of the 5527 iconoscope which leads to system simplification is that its signal plate is direct-coupled to the output circuit which aids the low-frequency response in the system.

³ R. W. Sanders, Industrial Television, *Radio-Electronic Engineering Edition of Radio and Television News*, **12**, 2 (February 1949). See also R. W. Sanders, Closed-Circuit Industrial Television, *Electronics*, **23**, 7 (July 1950).

We shall assume that interlaced scanning is not required, but a 60-cycle frame frequency is chosen to minimize flicker effects. No attempt will be made to lock in the vertical sweep with power-line frequency. The standard aspect ratio of 4 to 3 is to be used because it coincides with the shape of the iconoscope mosaic. Notice, then, that we have the following requirements:

$$\begin{aligned}
 A &= \text{aspect ratio} = 4 \text{ to } 3 \\
 f_p &= 60 \text{ cycles} \\
 n &= 250 \text{ lines} \\
 f_l &= \text{line frequency} = n f_p = 15 \text{ kilocycles}
 \end{aligned}$$

reasonable values for p_h and p_v , the horizontal and vertical flyback ratios, respectively, are 1 to 9 and 1 to 19. We may now set up the over-all block diagram for our closed system, which is given in Fig. 8-1.

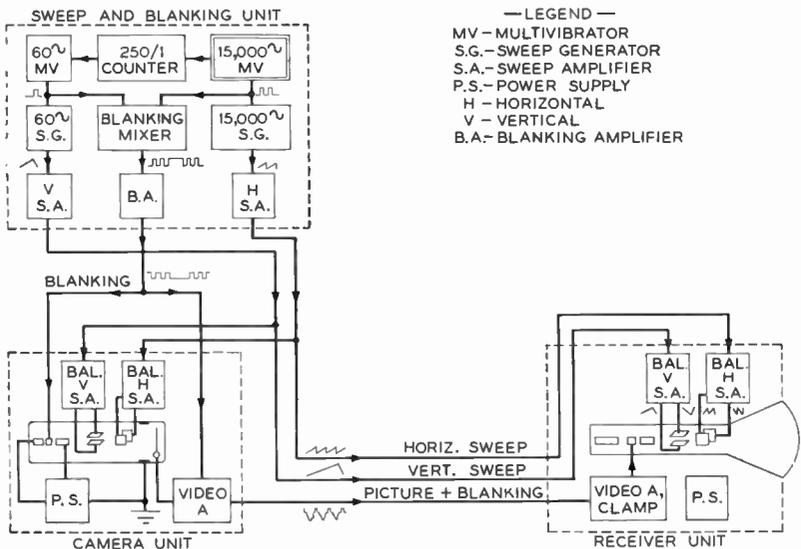


Fig. 8-1. Type I closed television system. Common sweep voltages are fed to both ends of the system.

In Chapter 1 we saw that the three basic units of a television system are the pick-up unit, the reproducing unit, and the link interconnect-

ing them. Subsequent work showed the need for synchronizing the sweep at both ends of the system, a requirement that may be met by providing a master synchronizing or sweep-voltage source. These components may be identified in the diagram as we discuss each one.

The heart of the whole sweep system is a multivibrator, free-running at the line frequency of 15,000 cycles. Following through the horizontal sweep system we see that this multivibrator triggers a saw-tooth sweep generator whose output is amplified and fed to the horizontal sweep line. Normally coaxial cable would be used for the various signals; a single line is shown on the diagram to avoid confusion.

In order for the vertical and horizontal sweeps to be locked in with each other, provision is made in the diagram for the master multivibrator to synchronize the 60-cycle multivibrator which heads the vertical deflection chain. A step down ratio of 250 to 1 is provided in the sync signal by a frequency-dividing or counter circuit located between the two multivibrators. The operation and analysis of the counter circuit is given in Chapter 11. For the moment we may summarize its operation by stating that it delivers one output pulse for every 250 pulses applied to its input. It therefore serves to reduce the line multivibrator output to a 60-cycle synchronizing pulse which may be used to lock in the low-frequency system. The 60-cycle multivibrator feeds a sweep-generation and -amplifier chain similar to that already mentioned. Any number of variations of the basic circuit are possible, of course. For example, a multivibrator and its sweep generator may be combined into the circuit shown in Fig. 4-22*b*. We have already developed the necessary design equations for these units, which may be evaluated quite readily for the standards which have been listed. Our next consideration is of the means provided for generating the blanking signals.

The third chain in the sweep and blanking unit furnishes the necessary composite blanking signal which blanks out the iconoscope trace and the video signals during the vertical and horizontal flyback intervals. In the interests of simplicity we shall assume the blanking and flyback intervals to be identical. The various blanking signals that are required are shown for an interval of time during which the relatively long vertical blanking pulse occurs in Fig. 8-2. The various time quantities involved may be calculated as follows.

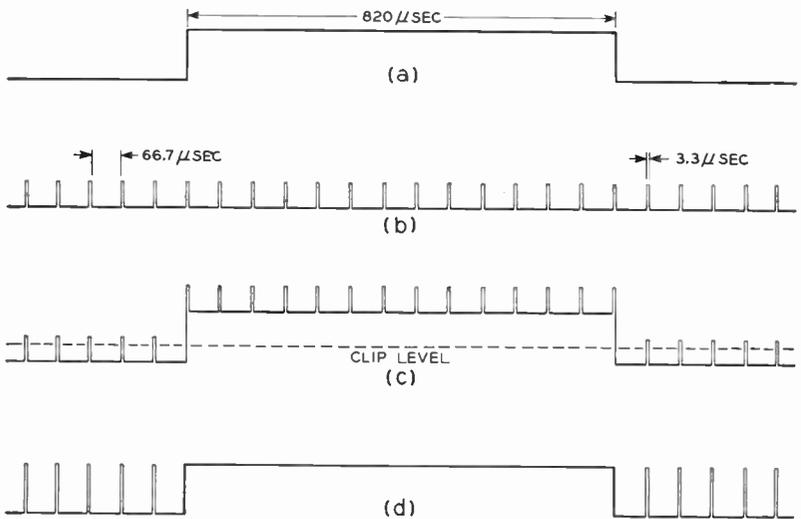


Fig. 8-2. Development of the composite blanking system.

For the vertical system

$$\begin{aligned}
 V_p &= \text{frame interval} = \frac{1}{f_p} = \frac{1}{60} \\
 &= 16.67 \text{ milliseconds}
 \end{aligned} \tag{8-1}$$

From equation (3-26)

$$V_p = (\tau_s)_v + (\tau_f)_v$$

and

$$p_v = \frac{(\tau_f)_v}{(\tau_s)_v}$$

$$\text{Therefore, } (\tau_s)_v = \frac{V_p}{1 + p_v} = \frac{19}{20} (16.67) = 15.85 \text{ millisecc} \tag{8-2}$$

$$\text{and } (\tau_f)_v = V_p - (\tau_s)_v = 0.82 \text{ millisecc} = 820 \mu\text{secc} \tag{8-3}$$

Similarly for the horizontal system

$$H = \text{line interval} = \frac{1}{nf_p} = 66.7 \mu\text{secc} \tag{8-4}$$

$$(\tau_s)_h = 60.03 \mu\text{secc} \tag{8-5}$$

and

$$(\tau_f)_h = 6.67 \mu\text{secc} \tag{8-6}$$

Since we have decided to let flyback and blanking times be identical in each case, (8-3) and (8-6), respectively, give the vertical and horizontal blanking intervals. These are indicated in Fig. 8-2.

It should be observed that it is useless to preserve the identities of the two separate blanking signals during the vertical blanking interval; if the vertical signal has blanked the iconoscope, certainly the superposition of another blanking signal will not affect the reproduced picture. Thus provision is made in the system to combine the signals into a composite blanking signal shown at *d* in the figure. The two signals are combined in the sweep and blanking unit rather than at the camera unit in order that an extra cable length may be eliminated.

The method of combining the two blanking waves is shown in Fig. 8-3. The two pulses which are derived from the multivibrators are fed to the control grids of V_1 and V_2 . Since these two tubes have a common plate load, they add the signals as indicated in Fig. 8-2c,

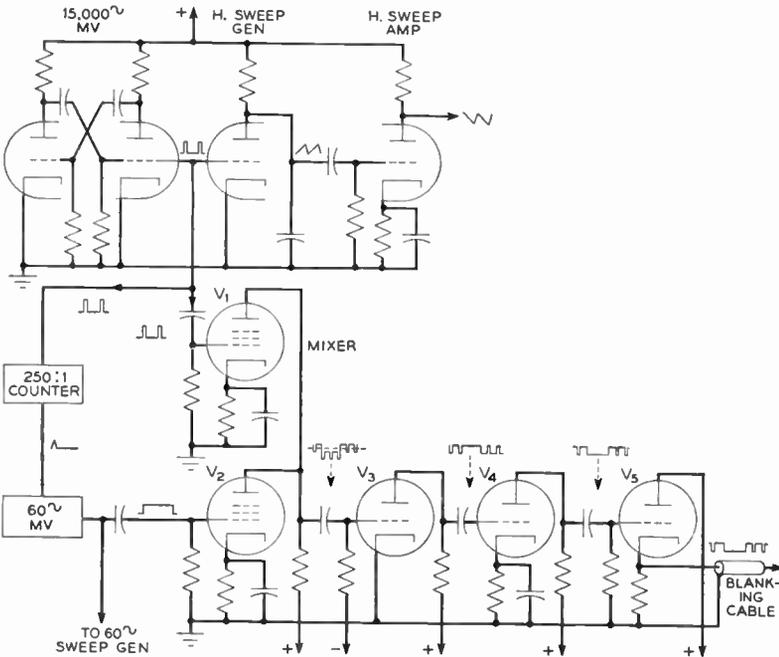


Fig. 8-3. Detail of the blanking signal mixer circuit.

and as described in connection with Fig. 6-26a.⁴ The bias on V_3 is adjusted so that all portions of the combined signal below the clip level drive V_1 below cutoff, and an amplified, positive-going composite blanking signal appears at the plate.

A negative-going signal is required to blank out the iconoscope. V_4 produces the required phase inversion and V_5 , in a cathode follower circuit, serves to match the characteristic impedance of the coaxial cable leading out from the sweep and blanking unit to the rest of the system.

It should be noticed in the block diagram (Fig. 8-1) that no blanking signal cable connects to the receiver unit. The need for this cable is eliminated by combining the blanking and picture signals, a reasonable procedure since both have the same ultimate destination, the CRT control grid. Combination of the two signals may be effected in the video amplifier chain by using the mixer circuit just described or some equivalent circuit. Since we have delivered a negative-going blanking signal to the blanking cable, some other circuit is implied. One such possible circuit is given in Fig. 8-4.

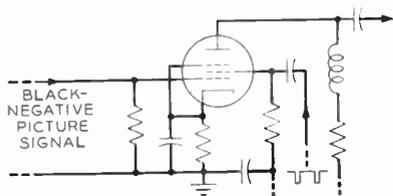


Fig. 8-4. Picture and blanking signals may be mixed in a single pentode.

The mixing is made to take place in a stage which is fed a black-negative picture. The negative-going blanking signal is applied to the screen grid of the tube. The effect of the negative blanking signal applied there is to decrease the plate current in the stage. If an amplitude control for the blanking

pulse amplitude is provided, the height of the pedestal level may be set relative to the cutoff voltage of the following stage. Thus the proper background level of the picture may be set. The action here is quite similar to that described in connection with Fig. 6-26, except that a different mixing system is employed and the effective clip level is controlled in a different manner. The signal available at the output of the mixer stage consists of the picture signal plus the blanking signals (black level) at the proper time intervals. This composite picture signal when fed to the CRT grid will ensure that the latter is blanked out during horizontal and vertical flyback intervals.

⁴ A graphical analysis of the mixer circuit is presented in section 11-16.

Consider next the balanced deflection amplifiers which are located in both the camera and receiver units. The sole function of these amplifiers is to provide deflection voltages of proper amplitude and which are *balanced with respect to ground*. The use of balanced deflection signals tends to give more satisfactory operation in both of the cathode-ray devices.

The difficulty encountered with unbalanced deflection where one plate is returned to ground is illustrated in Fig. 8-5. In this case the electrostatic field is not uniform.

When the upper plate is positive, the electron beam moves upward and sees a different accelerating voltage than it does when it is deflected downward. The net effect is that the beam suffers defocusing as it moves toward certain areas on the target.

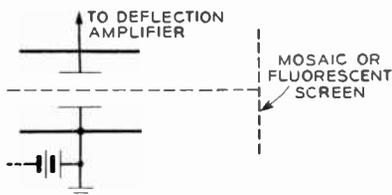


Fig. 8-5. Unbalanced deflection. One deflection plate is at ground potential.

With balanced deflection, on the other hand, neither plate is grounded and they are fed from a push-pull amplifier. Thus the average of the two deflection voltages is zero or ground potential and the electron effectively sees a uniform field as it moves toward the target.

Any of several forms of push-pull amplifiers may be used. One form, known as a paraphase amplifier, requires no separate tube for phase inversion.

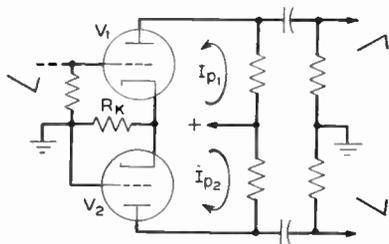


Fig. 8-6. Cathode-coupled paraphase amplifier.

The cathode-coupled version of the paraphase amplifier is shown in Fig. 8-6. Since the grid of V_2 is grounded, its cathode-to-grid voltage is that developed across the common cathode resistor, R_K . The polarities indicated in the diagram show that when a negative voltage is applied to the grid of V_1 , a voltage of opposite phase

and of almost the same magnitude is applied to V_2 . The difference in magnitudes is due to degeneration caused by R_K . The circuit may be analyzed by use of the equivalent plate-circuit theorem.

The required bandwidth of the several sweep amplifiers should be investigated to see if compensation is required. To do this we need to know how many harmonics of the saw-tooth waves are significant. To this end we shall make use of the Fourier expansion derived by Von Ardenne⁵ for a saw-tooth wave of finite flyback ratio, p .

$$e(t) = \frac{1}{\pi(P - P^2)} \sum_1^{\infty} \left(\frac{\sin n\pi P}{n^2} \right) \sin n\omega t \quad \text{where} \quad P = \frac{1}{1 + p} \quad (8-7)$$

For the standards we have chosen, $P = \frac{1}{2} \frac{9}{10} = 0.95$. The envelope of the relative amplitudes for the first ten terms of the series (8-7) are plotted in Fig. 8-7 as a function of n , the order of the harmonic.

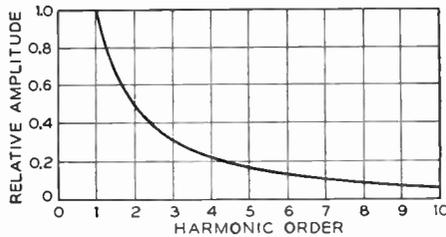


Fig. 8-7. Envelope of the relative magnitudes of the first ten harmonics of a saw-tooth wave of flyback ratio $p = 1/19$.

It will be observed that the change in magnitude for terms having an n greater than 6 or 7 is small and there is some question as to what harmonics may be neglected. We shall assume that the first 10 at least must be passed by the amplifier. Consequently the horizontal deflection amplifiers must have a pass band extending up to 150 kilocycles, which certainly presents no problems, and no compensation is required. For the vertical deflection system the fundamental component is 60 cycles and in general low-band compensation will be required.

We consider next the video amplifier system. The two prime aspects which must be considered here are the gain and bandwidth required. Let us first make an estimate of the required bandwidth by means of eq. (2-2)

$$f_r = \text{top video frequency} = \frac{1}{2} N_v^2 f_p$$

⁵ M. Von Ardenne, "Distortion of Saw-tooth Waveforms." *Electronics*, 10, 11 (November 1937).

To a first approximation we may assume N_v , the number of elements along a vertical line, to be identical to the number of active lines, thus

$$N_v \approx n_a = \frac{1}{2} \frac{9}{10} (250) = 237 \quad (8-8)$$

and, substituting, we have

$$f_v = \frac{1}{2} \left(\frac{1}{3} \right) (2.37)^2 10^4 (6 \times 10^3) = 2.24 \text{ megacycles} \quad (8-9)$$

To be on the safe side we shall design the amplifiers for an upper cutoff frequency of approximately 2.5 megacycles. The lowest frequency to be transmitted is 60 cycles, the picture-repetition rate. The design equations for an amplifier to meet these frequency requirements have been given in Chapter 7.

The gain required of the video amplifier system is determined by the voltages at three points in the system: the iconoscope output voltage, the voltage required at the input of the picture and blanking cable, and the value required at the kinescope grid. The first of these may be calculated with the aid of equations in Chapter 6 when the lighting level of the televised object and the constants of the pickup lens system are known.

Consider the voltage input requirements for the video cable. The cable must deliver to the receiver unit a voltage which is well above the noise level of the first video amplifier there. Call this required voltage E_2 . The cable will have an attenuation of α decibels per foot, a figure which is available from cable manufacturers' data. Then if l is the length in feet of the interconnecting cable, the required input voltage to the cable, E_1 , will be

$$E_1 = E_2 10^{\alpha l / 20} \quad (8-10)$$

Equation (8-10) is the result of the definition of attenuation expressed in decibels. Typical values of E_1 range from 1 to 5 volts peak to peak, the actual value being dependent upon the length and attenuation constant of the actual cable in use. The amplification required in the camera unit will be the ratio of E_1 calculated from (8-10) to the output voltage delivered by the iconoscope.

In a similar manner the total video amplification required at the receiver unit is the ratio of the kinescope grid voltage required for the full reproduction of the gray scale from black to white to E_2 . The former value may be obtained from the characteristics of the cathode-ray tube in use.

It is generally desirable to have some sort of picture gain and

brightness control at the receiving unit. The former, which effectively controls the contrast on the reproduced image, may be obtained with the conventional volume control of audio work. The brightness level control is illustrated in Fig. 8-8. C_c , R_2 , and the

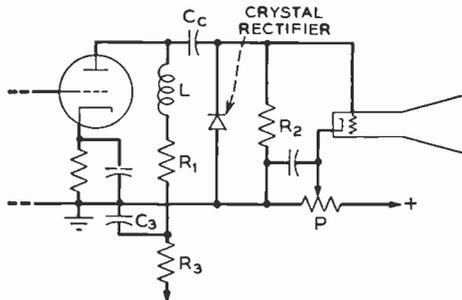


Fig. 8-8. P is the manual brightness control.

crystal rectifier may be identified with the clamping circuit described in the last chapter. Their action is to develop across R_2 a bias equal to the average value of the incoming signal. The potentiometer, P , allows an additional d-c component derived from the power supply to be added in series between the CRT cathode and grid. Thus P provides a means of controlling the d-c or average brightness component of the reproduced signal.

We have thus discussed the major components of the simple closed television system whose block diagram appears in Fig. 8-1. It should be immediately apparent that a number of variations are possible. We shall consider a few of these to indicate the flexibility which is available in the design of a closed television system. We follow a line of development which leads toward the open system, which employs radio transmitters, of the commercial telecasting industry.

8-2. Type II—Closed System

In the system just described, synchronization of the scanning rasters at both ends of the closed system was obtained by generating all sweep voltages at a common point. These saw-tooth voltages were then fed to the camera and receiver units. Another method of obtaining the same result would be to generate a set of master syn-

chronizing signals which, in turn, are used to trigger sweep-generator circuits located at both ends of the system. Such a system, which is illustrated in Fig. 8-9, provides a degree of flexibility not present in the type I network just described.

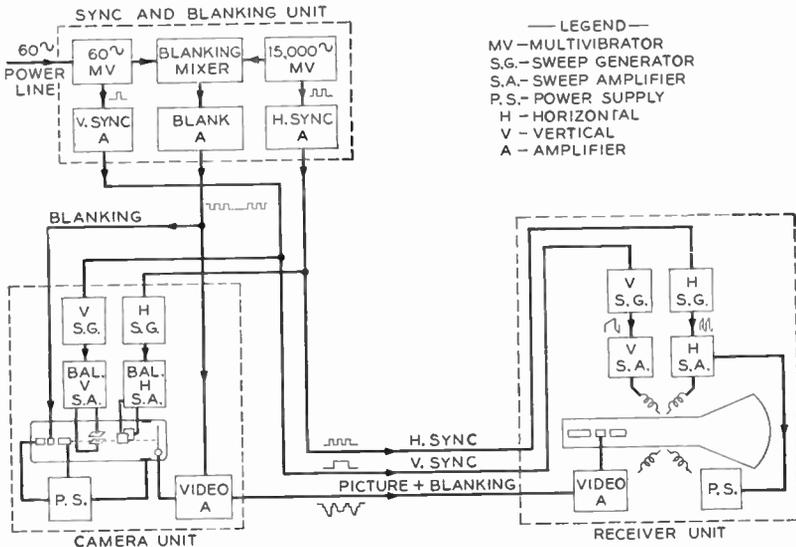


Fig. 8-9. Type II closed television system. Common sync voltages are fed to both ends of the system.

To illustrate the flexibility of the common sync system the circuit shows a CRT which employs magnetic deflection in the receiver unit. This is quite possible, even though the camera tube uses electrostatic deflection, because in each case the required sweep voltage is generated locally. The common synchronizing signal locks-in all the sweep generators and ensures that all the scanning rasters begin simultaneously and in phase. In the third type of system to be discussed we shall consider the idea of combining the horizontal and vertical synchronizing signals with the picture and blanking voltages. If such a combination is possible, two of the unit-interconnecting cables may be eliminated.

The sync and blanking unit of Fig. 8-9 merits special attention for it incorporates some features which depart from the pattern established in earlier chapters. It may be seen from the diagram that the

two base-frequency multivibrators, which operate at 60 cycles and 15 kilocycles, are not interconnected by a synchronizing link, *i.e.*, the maintenance of the correct frequency ratio between them depends solely upon the frequency stability of the two multivibrators, that is, the random scan is used. If, for example, the line-frequency multivibrator drifts from its design value of 15 kilocycles, the only effect will be a slight change in the number of lines and a corresponding change in the system resolution. The advantage of the free-running line-frequency multivibrator is that it eliminates the need for a counting circuit to provide lock-in between the low- and high-frequency scans.

Reference to Fig. 8-9 shows that the 60-cycle multivibrator is synchronized from the power line. We have already seen in Chapter 2 that this will reduce the effects of power supply hum voltage on the reproduced picture, because it ensures that the vertical sweep will be integrally related to the hum components and will have a fixed phase relationship to them.

The high-voltage power supply for the cathode-ray tube in the receiver unit is shown connected to the horizontal sweep amplifier. The implication is that a flyback-type high-voltage power supply is used.

The remaining elements of the system are similar to those described for the type I system. A complete design of the type II system which shows all the necessary circuit components has been made by Barrett and Goodman.⁶

8-3. Type III—Closed System

It has been pointed out that the number of cables feeding the receiver unit may be reduced to one if the synchronizing signals are combined with the picture and blanking signal. Although it may seem that the reduction in the number of cables is rather trivial, we shall outline the system because it lays the basis for the second section of this book which is concerned with the commercial telecasting system. If three independent channels were required for horizontal sync, vertical sync, and the picture, the problem of transmitting by

⁶ R. E. Barrett, and M. M. Goodman, "Simplified Television for Industry." *Electronics*, 20, 6 (June 1947). A similar system has been described which uses an image dissector in place of the iconoscope; see R. W. Sanders, *op. cit.*

radio would require three separate transmitters. If, on the other hand, all three channels are combined to give a "composite video signal," only one transmitter (*i.e.*, one interconnection channel) will serve. In setting up the type III closed system, then, we are pointing the way to what follows, nevertheless this must not obscure the fact that the development of the composite video signal to eliminate two cables in a closed system is of consequence itself. The block diagram of the type III network is shown in Fig. 8-10. The same

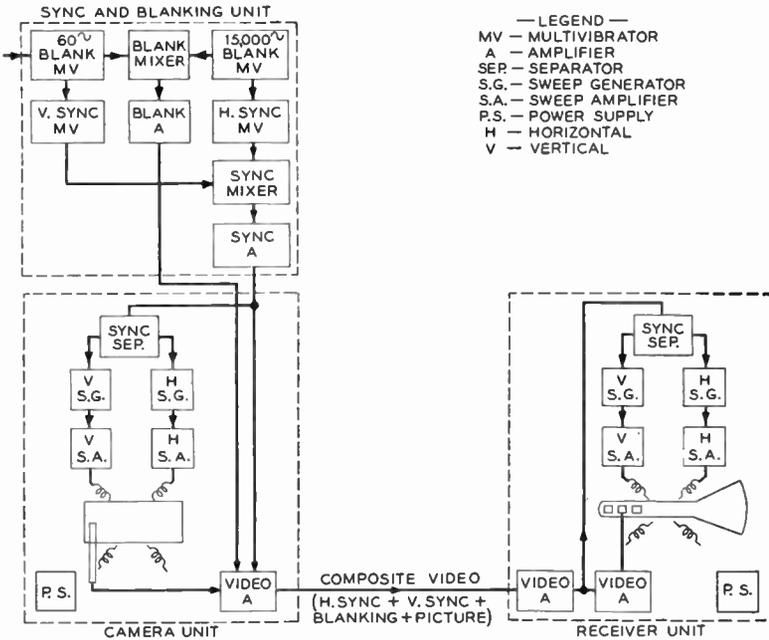


Fig. 8-10. Type III closed television system. Sync, blanking, and picture are combined into the composite video signal. Only one interconnecting channel is required between the two ends of the system.

standards of line and frame frequency which were proposed at the beginning of the chapter are assumed. By way of variation an image dissector is shown as the camera tube. Because of the marked similarity between this and the commercial system which is described at

length later, we shall only give a broad outline of the circuit operation. Each component in the sync and blanking unit will contain several subcomponents, some of which will be described. A complete synchronizing signal generator is described in Chapter 11.

The combination of the several voltages into a single composite video signal is possible because the synchronizing signals and picture do not occur simultaneously. We have previously pointed out that the picture is blanked out (*i.e.*, driven to or beyond black level) during the retrace intervals. The sync signals serve to initiate the retrace in both directions; hence we need only to ensure that the picture is blanked out just before any synchronizing pulse begins. If the synchronizing pulses extend in the black direction beyond the black level, they will have no effect on the picture information proper. Following this line of thought we see that, were it possible, the gray shade produced by the sync pulses on the kinescope would be "blacker than black," hence the notation that the sync pulses lie in the blacker-than-black region of the signal's amplitude.

Four principal topics confront us: (1) Combination of horizontal and vertical synchronizing pulses into a combined synchronizing or "supersync" signal; (2) Combination of the supersync and picture voltages into the composite video signal; (3) Separation of the composite video signal into its supersync and picture components; (4) Separation of the supersync into its horizontal and vertical sync components. We now consider these processes of combination and separation in order. We shall only outline the system because of its similarity to the commercial one which is covered at length in the subsequent chapters.

8-4. The Supersync Signal

It should be realized at the outset that a number of methods of combining the horizontal and vertical synchronizing signals are possible. We shall follow commercial practice and describe the system which requires that the horizontal and vertical synchronizing pulses differ considerably in width.

A moment's reflection will show that if the two sync signals are to be combined and then separated, they must be sufficiently different so that some sort of electrical network can distinguish between them. This requirement may be met by using a short horizontal pulse and a relatively wide vertical pulse. These may be separated by the use

of differentiating and integrating circuits as shown in Fig. 8-11.⁷ For example, it may be seen at *b* that if sufficient bias is placed on a sweep-generating stage which is fed by an integrating circuit, only the longer pulse will produce an output of sufficient magnitude to trigger the sweep generator; the integrating circuit can "protect" the vertical sweep generator from the narrow horizontal sync pulses. We see, then, that the choice of a narrow horizontal sync pulse and a wide vertical one meets the separation requirement. It would appear that a supersync signal of the form shown in Fig. 8-12*a* would meet the requirements for satisfactory synchronization.

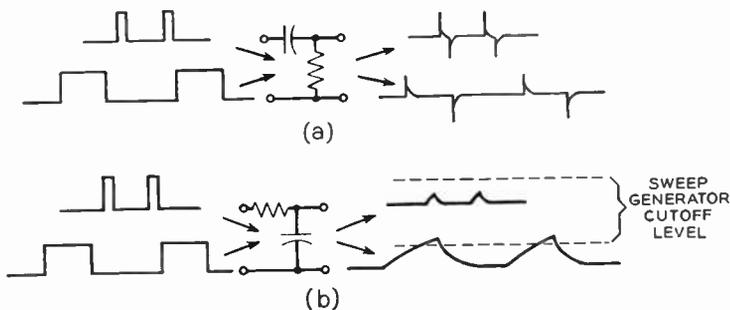


Fig. 8-11. Differentiating and integrating circuits can distinguish between narrow and wide pulses. (a) Response of a differentiating circuit to pulses of different width. (b) Response of an integrating circuit to pulses of different width.

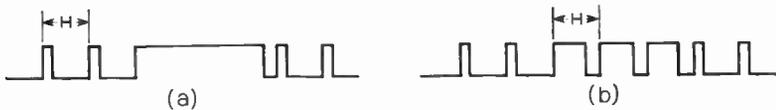


Fig. 8-12. Development of the supersync wave form. (a) Developmental form of supersync. (b) Final form of the supersync signal which permits horizontal synchronization during the vertical sync interval.

A more careful consideration of the problem, however, shows that if such a signal is applied to a differentiating circuit which serves to trigger the horizontal sweep generator, no differentiated pulses will occur for the entire duration of the long vertical pulse, a condition which results in the loss of horizontal sweep during the vertical re-

⁷ The operation of these circuits is analyzed in chap. 11.

trace. Even though no picture information is presented during this interval, it is advisable from the standpoint of horizontal sweep stability to keep the sweep going and in synchronism at all times. This we may accomplish by notching the vertical pulse so that a leading edge occurs at each interval of H at all times. A wave of this form is shown at Fig. 8-12*b*. When it is differentiated in the horizontal sweep system, each leading edge produces a sync pulse for the horizontal sweep generator. The narrow notches in the vertical sync pulse have negligible effect on the integrated output which triggers the vertical sweep circuit. We shall assume, then, that in combining the horizontal and vertical sync pulses, the required output is that shown at *b* in the diagram.

One form of mixing circuit which can meet these requirements is shown in Fig. 8-13. The horizontal sync *a* and the vertical sync *b* are applied respectively to the control and screen grids of the cathode

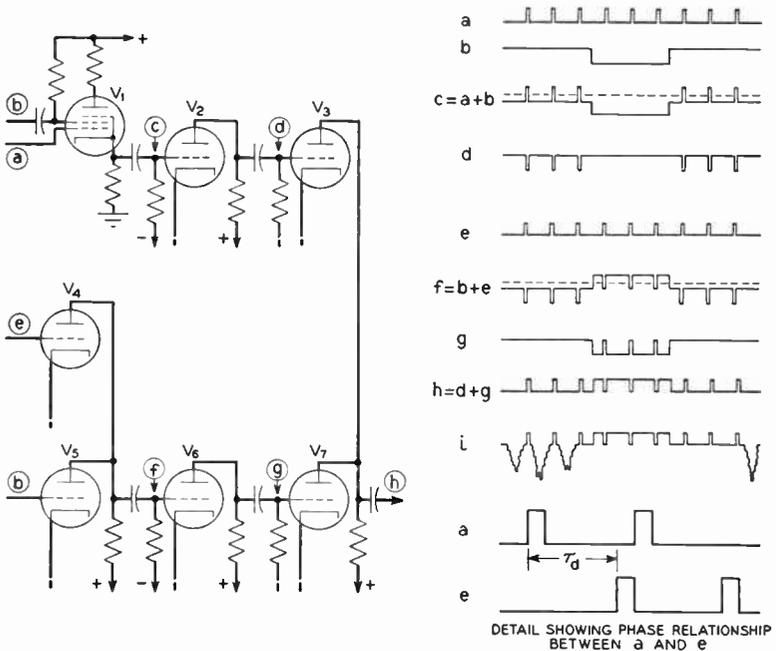


Fig. 8-13. Generation of the supersync signal. Detail showing phase relationship between *a* and *e*.

follower V_1 . The magnitude of b is sufficient to cut off the tube. The effect is to remove the three horizontal pulses (the number three is chosen arbitrarily) which occur during the vertical sync interval. V_2 is biased so that the notch due to b is eliminated. The resulting wave delivered to V_3 consists of a train of negative horizontal sync pulses, of which three are eliminated for the duration of each vertical sync pulse.

The wave shown at e consists of a train of horizontal sync pulses which have been delayed by an amount τ_d . From the detail in the diagram it may be seen that the trailing edges of e coincide with the leading edges of a . The reason for this relationship will become apparent as we continue. e and the vertical sync pulse b are combined in the common plate load of V_4 and V_5 to give f . The bias level on tube V_6 is adjusted to clip off the lower portion of the wave so that the only signal appearing on the grid of V_7 is the notched vertical pulse. The signals applied to V_3 and V_7 combine to give the required supersync voltage h .

We may now see the need for the delayed horizontal pulses which were applied to V_4 . Notice that the delay in e is chosen so that its trailing edge (which becomes a leading edge in the notched vertical pulse) corresponds to a leading edge in the sync pulses shown at a . Then in the final wave h the *rising* edges of successive pulses occur at intervals of H , the line period, thereby ensuring that horizontal sweep is maintained in synchronism throughout the vertical pulse duration. The single exception is due to the leading edge of the vertical pulse itself. This may cause one misfire, but the horizontal sweep will be in proper synchronism by the time the picture data are again presented.

The over-all action of the supersync mixing circuit may be summarized as follows. At the end of each picture or frame, three horizontal sync pulses are omitted. In their place is substituted a notched vertical pulse, the notches occurring in such a way that each leading edge occurs at the proper time to initiate the horizontal sweep.

8-5. The Composite Video Signal

Assume that the picture and blanking signals have been combined, say, as in Fig. 8-4. We must now mix this signal with the supersync. It has already been indicated that the supersync shall lie in

the blacker-than-black region of the final signal. The mixing can take place in a common plate load circuit which has been described before and is illustrated by V_3 and V_7 in Fig. 8-13. Thus if the positive-going supersync were applied to the grid of one mixer tube, and a black-positive picture and blanking signal to the other, the combined output voltage would be the composite video signal shown at i in the figure. It may be seen from Fig. 8-10 that the mixing of the composite video signal occurs in the video amplifier chain located at the camera unit. One other point should be mentioned. For the system whose block diagram appears in Fig. 8-10, the horizontal master multivibrator is free-running and hence may exhibit a slow drift over a long interval of time. This will show up as a slow change in the spacing between the leading edge of the notched vertical pulse and the horizontal sync pulses. We have previously observed that this lack of tie-in between horizontal and vertical sweep in the non-interlaced scan is not serious; in the particular circuit which has been the subject of study this shift in leading edges will occur only during the vertical blanking interval and will have little, if any, effect on the final picture.

8-6. Supersync and Picture Separation

We have seen how the several components are combined into the composite video signal at the pickup end of the system. This composite voltage is transmitted by cable to the receiver unit where it must be analyzed into its original components, each of which must be channeled to the proper location. This is our next consideration.

Inspection of the composite signal, i , in Fig. 8-13, shows that the supersync and picture components differ in time and in amplitude. It is this last fact which allows us to separate the two components. Thus if a sync-positive composite video signal is applied to a properly biased stage so that all of the picture components lie below cutoff, only the synchronizing signals will appear across the plate load; the supersync has been stripped from the picture. The stripped supersync may then be inverted and fed to differentiating and integrating circuits as shown in Fig. 8-14. The two outputs may then be used to trigger the horizontal and vertical sweep generators, respectively.

As far as the picture channel which feeds the grid of the cathode-ray tube is concerned, no picture stripping is required, the reason being that the presence of the supersync on that grid only tends to

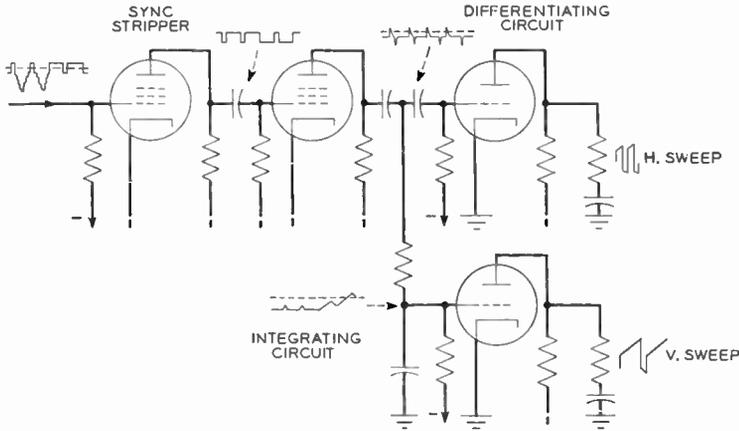


Fig. 8-14. Sync stripping, and separation of the supersync into its horizontal and vertical components.

drive the kinescope beyond cutoff. Since this occurs during the blanking interval when no light is being presented on the screen, the final picture is unaffected by the presence of the synchronizing signals.

We have seen, therefore, how at the expense of more terminal equipment the number of channels required for connecting the two ends of the closed television system may be reduced to one. Although the type III closed system which we have just described is interesting from an instructional point of view, the likelihood of its being used as a practical system is small. The reason is that in principle and in several details it resembles the commercial telecasting equipment. It would be more feasible to adopt the commercial standards and make use of the standard equipment which is available on the market.

Of closed television systems in general, we may say that they are coming to be of increasing importance. The design of such a system involves a choice of a set of standards which meet the requirements in hand, and the synthesis of the system from the several components which have been the subject of the first portion of our study.

Part II

THE COMMERCIAL TELECASTING
SYSTEM

CHAPTER 9

SPECIAL PROBLEMS OF TELECASTING

It is axiomatic that a single set of television standards is a basic requirement for a nationwide commercial telecasting system. Only the assurance that any receiver will be able to receive any program from any station within its operating range will ensure willingness on the part of the public to invest in a television receiver. It is fortunate for the rapid development of commercial telecasting in the United States that industry and the Federal Communications Commission alike recognized this need to overcome buyer reluctance. It was proposed to keep telecasting on an experimental basis until more or less universal agreement could be reached on a set of operating techniques and standards which at the same time would make available high-definition pictures and yet allow technical advances within that framework of techniques and standards.

It is beyond the scope of our work to trace the technical developments which have culminated in the present commercial system. We shall, however, list some of the principal standards, show the basis for their acceptance, and explain what electrical components are required in the system that operates under those standards.¹

The guiding principle in establishing standards is based on the "lock and key" relationship between the transmitter and receiver: the signal radiated from a transmitting station must be susceptible to reception and conversion into a high-definition reproduced image by any receiver which conforms to the accepted standards. A second principle is based on the preponderance of receivers over transmitters, and the technical ability of the personnel which operate them: whenever possible, the construction and operation of the receiver will be simplified at the expense of the transmitting facilities.

Some idea of the difficulty involved in getting industry-wide agree-

¹ An excellent summary of the establishment of present standards is available. See D. G. Fink, *Television Standards and Practice*. New York: McGraw-Hill Book Company, Inc., 1943.

ment on standards may be had by briefly outlining some of the questions which must be answered. To begin with, the commercial telecasting system must handle a picture program plus the accompanying sound. Shall separate transmitters be used for the aural and video signals, and if so what should be the separation between their carrier frequencies? Shall the transmitters employ frequency—or amplitude—or some other form of modulation? Shall the antennas employ vertical or horizontal polarization? How shall synchronizing information be transmitted? What shall be the picture shape, the frame frequency, the line frequency? What shall be the scanning pattern; should an interlaced scan be employed? It is interesting to review briefly the means which were used to answer these questions.

9-1. The Standardization Committees

As early as 1929, even before means of all-electronic pickup were developed, the radio industry through the Radio Manufacturers' Association saw the need for a clearing house for standardization, and set up the Committee on Television to serve as that agency. In the intervening years the committee followed developments in the field and in 1935 investigated the advisability of establishing standards based on the all-electronic scanning system which had been developed. The following year the committee recommended to the Federal Communications Commission a system based on the following standards: 441 lines, 60 fields per second using 2 to 1 interlace, aspect ratio 4 to 3. The picture transmitter was to employ conventional double-sideband amplitude modulation. The aural transmitter was to be amplitude-modulated and the two carrier signals were to be separated by 3.25 megacycles with a total bandwidth of 6 megacycles allotted per station. Primarily transmission would be in the VHF range around 60 megacycles.

In the late summer of 1936 the F.C.C. replied to these proposals by opening 6-megacycle-wide channels from 42 to 56 and from 60 to 86 megacycles for experimental operation only. No decision was made on the standards which were still incomplete and were not essential to experimental work. It is to the credit of the broadcasting companies who participated in this early work that they were willing to make large investments each year with no prospect of immediate financial return.

In the two following years the R.M.A. committee continued work

on the standards and set up a recommended operating practice which specified horizontal antenna polarization and filled in other standards chiefly concerned with the transmission of synchronizing signals. By the end of this period it was evident that the use of single side-band transmission of some sort would provide better utilization of the 6-megacycle channel width and the recommended separation between aural and visual carriers was raised from 3.25 to 4.5 megacycles. The enlarged list of standards was recommended to the F.C.C. Again the standards were not made final but two types of operation were permitted, type I, which was unscheduled and for experimental work, and type II, which was for scheduled programs of interest to the public. Financially the burden was still on the broadcaster; no commercial sponsorship of programs was permitted, although some sponsorship funds were permitted for developmental work.

By late 1940 a rift arose in the industry regarding the proposed R.M.A. standards. The National Broadcasting Company began limited commercial operation under the R.M.A. proposals, and its parent organization, the Radio Corporation of America, began to manufacture receivers for public use. The effect of the rift was the announcement by the F.C.C. that it would issue licenses for full-scale commercial operation whenever the industry at large would agree on a single set of standards.

The agreement was reached through the agency of the National Television System Committee. Drawing its membership from the R.M.A. and other interested organizations which were outside that body, the committee began to meet in July 1940. In its work it reviewed all the major aspects of the television system and recommended a unified set of standards to the Federal Communications Commission. By July 1, 1941 the Commission had issued the Standards of Good Engineering Practice Concerning Television Broadcast Stations, and full commercial operation of television stations was under way. However, after the declaration of war in 1941 the stations ceased commercial operation.

By 1945, after cessation of hostilities, the F.C.C. issued a new set of rules which differed from the prewar set in three prime features; the allocation of channels was increased to 13 with new frequency assignments, the number of scanning lines was raised from 441 to 525, and frequency modulation of the aural transmitter was specified.

These standards remain in effect at the time of writing with the prime exception that channel 1 is no longer in use. The commercial standards in the United States are a direct outcome of recommendations from the television industry itself.

9-2. The Commercial Television System, Block Diagram

With our background of closed television systems we may now intelligently set up the block diagram for the open system used commercially. Remembering that transmission of the program material between transmitter and receiver is by means of radio, we first consider a few of the problems involved in that portion of the system.

During the years when the R.M.A. committee on television was working on standards the techniques of sharing a single carrier between two or more modulating signals were not developed. It is only natural then, that if sight and sound could not be placed on a single carrier that two separate transmitters be used, the one for the picture signal and the other for the accompanying sound program. The two carrier frequencies are to be separated by 4.5 megacycles in conformance with the standard mentioned above.

At the receiving location it would be unnecessarily complex if two separate receivers were required, one each for the visual and aural programs. This complexity is circumvented by a superheterodyne receiver whose radio frequency and converter circuits are sufficiently broad-band to accept both programs. The intermediate frequency signals whose carriers are separated by 4.5 megacycles may then be separated in properly tuned IF amplifiers. In effect, the television receiver comprises two separate IF and low-frequency channels fed by a common RF and converter section, an arrangement which permits tuning to picture and accompanying sound in a single operation. The complete block diagram of the open system is shown in Fig. 9-1. We stress once again that the system shown is typical; many variations are possible which permit operation under the specified standards. Several of these variations are considered in subsequent chapters.

Consider the block diagram of Fig. 9-1. The aural transmitter is frequency-modulated to a maximum swing of ± 25 kilocycles. Its power output shall be within one-half to one and one-half times the peak power output of the visual transmitter. The diagram shows the FM transmitter feeding a diplexer, a unit which permits the out-

hum effects which were discussed in Chapter 2. A few other features of the block diagram are worthy of mention; for example, notice that two blanking outputs are shown from the sync and blanking generator. These are the "camera blank" output and the "line blank" output which feed the camera and line amplifiers, respectively. The implication is that these two blanking voltages differ in wave form; actually, as may be seen from the diagrams in Chapter 11, the line blanking pulses are of longer duration than the camera blanking pulses and overlap the leading and trailing edges of the latter. The line-blanking voltage is mixed with the picture signal in the line amplifier and, after passing through the entire system, appears on the control grid of the final cathode-ray tube. The effect of the wider line-blanking pulses, then, is to ensure that the cathode-ray tube at the receiver is blanked out before and after the camera blanking occurs. This guarantees that the camera will be delivering picture information when the CRT is unblanked.

This policy of making the camera end work a little faster than the playback end is also carried over into the sync signals. In the block diagram the two lower outputs from the sync and blanking generator furnish synchronizing signals to the studio sweep generators. These camera synchronizing signals, labeled "*H* drive" and "*V* drive" on the diagram, are in each case narrower than their counterparts which go to make up the supersync that is transmitted to the receiver. Once again this ensures that the camera retrace begins and ends within the duration of the CRT retrace, thus the pickup system is always ready to deliver picture information whenever the receiver is able to display it. The net effect is that two sets of blanking and sweep standards are in effect, one for the camera tube and one for the CRT, the former being the more rigid.

The line-blanking, supersync, and picture signals are combined in the line amplifier in the general manner described in the last chapter. Also shown in the diagram is a monitor whose sweep is obtained from the studio sweep generator. Such a monitor is termed a "driven monitor" because the source of its synchronizing signals is the *H* drive and *V* drive outputs of the sync generator. The monitor provides a constant check on the operation of the entire pickup system.

The principal standards for the studio facilities are listed below:

- (1) $A = 4$ to 3.
- (2) During the active scanning intervals, the *scene* shall be

scanned from left to right and from top to bottom with constant horizontal and vertical velocities.

- (3) A 2 to 1 interlaced scan pattern shall be used.
- (4) $f_p = 30$ cycles.
- (5) $f_f = 60$ cycles.
- (6) $n = 525$.
- (7) $f_i = nf_p = 15,750$ cycles.
- (8) The wave form of the composite signal is similar to that shown at i in Fig. 8-13 and covered in detail in Chapter 11.

It will be observed that standards 1 through 5 of the list given above have been discussed in Chapter 2. The choice of 525 lines is based on the bandwidth available and on the characteristics of the receiver, a subject covered in the next chapter.

The video transmitter proper comprises a crystal oscillator whose output is multiplied to the proper frequency and whose power level is raised to the value required at the input of the modulated amplifier. The modulated output is fed to the diplexer through a side-band filter which removes a large portion of the lower side band generated in the process of amplitude modulation. The transmitter is the subject of Chapter 13. The principal standards which apply are listed below:

- (9) Operation in the VHF band shall be confined to one of the following channels:²

| <i>Megacycles</i> | <i>Megacycles</i> |
|-------------------|-------------------|
| #2. 54 to 60 | #8. 180 to 186 |
| #3. 60 to 66 | #9. 186 to 192 |
| #4. 66 to 72 | #10. 192 to 198 |
| #5. 76 to 82 | #11. 198 to 204 |
| #6. 82 to 88 | #12. 204 to 210 |
| #7. 174 to 180 | #13. 210 to 216 |

- (10) The visual carrier shall be located 4.5 megacycles below the aural carrier frequency.
- (11) The aural carrier frequency shall be located 0.25 m.c. below the upper frequency limit of the channel.
- (12) The visual carrier shall be amplitude modulated with both picture and synchronizing signals, the two signals corresponding to different ranges in modulation amplitude.

² These are the assigned channels as of December 1950. It is anticipated that additional channels in the VHF region extending from 475 to 890 megacycles will be opened for television transmission.

- (13) A decrease in light shall cause an increase in radiated power.
- (14) The black level shall correspond to a definite carrier level, independent of picture content.
- (15) Pedestal level shall be transmitted at $75\% \pm 2.5\%$ of the peak voltage amplitude of the signal.
- (16) The maximum white signal shall modulate the carrier to $12.5\% \pm 2.5\%$ of the peak voltage amplitude of the radiated signal.
- (17) The radiated signals shall have horizontal polarization.

The required amplitude distribution in the composite video signal to meet these requirements is shown in Fig. 9-2. How the visual

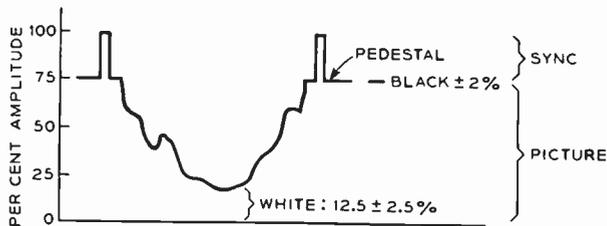


Fig. 9-2. Distribution of amplitude in the composite video signal.
The upper 25 per cent is reserved for the synchronizing signals.

and aural side bands are disposed in the 6-megacycle channel allotted to a station is discussed in Chapter 12.

The receiver shown in Fig. 9-1 comprises a common "front end" of approximately 6-megacycle bandwidth and capable of handling the combined visual and aural signals, followed by two independent IF and output chains. The upper or aural system is composed of the components of a conventional FM receiver which serve to demodulate the FM aural carrier. The audio voltage is amplified and fed to a loudspeaker.

In the video portion of the receiver, the IF signal is amplified and detected, the detected output being the composite video signal. The remainder of the system is similar to the receiver unit of the type III closed system previously described. The black negative composite video signal is fed to the control grid of the CRT where it serves to modulate the light output on the fluorescent screen. A sync-positive version of the same signal is fed to a sync stripper, which removes the picture information and sends the supersonic to the two sweep sys-

tems. Separation of the horizontal and vertical sync pulses is accomplished with the aid of differentiating and integrating circuits as previously described.

In the remaining chapters of this section we investigate the reasons governing the choice of some of the standards which have been listed. We then investigate the various components of the visual portions of the over-all commercial telecasting system.

CHAPTER 10

THE OPTIMUM NUMBER OF LINES

In the final analysis the choice of the optimum number of scanning lines for the commercial television system must be based on subjective tests. It is nevertheless illuminating to subject the various factors involved to an analysis which points in the right direction. An analysis of this sort lends to a better understanding of the whole system and demonstrates admirably some of the analytical methods that have been developed in the art. Such an analysis is the subject of the present chapter, and it will be shown how the optimum number of scanning lines may be determined when the bandwidth allowed for transmitting the picture signal is specified.¹

10-1. Some Factors Governing Resolution

In Chapter 2 an approximate equation was derived which related the highest video frequency transmitted to N_v and N_h , the number of picture elements transmitted along a vertical and a horizontal line, respectively. The term "resolution" of the picture was also used as a measure of the figure of merit, $M = N_v N_h$. By an extension of this concept we may define the resolution as that characteristic of a picture which makes it sharp and clear as opposed to smeared and blurred. We must now consider some of the factors in the over-all television system which affect the resolution of the final picture.

In following the line of reasoning used in Chapter 2 we see that N_v , a measure of the resolution in the vertical direction, is directly proportional to the number of scanning lines; the larger the number of

¹The problem here is essentially that faced by the F.C.C. when it changed the 441-line standard to 525 in 1945. A more basic problem when standards for a new system are set up is this: (a) What number of lines is required to produce a satisfactory image when the observer is some fixed distance, say 3 to 6 times picture height, away from the reproduced image; (b) what bandwidth is required to transmit the signal obtained when scanning with this number of lines? See, for example, E. W. Engstrom, "A Study of Television Image Characteristics," Part I. *Proc. IRE*, 21, 12 (December 1933).

lines, the greater will be the number of elements which can be reproduced in the vertical direction. There is an upper useful limit to this relationship, however, which is imposed by the limited visual acuity of the human eye; since the eye can only resolve a finite amount of detail at any given viewing distance, it is wasteful to increase n and N_v beyond the limit which the eye can discern. This factor, then, places an upper limit on the number of scanning lines.

A second factor which is related indirectly to the number of lines is N_h , the detail transmitted along a scanning line. Consider this relationship. In Chapter 5 it was shown that the variation of brightness over the surface of a picture could be considered as the summation of an infinite number of bidirectional Fourier components. It was also shown that the parameter k was a measure of the detail along a horizontal line, for k determines the number of sinusoidal brightness variations *across* the face of the picture. It follows, therefore, that N_h is proportional to k .

In the same chapter we derived the relationship between the Fourier components and the frequency spectrum of the electrical signal which results from scanning a transducer over the picture in the specified scan pattern. The resulting spectrum which is plotted in Fig. 5-7 shows that k , which determines the multiple of line frequency present, is the chief factor in determining the high-frequency limit of the electrical signal. We might sum up these facts in the following manner: N_h determines k , which in turn determines the bandwidth of the electrical signal. If, then, the television system had an infinitely wide bandwidth, there would be no limitation on the amount of detail transmitted along a scanning line, or we may say that N_h is limited by the system bandwidth provided that aperture distortion is eliminated.

It would appear, then, that N_h is independent of the number of lines. As a practical matter this statement is incorrect as we may see from the approximate equation preceding (2-2)

$$f_v = \frac{1}{2}N_vN_hf_p$$

It follows that for a given system bandwidth, f_v is fixed so that the product N_vN_h must be constant; hence a finite bandwidth relates N_h to N_v and the number of scanning lines. From standard 10 of the last chapter it follows that the video bandwidth of the commercial system cannot exceed 4.5 megacycles. Our problem, then, is to

choose the number of lines such that an optimum compromise between N_h and N_v is reached. It should be noticed that they may be traded, the one for the other.

We have previously observed that a person viewing a televised image naturally tends to adjust his viewing distance so that the coarsest detail blends into a smooth picture. It is patently wasteful of detail, then, to greatly increase the resolution in either direction over that in the other direction. Our criterion is to choose a number of scanning lines which, for the specified bandwidth of roughly 4.5 megacycles, gives equal vertical and horizontal resolution. We must next standardize the amount of overlap between adjacent scanning lines.

10-2. Overlap, Line Structure, and Resolution

In deriving the number of lines which gives approximately equal vertical and horizontal resolution it is convenient to consider the problem at the receiving end of the system where the video signal applied to the control grid of a cathode-ray tube modulates the light intensity on the fluorescent screen. To simplify the work we shall neglect the effects of aperture distortion at the pickup end which may be corrected electrically.

We consider first the effect of adjacent line overlap on the reproduced image in the presence of a constant white video modulating signal. We have seen in Chapter 5 that an electron beam of circular cross-section scanning across a fluorescent screen leaves behind it a line whose intensity normal to the direction of scan varies as the \cos^2 of the distance from the center, Figure 5-23. Let r be radius of the scanning spot and y be distance measured from the center line of the scanning line. Then the intensity across a scanned line is given by

$$I(y) = \hat{I} \cos^2 \left(\frac{y}{r} 90^\circ \right) \quad (10-2)$$

If, now, the spot diameter is adjusted such that the edges of two adjacent scanned lines just touch, eq. (10-2) may be applied directly to give the variation of intensity vertically across the face of the scanned area. The resulting image produced by this condition is shown at *a* in Fig. 10-1. Inspection of the diagram shows that a flat or uniform field is not reproduced, and that the line structure of the image is very apparent.

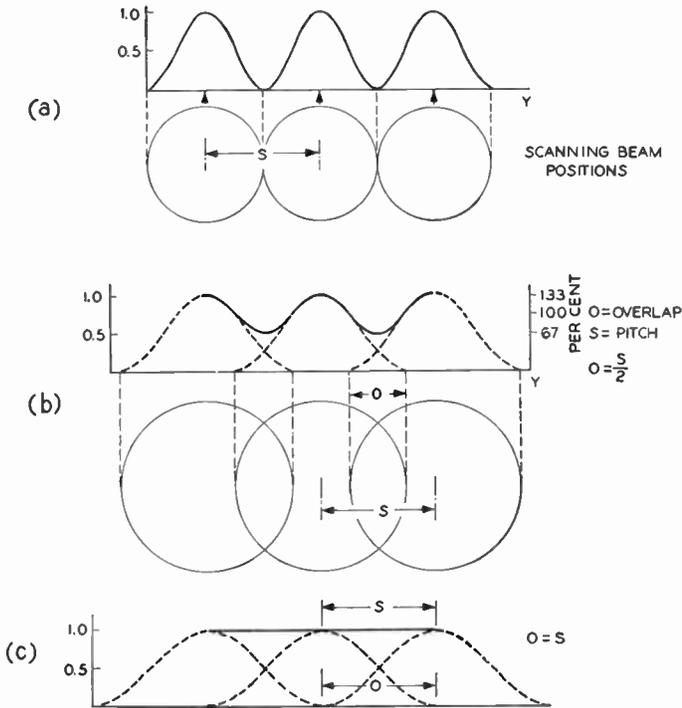


Fig. 10-1. Effect of adjacent-line overlap on field flatness. A cosine-squared distribution of intensity across the scanned line is assumed. (a) Reproduced intensity across the scanning lines, adjacent lines contiguous. (b) Reproduced intensity across the scanning lines, overlap = 0.5 pitch. (c) Reproduced intensity across the scanning lines, overlap = pitch.

This condition may be ameliorated by keeping the pitch or distance between adjacent scanning lines, s , constant and widening out the spot width. To illustrate this, Fig. 10-1b is plotted for the case where adjacent lines overlap by an amount which is 50 per cent of the pitch. The dotted curves show the intensity distribution across each scanning line. The solid curve is the result of summing up the individual curves point by point and represents the actual distribution across the reproduced image in the vertical direction. It will also be seen from the figure that the 50 per cent overlap case provides a flatter field than does the zero overlap case. The disadvantage of overlapping adjacent lines, however, is that with overlap a loss of

detail must result in the overlap region. For example, when the light is being modulated by a video signal, the image reproduced in the region of overlap will be the sum of the images of the two scanning lines, and a loss of resolution results. The condition is illustrated in Fig. 10-2, where the picture data is indicated by cross

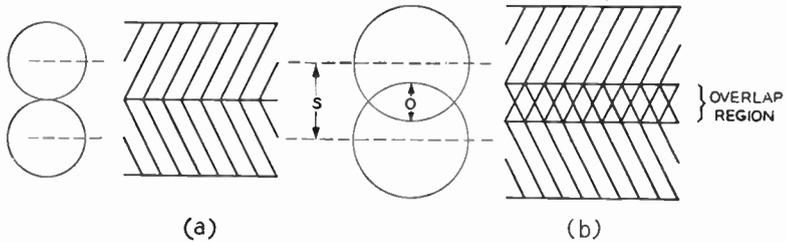


Fig. 10-2. When adjacent lines are overlapped to flatten the field, there is a loss of detail in the overlap region. (b) also illustrates the loss of detail due to "pairing."

hatches. At a where there is no overlap, the picture data along each line is kept separate from the adjacent ones. At b , the spot size has been increased to give an adjacent line overlap of one-half the pitch. It may be seen that where the lines overlap the data becomes confused with a resulting loss in detail. Actually the 50 per cent overlap condition will not give as poor results as indicated in the diagram for as we have seen, the intensity of the scanned image decreases as the \cos^2 of the normalized distance from a line center. This means that in the actual case the region where the resolution decreases will be narrower than the overlap band.²

In extending these ideas even farther, we see that if 100 per cent overlap be used, that is, if $o = s$, a perfectly flat field will result, as shown in Fig. 10-1c. The fact that the 100 per cent overlap yields a flat field may be verified by the well-known trigonometric relationship: $\sin^2 \theta + \cos^2 \theta = 1$. From the point of view of resolution, however, the $o = s$ condition is entirely out of the question.

² Figure 10-2 also illustrates the effect of "pairing" in interlaced scanning systems. If the lines of successive fields do not fall midway between each other adjacent pairs of lines, one from each field, overlap to some extent producing a loss of detail in the overlapped region. This grouping of adjacent lines is known as pairing.

The last few paragraphs have shown that spot size affects the resolution (a fact which we have previously observed in Chapter 5) and the flatness of a white field. It should be apparent, then, that in determining the optimum number of lines for the television system we must decide beforehand what overlap is to be assumed for the analysis and hold that value fixed. There is some difficulty involved in choosing a value of overlap because individual tastes differ. Some persons are willing to sacrifice detail to eliminate the line structure in the picture; others want maximum detail at all costs. As a compromise between these extreme points of view, the 50 per cent overlap case illustrated in Fig. 10-1b will be assumed in the remainder of the treatment. The results may be revised in either direction to conform to other choices in adjacent line overlap.

Reference to Fig. 10-1b shows that with a constant white signal applied to the grid of the cathode-ray tube the field intensity in the vertical direction fluctuates about a mean of $I = 0.75\hat{I}$. It is convenient for the work which follows to define this $0.75\hat{I}$ response as the 100 per cent intensity level. It is clear, then, that any value of I may be converted to a per cent-intensity value by the following equation:

$$\text{per cent intensity} = \frac{I}{0.75\hat{I}} \quad (10-3)$$

The scale corresponding to this equation is shown at the right in Fig. 10-1b.

We must next consider what happens when a varying, rather than a constant, modulating signal is applied to the control grid of the cathode-ray tube. If we assume that the latter is being operated in the linear region of its light-intensity-grid-voltage characteristic, a change in control grid voltage will produce a proportionate change in \hat{I} of the scanned line. We are able then to plot the intensity variation on the fluorescent screen when the control grid voltage is changed.

This principle is illustrated in Fig. 10-3 where we see the variation in intensity in the reproduced image along a vertical line c . It is assumed that the control grid voltage is such that a black-to-white transition should be reproduced as represented by the line OM . *The effect of aperture distortion in the horizontal direction along a scanning line is neglected* because it may be corrected by proper electrical net-

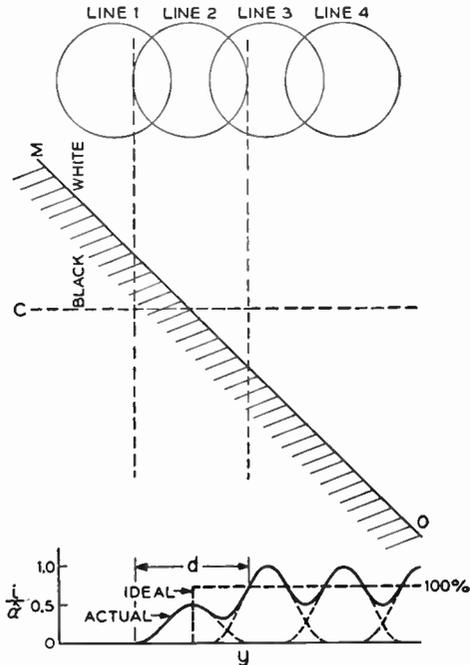


Fig. 10-3. The reproduced light intensity along a vertical line c when a black-to-white transition OM is reproduced.

works. When the beam is moving along line 1 it is blanked out at c by the signal voltage and the intensity on the tube face is zero as shown in the intensity plot in the diagram. As the spot moves along line 2 it straddles the black-to-white transition at c and hence the light intensity will be determined by a 50 per cent gray control grid signal. Thus for line 2 the peak value of I is one-half of the value which gives white, and the intensity curve for line 2 is the $\cos^2 \theta$ function modified by the factor $\frac{1}{2}$ as shown in the diagram. For the remaining lines the spot reproduces full white along c and the intensity curves have the shape shown, which corresponds to the $\cos^2 \theta$ function. It is this type of intensity curve that results when an abrupt black-to-white boundary is scanned, which serves as the basis for the analysis which follows.

10-3. Vertical Resolution³

We have stated earlier in the chapter that our goal is the choice of a number of scanning lines which in a video pass band of some 4.5 megacycles will give approximately equal vertical and horizontal resolution. Clearly, if we are to compare the resolutions in both directions, we must have means of measuring them. One such measure, N_v and N_h , may be used quite conveniently when the subject matter is of the checkerboard type illustrated in Fig. 2-4. It has been pointed out, however, by Kell, Bedford, and Fredendall that such a checkerboard pattern does not provide a useful measure of resolution because the alternate black and white squares do not serve as fundamental "building blocks" from which a more complex image may be synthesized. A basic pattern of greater utility is of the unit function form and consists of an abrupt black-to-white boundary of the type illustrated in Fig. 10-3. Reference to that figure shows that the abrupt transition along the vertical line c is reproduced as a gradual oscillating change in light intensity, which spreads over approximately the width of one scanning line for the case illustrated. We can, therefore, *take the distance, d , expressed in units of the pitch, required for the reproduced intensity to rise from zero to the 100 per cent intensity level as a measure of the vertical resolution.*⁴ Notice that the actual, abrupt transition is stretched out over the normalized distance, d , thus $1/d$ is a direct measure of the blur in the reproduced boundary and may be taken as the vertical resolution expressed in pitch units. The actual resolution, V , will be

$$V = \text{vertical resolution} = \frac{s}{d} \quad (10-4)$$

³ R. D. Kell, A. V. Bedford, and G. L. Fredendall, "A Determination of Optimum Number of Lines in a Television System." *RCA Review*, 5, 1 (July 1940).

⁴ It should be noticed here that the same approach is used as in the calculation of the transient response of an amplifier. See Chapter 7. It may be argued that if the scanning process can reproduce the abrupt transition in a step function, it can reproduce any less steep transition. Thus the step function is again taken to be the basic test signal.

An entirely different approach to the problem may be based on the "monoline" resolution, where the reproduction of a narrow white bar is considered. This approach lends itself to ready manipulation by means of the Fourier integral or Laplace transform. See, for example, V. K. Zworykin and G. A. Morton, *Television*. New York: John Wiley and Sons, Inc., 1940, chap. 6.

It might appear, then, that in drawing Fig. 10-3 we have completed the determination of V , but as a practical matter this is not true for we considered only a particular case, where the boundary line between the black and white areas crosses the center of the scanning line. To be more general, we should consider a number of cases, each for a different transition condition, and construct an average rise curve whose normalized rise distance is taken as d . In Fig. 10-4

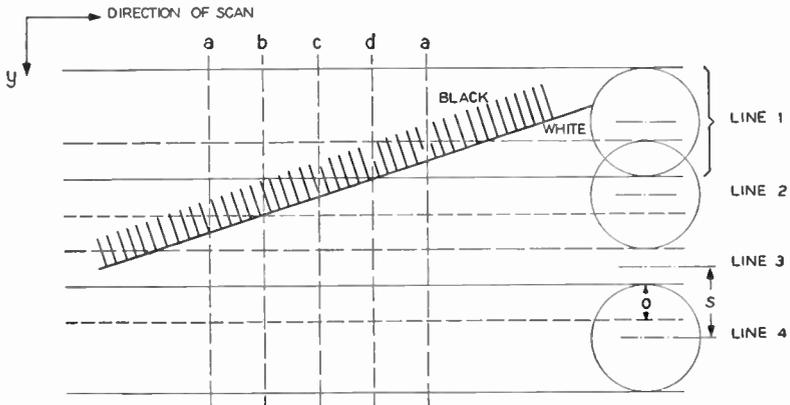


Fig. 10-4. Scanning a black-to-white boundary. Four vertical lines representing different transition conditions are chosen. The reproduced intensity along each vertical line is shown in Fig. 10-5.

the boundary line is shown as well as a number of vertical lines labeled a through d , at each of which a different condition of transition prevails. It should be observed that c represents the same condition as shown in Fig. 10-3 and that the assumed overlap value of 50 per cent is used.

The reproduced intensity along each of the sample vertical lines may then be plotted as in Fig. 10-5. Each of these curves exhibits a different rise distance, d , and hence corresponds to a different value of V . This fact, of course, ties in with our previously defined concept of utilization ratio; the resolution depends upon the relative orientation of the subject matter and the scanning lines.

It may be assumed that the human eye will integrate light along the black-to-white transition curve; hence it is valid to construct a mean rise curve, v , which is the average of the four rise curves shown,

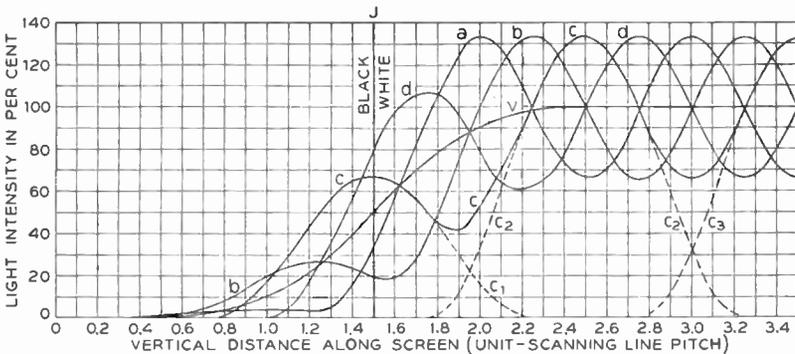


Fig. 10-5. Curves showing the variation of light intensity along the four vertical lines of Fig. 10-4. The rise distance from 0 to 100 per cent of the average curve, v , is a measure of the vertical resolution. (Courtesy of *RCA Review*.)

and we take d to be the normalized rise distance of this mean curve. By eq. (10-4) the vertical resolution will be the reciprocal of this value.

The relationship between the actual rise distance in linear units and n , the number of lines, comes about because s , the interline center pitch, is inversely proportional to n for a given picture height or

$$\text{actual rise distance in linear units} = ds = \frac{dh}{n} \quad (10-5)$$

Then as the number of lines increases, the actual rise distance decreases, indicating a higher value of vertical resolution. In consequence, a curve of vertical resolution plotted against the number of lines will be a straight line.

10-4. Horizontal Resolution

We have seen how vertical resolution may be measured as the reciprocal of the distance required for the scanning system to reproduce a step-function transition in brightness. We now must set up a similar definition for the horizontal resolution of the system. In this case we are interested in the reproduction of a black-to-white transition *along a scanning line*; thus we must orient the boundary line so that it is normal to the lines.

If we assume an infinitesimally narrow reproducing aperture, or an

aperture whose width has been compensated for, nothing in the scanning system proper will affect the response; rather the response will depend solely upon the transient response of the electrical networks which deliver the electrical signal to the CRT grid. Thus the first step in determining horizontal resolution is the evaluation of the transient response of the electrical portions of the system. Then, since the distance traveled by the scanning spot is linearly related to the scanning velocity, we may relate the rise time of the electrical system to the rise distance in the reproduced image. The horizontal resolution, H , may then be determined as the reciprocal of this rise distance.

We immediately are faced with a problem because there is no one single transient response for which all television receivers are designed. In fact, receivers will exhibit widely varying responses, depending to a large degree on their selling price. To circumvent this ambiguity we shall assume the five different idealized receiver characteristics, which are shown in Figure 10-6 as being typical of receivers

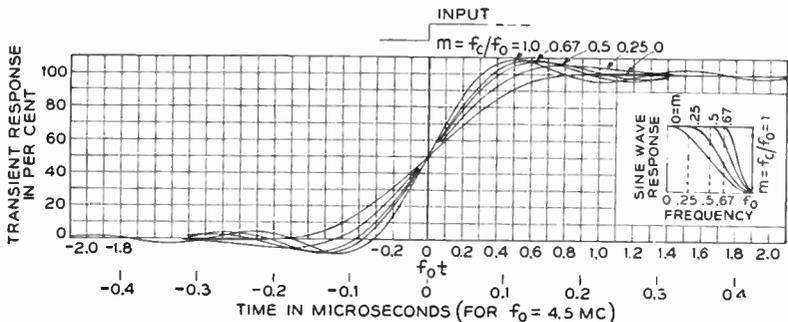


Fig. 10-6. Transient response of five idealized video transmission systems. The rise time of each curve is related to the corresponding horizontal resolution. (Courtesy of *RCA Review*.)

extending from the low- to the high-priced class. As a matter of convenience each curve may be identified by a parameter m defined by

$$m = \frac{f_c}{f_0} \quad (10-6)$$

where f_c and f_0 are, respectively, the upper limit of the mid-band and the zero-response frequency of the corresponding steady-state gain

characteristic. These latter characteristics are plotted in the insert of Fig. 10-6. In each case, a constant delay characteristic is assumed.⁵ Once the set of typical response curves has been assumed, the rise time of each may be measured and converted to rise distance on the CRT screen. To carry out this last step we must calculate the linear velocity of the scanning beam. Thus let it be assumed that the horizontal blanking interval decreases the length of a scanning line by 15 per cent and that the vertical blanking causes a loss of 10 per cent of the total number of lines as far as picture reproduction is concerned. We then have the following relationships:

$$\tau_s = \frac{0.85}{nf_p} \quad (10-7)$$

and

$v_h =$ horizontal scan velocity

$$= \frac{w}{\tau_s} = \frac{wnf_p}{0.85} = \frac{Anf_p}{0.85} h \quad (10-8)$$

and for the interline pitch

$$s = \frac{h}{0.9n} \quad (10-9)$$

Then by means of (10-8) the rise distance corresponding to any rise time in Fig. 10-6 may be determined as a percentage of the picture height, h . Similarly, the actual vertical rise distance may be expressed as a percentage of the picture height by the use of (10-9).

To illustrate the work involved we calculate a typical point for $n = 507$ lines, $f_p = 30$ cycles, and $m = 1$. From Fig. 10-6, the normalized rise time from black to white is 0.6, or the rise time is

⁵ Figure 10-6 demonstrates quite precisely the interrelationship that exists between transient response and the steady-state characteristics of a network, a subject which is discussed in Chapter 7. The steady-state curves shown in the insert are idealized in that (a) they are flat up to a frequency f_c , and (b) they follow a sine curve in decreasing from f_c to f_o , the latter being 4.5 megacycles. With the gain idealized in this manner and with an assumed linear phase characteristic, the corresponding transient response may be calculated by the Fourier integral.

It is instructive to notice the relationship between steepness of cutoff and ring in the transient response. The "ideal" case for which $m = 1$ is identical to the "ideal" amplifier discussed in Chapter 7 and illustrated in Fig. 7-8; it exhibits ring and overshoot. As m decreases, that is, as the cutoff becomes less steep, the transient response exhibits less overshoot and a longer rise time.

$$t = \frac{0.6}{f_0} = \frac{0.6}{4.5 \times 10^6} \quad (10-10)$$

Then the actual rise distance along the cathode-ray tube screen is

$$\begin{aligned} \text{actual rise distance} &= v_h t = \left(\frac{4}{3}\right) \frac{(507)(30)(0.6)}{(0.85)(4.5 \times 10^6)} h \\ &= 0.0032h \end{aligned} \quad (10-11)$$

Similarly, from Fig. 10-5, d is 1.8, then actual rise distance

$$\text{actual rise distance} = 1.8s = \frac{1.8h}{0.9(507)} = 0.0039h \quad (10-12)$$

These two distances may be identified in Fig. 10-7 as the horizontal ($m = 1$) and vertical rise distances, respectively. It follows of course

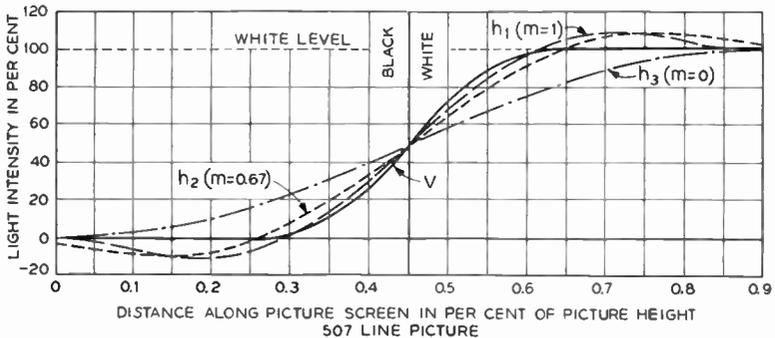


Fig. 10-7. When the scanning velocity is known, the rise times of Fig. 10-6 may be converted to rise distance along the screen, expressed as a percentage of the picture height. (Curves are for $n = 507$ lines. (Courtesy of *RCA Review*.)

that similar curves could be plotted for different values of n so that a direct comparison of the horizontal and vertical rise distance or blur could be made.⁶ The data from several such curves are summarized in Fig. 10-8. Notice that the curves of horizontal resolution which are given by the reciprocal of eq. (10-11) are of the inverse type as would be expected from the equation

$$\text{horizontal resolution} = H \propto \frac{1}{n} \quad (10-13)$$

⁶ For additional curves see R. D. Kell, A. V. Bedford, and G. L. Fredendall, *op. cit.*

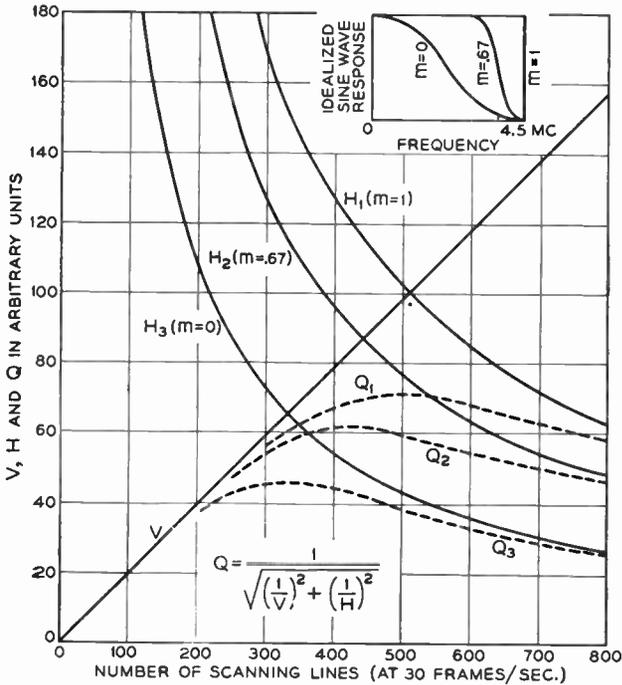


Fig. 10-8. *V* and *H* curves show vertical and horizontal resolution. Equal resolution in both directions occurs where the solid curves intersect. The *Q* curves indicate picture quality. (Courtesy of *RCA Review*.)

In Fig. 10-8 the quantities *V* and *H* are, respectively the reciprocals of actual vertical and horizontal rise distances expressed in arbitrary units rather than as a fixed percentage of picture height. As such, they are each a direct measure of the resolution in the appropriate direction.

Let us now examine the *V* and *H* curves of Fig. 10-8 with the view of picking that value of *n* which will give substantially equal values of vertical and horizontal resolution. Clearly from the curves no such single choice can be made because the receiver response enters in. For example, for the "best" receiver for which *m* = 1, the condition of equal *V* and *H* is provided by an *n* slightly in excess of 500 lines. Similarly for the *m* = 0.67 receiver *n* should be just under 450, and for *m* = 0, *n* should be roughly 330.

10-5. Over-all Resolution

In general, all types of receivers will be in use and on the V and H data alone we cannot make the optimum choice for n . To resolve the difficulty we seek some other quantity, which may give additional information that is independent of the receiver response. Kell, Bedford, and Fredendall have recommended a factor "which is indicative of the over-all picture quality," and which is related to V and H in the following manner: Since V is the vertical resolution, $1/V$ is the vertical blur or rise distance for the black-to-white transition, and similarly $1/H$ is the horizontal blur. Regarding these two blurs as vector quantities at right angles to each other, the magnitude of the resultant blur, B , will be

$$B = \sqrt{\left(\frac{1}{V}\right)^2 + \left(\frac{1}{H}\right)^2} \quad (10-14)$$

Since blur and resolution are reciprocally related, we may take $1/B$ to be the resultant over-all resolution or the picture quality, Q .

$$Q = \frac{1}{B} = \frac{1}{\sqrt{\left(\frac{1}{H}\right)^2 + \left(\frac{1}{V}\right)^2}} \quad (10-15)$$

The justification for treating the component blurs as vectors is that the results are in good agreement with those of subjective tests.

With the over-all resolution, Q , defined, it may be plotted against n for the various receiver types as shown by the dashed curves in Fig. 10-8. Inspection of these curves shows that a value of n slightly in excess of 500 gives equal values of V and H as well as maximum quality for the best type of receiver ($m = 1$). For the poorer receivers the vertical resolution will be higher than the horizontal resolution. We shall assume, then, that subject to all our assumptions, such as $\cos^2 \theta$ light distribution across a scanned line, 30-cycle picture frequency, corrected aperture effects, and 50 per cent adjacent line overlap, the optimum number of lines is in the vicinity of 500, a value which gives essentially equal V and H and good over-all quality for a good receiver. The precise value of n is affected by certain other considerations which are taken up in the next section.

10-6. Circuit Factors

Once the general order of the number of scanning lines has been determined from the consideration of resolution, we must look to the requirements of the over-all system to aid in choosing the actual value. In the last chapter it was specified that the 2 to 1 interlaced scan pattern is standard; hence *n must be an odd number*.⁷ It was also specified that a rigid lock-in between the line and frame frequencies is required. In the next chapter we shall see that this tie-in of the two frequencies is provided by an electronic frequency-dividing circuit which gives most satisfactory operation when the frequency division ratio consists of a small number of integral factors. From our previous work we know that the line and frame frequencies are related by the factor *n* for

$$f_l = nf_p$$

Thus *n* is the required division ratio of the counting circuit. The two requirements on *n*, that it shall be odd and consist of a small number of integral factors, are satisfied by the following numbers which lie in the general vicinity of 500:⁸

| <i>n</i> | <i>factors</i> |
|----------|---|
| 441 | $7 \times 7 \times 3 \times 3$ |
| 495 | $11 \times 5 \times 3 \times 3$ |
| 507 | $13 \times 13 \times 3$ |
| 525 | $7 \times 5 \times 5 \times 3$ |
| 567 | $7 \times 3 \times 3 \times 3 \times 3$ |

Of these several values 525 has been adopted as the standard number of lines for the commercial television system.

One final point should be mentioned. The choice of 525 lines for commercial practice is based, as we have seen, on a maximum video bandwidth of 4 to 4.5 megacycles. Some attention has been directed to the practice in several European countries of using a number of lines in excess of 800. It can be seen immediately from our work in the preceding sections that such a value with a 4.5-megacycle bandwidth would give a vertical resolution far in excess of the horizontal value. The European practice circumvents this shortcoming by

⁷ Cf. Section 3-8.

⁸ D. G. Fink, *Television Standards and Practice*. New York: McGraw-Hill Book Company, Inc., 1943, p. 229.

using a wider bandwidth, and the resulting broad-band high-resolution system is intended for large-screen theater reception rather than private reception in the home. The consensus in the United States is that 525-line 4.5-megacycle bandwidth standards provide adequate pictures on the screen sizes used in home-type receivers and provide room for improved picture quality resulting from technical advances under those standards.

10-7. Test Pattern and Resolution

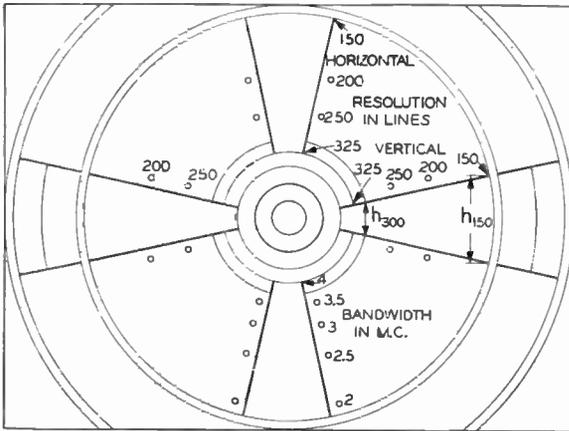
The accepted practice in telecasting stations calls for the transmission of test pattern for a period of at least 15 minutes prior to the transmission of regular program material. The test patterns shown in Fig. 10-9 and 10-10 serve to identify the transmitting station by means of call letters, channel number, and other insignia, and are designed to enable the operator of a television receiver to adjust his set for optimum reception of the transmitted signal. Our chief con-



Fig. 10-9. A typical station test pattern. (Courtesy of Columbia Broadcasting System.)



(a)



(b)

Fig. 10-10. Another form of test pattern. (a) As viewed on a receiver. (Courtesy of American Broadcasting Company.) (b) Outline of the pattern showing the calibration points used in determining resolution and receiver bandwidth.

cern at the moment is to see how the test pattern may be used to check the resolution of a television system. Since the pattern affords several other checks on system performance, we shall discuss them also for the sake of completeness.

Reference to the two figures shows that the pattern has two large concentric circles, the outer one white, the inner one black. The circles have diameters in the ratio of 4 to 3 and hence are proportional to the standard picture aspect ratio. Thus the proper picture height is obtained by adjusting the vertical sweep gain until the top and bottom of the black ring just touch the top and bottom of the mask surrounding the face of the cathode-ray tube. In a similar fashion the horizontal sweep gain is set to have the edges of the outer ring touch the vertical edges of the mask. This procedure ensures that the receiver sweep will give the proper aspect ratio. Furthermore since the shape of the circles and the wedges are known, they also serve as a check on the linearity of the receiver sweep circuits. For example, in Fig. 10-10a the circles appear as ellipses, indicating that the vertical sweep is not linear. Screwdriver adjustments can be made on the receiver linearity controls to remedy the situation.

The test patterns also provide a means of properly setting the receiver contrast control. This is done with the help of the five concentric rings which range in shade from white to black in four equal increments. The contrast control should be adjusted until the apparent difference in brightness between any two of the adjacent rings is constant.

Since the focus control provides a means of setting the resolution of the reproducing system, it should be adjusted so that the black and white lines which make up the four horizontal and vertical wedges can be distinguished as near to the center of the pattern as possible. Under this condition the resolution is maximum, as will be explained in the paragraphs which follow.

The four wedges are drawn in such a way that the vertical and horizontal resolution may be reckoned directly in terms of the number of elements transmitted along a vertical line, N_v ,⁹ or in terms of the number of effective lines with which the whole system is operating. In order to see this we first consider the two identical horizontal wedges which are located on either side of the bull's-eye in Fig. 10-10. Inspection of the diagram shows that each wedge is formed of 16

⁹ See chap. 2.

black and 15 white bars arranged alternately; thus along any vertical line across these wedges, the bars represent 31 picture elements. Furthermore, since the wedges are tapered toward the center of the pattern, the 31 picture elements are crowded into a smaller and smaller space as the center of the pattern is approached. It is this property of the wedges which permits their use as a check on the system resolution. The height of the wedges where they intersect the large black circle (h_{150} in Fig. 10-10b) is $1/4.84$ times the picture height, h . Since there are 31 picture elements in the distance h_{150} , the total number of elements along a vertical line of length h drawn through the intersections of the wedges and the black circle is

$$N_v = 31 \frac{h}{h_{150}} = 31(4.84) = 150 \text{ elements} \quad (10-16)$$

Hence, if the television system can just resolve the outer edges of the wedges, it is capable of displaying 150 picture elements along a vertical line, or we can say that it is effectively operating with 150 lines, even though the actual number of scanning lines is 525.

At the point where the wedges intersect the white ring of the bull's-eye, the 31 elements occupy a vertical distance which is one-half h_{150} ; hence if the system can just resolve the wedges where they intersect the white ring, it is handling N_v elements along a vertical line given by

$$N_v = 150 \frac{h_{150}}{h_{300}} = 150(2) = 300 \text{ elements} \quad (10-17)$$

or the effective system resolution is 300 lines.

It may be inferred from these data that the maximum vertical resolution is determined by that point along a horizontal wedge at which the black and white bars can just be distinguished from each other. Since the wedges taper at a uniform rate, the resolution is proportional to location of that point along the wedge. To aid in interpreting the results white dots are placed above and below the wedges at the 200 and 250 line points.

With this information in hand we may estimate the vertical resolution of the receiver pattern shown at a in Fig. 10-10. Along the right-hand wedge the black and white bars are clearly discernible to a point approximately midway between the 200 and 250 line calibration dots; hence the vertical resolution is about 225 lines.

Horizontal resolution may be determined in exactly the same

manner from the two vertical wedges which are identical to their horizontal counterparts, except for their orientation and the fact that they do not extend as far in toward the center of the bull's eye. Once again, resolution is expressed in terms of an equivalent number of scanning lines, or N_v , and the upper and lower pairs of dots on either side of the *upper* vertical wedge correspond to resolutions of 200 and 250 lines, respectively.

We have seen in our previous work that horizontal resolution is related to the receiver transient response which, in turn, is related to the receiver bandwidth; hence the resolution wedges may also be used to determine the effective receiver bandwidth. If N_v is the effective number of lines, then for a picture of aspect ratio 4 to 3, the number of elements along a horizontal line will be

$$N_h = \frac{4}{3} N_v$$

Then, using the methods of Chapter 2, we assume the frequency of the signal required to produce these elements to be

$$f = \frac{1}{2} \frac{N_h}{(\tau_u)_h} \quad (10-18)$$

but in the commercial standards, $(\tau_u)_h = 53.3 \mu\text{sec}$; therefore

$$f = \frac{1}{2} \left(\frac{4}{3} \right) \frac{N_v}{(53.3)10^{-6}} = 0.0125 \times 10^6 N_v \quad (10-19)$$

With this equation the effective bandwidth of the receiver may be determined from either vertical wedge. In the actual pattern four calibration points are provided along the lower wedge so that the bandwidth may be read directly. Reading downward these dots correspond to bandwidths of 3.5, 3, 2.5, and 2 megacycles, respectively. Then, reading from Fig. 10-10*a*, we see that the receiver illustrated has an effective bandwidth of approximately 3 megacycles. This value may be verified by eq. (10-19).

It must be stressed that all stations do not use the same details in the test pattern chart (Fig. 10-9). The effective number of lines and bandwidth may be calculated for each of them in the same manner that has been illustrated in this chapter.

CHAPTER 11

SYNCHRONIZATION

Frequent reference has been made in the preceding chapters to the use, shape, and generation of the synchronizing signals which serve to lock-in the vertical and horizontal sweep at both ends of the television system. In commercial telecasting practice, where picture quality is at a premium, the requirements for timing, duration, and shape of these synchronizing pulses are much more rigid than in any of the previously described closed systems. In the present chapter we shall investigate these requirements, synthesize the composite synchronizing signal which meets them, and consider the several circuits required to generate that signal.

11-1. Synchronization Requirements

In brief, the synchronizing signal must meet the following primary conditions:

- (1) It must provide positive synchronization of the sweep circuits.
- (2) The horizontal and vertical components of the composite signal must be susceptible to separation by simple electrical circuits.
- (3) The signal must be of such a form that it may be combined with the picture and blanking signals to modulate a common carrier at the transmitter.

Additional requirements demanded by the sweep circuits and by the need for stability of sweep are:

- (4) The synchronizing pulses must occur at the end of a sweep in order to initiate flyback.
- (5) There shall be a rigid lock-in between the horizontal and vertical sync components to reduce pairing in the interlaced scan pattern.
- (6) Provision shall be made to maintain horizontal synchronization at all times, even during the vertical synchronizing interval.

It will be observed that several of the above requirements are met

by the synchronizing signals used in the Type III closed system of Chapter 8. We shall, therefore, be able to draw on the material of that chapter.

11-2. Amplitude Components of the Composite Signal

The third requirement listed above demands that the sync, picture, and blanking signals be combined into a single composite video signal, which may be used to amplitude-modulate the carrier of the video transmitter. Under the current set of standards these components in the combined signal are kept apart by amplitude separation. Thus of the 100 per cent peak-to-peak value of the combined signal, the upper $25\% \pm 2.5\%$ is reserved for the blanking signals. The remaining 75 per cent are available for the picture signal proper. Black or blanking corresponds to the dividing line between sync and picture components. This distribution of the amplitude components is illustrated in Fig. 9-2. Since the signal must remain a single-valued function, the sync and picture components must be separated in time as well as in amplitude. Such a separation is quite feasible since a sync pulse occurs at the end of a trace when the picture is blanked out, a condition also illustrated in Fig. 9-2. The amplitude and time location of the synchronizing pulses may be designated as follows: They lie in the "blacker than black" amplitude region¹ and occur during the blanking intervals at the end of each line and each field when no picture signal is being transmitted.

11-3. Wave-form Components of the Composite Sync Signal

We have already seen in Chapter 9 that differentiating and integrating circuits are capable of distinguishing between wide and narrow pulses; thus the second primary requirement may be satisfied by pulses of this type. From the relationship between the line and field frequencies of the interlaced scan it follows that V , the field period, is $n/2$ times as long as H , the line period. It is only natural, then, that the longer pulses be used for synchronizing the vertical sweep, and the narrow ones, the horizontal.

We next consider in more detail the action of the circuits which may be used to distinguish between the wide and narrow pulses.

¹ See Section 8-3.

11-4. Differentiating Circuit²

Let a voltage e of any wave shape be applied to a condenser, then

$$q = Ce_c \tag{11-1}$$

and

$$i = \frac{dq}{dt} = C \frac{de_c}{dt} \tag{11-2}$$

that is, the current flowing through the condenser is proportional to the time derivative of the voltage across the condenser. If now a small resistor be placed in series with the condenser—small enough so that for every frequency component in the signal the resistance is negligibly small with respect to the reactance of the condenser—then to a good approximation the applied and condenser voltages are equal, or

$$e_c \approx e \tag{11-3}$$

and the drop across the resistor will be

$$e_R = iR \approx RC \frac{de}{dt} \tag{11-4}$$

Notice that the requirement $R \ll X_c$ implies a small resistance and a small capacitance. Therefore, in a series R - C circuit of short time constant, the voltage across the resistor is proportional to the time derivative of the applied voltage. Such a circuit is termed a “differentiating circuit” and is illustrated in Fig. 11-1a.³ If a square wave

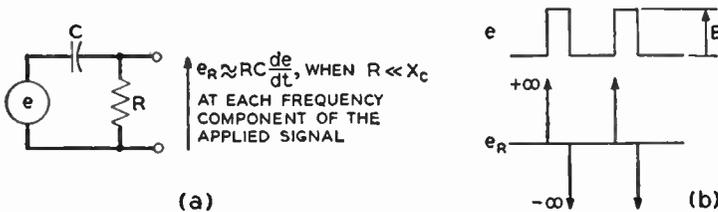


Fig. 11-1. The differentiating circuit. (a) Circuit. (b) Idealized wave forms.

² J. G. Clarke, “Differentiating and Integrating Circuits.” *Electronics*, **17**, 11 (November 1944).

³ A series circuit of inductance and resistance may also be used for differentiating. A short time constant, L/R , is required and the differentiated voltage appears across the inductor. The circuit finds little use in television applications because of the resistance associated with the inductor and because neither element can block a d-c component of applied voltage.

is applied at the input terminals of the network, eq. (11-4) indicates that the output will have the shape shown at *b* in the diagram. At each leading edge the slope of the input wave, de/dt , is high and causes a pulse of considerable height and of zero width extending in the positive direction. At each trailing edge a similar pip is produced, but in the negative direction.

The idealized wave form of Fig. 11-1*b* cannot be realized in a practical circuit, first, because it is impossible to produce a square wave with zero rise time and, second, because the restriction on the relative values of R and X_c cannot be met for the higher order frequency components of the square wave. For television applications, then, where a square pulse is used for synchronizing, it is convenient to consider the differentiating network in another light. We may apply the concepts of Chapter 4, where we discussed the charge and discharge of an R - C network with a pulse of finite duration, τ , applied. Thus if the square pulse of width τ be applied to the network, the output will be given by the equation

$$e_R = E\epsilon^{-t/RC}, \quad 0 \leq t \leq \tau \quad (11-5)$$

for the leading edge, and by

$$e_R = -E\epsilon^{-t/RC}, \quad t > \tau \quad (11-6)$$

for the trailing edge. In both of the last two equations E is the peak value of the applied square wave. If, now, RC be made small with respect to τ , say $RC = \tau/25$, the output pips will be of short duration. This condition is illustrated by the curves of Fig. 11-2. The actual shape of the output wave may be plotted by evaluating eqs. (11-5) and (11-6). Comparison of the idealized and actual wave forms shows that in the latter case the pip height is constrained to the value E and the duration is lengthened from zero to approximately $5RC$ seconds.

From the point of view of the synchronizing circuits, the important feature of the differentiating circuit is that each leading edge of the applied square pulse produces a positive pip, and each trailing edge a negative one. These conditions prevail even though the applied pulse is wider than that illustrated in Fig. 11-2*a*. The wave forms of Fig. 11-2*a* are still idealized inasmuch as physical square waves always have a finite rise and decay time, as shown at *b* in the diagram. Analysis of the differentiating circuit with such a non-ideal wave

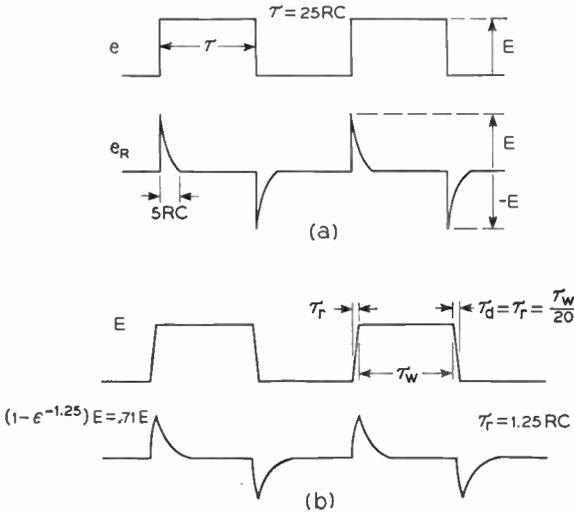


Fig. 11-2. (a) Wave form obtained from a differentiating circuit when an ideal (zero rise time) square wave is applied. (b) Differentiated output of a non-ideal square wave.

applied shows that during the linear rise interval, the output voltage builds up exponentially, resulting in a rounded leading edge of the differentiated output. The wave form shown in the figure assumes that the circuit time constant is $\frac{1}{25}$ times the duration of the flat portion of the square wave, τ_w , and that the ratio of rise time to τ_w is $\frac{1}{20}$. An even sharper output pulse may be obtained if a shorter time constant relative to the pulse width is used, but at the same time the amplitude of each pip is reduced. It may be observed that the net effect is the same as that previously described: each leading edge of the square wave produces a positive pip of short duration, which may be used for synchronizing purposes.

11-5. Integrating Circuit

Let a current i of any shape be passed through a condenser of capacitance C . Then

$$q = Ce_c = \int i dt \tag{11-7}$$

and
$$e_c = \frac{1}{C} \int i dt \tag{11-8}$$

Equation (11-8) indicates that the voltage across a condenser is proportional to the time integral of the current flowing through the condenser; hence if we can make this current proportional to e , an applied voltage, the condenser voltage will be proportional to the integral of e . Let a large resistor, R , be placed in series with C and let e be applied across the combination. Then for those components of the signal for which

$$R \gg X_c \quad (11-9)$$

we obtain
$$i \approx \frac{e}{R} \quad (11-10)$$

and
$$e_c \approx \frac{1}{RC} \int e dt \quad (11-11)$$

Such an R - C circuit, whose output voltage is derived from the condenser, is termed an "integrating circuit" and is illustrated in Fig. 11-3a. Notice that the inequality (11-9) implies the use of a large

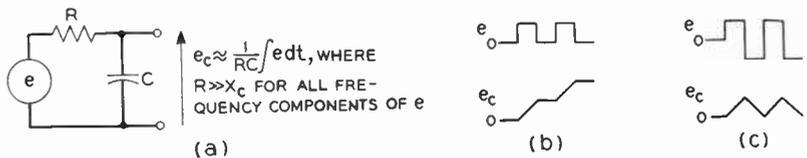


Fig. 11-3. The integrating circuit. (a) Circuit. (b) Idealized wave form with a positive square wave applied. (c) Idealized wave form with alternating square wave applied.

resistor and a large capacitance to which the reactance is inversely related. Thus the time constant, RC , of the circuit must be large in comparison to the width of the applied square pulse.⁴

Once again, when a square wave is applied to the input terminals of a practical network, the inequality (11-9) cannot be satisfied for all the frequency components present and it becomes more convenient to consider the circuit on a charge and discharge basis. Thus, by eq. (4-1), the condenser voltage for the duration of the applied square pulse is

$$e_c = E(1 - e^{-t/RC}), \quad t \leq \tau \quad (11-12)$$

⁴ A series circuit of L and R may also be used. A long time constant is required and the output voltage is developed across the resistor.

where E is the magnitude of the applied pulse. At $t = \tau$, e_c will have some value E_c , depending upon the ratio of τ to RC , and for the interpulse period the output voltage will be by equation (4-3)

$$e_c = E_c \epsilon^{-t/RC}, \quad t > \tau \tag{11-13}$$

The response of the network to pulses of different widths may now be illustrated. The pulse widths shown in Fig. 11-4 are chosen

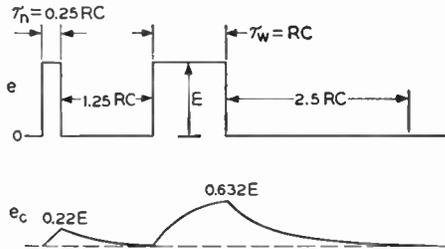


Fig. 11-4. Actual response of the integrating circuit to square pulses of different widths.

arbitrarily to be $\tau_n = 0.25RC$ and $\tau_w = RC$. Then for the narrow pulse the output rises to

$$e_c]_{\tau_n} = E(1 - \epsilon^{-0.25}) = 0.22E \tag{11-14}$$

During the first interpulse interval of duration $1.25RC$ the condenser discharges to

$$e_c]_{1.25RC} = 0.22E \epsilon^{-1.25RC} = 0.22E(0.286) = 0.063E \tag{11-15}$$

a value which is assumed to be negligible in the diagram. Upon application of the wide pulse the condenser charges from zero to

$$e_c]_{\tau_w} = E(1 - \epsilon^{-1}) = 0.632E \tag{11-16}$$

and thereafter the voltage drops exponentially to zero. In the diagram the terminal point is for a discharge interval of $2.5RC$ seconds and the corresponding voltage is

$$e_c]_{2.5RC} = 0.632E \epsilon^{-2.5} = 0.632E(0.082) = 0.0518E \tag{11-17}$$

which again may be assumed negligible.

The curves of Fig. 11-4 show that the peak value of the output voltage of the integrating circuit is dependent upon, but not necessarily directly proportional to, the width of the applied square pulse. This fact shows how the second basic requirement on the synchro-

nization system is met. The vertical synchronizing information may be separated from a mixture in time of horizontal and vertical pulses by use of an integrating circuit and a biased sweep generator. This is illustrated in Fig. 11-5, where the integrated output voltage is used to overcome the cutoff bias on the vertical-sweep generator. The narrow horizontal sync pulses produce an output of insufficient ampli-

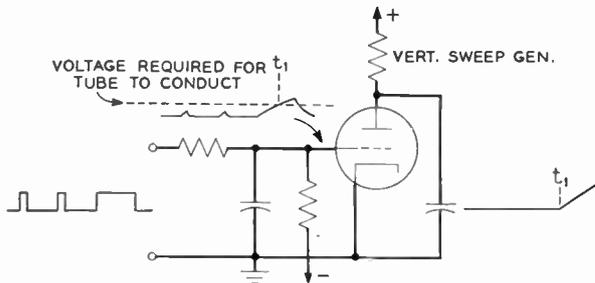


Fig. 11-5. The integrating circuit allows the vertical sync information to be separated from the composite sync signal.

tude to fire the discharge tube. The longer vertical pulse produces an output pip which can drive the tube above cutoff, thereby initiating the condenser discharge or flyback portion of the vertical-sweep cycle. We see, then, that the square pulses of different widths are susceptible to separation by electrical networks. We next consider how these pulses must be disposed in time in order to provide proper synchronization of both the horizontal- and vertical-sweep systems.

DEVELOPMENT OF THE COMPOSITE SYNC SIGNAL

We have seen that square pulses lying in the blacker-than-black region of the composite video signal are used for synchronizing purposes. We now consider the time distribution and width of the horizontal and vertical pulses.

11-6. Horizontal Sync Pulse

The period H of the commercial system is extremely short.

$$H = \frac{1}{f_t} = \frac{1}{nf_p} = \frac{1}{(5.25 \times 10^2)(3 \times 10^1)} = 63.5 \mu\text{sec} \quad (11-18)$$

In this short time interval the entire horizontal sweep cycle of scan and flyback must take place. The various sweep generators which we have discussed are synchronized by causing the sweep-generating condenser to discharge; consequently the horizontal sync pulse, which initiates this action, must occur during the H interval *at the end* of the sweep part cycle. Of the total interval approximately 8 per cent is

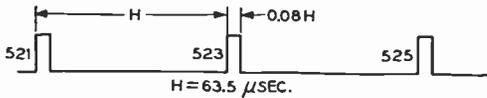


Fig. 11-6. Timing and width of the horizontal sync pulses.

reserved for the horizontal synchronizing pulse. Specifications for the shape and duration of the horizontal sync pulses are shown at E in Fig. 11-11. In order that a stable horizontal sweep be maintained, the *leading edges* of adjacent horizontal sync pulses must occur at intervals of H . The train of horizontal pulses required to meet these specifications is illustrated in Fig. 11-6.⁵

11-7. Vertical Sync Pulses

In comparison to H the vertical or field interval, V , of the commercial system is relatively long for

$$V = \frac{1}{f_f} = \frac{1}{2f_p} = \frac{1}{2(30)} = 16.67 \text{ millisecc} \quad (11-19)$$

Of this interval the amount allowed for the vertical synchronizing pulse is approximately $3H$ or

$$3H = 3 \left(\frac{2}{n} \right) V = 0.0114V \quad (11-20)$$

Since these pulses serve to initiate the vertical retrace, they must occur at the end of each field and the leading edges of adjacent pulses must be separated by intervals of V . The train of vertical sync

⁵ It should be realized that forms of synchronizing pulses other than the square wave could be used successfully in a television system. An excellent summary of certain alternate proposals made to the N.T.S.C. may be found in D. G. Fink, *Television Standards and Practice*. New York: McGraw-Hill Book Company, Inc., 1943. In the present chapter the discussion is limited to the synchronizing wave forms which have been standardized in the United States of America.

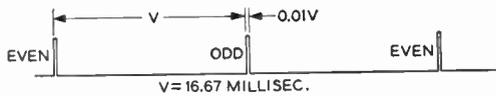


Fig. 11-7. Timing and width of the vertical sync pulses.

pulses is shown in Fig. 11-7. It should be noted that the time scale is *not* the same as that used for the horizontal pulses in Fig. 11-6.

11-8. Interlace Requirements

Having set up the requirements on the two types of synchronizing pulses, we must now determine how they are spaced relative to each other, that is, we must, in effect, combine the pulses shown in the last two figures onto a common time scale and adjust their lateral displacement so that they will properly produce the 2 to 1 interlaced scan. The basis for this lateral adjustment is that the first field must end

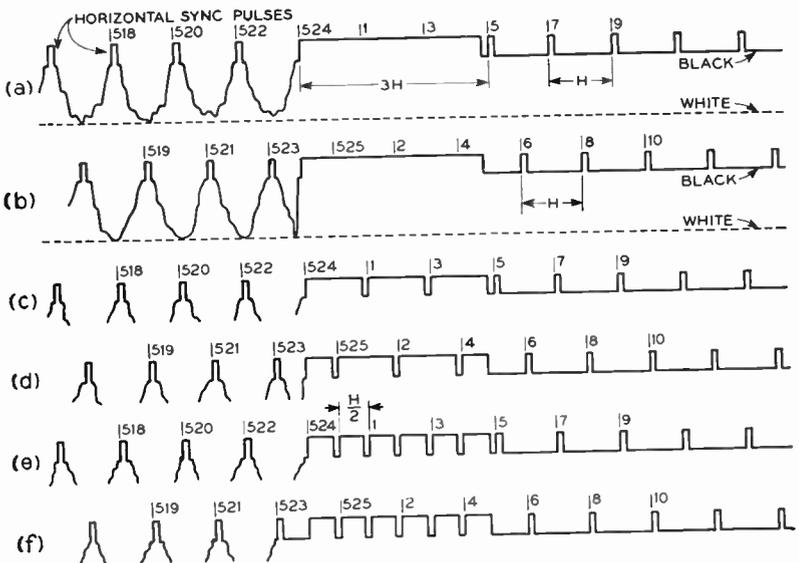


Fig. 11-8. Development of the composite sync signal. (a) End of an even field. (b) End of an odd field. (c) End of an even field. The vertical pulse is serrated to provide a rising wave front at intervals of H . (d) Same as (c), but for the odd field. (e) End of an even field. Serrations at twice the line frequency make all vertical pulses identical. (f) Same as (e), but for the odd field.

in the middle of a horizontal line interval. To aid in identifying the various pulses in the discussion which follows it is convenient to assign a number to each of them. The basis for the numbering system has been established in Fig. 3-24. We number the lines in order from 1 to 525, proceeding from top to bottom of the complete raster. The horizontal pulse which terminates a line will have the same number as that line.

We shall arbitrarily assume that the vertical sync pulse which terminates the first field occurs in the middle of the 525th line. The vertical sync pulse which terminates the second field will therefore occur at the end of line number 524. The first field will consist of odd-numbered lines only and may be termed the "odd field." Similarly, the second field contains all the even-numbered lines plus the second half of the 525th line and may be termed the "even field." Referring to Fig. 11-6 and 11-7 we see that an "odd" vertical pulse must begin at the end of line 524. These phase relationships between horizontal and vertical sync pulses are illustrated at *a* and *b* in Fig. 11-8. Notice that the pulses are drawn so that the leading edges of the odd and even vertical pulses lie along a common vertical line in the diagram.

11-9. Serration

It may be seen directly from the figure that if the waves shown are differentiated, no positive synchronizing pips will occur at intervals of H for the duration of the vertical pulses; hence horizontal synchronization is lost during the width of the vertical pulse, a condition which contradicts the sixth requirement listed at the beginning of the chapter. To overcome this difficulty we introduce notches, or serrations, into the vertical pulses as shown at *c* and *d* in Fig. 11-8, each notch being placed so that a leading or rising edge occurs at the end of each H interval. Then, when the wave is differentiated, a positive pip occurs at each leading edge and proper synchronization is provided for lines 1 and 3 at *c* and for lines 525, 2, and 4 at *d*, and the specified condition is satisfied.

Notice that the even vertical pulse is notched only twice, whereas the odd vertical pulse has three serrations: the serration requirements are different for successive vertical sync pulses. This situation produces a rather severe requirement upon the sync generator that must develop these waves. From the point of view of the circuitry in-

volved it would be desirable to have all the vertical pulses identical, whether they terminate an odd or an even field. It is more or less apparent that if each vertical pulse be serrated five times with a leading edge occurring at every interval of $H/2$, all the vertical pulses will be identical and will appear as shown at *e* and *f* in Fig. 11-8. We must now determine whether the additional notches will upset the horizontal synchronization.

In Chapter 4 we saw that it is desirable to drive a sweep generator from a synchronized impulse oscillator of some form in order to protect the kinescope screen. Then, if the synchronizing signals are lost for any reason, the sweep will continue at the frequency of the free-running oscillator. Consider the action of such an oscillator in the presence of the wave shown at *e*. A blocking oscillator is assumed and it is forced to operate at line frequency by the differentiated output of that wave. The various wave forms involved are shown in Fig. 11-9. Observe that as the result of differentiation a

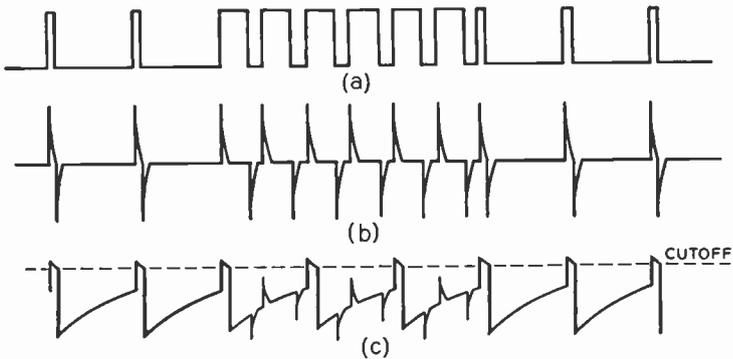


Fig. 11-9. With proper adjustment the pips occurring at the center of a line have insufficient amplitude to fire the blocking oscillator. (a) Composite sync at the end of an even field. (b) Differentiated composite sync. (c) Combined blocking oscillator grid voltage and differentiated pips.

positive pip occurs at each rising edge of the original wave. The grid voltage of the blocking oscillator consists of an exponential component rising toward cutoff, on which are superimposed the differentiated pips. With proper adjustment of the grid voltage the "in-between" pips have insufficient amplitude to fire the oscillator, and synchronization occurs only at the proper H intervals. A similar argument also obtains for the odd field. It is, therefore, quite feasible

to serrate the vertical pulses at twice the line frequency to ease the requirements on the sync generator circuits. Protection against misfire caused by the $H/2$ interval pips is afforded by the exponential characteristic of the oscillator grid voltage.

Having justified the need for the serrations in the vertical pulse, we now find it convenient to change our notation slightly: We shall now refer to each pulse between serrations as a vertical sync pulse, a nomenclature which has found more or less universal acceptance. On this basis vertical synchronization is provided by six vertical pulses which occur at the end of each field. In effect, we replace one long serrated pulse by a train of six narrower pulses, which occupy the same interval. The dimensions of the newly defined vertical sync pulses are shown at d in Fig. 11-11.

11-10. Equalization of Vertical Fields

Although the problem of maintaining horizontal synchronization during the vertical sync intervals has been solved by replacing a single wide pulse at field frequency by a train of six narrow pulses at twice

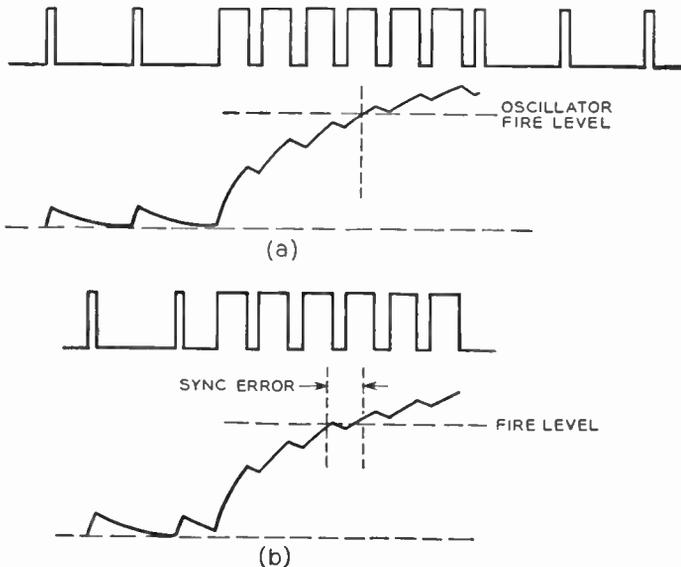


Fig. 11-10. Unequal integrated voltages at the end of successive fields may cause a shift in the point of synchronization. (a) Integrated voltage for the even field. (b) Integrated voltage for the odd field.

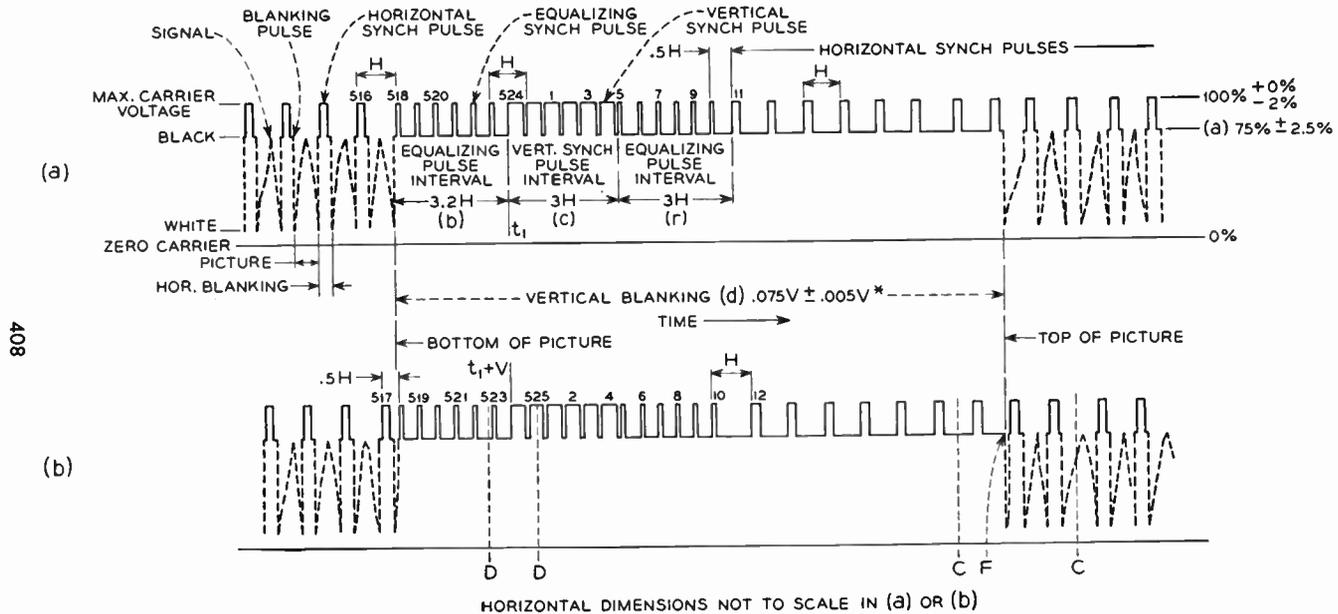
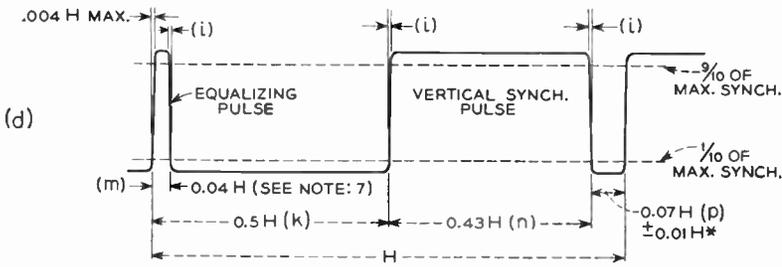
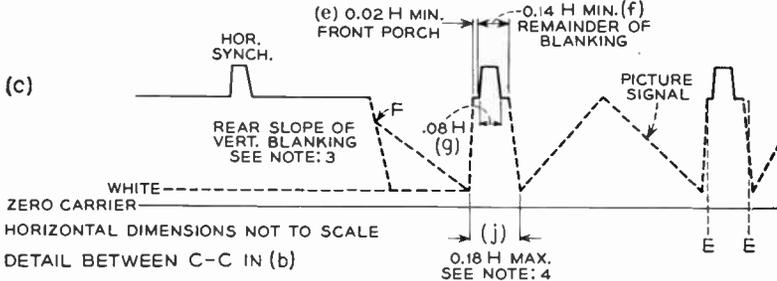
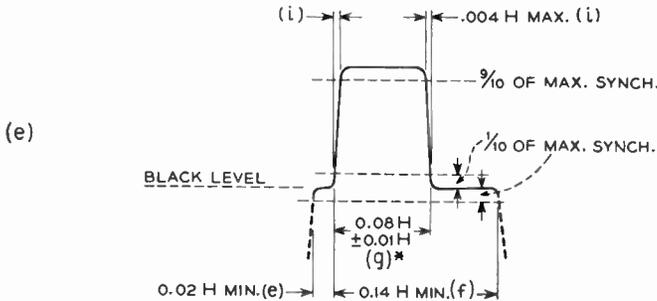


Fig. 11-11. The composite video signal, showing details of the synchronizing and blanking components.



DETAIL BETWEEN D-D IN (b)



DETAIL BETWEEN E-E IN (c)

Notes: 1. H = time from start of one line to start of next line. 2. V = time from start of one field to start of next field. 3. Leading and trailing edges of vertical blanking should be complete in less than $0.1H$. 4. Leading and trailing slopes of horizontal blanking must be steep enough to preserve min. and max. values (e + f) and (j) under all conditions of picture content. 5. *Dimensions marked with an asterisk indicate that tolerances given are permitted only for long time variations, and not for successive cycles. 6. For receiver design, vertical retrace shall be complete in $0.07V$. 7. Equalizing pulse area shall be between 0.45 and 0.5 of the area of a horizontal sync pulse.

Fig. 11-11 (cont.).

the line frequency, we are still faced with a difficulty in regard to vertical synchronization. From Fig. 11-8e and f we see that in the even field the vertical sync train lags horizontal pulse 522 by an interval H , whereas in the odd field the interval is only $H/2$. When these voltages are applied to the integrating circuit, the integrating condenser has twice as long to discharge the effects of horizontal pulse 522 in the even field as in the case of pulse 523 in the odd field. Thus the voltage on the integrating condenser is not the same at the end of the odd and even fields, and a slight shift in the vertical synchronizing point may occur and cause pairing of the fields in the raster. This condition is illustrated in Fig. 11-10. The solution to this difficulty is to make the synchronizing pulses identical for a period of, say, $3H$ before the start of each vertical pulse train. Then the integrating condenser will always have the same initial voltage at the start of the vertical sync interval of either field. Several methods of equalizing the pre-vertical pulse condition suggest themselves, the most obvious being the one which eliminates all pulses for an interval $3H$ at the end of each field. This is a poor solution because we would lose horizontal sync. One practicable solution is to insert a series of "equalizing pulses" in the pre-vertical sync interval as shown at *a* and *b* in Fig. 11-11. These equalizing pulses occur at twice the line frequency and affect the horizontal sweep system in the same manner as do the vertical sync pulses. Their width is one-half that of a horizontal synchronizing pulse, which ensures that the energy supplied to the integrating circuit per line interval remains constant, a fact which may be verified from the following considerations: The energy supplied by each square voltage pulse is proportional to the square of the area under the pulse. Thus the energy supplied during line 518 is proportional to the square of the area under a horizontal pulse. In line 520 it is proportional to the square of the area under two equalizing pulses. Then, since each equalizing pulse is of half the area of a horizontal pulse, the total areas in lines 518 and 520 are equal and so is the energy supplied.

We may summarize the effect of the six equalizing pulses preceding each vertical pulse train as follows: They equalize the voltage on the integrating condenser at the end of each field to ensure positive vertical synchronization at the proper intervals and proper interlace in the final scan pattern. It may be observed in the diagram that six additional equalizing pulses follow each vertical synchronizing train.

These serve in a similar manner to equalize the integrating condenser voltage after the vertical impulse oscillator has been triggered and the integrating condenser discharged. Furthermore, at the time the standards were set up it was believed that future development might lead to trailing-edge synchronization; so these last six equalizing pulses were inserted to provide a degree of flexibility which might be required.

Two additional points should be stressed about the composite sync wave form shown in Fig. 11-11. The line interval, H , is used to indicate the time interval between leading edges of successive pulses. The location of any given scanning line, however, is not from leading pulse edge to leading pulse edge but as shown in Fig. 11-12. The actual position of the line is displaced slightly to the right of successive leading horizontal sync pulse edges. The reason for this is, of course, that one line comprises a sweep and a retrace, the latter being initiated by the leading sync pulse edge.

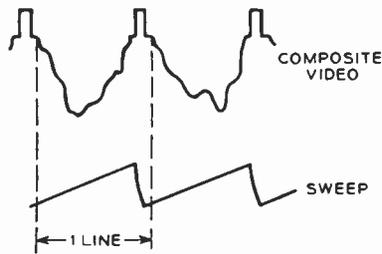


Fig. 11-12. Any single line is of duration H but is displaced to the right from the leading edges of successive horizontal sync pulses.

It should be stressed that receivers will operate with a composite sync signal which is less complicated than the R.M.A. signal shown in Fig. 11-11. With a simpler sync signal, however, either a more complicated receiver is required or less positive operation results. We have already seen that the wiser choice is to maintain performance and simplify the receiver at the expense of the transmitter facilities.

11-11. Blanking

It is common practice to develop the blanking signals in the same generator that forms the composite sync or supersync signal. It is therefore desirable to discuss the blanking signals at the present time. We may state the purpose of the blanking signals briefly: They serve to blank out or cut off the scanning beam for those portions of the sweep cycle which are devoted to flyback. As a measure of precaution the blanking signals extend from a short time before to some time after the duration of the corresponding sync pulse. This ensures

that the entire sweep system is in stable operation when the picture information is applied to the control grid of the cathode-ray tube. The dimensions of the horizontal and vertical blanking intervals may be read directly from Fig. 11-11. It will be noticed that some 18 lines per field are lost for picture presentation because of the vertical blanking. These "lost" lines are the inactive lines, n_i , defined in Chapter 3.

THE SYNCHRONIZING SIGNAL GENERATOR

The generation of the composite synchronizing signal described above requires a rather complex network of components. A number of different circuit configurations may be used to form a generator which can synthesize the required wave. To lend continuity to our discussion we shall consider primarily one form of sync generator which has been described in the literature.⁶

11-12. General Requirements

We first consider some general aspects of the sync generator. In Chapter 9 it was pointed out that commercial practice in the United States requires that the pickup equipment adhere to more rigid standards than the receiver; the camera blanking and sweep must occur in a shorter interval than they do at the receiver in order that picture data are always available when the receiver is ready for it. This practice requires that sync and blanking signals separate from those used in the composite video signal be developed by the synchronizing signal generator. In all, then, five principal voltages must be delivered by the generator, which is the basic clock for timing all the sweeps in the entire television system:

- | | | |
|-------------------------------|---|--|
| (1) Supersync | } | Add with the picture voltage to form the composite video signal. |
| (2) Composite line blanking | | |
| (3) Horizontal drive | } | Used to synchronize and blank the camera equipment. |
| (4) Vertical drive | | |
| (5) Composite camera blanking | | |

It is essential for a stable scanning raster that the frequencies of the

⁶ A. V. Bedford and J. P. Smith, "A Precision Synchronizing-Signal Generator." *RCA Review*, V, 1 (July 1940).

several pulse components in the supersync wave be maintained to a high degree of precision. Inspection of Fig. 11-11 shows that there are three basic repetition rates present in the wave: The vertical and equalizing pulses occur at twice the line frequency, or 31.5 kilocycles; the horizontal pulses occur at line frequency, or 15.75 kilocycles; and the group of vertical pulses at the end of each field occur at field frequency, or 60 cycles. These three basic frequencies are integrally related.

$$31.5 \text{ kilocycles} = 2(15.75) \text{ kilocycles} = n \text{ } 60 \text{ cycles} \quad (11-21)$$

It is possible, then, to lock in these frequencies by using a master oscillator operating at 31.5 kilocycles and by deriving the lower frequencies from it with the use of counters or frequency-dividing circuits.

We have also seen that the effects of hum on the scanning pattern may be minimized if the 60-cycle field frequency is locked in with the power line. This lock-in may be effected by the system diagrammed in Fig. 11-13. The 60-cycle component derived from the counter

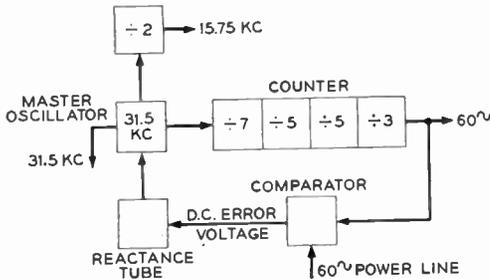


Fig. 11-13. The master timing system of the sync-signal generator.

circuit is compared with the power-line frequency in the comparator, which delivers a d-c voltage proportional to the error between the two frequencies. The error voltage controls a reactance tube which adjusts the master-oscillator frequency until the compared frequencies are identical. Since all the frequencies in the supersync are derived from the master oscillator, this system automatically corrects for drift and maintains the proper relationships between all the components.

11-13. Counter Circuit⁷

One of the primary components of the synchronizing-signal generator is the counter circuit or step-charge rectifier, which serves to divide the master-oscillator frequency of 31.5 kilocycles by 2 to give line frequency and by 525 to give the field frequency. We shall assume that the input pulses are of square shape and that the circuit to be considered is that of Fig. 11-14a. The peak value of the

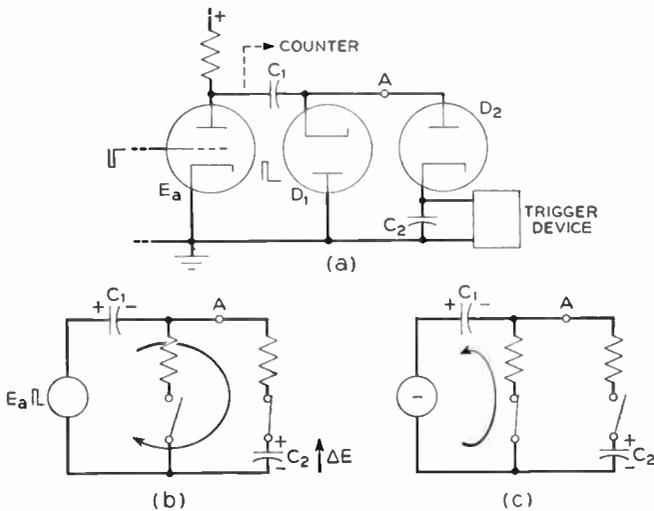


Fig. 11-14. The diode-clamped counter circuit. (a) Basic counter circuit. (b) Part cycle with C_1 and C_2 charging. (c) Part cycle with C_1 discharging. C_2 retains its charge.

amplified positive pulse from plate to ground is E_a . Consider the operation of the circuit. During the duration of the positive pulse, the point A is positive with respect to ground and the following condition obtains: The clamping diode, D_1 , will be nonconducting but D_2 , the charging diode, will conduct, allowing condensers C_1 and

⁷ Another form of counter which employs a resonant circuit to stabilize a blocking oscillator has been proposed to reduce the number of tubes required in the sync generator. See A. R. Applegarth, "Synchronizing Generator for Electronic Television." *Proc. IRE*, **34**, 3 (March 1946); M. Silver, "High Ratio Multivibrator Frequency Divider," *Radio Electronic Engineering Edition of Radio and Television News*, **13**, 1 (July 1949).

C_2 to charge. This condition of operation is illustrated by the equivalent circuit of Fig. 11-14*b*. Assuming that the charging time constant of the circuit, T_c , is very small compared to τ , the width of the applied pulse, we see that the two condensers in series will charge up to the full value of E_a . The distribution of voltage between the two condensers will be proportional to the inverse of their capacitances; hence during τ the condenser C_2 receives an increment of voltage

$$\Delta E = \frac{C_1}{C_1 + C_2} E_a \quad (11-22)$$

At the end of τ the applied voltage drops to zero. The positive voltage on C_2 causes the charging diode, D_2 , to open circuit, thereby isolating C_2 , which will hold the increment of voltage ΔE given by eq. (11-22). Since the point A is no longer positive relative to ground, D_1 conducts, allowing C_1 to discharge completely. Figure 11-14*c* illustrates this second part-cycle of operation. D_2 prevents C_2 from discharging, but D_1 clamps A to ground potential. The circuit is then ready for the next positive pulse.

On the second pulse C_1 and C_2 again receive charge through the charging diode D_2 but the *increment* of voltage received by C_2 will be less than that indicated by (11-22) because C_2 charges from an initial charge Q_0 , given by

$$Q_0 = \Delta E C_2 = \frac{C_1 C_2}{C_1 + C_2} E_a \quad (11-23)$$

After the second pulse D_2 opens, holding the charge on C_2 , but C_1 is discharged through the clamping diode. The process is repetitive; each applied pulse in succession adds a smaller and smaller voltage increment on C_2 . Between pulses C_1 is discharged. The operating cycle stops when the sum of the voltage increments on C_2 is sufficient to fire the triggering device shown at a in Fig. 11-14. The trigger, generally a blocking oscillator or multivibrator, serves a dual function; it discharges C_2 , readying the whole circuit for the next cycle of operation, and it produces an output pulse. If k input pulses are required to allow C_2 to fire the trigger device, the ratio of input to output pulse-repetition rate is k , the circuit count ratio.

To analyze the operation we must find in what manner the voltage on C_2 builds up as a function of the *number of applied pulses*. During any single pulse, the buildup is exponential but we assume $\tau \gg T_c$

so that during each pulse interval a steady-state condition is reached. Now let

- E_a = amplitude of the applied pulse
- i = number of pulses applied
- ΔE_i = increment of voltage on C_2 caused by the i th pulse
- q_i = charge circulating because of the i th pulse
- $= C_2 \Delta E_i$
- E_i = total voltage on C_2 after i pulses have been applied
- Q_i = total charge on C_2 after i pulses have been applied
- E_0 = initial voltage on C_2 before application of the first pulse⁸

Now under the assumed condition C_1 and C_2 charge to the full value of E_a during each pulse; hence for the i th pulse we may write

$$\frac{q_i}{C_1} + \frac{Q_{i-1} + q_i}{C_2} = E_a \quad (11-24)$$

Therefore $q_i \left(\frac{1}{C_1} + \frac{1}{C_2} \right) = E_a - \frac{Q_{i-1}}{C_2}$

or $q_i = \frac{C_1 C_2}{C_1 + C_2} (E_a - E_{i-1}) \quad (11-25)$

Then, by the definition of ΔE_i ,

$$\begin{aligned} \Delta E_i &= \frac{q_i}{C_2} = \frac{C_1}{C_1 + C_2} (E_a - E_{i-1}) \\ &= r_1 (E_a - E_{i-1}) \end{aligned} \quad (11-26)$$

where $r_1 = \frac{C_1}{C_1 + C_2} \quad (11-27)$

E_i is defined as the total voltage on C_2 after i pulses have been applied; hence we may write

$$E_i = E_{i-1} + \Delta E_i \quad (11-28)$$

and, substituting for ΔE_i from (11-26), we have

$$\begin{aligned} E_i &= E_{i-1} + r_1 (E_a - E_{i-1}) \\ &= r_1 E_a + (1 - r_1) E_{i-1} \end{aligned}$$

⁸Subsequent work will show that certain trigger devices are unable to discharge C_2 completely; hence we allow for this initial charge. A similar analysis, which neglects the E_0 term, has been published. See A. Easton and P. H. Odessy, "Design of Counter Circuits for Television." *Electronics*, 21, 5 (May 1948).

$$E_i = r_1 E_a + r E_{i-1} \quad (11-29)$$

where

$$r = 1 - r_1 = \frac{C_2}{C_1 + C_2} \quad (11-30)$$

In order to convert (11-29) into a form which may be recognized as a familiar mathematical function, we evaluate it for several values of i , a procedure which is straightforward, though lacking in elegance. Thus, for $i = 0$, no pulse has been applied and $E_{i-1} = E_0$; therefore,

$$i = 0 \quad E_i = E_0$$

$$i = 1 \quad E_i = E_1 = r_1 E_a + r E_0$$

$$\begin{aligned} i = 2 \quad E_i = E_2 &= r_1 E_a + r E_1 \\ &= r_1 E_a + r(r_1 E_a + r E_0) \\ &= r_1 E_a(1 + r) + r^2 E_0 \end{aligned}$$

$$\begin{aligned} i = 3 \quad E_i = E_3 &= r_1 E_a + r E_2 \\ &= r_1 E_a + r[r_1 E_a(1 + r) + r^2 E_0] \\ &= r_1 E_a(1 + r + r^2) + r^3 E_0 \end{aligned}$$

$$\begin{aligned} i = 4 \quad E_i = E_4 &= r_1 E_a + r E_3 \\ &= r_1 E_a + r[r_1 E_a(1 + r + r^2) + r^3 E_0] \\ &= r_1 E_a(1 + r + r^2 + r^3) + r^4 E_0 \end{aligned}$$

and, finally,

$$i = i \quad E_i = r_1 E_a(1 + r + r^2 + \dots + r^{i-1}) + r^i E_0 \quad (11-31)$$

The series within the parentheses of eq. (11-31) may be recognized as a geometric progression of common ratio r and the sum of its first i terms is⁹

$$1 + r + r^2 + \dots + r^{i-1} = \frac{1 - r^i}{1 - r} = \frac{1 - r^i}{r_1} \quad (11-32)$$

Thus E_i becomes

$$\begin{aligned} E_i &= r_1 E_a \frac{(1 - r^i)}{r_1} + r^i E_0 \\ &= E_a \left[1 - \left(\frac{C_2}{C_1 + C_2} \right)^i \right] + E_0 \left(\frac{C_2}{C_1 + C_2} \right)^i \end{aligned} \quad (11-33)$$

⁹ See, for example, W. L. Hart, *College Algebra*. Boston: D. C. Heath and Company, 1926.

and we have the final expression for the voltage on C_2 as a function of the number of applied pulses. The relative importance of the two terms depends upon the relative values of E_a and E_0 . We shall presently see that E_0 is determined largely by the type of trigger circuit used to discharge C_2 .

Curves of E_i v. i with r as the parameter and E_0 assumed zero are plotted in Fig. 11-15. Inspection of the curves will verify the follow-

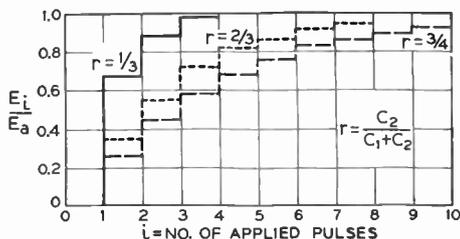


Fig. 11-15. Charging curves for the diode-clamped counter. E_0 , the initial voltage on C_2 , is assumed to be zero.

ing statement: As r approaches unity, the step heights become equalized, the voltage per step decreases, and the number of steps required to reach a given voltage increases.

In the design of a counter stage it is desirable that the k th step, which fires the trigger circuit, have a magnitude in the order of 1 or 2 volts in order to ensure positive firing action. It follows that the optimum capacitance ratio, r , depends upon E_a , the firing voltage, and the count ratio. Where receiving-type tubes are used, these factors limit the maximum count ratio to approximately 15. It is because of this limitation that the number of lines in the television system is chosen such that it may be broken down into a number of *small* factors.

It is of interest to note that one limitation on count ratio k in the counter circuit may be traced directly to the fact that the voltage increment per step decreases as the number of applied pulses is increased. This inherent limitation may be overcome by replacing the clamping diode, D_1 , of Fig. 11-14 by a triode. The modified triode-clamped counter is shown in Fig. 11-16.¹⁰ The primary differ-

¹⁰ C. E. Hallmark, "An Improved Counter-Timer for Television." *Radio-Electronic Engineering Edition of Radio News*, 9, 1 (July 1947).

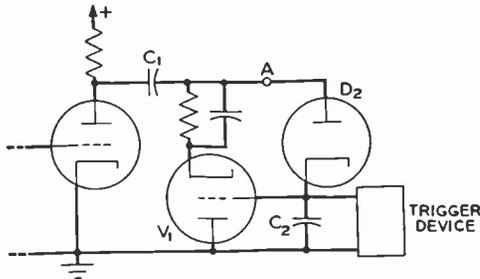


Fig. 11-16. Triode-clamped counter circuit. (Courtesy of *Radio and Television News*.)

ence in this, as compared to the diode-clamped circuit, is that the control grid of the clamping triode, V_1 , is connected to the positive end of C_2 . Because of this, C_1 does not fully discharge between input pulses but retains sufficient voltage so that ΔE on C_2 remains constant. Proper design of the circuit allows the full count of 525 which is required in the sync generator to be obtained in a single stage. It is inevitable, however, that the maximum voltage increment on C_2 per step will be less than $E_a/525$. For values of pulse magnitude that may be obtained in practice, this generally holds ΔE_k for the firing step below the 1-volt value, which is desirable from the standpoint of firing stability. Current practice retains a cascaded series of diode-clamped circuits.

Generally speaking, there are three principal types of trigger devices which may be used to discharge C_2 and to produce the output pulse. These are the thyatron, the blocking oscillator, and the modified multivibrator. Of the three which are illustrated in Fig. 11-17 the thyatron is the least expensive. It is seldom used in television applications because of the instability of its firing point. The blocking oscillator provides positive action but provides little, if any, control of the output pulse shape. The third circuit at *c* incorporates a separate discharge tube in conjunction with an electron-coupled multivibrator. The large number of components required is the price paid for stability plus control of the output wave form.

Consider the operation of the circuit at *a*, which employs the thyatron. The firing voltage of the tube is determined by the bias control P_1 . As a typical example, say that a count ratio of 3 to 1 is required, E_a is 50 volts, $C_2 = 3C_1$, and the tube drop across the

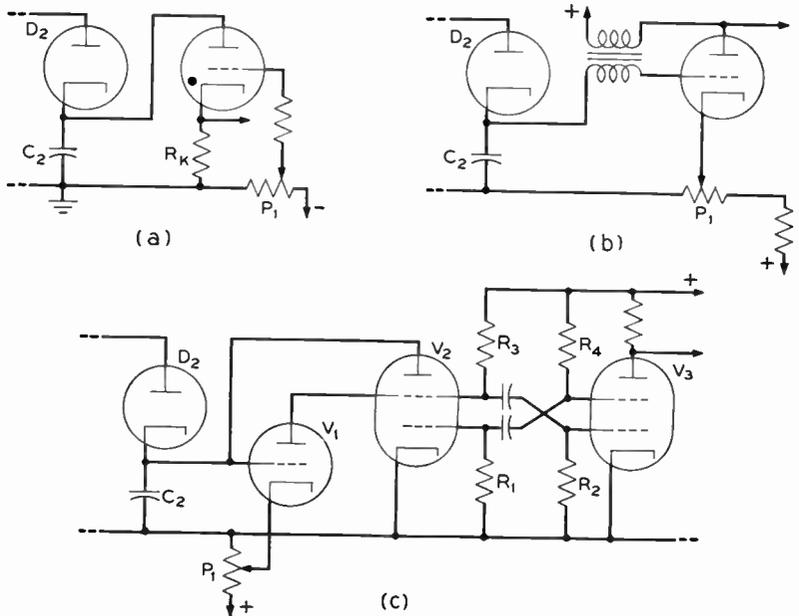


Fig. 11-17. Representative trigger devices. C_2 is the charging condenser of the associated counter circuit. (a) Thyatron trigger circuit. (b) Blocking oscillator trigger circuit. (c) Triggering is provided by V_1 in conjunction with the electron-coupled multivibrator.

thyatron, when it is conducting, is 20 volts. Then, from (11-30), $r = 0.75$. Let us calculate the value of voltage appearing on C_2 after the application of the second and third pulses. Since the lowest voltage which appears across the thyatron is the tube drop itself, $E_0 = 20$ volts then, from (11-33),

$$i = 2 \quad E_2 = 50[1 - (0.75)^2] + 20(0.75)^2 = 30.7 \text{ volts}$$

$$\text{and } i = 3 \quad E_3 = 50[1 - (0.75)^3] + 20(0.75)^3 = 37.4 \text{ volts}$$

In order to have a count ratio of three, the thyatron should fire between these values of voltage, say at 34 volts. Reference to the firing characteristic of the thyatron will then show what value of bias is required to have the tube fire at this voltage. The output pulse is developed by the discharge current flowing through R_K .

The operation of the blocking oscillator is quite similar to that just

described, except that the firing point is determined by the negative throw of the oscillator on the grid side. Since this is generally negative, C_2 will be left with its upper terminal negative and E_0 of eq. (11-33) will be a negative quantity.

In the third circuit of Fig. 11-17 V_1 is biased beyond cutoff by the proper adjustment of P_1 . Furthermore, the multivibrator is designed so that normally V_2 is also cut off. Then, when the k th pulse has been applied, the voltage on C_2 is just sufficient to overcome the bias on V_1 , which conducts. Notice that the plate current of V_1 flows through R_3 and causes a drop in the screen grid voltage of V_2 . This drop has two effects. The multivibrator throws, and V_2 conducts, causing C_2 to discharge. The multivibrator then completes its operating cycle, V_1 and V_2 are cut off, and the circuit is readied for the next triggering pulse from C_2 .

Observe that, in effect, we have a multivibrator which is synchronized by pulses delivered by C_2 and V_1 . The electron-coupled plate of V_2 serves the auxiliary purpose of discharging C_2 . It should be clear from a consideration of the synchronizing concept that the free-running inter-output-pulse period of the multivibrator must be greater than the period between the synchronizing pulses developed across C_2 . Stated in other terms this means that

$$\tau_2 < \frac{k}{f}$$

where τ_2 is the interval for which V_2 is cut off under free-running conditions, f is the frequency of the pulses applied at the counter input, and k is the count ratio.

In typical circuits of this type, the plate of V_1 almost completely discharges C_2 and E_0 may be taken as zero. Under these circumstances it is possible to develop the design, using the curves of Fig. 11-15 in place of eq. (11-33).

11-14. Reactance Tube

A second component of the complete sync generator which we shall discuss is the reactance tube, which serves to convert a varying d-c error voltage into a corresponding change in oscillator tuning capacitance. The need for such an element was described in connection with Fig. 11-13. Consider the circuit shown in Fig. 11-18. A triode is shown in place of a pentode to simplify the schematic

arrangement. Let a source of alternating voltage, E , be applied between plate and ground. The impedance of the R - C branch is assumed to be high so that I_1 is negligibly small in comparison to the plate current $I_p = g_m E_g$. Furthermore, R is assumed to be small relative to X_c , thus I_1 leads E by nearly 90° . Since E_g is the drop $I_1 R$, E_g and $g_m E_g$ are in phase with I_1 and so lead E by nearly 90° . The ratio of $g_m E_g$ to E is the admittance viewed by E , which may be seen

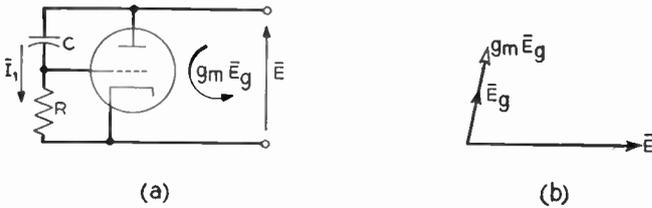


Fig. 11-18. The reactance tube. (a) One form of basic circuit. (b) Vector diagram. $g_m E_g$ is assumed to be very much greater than I_1 .

to have a capacitive component whose magnitude is proportional to g_m . Thus the magnitude of this capacitive component can be varied by controlling the bias of the stage, which, in turn, varies g_m . As the bias voltage becomes less negative, both g_m and the effective shunt capacitance increase. Thus a d-c voltage may vary the effective input capacitance between the plate and cathode of the tube. The quantitative relationship between the input impedance and g_m may be obtained by direct application of the equivalent plate-circuit theorem to the circuit of Fig. 11-18a.¹¹ It may also be demonstrated that the tube behaves as a variable inductance if C is replaced by an inductance. In the latter case a large blocking condenser is required in series with the inductance to isolate the grid circuit from the d-c plate voltage. The connection of the reactance tube and the tank circuit of its associated oscillator is shown in Fig. 11-19.

11-15. Comparator Circuit

A third basic component of the sync generator is the comparator circuit, which serves to deliver a d-c error voltage whose magnitude is proportional to the difference in frequency of the two input signals.

¹¹ See, for example, H. J. Reich, "Theory and Applications of Electron Tubes." New York: McGraw-Hill Book Company, Inc., 1944.

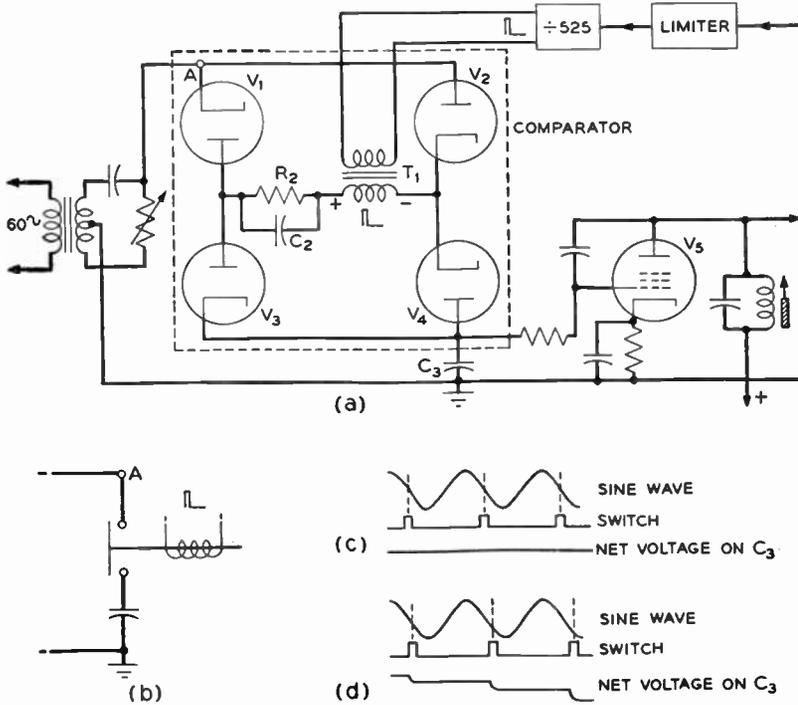


Fig. 11-19. The comparator circuit. (a) Circuit diagram. The comparator proper is enclosed by the dotted line. (b) Equivalent circuit of the comparator itself. (c) Wave forms when both inputs to the comparator have the same frequency and proper phase relationship. (d) Wave forms when the square wave has a lower frequency than the sine wave. (Courtesy of *RCA Review*.)

The basic comparator circuit is shown in Fig. 11-19 and may be identified as that portion of the circuit which is enclosed by the dotted line. Consider the operation of the circuit. With no voltages applied, the four switching diodes V_1 , V_2 , V_3 , and V_4 are nonconducting and there is an open circuit between the point A and ground. When a 60-cycle square pulse derived from the counter circuit is fed to the circuit through the transformer, all four of the tubes are forced to conduct and a conduction path is established from A through C_3 to ground. The flow of current during the same interval also charges C_2 , making its right-hand side positive. At the end of the 60-cycle pulse, C_2 places a reverse polarity on the switching diodes; they cease

conducting and the conduction path from *A* to ground is broken. Notice, then, that the action of the four tubes and the applied pulse may be replaced by a relay as shown in Fig. 11-19*b* for purposes of discussion. For the duration of the applied pulse the contacts are closed; at all other times they are open.

Now consider what happens when a 60-cycle sine wave derived from the power line is applied to the circuit between *A* and ground. Obviously, this sine wave will charge C_3 but only during those intervals when the relay contacts are closed. The amount of voltage received by C_3 during these intervals depends upon the difference in frequency of the square and sine wave inputs and their relative phase. At *c* in the diagram both waves are of the same frequency and, with the switching interval astride the crossover point of the sine wave, the net voltage on C_3 remains zero.

If for some reason the frequency of the master oscillator decreases, the square wave output of the counter chain will also drop in frequency. The wave forms corresponding to this condition are shown at *d*. Notice that the relay contacts are closed only during the negative portions of the sine wave, so C_3 will charge up negatively. It may be observed from the circuit that the voltage across C_3 effectively varies the bias on the reactance tube V_5 . We have already seen that a more negative bias lowers g_m and also the shunt capacitance presented by the reactance tube. Thus the negative voltage on C_3 raises the oscillator frequency and the counter output approaches its proper value until finally the stable condition indicated at *c* obtains.

Notice the importance of the relative phase of the sine wave. If it were 180° out of phase from the wave, as shown at *d*, a lowered counter output frequency would charge C_3 positively and drive the system even farther away from its stable point.

A similar line of reasoning will show that if the master oscillator shifts its frequency upwards, C_3 will charge positively and will lower the oscillator frequency by increasing the shunt capacitance.

While the comparator and frequency lock-in circuit exhibits excellent stability, it cannot easily be adjusted the first time it is placed in operation. Two principal adjustments must be made: The master oscillator frequency must be set very close to its correct value, and the phase of the sine wave must be set properly. To aid in the first adjustment a shorting switch is normally wired in shunt with C_3 . With this switch closed the reactance tube and oscillator are isolated from

the comparator circuit, the bias on V_5 is fixed, and the oscillator may be adjusted by means of the tuning slug in its tank coil. It will be observed that any percentage of change in oscillator frequency will reflect the same percentage of change in the counter output, but, nevertheless, from the viewpoint of adjusting the system it is convenient to compare the oscillator frequency directly with a standard of 31.5 kilocycles; it is easier to observe a difference of several cycles than of a fraction of a cycle, even though the percentage difference is the same in both cases.

Once the master-oscillator frequency is adjusted closely to its correct value, the phase of the counter output and sine wave should be compared. This may be done with the help of an oscilloscope. The potentiometer, P_1 , of the phase shift network permits a phase adjustment over a range of roughly 180° .¹² If a complete phase reversal is required, the leads from the transformer secondary to the shifting network may be reversed. Once adjusted, the entire feedback loop, comprising the oscillator, counter, comparator, and reactance tube, exhibits excellent stability and lock-in with the power supply.

11-16. Voltage-adding Circuit

A fourth basic component of the sync generator is the mixing or voltage-adding circuit, which functions to add two voltages point by point. Frequent reference has been made in our previous work to this circuit, which consists of two vacuum tubes operating into a common plate load. We now consider the action of the circuit in some detail.

Basically the operation is quite simple. In each tube the plate current is proportional to the applied grid voltage. Plate currents of both tubes flow through a common resistor; hence the output voltage is proportional to the sum of the two currents. If the operation is linear, the two input wave forms are effectively amplified, added, and inverted.

A practical consideration arises when the tubes used are triodes. From Fig. 11-20 it may be seen that V_1 sees an effective load consisting of R_L shunted by r_{p2} , the plate resistance of the second tube. This combined load will be always less than r_{p2} , a value which will

¹² For a discussion of the phase shifting network see section 11-22.

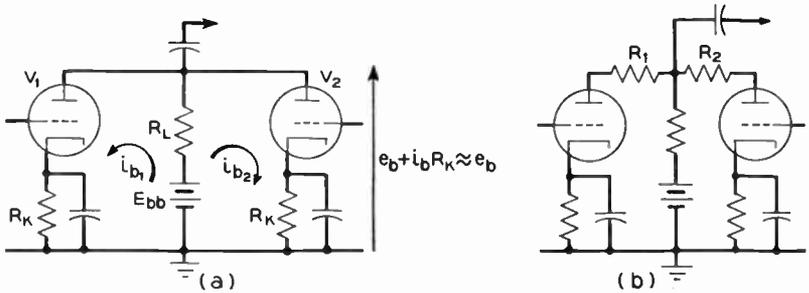


Fig. 11-20. The mixer, or voltage-adding, circuit. (a) Basic circuit. (b) Resistors R_1 and R_2 are added to increase linearity when triodes are used.

give poor linearity, for recall that in a triode the linearity improves as the load resistance is raised. Furthermore, as the grid voltage on V_2 is varied, r_{p2} will also vary. Thus V_1 works into a low-resistance variable load. This difficulty may be overcome by isolating the tubes from each other with linearizing resistances, R_1 and R_2 , shown in Fig. 11-20b. Notice that these resistors will act to reduce the magnitude of the output voltage. When pentodes are used in the circuit, r_p is large and R_L is only a small part of the total circuit resistance. Furthermore, in pentodes the linearity of operation is practically independent of the load resistance, so the linearizing resistors are not required.

We shall consider the analysis of the circuit from two points of view: first when the operation is linear, and second when the operation is on the curved portion of the tubes' characteristics. In the former case the use of the equivalent plate circuit is justified and the analysis is quite simple. We shall assume that pentodes are used, that R_1 and R_2 are not included, and that the two tubes have identical characteristics. Then, reading directly from the equivalent circuit of Fig. 11-21b, we see that the output voltage delivered across R_L is

$$E_o = -g_m(L_1 + L_2) \frac{\frac{r_p}{2}}{\frac{r_p}{2} + R_L} R_L \quad (11-34)$$

or if the constant voltage form is preferred, we note that $\mu = g_m r_p$, and the expression becomes

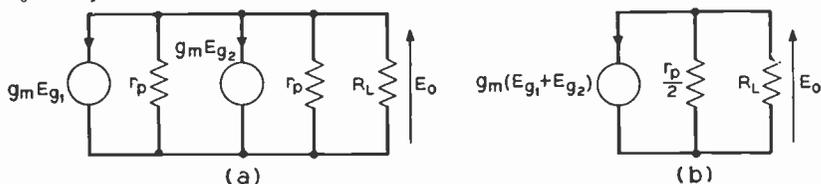


Fig. 11-21. Equivalent circuits for a voltage adder employing pentodes. (a) Constant-current equivalent circuit. (b) Simplified equivalent circuit.

$$E_o = -\mu(E_1 + E_2) \frac{R_L}{r_p + 2R_L} \tag{11-35}$$

Under certain circumstances, when the tubes are operated in the region of nonlinearity, a graphical analysis of the mixer circuit is preferred. We first determine the quiescent or operating point of the tubes when no signals are applied to the grids. From the basic circuit, Fig. 11-20a, we see that at any instant of time the voltage between the cathode and common plate connection of the tubes will be (neglecting the bias voltage developed across R_k)

$$e_b = E_{bb} - (i_{b1} + i_{b2})R_L \tag{11-36}$$

where

e_b = total instantaneous plate voltage,

i_{b1} = total instantaneous plate current of V_1 ,

i_{b2} = total instantaneous plate current of V_2 ,

E_{bb} = supply voltage.

Since the vacuum tube is generally a nonlinear device, we cannot express i_{b1} and i_{b2} as explicit functions of e_b , and we resort to the static characteristic curves, which may be represented by

$$i_b = f(e_b, e_c) \tag{11-37}$$

Equations (11-36) and (11-37) are simultaneous equations and may be solved graphically. Equation (11-36) is plotted on the same co-ordinates as the static characteristic curves. Then the current which corresponds to any set of grid and plate voltages is indicated by the intersection of the straight line (11-36) and the appropriate static curve. It is by this method that we determine the operating point, for since (11-36) is true at any instant, it must be true when the currents and voltages are at their quiescent values. Hence we may replace e_b by E_{bo} and i_b by I_{bo} , and the equation becomes

$$E_{bo} = E_{bb} - (I_{bo1} + I_{bo2})R_L \tag{11-38}$$

If we assume identical tubes and similar bias voltages,

$$I_{bo1} = I_{bo2} = I_{bo} \tag{11-39}$$

and (11-38) becomes

$$E_{bo} = E_{bb} - 2R_L I_{bo} \tag{11-40}$$

whence

$$I_{bo} = \frac{E_{bb} - E_{bo}}{2R_L} \tag{11-41}$$

Equation (11-41) plots as a straight line with voltage intercept equal to E_{bb} , current intercept of $E_{bb}/2R_L$, and slope $-1/2R_L$. This line is the usual d-c load line for a resistance of magnitude $2R_L$ and is shown in Fig. 11-22. Under quiescent conditions $e_c = E_{cc}$, the bias voltage, and the O point is located at the intersection of the load line and the

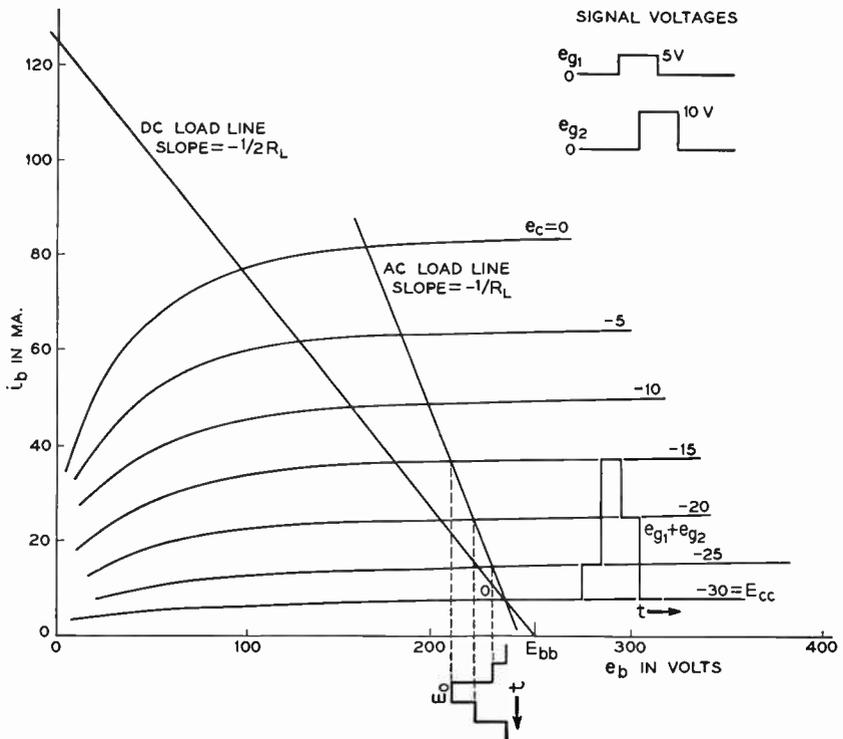


Fig. 11-22. Graphical solution of the voltage adder. The effect of nonlinearity may be observed.

static curve corresponding to the bias voltage. Notice that our results thus far may be interpreted in the following manner: The total d-c plate current from both tubes, $2I_{b0}$, flowing through a resistance R_L , is equivalent to the d-c plate current from a single tube flowing through a resistance of magnitude $2R_L$.

With the O point located we may then draw through it the a-c load line, which, by eq. (11-36), is seen to have a slope of $-1/R_L$. The output currents and voltage may then be determined in the conventional manner as shown in Fig. 11-22. The various voltages used in the diagram are chosen to illustrate how two square waves may be added in the circuit and to show the type of distortion encountered because of nonlinearity.

11-17. Delay Networks

In assembling the various components into the synchronizing-signal generator we shall find need for a unit which is able to delay a train of narrow square pulses by a time interval $n\tau$. It must be realized at the outset that repetitive waves which have widths in the order of microseconds exhibit an extremely wide frequency spectrum. If such a wave of width δ and of period T be expanded into a Fourier series, it may be shown that the amplitudes of the frequency components vary as the $(\sin x)/x$ function as shown in Fig. 11-23. The

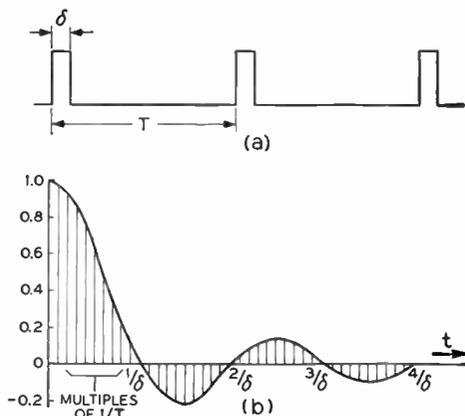


Fig. 11-23. Characteristics of the repeating square pulse. (a) A train of narrow square pulses. (b) The amplitudes of the square-pulse frequency spectrum are proportional to the $(\sin x)/x$ function. In the actual spectrum all terms are positive.

frequencies present are all multiples of $1/T$ and the amplitude envelope passes through zero at those frequencies which are multiples of $1/\delta$. It has been determined experimentally that the pulse may be reproduced by those components whose frequencies extend up to the fifth zero or to $f_{\max} = 5/\delta$. In considering the fifth component of the sync generator, the delay network, we must bear in mind that all frequencies up to $5/\delta$ must be delayed by a constant time interval. Two forms of delay network are in common use: the artificial line and the delay multivibrator. These will be considered in order.

It is well known that a lossless transmission line exhibits a delay time independent of frequency. Such a line may be used as a delay network. In practice the delay obtained per unit length of line is too small to be of value, so we contrive to replace the actual line by an artificial line formed of lumped parameters, which gives delays of the order required.

It is beyond the scope of our work to consider in detail the analysis of an artificial transmission line. We shall, however, briefly review some of its salient features. It is well known¹³ that a lossless low-pass constant- k filter exhibits the following properties in its pass band.

$$\begin{aligned}
 \alpha &= \text{attenuation constant} = 0 \\
 \beta &= \text{phase shift per section} \\
 &= 2 \operatorname{arc} \sin \frac{f}{f_c} \\
 f_c &= \text{cutoff frequency} \\
 &= \frac{1}{\pi \sqrt{LC}} \\
 R_0 &= \text{nominal characteristic impedance} \\
 &= \sqrt{\frac{L}{C}}
 \end{aligned}
 \quad \left. \vphantom{\begin{aligned} \alpha \\ \beta \\ f_c \\ R_0 \end{aligned}} \right\} (11-42)$$

where L = series inductance per section
and C = shunt capacitance per section

We have already seen from eq. (7-15) that phase shift and delay time, τ , are related, hence we may write:

¹³ See, for example, E. A. Guillemin, *Communication Networks*. John Wiley and Sons, Inc., 1935, Vol. 2, chap. 9.

$$\tau = \frac{\beta}{\omega} = \frac{2}{\omega} \arcsin \frac{f}{f_c} \tag{11-43}$$

If, now, we design the network so that all signal components of significant amplitude are of frequencies less than or equal to $0.5f_c$, then the ratio f/f_c will be less than 0.5 and as a good approximation we may set¹⁴

$$\arcsin \frac{f}{f_c} \approx \frac{f}{f_c}$$

and we finally have for the delay time per section

$$\tau \approx \frac{2f}{\omega f_c} = \frac{2}{2\pi f} f\pi\sqrt{LC} \approx \sqrt{LC} \tag{11-44}$$

We see, then, that for $f \leq 0.5f_c$, τ is approximately constant and hence a single section of the low-pass filter may be used to introduce a delay of τ seconds in the applied signal. Notice that in the same frequency band $\alpha = 0$ and the amplitudes of the various signal components will remain unchanged. When several such sections are cascaded to provide longer delay intervals, the whole network is referred to as an artificial line.

We may now consider the design of a low-pass section which is to delay a square pulse of width δ . Since all frequencies up to $5/\delta$ must be delayed by a constant amount, we may write

$$f_c = 2f_{\max} = 2\left(\frac{5}{\delta}\right) = \frac{10}{\delta} \tag{11-45}$$

Then, combining (11-42), (11-44), and (11-45),

$$\tau \approx \sqrt{LC} = \frac{1}{\pi f_c} = \frac{\delta}{10\pi} \tag{11-46}$$

and, from (11-42),

$$\sqrt{\frac{L}{C}} = R_0 \tag{11-47}$$

These two equations may then be manipulated to give the circuit constants L and C . Thus, multiplying them, we get

$$\left. \begin{aligned} L &= \frac{\delta R_0}{10\pi} \\ C &= \frac{\delta}{10\pi R_0} \end{aligned} \right\} \tag{11-48}$$

and dividing

¹⁴ By comparing θ and $\sin \theta$ for θ up to 30° or 0.524 radians we note that the maximum error in eq. (11-43) is less than 5 per cent.

Notice that the restriction $f_{\max} \leq 0.5f_c$ means that the section cannot be designed for a delay time greater than $\delta/10\pi$. Generally a single section will not provide sufficient delay, in which event several stages may be cascaded to give the required value. The value of τ for a section may be decreased by raising the factor 10 in the design equations. This may be done without violating the restriction that $f_{\max} \leq 0.5f_c$.

It is sometimes the practice to improve the operation of the circuit by providing the proper impedance match at both ends of the cascaded line. This may be effected by terminating each end of the cascaded artificial line in an m -derived half-section with $m = 0.707$.¹⁵ The complete line with the necessary design equations is illustrated in Fig. 11-24. This artificial delay line is frequently called a "stick."

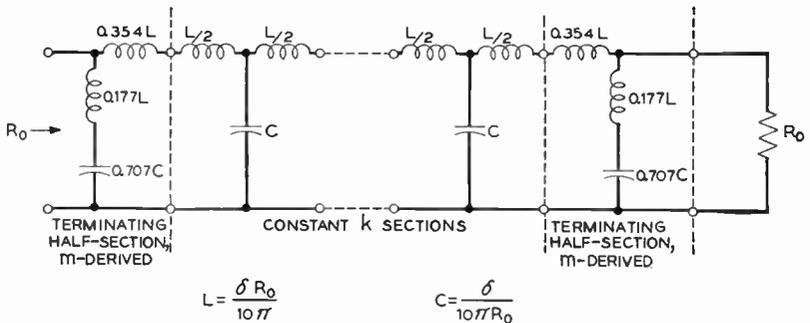


Fig. 11-24. The complete artificial delay line. The delay per section is $\tau = \delta/10\pi$ seconds.

Varying amounts of delay may be had in multiples of τ , the delay per section, by tapping off the line at the appropriate terminal pair. It should be realized that there are other types of delay lines than that just described. Superior performance may be had from lumped-parameter networks of the m -derived type where m is greater than unity, a condition which may be obtained by providing mutual inductance between adjacent coils along the stick. Kallman has recommended the value of $m = 1.27$.¹⁶

¹⁵ The value of $m = 0.707$ is used instead of the more usual 0.6 of conventional filter design because it provides a better impedance match in the band $0 \leq f \leq 0.5f_c$.

¹⁶ H. E. Kallman, "Equalized Delay Lines." *Proc. IRE*, **34**, 9 (September 1946).

Considerable development work has been carried out on delay lines, which employ distributed rather than lumped parameters.^{16,17} For example, the delay per unit length of a coaxial-type cable may be increased by replacing the conventional center conductor by a continuous spiral wound on a plastic core. For the present the lumped-parameter artificial line is more convenient where several taps are required, each corresponding to a different delay.

Generally speaking, for the narrow pulse widths encountered in the supersync signal, the delay per section which may be obtained with an artificial line is small, being in the order of 1 microsecond or less. Where relatively long delay times are required, the physical length of the "stick" may become excessive and some other sort of delay device becomes necessary. We consider now the delay-multivibrator circuit. In this case the original pulse does not appear at the output delayed by some interval, but the original pulse is used to initiate a new pulse which is delayed by the appropriate interval. The idea behind the delay system is illustrated in Fig. 11-25. The original

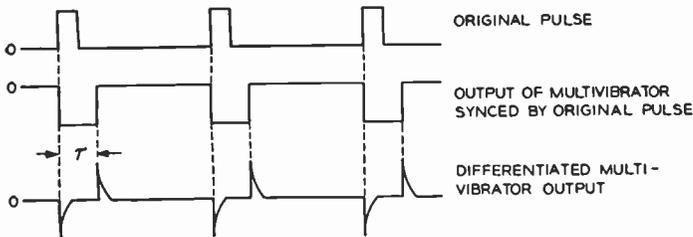


Fig. 11-25. Wave forms of the delay multivibrator. The positive pips in the differentiated wave lag the leading edges of the original wave by τ , the width of the multivibrator output pulse.

pulse is used to synchronize a multivibrator, which is designed to produce a negative output pulse of width τ . On differentiation the multivibrator output produces a positive pip, whose leading edge lags the original pulse by τ . The positive pips may then be used to synchronize another multivibrator. Inspection of the diagram shows that the delay between the original pulse and pip is τ , the width of the delay multivibrator pulse; hence the amount of delay is determined by the design of the delay multivibrator.

¹⁷J. P. Blewett and J. H. Rubel, "Video Delay Lines." *Proc. IRE*, **35**, 12 (December 1947).

11-18. Familiar Components

Some of the details of five basic components of the sync generator have been discussed in the last five sections. The remaining components are more or less familiar circuits and need only be mentioned: the limiting amplifier, the master oscillator, and power supply. The limiting amplifier consists of an amplifier stage, operating with low bias and low plate voltage. Under the influence of the applied signal the tube is driven to saturation and the magnitude of the output voltage is limited by that effect. The limiter is inherently a distortion-producing device and may be used to "square up" a sine wave. It may also be used to limit the amplitude of an applied square wave to a specified value.

The oscillator used to generate the base frequency of 31.5 kilocycles may take the form of any of the standard oscillator circuits and need not be discussed here. It is common practice to use electron coupling in the output to prevent any effects at the input of the counter chain from "pulling" the oscillator proper. The power supply for the entire sync generator is quite conventional. Electronic regulation is used to aid in maintaining a high degree of stability in the entire unit.

11-19. Operation

The principal components of the generator circuit have been described in the sections above. We are now in a position to see how these components are assembled into the complete generator. Consider the block diagram of Fig. 11-26. It is convenient to break down the operation into two units. The first or timing unit, which is enclosed by the dotted line in the diagram, serves to establish the three basic frequencies required in the supersync signal, namely, 31.5 and 15.75 kilocycles, and 60 cycles. It also serves to lock in these components with the power-line frequency. The system is seen to be identical with that shown in Fig. 11-13 with one exception: a limiting amplifier is placed between the master oscillator and the counter chains. This limiter serves to "square up" the sinusoidal output of the oscillator in order to provide a steep wave front pulse for the counter. This results in a more positive trigger action in the counters.

The remainder of the sync generator is the wave-shaping unit,

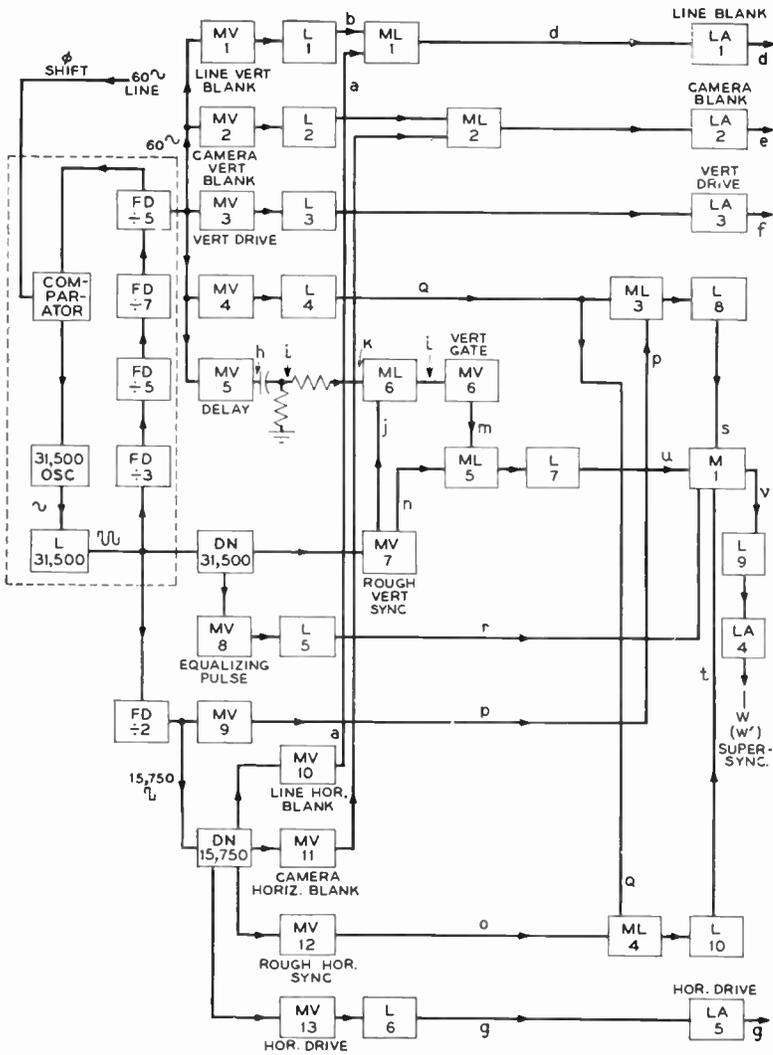


Fig. 11-26. Block diagram of a typical synchronizing-signal generator. (Courtesy of RCA Review.)

which forms the pulse components of the supersync signal and adds them in the proper phase. The principal wave forms at critical points in the unit are shown in Fig. 11-27. Notice that only one-half of the actual number of pulses are shown at each point; for example, in the complete supersync wave at w and w' only three rather than six vertical pulses are shown. This procedure is adopted to simplify the drawing. Consider first the generation of the composite line-

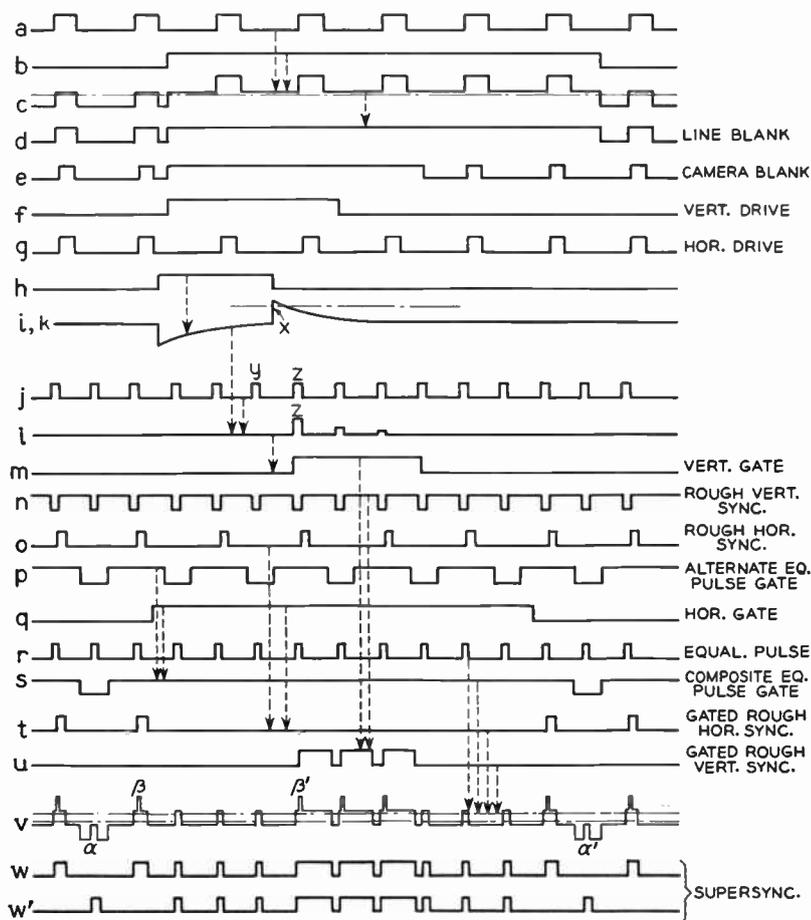


Fig. 11-27. Wave forms in the sync generator of Fig. 11-26. Only one-half the actual number of pulses are shown. (Courtesy of *RCA Review*.)

blanking pulses. A 15,750-cycle square pulse is suitably delayed in DN-15,750 and triggers a multivibrator MV-10, which develops the horizontal blanking pulses *a*. The 60-cycle or vertical blanking pulses are developed by MV-1, limited in L-1, and have the final form *b*. These two waves must be combined so that during the vertical blank intervals, the horizontal pulses are eliminated, a function which is accomplished in the following manner. The two waves, *a* and *b*, are added in ML-1 to produce *c*. Notice that during the vertical blank interval the horizontal pulse rides on top of the vertical pulse and hence may be clipped off by the limiter as indicated by the dotted line in the diagram. The final composite blanking wave is delivered through LA-1 and has the form shown at *d*.

The camera blanking signal is developed in a similar manner. Recall that in each case the camera is blanked for a shorter period than the receiver CRT, and each camera pulse must occur within the duration of the corresponding line pulse. The horizontal camera blanking pulses are therefore derived from a separate multivibrator, MV-11. The resulting pulses lag the line pulses because MV-11 taps in farther along the delay stick than does MV-10. The narrower camera vertical blank pulse is generated in MV-2. The two components are then combined, clipped, and delivered to the output through LA-2 and produce the wave *e*. Notice the relative timing in the components of *d* and *e*.

The vertical drive signals for the camera sweep are derived from MV-3, which is synchronized directly from the 60-cycle output of the main counter chain. The wave form is shown at *f*. The horizontal drive at 15,750 cycles is produced in MV-13. Proper phasing between these pulses and the camera horizontal blanking is obtained by tapping the input of the multivibrator at the appropriate point on DN-15,750. The limited output is delivered through LA-5 and is shown at *g*.

The remaining portions of the circuit serve to develop the supersync proper. The individual rough vertical sync pulses occur at twice the line frequency, or 31.5 kilocycles. They are keyed by pulses from LA-31,500, which are suitably delayed by the stick DN-31,500, and are generated in MV-7. Their shape is shown at *h*. Now these pulses occur in groups of six (shown as three in Fig. 11-27) and only at the end of each field; hence these groups must be gated in at the proper intervals. The vertical gate, *m*, is derived as follows: Refer-

ence to e and m in the figure shows that the gate pulse must lag the leading edge of the line vertical blanking pulse. The necessary delay is produced by the delay multivibrator, MV-5, and by ML-6. The approximate delay is determined by the width of the pulse produced in the delay multivibrator, as shown at h . h is differentiated and the rising edge x in i is used to key in ML-6. The inverse of n , shown at j , is also fed to the mixer. Because of the action of k at the mixer, the pulse labeled z appears across the output of ML-6 and is used to synchronize MV-6, the vertical gate generator. By this relatively complex method, the leading edge of the vertical gating pulse is made to coincide exactly with a trailing edge of the gated pulse, n , thereby ensuring perfect timing. The final group of six vertical pulses is shown at u .

We shall next consider the generation of the horizontal synchronizing components of the supersync wave. The rough horizontal sync pulses occur at line frequency and are developed in MV-12. Synchronism is again provided by tapping the input of the multivibrator to the appropriate point in DN-15,750. The resulting pulses shown at o are fed to ML-4. A study of the composite sync wave shows that these pulses must be keyed out for the duration of the total of 12 equalizing and 6 vertical sync pulses at the end of each field. Thus we require a gating pulse at 60 cycles and of the appropriate width. This horizontal gate, shown at q , is developed in MV-4, limited, and combined with the rough horizontal pulses in ML-4. The resulting wave, with the pulses absent during the equalizing and vertical sync interval, is shown at t .

The next problem is the generation and gating of the equalizing pulses. Suitably delayed synchronizing pulses at twice the line frequency are derived from DN-31,500 and fed to MV-8, which generates the equalizing pulses, r . In MV-9 a 15.75-kilocycle square wave shown at p is developed. When combined with q in ML-3, the composite equalizing pulse gate, s , results. The final addition of all the several supersync components is accomplished in the mixer M-1. Notice from the block diagram that the components to be added are r , s , t , and u . The resulting wave obtained at the output of MV-1 is shown at v . This wave requires some explanation. At α and α' the effect of the composite equalizing pulse gate, s , may be seen. The equalizing pulse, which occurs in the center of each horizontal line during the *active* portion of each field, is notched out. The effect

of adding the equalizing pulses to s , t , and u is indicated by the pips labeled β and β' . A careful inspection of the wave forms of Fig. 11-27 shows that the leading edges of r always precede the leading edge of a rough horizontal or rough vertical sync pulse. In the mixing or adding process the two pulses overlap as shown in Fig. 11-28. By

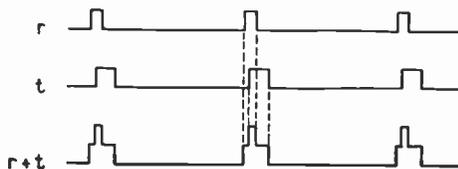


Fig. 11-28. The equalizing and rough horizontal sync pulses are "lap-joined" to ensure wave-front stability.

this process of "lap-joining" the waves, *the leading edge of every pulse in the supersync wave is furnished by a common source, MV-8.* Notice that each component, say, the final horizontal sync, consists of the rough horizontal pulse plus a leading edge furnished by the equalizing pulse generator. The advantage of this system is that the effects of relative drift between the several multivibrators are minimized and precise timing is maintained between all leading edges in the final supersync wave form. The limiter L-9 clips off the undesired portions of v , and the final wave w and w' is delivered at the output.

The Bedford-Smith synchronizing-signal generator just described has proved to be a highly stable source of synchronizing signals. It has found wide acceptance in the industry in applications where its large physical size can be tolerated. In recent years there has been a trend toward the wide use of remote pickup at large distances from the studio. To meet the demands of light equipment for portable use considerable work has been done to reduce the size of the sync generator. A chief contributing factor to the reduction in size has been the development of miniature tubes. Development and research have also led to greater flexibility of operation and performance to closer tolerances.¹⁸

¹⁸ For descriptions of recent design trends see E. Schoenfeld, W. Brown, and W. Milwitt, "New Techniques in Synchronizing Signal Generators." *RCA Review*, VIII, 2 (June 1947); Operating and Maintenance Manual, Du Mont Model TA-124-B Image Orthicon Chain, Allen B. Du Mont Laboratories, Inc., 1948; Instruction Book, Television Field Pick-up Equipment, R.C.A., Engineering Products Department; G. Zaharis, "Television Synchronizing Generator." *Electronics*, 23, 5 (May 1950).

SYNC SIGNAL TESTING¹⁹

The standards on the wave form of the supersync signal shown in Fig. 11-11 prescribe that the rise time and duration of the several pulses be maintained within certain tolerances. We must, therefore, discuss some of the means available for checking the extremely short intervals which are involved in the supersync wave form. A quick inspection of the problem indicates that some technique involving a cathode-ray tube is desirable because of the high writing speeds which may be obtained. Our chief concern is to determine what sort of horizontal sweep will permit the required accuracy in measurement.

11-20. Saw-tooth Sweep

The most obvious type of sweep to use is the conventional saw-tooth form, which is built into most commercial oscilloscopes. The method involved would be to observe the pulse as it is displayed on the face of the cathode-ray tube and to make the required measurements with a scale. Then, if the sweep speed is known, the measured value of distance on the screen may be converted to the corresponding time interval. As a practical matter, three objections may be raised against this method of pulse-width determination.

- (1) Any nonlinearity in the horizontal scan will introduce an error in the measurements.
- (2) The probable error in measuring with a scale may be greater than the tolerance permitted in the pulse width.
- (3) The saw-tooth scan does not permit efficient use of the screen width for measuring purposes.

This last fact comes about because some finite time is consumed during the retrace of the horizontal scan, a loss which prevents an entire cycle of the applied pulse from being viewed on the screen. To circumvent this effect it is necessary to halve the sweep frequency so that two of the pulse cycles appear on the screen with a resulting shrinkage of the time scale. These objections may be overcome by using a sinusoidal, rather than a saw-tooth, horizontal sweep voltage.

¹⁹ R. A. Montfort and F. J. Somers, "Measurement of the Slope and Duration of Television Synchronizing Impulses." *RCA. Review*, VI, 3 (January 1942).

11-21. Sinusoidal Sweep

The basic circuit for measuring pulse width with a sinusoidal horizontal sweep is shown in Fig. 11-29a. The frequency of the sinusoid is chosen to be equal to the repetition rate of the pulses and its effect on the reproduced pattern is to widen the pulse width, provided that the pulse occurs where the sine wave has a maximum rate of change as shown at *b* in the figure. To this end a phase-shift network is

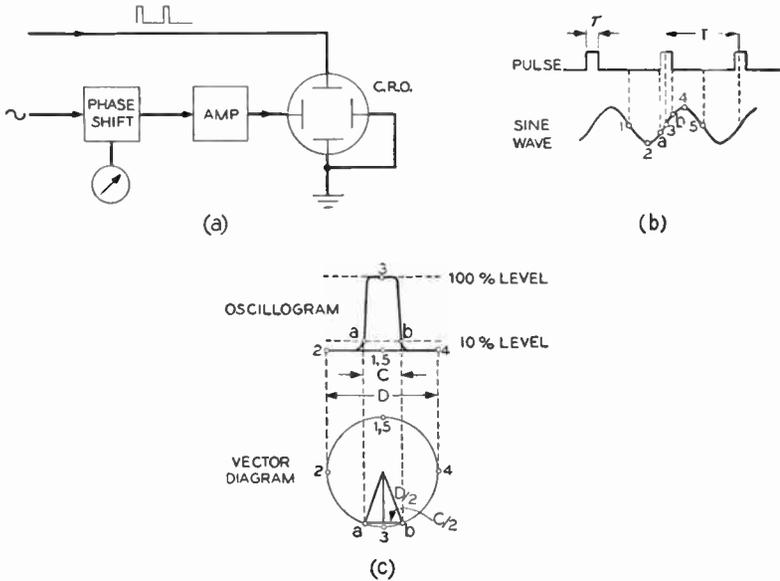


Fig. 11-29. Measurement of pulse width. A sinusoidal sweep voltage is used. (a) Basic circuit. (b) The deflection voltages. (c) Generation of the oscillogram. θ is the angle corresponding to D . (Courtesy of *RCA Review*.)

provided as shown. The appearance of the final image on the screen is shown at *c*, as well as the rotating-vector representation of the sinusoid. We now seek a relationship between the measured distances C and D and the relative pulse width. Since the rotating vector revolves with a constant angular velocity

$$\frac{\tau}{T} = \frac{\theta}{360^\circ} \tag{11-49}$$

and, reading directly from the diagram, we have

$$\theta = 2 \operatorname{arc} \sin \frac{C/2}{D/2} = 2 \operatorname{arc} \sin \frac{C}{D} \quad (11-50)$$

Then the per cent pulse width is

$$\% \text{ pulse width} = \frac{\tau}{T} \times 100 = \frac{\operatorname{arc} \sin \frac{C}{D}}{1.8} \% \quad (11-51)$$

Since the time-distance relationship is known, the per cent pulse width may be determined by direct measurement of C and D on the cathode-ray screen and the use of the last equation.

It should be noticed that in Fig. 11-28 τ is taken to be the time interval between the 10 per cent amplitude points on the pulse. The reason for this procedure may be seen from Fig. 11-11, where the several supersync pulse widths are specified in terms of these points. Since it is impossible to generate a pulse with zero rise time, a pulse cannot have constant width over its entire amplitude range. It has been convenient to standardize duration of the pulse at the 10 per cent amplitude level as the pulse width.

It should be apparent that an even greater expansion of the dimension C of Fig. 11-29 may be obtained on the oscilloscope screen

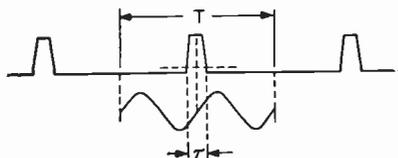


Fig. 11-30. Relationship between the pulse to be measured and the double-frequency sinusoid.

by doubling the frequency of the horizontal-sweep sine voltage, a procedure which is of considerable help in determining the rise time of a pulse which, in some instances, must be in the order of 0.25 microsecond. The time relationship between the pulse and the sinusoid under these conditions is shown in

Fig. 11-30. The corresponding expression for the per cent pulse width is

$$\% \text{ pulse width} = \frac{\tau}{T} \times 100 = \frac{\operatorname{arc} \sin \frac{C}{D}}{3.6} \% \quad (11-52)$$

In the measuring techniques which employ a sine-wave horizontal scan the accuracy of measurement depends upon the sweep being a

pure sine wave; hence extreme care must be exercised in the design of the horizontal-sweep amplifiers to ensure that nonlinear distortion is not introduced into the sweep.

11-22. Phase Shifter

While a number of types of phase shifter may be used with the sine-sweep measuring technique, the simplest is that shown in Fig. 11-31. If no current is drawn by the load across E_o and if the impedance of R and C is sufficiently high so that the transformer secondary acts as a constant-voltage source, then the following analysis is valid.

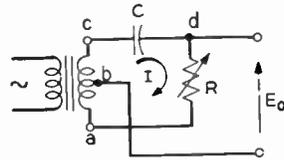


Fig. 11-31. Simple phase-shift network.

Let b be a center tap on the transformer secondary. Then

$$E_{ac} = E_{bc} = E_1 \tag{11-53}$$

By Kirchhoff's voltage law,

$$\begin{aligned} E_o &= IR - E_1 \\ &= \frac{2E_1R}{R - jX_c} - E_1 \end{aligned}$$

or
$$\frac{E_o}{E_1} = \frac{2R}{R - jX_c} - 1 = \frac{R + jX_c}{R - jX_c} \tag{11-54}$$

whence
$$\frac{E_o}{E_1} = 1 \sqrt{2 \operatorname{arc} \tan \frac{X_c}{R}} \tag{11-55}$$

Equation (11-55) shows that, subject to the assumptions, the magnitude of the output voltage is independent of the values of R and C and remains equal to one-half of the *total* transformer secondary voltage. Furthermore, the phase of the output voltage relative to the applied voltage may be controlled by varying R .

11-23. Dot Generator Technique

Still another technique for measuring pulse widths is provided by the dot generator mentioned in section 7-16. The key to pulse measurement is the establishment of a relationship between distance on the cathode-ray screen and time. With a 20-megacycle dot generator feeding the control grid of the test oscilloscope, the time base appears

directly on the face of the screen in the form of bright dots spaced at 0.05-microsecond intervals. The advantage of the method is that the precision of measurement is independent of the form of horizontal sweep and that errors resulting from parallax are eliminated. Use of the sine-wave type of sweep is desirable, however, because of its spreading action on the pulse, as viewed on the oscilloscope. The disadvantage of the method is that interpolation between the marker dots must be estimated.

11-24. Pulse Cross Testing

It is often desirable to have some means available in the control room of the television studio to check the supersync signal as a whole, rather than pulse by pulse. Such a means is afforded by the pulse cross test pattern, which provides an over-all view of the supersync wave in the region of the vertical blanking interval at the end of each field. Consider first the test circuit shown in Fig. 11-32. A black-

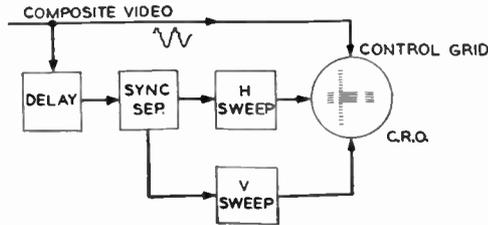


Fig. 11-32. Test equipment for producing the pulse cross test pattern.

positive composite video signal is fed to the test unit, where the sync components are stripped off and fed to the deflection circuits. Notice that a delay network is incorporated in the sweep system. It serves to delay the vertical sweep for one-half a field interval so that the vertical retrace portion of the composite video signal will appear in the center of the monitor screen. In a similar manner the horizontal sweep is delayed so that the horizontal sync and blanking pulses are centered on the screen.

Since the composite video signal applied to the control grid has a black-positive polarity, the sync portions of the signal tend to drive the reproduced image toward white. If, now, the d-c voltage on the cathode-ray tube grid is adjusted so that the pedestal level is just

visible on the screen, the picture components will be eliminated in the final image, which will appear as shown in Fig. 11-33. Inspection of the diagram shows the origin of the term pulse cross test pattern.

In the diagram each component pulse is labeled so that it may be identified with its counterpart in Fig. 11-11. Improper timing or duration of any of the pulse components may be recognized immediately by a corresponding displacement in the cross pattern.

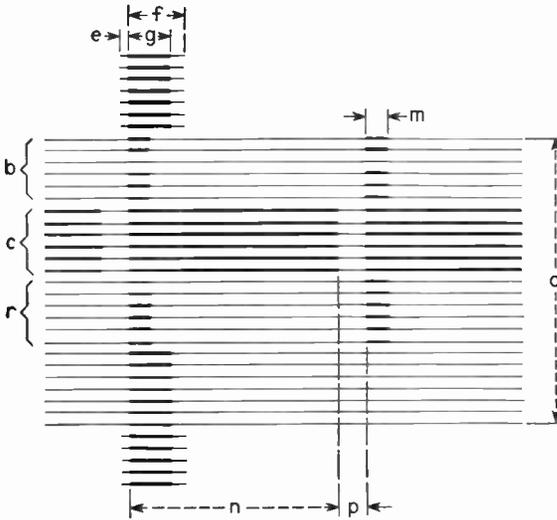


Fig. 11-33. The pulse cross test pattern. Dimensions shown are those of Fig. 11-11. The vertical scale is expanded for clarity.

CHAPTER 12

VESTIGIAL-SIDEBAND TRANSMISSION

Before we can discuss the video transmitter intelligently we must first investigate the nature of the signal which the transmitter is required to deliver to the antenna. We have already seen that a relatively broad band is required by the composite video signal. When this amplitude-modulates a radio-frequency carrier, an even broader bandwidth is produced. Our chief interest in the first part of the present chapter is to investigate what means may be employed to fit this modulated video signal and its associated frequency-modulated sound program into the 6-megacycle channel allowed by the Federal Communications Commission. In the analytical work which follows we shall assume the complex video signal to be replaced by a sine wave in order that a steady-state analysis may be used; this results in considerable simplification of the problem.¹

12-1. Amplitude Modulation

In conventional amplitude modulation the radio-frequency carrier wave

$$e_c = A \cos \omega t \quad (12-1)$$

is applied to the modulated amplifier stage. The modulating signal

$$e_m = E_m \cos \omega_m t \quad (12-2)$$

is also applied to that same stage but in such a manner that the amplitude of the output wave varies sinusoidally about the mean value A and at the modulating frequency ω_m . Under these conditions the R-F carrier is said to be amplitude-modulated, and the resulting wave is expressed by

$$e = (A + E_m \cos \omega_m t) \cos \omega t \quad (12-3)$$

¹ For a discussion of the transient aspect of the problem see R. D. Kell and G. L. Fredendall, "Selective Sideband Transmission in Television," *RCA Review*, IV, 4 (April 1940).

If e is plotted against time, the curve will be a cosine wave, whose amplitude varies cosinusoidally about the value A .

Whereas (12-3) is a convenient expression for the modulated wave as far as the transmitter is concerned, it may also be expanded to another form, which has greater utility in the work which follows. Thus, we may factor A out of the parentheses and get

$$\begin{aligned} e &= A \left(1 + \frac{E_m}{A} \cos \omega_m t \right) \cos \omega t \\ &= A(1 + m \cos \omega_m t) \cos \omega t \end{aligned} \quad (12-4)$$

where $m = \text{modulation index} = \frac{E_m}{A}$. (12-5)

If, now, the double cosine product of the last equation is expanded, there results

$$e = A \left[\underbrace{\cos \omega t}_{\text{carrier}} + \frac{m}{2} \underbrace{\cos (\omega + \omega_m)t}_{\text{upper side frequency}} + \frac{m}{2} \underbrace{\cos (\omega - \omega_m)t}_{\text{lower side frequency}} \right] \quad (12-6)$$

Equation (12-6) shows that the amplitude-modulated wave consists of three frequency components: the carrier and the upper and lower side frequencies. The last two components are identified by their frequencies, which are $(\omega + \omega_m)$ and $(\omega - \omega_m)$, respectively. Expressed in this form the wave allows us to investigate its bandwidth, which may be seen to be

$$\begin{aligned} &\text{bandwidth of amplitude-modulated signal} \\ &= (\omega + \omega_m) - (\omega - \omega_m) = 2\omega_m \end{aligned} \quad (12-7)$$

that is, the amplitude-modulated wave requires a bandwidth equal to twice the bandwidth of the modulating signal.

Let us see how such a double sideband wave could be fitted along with its audio program into a 6-megacycle channel. For the moment we shall adopt the following F.C.C. standards: 50 kilocycles allowed for the frequency-modulated sound signal and the audio carrier shall lie 0.25 megacycle below the upper limit of the channel. These standards place the lower limit of the aural signal 0.275 megacycle below the upper channel limit, thus the bandwidth left for the visual signal is $6 - 0.275 = 5.725$ megacycles. Since we are transmitting two sidebands, this bandwidth of approximately 5.7 megacycles must

be divided equally between the two sidebands, and the carrier will lie $\frac{1}{2}(5.7) = 2.85$ megacycles above the lower channel limit as shown in Fig. 12-1. Clearly, then, the maximum video modulating frequency which may be used is only 2.85 megacycles if the signal is to remain within its specified 6-megacycle channel. As we shall see, this represents a very poor utilization of the allotted bandwidth. We desire some form of transmission which will permit a 4.5-megacycle maxi-

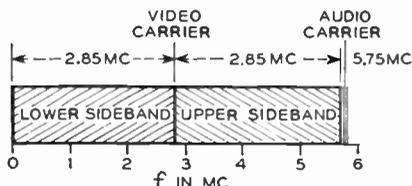


Fig. 12-1. Double-sideband transmission within a 6-megacycle channel permits a maximum video modulating frequency of only 2.85 megacycles.

imum modulating frequency to be used. For the moment, however, we shall continue our investigation of the double-sideband system, which will serve as a basis of comparison for the more efficient—as far as bandwidth utilization is concerned—systems to be described.

We next consider what happens to the signal (12-6) as it passes through the receiving equipment up to the final detector. In general, these predetector stages do not have ideal response characteristics and will modify the amplitude and phase of each component of the signal. This distortion, resulting from the predetector stages, may be represented in the following manner. Let

$B(f_i)$ = over-all amplitude response at frequency f_i of all the predetector stages, and

$\phi(f_i)$ = over-all phase shift at frequency f_i caused by all the predetector stages,

then the modulated wave delivered to the final detector will be e_1 .

$$e_1 = A \left\{ B(f) \cos [\omega t + \phi(f)] + \frac{m}{2} B(f + f_m) \cos [(\omega + \omega_m)t + \phi(f + f_m)] + \frac{m}{2} B(f - f_m) \cos [(\omega - \omega_m)t + \phi(f - f_m)] \right\} \quad (12-8)$$

We shall assume that the detector is of the diode type and ideal. Then, if m is not too small, the output of the detector will be simply the envelope of the wave applied to the detector. Our problem then is to manipulate (12-8) into the general form

$$e_1 = V \cos (\omega t + \theta) \tag{12-9}$$

where V is a function of time and, by comparison to (12-3), is seen to be the envelope which we seek. To simplify our work we let

$$W = \cos [(\omega \pm \omega_m)t + \phi(f \pm f_m)] \tag{12-10}$$

Then, adding and subtracting $\phi(f)$,

$$\begin{aligned} W &= \cos [(\omega \pm \omega_m)t + \phi(f \pm f_m) + \phi(f) - \phi(f)] \\ &= \cos \{[\omega t + \phi(f)] + [\pm \omega_m t + \phi(f \pm f_m) - \phi(f)]\} \\ &= \cos \left\{ [\omega t + \phi(f)] + \begin{array}{l} +\{\omega_m t + [\phi(f + f_m) - \phi(f)]\} \\ -\{\omega_m t + [\phi(f) - \phi(f - f_m)]\} \end{array} \right\} \end{aligned} \tag{12-11}$$

Equation (12-11) may be simplified by virtue of a characteristic common to almost all properly tuned radio-frequency amplifiers:

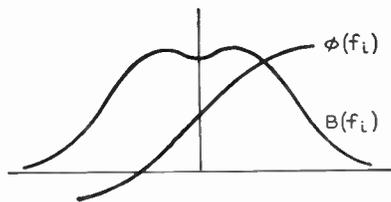


Fig. 12-2. Typical response characteristics of tuned amplifiers. The amplitude response exhibits even symmetry and the phase response exhibits odd or skew symmetry.

their amplitude and phase response characteristics exhibit symmetry about the carrier frequency, ω , as shown in Fig. 12-2. Taking advantage of this fact we may write

$$B = B(f + f_m) = B(f - f_m) \tag{12-12}$$

and $\phi = [\phi(f + f_m) - \phi(f)] = [\phi(f) - \phi(f - f_m)]$

and, substituting into (12-11), we get

$$\begin{aligned} W &= \cos \{[\omega t + \phi(f)] \pm [\omega_m t + \phi]\} \\ &= \cos [\omega t + \phi(f)] \cos [\omega_m t + \phi] \mp \sin [\omega t + \phi(f)] \sin (\omega_m t + \phi) \end{aligned} \tag{12-13}$$

Then, substituting (12-13) into (12-8), we get

$$\begin{aligned}
 e_1 &= A \left\{ B(f) \cos [\omega t + \phi(f)] \right. \\
 &\quad \left. + \frac{mB}{2} \cos [\omega t + \phi(f)] \cos (\omega_m t + \phi) \right. \\
 &\quad \left. + \frac{mB}{2} \cos [\omega t + \phi(f)] \cos [\omega_m t + \phi] \right\} \\
 &= A \{ B(f) + mB \cos (\omega_m t + \phi) \} \cos [\omega t + \phi(f)] \quad (12-14)
 \end{aligned}$$

and the envelope is

$$V = A \{ B(f) + mB \cos (\omega_m t + \phi) \} \quad (12-15)$$

This envelope which appears across the detector output consists of a d-c component $AB(f)$ and an a-c component

$$AmB \cos (\omega_m t + \phi) = BE_m \cos (\omega_m t + \phi) \quad (12-16)$$

The d-c component may be removed by a series blocking condenser and the final output consists simply of the a-c component given by (12-16). On comparing this output with the original modulating signal, we note that they are the same except that the amplitude has been changed by a factor B and the phase shifted by ϕ . When we think in terms of a single modulating component as presumed in the analysis, this seems to be of no importance. But remember, with the actual television signal we have a whole band of frequencies rather than a single frequency in the modulating signal; hence we must interpret (12-16) as a curve of amplitude and phase plotted against modulating frequency. These curves, of course, will have the same shape as those of Fig. 12-2, and are dependent upon those stages of the over-all system which lie between the modulated amplifier at the transmitter and the detector at the receiver. It follows from our previous work that if B is independent of frequency over the entire spectrum of the signal and if ϕ is linear with frequency over the same range, the wave form of the detected wave will be identical with that of the original modulating signal.

We may sum up these results. In amplitude modulation the modulated carrier requires a bandwidth equal to twice the highest modulating-frequency component. After passing through ideal am-

plifiers, the radio-frequency wave may be detected or demodulated by a linear detector. In the ideal case the shape of the detected output is identical to that of the original modulating signal. Theoretically, the double-sideband system of transmission is distortionless; its disadvantage is its failure to provide maximum utilization of the radio-frequency spectrum. We shall use these characteristics as a yardstick for measuring the performance of other systems of transmission.

12-2. Single-sideband Transmission

Inspection of eq. (12-6), which is the expression for the amplitude-modulated wave, shows that the quantities $mA = E_m$ and ω_m are contained in both of the sideband terms. Since E_m and ω_m uniquely specify the cosinusoidal modulating signal (12-2), it is immediately apparent that the modulation information is present in both of the sidebands. Mathematically, at least, it would appear that either sideband could be eliminated, with a 2 to 1 saving in bandwidth, and still sufficient modulation information would be present for satisfactory demodulation to take place at the second detector in the receiver. Let us investigate this possibility to see if an undistorted output signal can be obtained. As a matter of convenience we shall assume that the double-sideband output of the modulated amplifier at the transmitter is passed through some sort of a filter network so that the lower sideband is completely suppressed. The remaining signal, consisting of the carrier and upper sideband, is then delivered to the receiver and passed on finally to the second detector. Notice that we can accomplish these effects mathematically by letting $B(f - f_m)$ go to zero, which provides complete suppression of the lower sideband. If, then, we let e_2 be the single-sideband wave delivered to the detector, we have from (12-8) that

$$e_2 = A \left\{ B(f) \cos [\omega t + \phi(f)] + \frac{m}{2} B(f + f_m) \cos [(\omega + \omega_m)t + \phi(f + f_m)] \right\} \quad (12-17)$$

Then, utilizing our previous expansion for the second cosine term, we have

$$e_2 = A \left\{ B(f) \cos [\omega t + \phi(f)] + \frac{m}{2} B(f + f_m) \{ \cos [\omega t + \phi(f)] \cos [\omega_m t + \phi(f + f_m)] - \sin [\omega t + \phi(f)] \sin [\omega_m t + \phi(f + f_m)] \} \right\} \quad (12-18)$$

Notice that the argument $[\omega t + \phi(f)]$ occurs in each term of the expression, and the terms may be regrouped as follows:

$$e_2 = A \left\{ \left\{ B(f) + \frac{m}{2} B(f + f_m) \cos [\omega_m t + \phi(f + f_m)] \right\} \cos [\omega t + \phi(f)] - \left\{ \frac{m}{2} B(f + f_m) \sin [\omega_m t + \phi(f + f_m)] \right\} \sin [\omega t + \phi(f)] \right\} \quad (12-19)$$

Equation (12-19) is of the general form

$$e_2 = C \cos x + S \sin x \quad (12-20)$$

and since $\cos x$ and $\sin x$ are separated by 90° , the two terms may be combined and e_2 reduced to the form

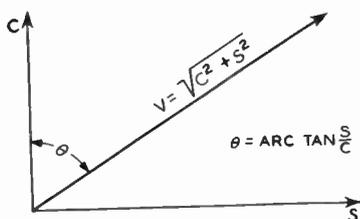
$$e_2 = V \cos (x - \theta) \quad (12-21)$$

where

$$V = \text{envelope} = \sqrt{C^2 + S^2}$$

and

$$\theta = + \text{arc tan } \frac{S}{C} \quad (12-22)$$



$$C \cos x + S \sin x = V \cos (x - \theta)$$

Fig. 12-3. The cosine and sine terms, C and S , may be combined into a single cosine term, which lags C by an angle θ .

The basis for these relationships is shown in Fig. 12-3. The envelope of (12-19), which is the demodulated output of the detector, will be

$$\begin{aligned}
 V &= A \sqrt{\left\{ B(f) + \frac{m}{2} B(f + f_m) \cos [\omega_m t + \phi(f + f_m)] \right\}^2} \\
 &\quad + \left\{ \frac{m}{2} B(f + f_m) \sin [\omega_m t + \phi(f + f_m)] \right\}^2} \\
 &= A \sqrt{B^2(f) + \left\{ \frac{m^2}{4} B^2(f + f_m) \right.} \\
 &\quad \left. + mB(f)B(f + f_m) \cos [\omega_m t + \phi(f + f_m)] \right\}} \quad (12-23)
 \end{aligned}$$

In general, m , the modulation index, will be less than unity and (12-23) may be simplified by means of the binomial expansion; thus,

$$(\alpha + \beta)^{1/2} = \alpha^{1/2} + \frac{1}{2}\alpha^{-1/2}\beta + \dots \quad (12-24)$$

or, applying this expansion to the radical in (12-23), we have

$$\begin{aligned}
 V \approx A \left\{ B(f) + \frac{m^2}{8} \frac{B^2(f + f_m)}{B(f)} \right. \\
 \left. + \frac{m}{2} B(f + f_m) \cos [\omega_m t + \phi(f + f_m)] \right\} \quad (12-25)
 \end{aligned}$$

Once again the d-c terms may be removed with a blocking condenser and the final demodulated a-c output is

$$\begin{aligned}
 \frac{AmB(f + f_m)}{2} \cos [\omega_m t + \phi(f + f_m)] \\
 = B(f + f_m) \frac{E_m}{2} \cos [\omega_m t + \phi(f + f_m)] \quad (12-26)
 \end{aligned}$$

The approximation in (12-25) is the result of the fact that the higher order terms in m are neglected because of their small amplitudes. Subject to this assumption, the output of the single-sideband system is seen to be identical to that of the double-sideband system except that the amplitude of the output in the former case is smaller by a factor of two. This result is quite reasonable, for (12-18) shows that the modulation power is divided equally between the two sidebands. If one of them is suppressed, only one-half of the output signal will result.

We may summarize the results for the single-sideband system of

transmission. If m is small so that the higher order terms in m of the expansion of (12-23) are negligible, the transmission of a single sideband of an amplitude-modulated signal and its subsequent detection will ideally yield a distortionless reproduction of the modulating signal wave form. The magnitude of the detected voltage will have only one-half the value obtained in the double-sideband system; the saving in bandwidth is 2 to 1.

It is interesting to observe how this saving in bandwidth affects the utilization of the 6-megacycle channel width specified by the F.C.C. Using our previous figures, we see that the entire spread of

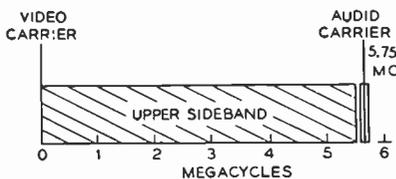


Fig. 12-4. Suppression of the lower sideband at the transmitter doubles the maximum video modulating frequency permissible in a 6-megacycle channel.

5.7 megacycles may be used for only one, rather than two, sidebands, and the maximum video-modulating component is raised from 2.85 to 5.7 megacycles. Channel utilization for the single-sideband system is illustrated in Fig. 12-4.

It is an unfortunate fact that the advantages of single-sideband transmission cannot be utilized in practice—at least at the present state of the television art—because it is impossible to design a filter structure which can provide a perfectly sharp cutoff and still maintain a constant delay characteristic throughout the entire pass band. It can be shown that sharp cutoff and constant delay are incompatible. Nevertheless, the advantage of the single-sideband system is sufficiently attractive to force a compromise of some sort. Such a compromise is the vestigial-sideband system, which is used in commercial practice.

12-3. Vestigial-sideband Transmission

In any system of transmission we try to adjust conditions so that the detected output has the same variation with time as the modulating signal. In building up the concept of the vestigial-sideband system we shall consider the effect of various types of sideband characteristics on the output signal from the detector. When we have found a combination which gives an undistorted output, we shall work backward toward the transmitter to determine the type of sideband suppression that gives these conditions. The final channel-utilization diagram may then be drawn.

We have stated that sharp-cutoff, linear-phase filters cannot be designed and that single-sideband transmission must be ruled out. As a starting point let us compromise: instead of cutting off the signal abruptly just below the carrier, let us design a filter that cuts off gradually, maintaining a reasonably constant delay characteristic. We shall increase the bandwidth of the transmitted signal slightly in order to permit the utilization of a filter, which is feasible. Let the characteristics of all the networks between the modulated amplifier and the detector be those shown in Figure 12-5a. Notice

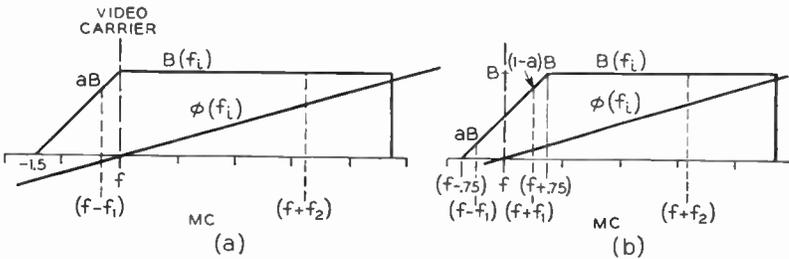


Fig. 12-5. Response curves for two types of vestigial-sideband transmission. (a) $B(f_i)$ and $\phi(f_i)$ for Case I. (b) $B(f_i)$ and $\phi(f_i)$ for Case II.

that the entire upper sideband plus a small part or vestige of the lower sideband is applied to the detector; hence the term vestigial-sideband system. How will this vestigial-sideband characteristic affect the final demodulated signal? Clearly, frequencies which lie below the visual carrier are affected differently from those above it; hence we shall carry through the analysis for two modulating components, one, f_1 , less than 1.5 megacycles, and the second, f_2 , greater than 1.5 megacycles. Reading directly from the diagram we have the following relationships:

$$\left. \begin{aligned}
 \text{For } f & \quad B(f) = B \\
 \text{For } f_1 & \quad B(f + f_1) = B \quad \phi(f + f_1) = \phi_1 \\
 & \quad B(f - f_1) = aB \quad \phi(f - f_1) = -\phi_1
 \end{aligned} \right\} \quad (12-27)$$

Then, substituting these values into (12-8), we have for e_2 , the voltage delivered to the detector when the modulating frequency is f_1 ,

$$\begin{aligned}
 e_2 &= A \left\{ B \cos \omega t + \frac{m}{2} B \cos [(\omega + \omega_1)t + \phi_1] \right. \\
 &\quad \left. + \frac{m}{2} aB \cos [(\omega - \omega_1)t - \phi_1] \right\} \\
 &= A \left\{ B \cos \omega t + \frac{m}{2} B \cos [\omega t + (\omega_1 t + \phi_1)] \right. \\
 &\quad \left. + \frac{m}{2} aB \cos [\omega t - (\omega_1 t + \phi_1)] \right\} \\
 &= A \left\{ B \cos \omega t + \frac{m}{2} B [\cos \omega t \cos (\omega_1 t + \phi_1) \right. \\
 &\quad \left. - \sin \omega t \sin (\omega_1 t + \phi_1)] \right. \\
 &\quad \left. + \frac{m}{2} aB [\cos \omega t \cos (\omega_1 t + \phi_1) \right. \\
 &\quad \left. + \sin \omega t \sin (\omega_1 t + \phi_1)] \right\} \quad (12-28)
 \end{aligned}$$

Since we once again will try to evaluate the envelope of (12-28), it is convenient to collect terms about $\cos \omega t$ and $\sin \omega t$; hence,

$$\begin{aligned}
 e_2 &= A \left\{ \left[B + \frac{m}{2} B \cos (\omega_1 t + \phi_1) + \frac{m}{2} aB \cos (\omega_1 t + \phi_1) \right] \cos \omega t \right. \\
 &\quad \left. + \left[-\frac{m}{2} B \sin (\omega_1 t + \phi_1) + \frac{m}{2} aB \sin (\omega_1 t + \phi_1) \right] \sin \omega t \right\} \\
 &= A \left\{ \left[B + \frac{m}{2} (1 + a)B \cos (\omega_1 t + \phi_1) \right] \cos \omega t \right. \\
 &\quad \left. - \left[\frac{m}{2} (1 - a)B \sin (\omega_1 t + \phi_1) \right] \sin \omega t \right\} \quad (12-29)
 \end{aligned}$$

By eq. (12-22) the envelope will be

$$\begin{aligned}
 V &= A \sqrt{B^2 + m(1 + a)B^2 \cos (\omega_1 t + \phi_1)} \\
 &\quad + \frac{m^2}{4} B^2 [(1 + a)^2 \cos^2 (\omega_1 t + \phi_1) + (1 - a)^2 \sin^2 (\omega_1 t + \phi_1)] \\
 &\hspace{15em} (12-30)
 \end{aligned}$$

Since the system is predominantly of a single-sideband type, m will

generally be small and the term involving $m^2/4$ in (12-30) may be assumed to be negligibly small. Then, applying the binomial expansion to the first two terms under the radical, we have finally for the envelope of the wave

$$V \approx A \left[B + \frac{m}{2} (1 + a) B \cos (\omega_1 t + \phi_1) \right] \quad (12-31)$$

and the final a-c component delivered by the detector is

$$\frac{Am}{2} (1 + a) B \cos (\omega_1 t + \phi_1) = \frac{E_m}{2} (1 + a) B \cos (\omega_1 t + \phi_1) \quad (12-32)$$

The last equation completes the analysis for the modulating signal of frequency $f_1 < 1.5$ megacycles. For the second modulating signal for which $f_2 > 1.5$ megacycles we may write the result directly, for note from Fig. 12-5a that at f_2 the system is precisely of the single-sideband type: the lower sideband is completely suppressed. From the diagram

$$\left. \begin{aligned} B(f + f_2) &= B & \phi(f + f_2) &= \phi_2 \\ B(f - f_2) &= 0 \end{aligned} \right\} \quad (12-33)$$

Substitution of these values into (12-26) yields the following result for the final demodulated a-c output:

$$\frac{E_m}{2} B \cos (\omega_2 t + \phi_2) \quad (12-34)$$

The results just derived indicate that the vestigial-sideband system is ideally capable of distortionless operation (subject to $m^2/4$ being negligible), except that the system is not "flat"; for every modulating frequency below 1.5 megacycles the output will be too large by a factor $(1 + a)$. This result is to be expected since for the f_1 type components two sidebands contribute to the output, whereas for the f_2 type the output is the result of the upper sideband alone. The net effect is that a form of pre-emphasis exists as shown in Fig. 12-6.

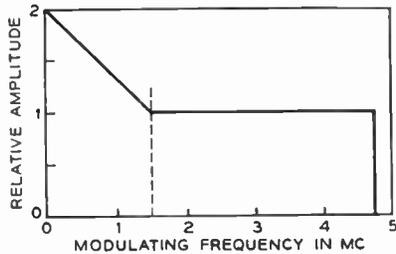


Fig. 12-6. The Case I system is not "flat"; the amplitudes of components below 1.5 megacycles are pre-emphasized.

Notice that if we can make the amplitude characteristic flat, the vestigial sideband will be quite satisfactory, provided that m remains small. Our immediate problem, then, is to so locate the cutoff portion of the $B(f_i)$ curve in Fig. 12-5a relative to the carrier that the relative amplitude of the demodulated signal will remain constant for all values of modulating frequency.

Consideration of the problem will show that the desired result may be obtained by centering the cutoff portion of the $B(f_i)$ curve on the carrier frequency as shown in Fig. 12-5b. Then, if the cutoff characteristic exhibits odd symmetry about the point $(B/2, f)$ for any modulating frequency $f_1 < 0.75$ megacycle, we have

$$\left. \begin{array}{l} \text{for } f_1 \quad \left. \begin{array}{l} B(f + f_1) = (1 - a)B \quad \phi(f + f_1) = \phi_1 \\ B(f - f_1) = aB \quad \phi(f - f_1) = -\phi_1 \end{array} \right\} \\ \text{and for } f \quad \left. \begin{array}{l} B(f) = \frac{B}{2} \\ \phi(f) = 0 \end{array} \right\} \end{array} \right\} \quad (12-35)$$

and the modulated signal delivered to the second detector will be from (12-8)

$$e_2 = A \left\{ \frac{B}{2} \cos \omega t + \frac{m}{2} (1 - a)B \cos [(\omega + \omega_1)t + \phi_1] + \frac{m}{2} aB \cos [(\omega - \omega_1)t - \phi_1] \right\} \quad (12-36)$$

This expression may be reduced by application of our previous methods to

$$e_2 = A \left\{ \left[\frac{B}{2} + \frac{m}{2} (1 - a)B \cos (\omega_1 t + \phi_1) + \frac{m}{2} aB \cos (\omega_1 t + \phi_1) \right] \cos \omega t + \left[-\frac{m}{2} (1 - a)B \sin (\omega_1 t + \phi_1) + \frac{m}{2} aB \sin (\omega_1 t + \phi_1) \right] \sin \omega t \right\} \quad (12-37)$$

Then, combining terms, we have

$$e_2 = A \left\{ \left[\frac{B}{2} + \frac{m}{2} B \cos (\omega_1 t + \phi_1) \right] \cos \omega t - \frac{m}{2} [(1 - 2a)B \sin (\omega_1 t + \phi_1)] \sin \omega t \right\} \quad (12-38)$$

By our previous methods the envelope V will be

$$V \approx A \left[\frac{B}{2} + \frac{1}{2} \frac{mB^2}{B} \cos (\omega_1 t + \phi_1) \right] \tag{12-39}$$

and the final a-c component is

$$\frac{AmB}{2} \cos (\omega_1 t + \phi_1) = \frac{E_m}{2} B \cos (\omega_1 t + \phi_1) \tag{12-40}$$

In order to see whether the system is flat, we must calculate the output corresponding to a modulating frequency $f_2 > 0.75$ megacycles. For such a frequency, the system is single-sideband and

$$\left. \begin{aligned} \text{for } f_2 \quad B(f + f_2) = B \quad \phi(f + f_2) = \phi_2 \\ B(f - f_2) = 0 \end{aligned} \right\} \tag{12-41}$$

The final a-c output voltage is

$$\frac{E_m}{2} B \cos (\omega_2 t + \phi_2) \tag{12-42}$$

Since the coefficients in (12-40) and (12-42) are equal, the system just described is flat, and theoretically may be made distortionless, provided that m is sufficiently small.

It is important to state once again the properties of the $B(f_i)$ curve which produces a flat system; the curve exhibits odd symmetry about its intersection with the carrier frequency axis for a frequency range extending from 0.75 mc below to 0.75 mc above the carrier. Under this condition

$$B(f + f_1) + B(f - f_1) = B \tag{12-43}$$

Thus the sum of the amplitudes of the two sidebands for any $f_1 < 0.75$ megacycle is always equal to the amplitude of the upper sideband for any modulating signal of frequency $f_2 > 0.75$ megacycle. This fact is demonstrated mathematically in the transition from eq. (12-37) to (12-38), where the coefficients $(1-a)$ and a add to give unity.

Remember, we have only considered the problem at the second detector. We know the type of vestigial-sideband signal, which, when applied to the second detector, gives the correct output-voltage shape. We must now determine what type of signal must leave the transmitter to give this condition in the receiver.

RECEIVER AND TRANSMITTER ATTENUATION

In general, there are two practicable methods of delivering the proper form of vestigial-sideband signal to the second detector in the receiver. In the first, or Transmitter Attenuation (T-A), system the required attenuation of the lower sideband, as specified by (12-43), is performed at the transmitter end of the television network. In the second, or Receiver Attenuation (R-A), system the necessary attenuation takes place at the receiver. Both systems will give the identical signal at the second detector. We next investigate each of them to determine which affords the optimum compromise for commercial telecasting.

12-4. Transmitter Attenuation

Of the two proposals made in the last paragraph the first or T-A system is the more obvious. Figure 12-5b shows the characteristics required of the signal; hence if a filter system having these characteristics is placed at the transmitter following the modulated amplifier, the lower sideband will be attenuated and the radiated signal will meet the specifications. Such a system requires of the receiver only that its several predetector stages be reasonably free of distortion. The characteristics required of the transmitter filter and of the receiver predetector stages are illustrated in Fig. 12-7.

To aid the discussion which follows it is convenient to define f_c as the upper cutoff frequency, whose value we have not as yet determined. Notice then that if we include the d-c component in our

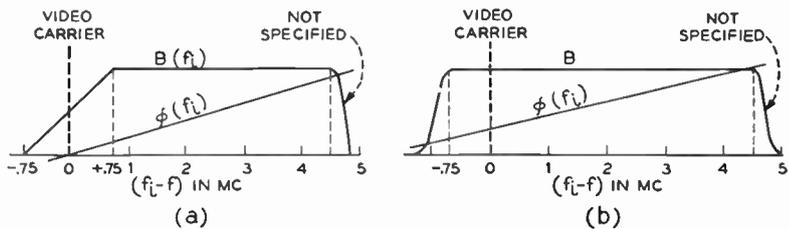


Fig. 12-7. Characteristics required for the T-A system of vestigial-sideband transmission. The position of the upper cutoff is not specified. (a) Characteristics of the transmitter filter. (b) Idealized characteristics of the receiver predetector stages for no distortion.

discussion, the total bandwidth occupied by the modulating signal is identically equal to f_c . This point is also of benefit in the work which follows. Let us list some of the principal features of the T-A system.

At the transmitter:

- (1) A filter of critical design which can handle transmitter power levels is required.
- (2) The carrier level is reduced to one-half the maximum signal level.

At the receiver:

- (1) The bandwidth required is greater than f_c .
- (2) No special cutoff characteristic is required.

12-5. Receiver Attenuation

The less obvious method of producing the necessary vestigial-sideband signal at the second detector requires the transmitter and receiver characteristics shown in Fig. 12-8. Notice that the filtering

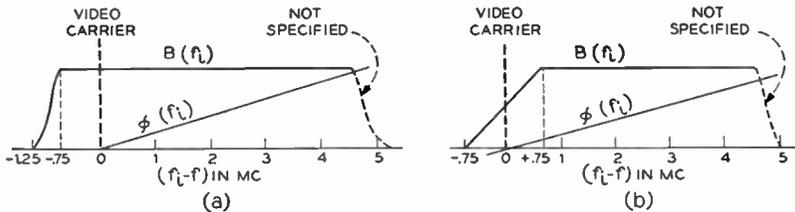


Fig. 12-8. Characteristics required for the R-A system of vestigial-sideband transmission. (a) Transmitter characteristics, (b) Receiver characteristics.

required at the transmitter is less critical than it was in the T-A case but at the same time the receiver design becomes more complicated. The chief features of the R-A system follow.

At the transmitter:

- (1) A filter of less critical design is required.
- (2) The carrier level is at the maximum signal level.

At the receiver:

- (1) The required bandwidth is greater than f_c but less than that of the T-A system.
- (2) A special cutoff characteristic is required.

12-6. Comparison

The two schemes of vestigial-sideband transmission may best be compared by superimposing their transmitter and receiver characteristics. In combining Figs. 12-7 and 12-8 we make the carrier levels the same in both cases. If we assume that the final transmitter stage in each system uses the same tube, the upper sideband amplitudes for the T-A plan have twice the amplitude or four times the power of the

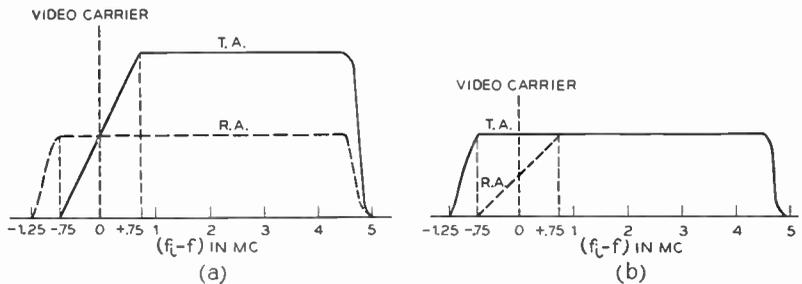


Fig. 12-9. The T-A and R-A characteristics are superimposed for comparison. (a) Transmitter. (b) Receiver.

R-A system. Since the output voltage at the detector is proportional to the sideband voltage, this means that, in theory at least, the T-A system provides a 4 to 1 apparent power gain over the R-A system. Notice that the apparent gain is not obtained by increasing carrier power, which is assumed to be the same in both cases; hence the apparent gain is obtained without added interference to other types of radio service in the same area. From the power point of view, then, the T-A system has the advantage.

In regard to transmitter filter design, the R-A system is simpler; in fact, experience seems to indicate that the T-A characteristic can be obtained only at relatively low power levels. This, in effect, means that T.A. requires low-level modulation and filtering, followed by broad-band linear power amplifiers, which serve to raise the level of the filtered signal. In R.A. high-level filters may be designed to meet the less rigid specifications; thus R.A. provides more flexibility in the design of the transmitter.

Consider next the requirements on the receiver. We shall see in Chapter 14 that any filtering action is confined almost exclusively

to the I-F amplifiers; thus the curves of Fig. 12-9*b* may be interpreted as the I-F pass characteristics. Notice that T.A. requires the broader I-F bandwidth, which shows up in performance as a lowered signal to noise ratio, which to a certain extent cancels out a substantial portion of the apparent power gain at the transmitter. The relative bandwidths have an even more important effect on the choice between the two systems. We shall find in Chapter 15 that the gain-bandwidth product for a given tube operating into a given type of plate load (*i.e.*, single-tuned or double-tuned) is constant. In T.A., each stage must have a wider bandwidth; hence the gain per stage will be less than for R.A. Thus a T-A receiver will in general require more I-F stages than will an equivalent R-A receiver. On this basis alone the R-A receiver should be cheaper.

In regard to the shape of the I-F characteristics, the R-A system requires that condition (12-43) be met, a fact that reduces the cost differential between the two systems. In T.A. the sharp cutoff at the lower end of the band must provide greater attenuation in order to protect the video signal from the sound program of the next lower telecast channel.

Faced with these conflicting factors, the National Television System Committee recommended that the R-A system be adopted as the standard. In spite of the fact that the need for the special filtering characteristic in R.A. violates the cardinal rule of simplifying the receiver at the expense of the transmitting equipment, the choice was predicated chiefly on the following factors: R.A. had been proved by numerous field tests, and at the time of the Committee meetings it was the consensus that the R-A receiver would be less expensive than its T-A counterpart. The R-A system of vestigial-sideband transmission has been adopted as standard in the United States. The bandwidth-utilization diagram is, of course, based on the transmitter characteristic and is shown in Fig. 12-10. Notice that the

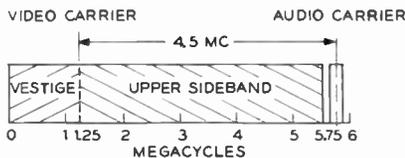


Fig. 12-10. Utilization of the 6-megacycle telecast channel with R-A vestigial-sideband transmission.

maximum permissible video signal is slightly less than 4.5 megacycles, a value which lies between the 2.85 and 5.7 megacycles of the double- and single-sideband systems, respectively. Thus the vestigial-sideband system of transmission provides the required compromise. The permissible video bandwidth has been reduced slightly in order that the design characteristics of the filter, which attenuates the lower sideband at the transmitter, are eased.

It should be realized that the R-A transmitter characteristic shown in Figs. 12-8 and 12-9 is idealized and it is desirable to set up standards to which a transmitter must conform within certain tolerances. This is most conveniently done in terms of an ideal detector, which may be used to check the transmitter output. In the early part of the chapter it was shown that, for small values of modulation index, the detector output at any modulating frequency is proportional to the sum of the sideband amplitudes corresponding to that modulating frequency. Thus, if an ideal detector were used to measure the characteristic of an ideal transmitter which meets the pass-band requirements of Fig. 12-8 exactly, the curve of rectified voltage *v.* modulating frequency would appear as shown in Fig. 12-11*a*. At first glance this curve seems to be of odd shape but consider the

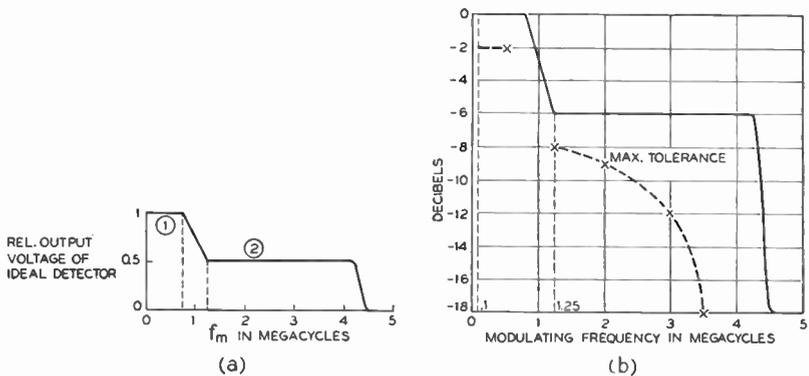


Fig. 12-11. The standard transmitter characteristic is specified in terms of the output of an ideal linear detector connected across the filter output. Notice that the phase response (and hence the transient behavior) is not specified. (a) Rectified voltage *v.* modulating frequency for a transmitter having the ideal R-A characteristic shown in Fig. 12-8. (b) Tolerances are specified relative to the ideal curve at five frequencies.

following: For modulating frequencies from 0 to 0.75 megacycle two sidebands of full strength are transmitted. For modulating signals in the range of 1.25 to 4 megacycles only the upper sideband is transmitted; hence the output in the upper range is one-half that of the lower range as indicated by the shelves labeled ① and ② in the diagram. For modulating frequencies between 0.75 and 1.25 megacycles the amplitude of the lower sideband tapers off with increasing modulating frequency. This tapering sideband added to the full-strength upper sideband gives the transition from the 1 to the 0.5 relative voltage level between 0.75 and 1.25 megacycles in Fig. 12-11a. The tolerances for the actual transmitter curve are expressed in terms of this idealized detector-output voltage curve. In general, the relative values are plotted in decibels, 0 decibel, or reference, level being that corresponding to the value between 0 and 0.75 megacycle on the ideal curve. The specified tolerances are indicated in Fig. 12-11b by the X's. It is further specified that in the actual transmitter the curve shall be smooth between the indicated points except for frequencies between 0.75 and 1.25 megacycles. The voltage of the lower sideband for modulating frequencies in excess of 1.25 megacycles must be at least 20 decibels below the reference level. It should be observed that current standards of the transmitter amplitude-frequency characteristic are quite mild. It should also be noticed that no specification is made on the phase characteristic of the transmitter; thus the transient response is not specified. The intention is to provide considerable flexibility at the present time and to raise the requirements in pace with progress in the art of transmitter design. A more complete discussion of the receiver characteristic is included in Chapter 14.

VESTIGIAL-SIDEBAND FILTERING

We have seen that the use of vestigial-sideband transmission of the R-A type permits the transmission of a 4- to 4.5-megacycle video bandwidth in the allotted 6-megacycle R-F channel, and that the radiated output of the transmitter must conform to the rather broad standards shown in Fig. 12-11. Since the modulated amplifier at the transmitter inherently generates two full sidebands, some sort of filtering action must be provided to suppress the lower sideband in the required manner. In general, this filtering action has been provided in two ways. (1) Modulation is carried out at a low power level and

lower sideband suppression is provided by off-center-tuned, linear R-F amplifiers. (2) Modulation is carried out at a high power level and a vestigial-sideband filter is interposed between the final R-F stage and the transmitting antenna system. We consider these two methods in order.

12-7. Off-center Tuning

It is well known that the magnitude of the output voltage of any network is equal to the product of its pass or amplitude characteristic times its input voltage. This fact may be utilized to suppress the lower sideband of an amplitude-modulated signal. Consider the curves of Fig. 12-12, which illustrate the principle involved. In passing through the tuned amplifiers, both sidebands are modified by the amplifier response. Since the amplifiers are tuned above the video carrier frequency, a large portion of the upper sideband is

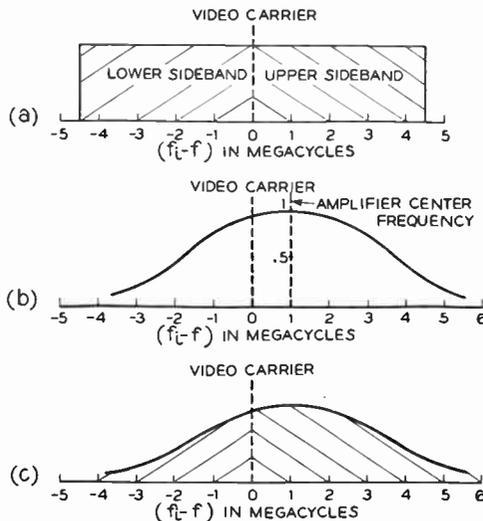


Fig. 12-12. Filtering by off-center tuning. A portion of the lower sideband is removed by tuning the linear amplifiers that follow the modulated stage to a frequency higher than the video carrier. Considerable control of the final pass band is available in the design of the amplifier stages. (a) Double-sideband output of the modulated amplifier. (b) The linear amplifiers are tuned to a frequency higher than the video carrier. (c) The final signal is the product of (a) and (b).

passed, and the lower sideband is reduced by a corresponding amount. By proper choice of the number of amplifier stages and proper design of their pass characteristics, the resulting signal of c in the diagram may be made to conform to the standards specified in the last section.² Use of this method of reducing the width of the lower sideband is discussed in Chapter 13.

12-8. Type A or Crossover Filter

The lower sideband from a modulated amplifier may also be removed by inserting a suitable filter between the transmitter and its antenna. In the present section we consider a typical filter of this type.³ In order that the energy of the suppressed portion of the lower sideband be prevented from causing reflections along the transmission line, it is good practice to dissipate that energy in a resistor separate from the antenna system. It was also considered important in the early development of the art to design a filter whose input impedance remains substantially constant over the entire band of operating frequencies, *i.e.*, over both the pass and reject bands of the filter, in order to eliminate reflections. Clearly, these characteristics are not met by conventional filters of the constant- k or m -derived category; in each case these filters present a reactive input impedance in the reject band of frequencies, a situation which results in multiple reflections between the transmitter and the filter.

The objections to these conventional filter types may be overcome by using a "frequency divider" or "crossover" network of the constant resistance type⁴. One basic form of such a crossover network is shown in Fig. 12-13*a*. We first demonstrate that the input impedance of the network is constant, independent of frequency. Reading from the diagram we have

$$Z_1 = j\omega L + \frac{R_o}{1 + j\omega C R_o} = \frac{R_o - \omega^2 L C R_o + j\omega L}{1 + j\omega C R_o} \quad (12-44)$$

and

$$Z_2 = \frac{1}{j\omega C} + \frac{j\omega L R_o}{R_o + j\omega L} = \frac{R_o - \omega^2 L C R_o + j\omega L}{-\omega^2 L C + j\omega C R_o} \quad (12-45)$$

² For information about response curves of multiple-tuned amplifiers and the effect of cascading several identical stages, see F. E. Terman, *Radio Engineering*, 3d ed. New York: McGraw-Hill Book Co., Inc., 1947, chap. 7.

³ G. H. Brown, "A Vestigial Sideband Filter for Use with a Television Transmitter." *RCA Review*, V, 3 (January 1941).

⁴ F. E. Terman, *Radio Engineers' Handbook*. New York: McGraw-Hill Book Co., Inc., 1943, p. 250.

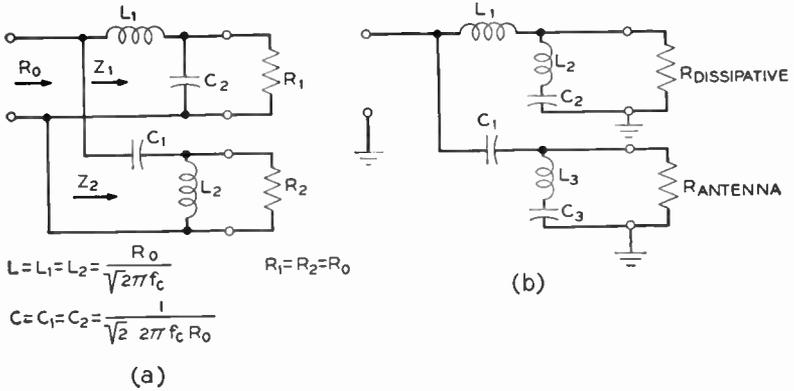


Fig. 12-13. Basic structures of the Type A vestigial-sideband filter. (a) A basic form of the constant-resistance crossover network. As viewed between the input and R_o , the structure behaves as a high-pass filter. (b) A sharper crossover may be obtained by utilizing series-resonant branches.

Then the input admittance, Y_{in} , will be

$$Y_{in} = \frac{1}{Z_1} + \frac{1}{Z_2} = \frac{1}{R_o} \left[\frac{(1 - \omega^2 LC) + j2\omega CR_o}{(1 - \omega^2 LC) + j \frac{\omega L}{R_o}} \right] \tag{12-46}$$

Now from the definitions for L and C in the diagram we have that

$$\frac{1}{\sqrt{2\pi f_c}} = \frac{L}{R_o} = 2R_o C$$

whence
$$\frac{\omega L}{R_o} = 2\omega CR_o \tag{12-47}$$

Thus the bracketed factor of (12-46) is unity, and the input impedance of the network is constant and equal to R_o .

We next consider the filtering action of the network. Let R_2 represent the transmitting antenna and R_1 a dissipating resistor of the same value. At low frequencies the reactance of L_1 is low and the reactance of C_2 is high and current flows through R_1 . In the Z_2 branch of the circuit C_1 offers a high reactance, and the reactance of L_2 is low, tending to short out R_2 . Thus at low frequencies little current flows through R_2 and most of the input power is dissipated in R_1 .

As the frequency of the input signal is raised, the two shunt networks gradually interchange roles until finally, at a sufficiently high

frequency, almost all of the input power is delivered to the antenna. At the crossover frequency, f_c , in Fig. 12-13a, the power is divided equally between R_1 and R_2 . Notice then that as the applied signal frequency is lowered, the input power is gradually transferred from the antenna to the dissipating resistor. As viewed from the antenna and input terminals, the entire crossover network appears as a four-terminal, high-pass filter. Thus, in general terms, the system of Fig. 12-13a meets the requirements of the vestigial-sideband filter. We have pointed out, however, that the crossover effect is gradual and unfortunately the cutoff in the equivalent high-pass filter is not sharp enough to meet the R-A requirements. This shortcoming may be remedied by modifying the basic structure slightly as shown at b in the diagram.

The sharper cutoff, which may be obtained by shunting the two resistances with series resonant circuits, is had at the expense of the constant input resistance. The series impedance branches can satisfy the design equations at a pair of frequencies only, and as a consequence the input impedance varies with frequency to a certain extent. By proper design, however, this variation may be minimized, and the circuit gives quite satisfactory results.

The actual design is based on three frequencies.

$$\left. \begin{aligned} f_1 &= \text{lower limit of the R-F channel} \\ f_2 &= f_1 + 1 \text{ megacycle} \\ f_3 &= f_1 - 1 \text{ megacycle} \end{aligned} \right\} \quad (12-48)$$

The circuit constants are chosen to meet the following conditions:

$$\left. \begin{aligned} L_1 \text{ and } C_1 & \quad \left. \begin{aligned} |X_{L1}| &= R_o \\ |X_{C1}| &= R_o \end{aligned} \right\} \text{ at } f_1 \\ L_2 \text{ and } C_2 & \quad \omega_2 L_2 - \frac{1}{\omega_2 C_2} = 0 \quad \text{at } f_2 \\ & \quad \omega_3 L_2 - \frac{1}{\omega_3 C_2} = -R_o \quad \text{at } f_3 \\ L_3 \text{ and } C_3 & \quad \omega_3 L_3 - \frac{1}{\omega_3 C_3} = 0 \quad \text{at } f_3 \\ & \quad \omega_2 L_3 - \frac{1}{\omega_2 C_3} = R_o \quad \text{at } f_2 \end{aligned} \right\} \quad (12-49)$$

It will be noticed that the six equations just stated may be combined to give the design values of the several lumped parameters. Actually we shall see that this step is unnecessary.

It is interesting to observe the behavior of the circuit at the two critical frequencies f_2 and f_3 . Figure 12-14 shows the equivalent circuits at these two frequencies. At f_3 , which is in the reject region of the lower sideband, L_3 and C_3 are series-resonant and short out the antenna to which no power will be delivered. Notice also that the

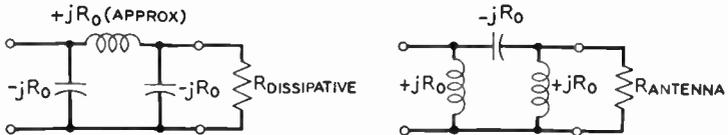


Fig. 12-14. Behavior of the Type A filter at the critical frequencies f_3 and f_2 . (a) At f_3 the circuit appears as a low-pass filter feeding the dissipating resistor. The antenna is shorted out by L_3 and C_3 (Fig. 12-13), which are series-resonant. (b) At f_2 the circuit appears as a high-pass filter feeding the antenna. The dissipating resistor is shorted out by L_2 and C_2 (Fig. 12-13), which are series-resonant.

right-hand end of C_1 will be grounded and the network appears as a low-pass filter feeding the dissipating resistor. Observe that the reactance values are shown for f_3 only. At f_2 , which lies in the pass band, the dissipating resistor is shorted out and the circuit appears as a high-pass filter feeding the antenna. Notice, then, that a complete crossover or cutoff occurs between f_2 and f_3 , a spread of 2 megacycles. In the final filter the cutoff is made even steeper by the addition of notching filters, which are described in a later section.

12-9. Calculation of the Distributed Constants

If any of the circuit inductance or capacitance values are evaluated for a typical R-F channel, it will be observed that the magnitudes are in the order of a few microhenries and tens of micromicrofarads. The problem of building a filter with these values of lumped circuit constants, which at the same time can be adjusted accurately and which can handle up to a kilowatt of power, is formidable. As a result the lumped parameters are generally replaced by sections of transmission line of the proper dimensions. A physical structure may be fabricated with less difficulty from the coaxial form of line than the parallel wire type. We next consider the reactance characteristics of such a line. It is well known that the input impedance of a lossless trans-

mission line of either the coaxial or parallel form with length l and characteristic impedance Z_o is given by^c

$$Z_{in} = Z_o \left[\frac{Z_R + jZ_o \tan \frac{2\pi l}{\lambda}}{Z_o + jZ_R \tan \frac{2\pi l}{\lambda}} \right] \tag{12-50}$$

where Z_R = terminating impedance,
 λ = wavelength of the applied signal on the line.

For the coaxial line with air dielectric, Z_o is a pure resistance equal to

$$R_o = 138 \log \frac{b}{a} \tag{12-51}$$

where b = inner diameter of the outer conductor,
 a = outer diameter of the inner conductor.

If the far end of the line is short-circuited ($Z_R = 0$), the input impedance becomes

$$Z_{in} = +jR_o \tan \frac{2\pi l}{\lambda} \quad \text{Short-circuited Line} \tag{12-52}$$

and if the far end is left open ($Z_R \rightarrow \infty$), Z_{in} becomes

$$Z_{in} = -jR_o \cot \frac{2\pi l}{\lambda} \quad \text{Open-circuited Line} \tag{12-53}$$

Notice that either the open- or short-circuited line behaves as a pure reactance and may be used to replace a lumped reactance, at least at any one particular frequency. Equations (12-52) and (12-53) are plotted in Fig. 12-15.

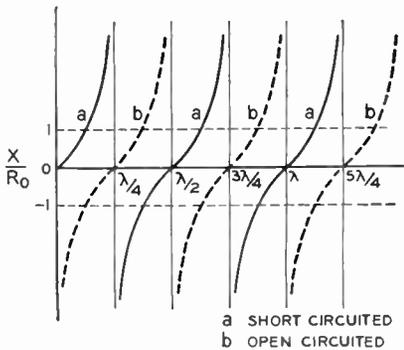


Fig. 12-15. Reactance curves for the short- and open-circuited lossless line.

^c See, for example, Terman, *op. cit.*

We must now consider this question: How shall the various lumped parameters be replaced by sections of transmission line? Consider first the capacitance C_1 . From (12-49) $|X_{C1}| = R_o$ at f_1 . Inspection of the curves of Fig. 12-15 shows that the minimum line length which satisfies this condition is $\lambda/8$ with an open-circuited line. The di-

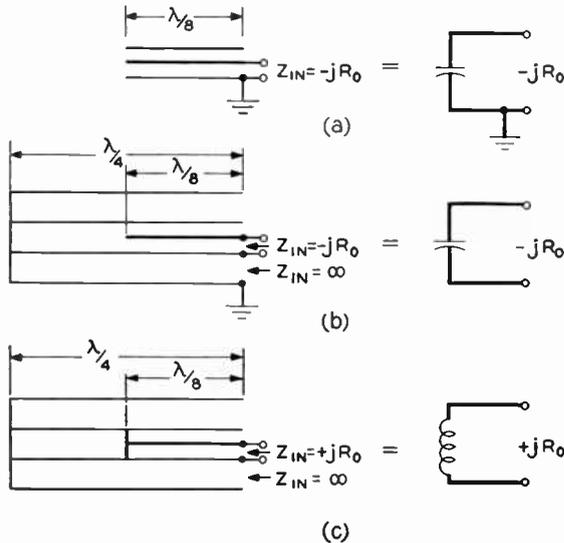


Fig. 12-16. Transmission-line equivalents of lumped circuit parameters. (a) An open-circuited stub of length $\lambda/8$ acts as a capacitive reactance of magnitude R_o , the characteristic impedance of the stub. (b) The equivalent capacitance is isolated from ground by extending the outer stub conductor an extra eighth-wavelength and surrounding it by a third conductor. The outer coaxial pair is short-circuited at the far end, is of length $\lambda/4$, and has an infinite input impedance. (c) Stub equivalent of an ungrounded inductance.

ameters of the conductor must, of course, be chosen to give the correct value of R_o . Thus, at f_1 , the capacitor C_1 may be replaced by the section of coaxial line illustrated in Fig. 12-16a, and the input impedance of the line becomes, since $l = l_1 = \lambda_1/8$

$$Z_{in} = -jR_o \cot \frac{2\pi\lambda_1}{8\lambda} = -jR_o \cot \frac{\pi\lambda_1}{4\lambda} \quad (12-54)$$

Since the argument contains λ , the input impedance of the line is a function of the applied frequency. In the lossless R-F line the phase velocity is constant so that

$$\frac{\lambda_1}{\lambda} = \frac{f}{f_1} \tag{12-55}$$

and the input impedance becomes

$$Z_{in} = -jR_o \cot \frac{\pi f}{4 f_1} \tag{12-56}$$

If a lumped condenser C_1 is used, its impedance at f_1 will be

$$\frac{1}{\omega_1 C_1} = R_o \tag{12-57}$$

and at any frequency

$$Z_{C_1} = \frac{-j}{\omega C_1} = -j \frac{\omega_1 R_o}{\omega} = -jR_o \frac{f_1}{f} \tag{12-58}$$

Equations (12-56) and (12-58) show that at a single frequency $f = f_1$ the line and capacitor have the same impedance, but at all other

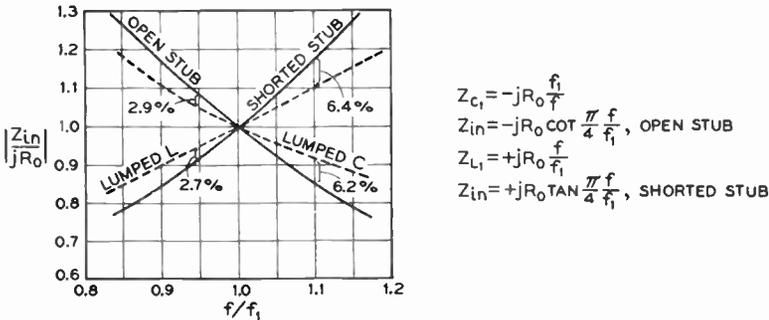


Fig. 12-17. Variation of Z_{in} with frequency for lumped constants and their stub equivalents. Both scales are expanded.

frequencies they differ. Equations (12-56) and (12-58) are plotted in Fig. 12-17.⁶ The maximum error incurred by replacing C_1 with

⁶ Equation (12-56) may be expanded in a Taylor's series about $\pi/4$.

$$\begin{aligned} \cot \frac{\pi f}{4 f_1} &= \cot \frac{\pi}{4} - \csc^2 \frac{\pi}{4} \left(\frac{f}{f_1} - 1 \right) + \dots \\ &= 1 - \frac{2\pi}{4} \left(\frac{f}{f_1} - 1 \right) + \dots \end{aligned}$$

or, to a first approximation,

$$Z_{in} = -jR_o \left(2.57 - 1.57 \frac{f}{f_1} \right)$$

The exact expression is plotted in Fig. 12-17, a fact which may be verified by the curvature in the line. The approximate equation plots as a straight line.

the open-circuited line depends, of course, upon the extreme limits of f/f_1 . Percentagewise these limits are greatest for the lowest transmission channel, No. 2, which extends from 54 to 60 megacycles. Allowing for a full video band of 4 megacycles the double-sideband limits will be 51.25 megacycles at the low end and 59.25 megacycles at the high. With an f_1 of 54 megacycles the limits of f/f_1 are 0.95 and 1.096. The total error incurred by use of the transmission line may be seen to be within 10 per cent over this range. For the higher channels, the f/f_1 limits, and hence the error, will be even less. We see, then, that the line may be used in place of the lumped capacitor.

One additional problem remains in replacing C_1 by the coaxial stub. The outer conductor of a coaxial cable is normally grounded. Reference to Fig. 12-13*b* shows, however, that both sides of C_1 or its equivalent must be above ground. This difficulty is overcome by using a "coax within a coax." Thus, the $\lambda/8$ open-circuited stub is placed within a larger tubing, which is grounded. The middle conductor is extended to a quarter wavelength and shorted to the outer shell. Then, the impedance, looking into the outer pair of tubes, is infinite as may be verified in Fig. 12-15. It should be noticed that the diameter of the outer shell is not critical for $Z_{in} = \infty$ for a quarter wave short-circuited line, regardless of its characteristic impedance. The final form of the line which replaces C_1 is shown at *b* in Fig. 12-16.

The inductance L_1 may also be replaced by a section of line, except that in this case the minimum length at which $|Z_{in}| = R_o$ is afforded by a short-circuited stub one-eighth of a wavelength long. The variation of the input impedance with frequency is given by⁷

$$Z_{in} = +jR_o \tan \frac{\pi}{4} \frac{f}{f_1} \quad (12-59)$$

and the corresponding impedance of L_1 is

$$Z_{L_1} = +jR_o \frac{f}{f_1} \quad (12-60)$$

The curves corresponding to these equations are plotted in Fig. 12-17. The final shorted stub with its outer shell is shown in Fig. 12-16*c*.

⁷ To a first approximation Z_{in} is given by Taylor's expansion

$$Z_{in} = jR_o \left(0.215 + 0.785 \frac{f}{f_1} \right)$$

Continuing the process of replacing the lumped parameters of Fig. 12-13*b* by their transmission-line equivalents, we next seek the equivalent of the series-resonant circuit comprising L_2 and C_2 . It may be seen from Fig. 12-15 that a short-circuited line appears as a series resonant circuit whenever its electrical length is an integral multiple of a half wavelength. Since L_2 and C_2 must be in resonance at f_2 , we may write that the required electrical length of the equivalent short-circuited line at f_2 is

$$l_2 = n \frac{\lambda_2}{2} \tag{12-61}$$

From the requirements set up in (12-49) we see that L_2 and C_2 must have an impedance $-jR_o$ at the frequency f_3 ; hence we may also write

$$Z_{in}]_{f_3} = jR_{o2} \tan \frac{2\pi l_2}{\lambda_3} = -jR_o \tag{12-62}$$

where R_{o2} is the characteristic impedance of the short-circuited section. The last two equations may be combined and simplified.

$$R_{o2} \tan \frac{2\pi n \lambda_2}{2\lambda_3} = R_{o2} \tan n\pi \frac{f_3}{f_2} = -R_o$$

Thus the design equation for the section is

$$R_{o2} \tan n\pi \frac{f_3}{f_2} = -R_o \tag{12-63}$$

where f_2, f_3 , and R_o are all known quantities. The difficulty in solving (12-63) is that two unknowns are involved, namely R_{o2} and the integer n and that the tangent function is periodic; hence a large number of solutions to the equation are possible. Three of these solutions are listed below.

| | | | | |
|-----|--------------|----------------|---|---------|
| n | R_{o2}/R_o | l_2 | } | (12-64) |
| 1 | 8.86 | $\lambda_2/2$ | | |
| 2 | 4.32 | λ_2 | | |
| 3 | 2.82 | $3\lambda_2/2$ | | |

If the reactance curves for these three values are plotted against frequency and compared with the corresponding curve for the lumped constants L_2 and C_2 , it will be found that the shortest line ($n = 1$) gives the closest approximation to the lumped case. However, an-

other factor must be considered. From eq. (12-51) we observe that as R_{o2} increases, the ratio of diameters of the coax outer and inner conductors must increase. In general, then, the larger is the required R_{o2} , the larger will be the tubing from which the filter section must be fabricated. Notice that by changing the value of n we can trade conductor diameter for conductor length and vice versa. In general, it is easier to assemble the filter if the diameters of all the sections are approximately equal. Brown has recommended that the design be based on $n = 3$.⁸ The same design procedure may be applied to the resonant branch comprising L_3 and C_3 . A schematic diagram of the complete Type A filter is shown in Fig. 12-18a. By cascading three

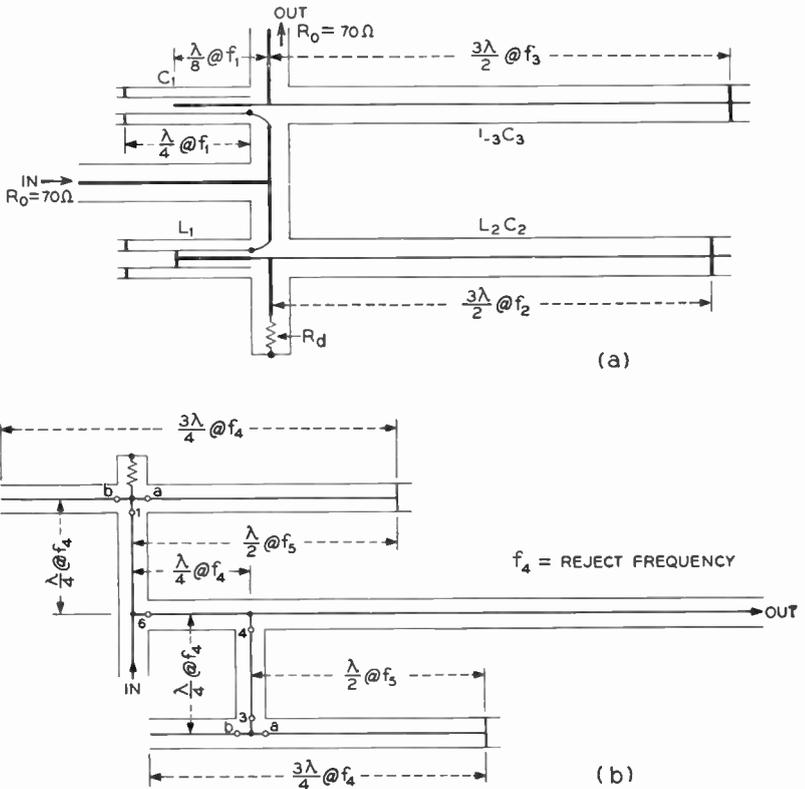


Fig. 12-18. Basic filter structures. (a) Type A filter. (b) Type B filter. (After Brown. Courtesy of RCA Review.)

⁸ G. H. Brown, *op. cit.*

such filters the amplitude response is cubed and a much sharper cutoff is obtained.

12-10. Type B, or Notching, Filter

For certain R-F transmission channels, for example, No. 3, the sound carrier of the next lower channel lies only 0.25 megacycles below f_1 , the lower frequency limit of the channel. Where such a situation exists it is desirable to have even greater attenuation than is provided by three cascaded Type A sections at this adjacent sound carrier frequency, say f_4 , in order to reduce the possibility of interference. The additional rejection which is required may be provided by the "notching" filter, shown at b in Fig. 12-18. This filter is of the constant-resistance type and provides an extremely narrow reject band. Consider its operation at the reject frequency which is defined by

$$\begin{aligned} f_4 &= \text{reject frequency of the type B filter} \\ &= f_1 - 0.25 \text{ megacycles} \end{aligned}$$

and let $f_b = f_1 + 2 \text{ megacycles}$ (12-65)

The calculations which follow are based on channel No. 2; hence

$$\left. \begin{aligned} f_1 &= 54 \text{ megacycles} \\ f_4 &= 53.75 \text{ megacycles} \\ f_b &= 56 \text{ megacycles} \end{aligned} \right\} \text{ (12-66)}$$

For the input impedances at the points marked a in the diagram we have

$$Z_a = +jR_o \tan \pi \frac{f}{f_b} \tag{12-67}$$

and at f_4 $Z_a]_{f_4} = +jR_o \tan \pi \frac{53.75}{56} = -j0.123R_o$ (12-68)

In order to calculate the input impedance at the points marked b we must first calculate the electrical length of the open-circuited section. Notice that as the frequency of the applied signal is decreased, the wavelength along the line increases, and the number of wavelengths which can fit into a line of fixed physical length decreases; hence the electrical length, l , which is the number of wavelengths along the line, is directly proportional to the frequency of the applied signal. Thus,

at f_4 the electrical length of the short-circuited line to the right of a is

$$l_a]_{f_4} = l_a]_{f_3} \frac{f_4}{f_3} = \frac{\lambda}{2} \frac{53.75}{56} = 0.96 \frac{\lambda}{2} \quad (12-69)$$

and the length of the open-circuited section is

$$l_b]_{f_4} = \frac{3\lambda}{4} - \frac{0.96\lambda}{2} = 0.54 \frac{\lambda}{2} \quad (12-70)$$

The impedance looking into b will be

$$Z_b = -jR_o \cot 0.54\pi \frac{f}{f_4} \quad (12-71)$$

and at f_4 this has the value

$$Z_b]_{f_4} = -jR_o \cot (0.54\pi) = +j0.13R_o \quad (12-72)$$

Notice that for all practical purposes Z_a and Z_b are equal in magnitude but of opposite sign; hence their parallel combination forms an anti-resonant circuit. Z_3 is therefore infinite or appears as an open circuit. Looking downward at point 4 we see an open-circuited section a quarter of a wavelength long which presents a short circuit. Thus the lower inverted \bar{T} short-circuits the output lead and no power can flow to the antenna. The short circuit at 4 reflects as an open circuit at 6. All the input power must flow into the upper \bar{T} section.

In the upper \bar{T} , Z_1 consists of the antiresonant circuit shunted by the dissipating resistor and hence has the value of that resistor. If its value is R_o , the characteristic impedance of the line, the quarter wave section between 6 and 1 acts as a 1 to 1 transformer and the input impedance of the filter is R_o . Thus at f_4 , the reject frequency, all of the input power is dissipated in the resistor and ideally no power is delivered to the antenna system.

The behavior of the circuit at f_5 , which lies in the pass band, may be checked in a similar manner. It may be shown that at f_5 Z_a and Z_3 are zero, making Z_4 infinite. In the upper \bar{T} , Z_a shorts out the dissipating resistor, making Z_1 zero. The quarter-wave section between 6 and 1 reflects this short circuit as an open circuit at the input; all the power is delivered to the antenna system.

The curves for a composite filter that employs both the Type A and Type B sections designed for Channel 1 are shown in Fig. 12-19.

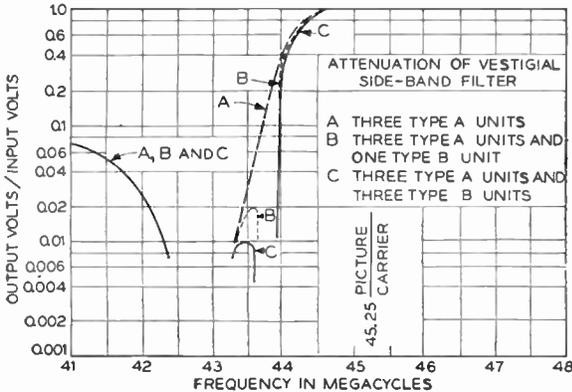


Fig. 12-19. Performance curves of the Type A and B filters.
(Courtesy of *RCA Review*.)

12-11. The Reactive Vestigial-sideband Filter

At the beginning of the last section it was stated that the constant-resistance type of network is desirable for use in the vestigial-sideband filter because it maintains a constant input resistance even in the reject frequency band and provides a separate resistor other than the load to dissipate the power of the rejected signals. More recent work has shown, however, that such conservative design is not necessary. In fact, in a video transmitter rated at 5-kilowatts peak power, the power in the reject band is of the order of 100 watts⁹ and this small amount of power may be dissipated in the plate circuit of the final stage of the transmitter without any difficulty. Furthermore, if the transmission line between the transmitter and filter is short, say 0.01 microsecond or less, in terms of delay time, the effect of reflecting this energy back to the transmitter is negligible. With these facts established it is apparent that a purely reactive filter of the constant- k or m -derived types may be used for vestigial-sideband filtering: the reject band energy is reflected from the filter input to the final transmitter stage, where it is absorbed. The chief advantages of the reactive filter over the constant-resistance type are twofold. First, sharper cutoff may be obtained with fewer filter sections, and second, the resulting coaxial structure is simpler and more economical.

⁹ E. Bradburd, R. S. Alter, and J. Raeker, "Vestigial Sideband Filter Design." *Tele-Tech*, 8, 10 (October 1949).

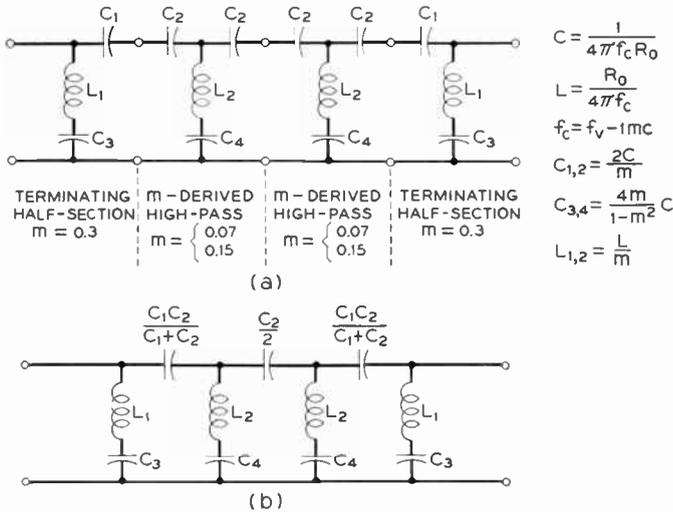


Fig. 12-20. The reactive vestigial-sideband filter. (a) The basic m -derived structure. (b) Adjacent series elements may be combined to simplify the structure. (Courtesy of *Tele-Tech.*)

A typical filter of this type consists of two m -derived $\bar{\Gamma}$ high-pass sections, plus two m -derived terminating π half-sections, which provide the proper impedance match at the input and output terminals. The basic structure is illustrated in Fig. 12-20a. As is usual in cascaded filters adjacent series elements may be combined to give the structure shown at *b* in the diagram. Since lumped capacitors may be used for the series elements, only four transmission-line elements, one for each resonant shunt arm, are required, and the circuit may be translated into a very simple mechanical structure, as shown in Fig. 12-21.

The two $\bar{\Gamma}$ sections are designed with $m = 0.07$ for channels No. 2 through No. 6, and with $m = 0.15$ for the upper channels No. 7 through No. 13. The terminating half-sections are designed with an m of 0.3 to provide proper termination at the input and output terminals. The design equations for the network in terms of lumped parameters are shown in the diagram. The shunt arms may be replaced by coaxial stubs as explained in the last few sections.

For one familiar with conventional filter theory an unusual feature of the filter will be immediately apparent: no prototype ($m = 1$)

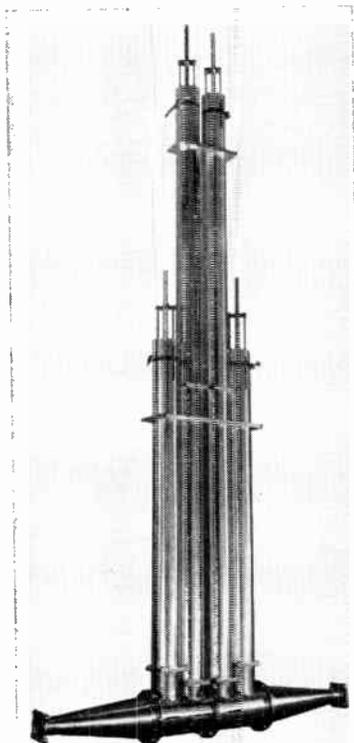


Fig. 12-21. The complete m -derived vestigial-sideband filter. The tapered units at the bottom are impedance-matching sections. Over-all height is approximately 6 feet. (Courtesy of Federal Telecommunication Laboratories, Inc.)

section is included. Its elimination results in some intersection impedance mismatch but the reflection losses caused by these mismatches are small enough to be neglected in the present case. Furthermore, since the pass band is less than 6 megacycles, the attenuation characteristic of the prototype section would tend to extend the reject region beyond the required limit, and so it is not included in the final design.

The attenuation and delay characteristics of the composite filter are shown in Fig. 12-22. The lack of cutoff in the region of 4 megacycles is corrected by controlling the bandwidth of the video ampli-

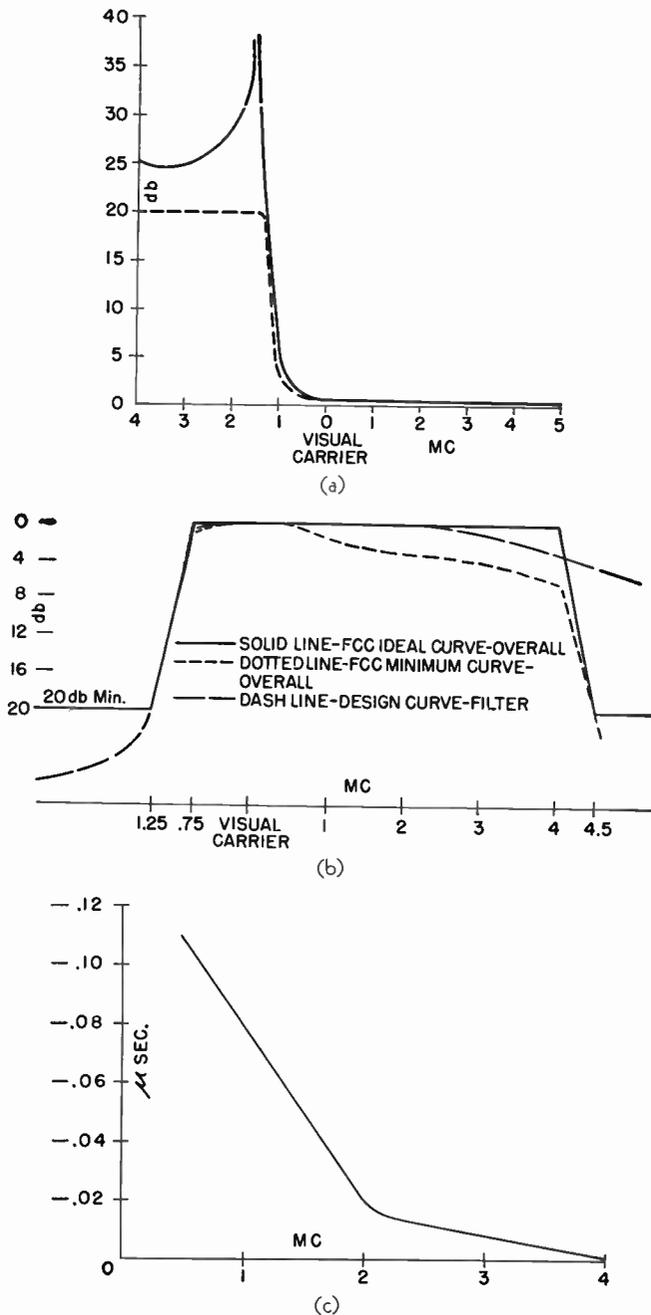


Fig. 12-22. Performance curves of the m -derived vestigial-sideband filter. (a) Attenuation versus frequency. (b) Measured and specified attenuation characteristics. (c) Time delay versus frequency. (Courtesy of Federal Telecommunication Laboratories, Inc.)

filters preceding the modulator. Correction of the delay characteristic is provided by an equalizer, which may also be included in the video section of the transmitter.

12-12. The Balun¹⁰

One final detail must be mentioned in connection with the vestigial-sideband filters just described. In constructing the filter from coaxial cable the outer conductor is grounded, causing the filter output to be unbalanced with respect to ground. The antenna system, which is fed from the filter, must be balanced to ground; hence some sort of device is needed to convert from one condition to the other. There are several such devices which are commonly referred to as "Bazookas" or "Baluns" (*balanced to unbalanced converter*). One basic form of the Bazooka is illustrated in Fig. 12-23. Use is made of the fact that the inner and outer surfaces of the conductors are not at the same potential and that the input impedance of a short-circuited quarter-wave line is infinite. Thus, even though the outer surface is grounded, terminal *b* is above ground by virtue of the fact that Z_{in} is infinite. Methods of broadbanding the Balun are discussed in the next chapter.

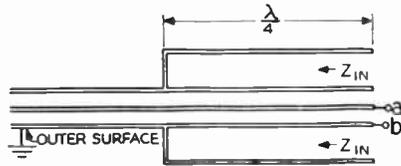


Fig. 12-23. The Balun, or Bazooka, a coaxial transformer for converting from unbalanced to balanced-to-ground transmission.

Use is made of the fact that the inner and outer surfaces of the conductors are not at the same potential and that the input impedance of a short-circuited quarter-wave line is infinite. Thus, even though the outer surface is grounded, terminal *b* is above ground by virtue of the fact that Z_{in} is infinite. Methods of broadbanding the Balun are discussed in the next chapter.

¹⁰ R. C. Paine, "The Balun—A Transmission Line Transformer." *Radio-Electronic Engineering Edition of Radio and Television News*, 11, 6 (December 1948).

CHAPTER 13

THE VIDEO TRANSMITTER

We have already set up most of the standards which govern the transmitter portion of the pickup end of the television system. The channels have been defined, the R-A transmission characteristic has been described, and we know that the video and sync information is to amplitude-modulate the radio-frequency carrier. One more important point must be decided before we may consider the video transmitter proper. Since the composite video signal is not symmetrical about its average value, we must decide whether the power peaks of the modulated wave shall correspond to the sync signals or to a maximum white picture signal. The former method is called "negative transmission" because an increasing light intensity at the camera tube causes a decreasing output power at the transmitter. The second method, where a white signal gives maximum power out from the transmitter, is known as "positive transmission."

The importance of this standard in receiver design is that the designer must know if an odd or even number of stages should occur between the second detector and the cathode-ray-tube control grid. Since the reproduced image is sensitive to phase, the incorrect number of stages will cause a negative image to be reproduced, *i.e.*, the blacks and whites will be interchanged. Other factors also enter in the choice between positive or negative transmission. We consider some of them in the following section.

13-1. Positive or Negative Transmission¹

It is interesting to note that in the United States negative transmission has been standardized while in Great Britain positive transmission is used. Some of the reasons which governed the American choice of having the sync pulses produce maximum radiated power will now be given.

¹ D. G. Fink, *Television Standards and Practice*. New York: McGraw-Hill Book Company, Inc., 1943.

The envelopes of the modulated wave for both types of transmission are shown in Fig. 13-1. Consider what effect a burst of noise received with the television signal will have on the reproduced image with both types of transmission. Assume that the noise occurs in the form of a short pulse during the scanning interval when picture

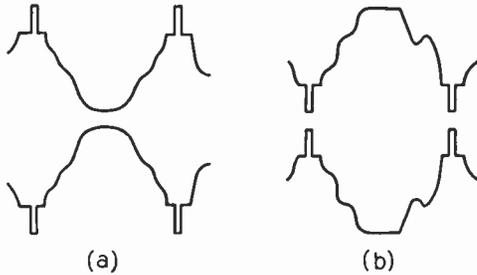


Fig. 13-1. Envelope of the modulated signal for negative and positive transmission. (a) Negative transmission. Maximum power is radiated for the duration of the sync pulses. (b) Positive transmission. Maximum radiated power corresponds to a white picture signal.

information is being transmitted. In either system this pulse will add to the amplitude of the received signal; at *a* in the diagram the noise will drive the signal toward the blacks, and at *b* toward the whites. Thus, in negative transmission, R-F noise shows up as narrow black streaks in the reproduced image. It was the opinion of the standardizing committee that black-appearing noise is less objectionable to the spectator than the white-appearing noise that would be present with the positive type of transmission.

Negative transmission also reduces the peak-power requirements on the transmitter as compared to the other system. This may be verified from the following considerations. In the negative system the transmitter delivers peak power only during the several sync intervals which are narrow and of known duration. On the other hand, the transmitter must be designed to deliver maximum power output whenever a white signal is being transmitted in the positive transmission system. The most severe condition that can occur here is that when several consecutive all-white lines are transmitted. Let us calculate the relative power demands on the final transmitter stage under both systems.

Consider first the negative system of transmission. By referring to Fig. 11-11 we may calculate the total duration of all the horizontal sync, vertical sync, and equalizing pulses over an entire *frame* interval. Thus, there are,

24 equalizing pulses of duration $0.04H$:

$$\text{total duration} = 24(0.04)H = 0.96H$$

12 vertical sync pulses of duration $0.43H$:

$$\text{total duration} = 12(0.43)H = 5.16H$$

Over the entire frame 18 horizontal sync pulses are lost for equalizing and vertical sync; hence there are

$525 - 18 = 507$ horizontal sync pulses of duration $0.08H$:

$$\text{total duration} = 507(0.08)H = 40.56H$$

Thus the total sync pulse duration over one frame is $46.68H$. Now

$$H = \frac{1}{f_l} = \frac{1}{nf_p} \quad (13-1)$$

so that the per cent duration of the sync pulses is

$$\% \text{ pulse duration} = \frac{46.68H10^2}{1/f_p} = \frac{(46.68)10^2}{n} = \frac{4668}{525} = 8.9\% \quad (13-2)$$

or approximately 9 per cent of the frame interval is devoted to transmitting the peak signal. Using the approximate figure of 9 per cent we may calculate the contribution of the several pulses to the average power when an all-black signal is being transmitted. The voltage signal is shown at *a* in Fig. 13-2. Black level is at 0.75 peak voltage, as specified in the standards. At *b* the corresponding power is shown. Since power is proportional to the square of voltage, black level is shown at $(0.75)^2 = 0.56$ peak amplitude. Then the pulse power-amplitude is $1 - 0.56 = 0.44$, and

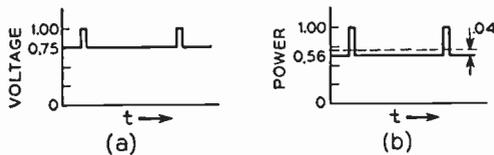


Fig. 13-2. Voltage and power for an all-black signal in negative transmission. (a) Voltage signal. (b) Corresponding power.

average power contributed by the pulses

$$= 0.44(0.09) = 3.96\% \quad (13-3)$$

and the average power of the signal is

$$\text{average power} = 0.56 + 0.0396 \approx 60\% \quad (13-4)$$

Hence, with negative transmission of an all-black signal, the average power is approximately 60 per cent of the peak power supplied during each pulse. Notice that in negative transmission an all-black picture places the most severe demands on transmitter power.

Next consider the most severe case in positive transmission, *i.e.*, that in which a white line is being sent, whose voltage and power diagrams are shown in Fig. 13-3. Notice that 25 per cent of the full-voltage amplitude has been reserved for the synchronizing signals.

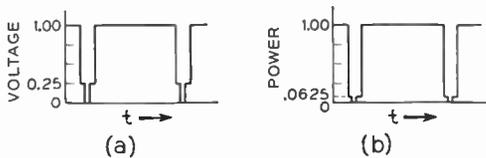


Fig. 13-3. Voltage and power for an all-white signal in positive transmission. (a) Voltage signal. (b) Corresponding power.

In order to calculate the average value of the signal shown at *b* we must calculate the areas resulting from blanking and the several synchronizing and equalizing pulses. The necessary dimensions may be obtained from Fig. 11-11. Then, following our previous procedure, we have

2 vertical blanking pulses of duration $0.075V$:

$$\text{total duration} = 2(0.075) \frac{(525)}{2} H = 39.4H$$

507 horizontal blanking pulses of duration $0.16H$:

$$\text{total duration} = 507(0.16)H = 81.1H$$

Expressed in terms of the frame interval, $1/f_p$, the total blanking duration is:

$$\frac{\text{total blanking duration}}{1/f_p} = \frac{120.5H}{1/f_p} = \frac{120.5}{525} = 22.9\% \quad (13-5)$$

From Fig. 13-3b the amplitude of the power blanking pulses is $1 - 0.0625 = 0.9375$, and their per cent area is

$$\text{average power of the blanking pulses} = (0.9375)(22.9) = 21.5\%$$

We have already calculated the sync and equalizing pulse duration to be 8.9 per cent, and with a power amplitude of .0625 they have an average power

$$\begin{aligned} \text{average power of the sync and equalizing pulses} \\ = (0.0625)(9) \approx .56\% \quad (13-6) \end{aligned}$$

The total average power represented by sync, equalizing, and blanking pulses is, therefore, roughly 22 per cent or under the most severe condition of positive transmission the average power supplied by the transmitter is 78 per cent of the peak power.

The two systems may now be compared. We shall assume that the final stage of a visual transmitter operates at constant plate-circuit efficiency so that the plate dissipation is proportional to the power output delivered. For a constant plate dissipation the average power output must remain unchanged; hence the peak power which may be delivered with negative transmission is $(0.78)/(0.6)$ times that delivered with positive transmission. The negative transmission system affords a 30 per cent increase in peak power over the positive transmission system for a given final stage.

Another advantage of negative transmission lies in the fact that regardless of picture content the maximum signal level is reached at least once during each line; hence the black level remains fixed and the average value of the picture component is carried along in the modulated wave. This also simplifies the automatic brightness or automatic gain control (A.G.C.) circuits in the video portions of the television receiver.

On the other side of the ledger, we see that the sync pulses of the composite video signal must drive the modulated amplifier into the upper part of its modulation characteristic, which generally tends to exhibit some curvature. If uncompensated, this effect tends to compress the sync pulses in the modulated wave. Satisfactory compensation may be provided by a "sync stretcher," which is described in a subsequent section in the present chapter.

It was felt that the reasons listed above justified the standardization of negative transmission in the United States. In building up the video transmitter we must make sure that the phase of the composite

video signal delivered to the modulated amplifier is such that the synchronizing pulses cause maximum power to be delivered to the antenna system.

13-2. Transmitter Block Diagrams

Our various transmission standards may now be summed up in terms of the functions required of the visual transmitter. In broad terms it must perform the following operations:

- (1) Generate an R-F carrier of proper frequency, stability, and power.
- (2) Accept the composite video signal from the pickup facilities and raise its power level to the value required at the modulated amplifier. Its sense or phase must be such as to produce negative transmission.
- (3) Amplitude-modulate the R-F carrier.
- (4) Deliver the R-A type vestigial-sideband signal to the antenna system.

At the time of writing three basic approaches are being used to accomplish these functions. For the lack of more suitable notation we shall identify the three types by the name of the manufacturers who make them, *i.e.*, Radio Corporation of America (R.C.A.), General Electric Company (G.E.), and Allen B. Du Mont Laboratories, Inc. (Du Mont). The three systems are illustrated in block diagram form in Fig. 13-4.

Inspection of the diagrams shows that the predominant difference in the three types of transmitters lies in the power level at which modulation is performed and in the manner of suppressing the lower sideband to meet the R-A vestigial-sideband requirement.² We shall discuss the three diagrams in ascending order of video-modulating power.

In the G.E. system modulation is performed at low power levels. As may be seen from the diagram the R-F power input to the modulated stage is 1 watt and an 80-volt peak-to-peak modulating signal is required. The advantage of excellent modulation linearity may be realized at this power level with a plate-modulated stage. The modulated R-F wave is fed through five linear off-carrier-tuned R-F

² D. G. Fink, "Design Trends in Television Transmitters." *Electronics*, 21, 1 (January 1948).

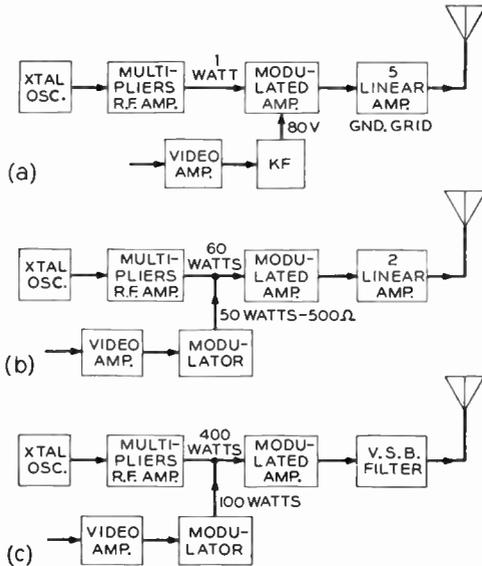


Fig. 13-4. Block diagrams of three commercial video transmitters. (a) The General Electric system. (b) The Du Mont system. (c) The R.C.A. system. (Courtesy of *Electronics*.)

amplifiers, which serve a dual function. They raise the power to the required level and clip the lower sideband to meet the vestigial requirements. This system of sideband filtering has been outlined in the last chapter and illustrated in Fig. 12-12. The low-level modulation system may be summarized in its essentials as follows: Low-level plate modulation presents no great problems; no separate vestigial-sideband filter is required; five linear amplifier stages are required. The saving entailed by eliminating the separate sideband filter is offset to some extent in that the linear R-F amplifiers are comparatively difficult to tune. Several adjustments are required in order to maintain linearity between input and output and to maintain the correct filtering characteristics.

In the Du Mont transmitter modulation takes place at an intermediate power level, the required input powers being 60 and 50 watts, respectively, for the R-F and composite video inputs to the modulated stage. At that power level either grid or plate modulation may be used but the particular circuit shown in the figure employs the former.

Since the power level is not excessive, the stage may be designed conservatively to give excellent linearity. Power amplification and clipping of the lower sideband is obtained in two off-carrier-tuned linear R-F stages. No separate vestigial-sideband filter is required. As compared to the G.E. transmitter, the Du Mont unit has a slightly more complicated modulation problem but fewer tuning adjustments because of the reduced number of linear amplifiers employed.

The R.C.A. transmitter represents a complete departure from the two other units in the method of vestigial-sideband filtering employed. Since the modulation is accomplished in the ultimate stage of the transmitter, the need for linear R-F amplifiers is eliminated and the lower sideband is clipped by a separate sideband filter of the general type described in the last chapter. At the present stage of development high-level modulation requires that grid modulation be used; the reason for this is covered in a subsequent section. The transmitter is easily tuned since no linear R-F amplifiers are required, but the problem of modulation is greater than that of the other two systems and the separate filter unit is required after the modulated stage.

It is not our purpose to weigh the commercial advantages of the three forms of transmitter design. Our emphasis of any one particular design is based solely on presenting some of the chief problems encountered in the television art. Since the high-level-modulation vestigial-sideband filter layout represents the greatest departure from conventional broadcasting techniques, our discussion will be confined primarily to that form of transmitter. It should be stressed that in the work which follows no single commercial transmitter is being discussed; wherever possible our discussion will be in general terms and will indicate only one of many possible variations in design.

PLATE V. GRID MODULATION

We have stated that the present stage of development places a restriction on the type of high-level modulation which may be used; if modulation takes place in the final stage of the transmitter, that stage must be grid-modulated. Since the problem of modulation is of prime importance in the video transmitter, we next consider the reason behind this statement. We shall briefly review the principles of plate and grid modulation for a background. In the discussion it

is assumed that the modulated stage is balanced, *i.e.*, of the push-pull type, though a single-ended stage may be used just as well.

The basic diagrams of the two circuits are shown in Fig. 13-5. In both circuits, direct coupling must be used between the modulating

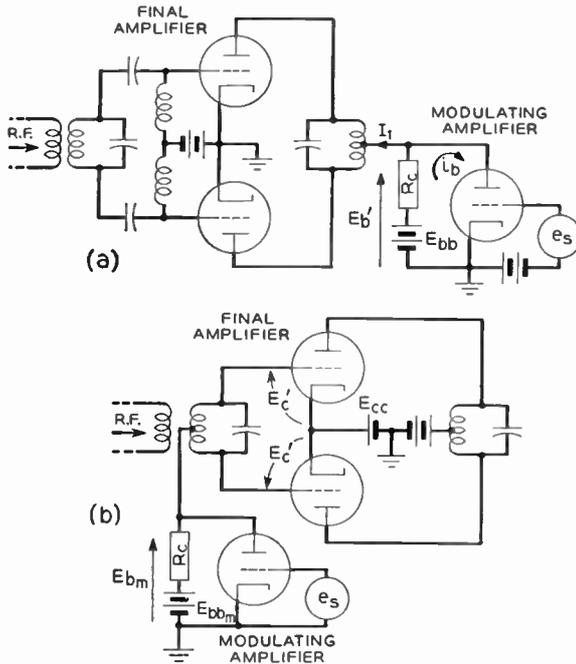


Fig. 13-5. Basic circuits for plate and grid modulation. Direct coupling must be used between modulating and modulated amplifiers to preserve the d-c component. (a) Plate modulation. (b) Grid modulation.

and modulated stages in order that the d-c component of the video signal be transmitted. Since transformer coupling cannot be used, the modulating stage must operate in Class A, as opposed to linear Class B, and hence will have a comparatively low efficiency.

13-3. Plate Modulation

The purpose of the modulated stage is to cause the envelope of the R-F carrier to vary in accordance with the modulating signal; hence for plate modulation we seek some device whose R-F voltage output is

directly proportional to its plate voltage. Such a device is the Class C R-F amplifier.^{3,4} When the push-pull R-F amplifier shown at *a* in Fig. 13-5 is biased at approximately twice the cutoff value, the required linear relationship is obtained and modulation may be obtained by making the effective plate voltage, E_b' , vary in proportion to the modulating signal. Notice from the diagram that

$$\begin{aligned} E_b' &= E_{bb} - (I_1 + i_b)R_c \\ &= (E_{bb} - I_1R_c) - i_bR_c \end{aligned} \quad (13-7)$$

where I_1 = average plate current of the modulated amplifier,

i_b = total plate current of the modulating amplifier.

If a single-ended final stage were used, a shunt-plate feed system would be required in order to isolate the coupling resistance R_c from the pulsating a-c components of modulated amplifier current. This problem does not arise with a push-pull final. Since i_b follows the modulating signal e_s , the proper linear relationship between E_b' and e_s is maintained. Notice that in order to obtain 100 per cent modulation, the drop i_bR_c must equal the d-c plate voltage, which is $(E_{bb} - I_1R_c)$ or, in other words, the output voltage of the modulating stage must be of the same order of magnitude as the d-c plate voltage on the modulated stage. This may not seem particularly serious until we consider the power requirements on the modulating amplifier. Since a 4.5-megacycle video bandwidth must be preserved, the value of coupling resistance, R_c , must be chosen with an eye toward the shunt capacitance across it and will in general be low. Since power is inversely proportional to resistance, this means that the power requirement on the modulator is large. We may illustrate this with an example. Let the modulated stage employ a Type 891 water-cooled tube, which has a rated power output of 5 kilowatts with a d-c plate voltage of 8 kilovolts.⁵ Then the modulating stage must furnish 8 kilovolts of modulating voltage across R_c . Assume a total shunt capacitance of 100 micromicrofarads across the resistor. We can now calculate the value of resistance required in order to maintain a constant load over the 4.5-megacycle bandwidth. From our

³ W. L. Everitt, *Communication Engineering*. New York: McGraw-Hill Book Company, Inc., 1937.

⁴ F. E. Terman, *Radio Engineering*, 3d ed. New York: McGraw-Hill Book Company, Inc., 1947.

⁵ R.C.A. Tube Handbook, Transmitting Types Section.

study of the resistance-coupled amplifier, we know that the impedance of a shunt combination of resistance and capacitance remains substantially constant up to one-tenth of the half-power frequency defined by

$$f_2 = \frac{1}{2\pi R_c C} \quad (13-8)$$

Then setting $f = 4.5$ megacycles $= 0.1f_2$ (13-9)

we have for the required value of coupling resistance⁶

$$R_c = \frac{1}{2\pi f_2 C} = \frac{1}{2\pi(4.5 \times 10^7)10^{-10}} = 354 \text{ ohms} \quad (13-10)$$

The power output required of the modulator during the sync peaks will be

$$P_o = \frac{(8 \times 10^3)^2}{3.54 \times 10^2} = 181 \text{ kilowatts} \quad (13-11)$$

In order for the modulating stage operating in Class A to deliver this power a great number of large tubes in parallel and a correspondingly large power supply would be required. Stated in other terms, the modulator would have to deliver a peak current of

$$i_{b\max} = \frac{8000}{354} = 22.6 \text{ amperes} \quad (13-12)$$

It is generally believed that a design which makes such excessive demands on the modulating stage is not advisable at the present stage of development, and plate-level modulation of the final stage is not used.

For the sake of completeness we consider the polarity of the modulating-stage driving signal, E_s , required to give negative transmission. Since the modulated stage delivers peak power when E_b' is maximum, the signal voltage developed across R_c must have a sync-positive phase. Since a 180° phase reversal occurs in the stage, E_s must be a sync-negative signal.

13-4. Grid Modulation

It is well known that an R-F amplifier biased at cutoff exhibits linearity between its output voltage and its grid signal voltage; hence

⁶ Normally compensation will be used and the calculated value of resistance will be greater. We assume no compensation in comparing the two types of modulation.

such an amplifier may be used as a grid-modulated stage. The basic circuit is shown in Fig. 13-5*b*. Again d-c coupling must be used between the modulating and modulated stages and the former must be operated in Class A. With the arrangement shown the effective bias on the push-pull stage is

$$E'_c = E_{cc} - (E_{bbm} - i_b R_c) \quad (13-13)$$

and by varying i_b , the modulator plate current, we can make the effective bias vary in proportion and hence control the amplitude of the output R-F voltage: grid modulation is obtained. In order to provide 100 per cent modulation, the modulating stage must furnish a voltage equal to the cutoff bias of the modulated amplifier which will be $1/\mu$ times as great as the driving voltage required in plate modulation. Thus in grid modulation the modulating voltage required is of the same order as the bias, rather than the plate voltage. This can result in a large saving in modulating power. For example, the 891 has a rated μ of 8. Then the ratio of modulating powers required for grid and plate modulation is $1/\mu^2 = 1/64$.

These facts may be demonstrated by carrying on our previous example. We shall assume the same tube in the modulated stage, *i.e.*, the 891, and the same coupling resistance of 354 ohms. Since the grid-modulated stage is biased at cutoff, its bias voltage must be

$$E_c = \frac{E_b}{\mu} = \frac{8000}{8} = 1 \text{ kilovolt} \quad (13-14)$$

and the required voltage across R_c will have the same value. Thus the peak power required of the modulating amplifier is

$$P_o = \frac{(10^3)^2}{354} = 2.83 \text{ kilowatts} \quad (13-15)$$

and the peak modulating current is

$$i_{b\max} = \frac{1000}{354} = 2.83 \text{ amperes} \quad (13-16)$$

Comparison of these figures with those calculated for the plate-modulated case shows the tremendous saving in modulating power afforded by grid modulation. In the latter case the required driving power may be furnished by two 891's in parallel. It may be shown quite readily that a sync-positive picture signal is required at E_s , the modulator input, in order to produce negative transmission.

Because of its less severe power demands grid modulation is used in the high-level-modulation vestigial-sideband filter type of transmitter. Further details of the circuit, particularly with reference to the methods of direct-coupling the modulating signal, are covered in a later section.

THE R-F SECTION

The over-all block diagram of the high-level-modulation vestigial-sideband transmitter may be broken down into four sections, each of which performs one of the functions listed earlier in the chapter. Of these we shall first consider the radio-frequency section, which serves to generate a radio-frequency carrier of the proper frequency, stability, and power. The components of the R-F section generally comprise an oscillator, frequency multipliers where necessary, and power amplifiers.

13-5. Oscillator

The F.C.C. standards on television transmitters require that the carrier frequency be maintained at its assigned value within a tolerance of ± 0.002 per cent; hence a prime design requirement on the R-F oscillator is frequency stability. The use of crystal control at the carrier frequency is not feasible because crystals generally do not perform satisfactorily at frequencies in excess of 10 megacycles,⁷ a value which is well below the assigned television channels. The operating frequency of a crystal is determined by its physical dimensions, and the crystal size is too small to be practicable when it is ground to operate above the 10-megacycle limit. To circumvent this difficulty the design of the oscillator may be centered about some form of stable resonant circuit.

We have already seen that sections of transmission line are more convenient than lumped circuit constants at television R-F frequencies; hence we seek the transmission line equivalent of a parallel resonant circuit. Inspection of Fig. 12-15 shows that the requirements may be met by a short-circuited line, one-quarter wavelength long. Figure 13-6 shows the schematic of a typical push-pull oscillator, employing a quarter-wave resonant section in the coaxial form as the grid tank circuit. The coaxial-, rather than parallel-, type line is used as the frequency-determining member because it may

⁷ F. E. Terman, *op. cit.*

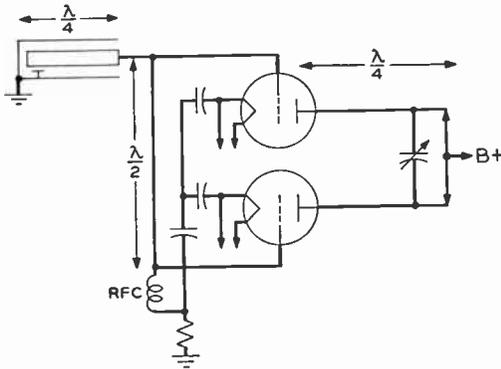


Fig. 13-6. Basic circuit of a push-pull oscillator that uses a short-circuited quarter-wave coaxial line as the grid tank circuit.

be built more rigidly and hence is less susceptible to frequency drift. In certain designs the effect of temperature change on the length of the coaxial section is minimized by forming the line out of some metal which has a low coefficient of thermal expansion, such as Invar. Notice that the closed end of the section is not short-circuited directly but through a small capacitor. Fine tuning may then be had by moving the inner conductor axially with a micrometer screw. In order to ensure good frequency stability the resonant section is designed to have an R_o of approximately 77 ohms, which value corresponds to a b/a ratio of 3.6. For this ratio the Q of the section is at its maximum value.⁸

While the coaxial line is desirable from a mechanical viewpoint, it introduces an electrical problem in that the two grids of a push-pull oscillator must be driven 180° out of phase. Since the outer conductor of the coax is grounded, it may not be connected directly to the grid of V_2 . This difficulty is overcome by utilizing the fact that a section of line one-half wavelength long serves as a 1 to 1 transformer, which provides a 180° phase reversal of voltage. Thus in the diagram, since the grid of the lower tube is one-half a wavelength "away" from the grid of the upper tube, the two grids are driven in phase opposition as required.

⁸ A. B. Bronwell and R. E. Beam, *Theory and Application of Microwaves*. New York: McGraw-Hill Book Company, Inc., 1947.

A parallel-type line of length $\lambda/4$ is used as the plate tank circuit. Coarse tuning is obtained by moving the shorting bar located at the far end of the line; fine tuning is accomplished by means of a small variable condenser connected across the line.

A much simpler and more stable oscillator may be designed by utilizing crystal control. We have already seen that it is not practicable to grind a crystal to operate directly at any of the television carrier frequencies; hence the crystal is chosen to operate at a subharmonic of the carrier frequency, and the crystal frequency is then raised to the required value by means of frequency multipliers. In the interests of efficiency of the multipliers it is desirable that the multiplication per stage be limited to two or, at the most, three. This means that the ratio of carrier to crystal frequency should be a number whose factors are two or three. One convenient ratio is $12 = 2 \times 3 \times 2$. Thus, for example, a crystal cut for 4.6042 megacycles, which is a practicable value, may have its frequency raised to the carrier frequency of 55.25 megacycles for channel No. 2 by two doublers and one tripler. Since the multipliers raise the power level as well as the frequency, the line-up of crystal oscillator plus frequency multipliers represents a good economic compromise. Its chief advantage is that use is made of the high stability provided by the special temperature-compensated crystal cuts which are now available.

13-6. R-F Amplifiers

The amplifiers which follow the R-F oscillator are tuned power amplifiers operating in Class C. Notice that only a single frequency is amplified in any stage and no attention need be paid to broad-banding.

In those stages where frequency multiplication takes place, the plate tank circuit is tuned to the proper harmonic of the applied frequency, *i.e.*, to the second harmonic in a doubler, and to the third harmonic in a tripler. Since the plate current, which flows in the form of nonsinusoidal pulses, is rich in harmonics, the tank circuit can deliver power at the harmonic to which it is tuned. The particular harmonic being selected may be strengthened by proper choice of the grid-bias voltage.⁹

⁹ F. E. Terman, *op. cit.*, chap. 7.

13-7. Grounded-grid Amplifier¹⁰

In those stages where power amplification and no frequency multiplication is needed, the grounded-grid circuit may be used to an advantage. This circuit differs from the more conventional grounded-cathode and grounded-plate (cathode follower) connections in that its grid is operated at a-c ground potential.

The basic circuit is shown in Fig. 13-7. The grounded-grid circuit offers three distinct advantages in circuit applications: (1) It is easy to neutralize. (2) With the newer tubes of the ring or disk seal type it is readily adaptable to resonant circuits of the coaxial line form. (3) It delivers a greater power output from a given tube.

We shall consider these features.

As may be seen from the diagram, theoretically no neutralization is required in the grounded-grid connection; since the ground plane extends between the cathode and plate structures in the tube, no feedback between the output and input circuits can occur through the plate-to-cathode interelectrode capacitance. This represents an ideal condition, however, for in most circuits the grid structure will be above ground because of the inductance of the wire running from grid to ground. Where this condition exists, neutralization may be accomplished by introducing a lumped reactance of the proper magnitude and sign in series with the grid lead. This may now be demonstrated. The actual interelectrode capacitances in a triode are Δ -connected, as shown in Fig. 13-8a. The purpose of neutralization is to reduce the admittance of the path between the plate and cathode to zero. This may be accomplished by introducing an effective ground plane between P and K .

Let the Δ network of capacitances be replaced by its equivalent T configuration as shown at b . Then the circuit will be neutralized if the point O is brought to ground potential, a condition which may be brought about by making the branch LC_3 series resonant. The value

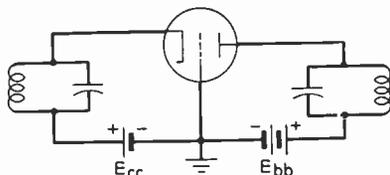


Fig. 13-7. Basic grounded-grid amplifier circuit.

¹⁰ C. J. Starnes, "The Grounded-Grid Amplifier." *Broadcast News*, No. 42 (January 1946).

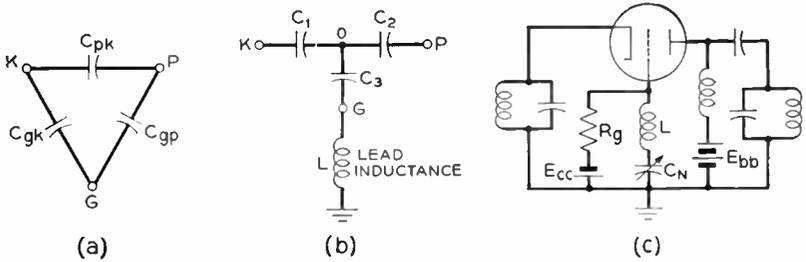


Fig. 13-8. Neutralization of the R-F grounded-grid amplifier. (a) The interelectrode capacitances in a triode are Δ -connected. (b) The Δ connection may be replaced by an equivalent T. (c) C_n is added to make the LC_n branch series-resonant.

of C_3 may be calculated in terms of the interelectrode capacitances by direct application of the $\pi = T$ transformation equations,¹¹ thus

$$C_3 = \frac{C_{gk}C_{pk} + C_{pk}C_{gp} + C_{gk}C_{gp}}{C_{pk}} \tag{13-17}$$

In most cases the reactance of C_3 will be greater than the reactance of the lead inductance, L , at the operating frequency and the series resonant condition is obtained by adding an additional variable neutralizing condenser, C_N , in series with the grid lead as shown at c in the diagram. C_N breaks the d-c grid return to ground which is re-established by introducing R_g as shown. The use of shunt-plate feed allows the lower end of the plate tank to be grounded, a distinct advantage where the tank takes the form of a resonant section of coaxial line. It should be noticed that the neutralizing circuit just described is good for narrow-band R-F applications only. In broad-band operation, the resonant condition of L , C_3 , and C_N cannot be maintained over the entire pass band and oscillation may occur.

The operation of the neutralization circuit just described may also be explained on the basis of filter theory: The network of the actual interelectrode capacitances, L , and C_N may be drawn as a bridged-T structure. On this basis it may be seen that at null the current through the lower T is advanced by 270° , while that through the bridging capacitor, C_{pk} , is advanced by only 90° ; hence cancellation occurs and no energy is fed from the plate to the grid.

¹¹ W. L. Everitt, *op. cit.*, chap. 2.

We have stated that a tube operating in the grounded-grid connection can deliver a greater power output than it can in the more conventional connections. To demonstrate this we shall first consider that the circuit of Fig. 13-7 is so biased that it is operating in Class A. By this device we may utilize the equivalent plate circuit theorem. The results will then be extended to include Class C operation.

At the operating frequency, the plate tank circuit appears as a pure resistive load, R_L , and the equivalent plate circuit of the amplifier is that shown in Fig. 13-9. By applying the superposition theorem directly to the circuit we may write

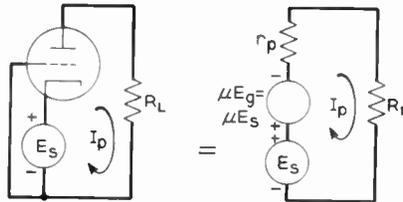


Fig. 13-9. Equivalent plate circuit of the Class A grounded-grid amplifier.

$$I_p = \frac{E_s}{r_p + R_L} + \frac{\mu E_s}{r_p + R_L} \tag{13-18}$$

and the power output will be

$$P_o = I_p^2 R_L = (1 + \mu)^2 \left[\frac{E_s^2}{(r_p + R_L)^2} R_L \right] \tag{13-19}$$

It will be observed that the bracketed term is precisely the power output of a conventional grounded cathode stage; hence the grounded grid circuit provides a power gain of $(1 + \mu)^2$ over the conventional circuit in Class A operation.

The mechanism by which this increase in power is delivered is revealed in eq. (13-18). Notice that two components of current are present in the plate circuit. The first, $E_s/(r_p + R_L)$, may be termed the “conductive” component, and it exists because the driving source E_s , being in the cathode return, is physically in series with the plate circuit; the driving source delivers power to the output circuit by conduction. The second term, $\mu E_s/(r_p + R_L)$, may be termed the “transfer” or “amplified” component, which is due to the equivalent generator, μE_o , acting in the plate circuit. Notice, then, that the increase in power output is furnished by the driving source rather than by the circuit itself. In effect, the circuit behaves like a series booster. So much for Class A operation.

In Class C operation, where the tube is biased beyond cutoff, the

general operation of the grounded-grid circuit remains the same but the analysis given above no longer applies. Since the driving source is in series with the load, it still delivers power to the output by conduction; but since the plate current is no longer sinusoidal, any equations based on the equivalent plate circuit are invalid. The power equations for Class *C* operation may best be derived by comparing the grounded grid circuit to its more conventional counterpart, the grounded-cathode Class *C* amplifier. We shall assume that in each case operation is to the diode line and that the same tube currents and voltage are present. The assumptions simplify the comparison of the two circuits. As a further simplification, the degenerative voltage developed across L and C_X will be neglected. The two circuits and their voltage diagrams are illustrated in Fig. 13-10.

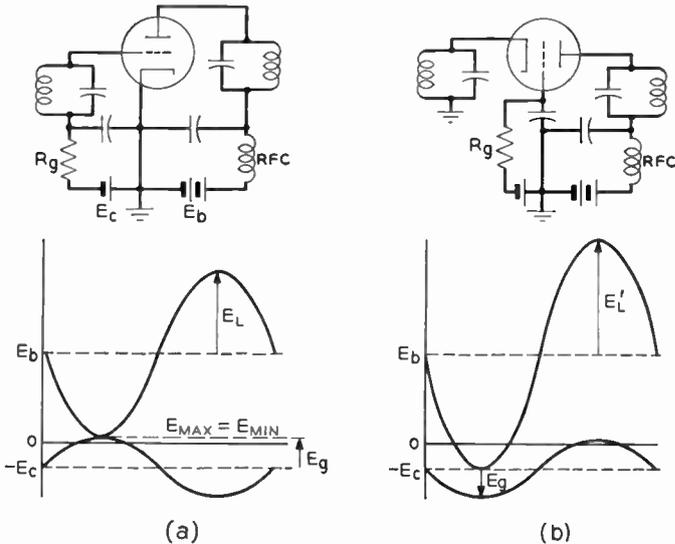


Fig. 13-10. Comparison of grounded-cathode and grounded-grid Class *C* amplifiers. (a) Grounded cathode. (b) Grounded grid. (After Starner. Courtesy of *Broadcast News*.)

Consider first the grounded-cathode Class *C* amplifier. Since we have assumed operation to the diode line (*i.e.*, the maximum grid voltage is equal to the minimum plate voltage, each being measured relative to ground), we may write

$$\begin{aligned} E_{\max} &= E_{\min} \\ &= E_g - E_c \end{aligned} \quad (13-20)$$

and the peak voltage across the plate load will be

$$E_L = E_b - E_{\min} \quad (13-21)$$

Let I_{p1} = peak value of the fundamental component of the plate current

Then the average power output of the stage will be

$$P_o = \frac{E_L I_{p1}}{2} \quad (13-22)$$

Next consider the grounded-grid stage. We shall use primes to indicate the counterparts of the quantities defined above. Once again the diode line condition is obtained when the maximum grid voltage is equal to the minimum plate voltage measured relative to ground. Since the grid voltage is fixed at $-E_c$, the bias voltage, we may express the condition as

$$E'_{\min} = -E_c \quad (13-23)$$

and the peak voltage developed across the load is

$$E_L' = E_b + E_c \quad (13-24)$$

From eqs. (13-20) and (13-21)

$$E_c = -E_b + (E_L + E_g) \quad (13-25)$$

and, substituting into (13-24), we get

$$E_L' = E_L + E_g \quad (13-26)$$

Equation (13-26) may be verified by comparing the voltage diagrams of the two circuits.

We have specified that the same current must flow in both stages; hence we must raise the load resistance in the grounded grid circuit to compensate for the increased load voltage. Thus, to equalize the currents, we set

$$R_L' = R_L \frac{E_L'}{E_L} = R_L \left(\frac{E_L + E_g}{E_L} \right) = R_L \left(1 + \frac{E_g}{E_L} \right) \quad (13-27)$$

and the power output delivered by the second circuit will be

$$P_o' = \frac{E_L' I_{p1}}{2} = \frac{(E_L + E_g)}{2} I_{p1} = \frac{E_L I_{p1}}{2} \left(1 + \frac{E_g}{E_L} \right) \quad (13-28)$$

$$= P_o \left(1 + \frac{E_g}{E_L} \right) \quad (13-29)$$

This confirms our results for Class A operation, namely, that the grounded grid circuit delivers a greater output power than the conventional stage does. By extension of the Class A treatment we know that the additional power is furnished by the driver; hence the driving power is

$$P_o' = P_o + P_o \frac{E_g}{E_L} \quad (13-30)$$

By definition the power gain of the stage is the ratio of output to driving power. Thus we have for the grounded-grid connection

$$\text{power gain} = \frac{P_o'}{P_o} = \frac{P_o \left(1 + \frac{E_g}{E_L}\right)}{P_o + P_o \frac{E_g}{E_L}} \quad (13-31)$$

The last equation indicates that the power gain may be controlled by varying the ratio E_g/E_L . To illustrate this, let the bias voltage be changed and E_g and E_L readjusted to maintain operation to the diode line. For example, if E_c is raised V volts, E_L and E_g must each be increased V volts. Since E_L is much larger than E_g , the percentage change in E_L will be much less than the percentage change in E_g ; hence a change in bias alters the E_g/E_L ratio and the power gain may be controlled by the bias.

In practice power gains in the order of 5 to 7 may be obtained. It is convenient to utilize the booster concept to operate with a lower power gain and let the driver furnish more of the output power. With a power gain of three, approximately one-quarter of the output is furnished by the driver stage and the tube size may be reduced.

We have stated that the ring seal type of tube is admirably suited to the grounded grid circuit, where resonant coaxial lines are employed as tank circuits. The reason for this may be seen from Figs. 13-11 and 13-12. Excepting the filament connections, the tube exhibits radial symmetry and may be fitted easily into a coaxial chamber.¹²

While the discussion above has referred specifically to triodes, the use of multielement tubes in the television transmitter is becoming more common.

¹² See S. Frankel, J. J. Glauber, and J. P. Wallenstein, "A Medium-Power Triode for 600 Megacycles." *Proc. IRE*, **34**, 12 (December 1946).

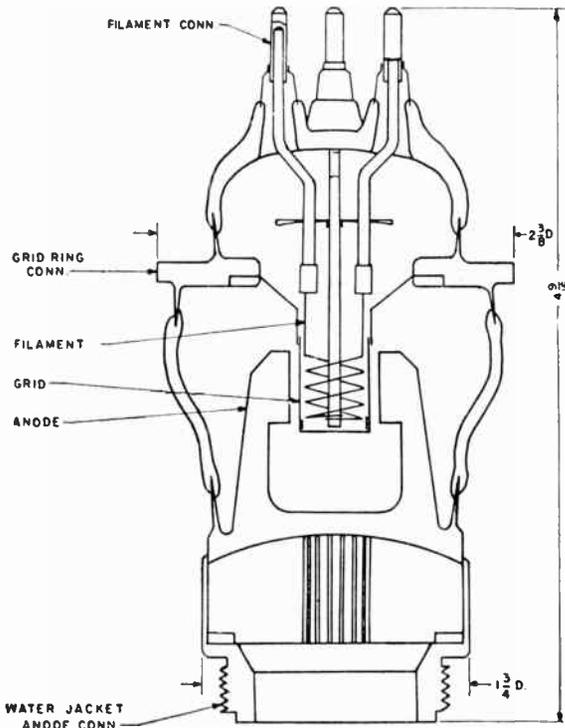


Fig. 13-11. Cross section of a typical ring-seal triode, the 6C22. (Courtesy of Federal Telephone and Radio Corporation.)

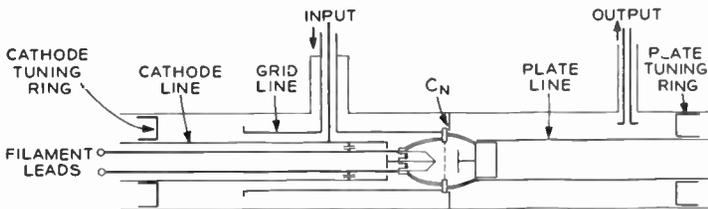


Fig. 13-12. Cross section of a grounded-grid coaxial amplifier employing a ring-seal triode. Notice how the neutralizing condenser is built into the coaxial structure. (After Frankel, Glauber, and Wallenstein. Courtesy of *Proc. IRE.*)

THE VIDEO SECTION

A second function which must be performed by the visual transmitter is the acceptance of the composite video signal from the pickup facilities and the raising of its power level to the value required at the modulator. That portion of the transmitter which performs this function may be termed the video section. The block diagram of a typical transmitter video section is shown in Fig. 13-13.

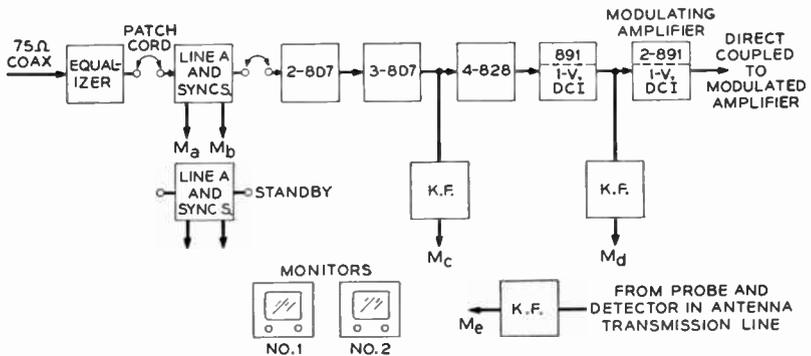


Fig. 13-13. Block diagram of the video section of a television transmitter.

13-8. Block Diagram

Reading from left to right in the diagram the first unit is the line equalizer. This is a passive network designed to equalize the amplitude and delay characteristics of the coaxial cable, which brings in the composite video signal from the studio facilities. The second unit, the line amplifier and sync stretcher, serves to raise the voltage level of the signal so that it may drive the chain of power amplifiers which follow. In certain designs the sync portion of the signal is subjected to additional amplification, and hum is removed by a special form of clamping circuit. This is explained in the following paragraphs. In conformance with commercial practice a stand-by line amplifier is provided to decrease the likelihood of shutdown in case of failure. Either unit may be connected into the video chain by means of patch cords. The five amplifier stages serve to develop the necessary modulating power for the final grid-modulated stage.

Two monitors are provided so that the signal may be checked at any one of a number of points throughout the transmitter. The several test points, labeled M_a through M_e in the diagram, are isolated by means of cathode followers. Switching means are normally provided so that any of the test points may be checked. Notice that the composite video signal is sampled; hence the supersync may be separated from the picture components and used to drive sweep generators in the monitors. Thus the actual picture program may be viewed. Under certain circumstances a conventional oscilloscope is provided and the composite signal is fed to its vertical deflection plates. Where this type of monitoring is used the line-by-line or field-by-field *wave form* of the signal may be checked, depending on whether the horizontal saw-tooth sweep is run at line (15.75 kilocycles) or field (60 cycles) frequency. Proper use of the monitors greatly reduces the time required to locate troubles in the entire transmitting system and aids in maintaining optimum operating conditions. We now consider some of the details of the units which have been mentioned.

13-9. Line Amplifier

The clamping circuit in the line amplifier is different from the clampers discussed in earlier chapters. Its chief purpose is to remove low-frequency variations in the incoming signal which are the result of hum and pickup. The basic circuit is shown in Fig. 13-14.^{13,14} It will be seen that the clamping action is controlled by separate keying signals rather than by the clamped signal itself and hence may be called "keyed" or "synchronized" clamping. We may consider the operation of the circuit from two points of view; first, how the clamping diodes V_3 and V_4 behave, and second, how their behavior affects the signal appearing at point A , which is connected to the grid of V_2 . A locally generated square pulse or key, which occurs during the duration of the sync "back porch" (see Fig. 11-11) is applied to the grid of V_5 . Since R_4 and R_5 in the plate circuit of V_5 are equal, two pulses of equal magnitude but opposite sense are applied to the bridge circuit shown. These two pulses appear in series across the

¹³ Instructions, TA-5A Stabilizing Amplifier, R.C.A., Engineering Products Department, 1946.

¹⁴ J. L. Schultz, "Television Stabilizing Amplifier." *Radio-Electronic Engineering Edition of Radio and Television News*, 12, 5 (May 1949).

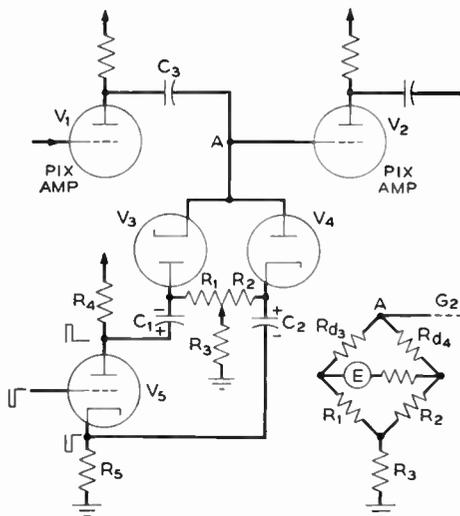


Fig. 13-14. A synchronized clamping circuit. (Courtesy of Radio Corporation of America.)

diodes which conduct, causing the condensers C_1 and C_2 to charge with the polarities indicated in the diagram. During the relatively long interval between successive pulses the condensers discharge slowly through R_1 , R_2 , and V_5 . During this same interval the diodes are cut off because of the reverse polarity appearing across them. When the next key is applied, the diodes conduct again, re-establishing the charge lost by the condensers between successive pulses.

With this action of the diodes in mind we may now consider their effect on the voltage at point A , which is connected to the grid of V_2 , the second picture amplifier. During the duration of the key, the diodes conduct because of the action of the two pulses delivered by V_5 . Since the pulses are of equal magnitude but opposite sense, they cancel between A and ground and hence have no direct effect on the voltage at A . For their duration, however, a conduction path is established between A and the tap between R_1 and R_2 ; hence A is clamped to a fixed voltage for their duration. Notice that the actual value of the clamp level may be adjusted by the position of the tap.

Between successive pulses, when the diodes are cut off by the voltages on C_1 and C_2 , they appear as an open circuit between A and ground. In other words, between pulses the grid return resistor of V_2

is infinite, providing an infinite time constant of the coupling circuit, which, in turn, effectively gives a flat response down to zero frequency, even though the coupling condenser, C_3 , may be of the order of 500 $\mu\mu\text{f}$. Thus the clamping circuit provides an excellent low-frequency response.

More important, the clamping circuit stabilizes the composite video signal. Since the grid of V_2 is clamped to a fixed voltage once each line interval, the maximum effect that a spurious low-frequency component can cause at the grid is that which occurs during one line interval. This clamping effect is illustrated in Fig. 13-15. The

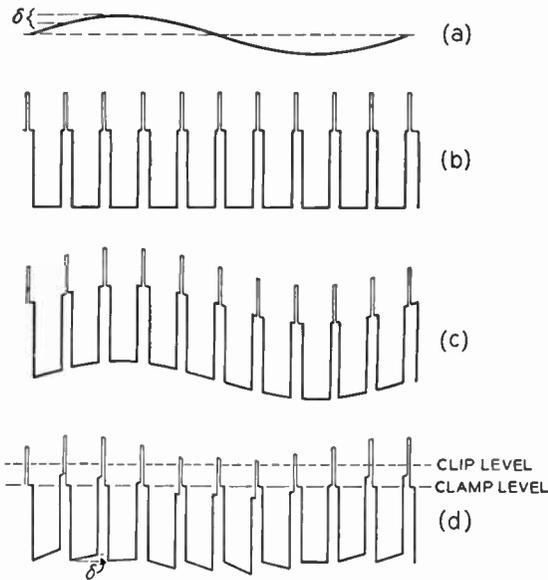


Fig. 13-15. The effect of hum is reduced by clamping. Since the signal is clamped at a fixed level once each line, the maximum hum that can appear in one line is δ . The scale of the hum voltage is greatly exaggerated to illustrate the effect of clamping.

magnitude of the hum component is greatly exaggerated in the diagram for the purpose of illustrating the action. Notice in the final clamped voltage that the error in each line is equal to the *change* in hum voltage during that line. This is represented in the diagram by the quantity δ .

It may be seen from the diagram that the height of the various sync pulses after clamping is not constant. This defect may be remedied by applying a sync-negative version of the clamped signal to a biased amplifier stage so that all the pulses are clipped to a common level, indicated at *d* in the figure. Notice that the sync pulses in the final wave are of smaller relative amplitude than in the wave shown at *c*. In order to preserve the proper sync amplitude, sync stretching may be employed before clamping takes place.

The basic circuit of the sync stretcher and its associated components is shown in Fig. 13-16. V_1 and V_2 are the same tubes shown

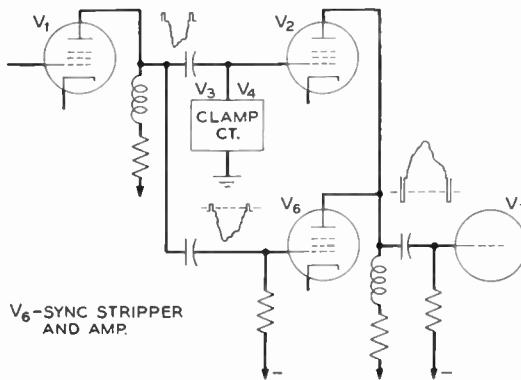


Fig. 13-16. The sync stretcher shown in relation to the clamping circuit.

in the preceding figure. The sync-positive composite video signal is fed to the biased sync stripper V_6 . With proper adjustment of the bias, the picture components of the signal lie below cutoff and only the sync is amplified and inverted. Simultaneously, the composite video is clamped on the grid of V_2 in the manner previously described. Having a common plate load, V_2 and V_6 serve as a mixer and add the amplified sync pulses to the composite video as indicated; the sync has been stretched. The problem of phase shift is eliminated because both signals go through a single tube and add in a common impedance. The relative height of the sync in the combined signal is adjusted by controlling the clip level on the grid of V_7 . Thus the problem of insufficient sync height, which was pointed out in connection with Fig. 13-15*d*, has been corrected.

It should be realized that the basic circuit shows the principles involved in the line amplifier. In commercial versions of the equipment, additional clamping is employed at the sync stripper input and other critical points in order that even greater stabilization of the signal is obtained. All stages in the unit are compensated out to eight or even 10 megacycles in order to ensure that a high-quality signal is delivered to the transmitter proper.

13-10. The Video Amplifiers

Figure 13-13 shows five stages of video amplification in the video section of the visual transmitter. These serve to develop the necessary modulating power for the final modulated stage. We have already discussed the problem of compensating amplifiers for broadband operation; our remarks here will be confined chiefly to the problem of developing the necessary power or voltage in each stage.

All of the amplifiers in the video section are operated in Class *A* to minimize distortion and, hence, in theory require no driving power. Each stage, then, with the exception of the modulator serves to raise the level of the driving voltage for the succeeding tube. Since the voltage levels are high, approaching 1 kilovolt for the last stage, large tubes of the power variety are necessary. Generally, tubes of this sort have comparatively low values of μ and the five stages shown in the diagram are required. Another factor points to the use of power tubes. We shall see that relatively low values of plate load resistance must be used. Then as the voltage level is raised, E^2/R increases, requiring a tube of large power-handling ability.

It is desirable to use small tubes of the same type wherever possible because of the low initial cost and the resulting reduction in spare tube stock required. Thus, in the first two stages of Fig. 13-13 two or three 807's are operated in parallel in order that sufficient power may be developed with the comparatively small tubes. We shall now consider the effect of shunt capacitance on output voltage when two tubes are operated in parallel. To simplify the discussion we shall assume that simple shunt compensation is used. Let I be the output current per tube and R_1 the load resistance for a single tube operating alone. Its output voltage will be

$$E_o = IR_1 \quad \text{ONE TUBE} \quad (13-32)$$

where R_1 is determined by C_s , the total shunt capacitance across the plate load.

If, now, two identical tubes are placed in parallel the total shunt capacitance is increased by C_o , the output capacitance of the second tube. If the bandwidth is to remain unchanged, a new value of R_1 will be required, namely,

$$R_1' = R_1 \frac{C_s}{C_s + C_o} \quad \text{TWO TUBES IN PARALLEL} \quad (13-33)$$

and the new output voltage will be

$$E_o' = 2IR_1' = 2IR_1 \frac{C_s}{C_s + C_o} = 2E_o \frac{C_s}{C_s + C_o} \quad (13-34)$$

Notice that if C_o is the larger part of C_s , $C_s \approx C_o$ and there is no advantage gained in paralleling the tubes; $E_o' \approx E_o$. If, on the other hand, the strays and input capacitance are large in comparison to C_o , parallel operation is practicable and will give a larger output voltage than may be obtained with a single tube.

Naturally the use of compensation is mandatory. We have seen in Chapter 7 that series-shunt compensation permits a higher value of load resistance and voltage amplification than shunt compensation; hence, series-shunt compensation is to be preferred. It is important to note that the various design equations which were derived for video amplifier compensation must be modified when applied to power amplifiers. In our earlier work which centered on voltage amplifiers, it was assumed that the plate resistance, r_p , was very large as com-

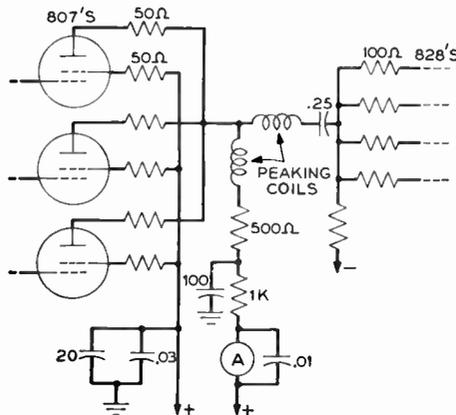


Fig. 13-17. A typical coupling network for the intermediate power level stages of the video amplifier.

pared to R_1 and had negligible shunting effect on the plate load. In triode power amplifiers r_p may be of the same order as R_1 , in which event R_1 must be replaced by R_1' , the parallel combination of the two resistances

$$\frac{1}{R_1'} = \frac{1}{R_1} + \frac{1}{r_p} \tag{13-35}$$

in all of the design equations. A typical coupling network for use between stages 2 and 3 of Fig. 13-13 is shown in Fig. 13-17.

13-11. The Constant-resistance Network

As the power level is raised along the video amplifier chain the power dissipation requirements of the several plate load resistors are also increased. It is unfortunate that shunt capacitance is inevitably associated with resistors and its value increases with the size of the resistor.¹⁵ The constant-resistance network may be employed to permit the use of smaller resistors of lower power capacity for a given plate load requirement. The two networks shown in Fig. 13-18*b* and *c* are alternate forms of a constant-resistance network and are equivalent in their operation.¹⁶

Consider the circuit of Fig. 13-18*a*. We desire that the input impedance shall be a pure resistance and independent of frequency. What are the requirements on X_1 and X_2 to meet this condition? The input admittance of the network is

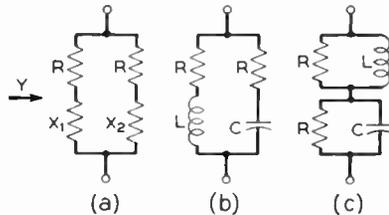


Fig. 13-18. Constant-resistance networks.

$$\begin{aligned} Y &= \frac{1}{R + jX_1} + \frac{1}{R + jX_2} = \frac{R - jX_1}{R^2 + X_1^2} + \frac{R - jX_2}{R^2 + X_2^2} \\ &= R \left(\frac{1}{R^2 + X_1^2} + \frac{1}{R^2 + X_2^2} \right) - j \left(\frac{X_1}{R^2 + X_1^2} + \frac{X_2}{R^2 + X_2^2} \right) \end{aligned} \tag{13-36}$$

¹⁵ A convenient rule of the thumb is that in small carbon resistors of the receiver type, the shunt capacitance, in micromicrofarads, is numerically equal to the wattage rating of the resistor.

¹⁶ A third form of the network has been shown in Fig. 12-13*a*. The method of operation is quite similar in all three forms.

If the input admittance is to be real, the imaginary term of (13-36) must be zero; thus,

$$\frac{X_1}{R^2 + X_1^2} = -\frac{X_2}{R^2 + X_2^2} \quad (13-37)$$

Clearing and combining terms, we finally have

$$R^2 = -X_1 X_2 \quad (13-38)$$

Equation (13-38) may be satisfied if

$$X_1 = +j\omega L$$

and

$$X_2 = -\frac{j}{\omega C}$$

} (13-39)

Substituting these values into (13-38) we get for the design condition of the constant-resistance network:

$$R = \sqrt{\frac{L}{C}} \quad \text{CONSTANT-}R \text{ NETWORK} \quad (13-40)$$

Substituting (13-38) into (13-36) we get for the input admittance

$$\begin{aligned} Y &= R \left[\frac{R^2 + X_2^2 + R^2 + X_1^2}{(R^2 + X_1^2)(R^2 + X_2^2)} \right] + j0 \\ &= \frac{R}{R^2} \left[\frac{2R^2 + (X_1^2 + X_2^2)}{2R^2 + (X_1^2 + X_2^2)} \right] = \frac{1}{R} \end{aligned} \quad (13-41)$$

and we see that the impedance of the network is R at all frequencies. It may also be shown that the network behaves like a pure resistance for transients. A similar analysis shows that the same design condition and input impedance obtain for the series form of the network, which is illustrated at c in the figure. Notice that the shunt capacitance across each resistor has been neglected; hence the results are approximate.

We may show quite readily that the network is of the crossover type. At the low frequencies C has a high reactance and most of the power is dissipated in the left-hand resistor. At the high frequencies, L offers a high impedance and most of the power is delivered to R in series with C . At the crossover frequency, f_c , the power is divided equally. Thus at f_c the currents in the two branches are equal and

$$\left| R + j\omega_c L \right| = \left| R - \frac{j}{\omega_c C} \right|$$

or
$$\omega_c I_c = \frac{1}{\omega_c C}$$

whence
$$f_c = \frac{1}{2\pi\sqrt{LC}} \quad (13-42)$$

Notice, then, that if we apply a signal covering a 5-megacycle bandwidth to the network, the power will be shared by the two resistors. Each resistor may be smaller than a single resistor of the same value of R . By this artifice the effects of shunt capacitance across the load resistor proper may be reduced.

It is of interest to note that for a typical television signal the power will not divide *equally* between the two resistors in the network because the distribution of energy is not constant over the entire video band.¹⁷ In practice, however, the two network resistors are chosen with equal wattage ratings.

In practical design of the constant-resistance network the shunt capacitance associated with each resistor is neglected. The design values of L and C are then pruned to give proper operation. It should be noticed that the series-type network of Fig. 13-18c has a slight advantage in that the shunt capacitance associated with the lower resistor is compensated by decreasing C by the value of the shunt capacitance. This technique of "washing out" an undesirable capacitance by making it part of the constant-resistance network will be used again in a later section.

Notice that the network does not compensate for the input and output capacitances of the tubes on either side of the network; hence some form of compensation is still required as indicated in Fig. 13-17.

13-12. The D-C Component

We have previously shown in Chapter 7 (see Fig. 7-28) that the grid swing requirements for a video amplifier may be reduced by retaining the d-c component of the signal. This component may be restored at any grid with the help of a d-c insert or clamping circuit. The use of such a circuit is indicated in the fourth video stage in Fig. 13-13, which stage employs an 891. Once inserted, the d-c component may be retained by clamping in each successive stage or

¹⁷ See G. H. Brown, "A Vestigial Side-Band Filter for Use with a Television Transmitter." *RCA Review*, V, 3 (January 1941). Brown shows the energy distribution in the sidebands of a modulated wave.

by the use of direct coupling of some sort. Clamping cannot be used in the final grid-modulated amplifier; hence direct coupling is the rule between the modulating and the modulated amplifiers. We shall consider two forms of direct coupling which are applicable in the high-power video stages of the transmitter.

The first system to be described was used in the Columbia Broadcasting System color television transmitter built by the Federal Telephone and Radio Corporation.¹⁸ The basic circuit of a direct-coupled amplifier is shown in Fig. 13-19*a*. The coupling between the plate

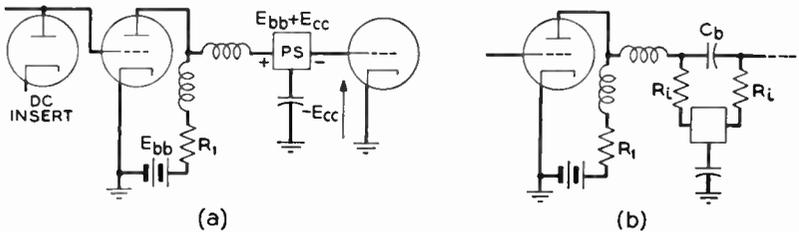


Fig. 13-19. A direct-coupled video interstage network. The d-c plate voltage is bucked out by a power supply isolated from ground. (Courtesy of Federal Telecommunication Laboratories, Inc.)

load and grid of the two tubes is through a well-regulated d-c power supply, which has a terminal voltage of $(E_{bb} + E_{cc})$. With the polarity shown this bucking supply leaves the grid with a voltage relative to ground of $-E_{cc}$, the bias.

The basic circuit has an inherent difficulty: the bucking supply must be ungrounded with the result that both of its terminals will have shunt capacitance to ground. This shunt capacitance adds to C_s across the plate load and impairs the high-frequency response of the system. To remedy this difficulty the two isolating resistors, R_i , may be used as shown at *b* in the diagram. These resistors are chosen to have a resistance much greater than R_1 so that the plate load is unaffected by their presence, and they isolate the plate and grid from the bucking-supply capacitance. Since no grid current flows, the d-c signal component is unaffected by the resistors. As far as the high

¹⁸ *Federal's High Power Transmitter built for the Columbia Broadcasting System for Color or Fine Line Television*. F. T. and R. Corp. Material used through the courtesy of the Federal Telecommunication Laboratories, Inc.

frequencies are concerned, however, a return path is completed through the input capacitance of the second stage and the IR drop in the isolating resistors becomes intolerable. To eliminate this condition a condenser C_b is bridged across the d-c network to provide a low-impedance path for the high-frequency components. Thus in the upper frequency ranges the coupling is through C_b . The d-c and low-frequency components, which represent changes in background level, pass through the bucking pack; since no d-c current flows, both terminals of the pack rise and fall by the same amount as the low-frequency changes in plate potential.

The second system of direct coupling to be described is employed in the R.C.A. TT-5A transmitter. It differs from the first system in the manner of handling the shunt capacitance which exists between the bucking-supply terminals and ground. We shall develop the circuit step by step, the first of which is shown at *a* in Fig. 13-20.

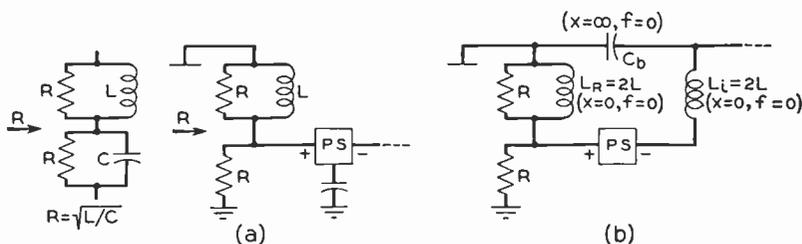


Fig. 13-20. The video interstage coupling network used in the R.C.A. TT-5A video transmitter.

Observe that the design concept is to wash out the shunt capacitance on the positive side of the bucking supply by making it part of a constant-resistance network. In order to isolate the grid from the capacitance between the negative terminal of the bucking supply and ground, an isolating inductance, L_i , is used as shown in Fig. 13-20*b*. In the actual design L_i must be chosen so that the constant- R network operates properly. Notice that at the higher video frequencies the internal impedance of the bucking supply is practically zero so that L_R and L_i are in parallel; hence their parallel combination must have the value required by the design equation of the constant- R network, or

$$\frac{1}{L_R} + \frac{1}{L_i} = \frac{1}{L} \tag{13-43}$$

Clearly (13-43) may be satisfied by setting

$$L_i = L_R = 2L \quad (13-44)$$

We may summarize the operation of the coupling network as follows. For the d-c component, C_b presents an open circuit so that L_i does not shunt L_R . L_R has zero reactance and shorts out the upper resistance; the entire plate load is furnished by the lower resistance. Coupling of the d-c component is through the bucking supply, both sides of which rise and fall by the same amount as the change in d-c plate voltage. For the high video frequencies C_b and the bucking supply have negligible impedance and L_R and L_i combine in shunt to give L ; the plate load is the entire constant-resistance network. Coupling of the high-frequency components is through the bucking condenser, C_b , which prevents them from passing through the high series impedance of $L_R + L_i = 4L$.

The constant-resistance concept has been extended in an extremely ingenious manner in the TT-5A transmitter; it is used as a filter network for the plate power supply. The development of the filter system is illustrated in Fig. 13-21. The original plate load is shown at *a*. At *b* the lower resistance is replaced by a second constant-resistance network and the terminal impedance of the entire circuit remains unchanged. This procedure may be used again and again until the required amount of filtering is obtained for the d-c power supply, which replaces the final resistance in the series as shown at *c*.

It may be seen from the diagram that the plate power supply views the entire network as a condenser input filter, while the video circuit sees a constant impedance, R . Since better regulation of the d-c voltage is obtained if a choke input filter is used, it is desirable to modify the network slightly. This may be accomplished by using the parallel form of constant-resistance network shown at *d* in the figure. Thus if we replace the power supply by the parallel network and move the supply one section to the right, we have the network shown at *e* and the desired condition is obtained. The constant-resistance network can therefore be used to provide the proper plate load and power supply filtering as well. The basic diagram of the complete network that may be used to couple the modulating to the modulated stage is illustrated in Fig. 13-21*f*. In commercial practice more sections are provided to ensure adequate filtering. The

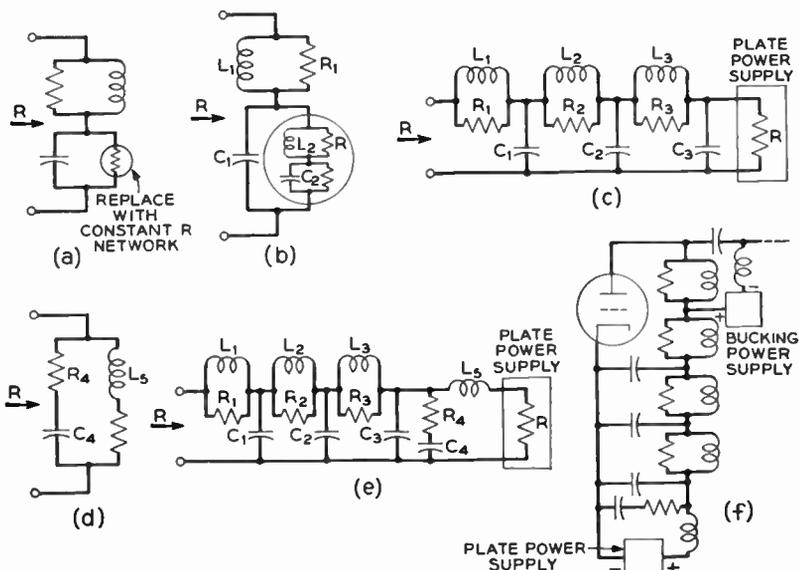


Fig. 13-21. Development of the R.C.A. interstage coupling network. (a) The basic constant-resistance network. (b) The lower resistor is replaced by a constant-resistance network. (c) One more stage has been added and the final resistor replaced by the internal resistance of the power supply. (d) The parallel form of the constant- R network may be used to provide a choke input for the power supply. (e) The network shown at (d) is used as a terminal section to provide a choke input filter for the plate power supply. (f) The complete coupling network. Additional filter sections may be added for better filtering.

initial ripple at the input of the filter is held to a low value by using polyphase rather than single-phase rectifiers.

THE MODULATING SECTION

A third function required of the visual transmitter is the modulation of the R-F carrier in such a manner that the output wave is of the negative transmission type. The section of the transmitter which performs this function may be termed the modulating section; it comprises the modulating and modulated amplifiers. Notice that we are borrowing the former from the video section, for the two stages may best be discussed as a unit.

A number of features of the modulating section have already been discussed. We have seen that grid modulation must be used in the high-level modulation transmitter and that the necessary modulating power across a low load impedance may be obtained by operating two or more tubes in parallel. We have also discussed means of providing the direct coupling between the modulating and modulated stages. In the present section we shall investigate certain other factors of concern in the modulating section.

13-13. The Modulated Stage

In general, the size of tube required for a given output power in the final or modulated stage may be reduced by using two tubes in a push-pull connection. In contrast to the video stages the modulated stage may take advantage of this connection because the modulated output wave is symmetrical about the zero axis. Since linearity is required between the R-F output and the video modulating voltages,

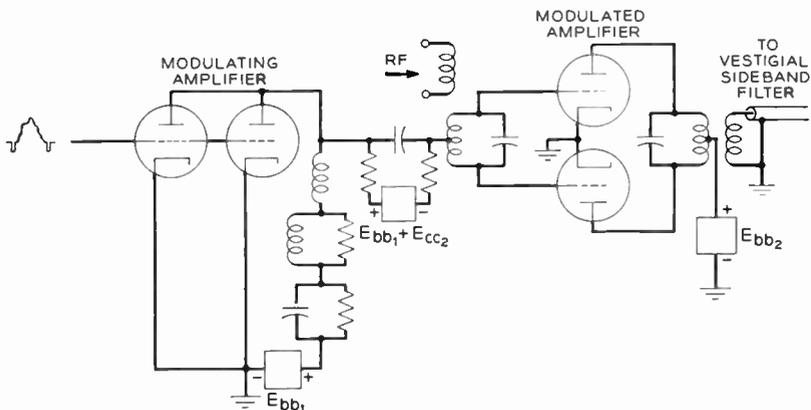


Fig. 13-22. A representative modulating section.

the final stage will operate in Class *B* with the bias adjusted at the cut-off value. The diagram of a representative final stage is shown in Fig. 13-22. Notice that direct coupling is employed and that the instantaneous effective bias on the final amplifier is $-E_{cc2}$ (the d-c bias) plus the sync-positive video signal. During the R-F peaks grid current flows in the final stage; hence the interstage coupling re-

sistance must be chosen low enough so that the flow of grid current has a minimum effect on that resistance. A typical modulation characteristic for the final stage is illustrated in Fig. 13-23. Notice that the flow of grid current on the modulation peaks introduces non-linearity. This may be minimized by utilizing less modulating signal so that operation is confined to the left of the vertical line on the diagram. This procedure inevitably lowers the output and the efficiency of the stage. A compromise may be struck by noting that for negative transmission the modulation peaks are produced by the synchronizing pulses; hence, we may compensate for curvature in the characteristic by stretching the sync amplitude in the video section so that it is too great in the modulating signal. Then it is compressed again to its proper value by the modulation characteristic. We have already discussed the sync stretcher, which may be adjusted properly by viewing the wave form of the demodulated output of the final stage. Great care must be taken to ensure that the sync in the radiated signal is of the proper amplitude, for the maximum tolerance allowed by the F.C.C. is $\pm 2.5\%$.

Notice that some slight compression of the whites will occur because of the curvature in the lower portion of the modulation characteristic. This is not serious, though it will result in a slight unbalance in the contrast of the final reproduced image.

13-14. The Plate Load

The design of the plate load impedance of the final stage is extremely important. Assuming that a modulating signal range of 4.5 megacycles is used, the plate load would be required to have a pass band twice as wide, or 9 megacycles, for double-sideband transmission. For television signals a portion of the lower sideband must be suppressed so that it is desirable to off-carrier-tune the tank circuit. The lower sideband rejection obtained in this manner can ease the requirements on the vestigial-sideband filter. A satisfactory procedure is to design the tank circuit for a half-power bandwidth of 6 megacycles. Reference to Fig. 12-8a shows that the limits of this half-power bandwidth should be located as follows.

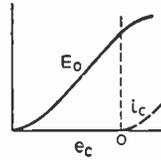


Fig. 13-23. Modulation characteristic of the grid-modulated Class B stage.

$$\begin{aligned}
 f_2 &= \text{upper half-power frequency} \\
 &= f_c + 4.5 \text{ megacycles} \\
 f_1 &= \text{lower half-power frequency} \\
 &= f_c - 1.5 \text{ megacycles}
 \end{aligned}
 \tag{13-45}$$

Now, even in the lowest channel, No. 2, the fractional half-power bandwidth is roughly 0.1, or the required Q is 10; hence we may assume that the pass characteristic of the tank circuit displays arithmetic symmetry, and the required center, or resonant, frequency may be taken to be the arithmetic mean of the half-power frequencies.

$$\begin{aligned}
 f_0 &= \text{resonant frequency of the tank circuit} \\
 &= \frac{f_1 + f_2}{2} = \frac{f_c - 1.5 + f_c + 4.5}{2} \\
 &= f_c + 1.5 \text{ megacycles}
 \end{aligned}
 \tag{13-46}$$

It follows directly that

$$\begin{aligned}
 f_1 &= f_0 - 3 \text{ megacycles} \\
 \text{and} \quad f_2 &= f_0 + 3 \text{ megacycles}
 \end{aligned}
 \tag{13-47}$$

The power output and linearity of the modulation characteristic are dependent upon the load on the modulated tubes. Let R_L be the required value of this load for a single tube. For the push-pull connection, the total resistance required from plate to plate is four times this value and we shall designate this total resistance as R_{pp} . Then

$$R_L = \frac{R_{pp}}{4}
 \tag{13-48}$$

With these data given we can calculate the constants for the plate tank circuit. At resonance, f_0 , it must present a real impedance of magnitude R_{pp} , and its half-power bandwidth must be 6 megacycles. The diagram of a lumped-constant resonant tank circuit is shown in Fig. 13-24a. Consider, first, the value of C required. We have seen

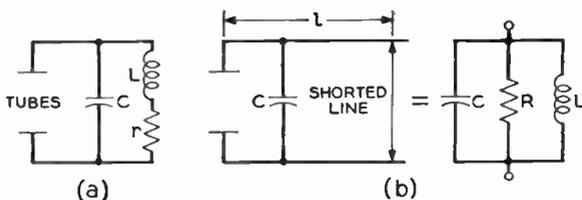


Fig. 13-24. Plate tank of the final stage. (a) Lumped constants. (b) Distributed constants.

that in a tuned amplifier the gain-bandwidth product is inversely proportional to the shunt C present. Thus, for a given half-power bandwidth, the maximum output will be obtained from the amplifier if C is at its minimum possible value. For this reason C should be the output capacity of the tube itself plus whatever stray capacitance is present. In the larger water-cooled tubes the water jacket contributes to the total shunt capacitance and its contribution must be taken into account. In typical commercial transmitters the total shunt capacitance may be held below $10 \mu\mu\text{f}$.

We next calculate the value of L that is necessary to meet the design conditions. By its definition, the circuit Q is

$$\left. \begin{aligned} Q &= \frac{f_0}{\Delta f} \\ \text{where } f_0 &= \text{resonant frequency} \\ \Delta f &= \text{half-power bandwidth} \\ &= f_2 - f_1 \end{aligned} \right\} \quad (13-49)$$

It may be shown that for all the assigned television channels the necessary value of Q is approximately 10 or greater. For example in channel No. 2 we have

By (13-46)

$$f_0 = f_c + 1.5 = 55.25 + 1.5 = 56.75 \text{ megacycles} \quad (13-50)$$

$$\text{and by (13-49)} \quad Q = \frac{56.75}{6} = 9.45 \quad (13-51)$$

In the higher channels $f_0 > 56.75$ megacycles and Δf remains constant at 6 megacycles; hence the value of Q for channel No. 2 is a minimum.

With a Q in the vicinity of 10 or more, the condition of antiresonance in the tuned circuit is given by

$$\omega_0 L = \frac{1}{\omega_0 C} \quad (13-52)$$

and the required value of L is

$$L = \frac{1}{\omega_0^2 C} \quad (13-53)$$

At antiresonance, the impedance of the network must be equal to R_{pp} and we may write

advantage, for it was demonstrated that a shorted line of length $\lambda/8$ exhibits an inductive reactance equal to that of a lumped inductance over a limited frequency range. It may be shown from our previous equations that for channel No. 7 $f_2/f_0 = 1.015$ and $f_1/f_0 = 0.983$; hence from Fig. 12-17 we see that a negligible error is introduced by using the line in place of the coil.

With l chosen as $\lambda_0/8$ we next must calculate the R_0 of the line. Thus, from eq. (12-52),

$$\frac{1}{\omega_0 C} = R_0 \tan \frac{\pi}{4} \quad (13-59)$$

$$R_0 = \frac{1}{\omega_0 C}$$

where C is the output capacitance of the final stage.

Since the output circuit of the push-pull amplifier is balanced to ground, it is more convenient to use the parallel wire rather than the coaxial form of transmission line. For the former, R_0 is given by¹⁹

$$R_0 = 276 \log \frac{b}{a} \quad (13-60)$$

where b is the spacing between centers of the conductors and a the radius of the conductors. The physical length of the short-circuited transmission line may be calculated quite readily. In general, the phase velocity along the line is roughly 2.5 per cent less than it is in free space; hence we may write

$$v_p = (0.975)(3 \times 10^{10}) \text{ cm/sec} \quad (13-61)$$

$$\text{or} \quad l = \frac{\lambda_0}{8} = \frac{(0.975)(3 \times 10^{10})}{8(2.54)10^6 f_{0m.c.}} = \frac{1440}{f_{0m.c.}} \text{ inches} \quad (13-62)$$

For channel No. 7 this represents a length of 8.15 inches. Caution should be exercised in using this figure, for it represents the total effective length of the plate line and must include the length of the connecting leads plus the length of the plate leads within the tubes themselves. In the higher channels this may bring about considerable difficulty because enough of the plate line must be external to the tubes so that coupling to the transmission line may be effected. This trouble may be overcome by extending the line length, l , to some odd

¹⁹ F. E. Terman, *op. cit.*, chap. 4.

multiple of $\lambda_0/8$. That this procedure does not affect the electrical properties of the line may be verified from the short-circuited line reactance curves of Fig. 12-15a.

It may be observed from eq. (13-59) that the required value of transmission line characteristic impedance, R_0 , increases as the resonant frequency, f_0 , is lowered; hence in the lower television channels a large ratio b/a is required for constructing the parallel wire transmission line and it may not be possible to fit the required dimensions to the plate connections. Furthermore, in these lower channels the between-conductor spacing may approach the physical length of the line, thereby rendering the R_0 equation invalid. Where these conditions exist, an alternate design procedure is necessary. Since the required value of R_0 may be reduced by changing the line length from the $\lambda/8$ value, a new and longer value of l may be chosen to lower R_0 to a practicable value. If this procedure is followed, care must be taken to ensure that the line presents the proper reactance within a few per cent over the entire pass band. As in the lumped circuit case, the necessary value of r may be reflected in from the transmission line.

TRANSMISSION LINE AND ANTENNA SECTION

The fourth function required of the television transmitter is that it deliver the vestigial-sideband signal of the R.A. type to the antenna system. Thus, in this last section we shall discuss the various components which perform this function.

13-15. The Coupling Loop and Balun

As a first step in delivering the output of the final modulated stage to the antennas we must couple the modulated output to the transmission line system. Mutual inductance coupling is used in order to isolate the transmission line system from the high d-c voltage on the plates of the final amplifier. A further problem arises in that the final output, which is balanced to ground, must be converted to an unbalanced signal because the coaxial form of transmission line is used throughout the entire filter and antenna feed system. One form of coupling network and Balun which has been used successfully is shown in Fig. 13-26. The left-hand member of the inverted U is a quarter of a wavelength long at f_0 and the impedance between X and

ground is infinite. For this reason the extended portion of the coax center conductor may be fastened at X for physical support and still remain ungrounded to a-c. Coupling between this extended center conductor and the final stage is provided by placing the entire U close to the plate line of the final stage, the actual distance between them being adjusted so that the proper value of resistance is reflected back into the tank circuit. The single-ended output appears across the coaxial line and is delivered directly to the vestigial-sideband filter.

Fig. 13-26. The Balun pickup loop. An actual loop of this type is shown in Fig. 13-25.

The sideband filter which is the second major component of the output section has been covered in Chapter 12. We next consider how its output, which is unbalanced to ground, may be converted to the balanced condition required by the antenna system. A simple Balun such as the one shown in Fig. 12-23 may not be used. Current practice employs a variation of the device in order to provide satisfactory operation over the full radiated bandwidth of approximately 6 megacycles. We shall first consider why the simple Balun falls short in operating over this relatively broad bandwidth. Its equivalent circuit is shown at a in Fig. 13-27. At f_0 the Balun sleeve has a length of $\lambda_0/4$ and behaves like an antiresonant circuit. At all other frequencies, however, the

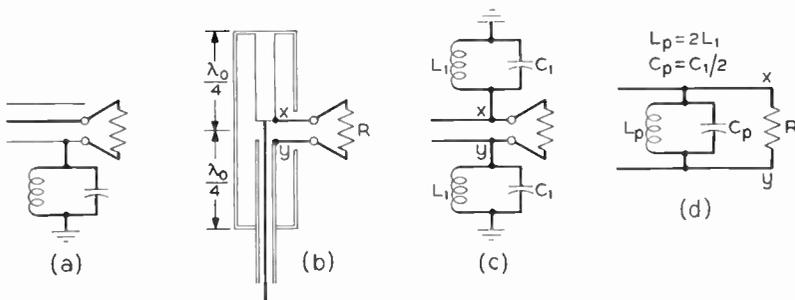


Fig. 13-27. Broadbanding the Balun. The two output leads from the Balun may be coaxial lines. (a) Equivalent lumped-constant circuit of the Balun shown in Fig. 12-23. (b) Cross section of the broadband Balun. (c) Equivalent circuit of (b). (d) Simplified equivalent circuit.

length is not $\lambda/4$, the impedance from y to ground is no longer infinite, and the condition of balance is upset.

This condition is remedied by the modified Balun shown in cross section at b . Notice that the center coax conductor is extended and anchored to an additional quarter-wave stub and that the outer sleeve has been extended upward to enclose this stub. By this device lines x and y are both connected to ground through equal impedances, and balance is maintained. The stub lengths are each $\lambda_0/4$ and present infinite impedance to ground at f_0 . At other frequencies the impedance departs from its infinite value but balance is maintained, for both the upper and lower stubs exhibit the same frequency characteristic.²⁰

13-16. Diplexer

The signal is push-pull at the Balun output and ready to feed the antenna system. Before discussing the transmitting antenna, however, it is well to consider the diplexing unit, which is indicated in the block diagram of the over-all television system (Fig. 9-1). Let us digress for a moment to see what need, if any, exists for such a unit.

In our discussion of the transmitting facilities we have directed our attention wholly to the visual transmitter. In an actual station two transmitters must be used, the visual unit and the aural unit, which transmits the sound program accompanying the television picture. Normally one might expect that each transmitter feeds its own separate antenna system; such a procedure has been used in several television stations. Another approach to the problem is possible, however, because the audio carrier lies just above the upper limit of the video sidebands. Thus, if an antenna can be devised which can handle the full 6-megacycle bandwidth allotted to aural and video signals and if a feed system can be devised to prevent interaction between the two transmitters, they both may feed a single antenna system with a large saving in initial installation cost. The second condition is satisfied by the diplexer, which permits both transmitters to feed a single output without interaction.

²⁰ An alternative form of broadband Balun, which gives a standing-wave ratio of less than 1.25 over a 4 to 1 frequency range, is described by E. G. Fubini and P. J. Sutro in "A Wide-Band Transformer from an Unbalanced to a Balanced Line." *Proc. IRE*, **35**, 10 (October 1947).

Consider the Wheatstone bridge circuit shown in Figure 13-28a. Since the two resistances are equal and the two reactances are equal, the bridge is balanced. Both generators deliver power to the resistors, but they are completely isolated from each other. No voltage appears between U and V due to E_2 , and E_1 produces no voltage

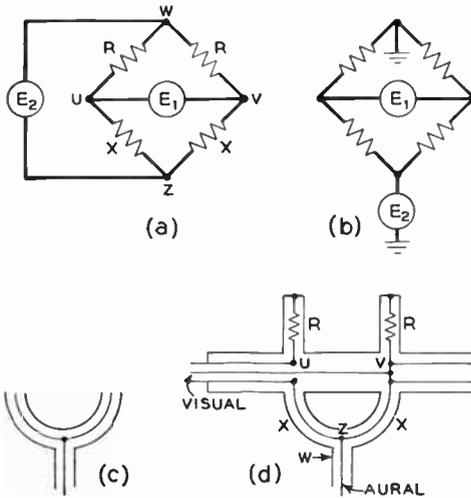


Fig. 13-28. Development of the diplexing unit. (a) Wheatstone bridge with two applied voltages. (b) The return path for E_2 is completed through ground. (c) The lower half of the bridge is replaced by a coaxial Y section. (d) The Y is combined with a Balun to give the complete circuit.

between W and Z . By using ground as a common return we may draw the circuit as shown at b . We have already seen that we may replace a lumped reactance by a section of transmission line; hence we may replace the lower half of the bridge by a coaxial Y shown at c .

Now let E_1 be the output of the aural transmitter, and E_2 the output of the visual transmitter after it has been converted to a push-pull form by a Balun like that shown in Fig. 13-27b. If this unit is combined with the Y section, we have the diplexer in its final form as at d . Corresponding points in the diplexer and the basic Wheatstone bridge are labeled similarly to simplify comparison of the two circuits, which are equivalent to each other.

It must be emphasized that the diplexing unit is not essential. Its use does eliminate the need for two separate antenna systems for the

visual and aural programs; it is a standard feature with the RCA TT-5A transmitting equipment.

TRANSMITTING ANTENNAS

It is beyond the scope of our work to study transmitting antennas in detail.²¹ We shall, however, state a few of the principal factors in their operation and describe a few typical antenna systems. Before proceeding to these factors we must first decide upon the direction of polarization of the radiated signal and the shape of radiation pattern required.

13-17. General Requirements

An electromagnetic wave traveling in space has two components, an electric and a magnetic field. If each of these two fields be represented by a vector, the electric field vector and magnetic field vector will be normal to each other and to the direction of propagation. The direction of polarization of the wave is taken to be the direction of the electric field.

A horizontal antenna radiates a wave whose electric field is horizontal; hence we may say that the wave is horizontally polarized. On the other hand a vertical antenna radiates a vertically polarized wave. Thus in deciding the direction of polarization of the radiated signal we are in effect deciding whether a vertical or horizontal antenna shall be used. The National Television System Committee in its report to the Federal Communication Commission recommended that horizontal polarization be standardized, for in comparison to vertical polarization it gives a better signal to noise ratio, is less susceptible to multipath reflections, and simplifies the design of the receiving antenna.²² It is interesting to note that vertical polarization permits a simpler transmitting antenna; the choice of horizontal polarization favors the receiver at the expense of the transmitter.

In reference to the location of the television transmitter the F.C.C. specifies that "the transmitter location should be as near the center of the proposed service area as possible, consistent with the appli-

²¹ See G. E. Hamilton and R. K. Olsen, "Television Antenna Design." *Communications*, February and March 1947.

²² For a complete comparison of the relative merits of the two types of polarization see D. G. Fink, *Television Standards and Practice*. New York: McGraw-Hill Book Company, Inc., 1943.

cant's (station owner's) ability to find a site with sufficient elevation to provide service throughout the area. Location of the antenna at a point of high elevation is necessary to reduce to a minimum the shadow effect on propagation due to hills and buildings, which may reduce materially the intensity of the station's signals in a particular direction."²³ In compliance with this directive most transmitters are located near the center of their service areas and, hence, require an antenna that radiates equal power in all directions in the horizontal plane. Such an antenna is described as omnidirectional.

It is convenient to show the directional characteristics of an antenna by means of its radiation pattern, which is a polar plot of relative radiated field strength v. direction in the horizontal plane. In choosing a transmitting antenna, then, we seek a horizontally polarized antenna system whose radiation or directivity pattern in the horizontal plane is a circle.

Certain other mechanical features are required of the transmitting antenna: it should be sufficiently rigid to withstand wind and ice loads and should be easy to mount at a high elevation.

13-18. The Dipole

The basic radiating element of most television transmitting antennas is the horizontal dipole, or half-wavelength-long antenna, fed at the center and placed parallel to the earth. For such a radiating element the field strength \mathcal{E} at any angle θ measured from the axis of the antenna in the horizontal plane is given by²⁴

$$\mathcal{E} = 60 \frac{I}{d} \left[\frac{\cos \left(\frac{\pi}{2} \cos \theta \right)}{\sin \theta} \right] \text{volts/meter} \quad (13-63)$$

where I is the current at the antenna midpoint in amperes and d is the distance from the antenna in meters. The field pattern or relative field strength may be plotted directly from the bracketed factor in eq. (13-63). The results are shown in Fig. 13-29. This "figure eight" radiation pattern indicates that the simple dipole alone falls short in meeting the omnidirectional requirement on the television antenna.

²³ Federal Communications Commission, *Standards of Good Engineering Practice Concerning Television Broadcast Stations*, 1945.

²⁴ F. E. Terman, *Radio Engineers' Handbook*, New York: McGraw-Hill Book Company, Inc., 1943.

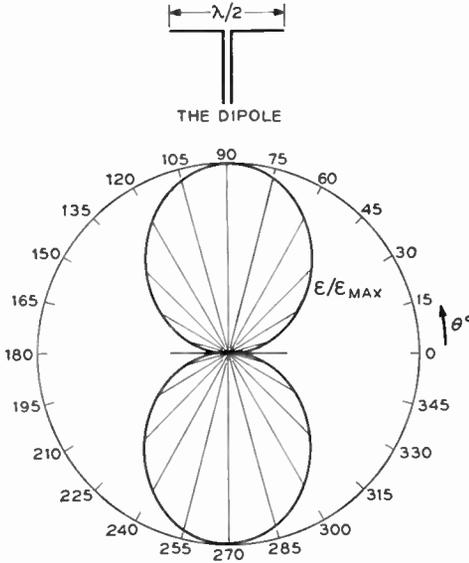


Fig. 13-29. Relative field strength of a dipole in the horizontal plane.

$$\epsilon_{\text{max}} = \frac{60I}{d} \text{ volt/meter}$$

13-19. Crossed Dipoles

If, now, two dipoles be placed normal to each other and fed in time quadrature (*i.e.*, the two are driven 90° out of phase), each dipole will exhibit a figure-eight field, but the resultant field will have the shape shown in Fig. 13-30. This resultant characteristic is derived on the following basis. In any direction θ the fields from the two antennas are in time quadrature; hence the resultant field is

$$\epsilon_r = \sqrt{\epsilon_1^2 + \epsilon_2^2} \quad (13-64)$$

Inspection of the diagram shows

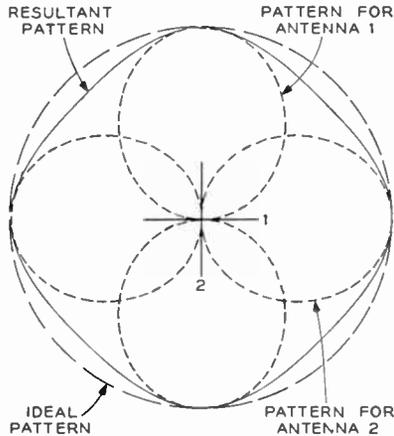


Fig. 13-30. Radiation pattern of a pair of crossed dipoles fed in time quadrature.

that the crossed dipoles produce a field which closely approximates the ideal circular pattern, and such a combination may be assumed satisfactory for television work insofar as the horizontal directivity is concerned.

Under certain circumstances the configuration of the supporting structure is such that the pair of crossed dipoles cannot be used. As a case in point, consider the tower of the Chrysler Building in New York City, which at one time housed the transmitting facilities of WCBS-TV. As may be seen from Fig. 13-31 the tower proper is

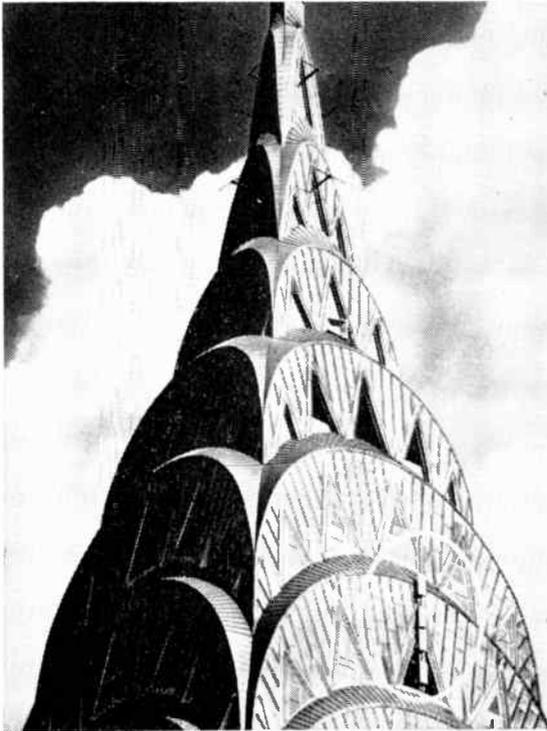


Fig. 13-31. Transmitting antennas located on the tower of the Chrysler building in New York City. The antenna circled in white is an experimental one, used for color transmission in the vicinity of 500 megacycles. At the top of the picture are four dipoles for transmission of the aural program. On the next lower level are the video antennas of the folded dipole type. (Courtesy of Columbia Broadcasting System.)

topped by a long aluminum spire. Were the crossed dipoles to be used, they would have to be mounted on the spire itself. Installation costs render such an installation impracticable and certainly the appearance of the building would not be enhanced by the protruding antenna elements and their associated transmission lines. To overcome these difficulties an antenna array consisting of four dipoles, one on each of the four tower sides, has been used.²⁵ With these four elements fed in phase the field pattern has eight lobes, as shown in Fig. 13-32, and offers a satisfactory compromise. The particular pattern shown in the figure is for an early prewar installation in which diplexing was not used. The several antennas which may be seen on the tower are identified in the caption of Fig. 13-31.

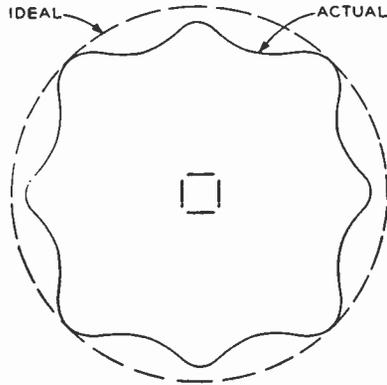


Fig. 13-32. Radiation pattern of four dipoles arranged in a square and fed in phase. The rear lobe of each dipole is reflected in the forward direction by the tower on which the antennas are mounted.

It should be noticed that in the Chrysler Tower installation each dipole is backed by the surface of the tower so that the energy radiated in the back lobe of the figure eight is reflected in the forward direction and eq. (13-63) no longer obtains. It is interesting to note that along the diagonals through the tower the fields from adjacent dipoles overlap, causing the field to be a maximum.

13-20. Vertical Directivity

We have observed that the use of more than one dipole in the horizontal plane modifies the radiation pattern in that plane. In a similar manner, the stacking of dipoles one above the other modifies the radiation pattern in the vertical plane. To illustrate this principle we consider the radiation pattern of the dipole in the plane perpendicular to the dipole axis. Since rotational symmetry prevails

²⁵ P. C. Goldmark, "Problems of Television Transmission." *J. Appl. Physics*, 10, 7 (July 1939).

in the plane, the radiation pattern is a circle. Notice, then, that a vertical dipole would be an ideal radiating element because of its omnidirectional characteristic in the horizontal plane. We have already seen, however, that it would produce a vertically polarized wave, which is deemed less desirable for the television system as a whole.

With the dipole placed parallel to the earth, its circular vertical pattern represents a waste of radiated energy because all the energy radiated above the horizon is lost as far as receiving sets are concerned. It is desirable, therefore, to modify the vertical radiation pattern so that more energy is concentrated toward the horizon. This may be accomplished by stacking two dipoles one above the other and feeding them in phase. For this simple array of two dipoles, spaced a half-wavelength apart and driven by two equal in-phase currents the field in any direction ϕ is given by

$$\varepsilon_r = 2\varepsilon \cos\left(\frac{\pi}{2} \cos \phi\right) \quad (13-65)$$

Evaluation of (13-65) shows that the resultant pattern is double-lobed, as shown in Fig. 13-33, and that the maximum field is in the horizontal direction ($\phi = 90^\circ$).

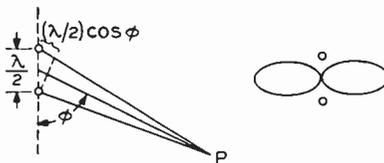


Fig. 13-33. Two dipoles are stacked vertically and fed in phase to increase the radiated energy in the direction of the horizon.

Notice also that in this direction the resultant field is twice that produced by a single dipole acting alone. The required increase in directivity has been obtained.

It is instructive to reason out the shape of the double-lobed pattern on a physical basis.

Since the pattern is symmetrical about the vertical and horizontal axes, we need consider only the lower right-hand quadrant, defined by $0 \leq \phi \leq 90^\circ$. Consider a point P far removed from the antenna and in the direction $\phi = 90^\circ$. The two antennas are excited in phase and are the same distance away from P ; hence the fields from both dipoles are in phase at P and add algebraically to give $\varepsilon_r = \varepsilon + \varepsilon = 2\varepsilon$.

Next consider a point P below the array at $\phi = 0^\circ$. Since the wave from the upper antenna travels one half-wavelength farther

horizontal-sweep system to these false sync pulses, a number of methods of stiffening the horizontal-sweep oscillator have been devised. The circuit shown in Fig. 14-23 is representative of one method which is used. A sine-wave horizontal oscillator, V215, is under control of the reactance tube V220. V214 serves in a discriminator circuit, which compares the oscillator and sync frequencies. Any difference between the two shows up as a d-c error voltage across the plate load resistors, R263 and R264; this voltage is fed to the grid of the reactance tube. Coarse control of the oscillator frequency is provided by the tuning slug in the secondary of Z204, which is also the oscillator tank inductance. Fine control is provided by the feedback loop, consisting of the discriminator and the reactance tube. The detailed operation of the control circuit may be made clearer with the aid of Fig. 14-24, which shows a similar unit that was used

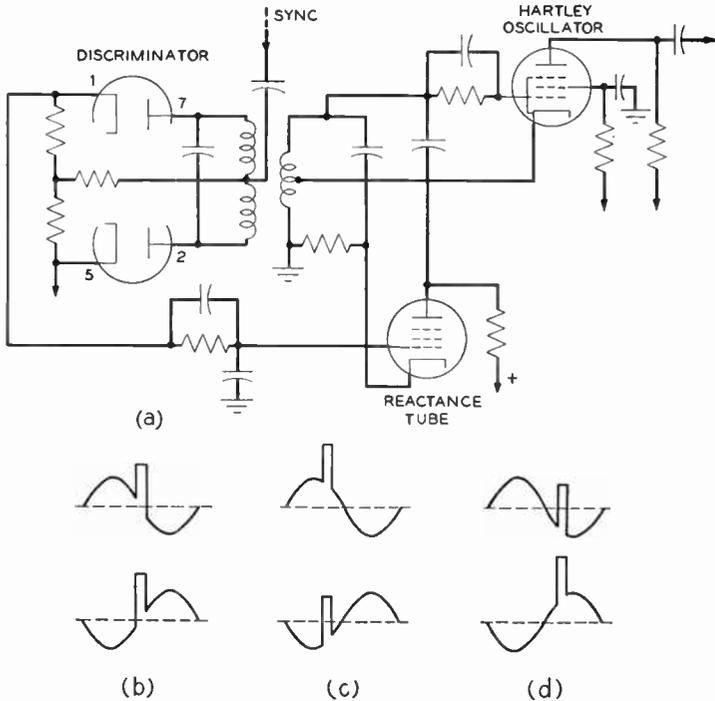


Fig. 14-24. Detail of an automatic frequency control for the horizontal sweep system. The circuit is similar to that used in Fig. 14-23. (Courtesy of *RCA Review*.)

In the Du Mont receiver shown in Fig. 14-23 separation of the sync from the picture components is handled in a manner similar to that described in the preceding paragraph: The diode, V204B, serves as the d-c restorer for the Teletron²⁹ and as a diode sync separator. Two stages of sync clipping are afforded by tubes V212-A and V213. Notice that the signal is inverted in the first clipper stage; hence both the top and bottom of the sync pulses are clipped. The reason for this double clipping is seen from the following considerations. The magnitude of the complete composite video signal is under adjustment of the contrast control, R232, which feeds bias voltage to the control grids of the first two video I-F amplifiers. This control is normally adjusted to give proper contrast in the reproduced image. As the I-F gain is varied, the sync signal amplitude varies. The double clipping keeps the magnitude of the sync output from V213 constant over a large range of those variations and so provides more stable control of the deflection circuits, particularly in the horizontal system which will be described.

The vertical-sweep system is quite conventional. The clipped sync pulses are fed through a buffer or isolating amplifier to a two-stage integrating network, whose output is inductively coupled to the vertical blocking oscillator. V216B serves as both blocking oscillator and discharge tube for the sweep-generating circuit, C257 and R279. The bias on the vertical deflection amplifier is controlled by R281. This resistor controls the operating point and hence the curvature of the tube's transfer characteristic, which is used to compensate for curvature due to the output circuit. Thus R281 adjusts the linearity of the vertical sweep and is labeled the vertical-linearity control.

The horizontal-sweep channel represents a departure from the other circuits which have been described. Bursts of R-F noise, either man-made or natural, which are picked up by the receiving antenna, add to the composite video signal and are amplified along with it. If these noise pulses are of sufficient amplitude to lie in the blacker-than-black region, they pass on to the synchronizing circuits. Since they are spaced in a random manner, they appear as false sync signals and on differentiation may upset the horizontal sweep. Pulses of this type have little effect upon the vertical system because they are usually quite narrow and contribute a negligible amount to the output of the integrating system. To overcome the susceptibility of the

²⁹ Teletron is the trade mark of the Du Mont picture tubes.

multivibrator circuit eliminates the need for an additional discharge tube; the charging condenser, which develops the saw-tooth sweep voltage, is connected directly from the right-hand plate of the tube to ground (0.1 μf and 500 $\mu\mu\text{f}$, respectively, for the vertical and horizontal sweeps). Push-pull deflection is provided by plate-coupled paraphase amplifiers in both sections. It may be observed that d-c centering or positioning voltages are fed to one horizontal and one vertical deflection plate in the 7EP4.

The use of a flyback power supply is not feasible because electrostatic deflection is used. The high voltage is supplied from a 60-cycle high-voltage transformer and a 2X2 rectifier. A separate 2.5-volt insulated winding is provided for the 2X2 filament, which is at high voltage with respect to ground. The negative lead of the high voltage system is connected to $B+$ of the low-voltage power supply.

The 12-inch magnetic deflection set shown in Fig. 14-22 has a slightly different deflection system from that just described. Notice first that the output of the 6AG7 video amplifier is capacitance-coupled to the picture tube grid and so a d-c insert is required at that grid. The necessary action is furnished by the right-hand section of the video detector 6AL5, which is labeled X-10 in the diagram. One advantage gained here is that the sync components of the composite video signal are separated in the same tube because it serves as a rectifier with only the sync components appearing across the 22-kilohm resistor from pin 7 of X-10 to ground. X-8 and X-9 serve to amplify the sync signals and to square them up by clipping. The fourth section operates as a cathode follower so that the proper polarity is maintained in the outputs feeding the two sweep channels. The remainder of the sweep channel comprises units which have already been described at some length. The differentiated or integrated sync wave is used to synchronize a blocking oscillator, which, in turn, controls the discharge tube of the sweep-generating circuit. Notice that in both horizontal and vertical channels the sweep voltage is generated across a resistor and condenser in series in order that the developed voltage be trapezoidal. The high voltage for the picture tube is obtained from a flyback supply on the horizontal-sweep output tube. Notice also that the 6BG6-G plate returns to $B+$ through the damper and yoke system, thereby indicating that reaction scanning is used.²⁸

²⁸ See Chapter 4.

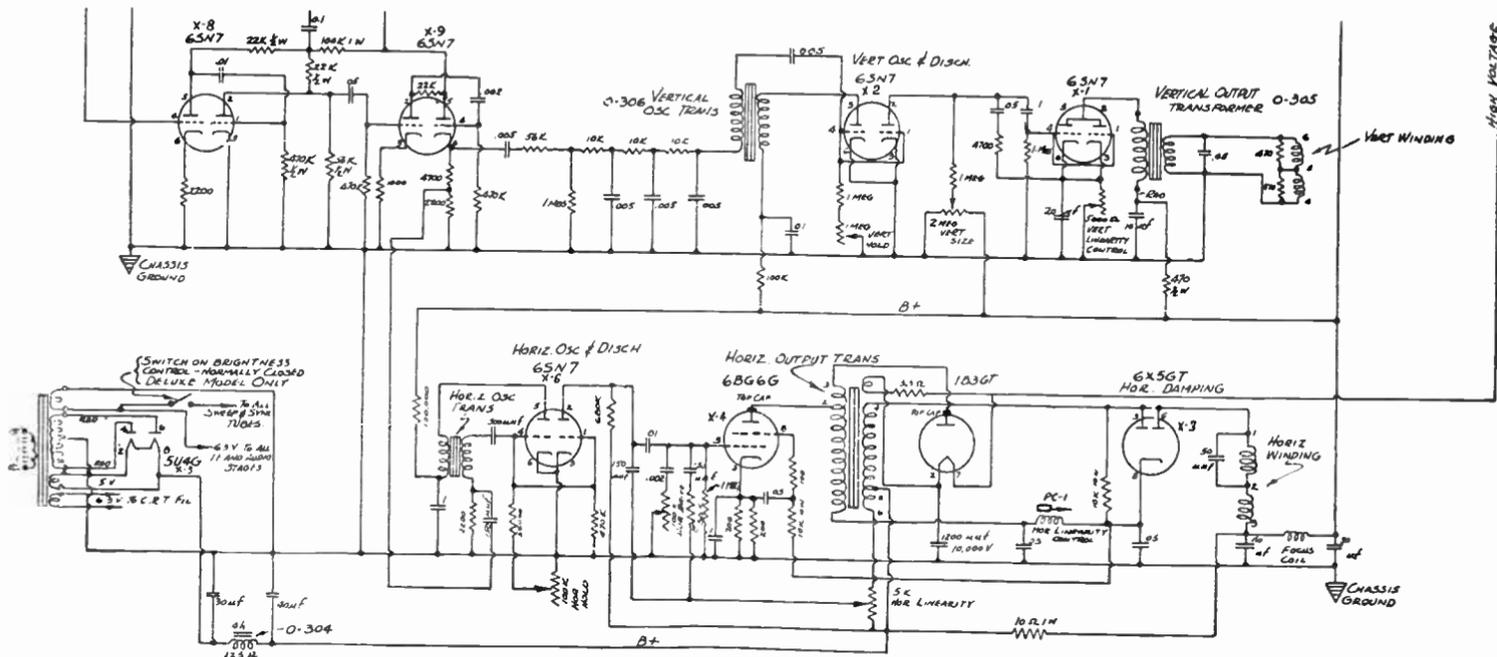


Fig. 14-22. A 12-inch magnetic-deflection receiver. The front end is not shown. (Courtesy of Transivision, Inc.)

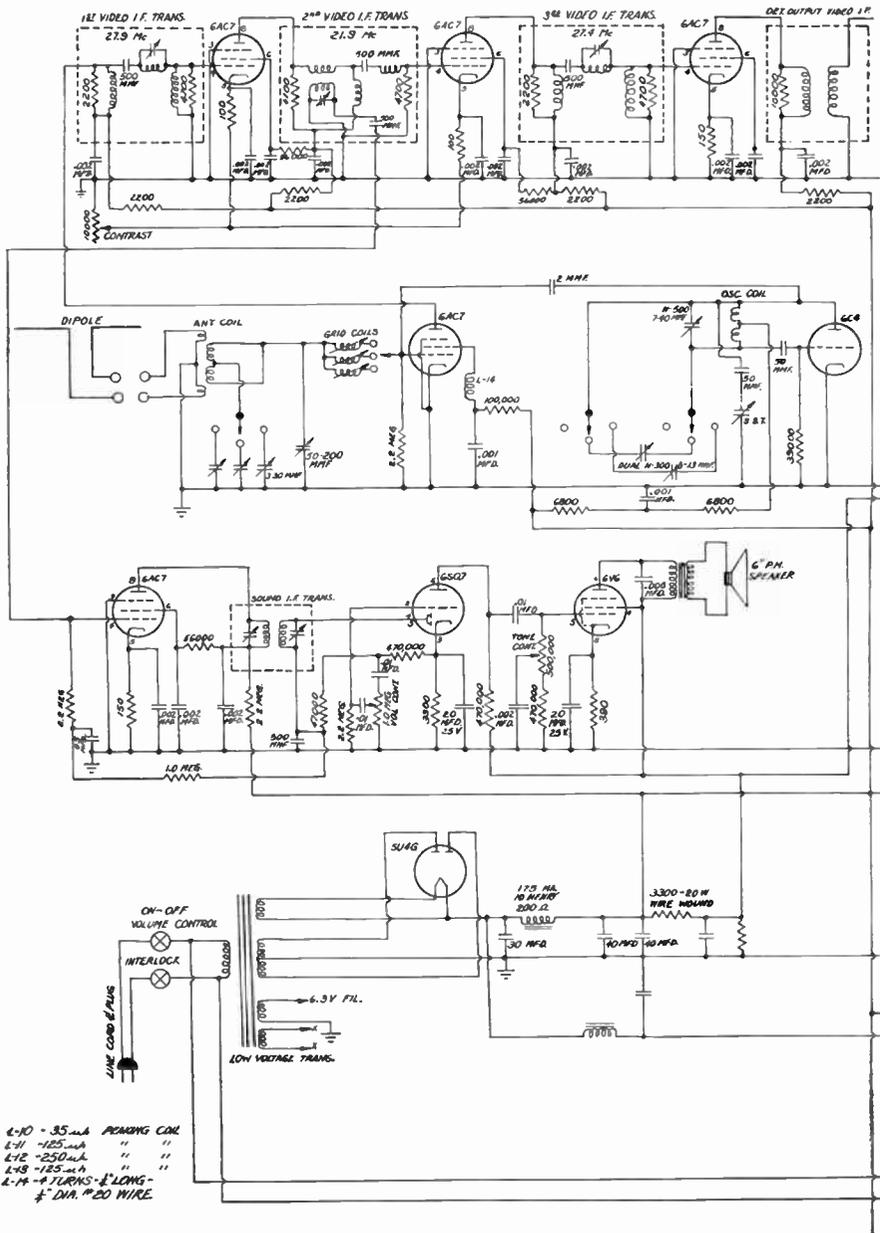


Fig. 14-21. A 7-inch electrostatic-deflection television receiver. Provision for receiving only three channels is shown. (Courtesy of Transvision, Inc.)

pedance in order to ensure a full response over the entire video band.

The video amplifiers and restorer circuit²⁷ have been discussed in previous chapters and though many variations are employed in commercial models, they will not be discussed here. Several typical video systems are shown in the accompanying diagrams. In each case the several components may be identified quite readily. The operation of the manual brightness control, which sets the background level in the reproduced image, has been described in Chapter 7.

14-9. The Deflection Circuits

Whereas the basic forms of deflection circuits have been studied in our previous work, we shall consider some of the details which are found in typical commercial receivers. Frequent reference will be made to the four diagrams of Figs. 14-21 to 14-24. All four of the receivers employ magnetic deflection and a flyback power supply except the 7-in. model of Fig. 14-21, which utilizes an electrostatic deflection picture tube Type 7EP4. In this set the composite video signal, consisting of picture and supersync components, is amplified as a unit in a single stage that employs a 6AG7. Since the tube has no fixed bias, clamping occurs on its grid; the d-c component is retained by direct-coupling the plate to the picture tube grid. The polarity of the output voltage is sync-negative; hence the supersync signals drive the picture tube beyond cutoff and so do not affect the final picture.

A portion of the output voltage is tapped off the video amplifier plate load and fed to the sync amplifier stage, which amplifies and inverts the composite video signal. The sync-positive signal is clamped by the right-hand half of the video detector diode and delivered to the sync separator from which two outputs are derived. The plate feeds a two-section integrating network, whose output, in turn, synchronizes the vertical oscillator. The second output, which is developed across the cathode return, is fed through a 150 μf differentiating condenser to the cathode of the horizontal oscillator.

In both horizontal and vertical deflection systems the impulse oscillator is a modified form of multivibrator in which the coupling from the right- to the left-hand section of the tube is accomplished through a common cathode resistor. Use of this cathode-coupled

²⁷ D. E. Foster and J. A. Rankin, "Video Output Systems," *RCA Review*, V, 4 (April 1941).

the amplifier gain at that frequency, thereby providing the required action. A variation of the degenerative trap circuit is shown at *e* in the diagram, where the actual trap is coupled by mutual inductance to L_1 , in the cathode return. Direct application of coupled-circuit theory to the network shows that the impedance of the network as viewed from the terminals of L_1 is

$$Z = j\omega L_1 - \frac{(j\omega M)^2}{Z_{22}} = j\omega L_1 + \frac{(\omega M)^2}{Z_{22}} \quad (14-75)$$

where Z_{22} is the loop impedance of the L_2C_2 circuit; hence Z will be maximum at the series resonant frequency of the trap circuit when Z_{22} will be a minimum. Thus degeneration occurs at the trap frequency, just as it did for the circuit of Fig. 14-19*d*.

The absorption type of trap, where a resonant circuit is coupled into the conductive circuit by means of mutual inductance, need not be restricted to use in the cathode return of the amplifier. In fact, the absorption-type trap has been mentioned in connection with Fig. 14-17, where the coupling is to the secondary side of the inter-stage transformer. Several variations of this circuit may be seen in the diagrams of the complete receivers (Fig. 14-21 to 14-24). It is stressed once again that other forms of trap circuits, such as the bridged T network, may be used; we have mentioned only a few representative forms to indicate some of the methods in common use.

The last feature to be mentioned in regard to the I-F amplifier system is the method of gain control used. If the amplifier tubes are of the remote cutoff type, their gain may be controlled by means of their bias voltage. Thus in some commercial receivers the grid return of one or more of the I-F stages is connected to a d-c source under control by a potentiometer. The effect of such a control on the reproduced picture is to increase the spread of the reproduced picture along the gray scale and so is known as the contrast control. A similar effect may be obtained by varying the bias on the input stage of the receiver.

14-8. The Video System

The remaining portion of the visual channel of the television receiver comprises the second detector, video amplifier stages, clamping or d-c restorer circuit, and the picture tube proper. The detector is a diode type of linear detector designed with a compensated load im-

values of I-F frequencies, there should be zero attenuation at 21.5 megacycles; thus we set

$$f_1 = 21.5 \text{ megacycles} \quad (14-70)$$

and, since the associated aural carrier is to be rejected,

$$f_2 = 21.25 \text{ megacycles} \quad (14-71)$$

Then at f_1 , L and C must be anti-resonant or

$$\omega_1 L = \frac{1}{\omega_1 C}$$

$$\omega_1^2 = \frac{1}{LC} \quad (14-72)$$

At f_2 the entire combination must be in series resonance, thus

$$-\frac{1}{\omega_2 C_c} = \frac{\omega_2 L}{\omega_2^2 LC - 1} \quad (14-73)$$

and, substituting for LC from (14-72),

$$-\frac{1}{\omega_2 C_c} = \frac{\omega_2 L}{\left(\frac{\omega_2}{\omega_1}\right)^2 - 1}$$

whence

$$L = \frac{1 - \left(\frac{\omega_2}{\omega_1}\right)^2}{\omega_2^2 C_c} \quad (14-74)$$

C_c is usually taken to be roughly $5 \mu\mu\text{f}$ and L and C may be evaluated from eqs. (14-74) and (14-72), respectively. The advantage of the circuit over the simple resonant circuit type of Fig. 14-19b is that the rise in the impedance at f_1 provides a much steeper characteristic in the vicinity of rejection, and a smaller portion of the visual band is rejected along with the aural signals.

In the circuit of Fig. 14-19d the trap takes the form of an antiresonant circuit placed in the cathode return of the I-F amplifier stage. At the antiresonant frequency the trap presents a high impedance to the flow of plate current. Thus a large degenerative voltage is developed between the cathode and ground, which lowers

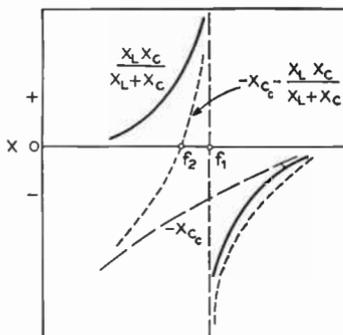


Fig. 14-20. Reactance curves for the trap circuit of Fig. 14-19(c).

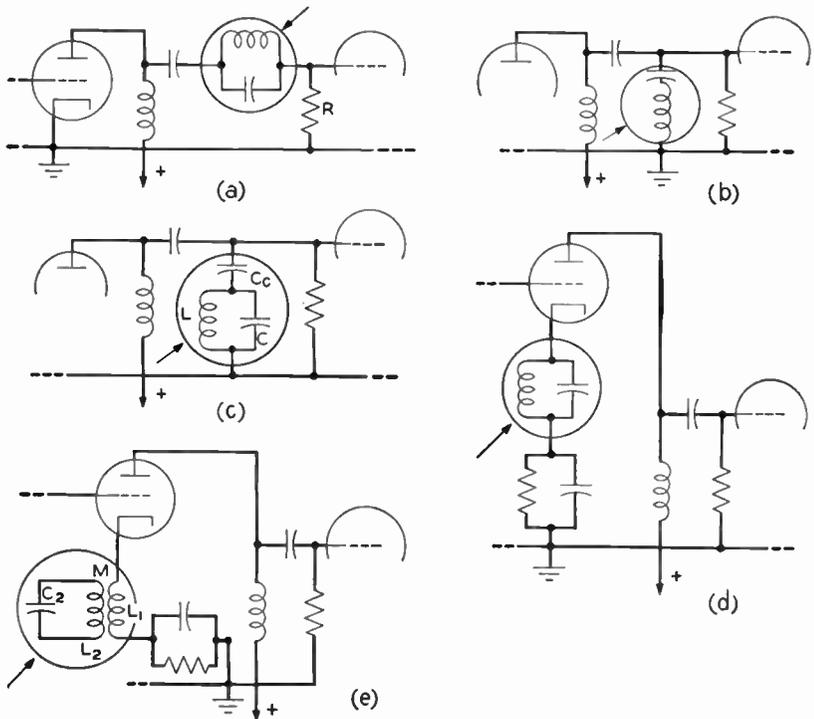


Fig. 14-19. Common forms of trap circuits. (a) Series. (b) Shunt. (c) Shunt. (d) Degenerative. (e) Degenerative-absorption.

Actually, however, the condenser C_c , which couples the trap to the grid lead, enters into the operation of the circuit, which may most conveniently be analyzed by means of reactance diagrams as shown in Fig. 14-20.²⁶ These curves show that the combination is series-resonant at some frequency f_2 and antiresonant at f_1 . By proper choice of the components f_1 and f_2 may be made quite close and form an excellent combination for providing a sharp attenuation of the associated aural carrier with minimum effect on the visual pass band (cf. Fig. 14-18b). The values of the three circuit components may be calculated on the following basis: For the previously specified

²⁶ Direct application of Foster's reactance theorem yields the same results.

In the two other common forms of I-F amplifiers which employ band pass filter coupling or stagger tuning, the design philosophy is similar to that just described. The basic pass characteristic is designed to be symmetrical, and trapping is used to distort it into the appropriate shape. Since the use of stagger tuning is comparatively new, it is covered at some length in the next chapter. In essence the method employs two or more stages, each of which has a properly loaded single-tuned circuit. The wide flat-topped pass characteristic, which is not characteristic of single-tuned circuits, is obtained by tuning the successive stages to different frequencies.²⁵ The advantages of the stagger-tuned amplifier are noteworthy. The interstage coupling networks are simplified because one, rather than two, tuned circuits is used. The coupling circuit coils are cheaper because fewer windings are required. High gains may be obtained with good stability; since successive stages are tuned to different frequencies, the likelihood of oscillation is minimized. The amplifier is easy to adjust; only a single tuning adjustment is required for each stage.

Any of a number of trap circuits may be used to provide attenuation at the several carrier frequencies that must be rejected. It is convenient to name the traps by either their location in the circuit or by their mode of operation. Thus in Fig. 14-19*a* the trap is in series with the connection between plate and grid of the adjacent amplifier stages and so is termed a series trap. Its manner of operation is quite simple and analogous to that in the series compensated video amplifier. At the reject frequency the trap presents a high impedance and the grid voltage developed across R is negligible. At virtually all other frequencies the trap impedance is small and, essentially, the full a-c plate voltage is delivered to the grid. In other words, the trap and grid return resistor behave like a frequency-sensitive voltage-dividing network.

In Fig. 14-19*b* the trap circuit shunts the interstage coupling network and so is called a shunt trap. At the reject frequency the trap is in resonance and shorts the grid to ground. A second form of shunt trap is shown at *c*. At first glance the trap appears as a parallel resonant circuit, which would not provide a low shunt path to ground.

²⁵ An alternate form of I-F amplifier that employs single-tuned circuits has been designed; it utilizes feedback in place of stagger tuning. See E. H. B. Bartelink, J. Kahnke, and R. L. Waters, "A Flat-Response Single-Tuned I-F Amplifier." *Proc. IRE*, **36**, 4 (April 1948).

It will be recognized that the percentage bandwidth of the I-F system is large and that the effective Q of the system is low; hence the narrow-band approximations usually applied to broadcast-band tuned amplifiers are invalid and the networks must be designed on the exact basis that recognizes that their response displays geometric symmetry.²³ Thus the center frequency of the coupling networks, or the frequency to which they are tuned, will be

$$\begin{aligned} f_m &= \text{geometric mean of the pass band} \\ &= \sqrt{(21.5)(25.25)} = 23.4 \text{ megacycles} \end{aligned} \quad (14-69)$$

With f_m and the width of the flat top specified, sufficient data are available to permit design of the double-tuned circuits. The resulting pass characteristic is indicated by the dashed line at b in Fig. 14-18.

It will be observed at once that the pass band includes the aural carrier of adjacent channel No. 7, the visual carrier of adjacent channel No. 9, and the associated aural carrier of channel No. 8. These are eliminated by traps which may be of the form represented by L_2 and C_2 in Fig. 14-17. In the over-all amplifier, then, at least three such traps must be provided, which are tuned, respectively, to 19.75 megacycles (adjacent visual carrier), 21.25 megacycles (associated aural carrier), and 27.25 megacycles (adjacent visual carrier). The linear portion of the final pass characteristic in the region of the visual carrier is obtained by means of the trap L_1 and C_1 of Fig. 14-17. Thus the symmetrical pass characteristic of the basic amplifier is literally distorted into the proper shape by means of the trap circuits.

Although the double-tuned form of coupling circuit, which is illustrated in Fig. 14-17, is used quite frequently in modern receiver design, it has one drawback from the manufacturing standpoint: since the coupling between the two coils is by mutual inductance, considerable care must be exercised during manufacture to ensure the proper spacing between the two coils. This difficulty may be overcome by using some other form of coupling. Larsen and Merrill²⁴ have described an amplifier of this type which employs capacitive coupling and is structurally quite simple.

²³ See Chapter 15.

²⁴ M. J. Larsen and L. L. Merrill, "Capacitance-Coupled Intermediate-Frequency Amplifiers," *Proc. IRE*, **35**, 1 (January 1947). This paper includes a derivation of the design equations for a capacitance-coupled trap circuit.

overcoupled and, in the absence of loading, would yield a double-humped response curve. This curve may be flattened out by making R_2 , the grid return resistor, sufficiently low to provide the required amount of shunt loading. The coupling between the several stages of the I-F amplifier is identical to that described.

In order to determine the required over-all pass characteristic of the amplifier without trapping it is well that we fix our attention on a particular R-F channel and observe what happens to the frequency relationships in the mixing process. Let the receiver be tuned to channel No. 8 which extends from 180 to 186 megacycles. The envelope of the transmitted signal is illustrated in Fig. 14-18a along

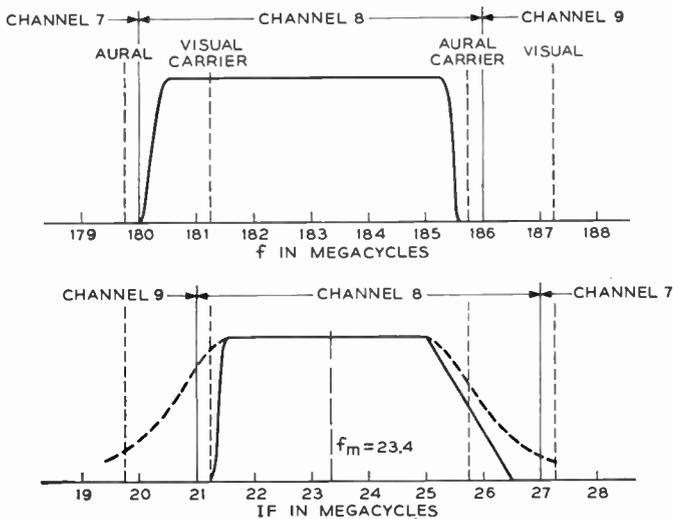


Fig. 14-18. The R-F and I-F characteristics for channel No. 8.

with portions of the two adjacent channels, No. 7 and No. 9. When these several signals pass through the mixer stage, the previously described frequency inversion takes place and the frequencies for the three channels in the I-F region are shown at *b* in the figure. The final R-A characteristic which must be obtained is indicated by the solid line. Since this curve is flat-topped from approximately 21.5 to 25.25 megacycles, the I-F coupling networks are designed to give a flat response over that same frequency range.

The majority of commercial television receivers employ one of three types of I-F amplifier circuits: double-tuned, band-pass filter, or stagger-tuned types. As a practical matter, it is generally most convenient to design the interstage coupling networks so that they provide a symmetrical pass characteristic about the mean of the visual band and then to modify the symmetrical characteristic by means of trap circuits so that it conforms to the ideal curve of Fig. 14-16. Additional trapping is also provided in

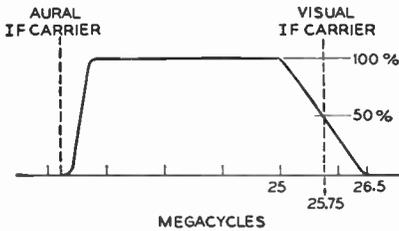


Fig. 14-16. The I-F response required for R-A transmission. Frequency inversion has occurred because the local oscillator frequency is higher than that of the incoming R.F.

order to suppress possible interference from the two channels on either side of that channel to which the receiver is tuned. This will be illustrated in the following paragraphs.

An I-F amplifier that employs double-tuned circuits in the interstage coupling networks is illustrated in Fig. 14-17. Consider the network between the first and second I-F amplifiers. The primary and secondary windings of T_2 are each shunted by the stray and interelectrode capacitances to ground and hence form two parallel resonant circuits which are coupled by mutual inductance. The two windings on T_2 are spaced in manufacture so that the two circuits are

in order to suppress possible interference from the two channels on either side of that channel to which the receiver is tuned. This will be illustrated in the following paragraphs.

An I-F amplifier that employs double-tuned circuits in the interstage coupling networks is illustrated in Fig. 14-17. Consider the network between the first and second I-F amplifiers. The primary and secondary windings of T_2 are each shunted by the stray and interelectrode capacitances to ground and hence form two parallel resonant circuits which are coupled by mutual inductance. The two windings on T_2 are spaced in manufacture so that the two circuits are

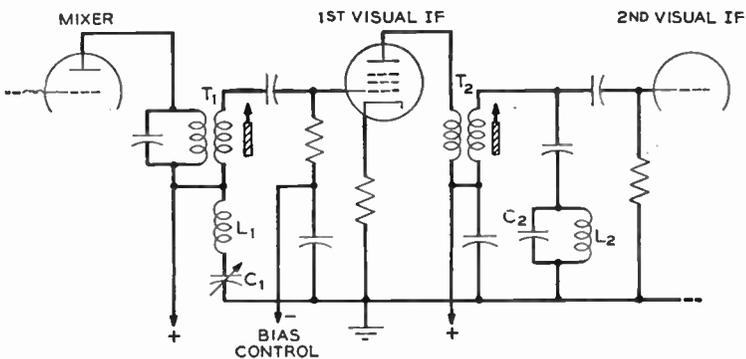


Fig. 14-17. I-F amplifier with double-tuned interstage coupling network.

than frequency changes on its control grid. The fallacy in this belief is that the frequency-modulated signals suffer slope detection at the visual second detector and, in their demodulated form, can cause considerable interference in the final image; hence, care is taken to prevent any of the F-M signals from reaching the video second detector.

The second representative form of separation circuit, shown at *b* in Figure 14-15, is similar to that just described except that the aural tuned circuit is placed in the cathode return of the mixer. The third circuit, shown at *c*, is used in the R.C.A. receivers and operates in a slightly different manner: the secondary circuit of the transformer *T* serves both as the aural take-off and as an aural trap for the visual system. The primary inductance of *T* is adjusted by means of a tuning slug so that it forms a broadband antiresonant circuit with the stray capacitance at f_0 , the mean of the visual band. The secondary circuit is tuned to 21.25 megacycles and hence removes the aural components by absorption; these are then tapped off from the secondary winding of *T* and delivered to the aural I-F channel. The remaining visual components, which appear across the primary of *T*, are coupled to the grid of the first visual I-F amplifier. The system just described does not eliminate the need for additional trapping at 21.25 megacycles in the visual system. One of the principal advantages afforded by the system is that only stray capacitance is needed to tune the primary circuit; we have already seen that the gain-bandwidth product of the amplifier is limited by the total shunt capacitance across the plate load; thus by holding *C* to its minimum possible value, we maximize the gain-bandwidth product.

We consider next the amplitude response of the visual I-F amplifier and the circuits of those amplifiers. We have previously stated that the I-F amplifiers must filter the lower sideband of the received signal in a particular manner so that the proper input is delivered to the second detector. In this connection we must take note of the fact that in the mixing process a frequency inversion has taken place so that the position of the two sidebands in the R-F signal is reversed in the I-F signal; hence the critical portion of the attenuation characteristic lies on the upper I-F sideband. Thus the required pass characteristic is that shown in Fig. 14-16. The over-all response of the I-F system is made to approach this characteristic by the use of tuned amplifiers and traps.

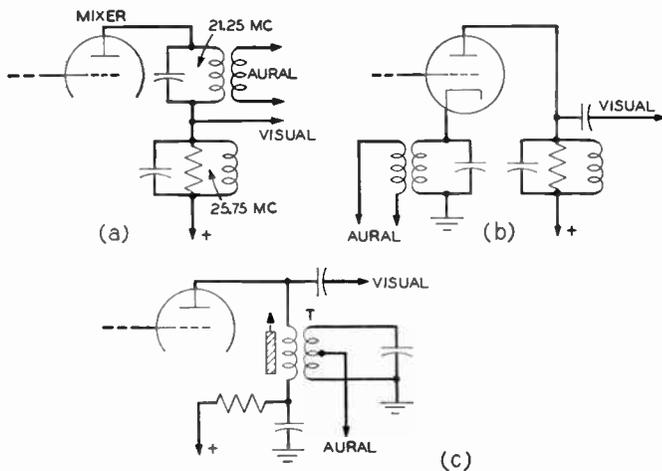


Fig. 14-15. Typical forms of I-F separation. (a) Series. (b) Cathode. (c) Separation by absorption.

of the several cross-modulation components, including the aural and visual I-F signals that encompass a 6-megacycle bandwidth, flows through a load consisting of two antiresonant circuits in series. Since the upper circuit is tuned to 21.25 megacycles, the aural components of the plate current produce an appreciable signal across the circuit, which signal is coupled off to the remainder of the aural system. Since Q of the tuned circuit is high, the response is narrow and all but the highest visual I-F components are virtually absent. Those visual components which are passed by the resonant filter system may be removed by traps introduced farther along in the aural channel.

The same sort of explanation, with one exception, holds for the lower tuned circuit, which is tuned to the visual I-F carrier of 25.75 megacycles. Since the visual signals cover some 5.25 megacycles, the tuned circuit is broadbanded with a shunt resistor, and the aural signals are rejected to only a small degree. This condition is remedied by using several high- Q rejection traps, tuned to 21.25 megacycles in the remainder of the visual I-F system.

It is often believed that the presence of the frequency-modulated signals in the visual system will have no effect on the reproduced picture because the cathode-ray tube is sensitive to amplitude rather

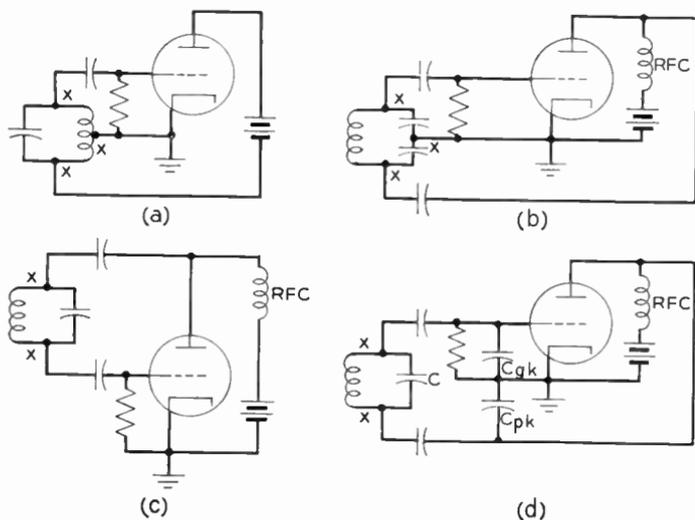


Fig. 14-14. Common forms of oscillator circuits. (a) Hartley. (b) Colpitts. (c) Ultraudion. (d) Ultraudion redrawn as a Colpitts oscillator.

The three sections of the receiver which have been discussed are normally built and considered as a single unit, termed the "front end" of the receiver. In this front portion of the receiver the two programs, aural and visual, are handled as a single signal. In the I-F system they are separated and thereafter handled independently. In our work, the discussion will be confined primarily to the visual channel. The aural I-F discriminator and audio sections follow conventional frequency-modulation practice. We next consider the I-F portion of the receiver.

14-7. I-F System

The first problem to be considered in the intermediate frequency system is the separation of the aural and visual programs. In the discussion which follows we shall assume the visual and aural carriers to be 25.75 and 21.25 megacycles, respectively. Several forms of separation circuit may be used. Three representative types are shown in Fig. 14-15. In the first two, shown at *a* and *b*, separation is based on the resonance rise of voltage across an antiresonant circuit. Thus, in the circuit at *a* the plate current of the mixer, which consists

Since all the R-F signal frequencies are lowered by the same amount in the heterodyning or mixing process, the 4.5-megacycle separation between the two carriers is preserved in the I.F., a fact which may be observed in the list of commonly used frequencies just given. It may also be observed that an inversion has been brought about because the oscillator operates at a higher frequency than the R.F., thus the visual I-F carrier is higher than the aural I-F carrier. In the postwar period there has been a definite trend toward higher values of I.F., or to a greater separation between oscillator and signal frequencies, and it is probable that the unassigned channel No. 1, which extends from 44 to 50 megacycles, will be used for the television I-F band. This trend is the result of the comparatively large amount of image interference which is obtained with 25- to 30-megacycle I.F.'s.

The third factor which we consider in reference to the local oscillator has to do with the type of circuit to be used. It is immediately apparent that the circuit should require a minimum number of tuned circuits in order to reduce the number of circuits which must be switched as the receiver is tuned from channel to channel; hence the tuned-plate tuned-grid circuit is not used. Of the remaining common stable circuits, namely, the Ultraudion, Colpitts, or Hartley, the first is the most widely used. The reason for this may be seen from the following considerations. The three basic circuits are shown in Fig. 14-14, and in each case the tie points between the tuned circuit and the remainder of the circuit are represented by x 's. As the receiver is switched from channel to channel, the oscillator tank must be changed; hence the number of switch contacts per channel is equal to the number of tie points. It may be seen that the Ultraudion circuit permits the simplest construction, for it requires only two, rather than three, switching points, and so is commonly used in receivers which do not employ continuous tuning.

The Ultraudion is redrawn in Fig. 14-14*d* in order to demonstrate that it is a modified form of the Colpitts circuit. Notice that the tank circuit condensers of the circuit at b are replaced by the inter-electrode capacitances C_{gk} and C_{pk} . The condenser C is usually the fine tuning control.

In the R.C.A. 6SO TS R-F unit the artificial line form of tank circuit is carried over into the oscillator section, which is also operated in push-pull. The circuit operates as a push-pull tuned-plate oscillator, with feedback provided to the grids through external condensers.

corporate a fine tuning control, which may be used to compensate for local oscillator drift. On the basis of the example in the last paragraph we see at once that *the local oscillator should always be adjusted to give optimum sound* rather than optimum picture reproduction.

The example also illustrates the basis for automatic frequency control of the television receiver local oscillator. It may be shown that the frequency-modulation discriminator is sensitive to drift of the center or carrier frequency and delivers a d-c error voltage, whose magnitude and polarity are proportional to that drift. Hence an automatic frequency control system may be set up, comprising a reactance tube shunted across the local oscillator tank, the reactance-tube bias being determined by the error voltage developed by the sound discriminator. The operation of the circuit is identical to that of the automatic frequency control circuit described in Chapter 11. Notice that excellent control is obtained for the picture because the error voltage is derived from the sound channel, which is more vulnerable to oscillator drift than is the broadband picture channel. The validity of the last statement presumes that the sound and picture I-F systems are in proper alignment.

The choice of the oscillator frequency relative to the frequency of the incoming signal is based on three principal factors; image interference, oscillator tuning range, and video I-F amplifier design considerations. For the first two factors it is desirable to have the oscillator frequency higher than that of the signal. Where double-tuned I-F amplifiers are used, the same condition is desirable because the gain-bandwidth product increases as the center frequency is raised.²² On the other hand, the problems of local oscillator stability become more difficult at higher frequencies but the consensus is that this point is outweighed by the other advantages, and the choice of the high local oscillator frequency is almost universal. The actual value of oscillator frequency for any channel is based on the value of intermediate frequency used in the particular receiver. Three common sets of I-F frequencies are listed below.

| | <i>Visual I-F Carrier</i> | <i>Aural I-F Carrier</i> |
|---|---------------------------|--------------------------|
| A | 25.75 megacycles | 21.25 megacycles |
| B | 26.40 | 21.90 |
| C | 37.30 | 32.80 |

²² See, for example, F. E. Terman, *op. cit.*, chap. 7.

of the incoming signal by the grid circuit, a series resonant trap is provided between the far end of the mixer grid line and ground. This series circuit is tuned to the I-F frequency and so eliminates the possibility of I-F frequency signals appearing in phase on the mixer grids.

14-6. The Local Oscillator

In the superheterodyne receiver the beat note of intermediate frequency is produced by mixing the R-F signal with the output of a local oscillator, operating at the proper frequency. In the present section we shall consider three of the principal factors which are of concern in the design of the local oscillator section of the receiver: (1) stability, (2) frequency of the oscillator relative to the input signal, and (3) type of oscillator to be used. We shall consider these in order.

It is an interesting fact that the stability limits on the local oscillator frequency are determined by the audio, rather than the video, portion of the complete television program. This may be seen from the following example: Recall that the television transmission standards permit a total swing of ± 25 kilocycles in the frequency-modulated sound carrier and a bandwidth of approximately 4.5 megacycles for the video signal. Now consider that a local oscillator, operating at 100 megacycles (this figure is chosen arbitrarily to simplify the calculations) drifts 0.03 per cent. Since the input signal frequency remains constant, the corresponding change in the I-F frequency will be

$$0.0003(100) = 30 \text{ kilocycles} \quad (14-68)$$

Clearly a change of 30 kilocycles in 4.5 megacycles will produce but little effect on the video signal. On the other hand, the 30 kilocycles drift in the audio carrier exceeds the change in frequency caused by 100 per cent modulation of the audio carrier and so will cause an intolerable degree of distortion in the demodulated audio signal. Thus, the stability of the oscillator must be such that the frequency-modulated sound signal remains on the linear portion of the discriminator characteristic. To this end the discriminator is normally designed to be much wider than 50 kilocycles, as might be expected from the transmission standards, 300 kilocycles being a typical value.

We have already observed that most commercial receivers in-

Thus,

$$\left. \begin{aligned} i_{b1} &= ae_{g1} + be_{g1}^2 \\ i_{b2} &= ae_{g2} + be_{g2}^2 \end{aligned} \right\} \quad (14-62)$$

and

We first consider the output when the incoming signals are push-pull. Substituting for e_{g1} and e_{g2} in terms of the signal and oscillator voltages, we have:

$$\begin{aligned} i_{b1} &= a(E_s \cos \omega_s t + E_o \cos \omega_o t) + b(E_s \cos \omega_s t + E_o \cos \omega_o t)^2 \\ &= a(E_s \cos \omega_s t + E_o \cos \omega_o t) \\ &\quad + b(E_s^2 \cos^2 \omega_s t + 2E_s E_o \cos \omega_s t \cos \omega_o t + E_o^2 \cos^2 \omega_o t) \\ &= a(E_s \cos \omega_s t + E_o \cos \omega_o t) + b(E_s^2 \cos^2 \omega_s t + E_o^2 \cos^2 \omega_o t) \\ &\quad + bE_s E_o \cos (\omega_s + \omega_o)t + \underbrace{bE_s E_o \cos (\omega_o - \omega_s)t}_{\text{I-F term}} \quad (14-63) \end{aligned}$$

In a similar manner we calculate i_{b2} to be

$$\begin{aligned} i_{b2} &= a(-E_s \cos \omega_s t - E_o \cos \omega_o t) + b(-E_s \cos \omega_s t - E_o \cos \omega_o t)^2 \\ &= -a(E_s \cos \omega_s t + E_o \cos \omega_o t) + b(E_s^2 \cos^2 \omega_s t + E_o^2 \cos^2 \omega_o t) \\ &\quad + bE_s E_o \cos (\omega_s + \omega_o)t + \underbrace{bE_s E_o \cos (\omega_o - \omega_s)t}_{\text{I-F term}} \quad (14-64) \end{aligned}$$

Since the I-F components in both plate currents are of the same sign, they add in the output. Notice also that the ω_s terms will cancel in the output because they are of opposite sign; hence direct feed-through of an incoming signal of I-F frequency is prevented.

The converter is unable to cancel out signals of the I-F frequency which arrive at the two grids in phase, for, in this case,

$$e_{g2} = e_{s1} \quad (14-65)$$

and

$$\begin{aligned} i_{b1} &= ae_{g1} + be_{g1}^2 \\ &= a(E_s \cos \omega_s t + E_o \cos \omega_o t) + be_{g1}^2 \quad (14-66) \end{aligned}$$

and

$$\begin{aligned} i_{b2} &= ae_{g2} + be_{g2}^2 \\ &= a(E_s \cos \omega_s t - E_o \cos \omega_o t) + be_{g2}^2 \quad (14-67) \end{aligned}$$

Since the two terms in ω_s are of the same sign, they will add in the output, and the desired cancellation is not obtained. To prevent feed-through of this type of signal on the mixer grids, which may be the result of feedback from the mixer plate circuit or of direct pickup

on the grid of the mixer. Even though the envelope may be observed with an oscilloscope, the beat note cannot be detected by a linear device. In order to produce the difference or I-F frequency term, we must apply e_p to a nonlinear device. To this end, the mixer may be operated as a plate circuit detector, whose output is proportional to the envelope of e_p . Terman²¹ has shown that the I-F voltage developed across the mixer plate load is

$$E_{I.F.} = g_c E_s Z_L \quad (14-59)$$

provided that Z_L , the plate load impedance, is small compared to the effective plate resistance of the mixer tube operating as a detector. The term g_c is the conversion transconductance, defined by

$$g_c = \frac{\partial(i_b)_{I.F.}}{\partial(e_c)_{R.F.}} \quad (14-60)$$

and in ordinary tubes is roughly one-fourth to one-third of the transconductance g_m . It follows that a high- g_m tube should be used as the mixer.

The mixer, which is used in the R.C.A. 630TS unit, represents an interesting departure from usual design. The push-pull feature of the R-F stage is carried over into both the mixer and oscillator sections. The signal and oscillator voltages are link-coupled into the mixer grid circuit and hence both appear in a push-pull fashion on the two mixer grids, *i.e.*, on the mixer.

$$\begin{array}{l} e_{g1} = e_s + e_o \\ \text{and} \quad e_{g2} = -(e_s + e_o) \end{array} \quad \left. \vphantom{\begin{array}{l} e_{g1} = e_s + e_o \\ e_{g2} = -(e_s + e_o) \end{array}} \right\} (14-61)$$

Since the resultant wave is symmetrical about the zero-voltage axis, the plate currents in the two sections of the 6J6 mixer are in phase and the two plates may be strapped together. This connection gives an important advantage in that any incoming signal of the I-F frequency that arrives at the mixer grids in push-pull fashion, *i.e.*, so that $e_{g1} = -e_{g2}$ at I-F frequency, will cause plate currents in phase-opposition that will cancel in the output circuit.

We may illustrate this analytically, and to simplify the work we shall assume that both sections of the mixer are identical and operating on the "square law" portion of their i_b - e_c characteristics.

²¹ *Op. cit.*, chap. 10.

difference frequency ($f_o - f_s$). This fact may be demonstrated as follows: The difference frequency, ($\omega_o - \omega_s$), is a small percentage of ω_o ; hence we may write

$$\omega_o - \omega_s = \delta$$

$$\text{or} \quad \omega_s = \omega_o - \delta \quad (14-55)$$

Then, from (14-54),

$$\begin{aligned} e_g &= E_s \cos(\omega_o - \delta)t + E_o \cos \omega_o t \\ &= E_o [m \cos(\omega_o - \delta)t + \cos \omega_o t] \end{aligned}$$

$$\text{where} \quad m = \frac{E_s}{E_o} \quad (14-56)$$

This equation may be simplified by expanding the first term within the brackets, thus,

$$\begin{aligned} e_g &= E_o [m(\cos \omega_o t \cos \delta t + \sin \omega_o t \sin \delta t) + \cos \omega_o t] \\ &= E_o [(1 + m \cos \delta t) \cos \omega_o t + (m \sin \delta t) \sin \omega_o t] \end{aligned} \quad (14-57)$$

Our experience in Chapter 12 shows that this expression may be reduced to the general form

$$e_g = V \cos(\omega_o t - \theta)$$

where V is the envelope of the combined voltages. Thus,

$$\begin{aligned} V &= E_o \sqrt{(1 + m \cos \delta t)^2 + (m \sin \delta t)^2} \\ &= E_o \sqrt{1 + 2m \cos \delta t + m^2 \cos^2 \delta t + m^2 \sin^2 \delta t} \quad (14-58) \\ &= E_o \sqrt{1 + 2m \cos \delta t + m^2} \end{aligned}$$

which verifies the statement that the envelope variation occurs at the difference frequency,²⁰ $\delta = (\omega_o - \omega_s)$.

It must be emphatically stated that this change in amplitude of e_g at the beat frequency does *not* mean that the beat frequency is present

²⁰ There is a strong temptation to further simplify (14-58) by setting $m = 1$, for then

$$V = 2E_o \cos \frac{\delta t}{2}$$

which seems to indicate that the envelope variation is at *one-half* the difference frequency. Actually in the television system 100 per cent modulation is not reached because $E_o > E_s$. It should also be stated that the plate detector does not perform satisfactorily when m approaches unity.

of the corresponding R-F component; therefore the relative amplitudes of the several R-F components are preserved in the I-F signal. We also note that the I-F amplitude is proportional to E_o and to k . Thus, in the interests of voltage gain, we want the oscillator voltage as large as the limits of linearity between g_m and E_c' allow. The factor k is simply the slope of the $g_m - e_c$ curve, and for a linear mixer we desire a tube which exhibits a steep g_m v. e_c characteristic.

As a practical matter, E_o is generally large enough so that the variation in g_m is not restricted to the linear portion of the curve and eq. (14-53) is no longer valid. Where this condition exists a point-by-point graphical analysis is required. Figure 14-13 shows two of

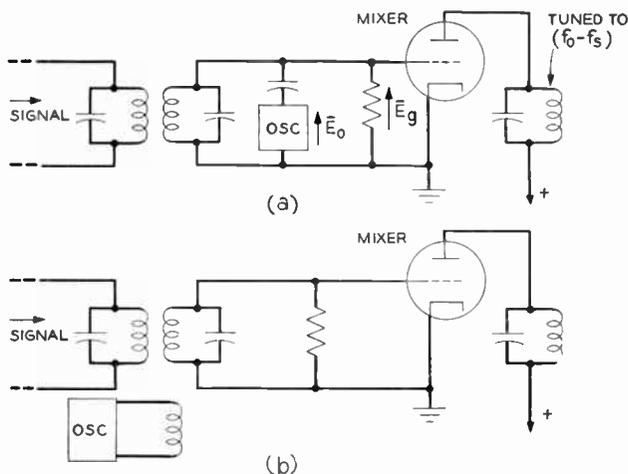


Fig. 14-13. Types of coupling between the local oscillator and mixer. (a) Capacitive coupling. (b) Inductive coupling.

the common methods of coupling the local oscillator to the mixer input. By the application of Thevenin's theorem to the circuit at the left of R_p it may be shown that in either case a fraction of the signal and oscillator voltages appear in *series* across R_p . Letting the primes indicate that fraction of the corresponding quantity, we may write

$$e_g = e_s' + e_o' \quad (14-54)$$

If the two sinusoidal voltages e_s' and e_o' are added point by point, it will be observed that the *envelope* of their sum, e_o , varies at the

the signal and oscillator voltages are so small that operation is confined to the straight-line portion of the curve, the following linear analysis obtains. We assume that the sinusoidal oscillator voltage causes the effective bias voltage to vary in the same manner, thus

$$E_c' = -E_{cc} + E_o \cos \omega_o t \quad (14-49)$$

where

E_{cc} = magnitude of the d-c bias voltage,

E_o = peak oscillator voltage,

ω_o = angular frequency of the oscillator.

Then, since the transconductance is related linearly to the bias voltage, we may write

$$g_m = g_{m_o} + kE_o \cos \omega_o t \quad (14-50)$$

where g_{m_o} is the transconductance at the quiescent point. When a signal voltage

$$e_s = E_s \cos \omega_s t \quad (14-51)$$

is applied to the grid, the alternating component of plate current will be

$$i_p = g_m E_s \cos \omega_s t \quad (14-52)$$

and, substituting for g_m , we obtain

$$\begin{aligned} i_p &= (g_{m_o} + kE_o \cos \omega_o t) E_s \cos \omega_s t \quad (14-53) \\ &= g_{m_o} + \frac{k}{2} E_o E_s \cos (\omega_o + \omega_s) t + \frac{k}{2} E_o E_s \cos (\omega_o - \omega_s) t \end{aligned}$$

The last term in (14-53) contains the difference between the oscillator and signal frequencies and is taken as the I-F or intermediate frequency. This I-F term is separated from the two other terms by the use of a tuned plate load, which has a resonant frequency equal to $f_o - f_s$.

Our previous work has shown that the modulated television signal contains a multiplicity of sideband components, each of a different frequency; hence we may interpret E_s and ω_s in the equation as each sideband component. It follows, then, that the entire R-F signal is reproduced at the output of the mixer, but each frequency has been lowered by an amount equal to the local oscillator frequency.

Inspection of the I-F term in eq. (14-53) shows that the amplitude of each I-F component is directly proportional to E_s , the amplitude

a significant contribution to the noise figure of the whole receiver. Thus, in considering the frequency mixer, which is the second stage, we seek a circuit that has a low noise figure. In general, frequency conversion in a superheterodyne receiver may be accomplished with a single multigrid converter tube, which serves as both local oscillator and mixer, or by means of two separate tubes, operating independently as oscillator and mixer. While the former system finds almost universal use in broadcast-band receivers, it provides poor results in television work, first because the partition noise caused by the several grids is high, and second because the converter tube fails to provide the high degree of frequency stability required in television work. Still another problem may arise when a particularly strong signal is developed by the oscillator section of the converter tube. Space-charge coupling to the signal grid may cause current to flow in that grid and the input circuit will be loaded by the resulting increase in input admittance. For these reasons, television-receiver design employs the alternate scheme of separate oscillator and mixer.

The separate mixer may take on either of two forms, single input or double input. In the former, both the signal and local oscillator output are applied to a single grid of a triode or pentode, while in the latter the two voltages are applied to separate grids in a multigrid mixer tube such as the 6I7. Clearly the latter method must be ruled out because of the multigrid partition noise. We are left with the single-input mixer and we prefer the triode to the pentode because of its lower equivalent noise resistance. In this connection it is interesting to note that the majority of prewar receivers employed a pentode, the 6AC7. The reason for this choice was that during that period the 6AC7 provided the highest value of transconductance ($g_m = 9000 \mu\text{mhos}$) available. We shall see presently that this is an advantage from the standpoint of conversion gain as well as of noise. The recent development of high- g_m triodes has eliminated the need for this choice of a pentode as the mixer; however, pentodes still enjoy wide use because they minimize neutralization problems.

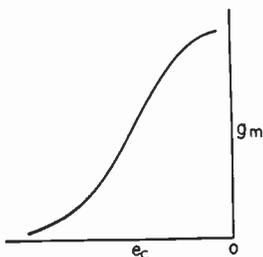


Fig. 14-12. Typical curve showing the variation of transconductance with grid voltage in a triode.

We next consider the operation of the single-input mixer circuit. A typical curve of g_m v. e_c is shown in Fig. 14-12. If both

is a compromise between these three factors. Maximum gain occurs at the center frequency when the coupling is critical.

The choice of tube used in the R-F stage has considerable bearing on the operation. In several of the previous sections we have noted that available power gain and voltage amplification increase, and the equivalent tube noise resistance decreases, as the g_m of the tube is raised. It has also been pointed out that the gain-bandwidth product of the amplifiers varies inversely as the tube's capacitance, the tube figure of merit being $g_m/2\pi C$. Thus we desire a tube with low input capacitance and a high transconductance. It is a fact that the cold input capacitance of a triode is inversely proportional to the grid-to-cathode spacing, while g_m is roughly inversely proportional to this spacing squared; hence by moving the grid closer to the cathode we find that g_m increases faster than the input capacitance and power gain; amplification and figure of merit all increase. Thus special tubes for this type of service have been developed, which have small grid-cathode spacing. Some typical examples of these special tubes are listed below:

| Tube | $C_{ek}(\mu\mu\text{f})$ | $g_m(\mu\text{mhos})$ |
|------|--------------------------|-----------------------|
| 6J4 | 4 | 12,000 |
| 6J6 | 1.6 | 5,300 |
| 6C4 | 1.6 | 2,200 |

Greater control of plate current by the grid voltage may also be had by reducing the spacing between adjacent turns on the grid structure. Thus a close-wound grid will also raise the tube figure of merit.¹⁹

14-5. The Frequency Mixer

It was shown in the discussion on available power gain that the triode R-F stage has a relatively low available power gain; hence the noise generated in the second stage of the television receiver makes

¹⁹ In the type 404A pentode a grid-cathode spacing of 2.5 mils and 400 turns per inch of 0.3-mil diameter wire on the grid are used. In this case, where the electrodes are plane and parallel, the figure of merit is

$$\frac{g_m}{2\pi C} = \frac{K}{2\pi} \sqrt{\frac{J}{a}} \left(1 - \frac{E_{c2}}{\mu^2} \frac{d\mu}{dE_g} \right)$$

where J = current density and a = grid-cathode spacing. See G. T. Ford, "The 404A, A Broadband Amplifier Tube." *Bell Lab. Record*, XXVII, 2 (February 1949).

tuning shaft, the plates of one set being displaced 180° from the plates of the other. Thus, the entire television band of 12 channels is covered by a single rotation of the shaft. Station selection is simplified by a detent mechanism, which drops into properly located slots on a disk tied to the tuning shaft. These slots are relatively wide to give rough tuning. Fine tuning within the width of the slot, that is, over one channel, is afforded by a vernier drive which is geared to the main shaft. An untuned input circuit is used, a fact which may be verified by inspection of the figure: only three tuning sections are present, one each for the R-F output, the local oscillator, and the mixer input. A single-ended version of the capacity-tuned R-F unit is shown in Fig. 14-11b.

By and large our discussion this far has been concerned with the mechanical features of tuning the R-F stage. We consider next the electrical characteristics of the stage. Since most tuners employ an untuned input circuit, the R-F selectivity is provided by the load between the R-F and mixer tubes. In general terms, the circuits for each channel are designed so that the center of the selectivity characteristic lies on the arithmetic mean of the R-F channel. This may be identified as the f_0 used in Chapters 12 and 13. The bandwidth is adjusted so that the half-power points of the response curve coincide approximately with the upper and lower limits of the R-F channel.

The coupling between the R-F and mixer stages may be used to a good advantage in controlling the shape of the selectivity characteristic. From our previous work we know that the response characteristic should be flat and the phase shift linear in the pass band if distortion is to be minimized. Furthermore, the sides of the amplitude response curve should be steep (good skirt selectivity) outside the pass band in order that the selectivity be good. None of these conditions is met by a single-tuned circuit. Since tuned circuits are used both as the R-F plate load and the mixer input, the advantages of double-tuned circuits may be utilized. By adjusting the coupling between the two tuned circuits to be nearly the critical value, all three factors of response, phase shift, and skirt selectivity are improved. Most constant delay time occurs when the actual coupling is slightly less than critical.¹⁸ The actual value of coupling used in any design

¹⁸ For curves of response, phase shift, and delay time of double-tuned amplifiers see F. E. Terman, *Radio Engineering*, 3d ed. New York: McGraw-Hill Book Company, Inc., 1947, chap. 7.

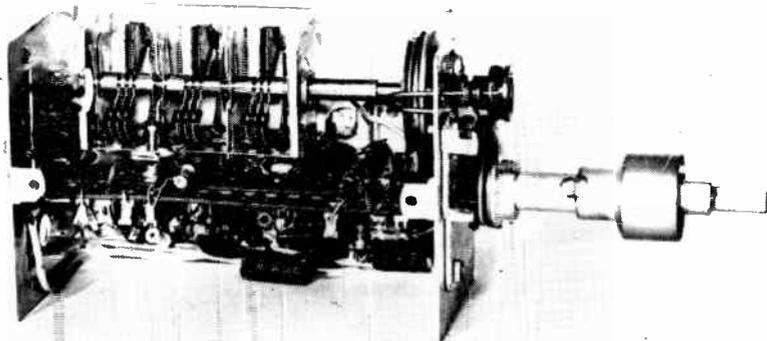
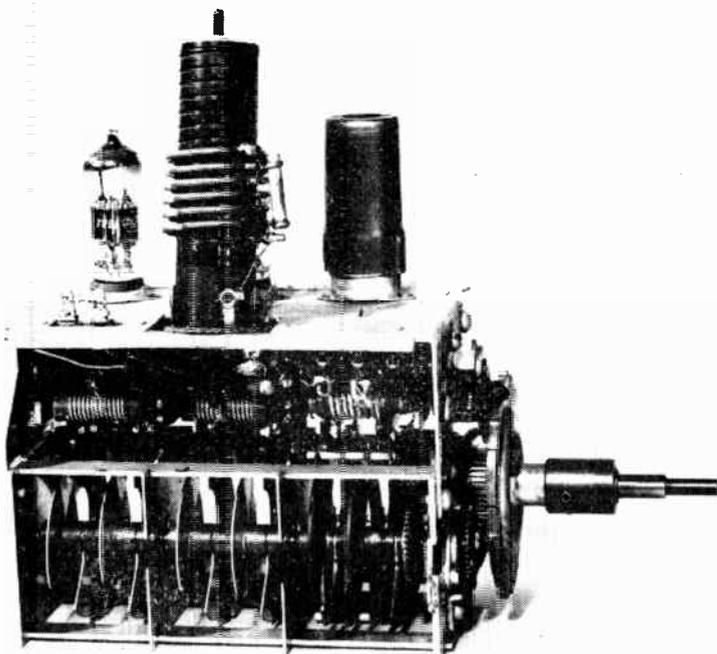


Fig. 14-11. Two forms of continuously tunable front ends that use capacitance as the variable element. (a) Push-pull circuits are used throughout. The tuning condensers may be readily identified. (b) A single-ended unit. Notice the simplification of the circuit. (Courtesy of General Instrument Corporation.)

control which operates on the main tuning shaft. Push-button station selection may be provided by an electric drive on the shaft. The complete Du Mont front end is shown in Fig. 14-23.

The unit shown at *b* in Fig. 14-10 represents a later design of the Inductuner in which the helical coils are replaced by flat spirals. The new design provides considerable saving in space and has certain advantages from the electrical standpoint. The required inductance range is covered in 6 rather than 10 turns, and the frequency v. turns characteristic is more linear.

The Belmont tuning unit also employs a variable inductance for station selection, but permeability tuning is used. Tuning is accomplished by moving tuning slugs into or out of two inductance coils. This motion is effected by means of a threaded tuning shaft and a rack to which the slugs are affixed. As the shaft is rotated, the rack is driven back and forth. It is impossible to obtain the required 4 to 1 inductance ratio with permeability tuning; hence the two sets of coils are provided, one each for channels No. 2 through No. 6, and for channels No. 7 through No. 13. Switching from one set to the other is handled automatically by a switch which is keyed to the tuning shaft. Again no separate fine tuning control is provided for the local oscillator.

A third continuously tunable R-F unit is that produced by the General Instrument Corporation. It represents a complete departure from the other two units which have been described in that it uses capacitance as the variable tuning element. Notice that the use of lumped capacitance in the tuned circuit contradicts the principle of maintaining the gain-bandwidth product by utilizing only stray capacitance. The unit was designed with this fact in mind and means are provided so that essentially constant bandwidth is maintained when the tuning capacitance is varied. This compensation is provided by link coupling between the R-F and mixer stages.¹⁶

Since it is not feasible to build a variable condenser which can tune over the required 4 to 1 frequency ratio, a band-switching scheme, similar to that of the Belmont unit, is employed, the switchover occurring between channels No. 6 and No. 7.¹⁷ As may be seen in Fig. 14-11a the condensers for both ranges are mounted on a common

¹⁶ J. A. Stewart, "Capacity-Tuned Television Tuner." *Radio-Electronic Engineering Edition of Radio and Television News*, 11, 3 (September 1948).

¹⁷ An extended-range unit which covers the 88- to 108-mega-cycle F-M band in addition to the 12 television channels has also been built.

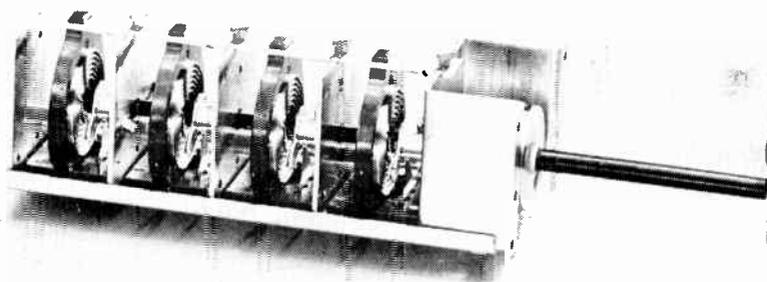
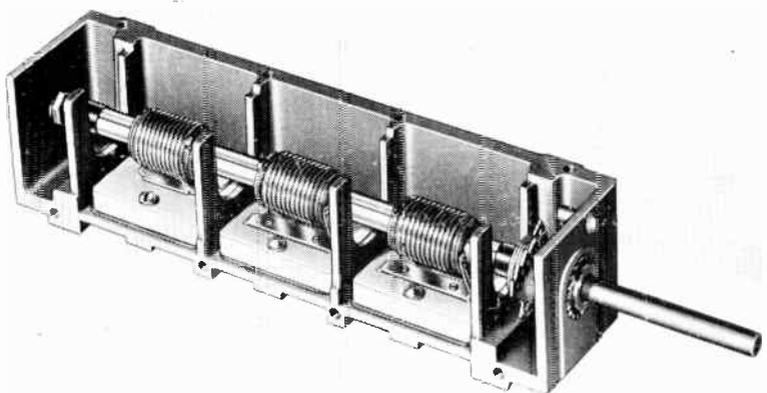


Fig. 14-10. Two forms of the Inductometer. (a) Three-section helical type. (b) Four-section spiral type. (Courtesy of P. R. Mallory and Co., Inc.)

means the inductance tapped off by the finger is made continuously variable from roughly 0.02 to 1.0 microhenry. The complete assembly, as used in television receivers, may be continuously tuned from 44 to 216 megacycles and so covers one amateur, two aeronautical, and the commercial frequency-modulation (88- to 108-megacycle) bands as well as the regular television channels. Details of the circuit design and operation are available in the literature.¹⁵ In contrast to the method mentioned previously, fine tuning is applied to all three sections of the variable inductance by means of a mechanical vernier

¹⁵ See, for example, P. Ware, "Inductive Tuning System for FM-Television Receivers." *Proc. Radio Club of America*, 23, 5 (March 1946).

of the untuned input is that the antenna lead-in is terminated in its characteristic impedance. To this end R_1 and R_2 add up to 300 ohms, the nominal characteristic impedance of a popular form lead-in line.¹⁴ With the receiver end of the lead-in properly terminated, a mismatch of 5 to 1 at the antenna may be tolerated. We shall see in Chapter 16 that this tolerance of antenna match is extremely helpful. In the later R.C.A. models, such as the 9T240, whose front end is shown in Fig. 14-8, the push-pull connection feeding triodes has been abandoned in the R-F and converter stages for the less expensive single-ended connection which feeds pentodes of the 6AG5 type. Notice that additional selectivity has been provided since both the grid and plate circuits of the R-F stage incorporate artificial lines for tuning. The input also has provisions for a balanced 300-ohm or an unbalanced 72-ohm antenna lead-in.

In those receivers that employ fixed-tuned circuits which are changed channelwise by some form of switch mechanism, no fine tuning control is used in the R-F stage proper. On the contrary, the fine control is provided by a small variable capacitor, shunted across the oscillator tank circuit where a high gain-bandwidth product is not required. The criterion for adjusting the fine tuning control on the local oscillator is discussed in a later section. It should be stated that channel-to-channel tuning for both the local oscillator and the mixer input is handled in the same fashion as in the R-F stage.

We have stated that the large majority of television set manufacturers use the fixed-tuned-circuit switch-mechanism system of tuning. Three notable exceptions are Du Mont, Belmont, and General Instrument Corporation, each of which utilizes some form of continuous tuning. One of the chief advantages afforded by these methods is that the need for switch contacts, which can cause poor-contact trouble, is held to a minimum. The basic Inductuner, which is the tuning unit in the Du Mont equipment, is shown in Fig. 14-10. The unit at *a* comprises three variable inductances, one each for the R-F, oscillator, and mixer sections. Each inductance comprises 10 turns of wire on the insulated shaft. Contact to the coil is made by a metal finger, which is free to move on a trolley along the axis of the coil. Thus, as the tuning shaft is rotated, the coil rotates, causing the finger to move progressively along the entire coil length. By this

¹⁴ E. O. Johnson, "Development of an Ultra Low Loss Transmission Line for Television." *RCA Review*, VII, 2 (June 1946).

by R.C.A. in their receivers and is illustrated in Fig. 14-8. As may be seen from the diagram, the various sections of line consist of inductance of one or two turns and, in some cases of a short link, of less than one turn,¹² each section being soldered between adjacent contacts on the rotary station-selector switch. Selection of the proper length of line is obtained by rotating the shorting bar (shown between contacts 2) to the proper position.

In the earlier R.C.A. 630TS receiver, a push-pull front end was used. The input portion of this unit is shown in Fig. 14-9. The

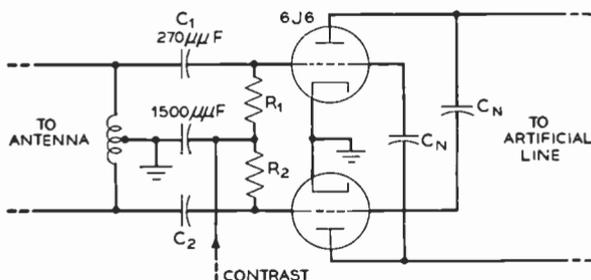


Fig. 14-9. The balanced input of an R.C.A. 630TS receiver.

presence of the push-pull stage allowed neutralization to be accomplished quite readily by means of cross-connecting the plates and grids of the tube. The neutralizing condensers were chosen to be approximately equal to the nominal grid-plate capacitance, which for the 6J6 is $1.6 \mu\mu\text{f}$.

Another advantage occurs from the push-pull connection. The antenna lead-in of the balanced type may be connected directly to the input. Two $270\text{-}\mu\mu\text{f}$ condensers, which were located between the lead-in and grids, serve to isolate the bias (from the contrast control) from the ground. The input circuit is virtually untuned, which is rather poor from the point of noise, but the over-all design holds the noise power to roughly 12 to 14 decibels above the thermal limit.¹³ Some small degree of selectivity is provided by T_1 , which tends to short out the lower frequency signals. The advantage

¹² For the short links, the effective value of L is computed on the basis of the distributed inductance along the link. The concept of "turns" as such becomes meaningless.

¹³ A. Wright, "Television Receivers." *RCA Review* VIII, 1 (March 1947).

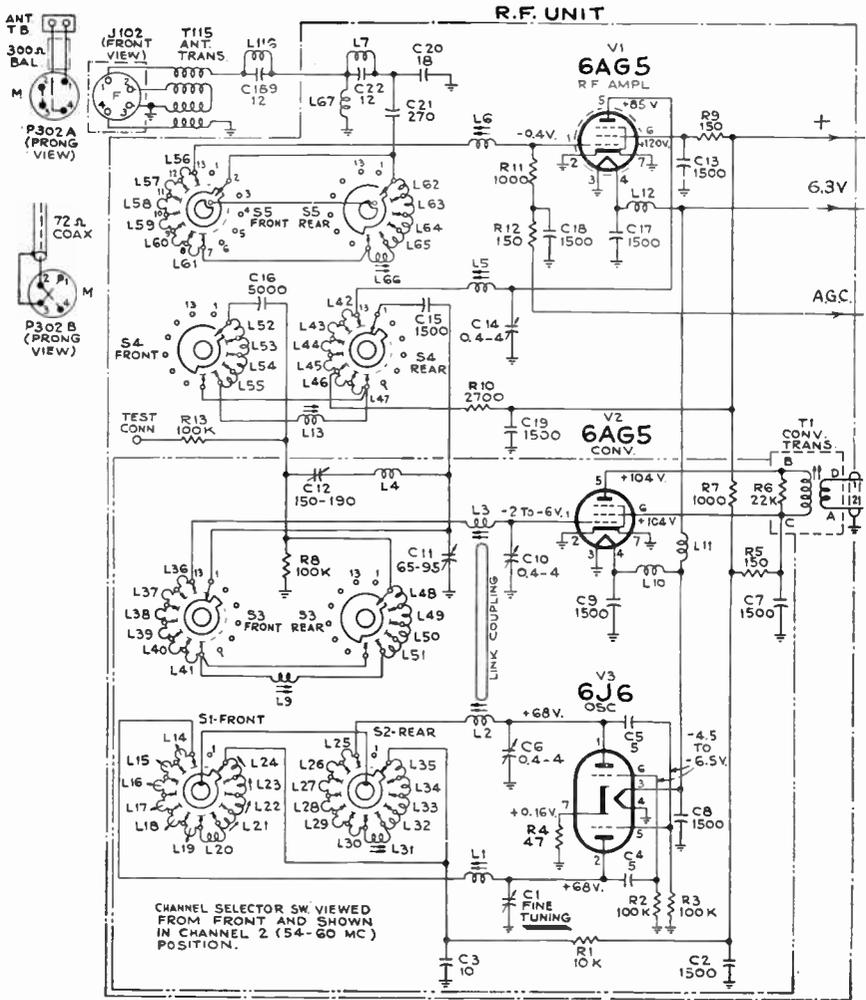


Fig. 14-8. The front end of an R.C.A. receiver. Tuning is provided by artificial lines, which are mounted on the switch wafers S-1 through S-5. Fine tuning of the local oscillator is accomplished by varying C1. (Courtesy of Radio Corporation of America.)

of the switch positions. Basically, of course, some form of L - C circuit is used, which may be tuned for the proper channel by presetting either L or C . We shall see in Chapter 15, however, that the gain-bandwidth product of an amplifier which employs a tuned load is inversely proportional to the total shunt capacitance in the tuned circuit; hence in common practice no lumped capacitance is used, and C is held to a minimum by utilizing only the circuit strays. The resonant frequency is determined by the value of inductance, and fine tuning is accomplished by moving a tuning slug in the inductance coil. In less expensive sets no tuning slug is provided and fine tuning is accomplished by squeezing together or separating the turns in the inductor. A circuit which employs the lumped- L stray- C form of resonant circuit is illustrated in Fig. 14-7.

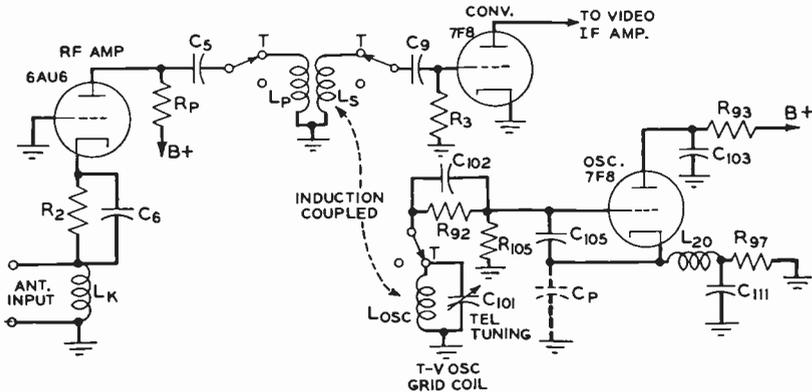


Fig. 14-7. Circuit diagram of a receiver front end that employs lumped inductance and stray capacitance for the tuned circuits. The blank switch contacts may be connected to components tuned for the commercial frequency modulation band. (Courtesy of General Electric Company.)

An alternate form of tuned circuit uses a resonant artificial transmission line in place of the lumped- L stray- C combination. We have already seen that a quarter-wave short-circuited transmission line behaves like an antiresonant circuit. The difficulty in using such a line in a television receiver is that it would be of excessive length, but this may be overcome by using an artificial line consisting of lumped L and stray C . This form of tuning system has been used

antenna, the set radiates a signal at local oscillator frequency which can cause considerable interference in other receivers in the vicinity.¹⁰ Fortunately, the large majority of present-day receivers do provide an R-F stage employing a vacuum tube, and these effects have been reduced considerably.

A large number of commercial television receivers employ a single-ended, rather than a push-pull, input stage. This brings about certain design difficulties because most receiving antennas work into a balanced-to-ground transmission line. Thus, some form of balanced-to-unbalanced transforming circuit is required. The problem is further complicated because a reasonably good impedance match must occur at the junction of the antenna lead-in and the receiver input terminals in order to minimize reflections resulting from an impedance mismatch. A number of solutions to the problem have been evolved and may be seen in the various circuit diagrams in this chapter. One word of caution must be mentioned. In those stages where the cathode operates above ground, the use of R-F chokes in the heater leads of the input tubes is imperative in order that the cathode not be grounded through the cathode-to-heater capacitance. Comparisons of the balanced and unbalanced input stages are available in the literature.¹¹

The problem of tuning the television receiver warrants special attention because of the extremely wide range of frequencies which is involved. An inspection of the twelve television channels listed in section 9-2 shows that the range of frequencies to be covered extends from 54 to 216 megacycles with two large gaps, one from 72 to 76 megacycles between channels No. 4 and No. 5, and the second from 88 to 174 megacycles between channels No. 6 and No. 7. Clearly, the problem of tuning over the entire 12 channels with a single, tuned antiresonant circuit is out of the question and special methods have been devised to overcome this difficulty. The large majority of manufacturers use separate tuned circuits, one for each channel, which may be selected by means of a rotary switch, push buttons, or some form of rotating turret mechanism.

It is interesting to note what forms of tuned circuits are used in each

¹⁰ E. W. Herold, "Local Oscillator Radiation and Its Effect on Television Picture Contrast." *RCA Review*, VII, 1 (March 1946).

¹¹ R. M. Cohen, "Radio-Frequency Performances of Some Receiving Tubes in Television Circuits," *RCA Review*, IX, 1 (March 1948).

Thus, the total noise at the receiver input is $N_A + N_i = 2kT \Delta f$ and the noise figure for the noiseless receiver is

$$F = \frac{N_A + N_i}{N_A} = 2 = 3 \text{ decibels} \quad (14-48)$$

The value of 3 decibels represents the minimum value which can ever be obtained with the lead-in line properly terminated. In practical receivers the noise factor may rise to as much as 20 or 30 decibels higher than this theoretical limit.

We have here the basis of one of the most important concepts of radio transmission. For satisfactory reception of the transmitted signal, the antenna must deliver to the receiver a signal which exceeds the receiver noise by a factor equal to the permissible signal-to-noise ratio. For signals below this value, the signal is lost in the noise and is of no value. Notice that at a fixed distance from the transmitter the actual signal-to-noise ratio may be improved by (1) raising the transmitter power or (2) lowering the noise figure of the receiver. Raising the transmitter power has the advantage of overriding noise generated in the receiver and outside noise as well, but the cost is high. In the absence of outside noise a 5-decibel improvement in the receiver noise figure has the same effect on signal-to-noise ratio as a 3 to 1 increase in power at the transmitter and the unit cost is low. The need for low-noise design in the receiver is self-evident.⁸

We next consider the several component stages of the television receiver.

14-4. R-F Stage⁹

Until early 1947 the chief problem in regard to the R-F stage which confronted television receiver manufacturers seemed to be whether or not the stage should be included. Several sets of this period appeared in which the antenna fed directly into the converter stage, which is poor practice for several reasons. First, the transconductance of a converter is invariably low so the noise figure of the receiver tends to be poor. Second, image rejection is very poor. Third, with no unilateral isolation provided between the local oscillator and

⁸ For additional information on the effect of noise in receiver design see W. A. Harris, "Some Notes on Noise Theory and Its Application to Input Circuit Design." *RCA Review*, IX, 3 (September 1949).

⁹ A. D. Sobel, "Television Front Ends." *Electronics*, 21, 9 (September 1948).

Grounded Cathode

$$A = \frac{-g_m}{g_p + Y_L} \quad (14-43)$$

Cathode Follower

$$A = \frac{g_m}{g_m + g_p + Y_L} \quad (14-44)$$

Grounded Grid

$$A = -\frac{(g_m + g_p)}{g_p + Y_L} \quad (14-45)$$

Inspection of these expressions shows that of the three the grounded-grid connection yields the greatest voltage amplification. This is quite consistent with our results for the grounded-grid stage in Chapter 13, where we observed that the driving source itself contributes a component to the plate current. Couple this advantage of high gain with that of no neutralization and it is easy to see why the grounded-grid stage seemed to be excellent for a receiver input stage. A special tube, the 6J4, ($g_m = 12,000 \mu\text{mhos}$) was designed especially for use in the grounded-grid input. Cognizance of the low available power gain which the circuit provides caused it to be abandoned in later designs in favor of other connections that provide a lower noise figure.

14-3. Minimum Noise Figure of an Ideal Receiver

We may define an ideal receiver as one in which no noise sources are present. From this definition and our previous work we may calculate the theoretical minimum noise limit of a receiver. We consider the antenna of radiation conductance G_A , connected to a lossless transmission line of characteristic admittance

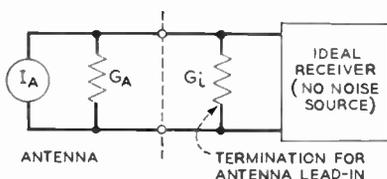


Fig. 14-6. Antenna termination with the ideal or noiseless receiver.

$$G_o = \frac{1}{R_o} = G_A \quad (14-46)$$

The line is properly terminated by setting $G_i = G_A = G_o$. Since the line is lossless, it contributes no noise and the system behaves as if the antenna were connected directly to the input terminals of

the receiver. From our previous work the available noise power from the antenna is

$$N_A = kT \Delta f$$

The terminating admittance G_i also has an available noise

$$N_i = kT \Delta f \quad (14-47)$$

grounded-grid stages following in that order. Therefore, as far as localizing noise in the first stage is concerned, the conventional stage is superior.

Other problems arise that do not appear in the analysis which has been given, namely, neutralization and broadband effects. Invariably with the conventional triode stage some form of neutralization is required in order to prevent oscillation caused by feedback through the grid-plate capacitance. In this respect, the two other stages are superior. Satisfactory means for neutralizing the grounded-cathode stage have been developed, however, as will be seen in the next section. Furthermore, the trend has been toward the use of pentodes to eliminate the need for neutralization.

The second factor causes quite a problem. In all broadband systems, such as in television where the pass band must cover a full 6-mega-cycle band, the tuned circuits must be loaded to provide the necessary low value of Q . Under this condition the value of Γ_B is low and the design of the second stage for low noise is of utmost importance. Thus, triodes are also used rather than pentodes in the second stage. A brief consideration of the fact that each stage may be connected any one of three ways, *i.e.*, grounded-cathode, cathode follower, or grounded-grid, shows that there are nine possible ways in which two triode stages may be connected in cascade. Among others, Wallman and his associates have investigated these nine possibilities and have recommended a "cascode," consisting of a grounded-cathode triode followed by a grounded-grid triode, as that combination which yields the lowest noise figure and at the same time provides the high voltage gain and stability which may be obtained with a pentode.⁷ The use of the cascode amplifier has not been adopted for television receivers. The general practice has been to utilize a single pentode R-F stage, which feeds directly into the converter tube that operates in the grounded-cathode connection.

It is interesting to note from an historical point of view that during the war considerable attention was devoted to the grounded-grid stage as the input for radar receivers, which in some cases are quite similar to television receivers. The voltage amplification of the three triode connections expressed in terms of admittances are listed below:

⁷ H. Wallman, A. B. Maenee, and C. P. Gadsden, "A Low-Noise Amplifier." *Proc. IRE*, **36**, 6 (June 1948).

network B will be the sum of all the available input powers calculated, or

$$N_{B'} = N_A + N_i + N_j' + N_s' \quad (14-37)$$

By eq. (14-14) the noise figure of the amplifier is

$$F_B = \frac{N_{B'}}{N_A}$$

which reduces to

$$F_B = 1 + \frac{G_i}{G_A} + \frac{|Y_A + Y_i|^2}{G_A} \left(\frac{G_L}{g_m^2} + R_t \right) \quad \left. \begin{array}{l} \text{GROUNDED-} \\ \text{CATHODE} \\ \text{STAGE} \end{array} \right\} \quad (14-38)$$

F_B and Γ_B for the grounded-plate and grounded-grid connections of the triode may be calculated in a similar manner. The results are given below.

$$\Gamma_B = \frac{g_m^2 G_A}{|Y_A + Y_i|^2 (g_m + g_p + G_L)} \quad \left. \begin{array}{l} \text{CATHODE} \\ \text{FOLLOWER} \\ \text{STAGE} \end{array} \right\} \quad (14-39)$$

$$F_B = 1 + \frac{G_i}{G_A} + \frac{|Y_A + Y_i|^2}{G_A} \left(\frac{G_L}{g_m^2} + R_t \right) \quad (14-40)$$

$$\Gamma_B = \frac{(g_m + g_p)^2 G_A}{[(G_A + G_i)g_p + (g_m + g_p + G_A + G_i)G_L](g_m + g_p + G_A + G_i)} \quad (14-41)$$

$$\approx \frac{G_A}{G_L} \quad \text{for } |Y_L| \gg g_p, \quad \left. \begin{array}{l} \text{GROUNDED-GRID} \\ \text{STAGE} \end{array} \right\} \quad (14-41a)$$

$$F_B = 1 + \frac{G_i}{G_A} + \frac{G_L}{G_A} + \frac{g_m^2 |Y_A + Y_i|^2}{(g_m + g_p)^2 G_A} R_t \quad (14-42)$$

NOTE: The third term in (14-42) is based on the approximation of equation (14-41a).

Since we have the expressions for available power gain and noise figure for the three triode connections, we are in a position to compare their relative merits as an input stage. Comparison of the several equations which are listed above shows that for identical values of circuit parameters all three connections yield essentially the same value of noise figure. Thus, no particular advantage is offered by any one of the three on that score. In regard to available power gain, a different situation exists. The conventional grounded-cathode stage yields the greatest value of Γ_B , with the cathode follower and

manner. Thus the mean squared Johnson noise current in the plate circuit is⁶

$$\overline{i_j^2} = 4kTG_L \Delta f \quad (14-31)$$

and the available Johnson noise power at the output terminals of the amplifier will be

$$N_j = \frac{\overline{i_j^2}}{4(g_p + G_L)} = \frac{kT \Delta f G_L}{(g_p + G_L)} \quad (14-32)$$

This quantity may also be referred to the input terminals by the use of (14-13); hence the available Johnson noise at the input of the amplifier is

$$N_j' = \frac{|Y_A + Y_i|^2 G_L}{G_A g_m^2} kT \Delta f \quad \begin{array}{l} \text{JOHNSON NOISE} \\ \text{IN } G_L \end{array} \quad (14-33)$$

The two remaining noise sources, which contribute to the total noise at the amplifier input, are the antenna itself and the conductive component of Y_i . The available power caused by the former is by (14-6)

$$N_A = kT \Delta f \quad \begin{array}{l} \text{JOHNSON NOISE} \\ \text{IN } G_A \end{array} \quad (14-34)$$

The mean squared Johnson noise current due to G_i will be

$$\overline{i_i^2} = 4kT \Delta f G_i \quad (14-35)$$

and, as may be seen from Fig. 14-5a, this current generator may be moved to the left of the dotted line to become part of the signal source, and hence will have an available power

$$N_i = \frac{\overline{i_i^2}}{4G_A} = \frac{G_i}{G_A} kT \Delta f \quad \begin{array}{l} \text{JOHNSON NOISE} \\ \text{IN } G_i \end{array} \quad (14-36)$$

Then the total available noise power at the input of the amplifier

⁶ Equation (14-31) is the constant-current equivalent of the constant-voltage expression for Johnson noise. This may be verified from eq. (6-13) and Fig. 14-2. From the diagram

$$i = \frac{e}{R_G}$$

then (6-26) becomes

$$\overline{i^2} = \frac{e^2}{R_G^2} = \frac{4kTR_G \Delta f}{R_G^2} = 4kTG_G \Delta f$$

By (14-14) the available signal power of the antenna is

$$S_A = \frac{I_A^2}{4G_A} \quad (14-26)$$

hence, by (14-5), the available power gain of the grounded-cathode stage becomes

$$\Gamma_B = \frac{S_B}{S_A} = \frac{g_m^2 G_A}{|Y_A + Y_i|^2 (g_p + G_L)} \quad \begin{array}{l} \text{GROUNDED-CATHODE} \\ \text{STAGE} \end{array} \quad (14-27)$$

We next consider the noise figure of the same stage. Two sources of noise are present in the stage itself, the shot effect in the plate current⁵ and thermal noise in G_L . We consider these in order. From eq. (6-24) the mean squared current in the plate circuit resulting from the shot effect is

$$\overline{i_s^2} = 2\epsilon i \Delta f P$$

and substituting for $2\epsilon i P$ from (6-31) we have

$$\overline{i_s^2} = g_m^2 R_t 4kT \Delta f \quad (14-28)$$

where R_t is the equivalent noise resistance of the tube. Since the internal conductance of the circuit is $(g_p + G_L)$, we have by (14-4) that the available shot noise at the output terminals of the amplifier is

$$N_s = \frac{\overline{i_s^2}}{4(g_p + G_L)} = \frac{g_m^2 R_t kT \Delta f}{(g_p + G_L)} \quad (14-29)$$

This noise may be referred to the input terminals of the amplifier by application of (14-13), thus

$$N_s' = \frac{N_s}{\Gamma_B}$$

and, substituting for Γ_B from (14-27), we finally have

$$N_s' = \frac{|Y_A + Y_i|^2}{G_A} R_t kT \Delta f \quad \text{SHOT EFFECT} \quad (14-30)$$

The noise due to the Johnson effect in G_L may be handled in a similar

⁵ Flicker effect is assumed to furnish a negligible contribution to the noise. See section 6-5.

admittances are used instead of impedances in order to simplify the algebraic work which follows. Consider first the more conventional grounded-cathode stage. The generator circuit at the left of the dotted line is the receiving antenna. Then, reading directly from the diagram, we have for the grid voltage on the tube

$$E_g = \frac{I_A}{Y_A + Y_i} \tag{14-20}$$

and, by the equivalent plate circuit theorem, the plate circuit of the tube proper may be replaced by an admittance $g_p = 1/r_p$ in parallel with a generator which produces a constant current

$$g_m E_g = g_m \frac{I_A}{Y_A + Y_i} \tag{14-21}$$

Now the amplifier stage of Fig. 14-5*a* may be identified with the four-terminal network, B , shown in Fig. 14-3. Thus, in order to calculate S_B , the available signal power at the output terminals, we terminate the output in a variable load Y_L' , which we vary until maximum power is delivered by the network. The maximum power condition will obtain when Y_L' is the complex conjugate of the net internal impedance $g_p + Y_L$. Under this condition, the total susceptance is zero and

$$\left. \begin{aligned} G_L' &= g_p + G_L \\ \text{where } G_L' &= \text{conductive component of } Y_L' \\ \text{and } G_L &= \text{conductive component of } Y_L \end{aligned} \right\} \tag{14-22}$$

The output voltage will be

$$E_o = - \frac{g_m E_g}{g_p + G_L + G_L'} \tag{14-23}$$

and, substituting from (14-21) and (14-22), we have

$$E_o = - \frac{g_m I_A}{(Y_A + Y_i) 2(g_p + G_L)} \tag{14-24}$$

and the available signal power will be

$$\begin{aligned} S_B &= E_o^2 G_L' = \frac{g_m^2 I_A^2 (g_p + G_L)}{4 |Y_A + Y_i|^2 (g_p + G_L)^2} \\ &= \frac{g_m^2 I_A^2}{4 |Y_A + Y_i|^2 (g_p + G_L)} \end{aligned} \tag{14-25}$$

The last equation is of extreme importance in the design of the front end or R-F section of a television receiver, for note the following: if the first stage has a large value of available power gain, Γ_B , the noise figure of the circuit is essentially equal to that of the first stage alone, and the noise contribution of the following stages may be neglected. If, on the other hand, Γ_B is low, the over-all noise factor is affected by the remaining amplifier stages. We may, therefore, set down a first requirement on the input stage of the receiver: it shall have as high an available power gain as feasible, and hence will generally be an active circuit employing a vacuum tube instead of being a passive circuit.

A second requirement on the input stage follows immediately. If Γ_B is high, F_B is the noise figure of the entire amplifier; hence the input stage should be designed to be as noise-free as possible. We have already seen that because of the partition effect, pentodes generate more noise than similar triodes. We have also seen that high values of g_m reduce tube noise; hence we expect the input stage of the television receiver to employ a high- g_m triode.

In general, a triode may be used in three connections which are named for the tube element that is grounded, *i.e.*, grounded cathode, grounded grid, or grounded plate (cathode follower). We next investigate the noise figure and available power gain for these three types of connection, which are illustrated in Fig. 14-5. Notice that

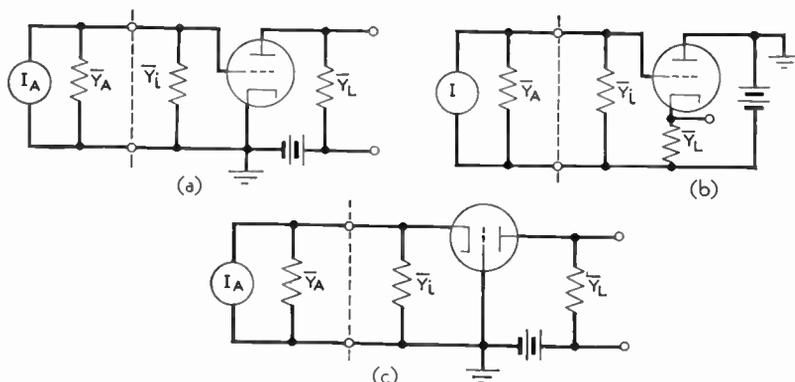


Fig. 14-5. The triode may be connected in three ways. (a) Grounded-cathode amplifier. (b) Grounded-plate circuit (cathode follower). (c) Grounded-grid amplifier.

Notice that if B is a noiseless network, the second term in (14-12) becomes zero, and the total available noise referred to the input becomes simply $N_A = kT \Delta f$. If B is a noise source, the total available noise referred to the input terminals is $N_{B'}$ of eq. (14-13). Therefore (14-14) may be interpreted in the following manner. The noise figure F_B is the ratio of the equivalent input noise power of the actual network B to that of the equivalent noiseless network B' .

Consideration of the work just presented indicates that $N_{B'}$ also consists of two components, the input from the signal source A and the noise generated in B , referred to the input terminals of B . The former is known to be $kT \Delta f$, so the second component may be evaluated by judicious factoring. Thus, from (14-14),

$$N_{B'} = F_B kT \Delta f = (F_B - 1)kT \Delta f + kT \Delta f \quad (14-15)$$

where $(F_B - 1)kT \Delta f =$ the available noise power generated in B , referred to the input terminals of B (14-16)

It should be observed that the whole concept in use here is that we replace the actual network by its noiseless equivalent and increase the available input power by such an amount that the actual available noise at the output remains unchanged. This concept may be extended to the case where two or more stages are cascaded. Thus let the output of network B feed the input of a four-terminal network C , which has a noise figure F_C . Then, by analogy to our previous work, we may refer the noise generated in C to the input terminals of C as $N_{C'} = (F_C - 1)kT \Delta f =$ available noise power generated in C , referred to the input terminals of C (14-17)

and this, in turn, is reflected back to the input terminals of B as

$$\frac{N_{C'}}{\Gamma_B} = \frac{(F_C - 1)kT \Delta f}{\Gamma_B} \quad (14-18)$$

Thus, the two-stage network may be replaced by two noiseless stages, B' and C' , fed by an available noise power

$$N_{B'} + \frac{N_{C'}}{\Gamma_B} = F_B kT \Delta f + \frac{(F_C - 1)kT \Delta f}{\Gamma_B}$$

and, by (14-14), the noise factor of the cascaded pair is

$$F_{BC} = F_B + \frac{(F_C - 1)}{\Gamma_B} \quad (14-19)$$

We have already observed in Chapter 6 that Δf may generally be taken to be the half-power bandwidth of the network. Where one or two single-tuned circuits occur in the network, Δf is 1.57 or 1.22 times the half-power value, respectively.

We now consider the concept of the noise figure F_B of network B . For the circuit of Fig. 14-3 let

$$\left. \begin{aligned} S_A &= \text{available signal power of network } A, \\ N_A &= \text{available noise power of network } A, \\ S_B &= \text{available signal power at the output terminals} \\ &\quad \text{of network } B, \\ N_B &= \text{available noise power at the output terminals} \\ &\quad \text{of network } B. \end{aligned} \right\} \quad (14-9)$$

Then, by definition,

$$\begin{aligned} F_B &= \text{noise figure of network } B \\ &= \frac{S_A/N_A}{S_B/N_B} = \frac{S_A}{S_B} \frac{N_B}{N_A} \end{aligned} \quad (14-10)$$

Since network B contains no signal source, S_B/S_A is precisely the available power gain of the network, and eq. (14-10) may be written

$$F_B = \frac{N_B}{\Gamma_B N_A} \quad (14-11)$$

Since network B is itself a source of noise, N_B/N_A is not equal to Γ_B but rather

$$N_B = N_A \Gamma_B + kT \Delta f \quad (14-12)$$

i.e., N_B consists of an amplified component $N_A \Gamma_B$, developed in network A , plus the component $kT \Delta f$, which is developed in B . We can, however, without changing N_B replace B by a noiseless equivalent network, say B' , of available power gain Γ_B , to which is applied an available input noise power $N_{B'}$, given by

$$N_{B'} = \frac{N_B}{\Gamma_B} \quad (14-13)$$

On this basis the noise figure becomes

$$F_B = \frac{N_{B'}}{N_A} = \frac{N_B'}{kT \Delta f} \quad (14-14)$$

depending upon the nature of the four-terminal network. This fact will be illustrated shortly.

The ratio of W_B to W_A , the available powers, respectively, of networks B and A , is defined as the available power gain, Γ_B , of the network B :

$$\Gamma_B = \frac{W_B}{W_A} = \text{available power gain of network } B \quad (14-5)$$

In the general case both networks of Fig. 14-3 will contain resistance and hence will be sources of Johnson or thermal agitation noise. Furthermore, if they contain vacuum tubes, they will produce additional noise resulting from shot, flicker, and partition effects, which have been described in Chapter 6.⁴ In that chapter we observed that, in each case, tube noise may be replaced by an equivalent resistance placed at the input of a noiseless amplifier. The value of this resistance is chosen so that its Johnson noise, when amplified by the ideal amplifier, will be equal to the output noise of the actual amplifier. Thus all noise sources may be reduced to an equivalent Johnson noise, which by eq. (6-26) has a mean squared voltage of

$$\overline{e_j^2} = 4kTR \Delta f$$

where R is the equivalent noise resistance which remains constant in the noise band Δf . From our definition of available power we see that the available noise power of the equivalent resistance R is

$$N = \frac{\overline{e_j^2}}{4R} = kT \Delta f \quad (14-6)$$

It is interesting to note that N is independent of the value of R .

For the sake of completeness we define Δf , the noise bandwidth of a network. The curve of Γ of a network is plotted against frequency in Fig. 14-4. Then, by definition,

$$\Gamma_{\max} \Delta f = \text{area under } \Gamma \text{ curve} \quad (14-7)$$

$$\text{or } \Delta f = \frac{1}{\Gamma_{\max}} \int_0^{\infty} \Gamma df \quad (14-8)$$

See section 6-4.

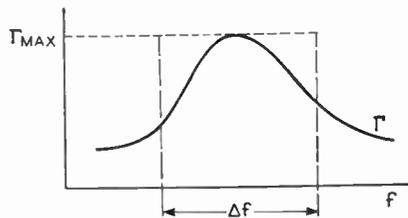


Fig. 14-4. Determination of the noise bandwidth, Δf .

and the power delivered to the load will be

$$W = I_{\max}^2 R_L = \frac{E^2}{4R_o} \quad (14-3)$$

The last three equations are simply a statement of the maximum power transfer theorem, and it follows that W is the maximum or "available" power of the generator. Since the available power is related to E , the generator may also be described in terms of W , its available power, and Z_o , its internal impedance. Of course, this concept may be applied to the generator whether it is a source of signal or of noise.

Still another form of equivalent circuit for a generator may be had by application of Norton's theorem. In this case the generator is represented by its internal impedance, Z_o , shunted by a constant-current generator which produces a current, I , equal to the short-circuit current. The corresponding equivalent circuit is shown at *b* in Fig. 14-2. Again we have by the maximum power transfer theorem that the maximum or available power of the generator will be

$$W = \frac{I^2}{4G_o} \quad (14-4)$$

where G_o is the conductive component of the internal admittance. We see, then, that either generator of Fig. 14-2 may be specified in terms of its available power and its internal impedance (or admittance).

Having defined the available power of a generator, we next consider the "available power gain" of a four-terminal network, which may be either active or passive.

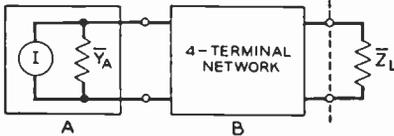


Fig. 14-3. A generator is connected to a four-terminal network.

In Fig. 14-3 such a network is driven by a generator of internal admittance Y_A and of available power W_A . By an extension of our previous work we may replace the entire network to the left of the dotted line by an

equivalent generator of available power W_B and equivalent internal admittance Y_B . Notice that W_B will be the power dissipated in the load only when $Y_L (= 1/Z_L)$ is the complex conjugate of Y_B . It should be noted that Y_B may or may not be dependent upon Y_A ,

impedance termination for the antenna lead-in line. Since the war the industry has become more aware of the problems of noise, local oscillator radiation and interference, and the trend has been toward front ends incorporating a vacuum tube R-F stage. We first consider the problem of noise at the input of the television receiver, for it is at this point that the maximum signal-to-noise ratio of the entire receiver system is established.

14-2. Available Power and Noise Figure^{2,3}

In order to establish the importance of the first receiver stage in setting the signal-to-noise ratio, we shall consider some basic concepts of noise and noise generators. First we consider a voltage source, which by Thevenin's theorem may be replaced, as far as the load is concerned, by a constant voltage E , equal to the open-circuit voltage of the generator, in series with an impedance Z_g , and equal to the internal impedance of the generator. The conventional representation of the generator is illustrated in Fig. 14-2a. If, now, a variable

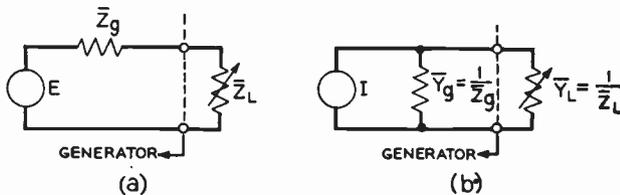


Fig. 14-2. Alternative representations of a generator. (a) Constant voltage (Thevenin's equivalent). (b) Constant current (Norton's equivalent).

impedance, Z_L , is connected to the generator, the current I flowing in the circuit will be maximum when Z_L is the conjugate of Z_g , *i.e.*, when

$$R_L = R_g \quad \text{and} \quad X_L = -X_g \tag{14-1}$$

When this condition is satisfied, the net reactance in the circuit is zero, and the maximum current which flows will be

$$I_{\max} = \frac{E_g}{2R_g} \tag{14-2}$$

² The author is indebted to Dr. Yardley Beers of New York University for an excellent summary of this material.

³ The concepts of available power, gain, and noise figure are given by Harold Goldberg, "Some Notes on Noise Figures." *Proc. IRE*, **36**, 10 (October 1948).

bandwidth product of an amplifier is a constant for a given tube type; hence, if the amplifiers are designed for the full 4.5-megacycle bandwidth, more stages will be required for a given gain. For the 5-in. receiver under consideration this would represent an uneconomical design. Modern television tubes of 10 in. or larger sizes are capable of reproducing the full video signal, and sets which incorporate them should be designed for a 4.5-megacycle bandwidth. In the work which follows we shall use this last figure.

With these two factors considered we may draw a functional diagram of the conventional receiver as shown in Fig. 14-1. The in-

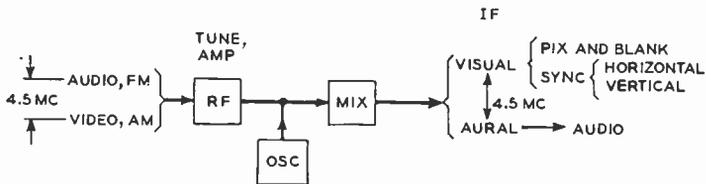


Fig. 14-1. Functional diagram of the conventional television receiver.

coming visual and aural programs are selected, amplified in the R-F stage, and combined with the local oscillator output in the converter to give the I-F visual and aural programs. Notice that the local oscillator operates at a frequency higher than the incoming carriers. This fact shows up in the diagram because the two carriers interchange positions: at the input the audio carrier is higher than the video carrier, whereas in the I-F systems the video carrier has the higher frequency. The reason for this carrier frequency change-over is discussed in a subsequent section.

The two I-F carriers are separated by suitably tuned I-F circuits and the remaining portions of the receiver handle the visual and aural programs in two independent channels. The functional diagram shows the breakdown of the composite signals into their component parts. We next consider the several sections of the receiver, which are illustrated in the block diagram of Fig. 9-1.

It is interesting to note that the complexity and performance of the input stage of the television receiver has been increasing. In the prewar sets, the R-F stage generally took the form of a multiple-tuned passive network, which served to provide selectivity and the proper

which amplify, detect, and convert each program to its proper medium.

14-1. General Discussion

It is well to consider some of the general aspects of receiver design before proceeding to a consideration of the circuit details. For example, we have observed that the R-A system of transmission, which is standard in the United States, requires that the receiver attenuate the incoming signal's lower vestigial sideband in a certain manner, as shown in Fig. 12-8*b*. We must decide in which section of the receiver this attenuation may most conveniently take place.

Clearly the R-A characteristic must be located at some point ahead of the second video detector because it must operate on a modulated, as contrasted to a detected, signal. It follows at once that the required filtering will take place in the R-F section, the video I-F section, or in both of them. Consider these possibilities. The I-F system always works over the same frequency band, regardless of the incoming signal frequency. Thus a single filter, placed in the I-F chain, may be tuned once to serve for all channels. On the other hand, if the filter is placed in the R-F chain, it will have to be retuned each time the set is switched from one channel to another. It is at once apparent that the simpler system will result when the filtering action takes place in the I-F chain alone. We shall discuss means of providing a pass characteristic which has odd symmetry about the 50 per cent response I-F visual carrier frequency point for a range of ± 0.75 megacycle in a later section.

A second point which must be considered in sets which employ a small cathode-ray tube is the maximum video bandwidth for which the set must be designed. The need for this consideration arises because the smaller tubes, operating with low second anode voltages, tend to have a scanning beam whose diameter is disproportionately large. As an example of this, consider a 5-in. tube whose beam diameter is $1/75$ in. For a standard 4 to 3 aspect ratio, the height of the reproduced image will be 3 in. and the maximum number of elements which can be reproduced along a vertical line in the picture will be roughly the height divided by the beam diameter, or $3 \times 75 = 225$ elements. In order that the horizontal resolution be equal to the vertical resolution, this means that a bandwidth of roughly 3 megacycles is required. We shall see in Chapter 15 that the gain-

CHAPTER 14

RECEIVERS

Although we have placed emphasis on the video portions of the television pickup and transmission facilities, it must be remembered that a complete television program consists of two separate groups of signals, one for the aural and one for the visual programs. As we have seen, each of these programs is radiated on a separate modulated carrier, the two carriers being separated by 4.5 megacycles, and in designing a television receiver we must devise a single unit which can accept both signal groups, separate them, and convert each into its appropriate medium of sight or sound. Two receiver design trends have developed during the postwar years, which result in receivers that, for the lack of more precise terms, we may designate as “conventional” and “intercarrier.” In the early part of the chapter we shall consider the former type; the intercarrier receiver is discussed at the end of the chapter.

It will be realized that a large number of manufacturers are active in the field¹ and it is not practicable to consider the many variations which are used in receiver design. We shall, therefore, confine our study to general principles and illustrate them with circuits used in representative commercial receiver models.

THE CONVENTIONAL RECEIVER

The first type of receiver to be considered is the conventional type whose functional diagram is shown in Fig. 9-1. As described in Chapter 9, it is a special form of superheterodyne receiver which comprises a common front end or R-F and converter section which selects the visual and aural programs from a single transmitting station, amplifies, and converts them to the intermediate or I-F frequency band; and two separate I-F and low-frequency systems

¹ As of the summer of 1949, television receivers were being produced by over 130 companies in the United States. Directory of Television Receiver Manufacturers, *Radio and Television News*, 41, 5, 46 (May 1949).

If, on the other hand, the generator and load are interchanged, the bridge will no longer be balanced and the meter reading depends upon the flow of energy from right to left along the transmission line. Thus the bridge circuit is sensitive to the direction of energy flow in the line. If the line is improperly terminated and reflections occur, the meter reading will depend upon the energy in the reflected wave only.

It may be shown quite readily that the incident wave energy will be measured if R and the meter are interchanged. Thus if a switching arrangement in the bridge is provided, the power in the incident and reflected waves, P_i and P_r , respectively, may be measured and the corresponding standing-wave ratio will be

$$\rho = \frac{1 + \sqrt{\frac{P_r}{P_i}}}{1 - \sqrt{\frac{P_r}{P_i}}} \quad (13-69)$$

and noting the corresponding voltmeter readings. When these two voltages are known, ρ may be calculated from their ratio.

Notice that the presence of the slot which must accommodate the movable probe in the outer coaxial conductor precludes the use of a pressurized coaxial line. Since it is desirable to maintain dry air or nitrogen under pressure in the line to reduce the danger of arc-over, we seek an alternative technique, which utilizes a fixed probe. If such a scheme is possible, the probe may be fixed in position with an insulating, pressure-tight collar and gasket, but some technique other than that of measuring $|E_{\max}|$ and $|E_{\min}|$ must be found in order that the standing-wave ratio may be determined.

Since the voltage maxima and minima along the transmission line are the direct result of the incident and reflected waves, it should be possible to measure the power in each of these components separately and then to determine ρ from these measurements. What is required, then, is a device which will measure the flow of energy in one direction along the line. Several forms of directional coupler have been devised which meet this requirement. One type, developed by Morrison and Younker is illustrated in Fig. 13-39*b*.^{35,36,37} The small slot in the outer conductor of the coaxial line behaves like a lumped inductance in series with the outer conductor and the probe is capacitively coupled to the center conductor. The whole measuring circuit may therefore be redrawn as a Maxwell bridge, as shown at *c* in the diagram. Consider the operation of the circuit when the transmission line is properly terminated in its characteristic impedance, R_0 . C may be adjusted by varying the probe-to-center conductor spacing until

$$RR_0 = \frac{L}{C} \quad (13-68)$$

which specifies the balanced condition of the bridge. When this condition obtains, no current flows through the meter, which will read zero. In other words, when the bridge satisfies (13-68), energy flowing from left to right along the transmission line produces no reading on the meter.

³⁵ *Op. cit.*

³⁶ For basic considerations in directional couplers see W. W. Mumford, "Directional Couplers." *Proc. IRE*, **35**, 2 (February 1947) and Correction, **37**, 6 (June 1949), p. 625.

³⁷ For other forms of directional coupler see F. E. Terman, *Radio Engineering*, chap. 4.

these techniques is available in the literature.^{30,31,32,33} We shall, however, consider one test which is quite representative, namely, that of the standing-wave ratio on the transmission lines which interconnect the transmitter, filter, and antenna system.

The standing-wave ratio, ρ , has been defined in eq. (13-67) and may be determined by measuring the voltage maximum and minimum along the transmission line. This may be done by the familiar slotted-line technique,³⁴ which requires that a section of the line, at least $\lambda_0/4$ in length, have its outer conductor slotted to allow the introduction of a probe into the coaxial line field as shown in Fig. 13-39a. Since the voltmeter reading is proportional to the voltage across the coaxial line at the point where the probe is inserted, $|E_{\max}|$ and $|E_{\min}|$ may be determined by sliding the probe along the slot

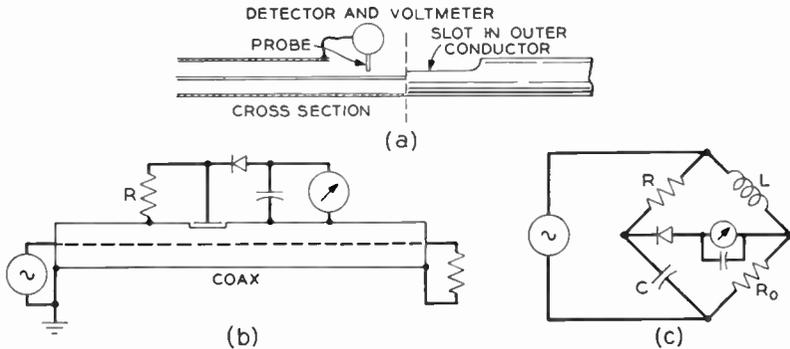


Fig. 13-39. Determination of voltage standing-wave ratio on a coaxial line. (a) The slotted line may be used for measuring ρ . (b) A directional coupler may also be used to determine ρ . (c) Equivalent bridge of the directional coupler. (Courtesy of *Proc. IRE*.)

³⁰ *Standards on Television*, "Methods of Testing Television Transmitters," The Institute of Radio Engineers, 1947.

³¹ *Standards of Good Engineering Practice Concerning Television Broadcast Stations*, Federal Communications Commission, 1945.

³² M. Silver, "Monitor for Television Broadcasting Stations." Proceedings of the National Electronics Conference, Vol. 3, 1947.

³³ J. F. Morrison and E. L. Younker, "A Method of Determining and Monitoring Power and Impedance at High Frequencies." *Proc. IRE*, **36**, 2 (February 1948).

³⁴ M.I.T. Radar School Staff, *Principles of Radar*. New York: McGraw-Hill Book Company, Inc., 1946, chap. 8.

broadband dipoles stacked vertically at a distance of $\lambda_0/2$, as illustrated in Fig. 13-33. The notched sheet equivalent of the dipole pair is quite satisfactory from an electrical standpoint, but mechanically it represents a poor design because the large sheet has considerable wind resistance. This last difficulty is overcome by using the compromise design shown in Fig. 13-37c where the sheet is replaced by a number of horizontal conducting bars. The number of these bars required to maintain the proper electrical characteristics has been determined experimentally to be seven. The resulting batwing exhibits excellent broadband characteristics and commercial design has centered on three batwing sizes, one to cover each of the following frequency bands: 54-66, 66-88, and 174-216 megacycles.

The batwing elements may, of course, be crossed in pairs and stacked vertically in a turnstile pattern to form a super turnstile antenna. A typical unit of this type before erection is shown in Fig. 13-38. The particular antenna shown employs six bays stacked



Fig. 13-38. A six-bay superturnstile antenna before erection. (Courtesy of American Broadcasting Company.)

vertically and provides a power gain of 6.4 as compared to a simple dipole located at the midpoint of the super turnstile array. *i.e.*, 6.4 kilowatts fed to the dipole would produce the same field as 1 kilowatt fed to the actual array.

13-24. Testing

It is beyond the scope of our work to consider the several techniques used in checking transmitter performance. A description of

approximately equivalent to a series resonant circuit. Let it be shunted by a parallel resonant circuit as shown in the diagram. Since the admittance functions of the two circuits are approximately complementary, they may be added to give a resultant admittance $Y = Y_p + Y_s$, which is approximately constant within a limited band range. Generally the compensating shunt network takes the form of a transmission line stub.

13-23. The Super Turnstile²⁹

The super turnstile or batwing antenna has been widely used in the installation of postwar television transmitting stations and is a good example of the broadbanding technique. Consider the development of the radiating elements of this array as shown in Fig. 13-37. At *a*

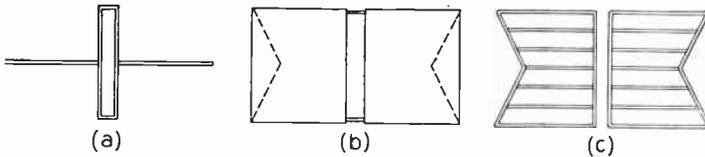


Fig. 13-37. Development of the batwing radiating element of the superturnstile antenna. (a) Dipole with compensating stub supports. (b) The dipole is expanded into a sheet and the vertical edges are notched to provide a better current distribution. (c) The sheet is replaced by a number of horizontal bars to reduce wind resistance.

we have a dipole and two shunt broadbanding stubs. Made of tubing, these stubs serve also as supporting elements for the dipole proper. Since the horizontal bar of each stub is located at a voltage node, these points may be grounded, thereby providing a rigid support. In order to provide additional broadbanding, we lower the Q of the dipole by expanding it into a vertical, conducting sheet, as shown at *b* in the diagram.

If now the vertical edges be notched as indicated by the dotted lines at *b*, the current distribution in the sheet is modified so that maximum current flows along the top and bottom edges and a minimum current flows across the center of the sheet. With this current distribution the radiating sheet is approximately equivalent to two

²⁹ R. W. Masters, "The Super Turnstile Antenna." *Broadcast News*, Number 42 (January 1946).

apex, two ellipsoids end to end, and other similar configurations. An example of a large-diameter, broadband dipole is shown in Fig. 13-35. A difficulty encountered with wide-band antennas of this



Fig. 13-35. A large-diameter broadband dipole formerly used by WCBS-TV. (Courtesy of Columbia Broadcasting System.)

type is that they must be mounted rigidly to withstand wind and ice loading.

The second broadbanding principle involves the use of a corrective

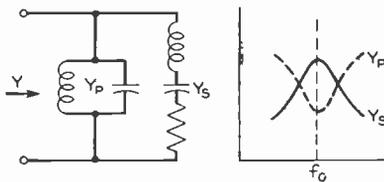


Fig. 13-36. Broadbanding a dipole with a compensating circuit. The admittance, Y_s , of the dipole is shunted by a complementary admittance, Y_p .

network shunted across the input terminals of the antenna. The design of such a network is not easy and generally resort must be made to graphical methods of solution.²⁸ A rather crude idea of the principle involved may be obtained from the following considerations. The radiating dipole element is

²⁸ Staff of the Radio Research Laboratory, *Very High Frequency Techniques*. New York: McGraw-Hill Book Company, Inc., 1947, vol. I. See also F. E. Terman, *Radio Engineering*. New York: McGraw-Hill Book Company, Inc., 1947, chap. 14.

of the radiated pattern will also change. In television transmission, however, the percentage bandwidth, $\Delta f/f_0$, is 0.1 or less, and these effects are generally considered to be of small importance.

On the other hand, the variation of the antenna impedance with frequency is of great concern. It is well known²⁷ that if energy is supplied to an improperly terminated transmission line, only part of the energy will be absorbed in the termination and the rest will be reflected back to the driving source. If a similar mismatch occurs at the source, some of the energy will be reflected again and will appear at some later time at the terminating end as a "ghost" or reflected image. This condition is serious in television work because the ghost shows up as a weak image, displaced to the right of the main image on the receiver's fluorescent screen.

When reflections occur on a transmission line, the incident and reflected waves add to give "standing waves" of current and voltage. The ratio of a voltage maximum to a voltage minimum along the line is known as the voltage standing-wave ratio

$$\rho = \left| \frac{E_{\max}}{E_{\min}} \right| \quad (13-67)$$

and is a measure of the impedance mismatch at the terminal end of the line. When the line is terminated in its characteristic impedance, the voltage is constant along the line

$$|E_{\max}| = |E_{\min}| \quad \text{and} \quad \rho = 1$$

As a practical matter, it has been found that ρ should be 1.1 or less through the whole television transmitter-to-antenna system. This condition requires that the antenna driving-point impedance remain constant within close tolerances throughout the entire 6-megacycle band. In general, two methods may be used to broadband the antenna impedance: the antenna may be designed to have a low Q , or compensating networks may be used. We shall consider these in order.

As the diameter of a radiating element is increased, the element's Q decreases; hence one general principle of broadbanding an antenna is to use a conductor of large diameter. An extension of this principle has resulted in broadband dipoles consisting of two cones apex to

²⁷ See, for example, J. D. Ryder, *Networks, Lines and Fields*. New York: Prentice-Hall, Inc., 1949, chap. 5.

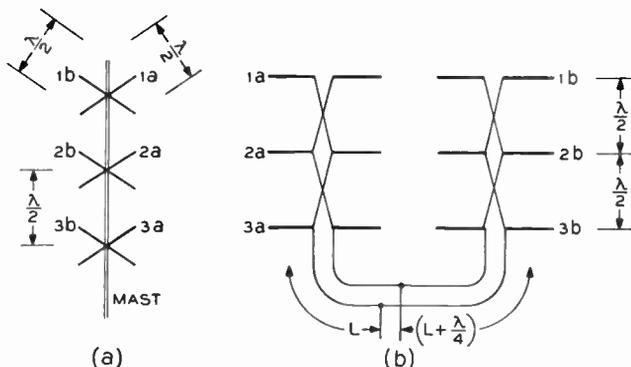


Fig. 13-34. The turnstile antenna. (a) Configuration. (b) Feeder system.

Now consider the phase relationships that exist vertically in a given stack. Dipole $2a$ is $\lambda/2$ farther away from the driving source than dipole $3a$; hence it would normally be driven 180° out of phase from $3a$. This condition is eliminated by transposing the interbay feeders as shown in the diagrams. The phase reversal, resulting from transposition, adds to the 180° shift caused by added line length, and the net shift is zero. Thus the in-phase relationship is maintained in each vertical stack.

13-22. Broadbanding

In discussing the radiation patterns of the antenna systems mentioned in the last few sections we have carefully avoided one of the principal problems present in the television system, that of bandwidth. It is well and good to specify the dimensions of an antenna array in terms of some wavelength, λ , but when the radiated signal covers a frequency band of roughly 6 megacycles, just what λ should be used and what is the performance of the system at wavelengths other than the one which is the basis for the design? To answer this, we must consider the effect of frequency on two factors: the radiation pattern and the impedance match between the transmission line and antenna array. As a starting point we may assume that the array dimensions are specified in terms of λ_0 , the arithmetic mean of the wave bandwidth to be transmitted. At other frequencies in that band, the electrical dimensions of the array change and the directivity

than the wave from the lower antenna in reaching P , the two waves arrive in phase opposition and give $\epsilon_r = \epsilon - \epsilon = 0$. At all other angles in the first quadrant the wave from the upper antenna travels an additional distance $(\lambda/2) \cos \phi$ and the two waves arrive at P with a phase difference β , given by

$$\beta = \pi \cos \phi \quad (13-66)$$

and the two fields may be combined by the law of cosines.

$$\begin{aligned} \epsilon_r &= \sqrt{\epsilon^2 + \epsilon^2 + 2\epsilon^2 \cos(\pi \cos \phi)} \\ &= \sqrt{2} \epsilon \sqrt{1 + \cos(\pi \cos \phi)} \\ &= 2\epsilon \cos\left(\frac{\pi}{2} \cos \phi\right) \end{aligned}$$

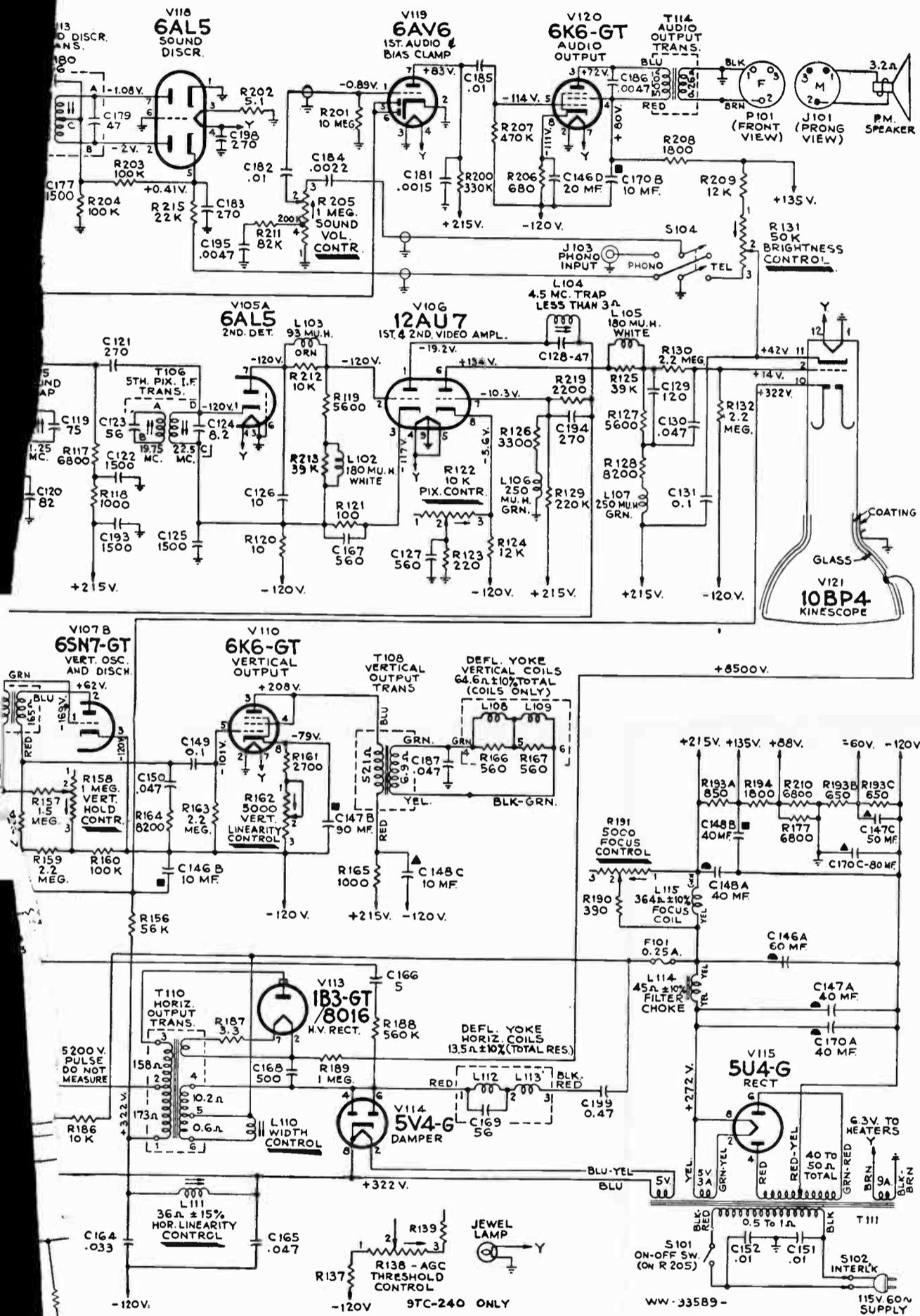
and we have derived eq. (13-65).

13-21. The Turnstile Antenna

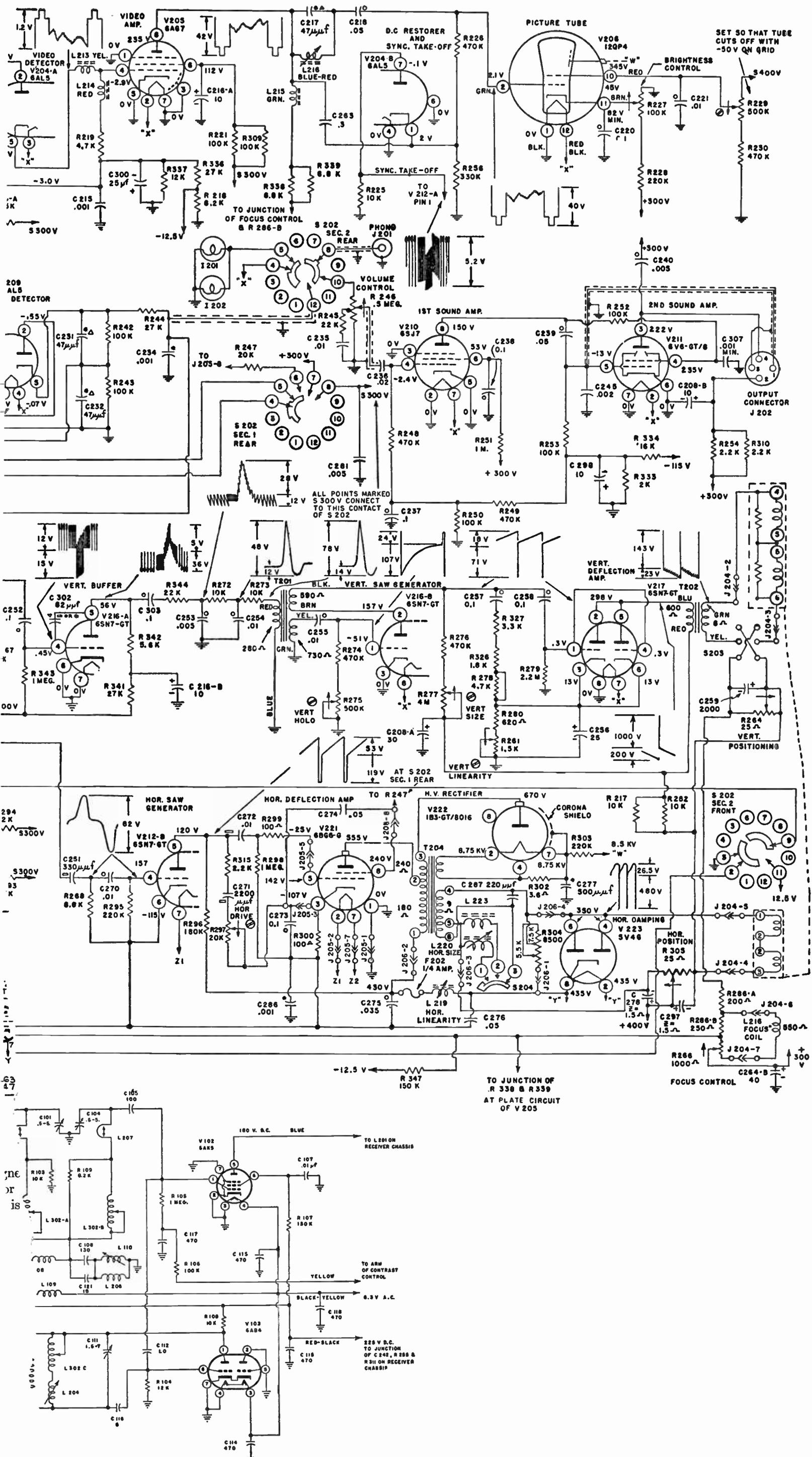
The method just described for increasing vertical directivity by stacking radiating elements above each other may also be applied to the crossed dipoles to give a turnstile antenna.²⁶ Such an array provides essentially an omnidirectional horizontal pattern with good vertical directivity. The latter may be increased even further by stacking more than two bays of crossed pairs. A three-bay turnstile antenna and its feeder system is shown in basic form in Fig. 13-34.

The feeder system shown in the diagram is of particular interest because all the dipoles in a vertical stack (*i.e.*, $1a$, $2a$, $3a$ or $1b$, $2b$, $3b$) must be fed in phase, but the two stacks must be fed in quadrature. These phasing conditions are met by proper adjustment of the feed lines. Since an electromagnetic wave travels with a finite phase velocity, it suffers a shift in phase which is proportional to the electrical distance traveled, a distance of one wavelength corresponding to a shift of 360° . Thus a line of length $\lambda/4$ may be used to delay a signal by 90° . We utilize this fact in the turnstile. By making the main feeder to stack b $\lambda/4$ longer than the stack a feeder, the former stack is driven 90° lagging with respect to stack a , and the required quadrature condition is obtained.

²⁶ Little agreement is present in the literature in the use of the term "turnstile," it being used in some cases to denote a single pair of crossed dipoles. We shall adopt the notation that turnstile refers to pairs of crossed, horizontal dipoles stacked vertically above one another.



tion receiver. A one-tube, VIII, horizontal scanning system is used. in Fig. 14-8. (Courtesy of Radio



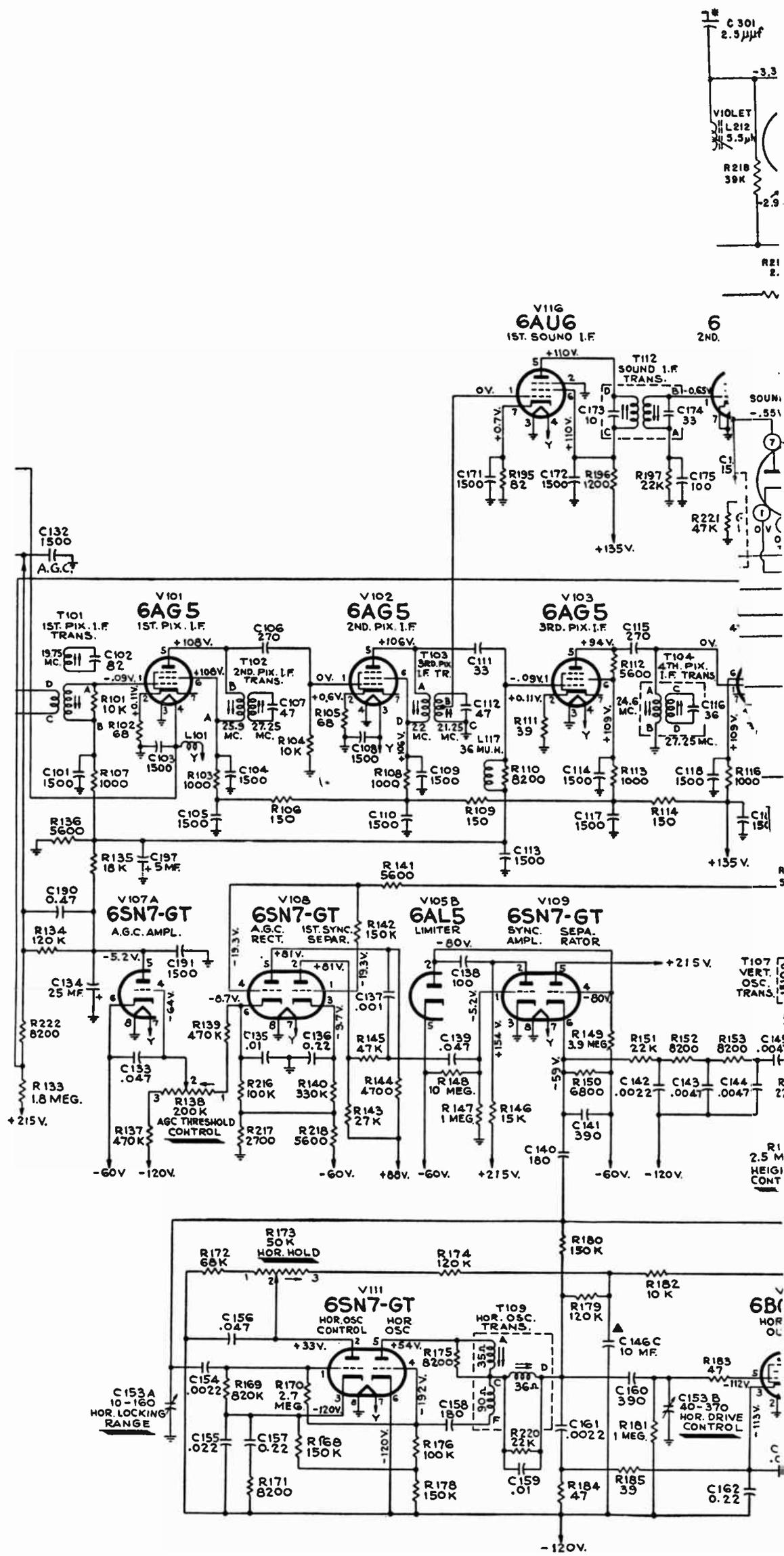


Fig. 14-25. A 10-inch medium frequency-control network for a receiver (The front end of the receiver is the property of the Radio Corporation of America.)

in the R.C.A. 630TS chassis. The circuit shown differs principally from that in the Du Mont receiver in that the discriminator tube is reversed with respect to the sine-wave and sync-signal inputs. It may be seen from Fig. 14-24 that a sine wave is coupled inductively to the plates of the discriminator from the tank circuit of the Hartley oscillator; hence the sine wave appears in phase opposition on the discriminator plates. On the other hand, the sync voltage is fed conductively to the center tap on the transformer primary and so appears in the same phase on the discriminator plates. With the local oscillator running at proper frequency and phase relative to the sync pulses, the voltages on the two discriminator plates will be those illustrated at *b* in the figure. The in-phase sync pulses gate on the two diodes, which conduct equal currents. These currents produce zero voltage across the cathode load resistor and no error voltage is delivered to the reactance tube.

If the oscillator drifts so that it lags the sync pulses, the discriminator plate voltages will have the form shown at *c*. It is clear that the upper diode has a greater applied voltage during the conduction interval than the lower diode; hence the upper plate current will exceed the lower and the net voltage developed across the load will cause pin 1 to be positive with respect to pin 5. This positive error voltage causes the reactance tube to compensate for the oscillator drift. The wave forms at *d* show the conditions which exist when the oscillator voltage leads the sync pulses. It should be noticed that a phasing control is provided on Z204 in Fig. 14-23. The purpose of this control is to allow proper adjustment of the phase of the sine wave fed to the discriminator so that when the oscillator is operating at the correct frequency, the horizontal retrace will occur during the horizontal blanking interval. If the control is out of adjustment, the entire picture is shifted to the left of the CRT tube and a black vertical blanking bar is visible.

Since, as we have seen, the comparison network is sensitive to phase differences in the two compared waves and delivers a d-c voltage proportional to the conduction time of the discriminator, a high degree of noise immunization is provided. Any random noise pulses arriving with the sync may cause the diodes to conduct but they have no instantaneous effect on the sweep, and their effect is averaged out in the d-c voltage delivered to the reactance tube.³⁰

³⁰ A. Wright, "Television Receivers." *RCA Review*, VIII, 1 (March 1947).

One of the chief disadvantages of the horizontal-sweep control circuit just described is that the control operates on a sine-wave oscillator rather than on the horizontal-sweep oscillator itself. To overcome this difficulty means must be provided so that the sine-wave oscillator controls the frequency of the trapezoidal generator. This is accomplished quite simply in the Du Mont receiver of Fig. 14-23, even though a number of additional circuit components are required. The grid signal on the sine-wave oscillator is large enough to drive the plate to saturation so that the output wave is flat-topped, as shown in the diagram. This wave, after being differentiated, produces a positive-going pulse of sufficient amplitude to trigger off the trapezoidal generator, V212B. The remainder of the horizontal-scanning system is similar to those which have been described previously; reaction scanning and a flyback power supply are utilized.

Figure 14-25 shows the diagram of an R.C.A. 9T240, 10-in. magnetic deflection receiver. One of the principal features in which the set differs from those already described is the frequency-control circuit on the horizontal-sweep system. Instead of using the discriminator-sine-wave-oscillator combination just described, the set employs a single control tube and blocking oscillator, V111, whose output feeds the horizontal output tube directly with a considerable saving in the number of components required. The right-hand half of V111 serves as a sine-wave-stabilized blocking oscillator, which, when cut off, allows the required saw-tooth voltage to build up on C161.

The frequency of the blocking oscillator is controlled by the d-c voltage developed between pin 3 of the control tube and ground; this voltage provides the bias for the blocking oscillator. The magnitude of the d-c voltage depends upon the relative phase of three voltages delivered to the grid of the control tube: (1) the sync voltage, which is delivered through C140, (2) the saw-tooth sweep voltage, delivered from C161 through R180, and (3) the integrated flyback voltage delivered through C166. The relative phase of the three components determines the length of time the control tube conducts and hence the d-c voltage on C155. Once again the control of frequency is determined by a d-c voltage and the effect of random noise pulses on the sweep circuit is averaged out.³¹

³¹ An excellent development of the R.C.A. single-tube control circuit is given by J. A. Cornell, "A Single Tube A.F.C. Circuit for TV Deflection Systems." *Radio and Television News*, **43**, 1 (January 1950).

A number of other systems of automatic control of the horizontal-sweep frequency have been developed and it may be expected that even simpler systems will be devised to reduce the number of required components without sacrifice in performance.³² It must be realized, however, that the discriminator or other phase comparator is not essential in providing protection from noise in the horizontal-sweep system. Schlesinger has described a system that employs a high- Q oscillator circuit, which provides excellent noise immunity.³³

14-10. Automatic Gain Control

Automatic gain control (AGC) is the television receiver's counterpart of automatic volume control (AVC) found in the usual broadcast radio receiver. The operation and function of the latter are well known. A d-c voltage which is proportional to the average carrier amplitude of the incoming R-F signal is derived from a detector and is used to automatically control the gain of amplifiers employing tubes of the remote cutoff type. A decrease in average carrier amplitude, resulting from, say, fading, reduces the magnitude of the control voltage and so raises the amplifier gain to compensate for the loss in the R-F signal. Thus the circuit compensates for changes in signal level and provides an output voltage of constant amplitude.

Basically the AVC circuit may be carried over into television work with one important exception. The concept of "average carrier amplitude" of an R-F wave modulated by the composite video signal is really without meaning because, as we have seen, the average carrier amplitude depends upon the background level of the picture component of the modulating signal and so depends upon the modulating signal. In short, the average carrier level is not a reference which may be used to measure the strength of the R-F signal. The obvious solution to this difficulty is to choose the peak, rather than the average, value of the modulated wave as the reference. This is accomplished by using a peak detector as the source of the AGC control voltage, which will then be proportional to the peak sync level in the R-F wave.

³² A number of basic systems based on the comparison of phase of two or more voltages have been described by E. L. Clark, "Automatic Frequency Phase Control of Television Sweep Circuits." *Proc. IRE*, **37**, 5 (May 1949).

³³ K. Schlesinger, "The Locked Oscillator in Television Reception." N. Y. Section, IRE, Television Symposium, January 22, 1949; also *Electronics*, **22**, 1 (January 1949).

The basic circuit of a simple AGC circuit is shown in Fig. 14-26. Separate diodes are provided for AGC and video detection, both being fed by the final I-F amplifier stage. The AGC detector develops a d-c voltage proportional to the peak carrier amplitude and that voltage, in turn, is used to control the bias on the several I-F amplifiers and the input stage of the receiver.

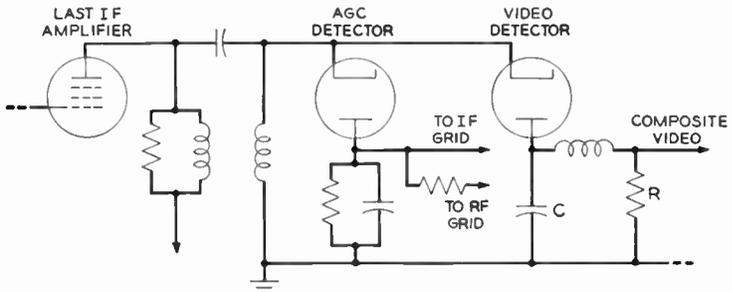


Fig. 14-26. Basic A.G.C. circuit. The A.G.C. voltage is developed by a peak detector.

The basic circuit which has just been described is quite susceptible to noise and so is seldom used in modern receivers. We have already observed that high-amplitude R-F noise impulses extend into the blacker-than-black regions of the detected signal and hence are in a position to upset the AGC system. Notice that the detector load condenser C can charge faster than it can discharge because during charge it is in series with the low detector impedance, while it must discharge through R , which is high in comparison. Thus the effect of any AGC voltage produced by a noise pulse is extended over a relatively long time interval. The situation may be improved, but not corrected, by feeding the modulated signal through a limiter before it reaches the AGC detector. The limiter then will clip the noise peaks to the level of the sync signals.

One's immediate reaction is to improve the noise immunity of the circuit by lowering the time constant RC of the AGC detector load. This is not feasible, however, because the sync pulses themselves occupy only a small percentage of a line interval and the R - C circuit must integrate the control voltage over that interval. What is needed is an AGC circuit which is immune to noise and still fast-acting to the point that the required control of gain is maintained.

In regard to the last point a large problem is presented by low-flying aircraft in the vicinity of the receiving antenna. Under certain circumstances the antenna may receive two signals, one direct from the transmitter and the second reflected from the airplane. As the latter moves in flight, the two signals go through a rapid cycle of reinforcing and canceling, which causes an extremely objectionable flicker or fluctuation of brightness in the reproduced image. To remove this effect an AGC network must be able to respond to signal variations at a very high rate.

Wendt and Schroeder³⁴ have described a number of fast-acting circuits which provide a high degree of noise immunity. One of the newer fast-acting circuits described in the reference is illustrated in simplified form in Fig. 14-27. The sync-negative video signal, which feeds the Kinescope grid, also drives V_1 , whose grid circuit acts as a

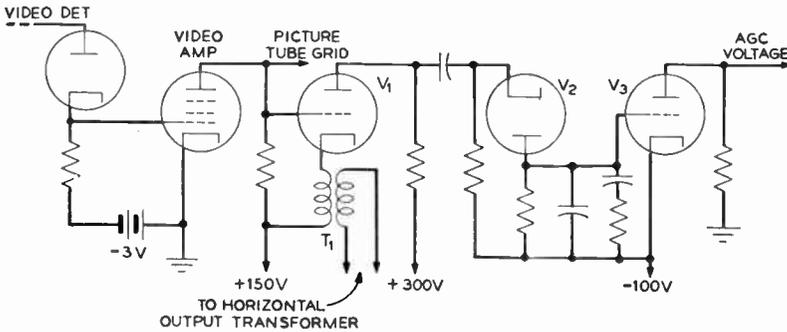


Fig. 14-27. A fast-acting A.G.C. circuit that also clamps the video signal. Peaking coils are not shown. (Courtesy of *RCA Review*.)

d-c restorer. V_1 will normally be nonconducting. Gating of the tube is accomplished through the transformer T_1 , which feeds a negative pulse from the horizontal output transformer to the cathode return of V_1 . The pulse amplitude is of sufficient magnitude to cause V_1 to conduct. During the conduction interval the output of V_1 will depend upon the magnitude of the composite video signal. When the video signal is small, the pulse through T_1 produces a large plate current; when the video signal is large, the grid voltage on V_1 bucks

³⁴ K. R. Wendt and A. C. Schroeder, "Automatic Gain Controls for Television Receivers." *RCA Review*, IX, 3 (September 1949).

the pulse and a smaller plate current flows. V_2 rectifies the amplified pulses and the grid voltage has magnitude and polarity determined by the video signal amplitude. V_3 is a d-c amplifier which inverts the signal and delivers the corresponding AGC voltage to the controlled amplifiers.

The circuit provides immunity to noise impulses in two ways. First, the AGC detector, V_2 , is driven only during the gated intervals which occur at the end of each sweep. Thus intersync noise pulses are completely eliminated from the gain-control system. Any noise riding on top of a sync pulse reduces the output of V_1 . Second, this reduced output caused by the coincident noise pulse will not be great enough to cause conduction in V_2 , and no change in the AGC voltage will be produced.

Notice that the AGC output from V_3 is completely isolated from $B+$ by returning the plate, through the load resistor, to ground. The d-c plate voltage for the tube is obtained by returning its cathode to 100 volts below ground.

14-11. Summary

The discussion up to this point in the chapter has covered some general considerations of television receivers and certain of the special features which are incorporated in typical commercial models of the conventional two-channel I-F type. Details of the sweep generators, sync separators, and video amplifiers that have been covered in previous chapters have not been repeated. We next consider the second general type of receiver, the intercarrier form, which uses a single I-F channel for both the visual and aural signals.

THE INTERCARRIER RECEIVER^{35,36}

An inherent difficulty present in the conventional television receiver is that it actually consists of two separate receivers, visual and aural, with only a single tuning control on the local oscillator associated with the common mixer circuit. If, for any reason, either of the two independent I-F systems drifts out of alignment, it is impossible to tune the receiver for both optimum picture and optimum sound.

³⁵ R. B. Dome, "Carrier-difference Reception of Television Sound." *Electronics*, 20, 1 (January 1947).

³⁶ S. W. Seeley, "Television Intercarrier Sound Design Factors." *Electronics*, 21, 7 (July 1948).

Since the sound system is narrow-band and more susceptible to mistuning than the wide-band video channel, the oscillator is adjusted to give the best sound reproduction—a strange procedure, indeed, in a system whose prime function is the reproduction of a picture. In the summer of 1946 R. B. Dome of the General Electric Company proposed a receiver system to the R.M.A. Subcommittee on V.H.F. Television Systems, which utilizes a common I-F channel for both audio and video signals. This system, which eliminates misalignment problems and results in a much simplified receiver, has been termed the Dome, intercarrier, or carrier-difference system of television reception. In contrast to other proposals for eliminating misalignment problems which employed some form of multiplexing or time-sharing between the sight and sound on a common carrier, the Dome proposal can operate under existing transmission standards, although certain slight modifications would be desirable to guarantee optimum performance. These modifications which would not affect existing receivers of the conventional type are discussed later in the chapter.

14-12. Principle of Operation

In order to understand the principle of operation of the intercarrier television receiver, let us examine some of the significant terms in the output of a nonlinear mixer. In general, three groups of input signals are applied to the mixer. These may be represented in the following manner: First, the amplitude-modulated visual carrier consists of a carrier plus several sideband frequencies. In order to reduce the number of terms in the equations we shall consider only one pair of sideband components corresponding to a modulating frequency ω_1 . Thus, for the visual signal we write

$$e_v = E_1 \cos \omega_v t + E_2 \cos (\omega_v + \omega_1)t + E_3 \cos (\omega_v - \omega_1)t \quad (14-76)$$

Second, the frequency-modulated aural wave of carrier frequency ω_a may be represented by the carrier and a single pair of sidebands

$$e_a = E_4 \cos \omega_a t + E_5 \cos (\omega_a + \omega_2)t + E_6 \cos (\omega_a - \omega_2)t \quad (14-77)$$

Third, the oscillator signal is

$$e_o = E_o \cos \omega_o t \quad (14-78)$$

We have previously seen that these signals effectively combine in series at the mixer input; hence for the signal voltage on the mixer

grid we may write

$$e_u = e_v + e_a + e_o \quad (14-79)$$

With the mixer operating in a nonlinear fashion, the plate current will be

$$i_p = ae_u + be_u^2 + ce_u^3 + \dots \quad (14-80)$$

where a , b , and c are constant terms which depend upon the mixer transfer characteristic. The several signal components that appear at the mixer output may be determined by substituting (14-79) into (14-80) and by expanding each term by means of trigonometric identities. Previous experience with expansions of this type shows that, among others, a number of heterodyne terms will be present whose frequencies are the sum or difference of the several component frequencies. For example, the oscillator beating against the visual signal components produces the terms

$$\left. \begin{aligned} E_1 E_o \cos (\omega_o - \omega_r) t \\ E_2 E_o \cos [(\omega_o - \omega_r) - \omega_1] t \\ E_3 E_o \cos [(\omega_o - \omega_r) + \omega_1] t \end{aligned} \right\} \quad (14-81)$$

and

It should be clear that

$$\begin{aligned} \omega_o - \omega_r &= \omega_i \\ &= \text{visual I-F carrier frequency} \end{aligned} \quad (14-82)$$

Hence, these terms may be rewritten as

$$\left. \begin{aligned} E_1 E_o \cos \omega_i t \\ E_2 E_o \cos (\omega_i - \omega_1) t \\ E_3 E_o \cos (\omega_i + \omega_1) t \end{aligned} \right\} \quad (14-83)$$

In a similar manner the visual carrier beats with the aural components and yields

$$\left. \begin{aligned} E_4 E_1 \cos (\omega_a - \omega_v) t \\ E_5 E_1 \cos [(\omega_a - \omega_v) - \omega_2] t \\ E_6 E_1 \cos [(\omega_a - \omega_v) + \omega_2] t \end{aligned} \right\} \quad (14-84)$$

Cross-modulation between (14-83a) and (14-84) gives still another frequency group

$$\begin{aligned}
 & E_1^2 E_4 E_o \cos [\omega_i - (\omega_a - \omega_v)t] \\
 & E_1^2 E_5 E_o \cos [\omega_i - (\omega_a - \omega_v) + \omega_2]t \\
 & E_1^2 E_6 E_o \cos [\omega_i - (\omega_a - \omega_v) - \omega_2]t
 \end{aligned} \tag{14-85}$$

The several components of (14-83) and (14-85) are in the general vicinity of ω_i and so are passed by the visual I-F system. *If, now, the amplitudes of (14-85) are attenuated such that they are of the same strength as the visual I-F sidebands, they will appear to the second detector as additional sidebands associated with the visual I-F carrier.* This concept is the basis for intercarrier reception. In the detecting process the sidebands are effectively "stripped off" the carrier and the signals present in the demodulated output will be

$$f_1 \quad f_a - f_v - f_2 \quad f_a - f_v \quad f_a - f_v + f_2 \tag{14-86}$$

Let us examine these frequencies with care. The first is simply the video signal proper. The third, $f_a - f_v$, is the *difference* between the two transmitter carrier frequencies, which by current standards must be 4.5 megacycles. Let us call this carrier-difference frequency the aural I.F. Then the second and fourth terms of (14-86) are the lower and upper sidebands, respectively, of the modulated aural I-F signal, which signal occupies a band of frequencies lying just above the upper limit of the video band. This aural I.F., centered on 4.5 megacycles, may then be removed from the video channel by a trap circuit, detected, and amplified in the usual manner.

This entire process may be summarized in the following manner: In the mixing process the aural signal is reproduced in a group of sidebands centered on $[f_i - (f_a - f_v)]$, which is the visual I-F carrier minus the difference between the aural and visual R-F carriers. If these signals are attenuated sufficiently, the visual I-F carrier will view them as more of its own sidebands. In the detection process the visual I-F carrier is demodulated; the output is the video signals plus a frequency-modulated aural carrier of 4.5 megacycles, which may be separated from the video by a trap circuit.

14-13. The Intercarrier Circuit

The block diagram of an intercarrier or carrier-difference receiver is shown in Fig. 14-28. One advantage of the carrier-difference system is immediately apparent: the number of stages in the receiver is reduced because no separate sound I-F channel, as such, is required.

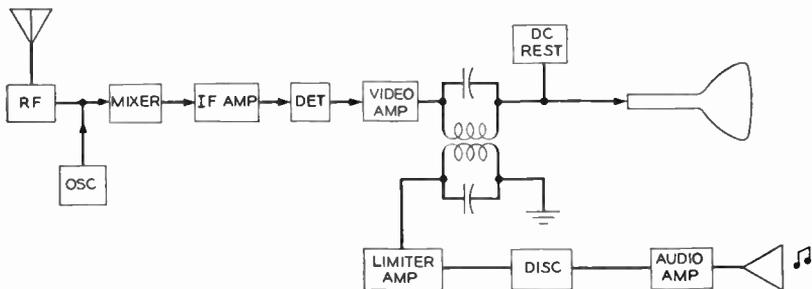


Fig. 14-28. Block diagram of the intercarrier receiver. Sweep circuits are not shown. Both visual and aural signals are amplified in a common I-F system. Separation of the two signals occurs after detection.

Notice, too, that the two programs are kept together through the video amplifier so that the latter contributes to the gain of the over-all aural system.

The stability afforded by the intercarrier system is even more important. The aural I-F frequency of 4.5 megacycles is almost completely independent of local oscillator drift, being determined solely by the R-F carriers at the transmitter which are subject to constant control. In tuning the receiver, one adjusts for optimum picture, and the sound reproduction is unaffected. We have shown in an earlier section that the video system is broad and fairly immune to small amounts of local oscillator drift. In the conventional receiver the stability requirements are set by the aural system. In the intercarrier receiver the aural I.F. is independent of local oscillator drift and a fair amount of drift may be tolerated. This fact is illustrated in commercial intercarrier receivers, which invariably have no provision for fine tuning of the local oscillator.

The price that must be paid for the advantages just listed is that further restrictions are required on the pass characteristics of the I-F amplifiers, which must provide the required attenuation of the aural components in the I.F. Let us consider these restrictions. The attenuation of the aural components may be accomplished by providing a low-amplitude pass characteristic in the range of I-F frequencies occupied by these aural components, which are centered on the frequency $[f_i - (f_a - f_v)]$. In writing the frequency in the form that was used in (14-85), we obscure the fact that this is precisely the

aural I-F frequency of the conventional receiver, a fact which may be verified by substituting for f_i from eq. (14-82). Thus,

$$\begin{aligned}
 f_i - (f_a - f_v) &= (f_o - f_v) - (f_a - f_v) \\
 &= f_o - f_a \\
 &= \text{conventional aural I.F.}
 \end{aligned}
 \tag{14-87}$$

Therefore the pass characteristic must have a "shelf" centered on 21.25 megacycles and extending roughly 25 kilocycles or more in each direction. In his report to the R.M.A. Dome recommended that the shelf height should correspond to the minimum level expected in the video carrier. For example, if maximum white modulation at the transmitter reduces the visual carrier to 15 per cent of its peak amplitude, the R-A characteristics will halve this in the receiver to 7.5 per cent. Then, if the aural carrier amplitude is equal to black level in the visual carrier, the shelf height must be

$$\frac{7.5}{75} = 10\%
 \tag{14-88}$$

In practical design a value ranging from 5 to 10 per cent is used. The ideal I-F pass characteristic for the intercarrier receiver is illustrated in Fig. 14-29a. The required attenuation in the vicinity of 21.25 megacycles is normally obtained by means of a trap circuit. Considerable care should be exercised to ensure that the shelf is reasonably level in order that slope detection of the frequency-modulated aural signal not take place.

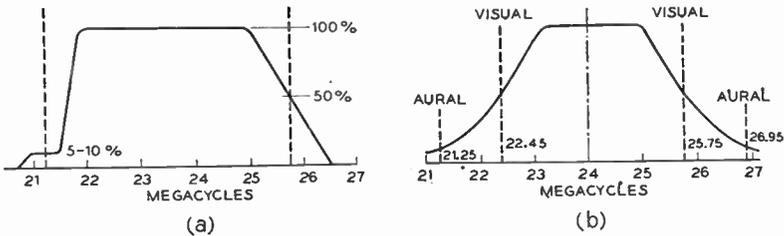


Fig. 14-29. I-F pass characteristics for the intercarrier receiver. (a) Ideal I-F characteristic. Notice the shelf at 21.25 megacycles, which is used to attenuate the aural signals. (b) Characteristic for a simplified receiver. Symmetry in the characteristic reduces design problems in the local oscillator, but a full 4.5-megacycle video bandwidth cannot be realized.

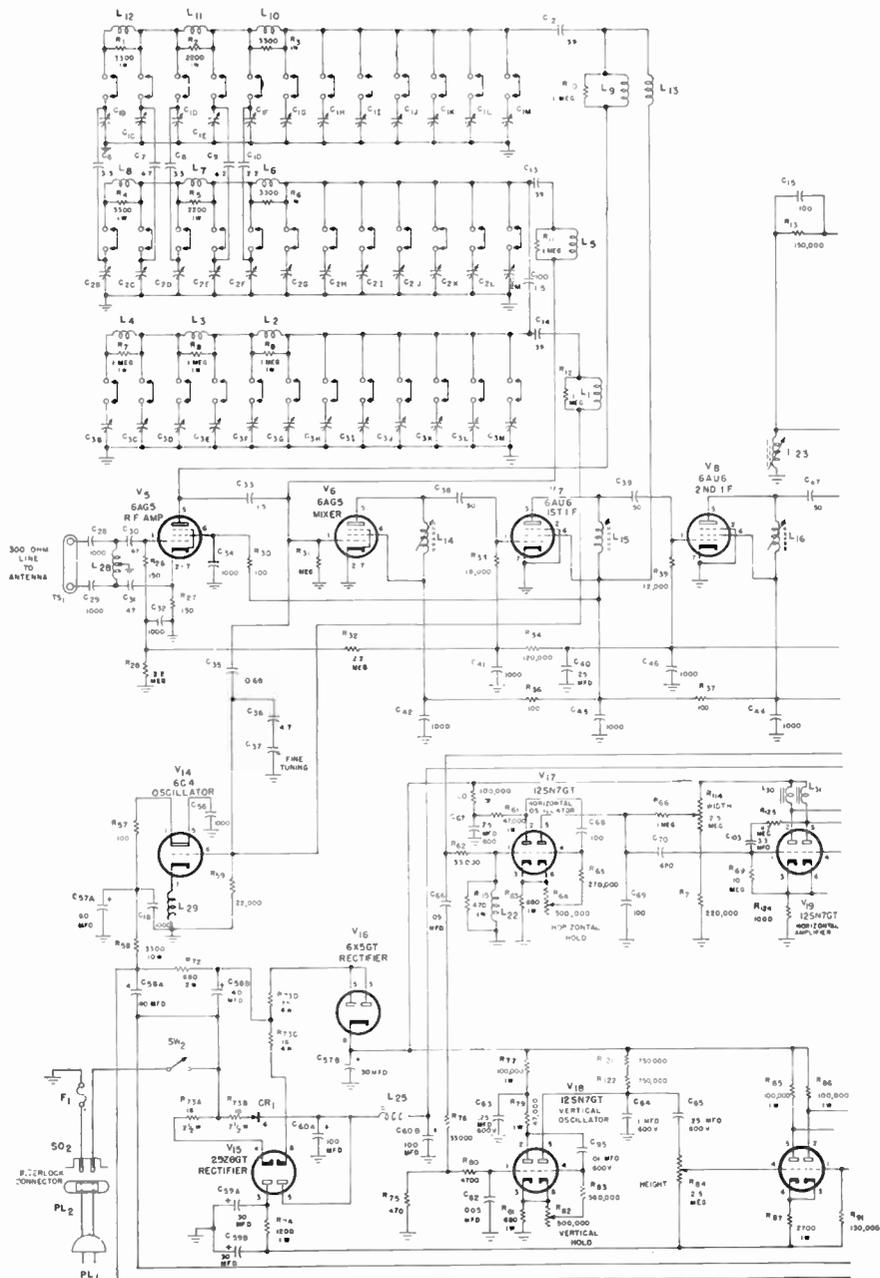


Fig. 14-30a. Circuit diagram of a 7-inch electrostatic-deflection receiver of the intercarrier type. Fig. 14-30b shows the I-F response. (Courtesy of The Hallicrafters Company.)

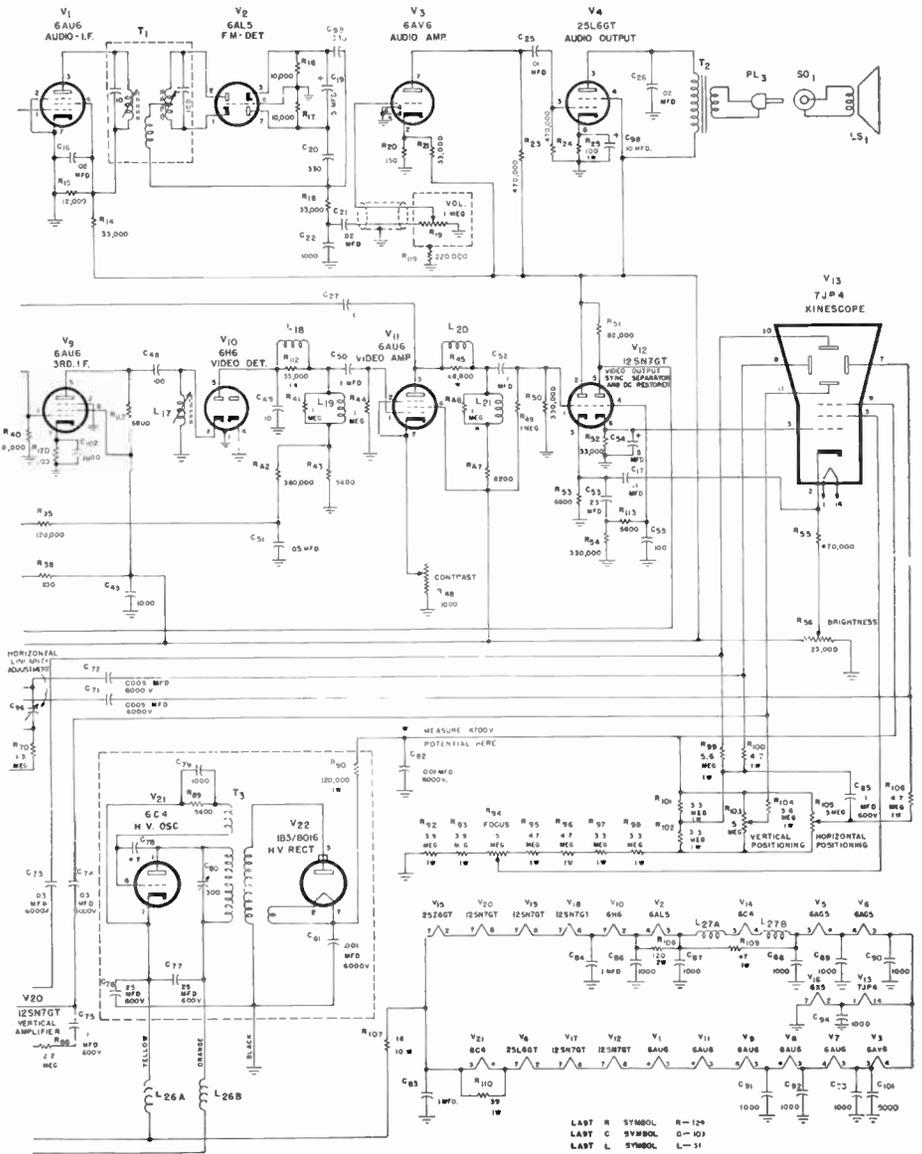


Fig. 14-30a. (cont.)

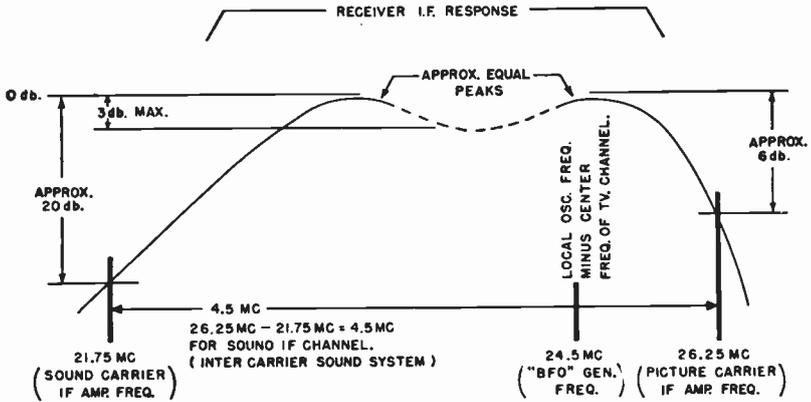


Fig. 14-30b.

Even further simplification of the intercarrier receiver may be had by utilizing a symmetrical I-F pass characteristic, but at the expense of performance. To illustrate, consider the symmetrical characteristic shown at *b* in Fig. 14-29. A response of this type may be obtained directly by means of the I-F interstage coupling networks; no special trapping circuits are required. This response curve further simplifies the oscillator design because two different sets of intermediate frequencies may be used. For example, for channels No. 2 through No. 6 the local oscillator operates at a frequency higher than the incoming R.F. and the I.F.'s are:

$$\left. \begin{array}{l} \text{aural I.F.} = 21.25 \text{ megacycles} \\ \text{visual I.F.} = 25.75 \text{ megacycles} \end{array} \right\} \text{ channels 2-6}$$

For the higher channels, the oscillator operates *below* the incoming R.F. and the corresponding intermediate frequencies are

$$\left. \begin{array}{l} \text{aural I.F.} = 26.95 \text{ megacycles} \\ \text{visual I.F.} = 22.45 \text{ megacycles} \end{array} \right\} \text{ channels 7-13}$$

By this device the operating range of the local oscillator is reduced by a considerable factor with a corresponding increase in its frequency stability.

The chief disadvantage of the simplified form of the receiver is that it does not provide a 4-megacycle bandwidth for the visual I-F signal. This may be seen directly from Fig. 14-29*b*. Ideally, the response

at the two aural carriers, 21.25 and 26.95 megacycles, should be 26 decibels below the flat top, and the sides should be reasonably linear to meet the R-A requirements. These two requirements are not compatible and the aural carrier level is generally raised to between -17 and -20 decibels. With this amount of attenuation, any audio signals which appear at the CRT grid by virtue of slope detection are so small that they have negligible effect in the reproduced image.

Figure 14-30 shows the diagram of a commercial receiver of the intercarrier type. It may be seen that the aural signals centered on 4.5 megacycles are removed at the plate of the first video amplifier. The shunt combination of resistance and capacitance on the grid of the aural I-F amplifier in conjunction with the unbiased grid circuit provides a clipping or limiting action, which serves to remove any spurious amplitude variations in the frequency-modulated aural signal. The remaining portions of the receiver are similar to those previously described, except for the high-voltage power supply. Since electrostatic deflection is used, the use of a flyback power supply is ruled out. To circumvent the need for a 60-cycle high-voltage supply, an R-F oscillator is used. With proper design the oscillator output may be made sufficiently large so that on rectification by the 1B3/8016 it furnishes the required high-voltage low-current supply for the cathode-ray tube.³⁷

14-14. Recommended Changes in Transmission Standards

In order to ensure correct operation of the intercarrier type of television receiver Dome has proposed that four principal modifications be made in the existing television transmission standards. A discussion of these proposals follows.

In the process of amplitude modulation, inadvertent frequency or phase modulation of the carrier may occur to some extent unless particular care is exercised in adjusting the modulating equipment. Since the final aural I.F. in the intercarrier receiver is developed by the beat between the visual and aural carriers, inadvertent modulation of this type will cause spurious modulation of the beat note,

³⁷ For design considerations of the R-F oscillator power supply see O. H. Schade, "Radio-Frequency Operated High-Voltage Supplies for Cathode Ray Tubes," *Proc. IRE*, **31**, 4 (April 1943); also R. S. Mautner and O. H. Schade, "Television High Voltage R-F Supplies," *RCA Review*, **VIII**, 1 (March 1947).

which, in turn, will be detected and reproduced in the audio channel of the receiver. The first proposal, then, is that the standards include a section requiring that no phase or frequency modulation of the visual carrier be permitted.

A second proposal has to do with the maximum allowable degree of modulation of the visual carrier. Since the aural I.F. is the beat note between the two carriers, its existence depends upon the presence of both carriers at all times. This fact may be verified from eqs. (14-85), which show that the amplitudes of the beat notes are proportional to the square of E_1 , the video carrier amplitude. The difficulty here is that current standards allow a maximum white signal to modulate the video carrier to 15 per cent *or less* of the peak voltage amplitude of the radiated signal. Thus, if a white signal modulates the visual carrier down to zero amplitude, no sound will be reproduced at the receiver for the duration of a white signal. For this reason Dome recommends that the phrase "or less" be eliminated so that the minimum carrier level will be 15 per cent of the peak.³⁸

The third proposal is that the frequency difference between the two transmitted carriers be incorporated into the standards. Existing standards specify tolerances on each carrier frequency separately. Thus, even though each carrier is held within the allowable tolerance, the beat note may drift from 4.5 megacycles, thereby impairing the reproduction of sound at the receiver. It may be supposed that a system of driving both transmitters from a single controlled oscillator will be used if this proposal is adopted by the Federal Communications Commission.

The fourth proposal for ensuring proper performance of the carrier difference receiver is that the ± 25 -kilocycle swing permitted for the frequency-modulated aural signal be raised to at least ± 75 kilocycles. The effect of this proposal would be to mask out inadvertent frequency modulation of the visual carrier.

Two facts should be noted in reference to the proposed modifications of the television transmission standards which have just been described. First, their acceptance will have no adverse effect on existing receivers of the conventional double-I.F. channel type which are properly designed. Second, the majority of television stations

³⁸ Since the time of Dome's proposal, the standard has been changed so that maximum white corresponds to 12.5 per cent \pm 2.5 per cent of the peak voltage amplitude. See standard No. 16, section 9-2.

maintain such excellent control over their transmission that inter-carrier receivers operate perfectly under the existing standards. The proposals are safety measures which would guarantee proper reception from all television stations. In support of this statement it is interesting to note that at the time of writing, receivers of the intercarrier type are produced by at least four large manufacturing organizations.

PROJECTION SYSTEMS^{39,40,41,42,43}

As black-and-white television has become more widely adopted, there has been an increasing demand on the part of the buying public for larger and larger images at the television receiver. To satisfy this demand direct-view tubes of 10, 12, 16, 19, and even 30-in. diameters have been developed, and the trend is to even larger sizes as better means are developed to overcome the large forces exerted by atmospheric pressure on the face of the evacuated tube envelope. These large tubes are expensive, however, and as a consequence resort has been made to various forms of optical systems which project an enlarged image of the fluorescent face of a small tube onto a viewing screen.

Basically two types of projection systems have been employed with cathode-ray tubes in television receivers: the refractive type, which utilizes a lens, and the reflective type, which employs a mirror. Of the two, the latter has found greater adoption in the United States. The reasons for this are discussed in the paragraphs which follow.

14-15. Refractive Projection Systems

The basic layout for a refractive optical projection system is shown in Fig. 14-31. The principle is quite simple, being similar to that used in a slide- or motion-picture projector: A real image of the illuminated face of the cathode-ray tube is focused on the viewing screen by a suitable lens system. Either an opaque or a translucent

³⁹ D. O. Landis, "Television Receiver." U. S. Patent No. 2,273,801, February 17, 1942.

⁴⁰ M. S. Kay, "Television Projection Systems." *Radio and Television News*, 41, 5 (May 1949).

⁴¹ E. L. Clark, "Projection Type Television Receiver." *Radio-Electronic Engineering Edition of Radio and Television News*, 9, 3 (September 1947).

⁴² V. K. Zworykin, "Reflective Optical System for Projection Television." *Radio News*, 38, 3 (September 1947).

⁴³ H. Rinia, J. de Gier, and P. M. van Alphen, "Home Projection Television Part I," *Proc. IRE*, 36, 3 (March 1948).

screen may be used, depending upon whether the image is projected from the front or the back of the screen. The primary problem involved

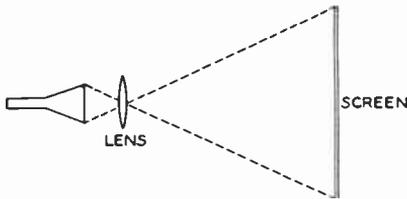


Fig. 14-31. Basic layout of a refractive projection system.

has to do with the brightness of the projected image. For example, let a 5TP4 projection kinescope be used and scanned with a raster size of 4 inches \times 3 inches. Further, let the projected image be 20 in. \times 15 in. If the lens system has a transmission coefficient of unity, the total light

flux in the raster and the image must be the same, or, in other words, the illumination on the viewing screen will be less by a factor of

$$\frac{4 \times 3}{20 \times 15} = \frac{1}{25}$$

than that of the kinescope image. This situation is aggravated by the fact that a typical lens system will have a transmission coefficient in the range of 0.06, so that the illumination ratio is reduced to roughly 1/400.

This situation may be eased by employing a larger kinescope, which permits reduction in the system magnification, but the lens size must also be increased. Alternatively, the tube size may be kept constant and a larger-aperture lens provided. To this end a commercially available $f/1.9$ lens has been developed. At present, larger stop sizes seem impracticable, first, because of high cost, and second, because of the lowered transmission coefficient. In the latter regard problems of aberration arise in the larger lens systems, which are corrected by increasing the number of elements in the lens system. We have already observed in Chapter 6 that the loss of light in transmission through a lens system occurs primarily at the air-glass boundaries; hence a system employing a large number of elements will generally have a lower transmission coefficient than a simple lens.

The use of the aluminum-backed 5TP4 projection kinescope and of the coated $f/1.9$ lens has permitted the development of a satisfactory refractive projection system, but this is less popular than the reflective type.

14-16. Reflective Projection.

The majority of reflective-projection television systems are based on some variation of the Schmidt optical system, which is used in certain astronomical telescopes. One form of the system is illustrated in Fig. 14-32*a*. In essence, the system utilizes a spherical mirror

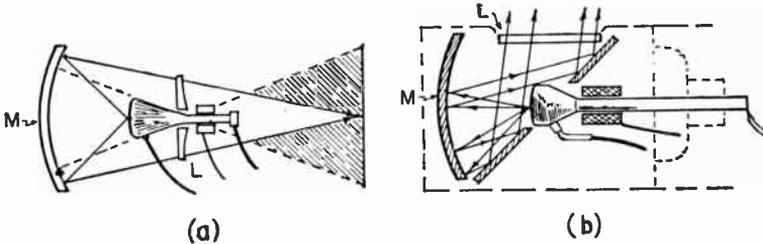


Fig. 14-32. Two forms of the Schmidt optical system of projection. (a) A straight system. M is a spherical mirror; L , a correction lens. The shaded area represents light lost because of interception by the deflection yoke. (b) A folded system. A flat mirror reduces the size of the optical system and no light is lost by interception. (Courtesy of North American Philips Co., Inc.)

that is silvered on its front surface to reflect a real image of the kinescope face onto the viewing screen. Since the mirror exhibits some spherical aberration (although to a lesser degree than the projection lens), a correction lens, L , is required.⁴⁴ As may be seen from the diagram this corrector is designed to fit over the neck of the projection tube. One interesting feature to note is that quite often the center portion of the curved mirror is cut away, the reason being that monochromatic aberrations are thereby eliminated.

The chief advantages of the reflective system are three: first, an over-all transmission coefficient of approximately 0.25 is obtained; second, of all mirror shapes the spherical is the easiest and least expensive to produce; third, even though of complex form, the correc-

⁴⁴ It is interesting to note that the separate corrector lens may be eliminated by use of a Mangin mirror in which the front surface is ground to provide the required correction. Both front and rear surfaces are spherical, but silvering must be on the rear surface. See D. H. Jacobs, *Fundamentals of Optical Engineering*. New York: McGraw-Hill Book Company, Inc., 1943.

To the author's knowledge the Mangin mirror has not been used in television work.

tor plate may be produced quite inexpensively. To this end techniques have been developed for heat-forming the corrector from a sheet of plastic.⁴⁵

The primary problems associated with the reflective system are collecting of dust on the optical surfaces, proper support of the re-

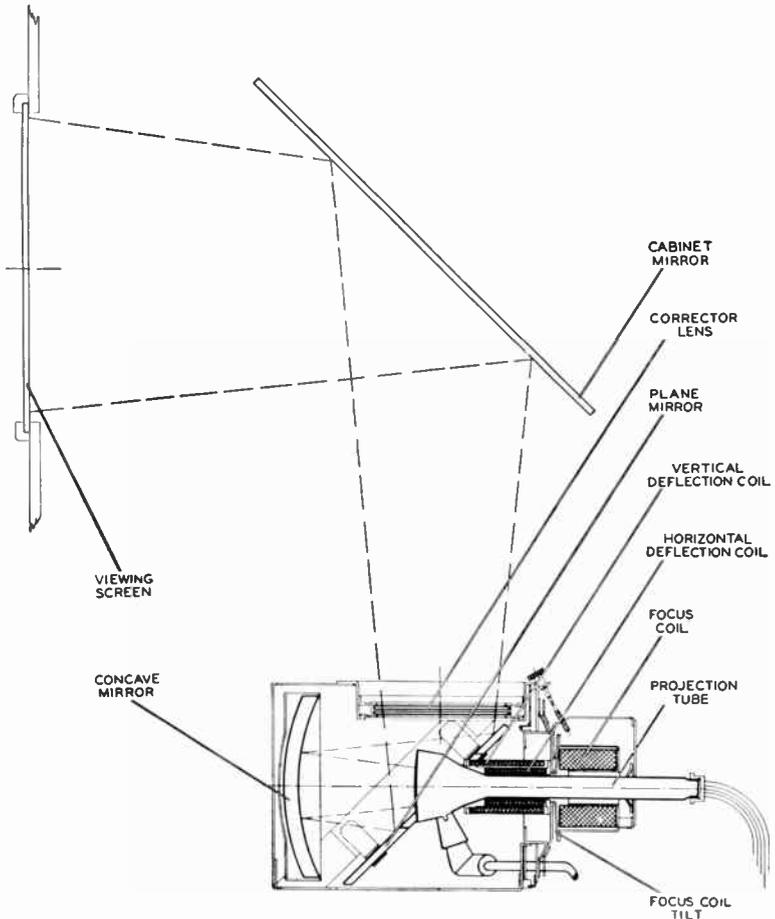


Fig. 14-33. Details of a folded Schmidt optical projection system. Notice how the use of two plane mirrors reduces the cabinet size required. (Courtesy of North American Philips Co., Inc.)

⁴⁵ An alternative system which utilizes a gelatine coating of proper form on a supporting layer of glass has been developed by the Philips organization. Rinia, de Gier, and van Alphen, *op. cit.*

flecting element, and sagging of the plastic corrector under its own weight.

Considerable ingenuity has been displayed in fitting the optical system into the limited cabinet space of the typical home television receiver. Shortening of the physical dimensions of the optical path is usually accomplished by folding the optical path by means of mirrors as shown in Fig. 14-32*b*. By way of example a projection system which employs two plane mirrors in addition to the curved one is shown in Fig. 14-33. The particular unit shown is a packaged unit, consisting of the tube, optical system, high-voltage power supply, and deflection equipment, and is manufactured under the trade name "Protelgram" by Philips. A noteworthy feature of the Philips system is the small (maximum screen diameter of 21.5 millimeters) projection tube used, the idea being that only small deflections are required, with a resulting saving in deflection and focusing power. Reference to the diagram shows how plane mirrors are used to fit the optical path into the receiver cabinet.

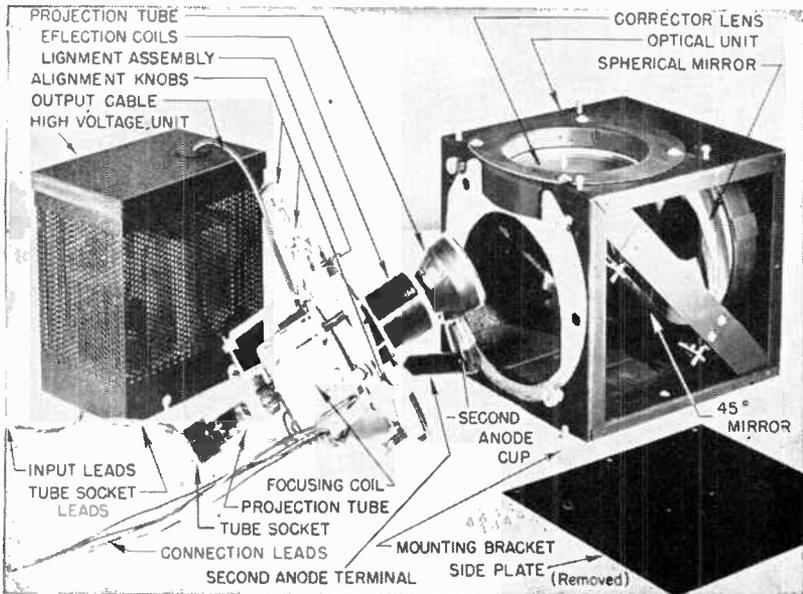


Fig. 14-34. Components of Protelgram, a packaged Schmidt optical unit with power supply and deflection system. (Courtesy of North American Philips Co., Inc.)

In a number of projection receivers plastic viewing screens are used, their surfaces being treated so that they concentrate the light into the region most commonly used for viewing, *i.e.*, they do not diffuse the light equally in all directions. In one Philco unit a reflecting type of screen has been used, whose viewing region is 60° in the horizontal plane and 20° in the vertical plane. Horizontal directivity is obtained by scoring the screen surface with a large number of vertical grooves of random shape. The screen is made concave in the vertical plane to produce the required vertical directivity. Further treatment of the screen surface provides uniform diffusion within the 60° and 20° angular limits, and a screen brightness of some 50 foot-lamberts is obtained.⁴⁶

14-17. The Skiatron System

In the types of projection television systems described thus far the brightness of the final image is limited by the amount of light which can be produced by the fluorescent screen of a cathode-ray tube. At least two other systems have been devised which eliminate this inherent limitation of the cathode-ray tube, the Scophony and Skiatron systems. In both of these, the light originates in an external source, such as a carbon arc, and its intensity is modulated by a light valve, which, in turn, is controlled by the video signal. In the older Scophony system, which is illustrated in Fig. 14-35, the light valve

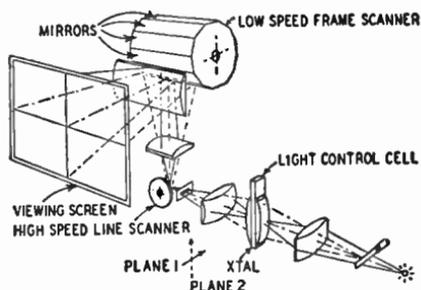


Fig. 14-35. Basic components of the Scophony reproducing system. Light from an external source, such as an arc lamp, is modulated by a light-control cell under control of the video signal. Scanning is provided by rotation of two polygonal mirrors. (Courtesy of Skiatron Corporation.)

⁴⁶ M. S. Kay, *op. cit.*

takes the form of a Kerr cell or, in later form, of an ultrasonic light valve. Scanning is provided by two polygonal mirrors mounted normal to each other. The Scophony system of reproduction has been used in England primarily.⁴⁷

The Skiatron projection system is an all-electronic counterpart of the Scophony system. The basic element is the Skiatron (*Skiashadow*) tube itself. Shown in Fig. 14-36 it has the general shape

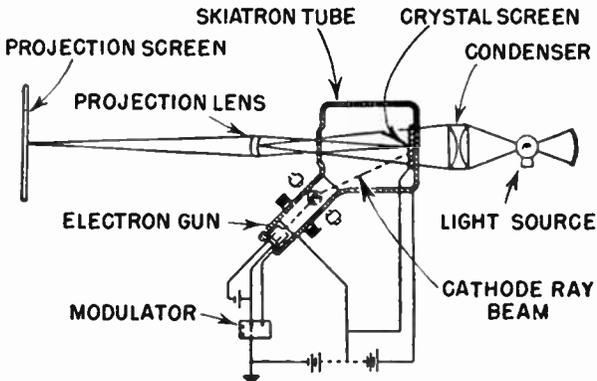


Fig. 14-36. The Skiatron projection system. Light from an external source is modulated by the variable-opacity screen in the Skiatron tube. Opacity is controlled by applying the video signal to the electron scanning beam. (Courtesy of Skiatron Corporation.)

of an iconoscope, but the familiar mosaic is replaced by a screen of materials which exhibit the property of "electron opacity," that is, the degree of opaqueness of the crystal screen may be controlled by electron bombardment. Thus, the electron beam which scans the screen is modulated by the video information and so controls the light which reaches the viewing screen on which the television picture is reproduced.⁴⁸ While the Skiatron tube found some use in radar projection systems during the war,⁴⁹ at the time of writing it has not been used to any great extent in television receivers manufactured in this country.

⁴⁷ For a detailed description of the light valve and its operation, see A. H. Rosenthal, "Problems of Theater Television Projection Equipment." *J.S.M.P.E.*, 45, 3 (September 1945); H. W. Lee, "The Scophony Television Receiver." *Nature*, 142 (July 1938).

⁴⁸ For further details of the Skiatron, see A. H. Rosenthal, *op. cit.*

⁴⁹ L. R. Ridenour, *Radar System Engineering*. New York: McGraw-Hill Book Company, Inc., 1947.

CHAPTER 15

STAGGER-TUNED AMPLIFIERS

It was observed in the discussion on I-F amplifiers in the last chapter that the rather wide pass band of those amplifiers may be obtained with single-tuned component stages, provided that each of the cascaded stages is tuned to a different frequency. In the present chapter we consider the design equations for a stagger-tuned amplifier of this type. Since the single-tuned stage is the primary "building block" of the system, we shall first derive its response characteristic. We shall then show how several stages may be stagger-tuned to synthesize a given symmetrical pass characteristic of the maximal flatness type.

15-1. Single-tuned Amplifier

The basic circuit of the single-tuned amplifier is shown in Fig. 15-1.

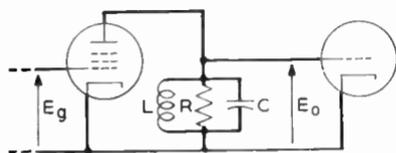


Fig. 15-1. The single-tuned amplifier. D-c returns and the blocking condenser are omitted.

Since a pentode is used, the voltage amplification of the stage at any frequency will be

$$A = -g_m Z_L \quad (15-1)$$

hence the frequency response, both of amplitude and phase, is determined solely by the variation of the load impedance, Z_L ,

with frequency. Let us investigate this impedance. Reading from the circuit diagram we have

$$\begin{aligned} Z_L &= \frac{1}{\frac{1}{R} + j\left(\omega C - \frac{1}{\omega L}\right)} = \frac{R}{1 + jR\left(\omega C - \frac{1}{\omega L}\right)} \\ &= \frac{R}{1 + jRC\left(\omega - \frac{1}{\omega LC}\right)} \end{aligned} \quad (15-2)$$

Let $\omega_o =$ center frequency $= \frac{1}{\sqrt{LC}}$ (15-3)

then $Z_L = \frac{R}{1 + jRC\left(\omega + \frac{\omega_o^2}{\omega}\right)} = \frac{R}{1 + j\omega_o RC\left(\frac{f}{f_o} - \frac{f_o}{f}\right)}$ (15-4)

Now, by definition, the effective Q of the parallel combination of R , L , and C is the ratio of the imaginary to the real part of the equivalent series impedance of R and L . Thus,

$$\begin{aligned} Z_s &= \text{equivalent series impedance of } R \text{ and } L \\ &= \frac{1}{\frac{1}{R} - \frac{j}{\omega L}} = \frac{R\omega L}{\omega L - jR} \end{aligned}$$

and, rationalizing,

$$Z_s = \frac{\omega^2 RL^2 + j\omega R^2 L}{R^2 + (\omega L)^2} \quad (15-5)$$

and, by definition,

$$Q = \frac{\omega R^2 L}{\omega^2 RL^2} = \frac{R}{\omega L} \quad (15-6)$$

At the center frequency, the effective value of Q will be

$$Q_o = \frac{R}{\omega_o L} = R\omega_o C \quad (15-7)$$

by (15-3), and the expression for Z_L becomes

$$Z_L = \frac{1}{1 + jQ_o\left(\frac{f}{f_o} - \frac{f_o}{f}\right)} \quad (15-8)$$

It may be demonstrated that $|Z_L|$ and hence A , exhibits geometric symmetry; that is, at any two frequencies, f_a and f_b , whose geometric mean is f_o , the center frequency, the function has the same magnitude, and equal phase angles but of opposite sign. To show this, let

$$f_a f_b = f_o^2 \quad (15-9)$$

Then at f_a

$$y_a = \left(\frac{f}{f_o} - \frac{f_o}{f}\right)_{f_a} = \left(\frac{f_a}{f_o} - \frac{f_o}{f_a}\right) \quad (15-10)$$

$$\text{Similarly at } f_b \quad y_b = \left(\frac{f_b}{f_o} - \frac{f_o}{f_b} \right) \quad (15-11)$$

and, substituting for f_b from (15-9), we have

$$y_b = \left(\frac{f_o^2}{f_a f_o} - \frac{f_o f_a}{f_o^2} \right) = - \left(\frac{f_a}{f_o} - \frac{f_o}{f_a} \right) \quad (15-12)$$

Since (15-10) and (15-12) are identical except for the minus sign, the condition for geometric symmetry is fulfilled. It may be seen, therefore, that a plot of Z_L or A on a logarithmic frequency scale will be symmetrical about the center frequency, f_o .¹

Consider next the half-power (3-decibel) bandwidth of the single-tuned amplifier stage. Let

$$\left. \begin{aligned} f_2 &= \text{upper half-power frequency} \\ f_1 &= \text{lower half-power frequency} \\ \Delta f &= \text{half-power bandwidth} \\ &= (f_2 - f_1) \end{aligned} \right\} \quad (15-13)$$

Since the half-power point occurs when the magnitude of the voltage amplification falls to 0.707 times its center frequency value, it is clear from (15-8) that at f_2

$$Q_o \left(\frac{f_2}{f_o} - \frac{f_o}{f_2} \right) = 1 \quad (15-14)$$

Then, factoring out $1/f_o$, we have

$$\frac{Q_o}{f_o} \left(f_2 - \frac{f_o^2}{f_2} \right) = \frac{Q_o}{f_o} (f_2 - f_1) = 1$$

$$\text{or} \quad \Delta f = \frac{f_o}{Q_o} \quad (15-15)$$

To obtain the expression for Δf in terms of the circuit parameters we substitute for Q_o from eq. (15-7); thus,

¹ It may be shown that Z_L displays arithmetic symmetry when Q_o is 20 or greater. Under this condition the expression for the voltage amplification becomes

$$A \approx \frac{-g_m R}{1 + j2RC(\omega - \omega_o)}$$

This approximate case is not considered here because the requirement on Q_o is not satisfied in the design of television I-F amplifiers operating in the vicinity of 25 megacycles.

$$\Delta f = \frac{1}{2\pi RC} \quad (15-16)$$

It should be noticed that the half-power bandwidth of the amplifier is independent of the center frequency to which the plate load circuit is tuned.

If (15-8) and (15-1) are combined, the final expression for the gain or voltage amplification becomes

$$A = \frac{-g_m R}{1 + jQ_o \left(\frac{f}{f_o} - \frac{f_o}{f} \right)} \quad (15-17)$$

It is convenient for the work which follows to rearrange this expression for the voltage amplification; thus,

$$A = -\frac{g_m R}{Q_o} \frac{1}{\frac{1}{Q_o} + j \left(\frac{f}{f_o} - \frac{f_o}{f} \right)} = -\frac{g_m R}{2\pi f_o C R} \frac{1}{\frac{1}{Q_o} + j \left(\frac{f}{f_o} - \frac{f_o}{f} \right)} \quad (15-18)$$

$$= -\frac{g_m}{2\pi C f_o} \left[\frac{1}{d + j \left(\frac{f}{f_o} - \frac{f_o}{f} \right)} \right] \quad (15-19)$$

where

$$\begin{aligned} d &= \text{dissipation factor} \\ &= \frac{1}{Q_o} = \frac{\Delta f}{f_o} \end{aligned} \quad (15-20)$$

It is of further convenience to normalize all frequencies in the bracketed factor with respect to the center frequency, f_o . Thus, let

$$F = \text{normalized frequency} = \frac{f}{f_o} \quad (15-21)$$

$$\text{then} \quad d = \frac{\Delta f}{f_o} = \Delta F \quad (15-22)$$

and (15-19) becomes

$$A = -\frac{g_m}{2\pi C f_o} \frac{1}{\Delta F + j \left(F - \frac{1}{F} \right)} = -\frac{\mathfrak{M}}{f_o} \frac{1}{\Delta F + j \left(F - \frac{1}{F} \right)} \quad (15-23)$$

We see, then, that the amplification is the product of a constant factor, \mathfrak{M} , which is identical to the figure of merit previously defined in Chapter 7,² the reciprocal of f_o , and a form factor which varies

²This is true provided that C consists only of the tube interelectrode capacitance, a condition which is normally fulfilled in television I-F amplifiers.

with frequency. We may, therefore, write the magnitude of the amplification as

$$A = \frac{\mathfrak{M}}{f_o} S_1 \quad (15-24)$$

$$\text{where } S_1 = \frac{1}{\left| \Delta F + j \left(F - \frac{1}{F} \right) \right|} = \frac{1}{\sqrt{(\Delta F)^2 + \left(F - \frac{1}{F} \right)^2}} \quad (15-25)$$

The normalized half-power bandwidth, ΔF , may be calculated in the following manner. At F_2 , the normalized upper half-power frequency, S_1 has the value $1/\sqrt{2}$; hence at F_2

$$S_1 = \frac{1}{\sqrt{(\Delta F)^2 + \left(F_2 - \frac{1}{F_2} \right)^2}} = \frac{1}{\sqrt{2}} \quad (15-26)$$

but

$$\frac{1}{F_2} = F_1 = \text{normalized lower half-power frequency} \quad (15-27)$$

$$\text{so } F_2 - \frac{1}{F_2} = F_2 - F_1 = \Delta F \quad (15-28)$$

$$\text{therefore } (\Delta F)^2 + (\Delta F)^2 = 2 \quad (15-29)$$

$$\text{or } \Delta F = 1$$

We may define the gain-bandwidth factor, Γ , of the amplifier as the product of the gain, normalized with respect to the tube figure of merit, times the half-power bandwidth

$$\Gamma = \frac{A_o}{\mathfrak{M}} \Delta f = \frac{1}{\Delta F f_o} \Delta f = 1 \quad (15-30)$$

that is, the gain-bandwidth factor of the single-tuned stage is unity.

It is clear from eq. (15-23) that the shape of the amplifier response is determined solely by S_1 , which is plotted to a normalized frequency scale in Fig. 15-2.³ Inspection of the S_1 curve shows that the re-

³ It may be observed that slide-rule calculation of the selectivity function for values of F near unity is not accurate because the function involves the difference between two nearly equal numbers. Professor Philip Greenstein of New York University has suggested the following transformation, which is of value in this regard (see bottom of opposite page.)

sponse of the single-tuned stage is quite unsatisfactory for use as a television I-F amplifier; the response is not flat-topped, and the skirt selectivity is poor. Hence we seek some other, more desirable response shape.

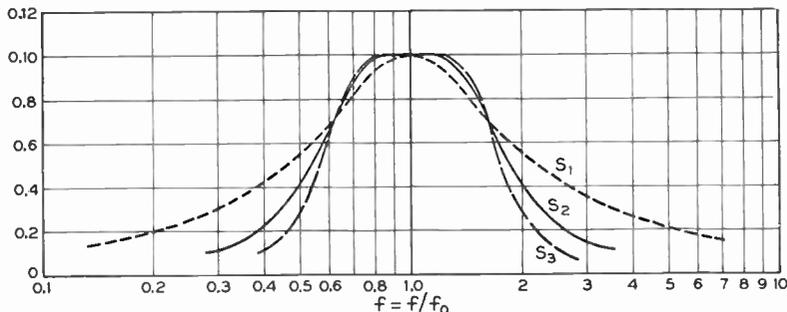


Fig. 15-2. Selectivity functions for tuned amplifiers. $\delta = 1$.

If m identical single-tuned stages are cascaded, each being tuned to the same center frequency, the selectivity function of the over-all amplifier will be the product of the m individual functions, or

$$S_1^m = \frac{1}{\left[(\Delta F)^2 + \left(F - \frac{1}{F} \right)^2 \right]^{m/2}} = \frac{1}{\left[1 + \left(F - \frac{1}{F} \right)^2 \right]^{m/2}} \quad (15-31)$$

Now the half-power points of the over-all selectivity function occur when the magnitude of the function is $1/\sqrt{2}$. Hence, we may write that at F_4 , the normalized upper half-power frequency of the over-all amplifier,

$$\left[1 + \left(F_4 - \frac{1}{F_4} \right)^2 \right]^{m/2} = 2^{1/2}$$

Let $y = \ln r$

then $-y = \ln \left(\frac{1}{r} \right)$

and $r - \frac{1}{r} = e^y - e^{-y} = 2 \sinh \ln r$

and $S_1 = \frac{1}{\sqrt{1 + (2 \sinh \ln r)^2}}$

$$\text{or} \quad 1 + \left(F_4 - \frac{1}{F_4} \right)^2 = 2^{1/m} \quad (15-32)$$

$$\begin{aligned} \text{Again} \quad \frac{1}{F_4} = F_3 \\ = \text{normalized lower half-power frequency} \\ \text{of the over-all amplifier} \end{aligned} \quad (15-33)$$

$$\begin{aligned} \text{and} \quad F_4 - F_3 = (\Delta F)_T = \delta \\ = \text{normalized half-power bandwidth of} \\ \text{the over-all amplifier} \end{aligned} \quad (15-34)$$

It follows at once that the normalized bandwidth of the m synchronously tuned stages in cascade is

$$\delta = (\Delta F)_T = \sqrt{2^{1/m} - 1} \quad (15-35)$$

Equation (15-35) shows that the over-all bandwidth of the cascaded stages *decreases* as more stages are added. It may, therefore, be seen that cascaded synchronously tuned stages are not satisfactory for use in the television I-F amplifier. Further, when m stages of amplification are cascaded, the gain-bandwidth factor is taken to be

$$\Gamma = \frac{(A_o)_T^{1/m}}{\mathfrak{M}} (\Delta f)_T = \frac{(A_o)_T^{1/m}}{\mathfrak{M}} f_o \delta \quad (15-36)$$

where

$$\left. \begin{aligned} (A_o)_T &= \text{over-all gain of the } m \text{ cascaded stages at} \\ &\quad \text{the center frequency} \\ \text{and} \quad \delta &= \text{normalized over-all half-power bandwidth} \end{aligned} \right\} \quad (15-37)$$

It may be seen, then, that $\Gamma \neq 1$ for *synchronously* tuned stages in cascade.

A closer approach to the desired I-F response is afforded by the selectivity function

$$S_n = \frac{1}{\sqrt{\delta^{2n} + \left(F - \frac{1}{F} \right)^{2n}}} \quad (15-38)$$

Curves of this type are plotted in Fig. 15-2 for $n = 2$ and $n = 3$, and they correspond, respectively, to transitionally coupled amplifiers employing two- and three-tuned circuits. In the remainder of this chapter we shall show how S_n may be synthesized by means of n pairs of staggered single-tuned amplifier stages.

15-2. The Staggered Pair⁴

It may be demonstrated that a single pair of staggered single-tuned stages, as shown in Fig. 15-3, may be designed to give the selectivity function

$$S_2 = \frac{1}{\sqrt{\delta^4 + \left(F - \frac{1}{F}\right)^4}} \tag{15-39}$$

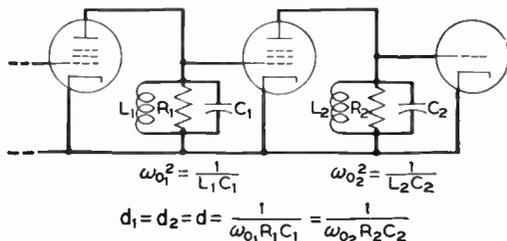


Fig. 15-3. The stagger-tuned pair.

The two tubes are assumed to be identical and both stages to have the same dissipation factor. Let the two stages have the center frequencies

$$\left. \begin{aligned} f_{o1} &= \alpha f_o \\ \text{and} \quad f_{o2} &= \frac{f_o}{\alpha} \end{aligned} \right\} \tag{15-40}$$

⁴ The method herein described for synthesizing the selectivity function is due to H. E. Wallman, "Stagger-Tuned I.F. Amplifiers." M.I.T. Radiation Laboratory Report 524, 1944. It has the advantage of being simple, straightforward, and requires only algebraic manipulation. It is admittedly less powerful than those procedures which are based upon geometrical distribution of the poles and zeros of the coupling networks in the complex plane. It may be shown that Wallman's results yield the Butterworth stagger, which corresponds to the transitional type of selectivity curve, S_n , of eq. (15-38). Synthesis with stagger-tuned stages of other selectivity curves, such as the overcoupled or Chebishev (Tchebyscheff) type are best handled by other methods. See, for example, W. H. Huggins, "The Natural Behavior of Broadband Circuits." Electronic Research Laboratories, Report No. E5013, Cambridge 1948. Design equations for both Butterworth and Chebishev staggers are also presented by M. Dishall, "Design of Dissipative Band-Pass Filters Producing Desired Exact Amplitude-Frequency Characteristics." *Proc. IRE*, **37**, 9 (September 1949). See also R. F. Baum, "Design of Broadband I.F. Amplifiers." *J. Appl. Physics*, **17**, 6 (June 1946) and **17**, 9 (September 1946).

such that their geometric mean is f_o , the center frequency of the pair

$$f_{o1}f_{o2} = \alpha f_o \frac{f_o}{\alpha} = f_o^2 \quad (15-41)$$

Then the over-all gain of the pair will be

$$\begin{aligned} A_T = A_1 A_2 &= \left[-\frac{g_m}{2\pi C_1} \frac{1}{f_{o1}} \frac{1}{d + j\left(\frac{f}{f_{o1}} - \frac{f_{o1}}{f}\right)} \right] \times \\ &\quad \left[-\frac{g_m}{2\pi C_2} \frac{1}{f_{o2}} \frac{1}{d + j\left(\frac{f}{f_{o2}} - \frac{f_{o2}}{f}\right)} \right] \\ &= \frac{g_m}{2\pi C_1} \frac{g_m}{2\pi C_2} \frac{1}{\alpha f_o} \alpha \left[\frac{1}{d + j\left(\frac{f}{f_{o1}} - \frac{f_{o1}}{f}\right)} \right] \times \\ &\quad \left[\frac{1}{d + j\left(\frac{f}{f_{o2}} - \frac{f_{o2}}{f}\right)} \right] \quad (15-42) \end{aligned}$$

Again the frequencies may be normalized with respect to f_o , the center frequency of the pair, thus,

$$\left. \begin{aligned} F &= \frac{f}{f_o} \\ F_{o1} &= \frac{f_{o1}}{f_o} = \alpha \frac{f_o}{f_o} = \alpha \\ \text{and} \\ F_{o2} &= \frac{f_{o2}}{f_o} = \frac{f_o}{\alpha f_o} = \frac{1}{\alpha} \end{aligned} \right\} \quad (15-43)$$

Substitution of these values into (15-42) yields

$$A_T = \frac{g_m}{2\pi C_1} \frac{g_m}{2\pi C_2} \frac{1}{f_o^2} \left[\frac{1}{d + j\left(\frac{F}{\alpha} - \frac{\alpha}{F}\right)} \right] \left[\frac{1}{d + j\left(\alpha F - \frac{1}{\alpha F}\right)} \right] \quad (15-44)$$

and the selectivity function of the staggered pair is

$$S_T = \left| \frac{1}{d + j\left(\frac{F}{\alpha} - \frac{\alpha}{F}\right)} \right| \cdot \left| \frac{1}{d + j\left(\alpha F - \frac{1}{\alpha F}\right)} \right| \quad (15-45)$$

Then, expanding the denominator, Δ , we have

$$\begin{aligned} \Delta &= \left| d^2 + jd \left[\alpha \left(F - \frac{1}{F} \right) + \frac{1}{\alpha} \left(F - \frac{1}{F} \right) \right] - \left[\left(F^2 + \frac{1}{F^2} \right) - \left(\alpha^2 + \frac{1}{\alpha^2} \right) \right] \right| \\ &= \left| d^2 + \left(\alpha^2 + \frac{1}{\alpha^2} \right) + jd \left[\left(\alpha + \frac{1}{\alpha} \right) \left(F - \frac{1}{F} \right) \right] - \left(F^2 + \frac{1}{F^2} \right) \right| \quad (15-46) \end{aligned}$$

Considerable simplification results from the use of the following identities

$$\left. \begin{aligned} \left(\alpha^2 + \frac{1}{\alpha^2} \right) &= \left(\alpha - \frac{1}{\alpha} \right)^2 + 2 \\ \left(F^2 + \frac{1}{F^2} \right) &= \left(F - \frac{1}{F} \right)^2 + 2 \end{aligned} \right\} \quad (15-47)$$

Then (15-46) becomes

$$\Delta = \left| d^2 + \left(\alpha - \frac{1}{\alpha} \right)^2 + jd \left[\left(\alpha + \frac{1}{\alpha} \right) \left(F - \frac{1}{F} \right) \right] + \left[j \left(F - \frac{1}{F} \right) \right]^2 \right| \quad (15-48)$$

It should be observed that the only variable in (15-48) is the frequency term $(F - 1/F)$, which we shall represent by x . We may, therefore, write the selectivity function of the staggered pair as

$$S_T = \frac{1}{\left| d^2 + \left(\alpha - \frac{1}{\alpha} \right)^2 + jd \left(\alpha + \frac{1}{\alpha} \right) x + (jx)^2 \right|} \quad (15-49)$$

We now ask the question: Is S_T of the same form as S_2 of the eq. (15-39)? If it is, the staggered pair may be used to synthesize the S_2 characteristic. To answer the question we shall use this procedure: Expand the denominator of S_2 by judicious factoring, set up a complex function whose magnitude is equal to the denominator of S_2 , and finally show that the complex function may be made identical to the denominator of S_T . Proceeding on this basis, then, we undertake to expand the square of the denominator of (15-39) by factoring; thus,

$$\begin{aligned} \delta^4 + x^4 &= (x^2 - j\delta^2)(x^2 - (-j)\delta^2) \\ &= (x^2 - \epsilon^{j(\pi/2)}\delta^2)(x^2 - \epsilon^{-j(\pi/2)}\delta^2) \end{aligned} \quad (15-50)$$

and, multiplying out the factors, we obtain

$$\begin{aligned}\delta^4 + x^4 &= x^4 - \delta^2 x^2 (\epsilon^{j(\pi/2)} + \epsilon^{-j(\pi/2)}) + \delta^4 \\ &= x^4 - \left(2\delta^2 \cos \frac{\pi}{2}\right) x^2 + \delta^4\end{aligned}\quad (15-51)$$

It is apparent that the identity is preserved because $\cos \frac{\pi}{2} = 0$.

By retaining the middle term, however, we can obtain the desired results, but we must be sure not to divide by the cosine term.

The second step in the synthesizing procedure requires that we set up a complex function of the general form of (15-48), whose magnitude squared is $(\delta^4 + x^4)$. Thus we try a function

$$\delta^2 + jkx + (jx)^2 \quad (15-52)$$

and stipulate that

$$|\delta^2 + jkx + (jx)^2|^2 = \delta^4 + x^4 \quad (15-53)$$

Let us expand the left-hand member of the equation.

$$\begin{aligned}|\delta^2 + jkx + (jx)^2|^2 &= |(\delta^2 - x^2) + jkx|^2 \\ &= (\delta^2 - x^2)^2 + (kx)^2 \\ &= \delta^4 + (k^2 - 2\delta^2)x^2 + x^4\end{aligned}\quad (15-54)$$

Then, equating (15-51) and (15-53), we have

$$\delta^4 + (k^2 - 2\delta^2)x^2 + x^4 = x^4 - \left(2\delta^2 \cos \frac{\pi}{2}\right) x^2 + \delta^4 \quad (15-55)$$

hence $2\delta^2 - k^2 = 2\delta^2 \cos \frac{\pi}{2}$

or $k^2 = 2\delta^2 \left(1 - \cos \frac{\pi}{2}\right) = 2\delta^2 \cdot 2 \sin^2 \frac{\pi}{4}$

and $k = 2\delta \sin \frac{\pi}{4}$ (15-56)

These equations show that $(\delta^4 + x^4)$ may be replaced by (15-52), provided that k has the value specified in eq. (15-56). We may, therefore, rewrite S_2 in terms of the complex function; thus,

$$S_2 = \frac{1}{\left|\delta^2 + j \left(2\delta \sin \frac{\pi}{4}\right) x + (jx)^2\right|} \quad (15-57)$$

With S_2 expressed in this manner, we can see that it is of the same form as S_T and that the two selectivity functions will be identical, provided that the coefficients of corresponding powers of x are equal. Then, equating coefficients, we have

$$\text{for } x^0 \qquad d^2 + \left(\alpha - \frac{1}{\alpha}\right)^2 = \delta^2 \qquad (15-58)$$

$$\text{and for } x^1 \qquad d \left(\alpha + \frac{1}{\alpha}\right) = 2\delta \sin \frac{\pi}{4} \qquad (15-59)$$

We now solve for d and α . Squaring (15-59),

$$d^2 \left(\alpha + \frac{1}{\alpha}\right)^2 = 4\delta^2 \sin^2 \frac{\pi}{4}$$

$$\text{but} \qquad \left(\alpha + \frac{1}{\alpha}\right)^2 = \left(\alpha - \frac{1}{\alpha}\right)^2 + 4$$

$$\text{hence} \qquad d^2 \left[\left(\alpha - \frac{1}{\alpha}\right)^2 + 4\right] = 4\delta^2 \sin^2 \frac{\pi}{4} \qquad (15-60)$$

Substituting from (15-58) for $\left(\alpha - \frac{1}{\alpha}\right)^2$,

$$d^2[\delta^2 - d^2 + 4] = 4\delta^2 \sin^2 \frac{\pi}{4}$$

$$\text{or} \qquad d^4 - d^2(\delta^2 + 4) + 4\delta^2 \sin^2 \frac{\pi}{4} = 0 \qquad (15-61)$$

$$\text{Therefore} \qquad d^2 = \frac{(4 + \delta^2) \pm \sqrt{\delta^4 + 8\delta^2 + 16 - 16\delta^2 \sin^2 \frac{\pi}{4}}}{2} \qquad (15-62)$$

The term under the radical may be simplified.

$$\begin{aligned} \delta^4 + 8\delta^2 + 16 - 16\delta^2 \sin^2 \frac{\pi}{4} &= 8\delta^2 \left(1 - 2 \sin^2 \frac{\pi}{4}\right) + \delta^4 + 16 \\ &= 16 + 8\delta^2 \cos \frac{\pi}{2} + \delta^4 \end{aligned} \qquad (15-63)$$

Hence the dissipation factor is given by

$$d^2 = \frac{4 + \delta^2 - \sqrt{16 + 8\delta^2 \cos \frac{\pi}{2} + \delta^4}}{2} \qquad (15-64)$$

and α may be determined from eq. (15-58). The solution of these equations is plotted in Fig. 15-4.

We have therefore shown that a pair of single-tuned circuits of the proper dissipation factor, d , may be stagger-tuned to give an over-all selectivity characteristic, S_2 , defined by eq. (15-39). The quantity δ is the normalized half-power bandwidth of the combination:

$$\delta = \frac{(\Delta f)_T}{f_o} \quad (15-65)$$

where $(\Delta f)_T$ is the half-power bandwidth of the combination, and f_o is the center frequency of the combination. With these facts in mind

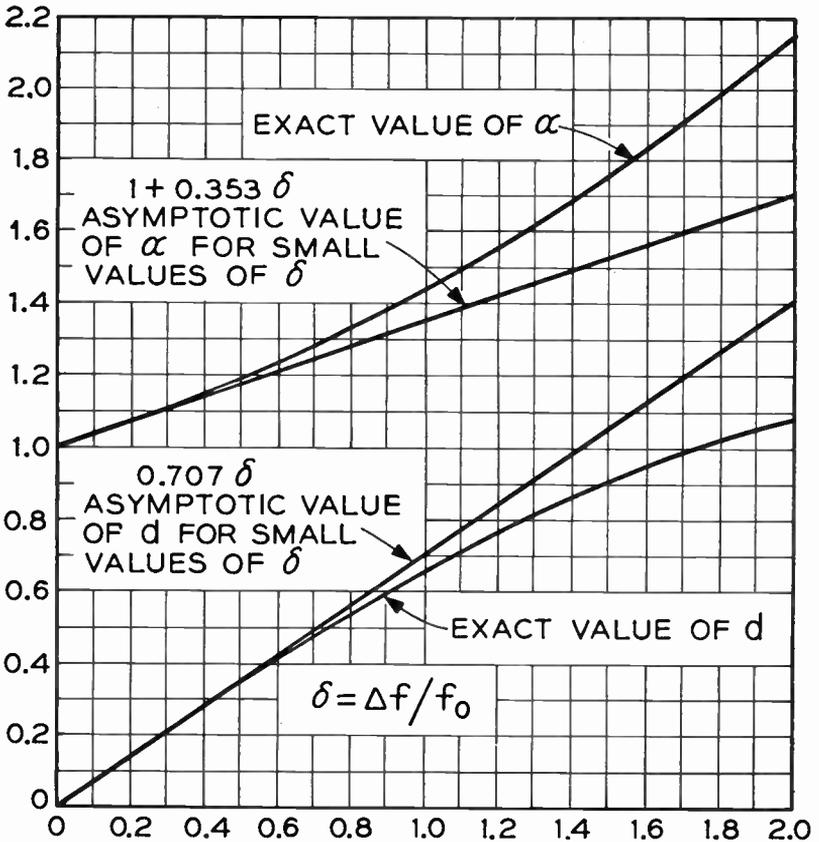


Fig. 15-4. Design curves for the staggered pair. (After Wallman.)

we are able to calculate the gain-bandwidth factor for the staggered pair; thus from (15-44)

$$A_T = 3\pi^2 \frac{1}{f_o^2} \frac{1}{\sqrt{\delta^4 + \left(F - \frac{1}{F}\right)^4}} \tag{15-66}$$

whence $(A_o)_T = 3\pi^2 \frac{1}{f_o^2 \delta^2}$ (15-67)

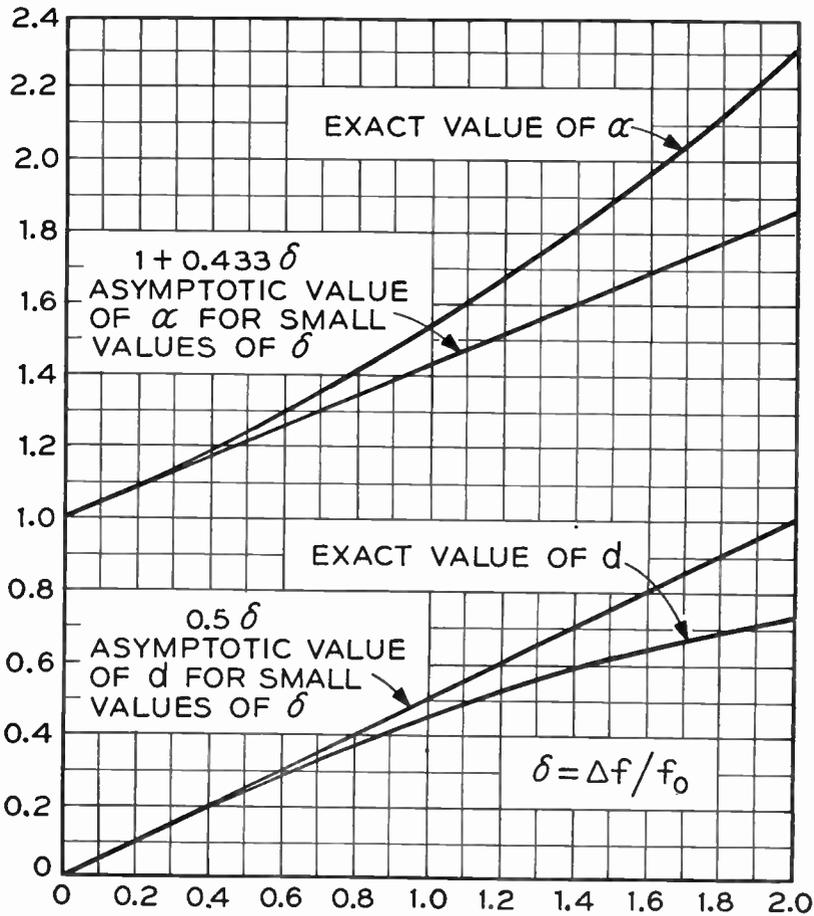


Fig. 15-5. Design curves for the staggered pair. (After Wallman.)

and, by (15-36),

$$\Gamma = \frac{(A_o)T^{1/2}}{\mathfrak{M}} f_o \delta = \frac{\mathfrak{M}}{f_o \delta} \frac{1}{\mathfrak{M}} f_o \delta = 1 \quad (15-68)$$

This latter relationship is of use in evaluating the gain of the staggered pair.

The use of the equations which have been derived may be illustrated by the following example. It is required to design a staggered pair centered at 30 megacycles and having a 25-megacycle half-power bandwidth. 6AK5's are to be used. $g_m = 5000 \mu\text{mhos}$ and $C_1 = C_2 = 12 \mu\text{mf}$. By eq. (15-65),

$$\delta = \frac{25}{30} = 0.834$$

Reading from Fig. 15-4, we obtain the design values

$$d = 0.565 \quad \text{and} \quad \alpha = 1.35$$

Hence we require two stages of $d = 0.565$ and stagger-tuned to the frequencies

$$f_{o1} = \alpha f_o = (1.35)(30) = 40.5 \text{ megacycles}$$

$$f_{o2} = \frac{f_o}{\alpha} = \frac{30}{1.35} = 22.2 \text{ megacycles}$$

From (15-7) and (15-20) the required loading resistors are

$$R_1 = \frac{1}{\omega_{o1} C_1 d} = \frac{1}{2\pi(40.5 \times 10^6)(12 \times 10^{-12})(0.565)} = 580 \text{ ohms}$$

and

$$R_2 = \frac{1}{\omega_{o2} C_2 d} = \frac{1}{2\pi(22.2 \times 10^6)(12 \times 10^{-12})(0.565)} = 1,058 \text{ ohms}$$

The values of inductance may be determined by means of eq. (15-3); thus

$$L_1 = \frac{1}{\omega_{o1}^2 C_1} = \frac{1}{[2\pi(40.5 \times 10^6)]^2(12 \times 10^{-12})} = 1.28 \mu\text{h}$$

$$L_2 = \frac{1}{\omega_{o2}^2 C_2} = \frac{1}{[2\pi(22.2 \times 10^6)]^2(12 \times 10^{-12})} = 4.3 \mu\text{h}$$

The gain of the pair at the center frequency, f_o , may be calculated in two ways. First, since $\Gamma = 1$ for the combination, we have from (15-36)

$$\Gamma = \frac{(A_o)_T^{1/2}}{\mathfrak{M}} (\Delta f)_T = 1$$

$$(A_o)_T = \left[\frac{\mathfrak{M}}{(\Delta f)_T} \right]^2 = \left[\frac{5 \times 10^{-3}}{2\pi(12 \times 10^{-12})(25 \times 10^6)} \right]^2 = 7.02$$

The second method involves the use of eq. (15-44); thus at f_o (*i.e.*, $F = 1$), the magnitude of the over-all gain is

$$\begin{aligned} (A_o)_T &= \frac{\mathfrak{M}^2}{f_o^2} \frac{1}{\sqrt{d^2 + \left(\frac{1}{\alpha} - \alpha\right)^2}} \frac{1}{\sqrt{d^2 + \left(\alpha - \frac{1}{\alpha}\right)^2}} \\ &= \left(\frac{\mathfrak{M}}{f_o}\right)^2 \frac{1}{d^2 + \left(\frac{1}{\alpha} - \alpha\right)^2} \\ &= \left(\frac{66.3}{30}\right)^2 \frac{1}{(0.565)^2 + (0.74 - 1.35)^2} = 7 \end{aligned}$$

It is clear that the first method, which notes that $\Gamma = 1$, is simpler. The amplification curves for the combination and the component amplifiers are plotted in Fig. 15-6. It may be seen that in contrast to the synchronously tuned case, cascaded stagger-tuned stages give an over-all bandwidth which is greater than the bandwidth of any

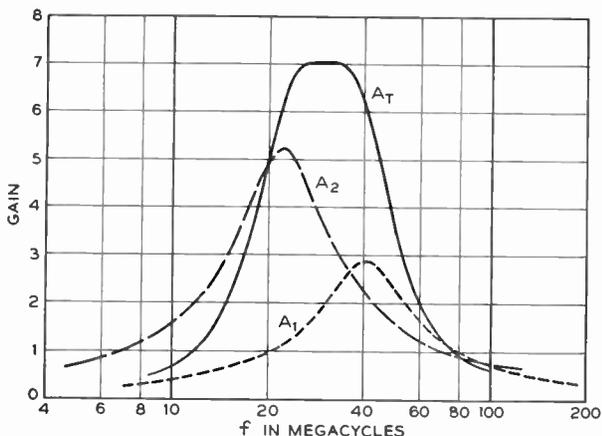


Fig. 15-6. Response of the staggered pair. $f_o = 30$ megacycles; $(\Delta f)_T = 25$ megacycles. A_1 and A_2 are the response curves of the component stages. A_T is the over-all response.

single component stage. The particular design shown in the diagram far exceeds the requirements of the television I-F amplifier in so far as bandwidth is concerned and illustrates the type of results which may be obtained with the stagger-tuning technique.

Mention should be made of gain control in the stagger-tuned amplifier. Reference to eq. (15-44) shows that the gain of the staggered pair is proportional to the product of the tube transconductances; hence, if one or more of the tubes is of the remote cutoff type, gain control is afforded by changing the bias voltage on those tubes. In this connection it is important to note that the variations in the tube g_m will not change the shape of the over-all response curve, provided that the Miller effect does not change the input capacitance of the tubes. This fact may be verified from eq. (15-44). The shape of the response curve is determined solely by the two bracketed form factors which are independent of g_m . Thus the curve shape is independent of the g_m factors. The same result may be obtained by reasoning physically from the curves of Fig. 15-6.

It is of interest to note that the stagger-tuned amplifier may be adjusted quite simply, provided that the Q_o or dissipation factor of each stage is correct. Since complete isolation of the tuned circuits is afforded by the intervening vacuum tubes, all adjustments are independent, and alignment is easily accomplished. For example, in the amplifier just designed a signal generator connected to the input terminals is set to $f_{o1} = 40.5$ megacycles and L_1 adjusted to give maximum output voltage. Either an oscilloscope or a voltmeter may be used to check the output. With the first stage tuned, the signal generator is reset to $f_{o2} = 22.2$ megacycles and L_2 adjusted to maximize the output once again.

15-3. The Staggered "n-uple"

The technique just described for synthesizing the S_2 selectivity function with a pair of stagger-tuned stages may be extended for higher values of n to give the function S_n of eq. (15-38). The analytical approach to the problem is a simple extension of the method used for the staggered pair; that is, we factor the denominator squared of S_n , namely, $(\delta^{2n} + x^{2n})$, in such a manner that each factor may be identified with the selectivity function of a single-tuned stage. As before, the factors are chosen so that the required

number of roots of -1 are forced into the expression. Thus, we factor the denominator of S_n in the following manner:

$$\delta^{2n} + x^{2n} = (x^2 - r_1\delta^2)(x^2 - r_2\delta^2) \cdots (x^2 - r_n\delta^2) \quad (15-69)$$

where r_l is the l th root of the n roots of -1 . A digression on the exponential form of these roots of -1 is in order. By Euler's identity we have that

$$e^{j\theta} = \cos \theta + j \sin \theta \quad (15-70)$$

Then, if $\theta = \pi$, the identity becomes

$$e^{j\pi} = -1 \quad (15-71)$$

and since the imaginary exponential is periodic in 2π , (15-71) may be rewritten as

$$-1 = e^{j(\pi+2l\pi)} = e^{j\pi(2l+1)} \quad (15-72)$$

where l is an integer.

Then, taking the n th root of both sides of the equation, we have

$$(-1)^{1/n} = e^{j\pi(2l+1)/n} \quad (15-73)$$

If, for example, $n = 5$, the corresponding roots of -1 will be

$$\left. \begin{aligned} l = 0 & \quad r_1 = e^{j\pi/5} \\ l = 1 & \quad r_2 = e^{j3\pi/5} \\ l = 2 & \quad r_3 = e^{j\pi} \\ l = 3 & \quad r_4 = e^{j7\pi/5} \\ l = 4 & \quad r_5 = e^{j9\pi/5} \end{aligned} \right\} \quad (15-74)$$

These five roots are shown plotted in the complex plane in Fig. 15-7. It is immediately apparent from the diagram that four of these roots are conjugate pairs; hence the five roots may be rewritten

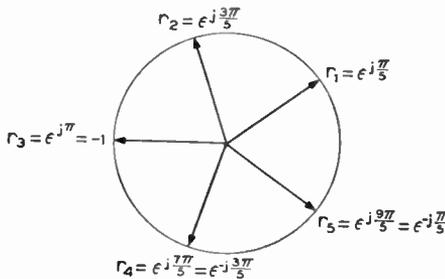


Fig. 15-7. The fifth roots of -1 .

Then, continuing the design procedure, we next seek a complex function of the form expressed in (15-52), which may be used to synthesize each of the bracketed factors. If this process be carried out as in the previous case, it may be shown that S_n is equal to

$$S_n = \frac{1}{\left| \left\{ \left[\delta^2 + j2\delta \sin\left(\frac{\pi}{2n}\right)x + (jx)^2 \right] \left[\delta^2 + j2\delta \sin\left(\frac{3\pi}{2n}\right)x + (jx)^2 \right] \right. \right. \\ \left. \left. \cdots \left[\delta^2 + j2\delta \sin\left(\frac{n-2}{2n}\pi\right)x + (jx)^2 \right] [\delta + jx] \right\} \right|} \quad (15-79a)$$

if n is odd;

$$S_n = \frac{1}{\left| \left\{ \left[\delta^2 + j2\delta \sin\left(\frac{\pi}{2n}\right)x + (jx)^2 \right] \left[\delta^2 + j2\delta \sin\left(\frac{3\pi}{2n}\right)x + (jx)^2 \right] \right. \right. \\ \left. \left. \cdots \left[\delta^2 + j2\delta \sin\left(\frac{n-1}{2n}\pi\right)x + (jx)^2 \right] \right\} \right|} \quad (15-79b)$$

if n is even

Now, we have seen that each of these factors except the last, where n is odd, may be synthesized by a properly designed staggered pair. We have also seen that $1/[\delta + j(F - 1/F)]$ is the response of a single-tuned stage. It follows at once that where n is odd, S_n may be synthesized by $(n - 1)/2$ staggered pairs plus a single-tuned stage centered on f_o , and where n is even, by $n/2$ staggered pairs. In either case n stages are required.

Expressed in terms of the component stages S_n is

$$S_n = \left| \left\{ \left[\frac{1}{d_1 + j\left(\frac{F}{\alpha_1} - \frac{\alpha_1}{F}\right)} \cdot \frac{1}{d_1 + j\left(\alpha_1 F - \frac{1}{\alpha_1 F}\right)} \right] \cdot \right. \right. \\ \left. \left[\frac{1}{d_3 + j\left(\frac{F}{\alpha_3} - \frac{\alpha_3}{F}\right)} \cdot \frac{1}{d_3 + j\left(\alpha_3 F - \frac{1}{\alpha_3 F}\right)} \right] \cdot \right. \\ \left. \cdots \left[\frac{1}{\delta + j\left(F - \frac{1}{F}\right)} \right] \right\} \right| \quad \text{if } n \text{ is odd;} \quad (15-80a)$$

$$S_n = \left\{ \left[\frac{1}{d_1 + j \left(\frac{F}{\alpha_1} - \frac{\alpha_1}{F} \right)} \cdot \frac{1}{d_1 + j \left(\alpha_1 F - \frac{1}{\alpha_1 F} \right)} \right] \cdot \left[\frac{1}{d_3 + j \left(\frac{F}{\alpha_3} - \frac{\alpha_3}{F} \right)} \cdot \frac{1}{d_3 + j \left(\alpha_3 F - \frac{1}{\alpha_3 F} \right)} \right] \cdot \dots \cdot \left[\frac{1}{d_{n-1} + j \left(\frac{F}{\alpha_{n-1}} - \frac{\alpha_{n-1}}{F} \right)} \cdot \frac{1}{d_{n-1} + j \left(\alpha_{n-1} F - \frac{1}{\alpha_{n-1} F} \right)} \right] \right\} \quad (15-80b)$$

if n is even

Comparison of (15-79) and (15-80) term by term shows the design constants to be

$$d_l^2 = \frac{4 + \delta^2 - \sqrt{16 + 8\delta^2 \cos \frac{l\pi}{n} + \delta^4}}{2} \quad (15-81)$$

and

$$\left(\alpha_l - \frac{1}{\alpha_l} \right)^2 + d_l^2 = \delta^2$$

where as before

$$\delta = \frac{(\Delta f)T}{f_o} \quad \alpha_l = \frac{f_{ol}}{f_o} \quad d_l = \frac{1}{\omega_{ol}RC} \quad (15-82)$$

The solutions of (15-81) for the staggered triple ($n = 3$) are plotted in Fig. 15-5.

15-4. Asymptotic Forms

Inspection of the design curves for the staggered pair and triple shows that as δ becomes 0.3 or less, d and α approach an asymptotic value. Under these conditions the solution of eqs. (15-81) becomes simpler and there is no need for the design curves. It is fortunate from the viewpoint of calculation that the values of δ required in television I-F amplifier design generally fall within the asymptotic

range. The corresponding solutions for d and α for this condition are now derived.

To illustrate the simplification of the design equations when $\delta \leq 0.3$ we shall carry through the calculations for the staggered triple. Thus, letting $n = 3$ and $\delta \leq 0.3$, eqs. (15-81) may be reduced as shown below. The δ^4 term under the radical is negligible; hence for $l = 1$

$$d_1^2 \approx \frac{4 + \delta^2 - \sqrt{16 + 8\delta^2(0.5)}}{2} \tag{15-83}$$

and, expanding the radical by means of the binomial theorem, we have

$$(16 + 4\delta^2)^{1/2} = 16^{1/2} + \frac{1}{2}(16)^{-1/2}(4\delta^2) + \dots \approx 4 + \frac{\delta^2}{2} \tag{15-84}$$

and (15-83) becomes

$$d_1^2 \approx \frac{4 + \delta^2 - 4 - \delta^2/2}{2} = \frac{\delta^2}{4} \tag{15-85}$$

or $d_1 \approx 0.5\delta$ (15-86)

Then $\left(\alpha_1 - \frac{1}{\alpha_1}\right)^2 = \delta^2 - 0.25\delta^2 = 0.75\delta^2$

$$\alpha_1 - \frac{1}{\alpha_1} = 0.865\delta$$

$$\alpha_1^2 - 0.865\delta\alpha_1 - 1 = 0 \tag{15-87}$$

whence

$$\alpha_1 = \frac{0.865\delta + \sqrt{0.75\delta^2 + 4}}{2} \approx \frac{0.865\delta + 2}{2} \approx 1 + 0.43\delta \tag{15-88}$$

Similarly, for $l = 3$,

$$d_3^2 = \frac{4 + \delta^2 - \sqrt{16 - 8\delta^2}}{2} \approx \frac{4 + \delta^2 - 4 + \delta^2}{2} \approx \delta^2 \tag{15-89}$$

therefore $d_3 \approx \delta$

and $\left(\alpha_3 - \frac{1}{\alpha_3}\right)^2 = \delta^2 - \delta^2 = 0$

whence $\alpha_3 = 1$ (15-90)

Equations (15-86), (15-88), (15-89), and (15-90) show that the selectivity function, S_3 , may be synthesized by three single-tuned stages designed as follows:

$$\text{Stage 1: dissipation factor} = 0.5\delta, \quad f_{o1} = (1 + 0.43\delta)f_o$$

$$\text{Stage 2: dissipation factor} = 0.5\delta, \quad f_{o2} = \frac{f_o}{(1 + 0.43\delta)}$$

$$\text{Stage 3: dissipation factor} = \delta, \quad f_{o3} = f_o$$

The asymptotic solutions for staggered n -uples up to $n = 5$ are listed in Table 15-1. Also shown are the design equations for the case when δ is so small that the selectivity function may be assumed to display arithmetic symmetry. Wallman has recommended the following ranges for each of the design formulas:

$$\delta < 0.05 \quad \text{Arithmetic case}$$

$$0.05 \leq \delta \leq 0.3 \quad \text{Asymptotic case}$$

$$0.3 < \delta \quad \text{Exact case}$$

To illustrate the use of the several design equations for the stagger-tuned I-F amplifier we shall consider the design of a stagger-tuned quintuple whose response is to be centered on 24 megacycles and which is to have a half-power bandwidth of 4 megacycles. 6AK5's are to be used. $g_m = 9000 \mu\text{mhos}$, and $C = 12 \mu\mu\text{f}$. We first calculate the normalized half-power bandwidth, δ , in order to find which design case obtains. Then, by (15-82),

$$\delta = \frac{(\Delta f)_T}{f_o} = \frac{4}{24} = 0.167$$

Thus, the asymptotic design equations may be used and the data obtained from Table 15-1, part B 5. The required center frequencies and dissipation factors of the five stages will be

$$f_{o1} = f_o(1 + 0.48\delta) = 24(1.08) = 25.9 \text{ megacycles}, \quad d_1 = 0.31\delta = 0.0516$$

$$f_{o2} = \frac{f_o}{1 + 0.48\delta} = \frac{24}{1.08} = 22.2 \text{ megacycles}, \quad d_2 = 0.31\delta = 0.0516$$

$$f_{o3} = f_o(1 + 0.29\delta) = 24(1.048) = 25.1 \text{ megacycles}, \quad d_3 = 0.81\delta = 0.135$$

TABLE 15-1

DATA FOR STAGGER TUNING (AFTER WALLMAN)

A. Arithmetic Symmetry, $\delta < 0.05$

| <i>n</i> | <i>Component stages</i> | |
|-------------|----------------------------|--------------------|
| | <i>Center frequency</i> | <i>Bandwidth</i> |
| 2 Pair | $f_o \pm 0.35(\Delta f)_T$ | $0.71(\Delta f)_T$ |
| 3 Triple | $f_o \pm 0.43(\Delta f)_T$ | $0.5 (\Delta f)_T$ |
| | f_o | $(\Delta f)_T$ |
| 4 Quadruple | $f_o \pm 0.46(\Delta f)_T$ | $0.38(\Delta f)_T$ |
| | $f_o \pm 0.19(\Delta f)_T$ | $0.92(\Delta f)_T$ |
| 5 Quintuple | $f_o \pm 0.48(\Delta f)_T$ | $0.31(\Delta f)_T$ |
| | $f_o \pm 0.29(\Delta f)_T$ | $0.81(\Delta f)_T$ |
| | f_o | $(\Delta f)_T$ |

B. Asymptotic Case, $0.05 \leq \delta \leq 0.3$

| <i>n</i> | <i>Center frequency</i> | <i>d</i> |
|-------------|------------------------------|-----------------------|
| | 2 Pair | $f_o(1 + 0.35\delta)$ |
| | $\frac{f_o}{1 + 0.35\delta}$ | 0.71δ |
| 3 Triple | $f_o(1 + 0.43\delta)$ | 0.5δ |
| | $\frac{f_o}{1 + 0.43\delta}$ | 0.5δ |
| | f_o | δ |
| 4 Quadruple | $f_o(1 + 0.46\delta)$ | 0.38δ |
| | $\frac{f_o}{1 + 0.46\delta}$ | 0.38δ |
| | $f_o(1 + 0.19\delta)$ | 0.92δ |
| | $\frac{f_o}{1 + 0.19\delta}$ | 0.92δ |
| 5 Quintuple | $f_o(1 + 0.48\delta)$ | 0.31δ |
| | $\frac{f_o}{1 + 0.48\delta}$ | 0.31δ |
| | $f_o(1 + 0.29\delta)$ | 0.81δ |
| | $\frac{f_o}{1 + 0.29\delta}$ | 0.81δ |
| | f_o | δ |

C. Legend

$(\Delta f)_T$ = over-all bandwidth

f_o = center frequency

$$\delta = \frac{(\Delta f)_T}{f_o}$$

$$d = \frac{1}{2\pi f_{on} RC}$$

f_{on} = center frequency of the *n*th stage

$$f_{o4} = \frac{f_o}{1 + 0.29\delta} = \frac{24}{1.048} = 22.9 \text{ megacycles,} \quad d_4 = 0.81\delta = 0.135$$

$$f_{o5} = f_o = 24 \text{ megacycles,} \quad d_5 = \delta = 0.167$$

The design of R_5 and L_5 for the fifth stage is calculated. From (15-82)

$$R_5 = \frac{1}{\omega_{o5} C_5 d_5} = \frac{1}{2\pi(24 \times 10^6)(12 \times 10^{-12})(0.167)} = 3,320 \text{ ohms}$$

$$\text{and } L_5 = \frac{1}{\omega_{o5}^2 C_5} = \frac{1}{[2\pi(24 \times 10^6)]^2(12 \times 10^{-12})} = 3.66 \mu\text{h}$$

The constants for the four remaining stages are computed in a similar manner.

We next calculate the selectivity function and the amplification. Our previous work shows that the normalized selectivity function of the staggered quintuple, S_5 , is

$$S_5 = \frac{1}{\sqrt{\delta^{10} + \left(F - \frac{1}{F}\right)^{10}}}$$

For purposes of calculation it is convenient to factor out the δ^{10} term from the radical; thus,

$$S_5 = \frac{1}{\delta^5 \sqrt{1 + \left[\frac{1}{\delta} \left(F - \frac{1}{F}\right)\right]^{10}}}$$

The calculations for evaluating this function are shown in Table 15-2.

TABLE 15-2

| $f_{m.c.}$ | F | $\frac{1}{F}$ | $\frac{f_o^2}{f_{m.c.}}$ | $\left(F - \frac{1}{F}\right)$ | $\frac{1}{\delta} \left(F - \frac{1}{F}\right)$ | $\left[\frac{1}{\delta} \left(F - \frac{1}{F}\right)\right]^{10}$ | S_5 |
|------------|-------|---------------|--------------------------|--------------------------------|---|---|-------|
| 24 | 1 | 1 | 24 | 0 | 0 | 0 | 1 |
| 25 | 1.040 | 0.961 | 23.1 | 0.079 | 0.475 | 0.00059 | 1 |
| 25.5 | 1.061 | 0.941 | 22.6 | 0.120 | 0.720 | 0.0375 | 0.984 |
| 26 | 1.082 | 0.924 | 22.2 | 0.158 | 0.950 | 0.599 | 0.792 |
| 26.5 | 1.102 | 0.906 | 21.75 | 0.196 | 1.177 | 5.1 | 0.405 |
| 27 | 1.124 | 0.890 | 21.4 | 0.234 | 1.405 | 30 | 0.179 |
| 27.5 | 1.144 | 0.874 | 20.95 | 0.270 | 1.621 | 125 | 0.090 |
| 28 | 1.168 | 0.856 | 20.58 | 0.312 | 1.875 | 540 | 0.043 |

These results are plotted in Fig. 15-8. Notice that if the local oscillator of the receiver which incorporates this I-F amplifier runs higher than the incoming R-F signal, the proper choice of visual I-F frequency is 26.35 megacycles in order that it may lie at the half-voltage point on the response curve. The aural carrier would be

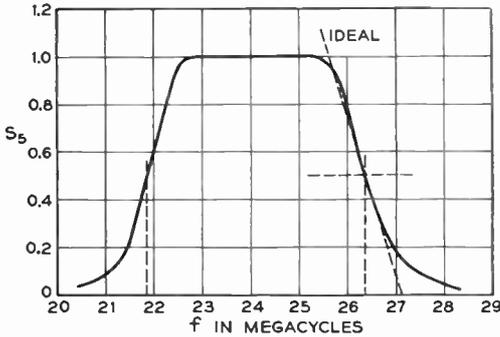


Fig. 15-8. Response of a staggered quintuple. $f_o = 24$ megacycles; $(\Delta f)_T = 4$ megacycles. The right-hand skirt matches the ideal I-F characteristic very well. Additional shaping may be obtained with traps.

$26.35 - 4.5 = 21.85$ megacycles. It should be noticed that the response in the region of the visual carrier closely approximates the ideal curve, which is linear in the range 26.35 ± 0.75 megacycles, and that the response is flat for approximately 3.60 megacycles. Trap circuits may be used to further shape the curve as required.

The gain of the five-stage amplifier may be calculated quite readily by noting that $\Gamma = 1$. Then, from eq. (15-36),

$$\Gamma = \frac{(A_o)_T^{1/5}}{\mathfrak{M}} (\Delta f)_T = 1$$

therefore $(A_o)_T = \left[\frac{\mathfrak{M}}{(\Delta f)_T} \right]^5 = \left[\frac{66.3}{4} \right]^5 = (16.58)^5$

or $\text{gain} = (5)20 \log 16.58 = 122$ decibels

In the usual case, one of the tuned circuits, say $L_5C_5R_5$, would serve as the load for the frequency mixer. In that case the gain will be less than calculated above because the conversion transconductance of the mixer is less than the g_m of the four other stages and hence the mixer figure of merit is less than that of the other stages.

CHAPTER 16

THE RECEIVING ANTENNA

Under current transmission standards the frequency of the radiated signal from the transmitting antenna lies in the very high frequency (V.H.F.) band. Some knowledge is therefore required of the propagation characteristics of this band of frequencies in order that the requirements on the receiving antenna may be understood. In the present chapter statement will be made of certain principal properties of V.H.F. propagation and of the properties of several common forms of receiving antennas used in that band.

16-1. Carrier Frequency Requirement

The theory of amplitude modulation presumes that the carrier frequency is so large compared to the frequency of the highest modulating-signal component that any one cycle of the carrier wave may be considered to be a true sinusoid. This requirement, which is tacitly assumed in writing the equation of the amplitude-modulated wave

$$e = A(1 + m \cos \omega_m t) \cos \omega t \quad (16-1)$$

may be met practically if the carrier frequency is at least ten times the highest modulating frequency. Then since the television video band extends up to approximately 4.5 megacycles, the minimum allowable television carrier frequency is some 45 megacycles, and television transmission must be confined to the V.H.F. or higher frequency bands.¹ The 12 channels authorized for commercial television transmission at present are located in the range from 54 to

¹ The designation of frequency bands used here is that adopted by the F.C.C. on March 2, 1943, namely,

| | |
|---------------------------|------------------------------|
| 3 to 30 megacycles | high frequency, H.F. |
| 30 to 300 megacycles | very high frequency, V.H.F. |
| 300 to 3000 megacycles | ultra high frequency, U.H.F. |
| 3000 to 30,000 megacycles | super high frequency, S.H.F. |

216 megacycles and thus are in the V.H.F. band. It is anticipated that some 40-odd additional U.H.F. channels in the vicinity of 500 to 900 megacycles will be opened for commercial transmission in the near future.

16-2. Elements of V.H.F. Propagation

In the general case of radio waves radiated from a transmitting antenna the radiated wave at a distance of several wavelengths from the antenna consists of two major components: a sky wave, which moves upward toward the ionosphere, where it may or may not be reflected back toward the earth at a point far removed from the antenna, and a ground wave, which stays close to the earth's surface. It is convenient to further subdivide the ground wave into space and surface components. In the broadcast band, where transmitting antennas are close to the ground in terms of wavelength the space component of the ground wave is canceled out and propagation in the normal service range of the transmitter results almost entirely from the surface component of the ground wave, which effectively travels along the earth's surface. The sky wave, which returns to the earth after reflection in the ionosphere, accounts for the sporadic reception of the signal at points beyond the normal range of the surface wave. It may be shown that the distance from the transmitter at which a given field strength may be received varies inversely as the frequency of the radiated signal; hence at higher frequencies, the surface wave is of negligible value in transmission.²

In the V.H.F. band the transmitted frequency is so high that the surface component of the ground wave has no effect at the receiving antenna. Furthermore, the bending effect of the sky wave in the ionosphere is not sufficient to reflect that wave toward the earth, so that the effect of the sky wave must be discounted in television transmission. It follows, then, that in the V.H.F. band transmission of the signal is by the space component of the ground wave. In contrast to the condition in the broadcast band, television antennas must be located several wavelengths above the earth's surface, and this space component is not canceled out except at certain points of interference.

² See, for example, W. L. Everitt, *Communication Engineering*. New York: McGraw-Hill Book Company, Inc., 1937, chap. 19.

Reference to Fig. 16-1 shows that the space wave may reach the receiving antenna by direct or reflected paths. Assuming a

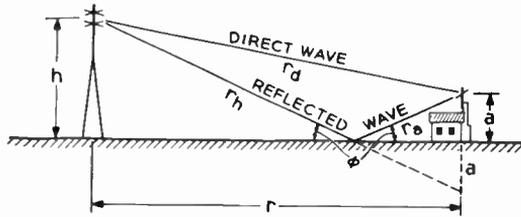


Fig. 16-1. The space wave may reach the receiving antenna by direct or reflected paths. Notice that the angles of incidence and reflection at the earth are equal.

flat-surfaced earth between the two antennas, the field strength at a receiving dipole, which is normal to the direction from the transmitting antenna, as a result of the direct wave component is

$$\epsilon_{d1} = \frac{7\sqrt{W'}}{r} \text{ volts/meter DIRECT WAVE (16-2)}$$

where W' is the radiated power in watts, and r is the distance between receiving and transmitting antennas.

Propagation by the direct wave is limited essentially to "line-of-sight" paths since the waves travel in a straight line, and the transmission is described as "quasi-optical." The maximum transmission distance for the line-of-sight propagation may be calculated quite easily if we assume a perfectly spherical earth with no intervening objects lying between the transmitting and receiving antennas. Thus, in Fig. 16-2, d , the optical distance to the horizon, may be calculated as follows: d is seen to be the length of the tangent to the circle between the transmitting antenna and the horizon. Then

$$(h + R)^2 = d^2 + R^2 \quad (16-3)$$

where R = radius of the earth \approx 4000 miles

Then $h^2 + 2hR + R^2 = d^2 + R^2$

and since $h^2 \ll R^2$

$$d = \sqrt{2hR} = \sqrt{2 \frac{h'}{(5280)} (4000)}$$

therefore
$$d = 1.23\sqrt{h'} \tag{16-4}$$

where d is the distance in miles, and h' is the antenna height in feet. It is usual practice in calculating the maximum distance of propagation to take into account the refraction of the wave by the earth's atmosphere. The refractive effect causes a slight curvature in the wave path so that the wave tends to follow the earth's surface. As a result the distance of propagation is slightly greater than the line-of-sight value and may be calculated by assuming the radius of the earth to be increased by some factor. Under normal atmospheric conditions the index of refraction in the atmosphere decreases linearly with height above the earth's surface, and this factor is taken to be $4/3$.³ Equation (16-4) then becomes

$$d = \sqrt{2h'} \tag{16-4a}$$

The same factor may also be applied to (16-5). If, now, the receiving antenna has a height of a' feet, the maximum line-of-sight transmission distance, r , may be seen from Fig. 16-2 to be

$$r = 1.23(\sqrt{h'} + \sqrt{a'}) \tag{16-5}$$

Returning to the flat-earth problem of Fig. 16-1 we see that if intervening objects do not prevent the reflected wave from reaching the antenna, the received field strength will depend on two, rather than on a single, components and eq. (16-2) will have to be modified. It might appear that any number of ground-reflection paths are possible. Since, however, a condition for reflection is that the angles of incidence and reflection be equal, one and only one point of reflection will allow the reflected wave to reach the antenna. Since the direct and reflected path lengths, r_d and $(r_k + r_a)$, respectively, are different, the two waves will not necessarily arrive in phase at the receiving antenna. Therefore we must calculate the phase difference produced by the path length difference, Δ . From Fig. 16-1,

$$r_d^2 = (h - a)^2 + r^2 \tag{16-6}$$

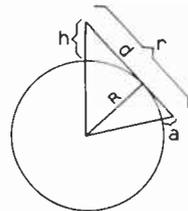


Fig. 16-2. Determination of the maximum distance of line-of-sight propagation on the spherical earth.

³ See, for example, D. G. Fink, *Radar Engineering*. New York: McGraw-Hill Book Company, Inc., 1947, chap. 4.

but since h and a are each much smaller than r , the first two terms of the binomial expansion for the square root may be used. Thus,

$$r_d = \sqrt{r^2 + (h - a)^2} \approx r + \frac{(h - a)^2}{2r} \quad (16-7)$$

Also, from the diagram,

$$(r_h + r_a)^2 = (h + a)^2 + r^2 \quad \text{or} \quad (r_h + r_a) \approx r + \frac{(h + a)^2}{2r} \quad (16-8)$$

Then the difference in path lengths is

$$\Delta = (r_h + r_a) - r_d \approx r + \frac{(h + a)^2}{2r} - r - \frac{(h - a)^2}{2r} = \frac{2ah}{r} \quad (16-9)$$

The path difference expressed in wavelengths is

$$\Delta_l = \frac{\Delta}{\lambda} = \frac{2ah}{\lambda r} \quad (16-10)$$

When r , the separation between the two antennas, is large, ϕ will be small and a 180° phase reversal of the wave will occur at reflection. Then the total difference in phase between the direct and reflected waves at the receiving antenna will be

$$\psi = 2\pi\Delta_l + \pi = \frac{4\pi ah}{\lambda r} + \pi \quad (16-11)$$

The resultant field strength at the receiving antenna will be the vector sum of the two components ϵ_r and ϵ_d , which differ in phase by ψ radians. We shall assume that the reflected wave suffers zero

attenuation in the reflecting process. Then, if the path length difference, Δ , were equal to zero, the direct and reflected waves would arrive 180° out of phase at the receiving antenna and would cancel out. This condition is illustrated in Fig. 16-3a. It is precisely this difference in path, Δ , which makes reception of the signal possible.

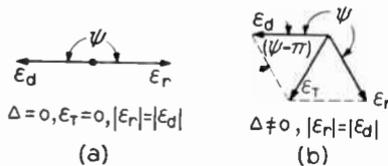


Fig. 16-3. Phase relationships at the receiving antenna. (a) The direct and reflected signals cancel if $\Delta = 0$. (b) ϵ_T is the vector sum of the direct and reflected components.

When Δ is different from zero or an integral multiple of the wavelength, the two components add vectorially, and the resultant field strength is

$$\varepsilon_T^2 = \varepsilon_r^2 + \varepsilon_d^2 - 2\varepsilon_r\varepsilon_d \cos(\psi - \pi) \quad (16-12)$$

and if, as before, zero attenuation is assumed at reflection,

$$|\varepsilon_r| = |\varepsilon_d| \quad (16-13)$$

and ε_T becomes

$$\begin{aligned} \varepsilon_T^2 &= 2\varepsilon_d^2[1 - \cos(\psi - \pi)] \\ &= 2\varepsilon_d^2 \cdot 2 \sin^2\left(\frac{\psi - \pi}{2}\right) \end{aligned}$$

or

$$\begin{aligned} \varepsilon_T &= 2\varepsilon_d \sin\left(\frac{\psi - \pi}{2}\right) \\ &= 2\varepsilon_d \sin\left(\frac{2\pi ah}{\lambda r}\right) \end{aligned} \quad (16-14)$$

It is clear from the last equation that maximum reinforcing of the two components will occur when Δ is some odd multiple of a half-wavelength.

Where Δ is small, the sine may be replaced by the angle and ε_T becomes

$$\varepsilon_T = 2\varepsilon_d \frac{2\pi ah}{\lambda r}$$

and, substituting from (16-2)

$$\varepsilon_T = \frac{88\sqrt{W}ah}{\lambda r^2} \quad (16-15)$$

It must be stressed that eqs. (16-2) and (16-15) presuppose a flat perfectly conducting earth and neglect any bending of the wave path due to changes in the refractive index of the atmosphere. Additional equations for field strength, which take into account the curvature of the earth's surface and finite ground conductivity, are available in the literature.^{4,5}

Certain practical factors modify all of the statements which have been made to some extent. For example, some bending of the space wave takes place as it moves away from the transmitting antenna and ranges up to 15 per cent greater than the line-of-sight value may be obtained. It is also found in practice that satisfactory re-

⁴ B. Trevor and P. S. Carter, "Notes on Propagation of Waves Below Ten Meters in Length." *Proc. IRE*, 21, 3 (March 1937).

⁵ Curves for predicting the ground-wave signal range for television are available in Appendix IV, *Standards of Good Engineering Practice Concerning Television Broadcast Stations*, F.C.C., 1945.

ception may be obtained even when the path between the two antennas is obstructed by a hill. The probable explanation here is that the waves are bent around the top of the hill and so tend to follow the earth contour to some extent.

The bending of the space wave also may give rise to what is known as anomalous propagation when the distance of transmission exceeds the calculated line-of-sight value by an extremely large margin. While such long-distance anomalous propagation is of little use for normal reception because of its irregularity of occurrence, it is becoming recognized as an important factor in station channel assignment because it may result in serious co-channel interference at certain times. We may illustrate this problem of co-channel interference by citing a specific example. WNHC-TV in New Haven, Conn., and WFIL-TV in Philadelphia, Penna., are both assigned to channel No. 6. The distance between the two cities is roughly 150 miles; therefore one would expect under normal conditions when eq. (16-4a) applies that no co-channel interference would be caused by these two stations. For example, a receiver located in the service area of New Haven would not receive any signal from the Philadelphia station because it lies beyond the horizon of the latter. Under certain atmospheric conditions, however, anomalous propagation may take place and the horizon of the Philadelphia station may move outward to include New Haven. As a result the New Haven receiver receives both signals with sufficient amplitude to cause serious interference.

Let us consider the mechanism of anomalous propagation in a qualitative fashion.⁶ The refractive index of any medium may be defined as the ratio of two velocities of wave propagation, namely, the velocity of propagation in a vacuum divided by the velocity of propagation in the medium. It may be seen at once that low values of the index in the vicinity of unity correspond to high velocities in the medium. The refractive index of the earth's atmosphere depends upon the densities of the several gases which constitute the atmosphere. In general, these densities decrease at higher elevations above the earth's surface; hence we might expect that the velocity of propagation increases with elevation. Thus, as a radio wave leaves an antenna, it moves away from the earth's spherical surface and into the less dense regions of the atmosphere where its velocity

⁶ For a quantitative treatment see, for example, D. G. Fink, *op. cit.*

increases. The net effect is that the wave is bent downward toward the earth. The degree of bending depends upon the exact manner in which the refractive index varies with altitude. Under normal conditions the rate of change is constant at a value of -1.2×10^{-8} per foot and, as we have stated, the horizon distance may be calculated by assuming a correction factor of $4/3$ for the radius of the earth. Under other atmospheric conditions, the rate of change of refractive index may have a larger numerical value, and the distance to the effective horizon increases. In fact, if the rate is -4.8×10^{-8} per foot, no horizon as such exists and the waves follow the earth's surface at a constant altitude. Between these extremes, the distance to the horizon may vary widely and anomalous propagation takes place. The same effect may also be observed when the rate of change is nonlinear with altitude.

Reflection of the space wave may also occur at surfaces other than that of the earth. For example, a tall building may present a surface sufficiently wide in terms of a wavelength that it serves as an efficient reflector of the television waves. When one or more of these waves reach the receiving antenna in addition to the direct wave, multipath reception is said to exist. Since the reflecting surface may be relatively far removed from both of the antennas, the difference in length between the direct and reflected paths may be so great that the image resulting from the reflected wave lags the main image by an appreciable portion of the horizontal line interval. Under this condition the secondary (reflected) image will appear displaced from the primary image on the receiver cathode-ray tube and is called a ghost image or simply a ghost. In general, the amplitude of the ghost is smaller than that of the primary image because of attenuation at the reflecting surface and over the longer path of transmission. Under certain conditions it is possible to locate the probable position of the surface that is causing the ghost.

Let the transmitting and receiving antennas be considered as the foci of an infinite family of ellipses, one of which is shown in Fig. 16-4. Now, by the definition of an ellipse,

$$r_T + r_R = \text{constant} = \overline{AA'} \quad (16-16)$$

hence it follows at once that the difference in direct and reflected paths is constant for reflections from any surface lying on the ellipse. Thus, any such reflecting surface on the ellipse will produce a ghost

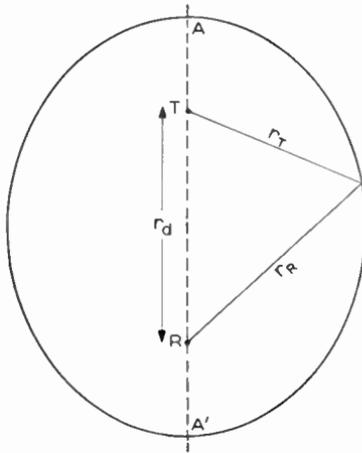


Fig. 16-4. An ellipse of reflection.
 $r_T + r_R = \text{constant} = \overline{AR}$.

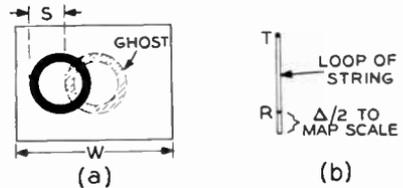


Fig. 16-5. An ellipse may be drawn on a map to locate the reflecting surface that causes the ghost. (a) Determination of s , the separation between the primary and ghost images on the cathode-ray tube screen. (b) Method of determining the length of string required for drawing the ellipse of reflection corresponding to s .

which lags the primary image by a certain definite amount. Since there are theoretically an infinity of these ellipses, the possibility of multiple-path reception is large. (Contrast this with reflections from the earth's surface.) In practice the number of bothersome reflections is small because the transmitting antenna is much higher than most objects in its vicinity. It should be apparent at once, however, that the effective number of reflected waves at the receiving antenna may be reduced by utilizing a directive, rather than an omnidirectional, receiving antenna.

The ellipse of reflection described in the preceding paragraph may be used to locate the probable source of reflection, and this, in turn, may help in proper orientation of the directive receiving antenna. Using the notation of Fig. 16-5 we let s be the separation on the cathode-ray tube screen between the primary and ghost images.

$$w = \text{picture width} \quad (16-17)$$

From Fig. 11-11, the horizontal unblank interval is

$$(\tau_u)_h = H - 0.16H = 0.84(63.5) \approx 53 \mu\text{sec} \quad (16-18)$$

and in this interval a radio wave travels a distance

$$d = c(\tau_u)_h = (186,293)(53 \times 10^{-6}) = 9.87 \text{ miles} \quad (16-19)$$

Then, if Δ is the difference between the direct and reflected paths which causes the image displacement s , we may write

$$\frac{\Delta}{s} = \frac{d}{w}$$

or
$$\Delta = 9.87 \frac{s}{w} \quad (16-20)$$

With Δ known the reflecting surface may be located on a map of the area. Two pins are placed on the map, one at the transmitter and one at the receiver locations. A loop of string is placed around the pins and its length adjusted as shown in *b* of Fig. 16-5. Then a pencil may be inserted into the loop and used to describe an ellipse. The reflecting surface will then be a large building, tower, or other large object which lies on the ellipse. When its location is known, the antenna may be oriented to minimize the ghost signal.⁷

16-3. General Requirements for the Receiving Antenna

We have seen that the use of a directive receiving antenna reduces the possibility of multiple-path reception. The directive antenna further has the advantage of reducing the strength of unwanted signals coming from directions other than that to the television transmitter. Thus, directivity is a general requirement of the receiving-antenna system.

Unfortunately in large urban communities the problem of directivity is complicated. When several transmitters are located throughout the city, a receiving antenna may be located between them and a single directive array cannot give optimum results. When this condition exists best results are obtained with a number of antennas, each oriented for a particular transmitter. Such a system is usually cumbersome and expensive, and a simple dipole is used as a compromise; in this case it is hoped that the field strengths will be sufficiently great from all the stations to provide a satisfactory picture.

In the suburban areas a more favorable condition exists because most of the transmitters will lie within a sector of 10° to 30° and a directive antenna system may be used to good advantage.

⁷ The method of using an ellipse to locate the reflecting surface was described to the author in 1946 by Mr. Robert Serrell, then of the Columbia Broadcasting System. It has subsequently appeared in print: J. R. Meagher, "Television Antennas and Transmission Lines," Part II. *RCA Radio Service News*, XIV, 1 (January-February 1949).

The second general requirement of the receiving antenna is that it be broadband, or rather *very* broadband, for in contrast to the transmitter bandwidth of 6 megacycles the receiving antenna must operate efficiently over the entire range of 54 to 216 megacycles, which covers the 12 authorized channels. In order to ease this requirement it has been the common practice to employ two antennas, one for the low (54- to 88-megacycle) and one for the high (174- to 216-megacycle) band.

The third general requirement is that a reasonably good impedance match be maintained between the antenna and the lead-in transmission line to the receiver input terminals. If a correct match is maintained at the receiver input, a mismatch of 2 to 1 at the antenna will still provide satisfactory results.

BASIC ANTENNA STRUCTURES

In the following sections we shall consider the properties of certain basic antenna structures. In discussing the directivity patterns we shall make use of the Rayleigh-Carson reciprocity theorem, which states that if a voltage E applied to antenna 1 causes a complex current I in an antenna 2, the same current I will flow in antenna 1 if E be applied to antenna 2. This generalization of the reciprocity theorem of linear-circuit theory effectively states that the directivity of an antenna receiving a signal is the same as the directivity of that antenna when it acts as a transmitter. This theorem is of considerable aid because it is frequently easier to derive the directivity pattern of a system when it is driven than when it is receiving.

16-4. The Dipole

The most basic antenna element in common use is the dipole or conductor of half-wavelength, which has been described in Chapter 13. It has a nominal radiation resistance (or internal resistance) of 72 ohms and has the figure-eight directivity pattern which is plotted in Fig. 13-29. In order to determine the type of problem involved with the dipole as a television receiving antenna, we shall consider that a single antenna of this type is to be used for all 12 television channels. Its electrical length should be chosen to be one half-wavelength long at f_m , the geometric mean of the bandwidth

to be received.⁸ Since the limits are 54 and 216 megacycles, the mean will be

$$f_m = \sqrt{(54)(216)} = 108 \text{ megacycles} \quad (16-21)$$

As the frequency of the applied signal is changed the electrical length of the antenna will also change. Thus, at 54 megacycles, the lower limit of the band, its length will be

$$l_{54} = \frac{\lambda}{2} \frac{54}{108} = \frac{\lambda}{4} \quad (16-22)$$

and at the upper limit of 216 megacycles the length will be

$$l_{216} = \frac{\lambda}{2} \frac{216}{108} = \lambda \quad (16-23)$$

Thus the dipole cut for 108 megacycles changes in electrical length from channel to channel and ranges from $\lambda/4$ to λ . Since the directivity pattern of the antenna depends upon its electrical length, the dipole's pattern will change from channel to channel. Three typical patterns are illustrated in Fig. 16-6.⁹ It may be seen at once that

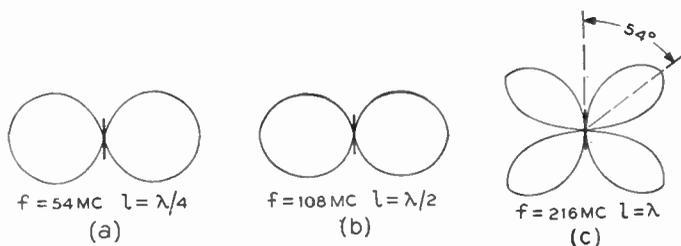


Fig. 16-6. The directivity pattern of a dipole of fixed length changes with the frequency of the received signal. The dipole is cut for 108 megacycles. (a) 54 megacycles: $l = \lambda/4$. (b) 108 megacycles: $l = \lambda/2$. (c) 216 megacycles: $l = \lambda$.

⁸ The choice of the geometric mean of the received band, f_m , as the design value for determining the antenna length is frequently encountered in the literature. It will be seen in Fig. 16-7, however, that such a design value discriminates against the lower-frequency stations because the antenna radiation resistance decreases as its electrical length is shortened. Since the radiation resistance is the ratio of received power to the square of the current at the dipole terminals, this means that the antenna is less efficient in converting received power to current. In general, then, it is desirable to cut the antenna to be $\lambda/2$ long at the lowest channel to be received. In the remainder of this chapter, however, we shall follow the common practice of specifying dipole length in terms of f_m .

⁹ Equations for calculating these patterns are available in a number of sources. See, for example, J. G. Brainerd *et al.*, *Ultra-High-Frequency Techniques*. New York: D. Van Nostrand Company, Inc., 1942, chap. 12.

if a station operating on channel No. 13 lies on a line normal to that of the antenna, the received signal will be small. Since the patterns at lower frequencies are relatively broad, some compromise may be effected by orienting the antenna so that one of the 216-megacycle lobes is aimed at the channel No. 13 transmitter.

Further complication results from the variation of the dipole impedance as the frequency is varied. Curves showing this effect are sketched in Fig. 16-7. The sharp peaks may be reduced by

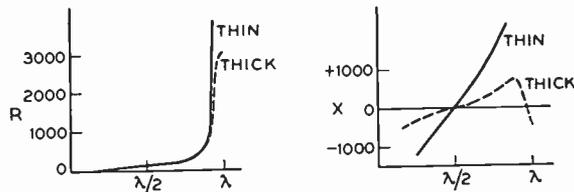


Fig. 16-7. Input impedance at the center of a dipole. (a) Resistive component. (b) Reactive component.

increasing the diameter of the dipole. Two ameliorative measures immediately suggest themselves. Either a dipole of relatively large diameter may be used or two or more dipoles may be used to cover the 12 channels. In regard to the first suggestion wind and ice loading effects generally restrict the dipole diameter to $\frac{3}{8}$ or $\frac{1}{2}$ in. Where the antenna is mounted indoors, for example in an attic, the dipole may be expanded into a biconical shape. An indoor antenna of this type, formed from metalized cardboard, is available commercially. Other systems for broadbanding the dipole without increasing its diameter are discussed in later sections.

The second solution to the broadband-directivity problem, that of using two antennas, has been used to a large extent. Since channels No. 6 and No. 7 are separated by a gap from 88 to 174 megacycles, which is assigned to commercial frequency modulation broadcasting and other nontelevision services, it is convenient to provide two antennas, one for the low band, 54 to 88 megacycles and covering channels No. 2 through No. 6, and one for high band, 174 to 216 megacycles and covering the remaining seven television channels. The mean frequencies for which the dipoles are cut are calculated below:

$$\begin{array}{ll}
 \text{Low-band} & \text{High-band} \\
 f_m = \sqrt{(54)(88)} & f_m = \sqrt{(174)(216)} \\
 = 69 \text{ megacycles} & = 194 \text{ megacycles}
 \end{array} \quad (16-24)$$

$$\begin{array}{ll}
 l_{64} = \frac{\lambda}{2} \frac{54}{69} = 0.392\lambda & l_{174} = \frac{\lambda}{2} \frac{174}{194} = .45\lambda \\
 l_{88} = \frac{\lambda}{2} \frac{88}{69} = 0.639\lambda & l_{216} = \frac{\lambda}{2} \frac{216}{194} = .557\lambda
 \end{array} \quad (16-25)$$

$$l_m = \frac{468}{f_{m(\text{m.c.})}} \text{ feet} \quad (16-26)$$

Equations (16-25) show that for either antenna the percentage change in electrical length over the frequency band is small and consequently the directivity pattern will retain a figure-eight shape with maximum pickup along the normal to the antenna. The small percentage change in l over the band also holds the antenna impedance variation within smaller limits than those for a single dipole covering both bands.

When the two-dipole system is used, some form of switching must be provided so that the receiver will be connected to the correct element for any channel being received. Where the use of two separate lead-ins is feasible, a double-pole double-throw switch may be used at the receiver proper. Alternatively, switching may be accomplished electrically at the antennas with a single lead-in connecting the switch to the receiver terminals. One form of this system applies a 24-volt, 60-cycle pulse to the lead-in to activate the remotely located switch. Isolation of the receiver from the 60-cycle switching pulse is accomplished by connecting the lead-in to the receiver input through two small isolating condensers.¹⁰ Similarly the television R-F signal is blocked from the 24-volt supply by means of R-F chokes.

A far more convenient system for the television-receiver owner completely eliminates the need for manual switching, the change-over from one antenna to the other being accomplished by a transmission-line crossover network.¹¹ One such self-switching system is illus-

¹⁰ D. A. Griffin, "Mast Head Antenna Switching." *Radio and Television News*, 42, 3 (September 1949).

¹¹ L. H. Finneburgh, "Evolution of a T.V. Antenna." *Radio and Television News*, 42, 1 (July 1949).

trated in Fig. 16-8. In the figure the subscripts h and l indicate the

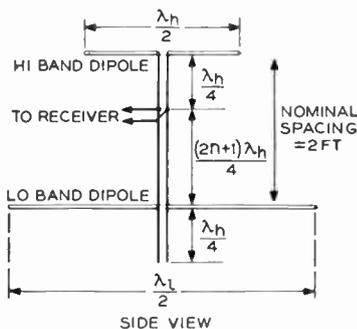


Fig. 16-8. A transmission-line crossover network for a hi-lo band antenna system. A 72-ohm line is used. The short, upper antenna is for channels No. 7 to No. 13; the lower, longer one for channels No. 2 to No. 6.

geometric mean of the high- and low-frequency bands, respectively. The crossover network is needed primarily to prevent the low-band dipole from marring high-band reception. For example, the length of the longer dipole at f_h is

$$l = \frac{\lambda}{2} \frac{194}{69} = 1.405\lambda \quad (16-27)$$

hence over a portion of the high band its pattern will break up into four or even six lobes, thereby increasing the chance of multipath reception. Thus, if the two antennas were connected directly to the lead-in, the low-band antenna would have an adverse effect on

high-band reception. On the other hand, the shorter antenna has a length at f_l

$$l = \frac{\lambda}{2} \frac{69}{194} = 0.178\lambda \quad (16-28)$$

It still exhibits the basic figure-eight directivity pattern and will not mar the low-band reception. The crossover network of Fig. 16-8 serves, then, to isolate the long dipole from the lead-in at the high-band frequencies. The design is a compromise based on f_h . The quarter-wave open-circuited stub, which is connected to the long dipole, shorts out that dipole at f_h . This short circuit, in turn, is reflected as an open circuit to the lead-in, which is located an odd number of quarter-wavelengths away.¹² The value of $3\lambda_h/4$ is normally used because a quarter-wavelength section does not have the necessary physical length to reach from the longer dipole to the lead-in connection. The connection between the lead-in and the high-band dipole is of length $\lambda_h/4$. If the transmission line has a characteristic impedance equal to the antenna resistance (72 ohms), this quarter-wave section serves as a 1 to 1 transformer at f_h . At other

¹² The impedance for the quarter-wave line may be checked in Fig. 12-15.

high-band frequencies the same condition will nearly be true. This may be verified by eq. (12-50) for the input impedance of a lossless transmission line, which is repeated below for convenience.

$$Z_{\text{in}} = R_o \left[\frac{Z_R + jR_o \tan \frac{2\pi l}{\lambda}}{R_o + jZ_R \tan \frac{2\pi l}{\lambda}} \right]$$

Under the conditions specified,

$$\left. \begin{aligned} Z_R &\approx R_o \\ l &\approx \lambda/4 \end{aligned} \right\} (16-29)$$

and the input impedance becomes approximately

$$Z_{\text{in}} \approx \frac{R_o^2}{Z_R} \approx R_o \quad (16-30)$$

Since the two dipoles are relatively close together, their driving-point impedances are affected by the mutual impedance between them. A number of compromises in antenna spacing and crossover networks have been used in commercial antennas. The dimensions shown in Fig. 16-8 are typical. In some instances a more complex crossover network is used so that F-M band signals received by the two antennas cancel at the lead-in, thereby reducing image effects caused by image interference from the 88- to 108-megacycle band.

A number of techniques for broadbanding the basic dipole have been incorporated in the manufacture of receiving antennas. These are discussed after the second basic receiving-antenna element, the folded dipole.

16-5. The Folded Dipole

The second basic element used in receiving antennas is the folded dipole. Illustrated in Fig. 16-9a, it consists of two simple dipoles, joined together at their ends, and in comparison to a dipole has a higher driving-point impedance and better broadband characteristic. Roberts¹³ has presented an analysis which demonstrates both of these properties of the folded unit. Because of the symmetry afforded by the ground plane, the actual folded dipole may be replaced by a

¹³ W. van B. Roberts, "Input Impedance of a Folded Dipole." *RCA Review*, VIII, 2 (June 1947).

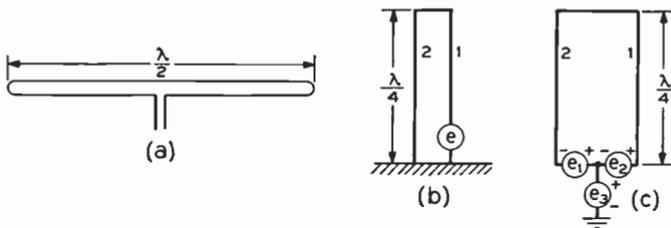


Fig. 16-9. The folded dipole. (a) Basic unit. (b) Equivalent folded unipole. (c) Equivalent circuit. (Courtesy of RCA Review.)

folded quarter-wave unipole above the ground plane. We shall consider both conductors of either folded unit to be of the same diameter and the spacing between them small as compared to a wavelength. Under these conditions the resistance R of each element is essentially the same as if that element were alone in free space.

Roberts has suggested the equivalent circuit shown at *c* in the figure to aid the analysis. Notice that the left-hand element, 2, remains grounded because the two equal voltages e_1 and e_3 are in phase opposition. The input resistance of the folded unit may now be calculated with the aid of the superposition theorem. Thus, let $e_3 = 0$. Under this condition we have a voltage $2e$ applied to a quarter-wave section, short-circuited at the far end. Since the input impedance of such a section is infinite, only a negligible current flows in element 1.

As a second step in the application of the superposition theorem we let e_1 and e_2 each be zero and calculate the current flow resulting from voltage e_3 . Under this condition the two elements are joined at both ends, appear in parallel, and so have an input impedance R . The current which flows in the combination will be

$$i_3 = \frac{e}{R} \quad (16-31)$$

and since the two elements in parallel are identical, i_3 will divide equally between them. Thus the current in element 1 because of e_3 is

$$(i_3)_1 = \frac{i_3}{2} = \frac{e}{2R} \quad (16-32)$$

Since zero current flows in element 1 because of e_1 and e_2 , $(i_3)_1$ of

(16-32) is the total current in that element produced by all three generators. The total voltage applied to element 1 is

$$e_2 + e_3 = 2e \quad (16-33)$$

hence the input impedance of element 1 is

$$Z_{in} = \frac{2e}{(i_3)_1} = \frac{2e \cdot 2R}{e} = 4R \quad (16-34)$$

Therefore, when both elements of the folded dipole are of the same diameter and are close together, the input impedance is four times that of a simple dipole, or approximately 300 ohms. Other values of input impedance may be obtained by increasing the diameter of one of the elements or by using more than two elements in parallel, but to date these variations have found little use in television work.

Roberts' analysis also gives the key to the broadbanding effect of the folded unit. By way of illustration, let the frequency of the applied voltage decrease so that the unipole length is less than $\lambda/4$. Then with $e_3 = 0$, e_1 and e_2 in series see a finite inductive impedance, and a lagging current will flow in element 1. Furthermore the magnitude of this current will increase as the frequency is lowered. With e_1 and e_2 both equal to zero, the length of the open-circuited section consisting of the two elements in parallel will be less than $\lambda/4$, and e_3 will see a capacitive impedance. Reference to Fig. 12-15 shows that as the frequency is lowered, the magnitude of this impedance will increase; thus $(i_3)_1$ leads the applied voltage and decreases with lowered frequency. The net effect is that the two reactive current components in element 1 tend to cancel and keep the input impedance more constant over variations in frequency. The effect is similar to that obtained with a broadbanding antiresonant circuit shunted across the input of the simple dipole.

As far as directivity is concerned, the two elements of the folded dipole are so close together that they appear as a single radiating element and so have the same directivity patterns as the simple dipole. Thus the techniques described in the last section of utilizing two separate antennas to cover the high and low channel bands apply equally well, even when the antennas are of the folded dipole form.

A distinct advantage of the folded-dipole receiving antenna is that it is readily amenable to lightning-protection methods. In a day when indoor broadcast reception antennas are the rule, the public

is prone to forget that an outdoor television antenna usually located above the highest point on a house provides a formidable lightning hazard and that adequate protection should be provided. In the folded dipole a voltage node occurs at the center of element 2 and so that point may be connected through the supporting mast to ground without impairing the receiving characteristics of the antenna. With the direct ground connection excellent lightning protection is afforded and no separate protective device is required.

16-6. The Parasitic Element

The figure-eight directivity pattern of the simple and folded dipoles leaves something to be desired, particularly where the antenna is used in a suburban area where all the transmitters lie within a relatively small sector. Under this condition the back lobe of the pattern, *i.e.*, the one in the direction opposite of that to the transmitters, serves no useful purpose and increases the possibility of interference from unwanted signals. Partial suppression of the rear lobe may be obtained with a parasitic or nondriven dipole element. The basic configuration is shown in Fig. 16-10.

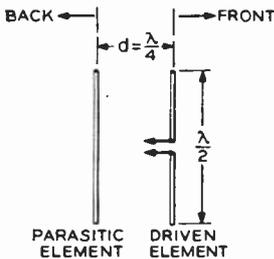


Fig. 16-10. A parasitic reflector may be used to improve horizontal directivity by suppressing the back lobe of the dipole radiation pattern. Top view of the array is shown.

In order to explain the operation of the array on a physical basis we shall assume zero mutual impedance between the two elements and make use of the Rayleigh-Carson reciprocity theorem. Thus we assume that power is supplied to the front, or driven, element.

The application of power to the front dipole causes a current I_d to flow and waves of equal intensity and in phase with I_d are radiated in the front and back directions. In traveling a quarter-wavelength to the parasite, the back wave, W_b , suffers a phase lag of ϕ_1 .

$$\phi_1 = \frac{2\pi d}{\lambda} = \frac{2\pi \lambda}{\lambda} \frac{1}{4} = \frac{\pi}{2} \text{ radians} \quad (16-35)$$

The back wave induces a voltage, E_p , in the parasite, which lags W_b by π radians. Under the assumption of zero mutual impedance

the parasitic element behaves like a series resonant circuit and I_p , the current due to E_p , is in phase with E_p . Thus I_p lags I_d by $3\pi/2$ radians. In effect, then, the system is equivalent to two dipoles, spaced one quarter-wavelength apart and driven 270° out of phase. I_p also causes equal waves to be radiated in both directions, those waves being in phase with I_p .

In the forward direction the parasitic's wave lags the driven wave by $3\pi/2$ because of the phase difference between I_d and I_p , plus $\pi/2$ due to the spacing; hence the two waves are in phase and add in the forward direction. Toward the rear the parasitic wave lags the driven wave by $3\pi/2$ minus $\pi/2$ due to the antenna spacing, and the two waves appear in phase opposition and cancel. By this simplified description we see that the parasitic reflector suppresses the rear lobe and acts to "reflect" the energy in the forward direction. By virtue of the Rayleigh-Carson theorem, the same action obtains when the array is used to receive the signal; maximum pickup occurs toward the front, whereas pickup from the back direction is minimized.

Practically speaking, the mutual impedance between the driven and parasitic elements cannot be neglected; in fact, the mutual impedance is the mechanism by which the parasitic element is excited, and proper phasing of the elements is obtained by adjusting the length of the parasite and the spacing between the two elements.¹⁴ Usually the reflector is made a little longer than the driven element.

Broadband requirements further complicate the problem because frequency changes affect not only the lengths of the elements, but the spacing between them as well. Simplification is again afforded by using two arrays, one for the low- and one for the high-band. Equations (16-25) show that the conditions are more severe in the low band, where the maximum deviation from the geometric band center is just less than 30 per cent. Thus, in seeking a compromise design for the parasitic reflector we seek a reflector length and spacing which give the most uniform results over a maximum frequency deviation of 30 per cent. It has been found that the curve of power gain of the dipole and parasitic array v. the spacing maximizes for a spacing of approximately 0.2λ with a relatively sharp decrease for

¹⁴ For curves of the effective impedance of the back element and for directivity patterns of the parasitic array see M.I.T. Radar School Staff, *Principles of Radar*. New York: McGraw-Hill Book Company, Inc., 1946, chap. 9.

spacings less than 0.15λ .¹⁵ Thus, if a distance of 0.225λ is used at the geometric mean of the band, a good compromise is obtained. The corresponding reflector length is taken to be approximately 0.525λ . Allowing a decrease of roughly 5 per cent for the wavelength on the antenna as compared to its value in air, we have the following design equations:

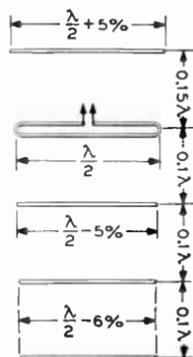
$$\begin{array}{ll} \text{driven dipole} & L = \frac{468}{f_{m(m.c.)}} \text{ feet} \\ \text{parasitic reflector} & L = \frac{492}{f_{m(m.c.)}} \text{ feet} \\ \text{spacing} & L = \frac{221}{f_{m(m.c.)}} \text{ feet} \end{array} \quad (16-36)$$

The directivity in the forward direction may also be increased by a parasitic element in front of the driven element, the requirement being that the "director's" length be less than $\lambda/2$. The curve of gain v. spacing is relatively narrow and peaks at a spacing of 0.1λ . If used at all, the director is used with a driven dipole and a reflector to form a three-element array.

In certain instances where extremely high directivity is required, the Yagi array may be used. Shown in Fig. 16-11, this array usually consists of a folded driven dipole, one reflector, and three directors. Since the performance depends upon several dimensions, the Yagi array is essentially a narrow-band device and as such is normally

Fig. 16-11. The Yagi array. One rear and three front parasitic elements are used. Top view is shown. (Courtesy of *Radio and Television News*.)

$$\begin{array}{ll} A = 0.15\lambda & B = 0.1\lambda \\ 1 = \frac{\lambda}{2} + 5\% & 2 = \frac{\lambda}{2} \\ 3 = \frac{\lambda}{2} - 5\% & 4 = \frac{\lambda}{2} - 6\% \end{array}$$



¹⁵ G. Grammer and B. Goodman, *The A.R.R.L. Antenna Book*, The American Radio Relay League, 1939, chap. 9.

used for a single channel which provides marginal reception. If the advantages of high gain and directivity are needed on several channels, one array should be provided for each channel along with some form of switching device. Since adjacent arrays interact, they should be separated by as large a distance as is practicable.

Under certain rare circumstances the economic factor may limit the number of Yagi arrays to two, one each for high- and low-band reception. Greenlee has recommended the following compromise values where this condition prevails:¹⁶

| | | | | | | | | |
|-------------|----------|----------|----|-----|----|-----|---|---------|
| <i>Band</i> | <i>A</i> | <i>B</i> | 1 | 2 | 3 | 4 | } | (16-37) |
| Low | 25½ | 17 | 84 | 80 | 76 | 75 | | |
| High | 8¼ | 6 | 30 | 28½ | 27 | 26¾ | | |

NOTE: All dimensions are in inches; notation is shown in Fig. 16-11.

It must be stressed that a Yagi array cannot provide anything like uniform performance over the entire low or high band and should only be used for bandwise operation in the absence of some other adaptable method.

BROADBANDING THE DIPOLE

The need for a broadband antenna element to provide uniform electrical characteristics over a wide frequency band has been pointed out in previous sections. It was also stated that some degree of broadbanding is afforded by the folded dipole. We next consider what means are available for broadbanding the simple dipole.

16-7. Fans and Vertical V's

Mention has been made of the fact that as the dipole diameter is increased, its resistance increases faster than its reactance, thereby producing the required effect by lowering the antenna *Q*. Practically speaking, the effective radiator size may be increased by employing several thin dipoles, all joined together at the feed points. Two antennas of this type, the vertical double V¹⁷ or conical array and the fan, are illustrated at *a* and *b* in Fig. 16-12. A third type, the vertical V, is broadbanded by bending its conductors upward in

¹⁶ L. E. Greenlee, "High-Gain Directional Array for Marginal TV Reception." *Radio and Television News*, 42, 2 (August 1949).

¹⁷ The term "vertical double V" is not in common use but is used here to distinguish the antenna from the conventional V, which is discussed later.

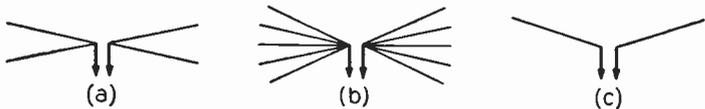


Fig. 16-12. Broadband dipoles. (a) The vertical, double-V, or conical unit, side view. (b) The fan, top view. (c) The vertical V, side view.

the vertical plane. In all three cases, the modified dipole structure retains the directivity characteristics of the basic dipole.

16-8. Spikes

Another broadbanding technique utilizes the principle of altering the current distribution in the receiving element and in effect allows the dipole to change its length in a manner which offsets the change resulting from variation in frequency. The end result is that the figure-eight radiation pattern is retained over both the high and low bands.¹⁸ The basic form of the modified dipole is shown at *b* in Fig. 16-13, where a lumped inductance is inserted in each arm of

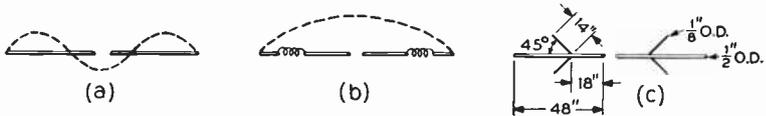


Fig. 16-13. The R.C.A. spiked dipole. (a) Current distribution in a dipole, $l = 3\lambda/2$. (b) Series inductance decreases the effective length, $l' = \lambda/2$. (c) The lumped inductance is replaced by V's or "spikes."

the dipole in order to modify the current distribution. The effect of the inductance is to decrease the distance between adjacent current minima;¹⁹ thus, if the basic dipole has an electrical length $l = 3\lambda/2$ with the distribution shown at *a*, the inductance may be used to give the distribution of *b* and the effective electrical length becomes $l' = \lambda/2$. It may be seen, then, that the added inductance makes the effective length l' less than the actual length l . As the frequency is raised, the added reactance rises and makes the shortening effect more pronounced.

¹⁸ O. M. Woodward, Jr., Reversible-Beam Antenna for Twelve-Channel Television Reception, *RCA Review*, X, 2 (June 1949).

¹⁹ F. E. Terman, "Radio Engineers' Handbook." New York: McGraw-Hill Book Company, Inc., 1943.

We have already observed that l of the basic dipole increases with rising frequency. Thus, with the proper choice of inductance, the decrease in l' may be made to compensate the increase in l and the net length of the dipole is maintained within relatively small limits. In practical form, as used by R.C.A., the dipole uses spikes or small V's in place of the lumped inductance, as shown at *c* in the diagram.

It has been determined that the spikes have negligible effect in the low band covering channels No. 2 through No. 6; hence the physical length is chosen to favor the low band and the spikes act to maintain the effective length at $\lambda/2$ in the high band as well. The spiked antenna may therefore be used to cover all 12 channels and may also be used with a similar spiked reflector for back suppression of the directivity pattern. Dimensions of the basic unit are shown in Fig. 16-13c.

16-9. Horizontal V

It has been mentioned that a broadbanding effect without change in the directivity pattern may be obtained by bending the sections of the dipole upward to form a vertical V as in Figure 16-12c. In a similar manner, if the sections are bent in the forward direction to form a V in the horizontal plane, broadbanding and change of directivity are obtained. It is known that the gain in the forward direction of a V is a function of the element lengths and the angle included between them. Since the electrical length of the element varies over the 12 television channels, a compromise design is required. The included angle is chosen to be 90° and the element lengths $\lambda/2$ (or if space does not permit, $\lambda/4$) at the lowest channel to be received. With these values, the directivity pattern is nearly uniform in all directions at the lower frequencies and gradually forms into a narrow figure eight, with some minor lobes, as the frequency is raised.

16-10. Vertical Directivity

The principle of stacking antenna elements in the vertical plane to increase directivity in that plane, which was discussed in Chapter 13, may also be used to an advantage in receiving antennas. This is particularly true when the receiving antenna is located in a region of high local R-F noise due to automobile ignition systems,

faulty neon signs, and other spark-producing devices. Consider the arrays shown in broadside view in Fig. 16-14. Simple dipoles are shown as the elements, but they may be replaced by V's, double

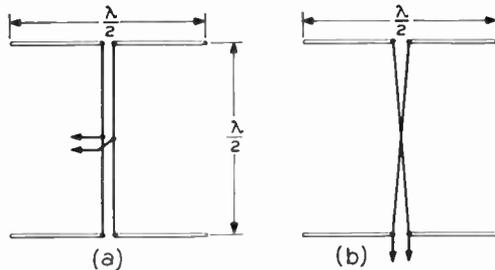


Fig. 16-14. Methods of connecting two in-phase elements stacked vertically to increase vertical directivity. (a) Center feed. (b) End feed.

vertical V's, or folded dipoles with or without reflectors. For a television signal which arrives broadside to the array, equal in-phase signals are developed in the upper and lower dipoles that add at the lead-in. For local noise generated directly below the array the voltage induced in the upper dipole will lag that in the lower dipole by 180° because of the difference in distance between the noise source and the two dipoles, and the noise signal will cancel at the lead-in. For noise sources in the immediate vicinity but not directly below the array, the phase difference between the two induced voltages will be less than 180° but more than 90° , so that the noise will cancel but not completely at the lead-in. The actual vertical directivity pattern is that shown in Fig. 13-33.

A number of receiving antennas which are available on the market are illustrated in Fig. 16-15. The broadband antenna illustrated

Fig. 16-15. Typical commercial television receiving antennas. (a) A broadband TV and FM antenna. (Courtesy of Frederick A. Kolster.) (b) A hi-lo band array suitable for use when all the stations to be received lie in the same general direction from the receiver. (c) A hi-lo band antenna, each band utilizing a folded dipole with a dipole reflector. Separate lead-ins from each unit, or a single lead-in with a crossover network similar to that shown in Fig. 16-8, may be used. (d) Two conical antennas stacked vertically. The reflectors are also conical units. (e) A hi-lo band antenna whose high-band unit is a Yagi array. (b)-(e) courtesy of Oak Ridge Antennas. (f) The Di-loop antenna. The loop is cut for channel No. 7 and the dipole for channel No. 2. Maximum pickup in the loop is normal to the plane of the loop. (Courtesy of Square Root Mfg. Corp.)

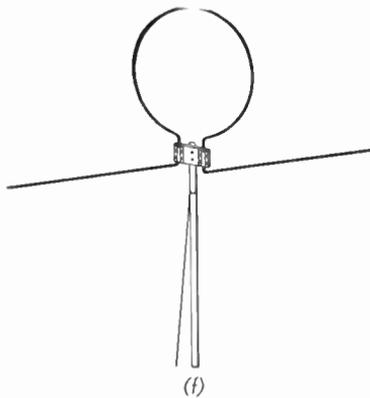
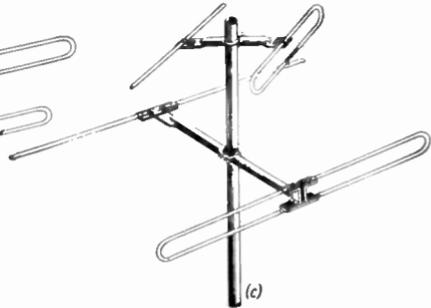
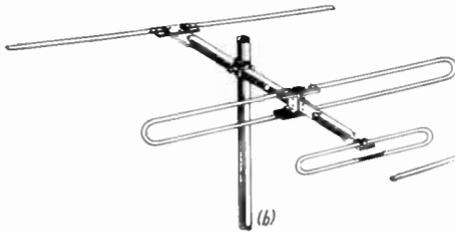
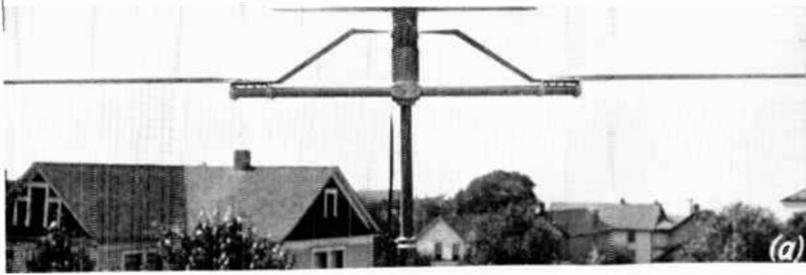


Fig. 16-15. (See caption on opposite page.)

at a is designed specifically for television and F-M reception and has been described in the literature.²⁰ The Di-loop shown at f is the basic element for the rotatable antenna described in the next section. The loop portion of the antenna is cut for channel No. 7 and the dipole is cut for channel No. 2. This particular combination gives excellent results. When compared to a standard-folded dipole cut for each channel, the performance of the Di-loop is as follows:²¹

| <i>Channel</i> | <i>Relative response, decibels</i> |
|----------------|------------------------------------|
| 2 | -1.2 |
| 4-6 | 0 |
| 7-13 | +1 |

One point in regard to the loop is of special interest: its diameter is in the order of one wavelength and maximum pickup of signals normal to the plane of the loop is obtained. This is in contrast to the more familiar loops used in direction finding. In the latter case the loop dimensions are small compared to a wavelength, and maximum pickup occurs in the plane of the loop.

16-11. Rotatable Antennas

When the receiving location is such that the several stations to be received lie in several directions, three general receiving installation alternatives are open.

(1) An omnidirectional may be used. This is generally avoided because of noise pickup problems.

(2) Individual antennas may be provided for each channel to be received.

(3) Rotatable antennas may be used.

In its most usual form the third alternative employs an antenna which may be physically rotated by a remotely controlled electric motor. A more recent form in the television field utilizes a well-known principle used in radio direction finding. Two receiving elements, Di-loops in the commercial model, are mounted at right angles to each other in a fixed position. Two separate lead-ins, one from each element, feed a broadband goniometer,²² whose output is

²⁰ Frederick A. Kolster, *Antenna Design for Television and FM Reception*, *Proc. IRE*, **36**, 10 (October 1948).

²¹ For further details on the directional properties of the antenna, see "Quadrature-Phased TV Receiving Antenna." *Tele-Tech*, **8**, 11 (November 1949).

²² See, for example, F. E. Terman, *Handbook, op. cit.*, section 12. Also footnote 21

connected to the receiver terminals. Rotation of the goniometer adjusts the phase relationship between the outputs of the two antenna elements; in effect, the antenna pattern is rotated electrically, even though the antenna position remains fixed.

Since the goniometer must be sufficiently broadband to cover all 12 television channels, a special design, differing from that of the conventional goniometer, is incorporated in the so-called Azimutrol unit. Shown in Fig. 16-16, it consists of four stator plates which are connected to two crossed Di-loops. Two resistance-coated rotor plates are connected to the input terminals of the television receiver. Rotation of the rotor allows the phase of the signals which are

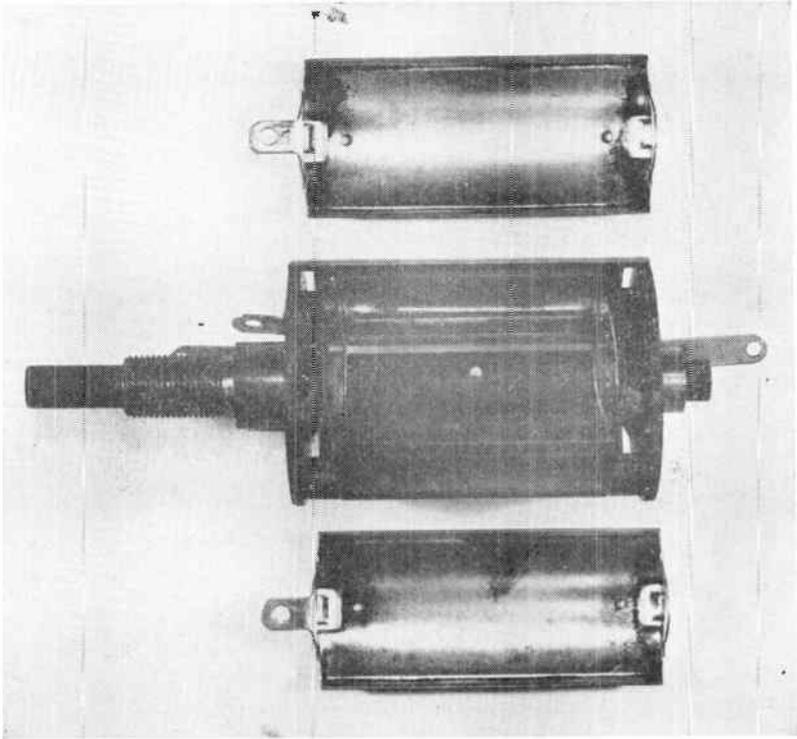


Fig. 16-16. The Azimutrol goniometer unit. Two of the stator plates are removed to show the rotor. A circuit diagram of the unit is shown in Fig. 16-17. (Courtesy of Square Root Manufacturing Corporation.)

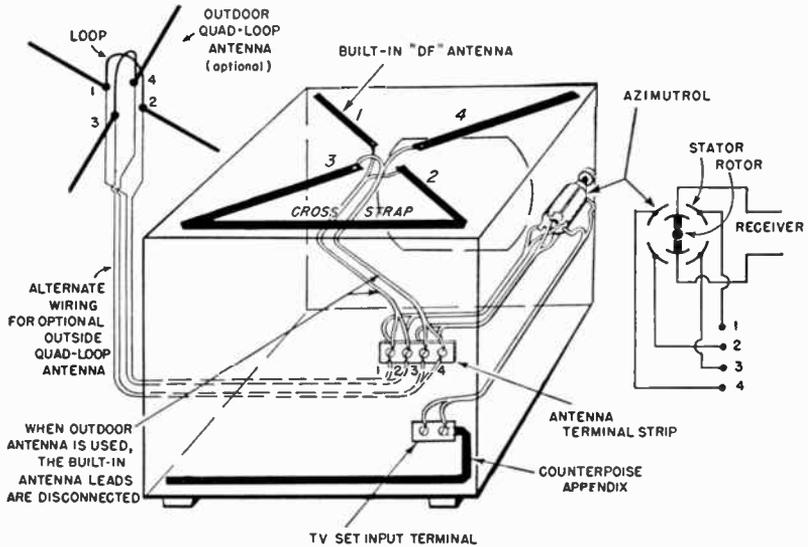


Fig. 16-17. A typical built-in antenna employing the Azimutrol. (Courtesy of Square Root Manufacturing Corporation.)

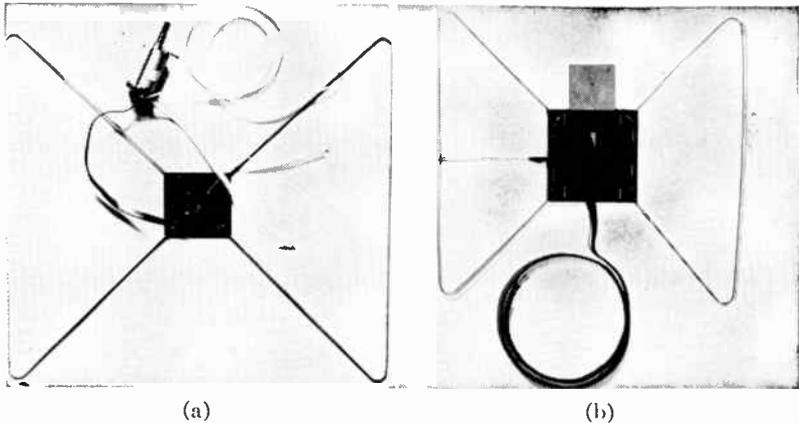


Fig. 16-18. Two versions of a built-in antenna. (a) A two-element unit using the Azimutrol for lobe rotation. The short sections of transmission line attached to the Azimutrol are matching stubs. (b) A simplified version of (a). Motion of the slide permits switching the lobe to either of two positions. (Courtesy of Square Root Mfg. Corp.)

delivered to the receiver from the crossed antennas to be varied so that electrical rotation of the antenna lobes is obtained. A circuit diagram of the Azimutrol is shown in Fig. 16-17.

Since the goniometer unit just described allows the antenna lobe position to be adjusted for optimum signal regardless of the direction to the television transmitter, it may be incorporated into an antenna system which is built into the cabinet of the receiver, an obvious advantage where it is not practicable or economically feasible to use an outdoor antenna. One form of such a built-in antenna is illustrated in Fig. 16-17. In practice the basic antenna elements have taken on a number of different forms, but those shown in the diagram illustrate the principle. Two crossed units are provided, each having a figure-eight pattern. With both elements fed to the Azimutrol, rotation of the lobes of the resultant pattern is obtained. Two forms of the built-in antenna are shown in Fig. 16-18.²³

The principal advantage of the rotary-antenna systems is that the receiving antenna may be installed in any convenient manner with no thought to orientation; with the receiver in operation the antenna may be oriented electrically to give the optimum results. As to disadvantages, the rotary systems are invariably more expensive than fixed antennas, and they require one more control to be adjusted by the television set owner.

MASTER ANTENNA SYSTEMS

It is an unfortunate fact that the majority of outdoor television antennas in use are large and, to some people, grotesque structures which seem to be a determining factor in the architecture of the American home. In particular, the roof of a multiple dwelling unit housing several television receivers appears like a forest of aluminum tubing and lead-in wires, and it is not only the appearance which is poor, but the performance of the several receiving antennas as well. It is inevitable that a large number of antennas located close together will interact, and where local oscillator radiation is present, it can amount to an intolerable situation as far as good television reception is concerned. In order to reduce the cross-interference in multiple

²³ Other forms of built-in antennas which do not employ the lobe-rotating feature have been used. One such unit, which may be adjusted channelwise for an optimum impedance match, is described in "Phileo's New Built-In TV Aerial." *Tele-Tech*, 8, 10 (October 1949).

dwelling units, a number of master antenna systems, designed to serve all receivers in the unit, have been developed. These have been built on two general principles: either video or R.F. is fed over the common distribution system. A few typical master systems are discussed in the next two sections.

16-12. Video Distribution

The first master antenna system to be described effectively employs one or more master television receivers, whose amplified composite video outputs are fed to a cable system connected to several sweep and video slave units. Such a system is quite satisfactory when the choice of program is under central control, such as in a hospital. The block diagram of such a common-video system is illustrated in Fig. 16-19. A single antenna and a master receiver are required

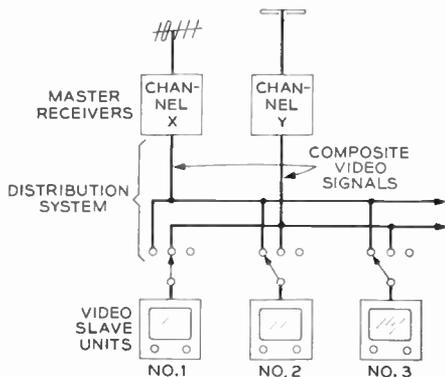


Fig. 16-19. Block diagram of a common-video type of master antenna system.

for each channel to be received, the antenna being cut and oriented to give optimum performance on that channel. The composite video output from each receiver is then fed to a cable distribution system. Program selection at each viewing position is by means of a switch, which connects the viewing unit to the desired video cable. Since no R.F. is involved at the viewing unit, it is really a slave to the master receivers; it merely accepts the composite video signal, separates picture and sync, develops its own sweep, and displays the picture. The system of Fig. 16-19 implies that the master receivers

are of the intercarrier type, so that the sound and composite video may use a common distribution cable.

Naturally any number of variations are possible; for example, the detected audio may be fed over a second distribution system, or some form of carrier system may be utilized so that all the receiver outputs may feed into a single, common distribution system. In any event, the common video system has one outstanding disadvantage: the individual viewing units cannot be commercial television receivers; special slave units are required. Such a limitation rules the system out for use in apartment houses, because any set owner wants to buy a commercial receiver of his own choice whose use is not limited to one particular type of distribution system.

16-13. R-F Distribution²⁴

The second major type of master antenna system feeds R.F., as opposed to composite video, to the distribution system. The entire system is designed so that any conventional television receiver connected to the system views the system as an ideal antenna delivering signals for each channel in the area. The block diagram of a typical system of this type is illustrated in Fig. 16-20. Again a single sep-

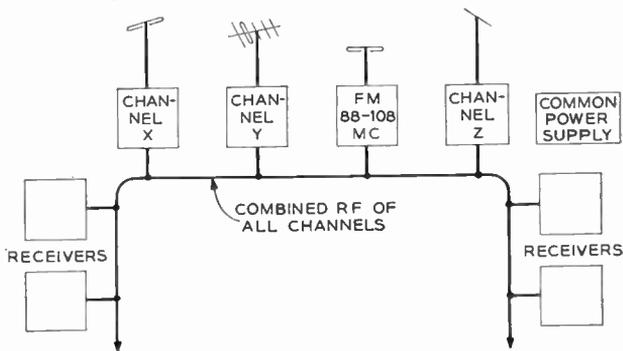


Fig. 16-20. Block diagram of a typical common-R.F. master antenna system.

arate antenna is cut and oriented for each channel to be received. Connection from each antenna to the distribution cable is through a five-tube R-F amplifier whose response is flat within 1 decibel over the 6-megacycle channel width, except for the F-M amplifier, which

²⁴ H. E. Kallmann, "Television Antenna and RF Distribution Systems for Apartment Houses." *Proc. IRE*, 36, 9 (September 1948).

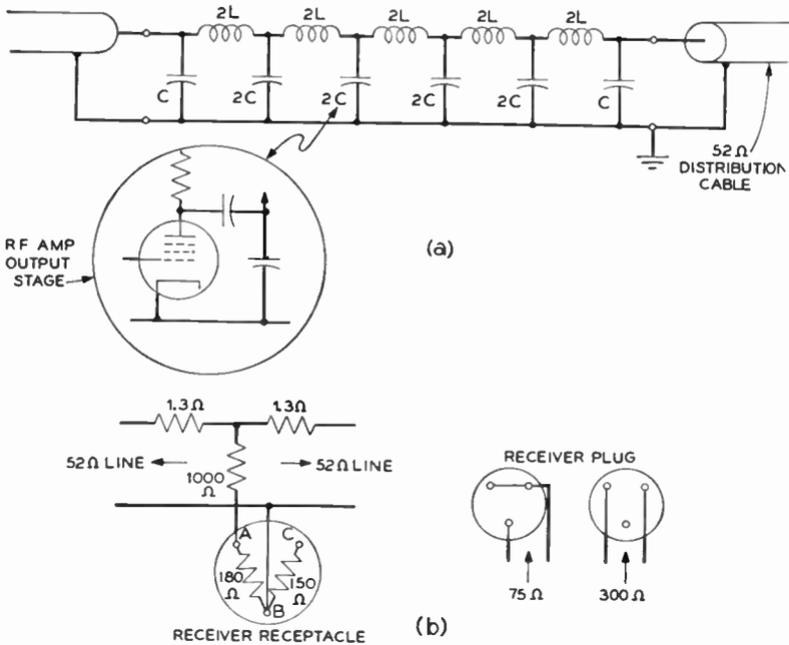


Fig. 16-21. Details of a common-R.F. distribution system. (a) The output capacitance of the R-F amplifier furnishes a portion of the shunt capacitance of the artificial line. (b) The receiver outlet isolation and matching system. (Courtesy of *Proc. IRE*.)

is essentially flat from 88 to 108 megacycles. The gain of each amplifier is under adjustment so that the signal strengths of all channels may be made equal in the distribution cable. With all channels equalized in this manner the need for contrast adjustment at the receiver is minimized as the receiver is tuned from one channel to another.

Considerable ingenuity is used in feeding the R-F amplifiers to the distribution cable without upsetting the impedance match. The principle used is illustrated at *a* in Fig. 16-21. The center of the 52-ohm cable is fed by an artificial line whose nominal impedance is also 52 ohms. Then the output capacitance of each amplifier is shunted across the line condenser so that the two in parallel furnish the design value of $2C$. By this means the R-F voltage is delivered to the line without upsetting the impedance conditions.

At each receiver receptacle three conditions must be met. First,

the distribution line termination must be correct; second, the receiver input impedance must be matched; and third, any local oscillator signal radiation from the receiver input terminals should be attenuated so that it will not interfere with other sets on the line. The receptacle connections shown at *b* in the figure meet these requirements. With the components shown a 30-decibel attenuation occurs between the receiver input and the distribution cable; thus a local oscillator signal suffers 60-decibel attenuation in going from one receiver to another through the cable. As far as the desired signal is concerned, the 30-decibel loss is more than compensated by the master R-F amplifiers which deliver a 30-millivolt signal to the distribution system.

The initial cost of the system just described is relatively high, primarily because of cable installation. Kallmann has estimated that the cost to 20 users would equal the cost of the same number of individual receiving antenna installations. A number of antenna and distribution systems of this general type are available commercially. Other systems of the passive type, *i.e.*, that use no amplifiers, have been developed for installations serving a maximum of twelve receivers.²⁵

²⁵ I. Kamen, "Television Master Antennas." *Radio and Television News*, 41, 4 (April 1949).

CHAPTER 17

TELEVISION MOTION PICTURES

The high programing costs of live-talent shows for the television medium make the frequent use of motion pictures as subject material almost mandatory. This is particularly true in those stations which provide programs throughout the day and rely upon films for spot announcements, news features, and general entertainment. In the present chapter we consider how the motion picture is adapted to the television medium.

17-1. The Basic Problem

Basically the problem in televising movies is that the frame and field frequencies are different in the two media and must be made compatible in some manner. We have already seen in the earlier chapters that the film uses a 24/48 set of standards, that is, the film is run at a speed such that 24 frames or individual still pictures appear in the film gate each second but each frame is displayed twice so that the flicker rate is 48 pictures per second. The television system, on the other hand, is based on a 30/60 standard by which 30 frames are scanned with a 2 to 1 interlace each second. Since the television rates are standard, the chief factor to be considered is how the two systems may be made compatible, and how 24 film frames may be "stretched" so that they appear 30 times each second.

The stretching cannot be accomplished by running the film at 30/24 (= 1.25) times its normal speed because the higher speed would upset the reproduction of the sound track, which is also on the film, and would impair the proper illusion of motion in the reproduced image. For these reasons, the film itself must be run at a 24 frame per second rate and the conversion produced by some means other than film speed-up.

A second problem has to do with the pull-down time of the film. If, for example, the film is exposed to the television pickup tube

during the active portion of the vertical scan interval, the film must be moved one frame in the vertical blanking interval of 1.25 milliseconds, a severe requirement indeed.

A third problem concerns the shutter mechanism, which must expose the film to the pickup system for short, precisely timed intervals with as much brightness as possible. These problems and their solutions are considered in the three following sections.

17-2. The Film Scanning Interval

In order to investigate the pull-down requirements in the cinema projector we must first consider the running speed of standard motion-picture film. Since sound film is generally used for television work, two sets of data are required: one for 35-millimeter film and one for 16-millimeter sound film. The pertinent data are listed in Table 17-1.

TABLE 17-1
DATA FOR 35- AND 16-MILLIMETER SOUND FILM

| | 35 mm | 16 mm |
|---------------------------|-------|-------|
| Average speed, ft per min | 90 | 36.3 |
| Frames per ft | 16 | 39.7 |
| f_p , frames per sec | 24 | 24 |
| h , frame height in. | 0.75 | 0.32 |
| A , aspect ratio | 4/3 | 4/3 |

The careful student will notice an apparent discrepancy in the figures presented in the table in that the frame height is not equal to the film width divided by the picture-aspect ratio. The reason for this discrepancy is, of course, that the entire film width is not used for the picture; space must be reserved for the sprocket holes, which are used for driving the film, and for the associated sound track.

If, now, a system is adopted which requires a frame change during the television vertical blanking interval, the pull-down of one frame height is required in 1.25 milliseconds. During this short period the film must start from rest, move downward one frame height, and come to rest again. It has been found experimentally that the film breaks under these conditions and some manner of easing the requirements, of lengthening the pull-down time, must be found if the operation is to be satisfactory.

The solution to the difficulty lies in the use of a storage tube such

as the iconoscope for the television pickup. Since a tube of this type stores information, the roles of scan and blanking intervals may be interchanged in so far as picture projection and picture pull-down are concerned. Thus with the storage tube the film is projected onto the mosaic during the vertical blanking interval and pull-down may occur during the relatively long (15.42 milliseconds) interval when the mosaic of the iconoscope is being scanned. It may be seen, then, that the storage camera tube is used to ease the requirements on film pull-down. In the usual film projector modified for television use a pull-down time of 14.9 milliseconds is used.

17-3. 24 Frames to 60 Fields

Since the film must run at its proper speed of 24 frames per second, it is apparent that the required conversion must be obtained by projecting each film frame onto the mosaic more than once. If this be done so that each film frame remains in the projection gate for an equal length of time, each frame must be projected onto the mosaic $60/24 (=2.5)$ times, clearly an impossible condition. A compromise may be reached, however, by exposing alternate frames an unequal number of times. For example, if the sequence 2,3,2,3, . . . is used, the requirements will be met. Half of the frames on each second of film will be exposed twice, and the other half three times; then

$$\frac{24}{2} (2) + \frac{24}{2} (3) = 24 + 36 = 60 \quad (17-1)$$

and the required conversion is obtained; with the film running at an average speed of 24 frames per second, the mosaic may be illuminated 60 times a second to meet the requirements of the interlaced scan.

Since the exposure time of the mosaic is extremely short, it is desirable that a light source of high intensity be used. Until 1948 a carbon arc was used for this purpose and the exposure time of the mosaic was controlled by a rotating shutter, which alternately opened and closed the optical path to the camera tube. On the basis of what has been said thus far, then, the film projector itself must be designed to provide the unbalanced film-in-gate periods and the proper open-shutter intervals. The required timing sequences to accomplish this in a 35-millimeter projector, which can have a 14.9 millisecond pull-down time, are shown in Fig. 17-1.

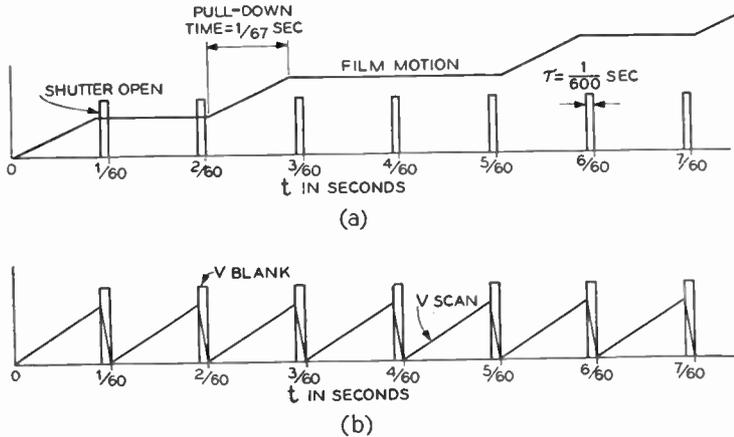


Fig. 17-1. Scanning sequence for the 35-millimeter film projector. (a) Film motion and shutter operation. (b) Vertical scanning and vertical blanking.

17-4. The 35-millimeter Projector

In order to provide the 2,3,2,3, . . . scanning sequence a modified intermittent movement of the Powers type may be used.¹ A schematic of the modified unit is shown in Fig. 17-2. The constant-speed drive is provided by a 3600 revolutions per minute synchronous motor, which is fed from the same power line to which the television synchronizing signal generator is locked. By this means the rotating shutter is locked in with the scanning sequence in the television camera tube. Since a phase ambiguity is present in the conventional synchronous motor, the motor used here is provided with a special d-c phasing winding, which provides positive lock-in on the proper phase.

The rotating shutter, which has a single exposure notch, is connected directly to the motor shaft and so provides one exposure per 1/60 second or per vertical field. The intermittent itself is driven through a 5 to 1 step-down gear train so that it rotates at 12 revolutions per second. Since two "throw holes," which initiate the pull-down motion of the drive sprocket, are provided, the number of pull-downs per second is $12 \times 2 = 24$, as required.

¹ E. W. Engstrom, G. L. Beers, and A. V. Bedford, "Application of Motion-Picture Film to Television." *J. Soc. Mot. Pic. Eng.*, **33**, 3 (July 1939).

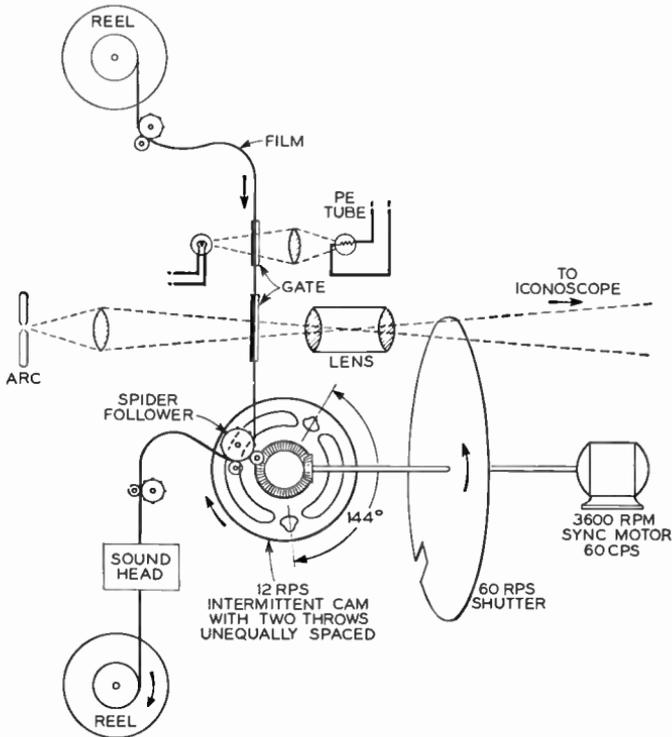


Fig. 17-2. Schematic diagram of a 35-millimeter film projector used with an iconoscope. Notice the unequally spaced throw holes in the intermittent cam. The phototube to the right of the upper film gate is used for setting the average value or background level of the video signal. (Courtesy of *J. Soc. Mot. Pic. Eng.*)

The 2,3,2,3, . . . exposure sequence is provided by allowing successive film frames to remain in the gate unequal time intervals. This action is made possible by unequal spacing of the two throw holes as shown in the diagram. The angular displacements between the two throws are 144° and 216° ; thus the relative exposure time of successive film frames is

$$\text{relative exposure time} = \frac{144}{216} = \frac{2}{3} \quad (17-2)$$

and again the requirement is met.

In more recent types of 35-millimeter projectors the rotating shut-

ter is placed between the light source and the film gate instead of as shown in the diagram in order to reduce the fire hazard always present when a carbon arc is used for film illumination.

17-5. The 16-millimeter Projector

With a large amount of program material now available on 16-millimeter sound film, which is much less expensive than the larger 35-millimeter size and which has a resolution comparable with standard television, we must consider these several problems for the smaller size film. It should be clear from the figures of Table 17-1, which show that the frame height in 16-millimeter sound film is less than half of that in the 35-millimeter size, that a much shorter pull-down time is possible, and the whole problem becomes much simpler. In fact, a standard 16-millimeter projector is available commercially which utilizes approximately one-tenth of the frame interval for pull-down. With this type of projector the standard intermittent movement may be used without modification, and the timing sequence used is that shown in Fig. 17-3.² It may be seen at once that each frame remains in the film gate for the same length of time, which indicates that an unmodified intermittent is used. The

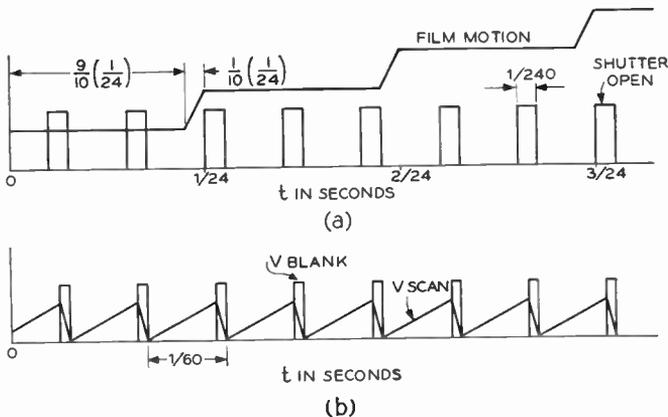


Fig. 17-3. Scanning sequence for the 16-millimeter film projector. (a) Film motion and shutter operation. (b) Vertical scanning and vertical blanking.

²The author is indebted to Mr. Bernard Erde of the Columbia Broadcasting System for the data on the 16-millimeter scanning sequence.

2,3,2,3, . . . sequence is provided by the rotating shutter alone, a method which is possible solely because of the short pull-down time. Thus the proper sequence is provided by the setup of Fig. 17-2, except that the two throw holes are spaced 180° apart. Inspection of Fig. 17-3 shows that with the 16-millimeter system described a relatively long exposure time for the mosaic may be provided; that is, exposure occurs during the vertical retrace interval plus a portion of the scanning interval. The increase of roughly 2.5 to 1 in exposure time plus the smaller film size as compared to the 35-millimeter case permits the use of a lower light intensity, and quite satisfactory results have been obtained with a 1-kilowatt projection bulb in place of the arc.

17-6. The Shutterless 16-millimeter Projector

The recent development of pulsed gas-filled light sources has resulted in the design of a television film projector which utilizes no rotating shutter, the shutter action being replaced by electrical control of the light source. The timing sequence used in a projector

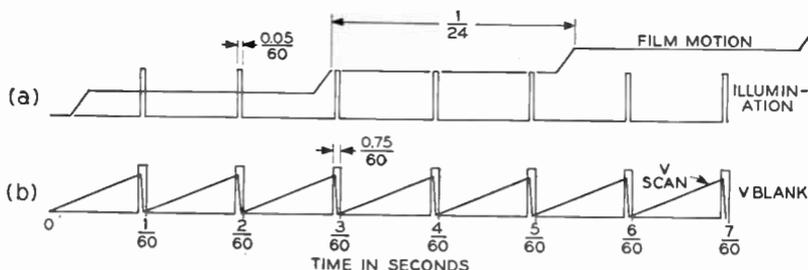


Fig. 17-4. Timing sequence for the shutterless film projector. (a) Film motion and film illumination. (b) Vertical scanning and vertical blanking.

of this type, which is produced by General Electric, is shown in Fig. 17-4.³ It will be observed that a very short mosaic exposure period is used (0.8 millisecond). This is possible because of the high light output of the type FT-230 krypton-filled lamp.

³ L. C. Downes and J. F. Wiggin, "Shutterless Television Film Projector." *Electronics*, 22, 1 (January 1949).

The circuit used for pulsing the arc tube is shown in Fig. 17-5. The tube consists of two tungsten electrodes in a krypton atmosphere and requires a peak-to-peak voltage of approximately 15 kilovolts to break down. This voltage is furnished by the H-V pulser, which is keyed by a blocking oscillator synchronized by the vertical blanking pulse from the television system. The form of the pulsing voltage

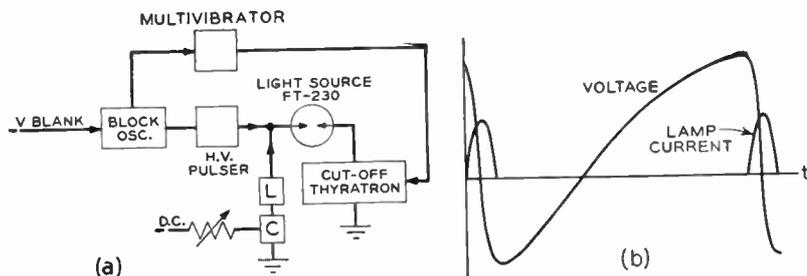


Fig. 17-5. Lamp-pulsing circuit for the shutterless film projector.
 (a) Pulsing circuit. (b) Approximate wave forms, not to scale.
 (Courtesy of *Electronics*.)

is shown at *b* in the diagram, and it is developed across the resonant circuit *L* and *C*. Once the arc has been established, a lower d-c voltage may be used to sustain it. This sustaining voltage is furnished by a three-phase selenium rectifier circuit. The d-c sustaining current may be varied between the limits of 1.5 and 2 amperes by means of the series regulating resistor *R* in order to provide control of the output. The return path for the FT-230 is through a thyatron, which cuts off when the oscillatory voltage produced by the *L-C* circuit swings negative. Thus the thyatron serves as a cutoff switch for the illuminating cycle. Proper design of the resonant circuit limits the illumination interval to approximately 5 per cent of the field interval. Notice that proper phasing of the exposure with respect to the television scanning cycle is ensured by using the vertical blanking pulse to trigger the lighting circuit.

It should be observed that proper operation of the equipment also requires that the film motion be synchronized with the television scan in order that pull-down occur in accordance with Fig. 17-4. Instead of relying on a synchronous motor with a phasing winding to perform this function, use is made of the circuit shown in Fig. 17-6. The heart of this system is the phase comparator, which we have

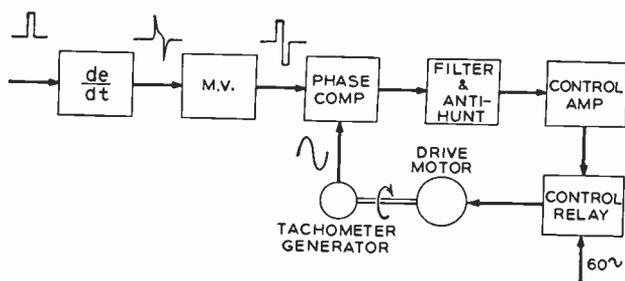


Fig. 17-6. Film motion phasing circuit for the shutterless projector. An error voltage derived from a comparison of a square wave and the tachometer output controls the drive motor through a feedback loop. (Courtesy of *Electronics*.)

previously described and which delivers a d-c error voltage whose magnitude and polarity depend upon the phase difference between the two voltages being compared. In the present case the compared voltages are a square wave derived from the television vertical blanking pulse and the sinusoidal output of a tachometer, which is driven by a film projector motor of the universal type. The former voltage, which serves as the standard, is a 50-50 square wave produced by a multivibrator that is synchronized by the differentiated vertical blanking voltage.

Under normal operation, that is, when the film drive speed and phase are correct, zero output is delivered by the comparator and the speed-control circuit is passive. Any drift in the drive motor produces an error voltage which, in turn, operates through the control circuit to adjust the motor current until the drift is no longer present and the film is properly phased.

Three broad advantages are claimed for the shutterless television film projector. First, the elimination of the rotating shutter reduces the driving power required, and the intermittent lamp draws an average of 400 watts; thus the entire unit requires less power than the light source alone in the shutter type of projector. Second, the temperature of the film gate is greatly reduced because of the low average power of the light source; hence the fire hazard is low. Third, the problem of synchronizing film-motion and illumination is simplified by the electronic control circuits, and the present unit is fast enough to operate under more rigid blanking requirements which may be standardized in the future.

Part III

COLOR TELEVISION SYSTEMS

CHAPTER 18

COLOR TELEVISION

Our study of television systems to this point has been confined to the so-called monochrome or black-and-white systems, in which the image is reproduced in shades of gray. Such systems may be said to be color-blind, a fact that is illustrated in Fig. 6-5, which shows the monochrome equivalents of several typical colors. The situation here is similar to that in black-and-white photography, where each color is reproduced by a gray tone. If, then, we desire to have a color television system capable of producing a final image in color, we must devise some manner of giving color perception to a color-blind system. Basically the color television system will consist of a black-and-white system on which are superimposed additional color-perceiving and color-reproducing components. In the present chapter we shall consider some of the schemes which have proved more or less successful in providing a reproduced television image in full color.

Our study will consist of three main divisions: Elements of color, basic means of analyzing and synthesizing color in a television system, and details of several practical color television systems.

ELEMENTS OF COLOR¹

In order to find means of superimposing color perception on the color-blind television system we must first consider some of the basic concepts of color and color reproduction. The starting point is light itself, which is defined as that portion of the electromagnetic spectrum to which the eye responds. White light comprises the whole visible spectrum as shown by the use of a prism or a diffraction grating; thus it may be demonstrated that the visible spectrum consists of a continual distribution of colors ranging from red, through

¹ For a general reference on color see W. Peddie, *Colour Vision*. Edward Arnold and Co., 1922; and Parry Moon, *The Scientific Basis of Illuminating Engineering*. New York: McGraw-Hill Book Company, Inc., 1936, chap. 13.

yellow, green, and blue, to violet, and each of these spectral colors may be identified with a certain wavelength or frequency of electromagnetic radiation, red lying at the low, and violet lying at the high end of the frequency spectrum. The color of any object is determined by the mixture of all the frequency components of the light reflected or transmitted by the object.

18-1. Response of the Eye

The response of the eye to color is largely a psychological matter, and a number of theories have been advanced to explain the phenomenon of color perception. One such theory, proposed by Young and Helmholtz, presumes that the retina, the light-sensitive surface in the eye, consists of three separate sets of nerve fibers, each of which has its own particular frequency response characteristic. The response curves for these fibers, which were investigated by Koenig, are illustrated in Fig. 18-1.² On the basis of the Young-Helmholtz

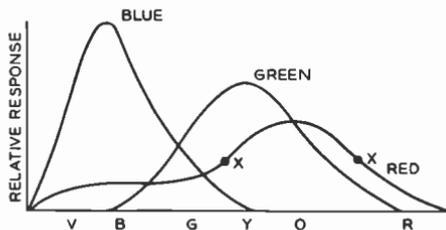


Fig. 18-1. Color response curves of the eye.

theory it is supposed that any color, say a green, is received by the brain as three separate stimuli, one from each set of color-sensitive fibers. Since each frequency in the spectrum corresponds to a unique set of responses, the brain is able to interpret any given set as a particular color.

It is interesting to note the meaning of monochrome response in this connection. Let it be supposed that the eye or any other device, such as an iconoscope for that matter, has only a single set of color-sensitive elements, say the red. Then in the absence of the blue and green elements, any color would be interpreted by that eye as a shade of red alone. Under this condition an ambiguity exists

² See, for example, H. M. Reese, *Light*. Columbia, Missouri: Lucas Brothers, 1927.

between any two colors which have the same relative response on the curve. For example, in Fig. 18-1 the two points marked x have the same relative response so that the yellow-green and red-orange indicated would have the same monochrome values.

18-2. Color Matching

It would seem, then, from the color response curves of the eye that any color may be synthesized as far as the normal eye is concerned by the proper mixture of the three primary colors red, blue, and green. This fact has been demonstrated very well in color photography and in the several color-printing processes. We shall see in the next few paragraphs, however, that there is nothing unique about the three primaries red, blue, and green, and that any color may be synthesized from any three primaries which differ from each other and no two of which can synthesize the third.

Before demonstrating this last statement it is important that we make careful distinction between color stimuli and color sensations. On the one hand we have the physical attributes of the radiant energy reaching the eye, the intensity of the radiation, and its wavelength (or the intensity and wavelength of each component in the radiation). These are properties of the stimulus and, as will be shown, it is difficult to directly relate these attributes to the psychological response which we know as color.

On the other hand, we have the psychological attributes of color, which may be conveniently described in terms of three general qualities: hue, chroma (or saturation), and value (or brilliance). These qualities form the basis for the Munsell system of color notation.³ Hue is that sensation which distinguishes one color from another; for example, we recognize a reddish or a yellowish sensation. Value is that sensation which makes color perception possible. For example, an orange viewed in bright sunlight has a typically orange sensation of great brilliance, whereas the same orange viewed at night may not even be visible. Thus we might say that the value of a color is related to its position along the gray scale; the sensation of value associated with the orange may be lowered by viewing it through a neutral gray filter. The third term, chroma, describes the degree of saturation of the color or, in other words, its mixture with white; that is, we can recognize the difference between an un-

³ See *Life*, July 3, 1944, for a reproduction in color of the Munsell chart.

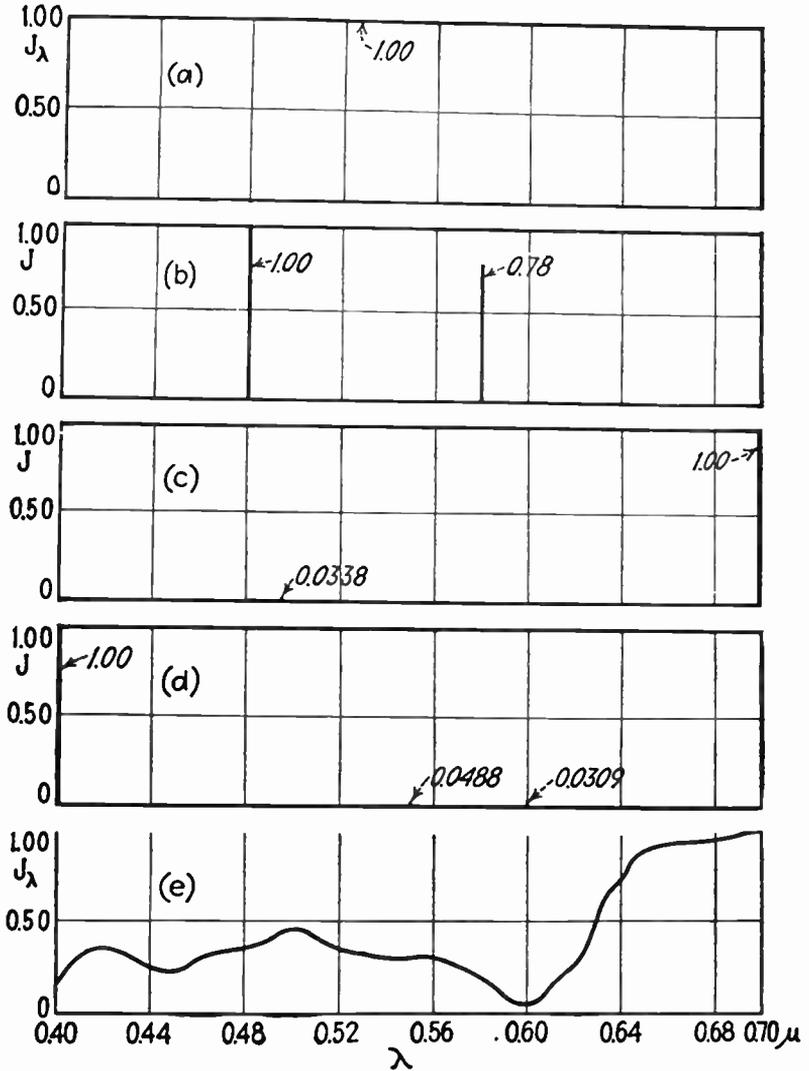


Fig. 18-2. The spectral distributions that cause a given color sensation are not unique. The three distributions of (a), (b), and (c) all appear as the same neutral gray. In a similar manner the three primaries at (d) give the same color as the continuous spectrum at (e). (By permission, from Parry Moon, *The Scientific Basis of Illuminating Engineering*, copyright 1936, McGraw-Hill Book Company, Inc.)

saturated red (or pink) and a saturated red. It is important to remember that these terms, hue, value, and chroma are general ones, which are convenient names for our visual sensations. They have no quantitative significance (except when used with a specific color chart such as the Munsell) and must not be used to describe the radiation which produces the sensation of color.

How, then, shall we describe a given color quantitatively? First of all it may be described by a curve of intensity *v.* wavelength, a spectroradiometric curve, for the radiation which produces the sensation of color. Actually a curve of this type, which may be obtained with a spectrophotometer, is of little use because it has been found experimentally that different intensity-wavelength curves can produce identical sensations of color. This is illustrated at *a*, *b*, and *c* in Fig. 18-2. The three distributions shown are markedly different; yet all three produce the same sensation of a neutral gray.

A more useful approach to the problem is based on specifying the intensities of three arbitrary but known primaries, which are able to match the unknown color, that is, to produce the same sensation as the unknown. The required intensities are determined experimentally by color matching. Figure 18-3 shows the basic equip-

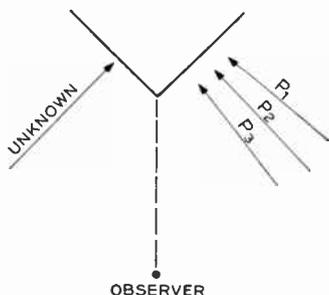


Fig. 18-3. The basic equipment required for color matching. The intensity of each of the primaries that illuminate the right half-field of the screen may be varied. Negative values of a primary are obtained by moving the primary source so that it illuminates the left half-field.

ment required, which is quite similar to the usual photometer except that three, rather than a single, known sources are used to illuminate the right-hand half of the field. It has been determined experimentally with this equipment that at the most only three controls need be adjusted to obtain a match of the colors appearing on both half-fields. Once a match is obtained, it may be said that the

stimulus $E(\lambda)$, which illuminates the left half-field is equivalent to the sum of the three primaries, P_1 , P_2 , and P_3 , of intensity $I_1(\lambda)$, $I_2(\lambda)$, $I_3(\lambda)$, respectively, or we may write

$$E(\lambda) = I_1(\lambda)P_1 + I_2(\lambda)P_2 + I_3(\lambda)P_3 \quad (18-1)$$

where the equality indicates that both sides of the equation evoke the same color sensation. A typical example of such a match is illustrated at *d* and *e* in Fig. 18-2. The three primaries of intensity shown at *d* produce the same color sensation as the continuous spectrum as shown at *e*. Thus, the particular $E(\lambda)$ may be specified in terms of the three primaries and their intensities.

Two other important facts may be determined with the color photometer. First, a color match may be obtained with *any* three primaries P_1 , P_2 , and P_3 , provided only that no two are identical and that no combination of two can match the third. Second, with some sets of primaries it may be necessary to move one of them so that it illuminates the left half-field along with the unknown in order to obtain the color match. When this is true, the primary which is moved is said to be a negative primary and a negative sign must be used with it in eq. (18-1).

Another piece of experimental datum obtained with the color photometer, which is of importance in our subsequent work, is that a match obtained at one value of intensity of the unknown is also a match over a wide range of intensities. In other words, a given match holds true over a wide intensity range.

18-3. The Color Triangle

Since, as we have seen, any color may be specified in terms of three primaries, any radiation may be represented by a point in space relative to three axes representing the three primaries used. Once again it must be stressed that any three primaries may be used, in fact, they may even be imaginary. While this may seem to be a ridiculous procedure, it nevertheless has the advantage that a suitable choice of imaginary primaries eliminates the need for negative coefficients in eq. (18-1). Furthermore, the fictitious primaries may also be chosen such that the point representing white has equal values for the coefficients $I_1(\lambda)$, $I_2(\lambda)$, and $I_3(\lambda)$. This is advantageous, for one can tell from the position of any point relative to the axes, whether

it has a red, blue, or green hue. Such a system of arbitrary primaries was adopted in 1931 and is called the C.I.E. system.⁴

In setting up the C.I.E. system three color-match curves showing the values of $I(\lambda)$ for three primaries required to match homogeneous radiation throughout the visible spectrum are obtained. Then, through a series of linear transformations these curves are transformed into a new set of curves of shape similar to those shown in Fig. 18-1.⁵ The ordinates of these transformed curves are designated \bar{x} , \bar{y} , and \bar{z} , the trichromatic coefficients. Thus any homogeneous radiation may be specified directly in terms of the trichromatic coefficients. If the radiation is not homogeneous, integration is required. It is, of course, possible to represent any radiation as a point in space relative to three axes which represent the trichromatic coefficients.

We have previously stated that a color match is essentially independent of the value of the radiation. This fact allows one of the co-ordinates required for specifying a homogeneous radiation to be eliminated so that the latter may be represented by a point on a two-dimensional plot or color map. To accomplish this in the C.I.E. system, a set of unified trichromatic coefficients are defined such that

$$\left. \begin{aligned} X &= \frac{\bar{x}}{\bar{x} + \bar{y} + \bar{z}} \\ Y &= \frac{\bar{y}}{\bar{x} + \bar{y} + \bar{z}} \\ Z &= \frac{\bar{z}}{\bar{x} + \bar{y} + \bar{z}} \end{aligned} \right\} \quad (18-2)$$

Then, since

$$X + Y + Z = 1$$

only two of the unified coefficients X , Y , Z are independent and the third, generally Z , may be omitted. Thus, any homogeneous radiation may be represented as a point in the X - Y plane. The resulting plot is known as the color triangle or the unified trichromatic coefficient diagram and is shown in Fig. 18-4.

⁴ Commission Internationale de l'Éclairage, *Compte Rendus des Seances 1931*. London: Cambridge University Press, 1932.

⁵ See P. Moon, *op. cit.*, pp. 470ff.

The outer 45° triangle shows the location of the three imaginary primaries of the C.I.E. system, red, blue, and green being located at the three points of the triangle as shown on the diagram. Our previous work has stated that any homogeneous radiation (or spectral radiation) may be plotted as a point in the diagram. If this is done

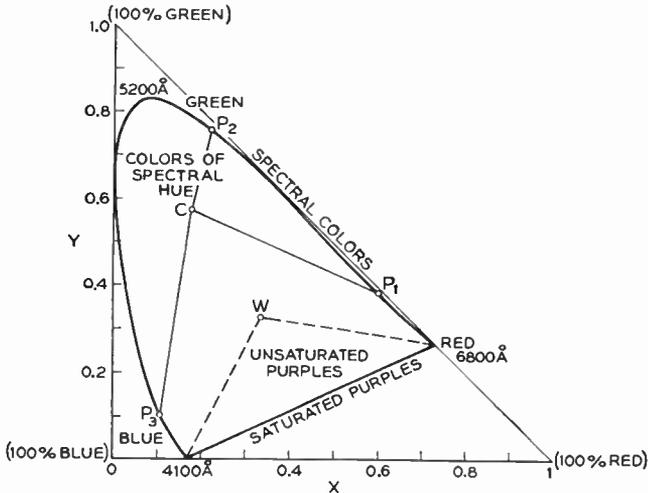


Fig. 18-4. The color triangle or unified trichromatic coefficient diagram. The imaginary C.I.E. primaries lie at the corners of the 45° triangle. The colors of the spectrum and the saturated purples lie on the inner triangle. The position of white is marked W . A color of any hue and degree of saturation may be specified in terms of the unified coefficients X and Y .

for the entire spectrum, the curved line starting at red and going up through green down to blue results. The solid line, then, represents the spectral colors. The sensations called saturated purples which do not appear in the spectrum of white light, result from mixing reds and blues and are shown by the solid line which closes the color triangle. Our results thus far show that all the spectral colors plus the saturated purples lie on the color triangle itself and, further, that the $E(\lambda)$ producing any of these sensations may be specified in terms of the unified coefficients X and Y .

Nonspectral colors which result from nonhomogeneous radiation

may also be specified in terms of the unified coefficients. This is accomplished by performing the integrations⁶

$$\left. \begin{aligned} x' &= \int_0^{\infty} E(\lambda)\bar{x}(\lambda)d\lambda & y' &= \int_0^{\infty} E(\lambda)\bar{y}(\lambda)d\lambda \\ z' &= \int_0^{\infty} E(\lambda)\bar{z}(\lambda)d\lambda \end{aligned} \right\} \quad (18-3)$$

and by finding the corresponding unified coefficients:

$$X = \frac{x'}{x' + y' + z'} \quad Y = \frac{y'}{x' + y' + z'} \quad Z = \frac{z'}{x' + y' + z'} \quad (18-4)$$

These equations are identical to (18-2), except that \bar{x} , \bar{y} , and \bar{z} are replaced, respectively, by x' , y' , and z' ; hence, the sensations caused by nonhomogeneous radiations also appear in the color diagram and lie within the triangle defined by the spectral colors and the saturated purples. Thus, any light source $E(\lambda)$ may be represented by a single point on or in the triangle;⁷ that is, the entire gamut of possible colors lies within the area of the color triangle. The regions lying between the inner color triangle and the outer 45° triangle defined by the three fictitious primaries have no physical significance.

In summary, then, we see that any radiation may be specified quantitatively by means of the unified trichromatic coefficients and may be mapped as a point on the color triangle.

The color triangle is also of help in synthesizing the sensation of color corresponding to some $E(\lambda)$. For example, let it be required to synthesize some color C from the three primaries P_1 , P_2 , and P_3 shown in Fig. 18-4. Then the relative intensities required of the primaries are inversely proportional to the lengths of line joining C to each primary.

⁶ See P. Moon, *op. cit.*, or P. C. Goldmark, J. N. Dyer, E. R. Fiore, and J. M. Hollywood, "Color Television, Part I," *Proc. IRE*, **30**, 4 (April 1942). For use of the color triangle in color mixture, see Frank J. Bingley, "The Application of Projective Geometry to the Theory of Color Mixture." *Proc. IRE*, **36**, 6 (June 1948).

⁷ The reader is referred to *Life*, July 3, 1944 for an excellent reproduction in color, within the limitations imposed by printers' inks, of the color triangle. The same article shows representations in color of several of the statements made in this chapter.

alyzed into three positive primaries and, conversely, the three positive primaries may be mixed in some fashion to reproduce the original color. It is this fact which is the *sine qua non* of any color television system. The system itself is inherently color-blind, being capable only of interpreting colors as levels along a monochrome scale. By virtue of eq. (18-1), however, the colors in the original object may be analyzed into three primary components, each of which may then be handled as a monochrome channel in the television system. At the reproducing end two basic functions must be performed. First, the output of each monochrome channel must be converted back into corresponding values of the correct primary hue and second, the three resulting primary images must be merged together in some satisfactory manner so that they are added together to synthesize the colors in the original object.

This broad principle for transmitting images in color is illustrated in Fig. 18-6. At the pickup end the light reflected from the televised

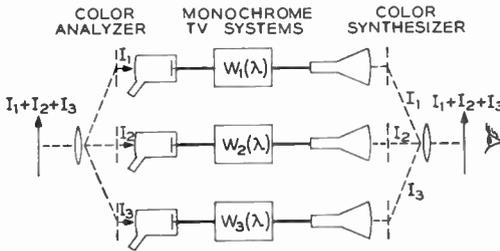


Fig. 18-6. The basic simultaneous color-television system. Reflected light from the object is analyzed into its three primary coefficients. Each primary signal is transmitted to the receiver by a separate monochrome television channel. The primary outputs are then mixed to synthesize the color of the original object.

object is sent through a color analyzer such that each camera tube sees only one primary color. Thus, for example, the upper channel develops a voltage proportional to $I_1(\lambda_1, t)$, the coefficient of its primary P_1 . The other two channels behave in a similar manner, so that each one carries a black-and-white signal whose amplitude is proportional to the coefficient of its primary in the image at any instant. In effect, then, the color image is transmitted to the receiver as three independent signals, each of which tells the strength of one of the primary colors into which the image has been analyzed.

At the receiving end a black-and-white image is reproduced on each cathode-ray tube. The color synthesizer converts each black-and-white image into its proper primary hue, and the three primary images are then combined to synthesize the color of the original object. The exact means of analyzing and synthesizing are discussed in subsequent sections. It will also be shown that three independent monochrome channels need not be used.

18-6. The Color Filter

It is apparent from the last section that a color filter of some form is necessary in the analyzing process in order to break up the original image into its three primary components. A filter is also required at the reproducing end to convert the shades of gray which appear on the cathode-ray tubes into corresponding values of the proper primary hue. Such a filter consists of a dye or pigment supported on a gelatin, acetate, or glass film and exhibits a band-pass response characteristic. The response of a typical filter of the Wratten series is illustrated in Fig. 18-7.⁹ It may be seen from the

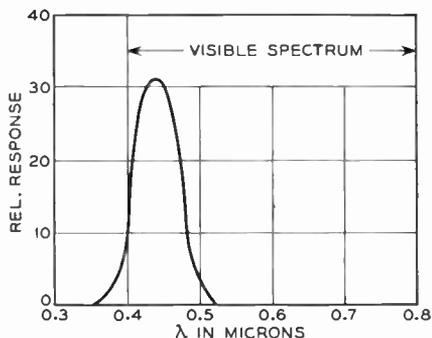


Fig. 18-7. Relative response of a color filter (Wratten 47-A).

response curve that the particular filter illustrated excludes all wavelengths except those in the vicinity of 0.44 microns, which is at the blue-violet end of the visible spectrum. Thus, if white light, which consists of all spectral colors, is shined on the filter, only blue-violet light will be transmitted and all colors corresponding to wavelengths outside of the pass band will be excluded. Hence the color analyzer

⁹ Data for the curve are from C. D. Hodgman (Ed.), *Handbook of Chemistry and Physics*, 22nd ed., Chemical Rubber Publishing Co., 1936, p. 1692.

in Fig. 18-6 may consist of three color filters, each of which passes one of the primary colors P_1 , P_2 , and P_3 .

At the risk of oversimplification we may assume for all practical purposes that white light consists of a mixture of the three primaries, red, blue, and green. Thus when white light shines on the blue filter, the red and green components are "subtracted" from the transmitted light which consists of the blue component alone. Following this line of reasoning we may describe the action of the three primary filters in equation form.

$$B = W - (R + G) \quad (R + G) = Y \quad (\text{yellow}) \quad (18-5)$$

$$G = W - (R + B) \quad (R + B) = M \quad (\text{magenta}) \quad (18-6)$$

$$R = W - (B + G) \quad (B + G) = C \quad (\text{cyan}) \quad (18-7)$$

It may be seen from these equations that the combinations blue and yellow, green and magenta, and red and cyan each add to give white. These combinations are defined as complementary colors.

It should be stressed that the last three equations give a highly simplified picture of the action which takes place at a color filter, but it is sufficient for our purposes. A more rigorous approach would involve the use of the filter response curve such as that shown in Fig. 18-7 in conjunction with a curve of intensity $v.$ wavelength of the incident radiation. The light output from the filter would then be given by the product of two curves. This procedure is analogous to that used in electrical circuits. The output of a network is given by the product of the input signal times the response of the network itself.

Still another form of color filter is available, the dichroic mirror.¹⁰ In contrast to the last mentioned filter type, which relies on the color-absorbing properties of a dye or pigment, the dichroic mirror exhibits a filtering action resulting from interference phenomena. Mirrors of this type are coated with very thin layers of materials having different indexes of refraction and exhibit the property of reflecting one color of the incident light and transmitting all the other colors present. Thus the filtering action occurs in the reflected light rather than in the transmitted light as is the case with the more familiar filter. Suitable techniques have been developed so that mirrors which reflect any one of the three primaries, red, blue, or green, may be manufactured.

¹⁰ G. L. Dimmick, A New Dichroic Reflector, *J.S.M.P.E.* 38 (January 1934).

The advantage of the dichroic mirror over the conventional filter as applied to a color television system is that it provides a more efficient utilization of the light flux available. This will become apparent when we discuss methods of optically splitting the original object into its three primary components.

18-7. Color Synthesis by Addition

Given the three primary colors, red, blue, and green, we may synthesize other colors lying within their triangle of reproduction by mixing or adding them, in which case we have an additive color process. In general, the primaries may be added in three different ways. First, small dots of the primaries may be interspersed over an area so that the eye merges them into a color mixture. This method is used in color printing, was utilized extensively by the painters of the impressionistic school and forms the basis for the dot-sequential color television system. Second, beams of light of the three primary colors may be merged by some form of lens system. This second method of color mixture is implied in Fig. 18-6: The three monochrome images which result from a filter in front of each cathode-ray tube are merged by an optical lens system so that the eye views the three primary images superimposed on a common field. As applied to television this scheme requires that three monochrome images be transmitted simultaneously and gives rise to a simultaneous color-television system.

The third system of color addition displays the primary monochrome images in sequence at a high enough rate of speed that retentivity of the retina in the eye provides the necessary superposition of the three colors. In subsequent sections it will be shown that this scheme of color reproduction forms the basis of the field-sequential color television system.

A combination of the second and third methods of color addition is used in the line-sequential color television system.

18-8. Color Synthesis by Subtraction

Colors other than the primaries may be synthesized by a subtractive process as well. The general principle is illustrated in Fig. 18-8, where the filters are of the complementary primary colors, cyan, magenta, and yellow. It may be observed that color repro-

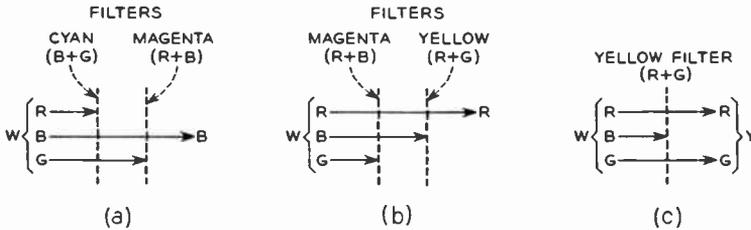


Fig. 18-8. Formation of colors from white light by subtractive-primary filters.

duction by the subtractive process requires a source of white light and the simultaneous superposition of three layers of color filters, and further that the density of these layers be variable if the entire gamut of colors in the reproduction triangle is to be reproduced. While the subtractive color process has enjoyed widespread use in color photography where the three filter layers may be deposited directly onto the film, its application to the field of television has not received as much attention as the additive color-reproducing process.

One color television system which is based on the subtractive system of color reproduction at the receiving end has been developed by the Skiatron Corporation. The basis of the system is the physical phenomenon that certain crystalline structures exhibit a varying opacity to transmitted light when they are scanned by electron beams of varying electron density. Crystals of this type have been dyed in the complementary primary colors mentioned above and incorporated in the Skiatron tube.¹¹ The proposed color-reproduction equipment employing the Skiatron is shown in diagrammatic form in Fig. 18-9. The output of each of the three monochrome channels controls the opacity of a dyed crystal screen in a Skiatron. Essentially white light from a carbon arc is focused so that it passes in series through the three tubes which behave like variable-density color filters. Since each filter subtracts one primary color in some degree from the white light, the final image which is viewed on a projection screen appears in color.

The chief advantage of the system is its theoretically high bright-

¹¹ See section 14-17; also U. S. Patent No. 2,330,172.

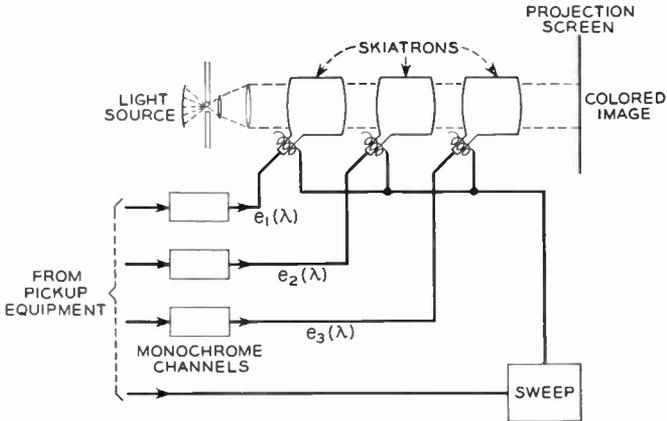


Fig. 18-9. The Siatron subtractive color-television system. Three simultaneous signals are required to synthesize colors properly by the subtractive process.

ness level since an arc, rather than a cathode-ray-tube phosphor is the initial source of light. The primary disadvantage of the system, at least at the time of writing, is that the system has not worked successfully on anything like a commercial scale. Even if the operation may be improved so that its commercial use is feasible, two additional limitations are present. First, by the very nature of subtractive color reproduction all three filter tubes must operate at the same time; hence a simultaneous system of transmitting the three monochrome signals is mandatory and no opportunity is available for utilizing a time-sharing scheme between the three primary-color signals. The second limitation, which may be reduced with care, is that the resolution is limited by the degree of register that may be obtained between the scanned images in the three tubes. That is, the final resolution depends upon how well the tubes are lined up optically and how well their scanning patterns coincide. In fairness it must be stated that this second limitation is not peculiar to the Siatron subtractive color reproducing system; it is present in any system which requires the merging of primary hue images from different sources.

At present it seems that the application of the subtractive color process is limited in television, and most of the development work has been toward systems based on the additive color principle.

COLOR TELEVISION SYSTEMS

The last few sections show that design trends for color television systems are toward the use of color synthesis by the addition of three primary colors. The principle behind these systems is that the televised image is analyzed by filters into three images, one for each of the primary colors, and that these three monochrome images are delivered to the receiving end where they are combined again to reproduce the original gradations of color. It has been pointed out that the combination of the monochrome images may be accomplished in two ways. First, the three separate images, which appear simultaneously, may be superimposed by an optical system, and second, the three images may be presented to the eye in such rapid sequence that retina retentivity accomplishes the necessary color addition. These two schemes have been called, respectively, the simultaneous and the sequential systems of color television. In the next sections we shall consider some of the basic forms of a color television system.

18-9. Basic Color Analyzing Systems

The most obvious method of breaking up a colored image into its three primary components has been illustrated in Fig. 18-6. We shall now consider some of the optical details of the method. One method of obtaining the three required images is illustrated in Fig. 18-10*a*, which shows a typical flying-spot system used for scanning color transparencies. Scanning is provided by the CRT shown at the left. Essentially white light from the fluorescent screen is passed through the color slide, and the resulting colored image is focused onto a pair of dichroic mirrors mounted normal to each other. Mirror *A* reflects only the red components of the image. The green and blue components are transmitted through the mirror.

In a similar manner mirror *B* reflects the blue components of the image to the lower phototube, which then delivers a "blue" signal of voltage. Mirror *B* is capable of transmitting the red and green components. Reference to the figure shows, however, that the two mirrors are so placed that only the green component is able to pass through to the right-hand phototube, hence it delivers a "green" signal of voltage. Thus the two crossed dichroic mirrors break up the image into its primary components and each causes a correspond-

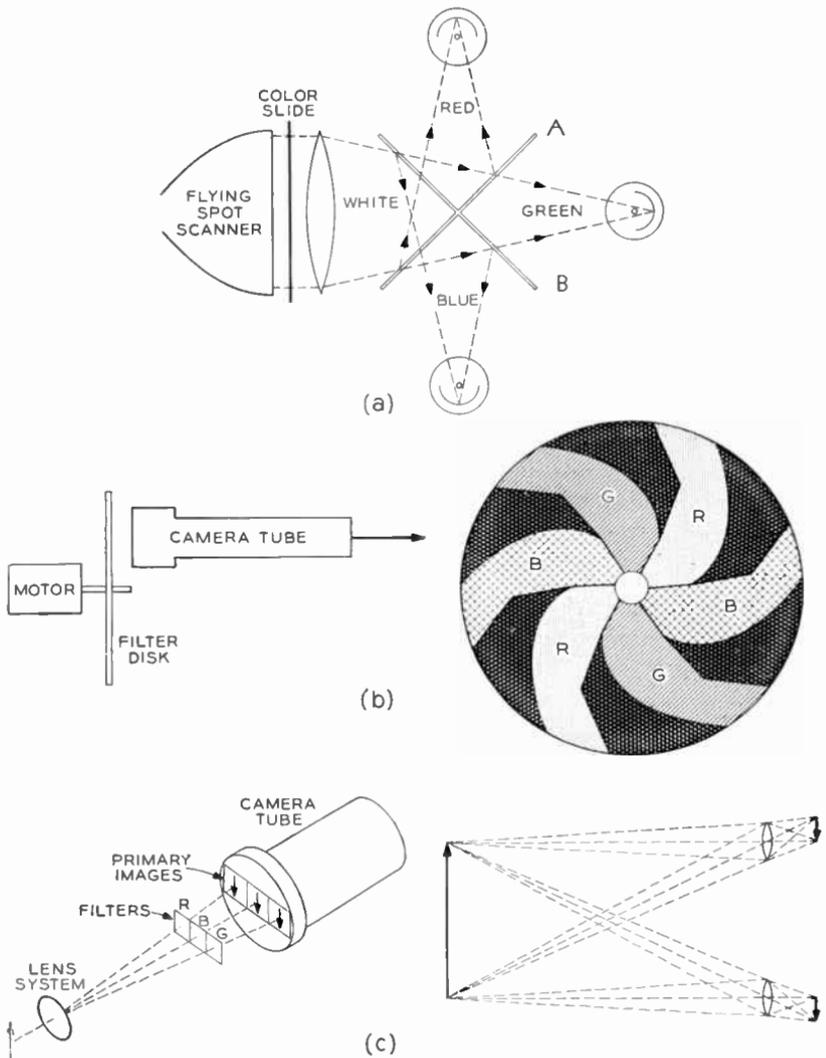


Fig. 18-10. Basic color-analyzing systems. (a) Two crossed dichroic mirrors divide the image into its three components. Each phototube delivers a voltage proportional to $I(\lambda, t)$ for a single primary component. A is a red-reflecting mirror; B , blue-reflecting. (b) A field-sequential analyzer. A rotating color filter disk causes subsequent fields to be shined on the mosaic in the three primary components of the image. The output signal is proportional over intervals of V to $I(\lambda, t)$ for each of the primaries in succession. (c) A line-sequential analyzer. The optical system projects three images onto the mosaic, each being passed through a filter. All three images are scanned by a single raster. The output signal is proportional over intervals of H to $I(\lambda, t)$ for each of the primaries in succession. A detail of two lenses in the optical system is shown at the right.

ing voltage signal to be produced by a phototube. It is a basic characteristic of this color-analyzing system, which has been used by R.C.A., that the three primary component images are produced at the same time; hence it may be termed a simultaneous or parallel color-analyzing process.

A number of variations of the simultaneous method are possible. For example, it need not be restricted to a flying-spot method of scanning. Thus, for live pickup, the scanner and slide of Fig. 18-10*a* are replaced by the usual camera lens, and the three phototubes replaced by three conventional black-and-white camera tubes such as image orthicons. With these substitutions, the optical action of the system remains unchanged. Electrically, the single raster on the flying-spot scanner is replaced by three rasters, one on each of the individual camera tubes. Extreme care must be taken to keep the rasters as nearly identical as possible, and to maintain proper registry of the three colored images relative to the scanning rasters.

Another variation of the basic simultaneous color analyzer is obtained by replacing the two dichroic mirrors, *A* and *B*, in Fig. 18-10*a*, by half-silvered mirrors and interposing a color filter sensitive to one primary hue between the mirrors and each of the corresponding phototubes. That this arrangement gives a less efficient utilization of the incident light flux may be seen from the following considerations. The two half-silvered mirrors divide the total flux equally between the three phototubes. Thus, only one-third of the total flux is directed toward, say, the red phototube. This coupled with the relatively low transmission coefficient of the red filter yields a low value of illumination on the tube photocathode and a correspondingly low value of output voltage. The same argument also holds for the two other channels.

With the dichroic mirrors, on the other hand, the division of light flux takes place on the basis of color. That is, *all* of the red light reaches the upper phototube. Furthermore, the filter transmission-loss is eliminated because no filter is used. It is for these reasons that the dichroic mirrors are superior to the half-silvered mirror and filter combination.

A second basic type of color analyzer samples the primary components of a colored image in some predetermined sequence, and hence may be termed a sequential color analyzer. One form of the sequential analyzer, which is shown in Fig. 18-10*b*, utilizes a single

camera tube, say of the image-orthicon type, in conjunction with a filter disk which is rotated by a motor. The heart of this analyzer, which forms the basis for the C.B.S. color television system, is the rotating filter disk. Shown in the diagram it consists of a circle of transparent plastic on which are mounted segments of filters in the sequence red, blue, and green. The disk is revolved at such a speed and the filter segments are so shaped that each successive field of the televised picture is exposed to the camera tube through a different filter. Thus, during one field the camera tube sees only the red components of the original image, during the second field it sees only the blue components, during the third field only the green components. The reasons for switching the filters in synchronism with and at the scanning field frequency are discussed under the C.B.S. system in a later section. For the moment we need only note that the primary components are sampled at field frequency, and we have a field-sequential color-analyzing system. In comparison with the simultaneous analyzer this system has the advantages of requiring only one camera tube and having a less complex optical system. Since only one camera tube is required, there is no problem of synchronizing three identical scanning rasters and maintaining proper register between three images and their rasters. On the other side of the ledger, the field-sequential analyzer requires a motor-driven disk or drum—a bulky combination. Furthermore, the motor must be synchronized with the electronic scanning system of the television channel, and the revolving disk displays some unfortunate (in this application) characteristics of gyroscopic action.

Still another basic form of the sequential type of color analyzer is shown at *c* in Fig. 18-10. In this case a multiple-lens system divides the incident light into three identical images side by side on the face of the camera tube. The basic layout for two components of the lens system is also illustrated in the figure. All of the small lenses produce identical images, which by proper adjustment may be made to fall side by side. Color analyzing is made to take place by interposing a color filter between each of the lenses and the camera tube face. Thus the final result shows three images, each of a different color, and all occurring simultaneously at the camera tube. The sequential nature of the analyzer is due to the method of electronic scanning used. The width of the scanning lines is increased until the raster covers all three images. Under this condition during

each line interval of the raster the output signal is divided equally between the red, blue, and green images. As viewed from the original image, the filter through which the information is sampled changes after each line; hence the system may be termed a line-sequential color analyzer. As described, this analyzer forms the basis of the color system proposed by Color Television, Inc. Its advantages are that only a single camera tube is used and no rotating filter disk is required. It has the same disadvantage as the simultaneous analyzer; it requires excellent optical registration, and further, it has poor utilization of the camera tube photocathode surface.¹²

18-10. Basic Color Transmission Systems

In the last section we studied the basic methods used for analyzing a colored image into its three primary components so that three corresponding monochrome voltage signals resulted, one for each of the primary colors. In the present section we consider the fundamental methods of transmitting these three voltage signals from the color-analyzing and scanning end to the reproducing end of the overall system. The problem here is essentially that discussed in Chapter 1. We have three sources of information. The information from all three sources may be transmitted simultaneously over three parallel channels, or the three sources may be sampled in a predetermined sequence and these samples delivered to a single communication channel. Let us see how these methods of transmission may be connected with the color analyzing systems which have just been described. Several possible combinations are illustrated in Figs. 18-6 and 18-11. Consider the former, where a simultaneous color analyzer is used. As we have already seen, three electrical signals are generated at all times that the system is in operation; hence one might conclude that it is only natural to feed each signal into an independent channel so that all three are transmitted simultaneously. Such a simultaneous-analyzer simultaneous-transmission system was used by R.C.A. in 1946. Some mention should be made of the three communication channels. If color is to be superimposed on a closed system, it is perfectly feasible to employ three separate cables as

¹² Another color analyzer quite similar to that of C.T.I. but which requires an intermediate photographic process has been proposed by Thomascolor, Inc. See, for example, John H. Battison, "Color Television Transmission Systems." *Tele-Tech*, 8, 10 (October 1949).

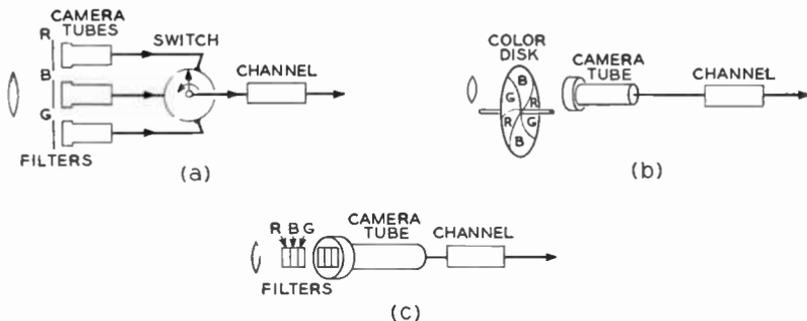


Fig. 18-11. Basic methods of color transmission. (a) Simultaneous analyzer—sequential transmission. (b) Field-sequential analyzer—sequential transmission. C.B.S. field-sequential system. (c) Line-sequential analyzer—sequential transmission. C.T.I. line-sequential system.

independent channels. If, on the other hand, an open system is to be employed, it is not practicable, as we saw in Chapter 8, to employ three separate transmitters; hence it is desirable to superimpose the three signals onto a single carrier and yet maintain them in such a form that they can be separated at the reproducing end. This may be accomplished by applying the techniques of carrier telephone operation so that each signal is confined to a predetermined band of modulating frequencies. The details for doing this are discussed later in the R.C.A. simultaneous-system section.

While it may seem natural to feed the three outputs of the simultaneous analyzer into three separate channels, this is by no means necessary. A perfectly reasonable alternative is that shown in Fig. 18-11a, where the outputs are sampled by a switching mechanism and these samples are fed to a single communication channel. In actual use the mechanical switch shown in the figure is replaced by a high-speed electronic switching circuit. This simultaneous-analyzer sequential-transmission system was considered by C.B.S. in 1940 but was rejected in favor of an all-sequential system with a color-switching rate equal to the scanning field frequency. It remained for the dot-sequential system of R.C.A. to put the simultaneous-analyzer sequential-transmission system to good use. The key to that system lies in the high rate of color switching which is used, namely 11,400,000 times a second. Thus each sample is over a small dot or area, there being some 724 dots to a single scanning

line. Further details of the simultaneous-analyzer sequential-transmission system are discussed in the section on the R.C.A. dot-sequential color television system.

Continuing the discussion of color-analyzer and channel connections we have the two remaining practicable combinations that are shown at *b* and *c* in Fig. 18-11. In each of these cases the output from the camera tube already contains the sampled information and only a single communication channel need be provided. Both of these systems are of the sequential-analyzer sequential-transmission type.

18-11. Basic Color-reproducing Systems

We have seen in the last section that there are two possible ways for the color information to arrive at the reproducing end of the color television system. First the signals from the three color images may arrive distinct from each other in three parallel channels or they may arrive in sequence over a single channel. The general problem at the reproducing end of the system, then, is to have each signal control a light source of the proper primary hue so that the three primary images are reconstituted; then these images must be merged or added together to form the final image in full color.

Where the signals arrive over three parallel channels, these functions are performed quite simply. Each channel may feed a separate CRT over each of which is placed a filter of the proper hue. The three resulting primary images then may be merged by a lens system of the type shown in Fig. 18-10*c*. It should be remembered in this connection that the lens system is bilateral in that it will form a single image from three separate light sources just as well as it will form three separate images from a single light source. Such a packaged unit of three CRT's with their associated filters and lens systems was developed by R.C.A. under the trade name "Trinoscope." The chief problem associated with the unit is the maintenance of proper registry between the component images. A misalignment in the amount of one line thickness can result in severe degradation of picture quality.

A number of other physical arrangements of the three CRT's may be used. For example, consider the direct-viewing arrangement of Fig. 18-12*a*. Here each tube utilizes a phosphor which phosphoresces in one of the primary colors. Two dichroic mirrors used in conjunc-

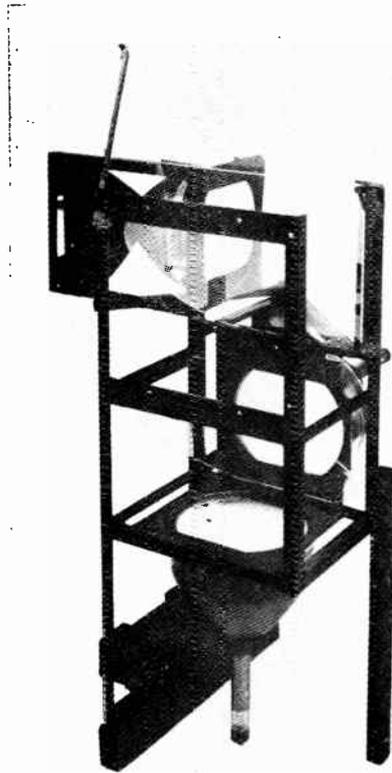


Fig. 18-12a. 3-CRT production of a color image, using 10-inch direct-view tubes. (Courtesy of Radio Corporation of America.)

tion with these tubes cause the final full-color image to be reflected from the silvered mirror at the top of the whole assembly.

A similar but more compact unit which utilizes small projection tubes is illustrated in Fig. 18-12*b*. Once again, colored phosphors are used in conjunction with crossed dichroic mirrors, and a lens is used to project a complete magnified color image onto a separate viewing screen. While the two systems of Fig. 18-12 provide better utilization of light than does the Trinoscope, they share its disadvantage of requiring extremely good registration of the separate images.

Consider next how the image-reproducing problem is handled in the C.B.S. all-sequential system. Since, as may be seen from Fig.

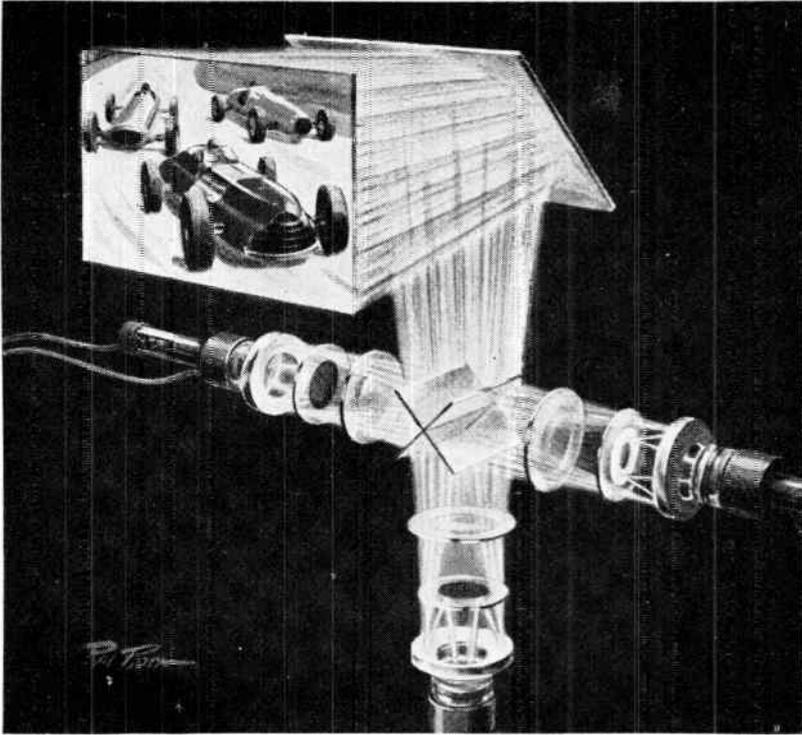


Fig. 18-12b. 3-CRT production of a color image, using projection tubes. (Courtesy of Radio Corporation of America.)

18-11b, the color sampling is produced by the rotating color filters, it is only necessary to have a similar filter wheel in front of a single CRT which receives the sequential signals from the communication channel. One important advantage is immediately apparent: Only a single reproducing tube is required since the filters are rotated. No problem of registry is present but the rotation of the color disks at both ends of the system must be in synchronism with each other and with the electronic scanning process.

In reproducing the image delivered by the C.T.I. system of Fig. 18-11c, again only a single CRT is required. Then, if the wide three-image horizontal sweep is used at the receiver, the three component images will appear side by side. These with three primary filters and the image-merging lens system reconstitute the final full-

color picture. The need for good registration of the three images is at once apparent.

The remaining problem to be discussed in this section is the manner in which the dot-interlace signal delivered by the system of Fig. 18-11a is handled. This is done by having a switching mechanism at the receiver which is identical to and operates in synchronism with the sending-end switch. The output from this mechanism consists of three separate signals, one corresponding to the dot samples of each primary hue in the original image.¹³ Once the three channels are separated, they may be fed to a Trinoscope or to either of the reproducing assemblies illustrated in Fig. 18-12.

All of the reproducing systems described thus far in this section are bulky, require moving parts, or some sort of filter external to the cathode-ray tube. It has long been the goal of the industry to produce a single cathode-ray tube which would be capable of generating all three primary color images and merging them within itself. The advantage of such a device in reducing the size and physical complexity of the color television receiver is obvious. One of the earliest attempts at building such a tube was made by John L. Baird before the war in England. Shown in Fig. 18-13a Baird's Telechrome tube utilized three separate electron guns, all of which were directed toward a multiple-colored phosphor. The tube never gave satisfactory operation because of the extreme difficulty in making the three guns produce identical rasters on the composite phosphor. It can be seen from the gun locations that all three require keystone correction. As a result of the Telechrome tube it was recognized quite early in the art that a single-gun three-phosphor tube would be required, or at least a multiple-gun tube, the guns being so arranged that all three electron beams would be deflected by a single common deflecting yoke. This poses a severe problem. How shall each phosphor be caused to phosphoresce at the proper time?

One answer to the problem lies in phosphor separation on a voltage basis. One tube, the Chromoscope,¹⁴ which utilized this principle

¹³ It is important to notice here that even though the switch feeds three parallel channels, no single channel carries a continuous signal at all times. Each channel carries sample signals of a single primary color, which occur at a rate of 5,800,000 per second.

¹⁴ See, for example, A. Bronwell, "New Viewing Tube for Color TV." *Tele-Tech*, 7, 3 (March 1948).

is shown in Fig. 18-13*b*. Four screens are placed in the tube at the viewing end. The first screen, *P*, is transparent to the electrons and maintained at a high positive accelerating potential. The three remaining screens are coated, respectively, with red, blue, and green phosphors and are manufactured so that a third of the total beam electrons goes to each of them. Color selection is provided by apply-

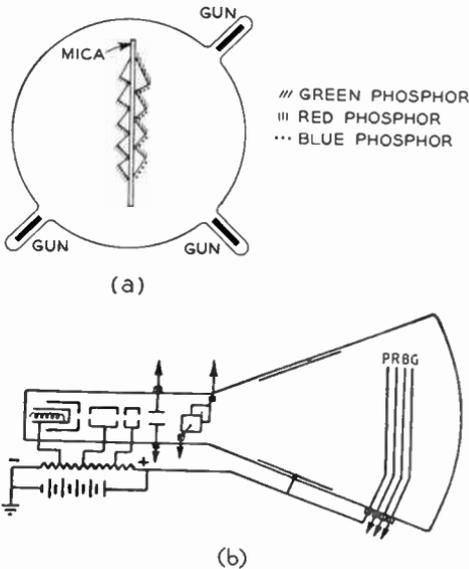


Fig. 18-13. Reproducing tubes that utilize multiple-color phosphors. (a) Telechrome tube. (b) The Chromoscope. (Courtesy of *Tele-Tech.*)

ing a high positive voltage to each screen in sequence. For example, during the interval that a "red" signal sample is applied to the gun, a high voltage is applied to the red-glowing phosphor screen and a red image is reproduced. During the blue sample interval the high voltage is applied to the blue-glowing screen and so on. Thus color selectivity is voltage-controlled.

Some registration problem is present in the Chromoscope because of parallax since the three color screens do not lie in the same plane. Thus in direct-view applications the position of the viewer is critical.

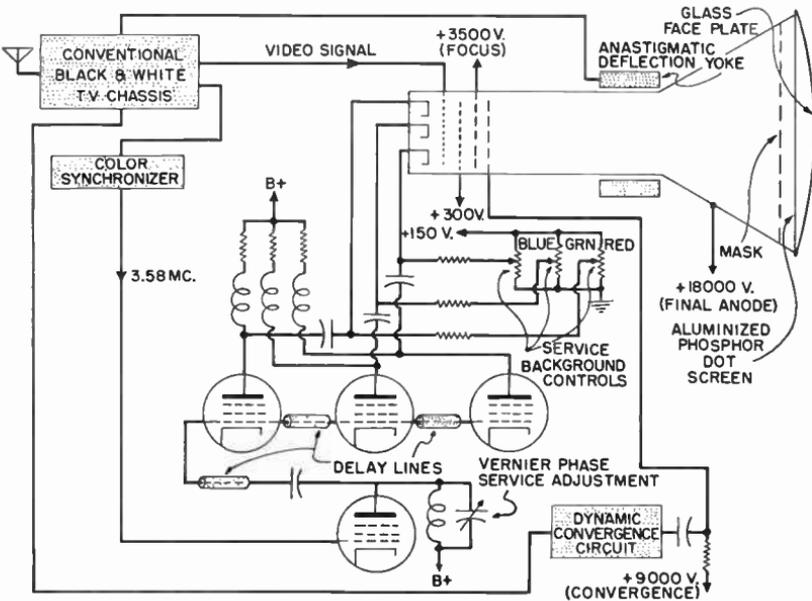
Parallax may be eliminated if the tube is used in conjunction with a projection system.

It should be immediately apparent that all three screens do not phosphoresce simultaneously; hence the Chromoscope is adaptable only to a system which delivers the color signals in sequential form. The tube is adaptable to either field or line sequential systems because the switching of voltage between screens may be handled at a high rate electronically.

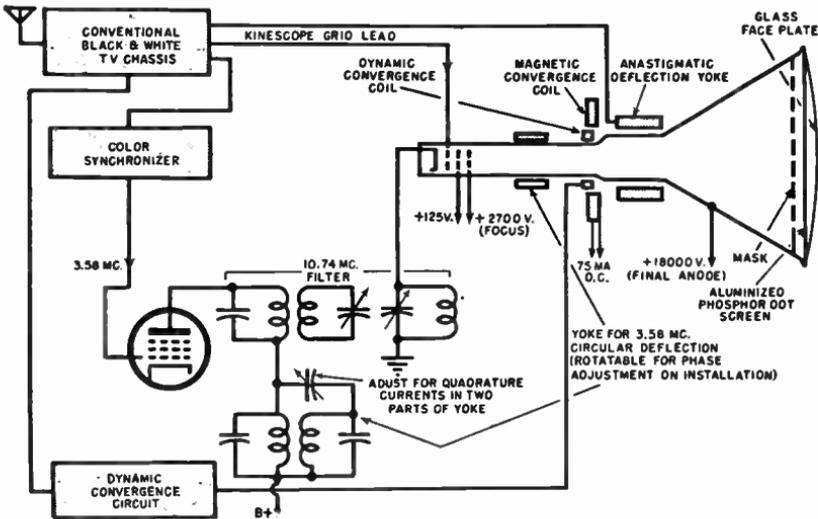
Two other three-color tubes which were announced by R.C.A. during the F.C.C. color television hearings in 1950¹⁵ are illustrated at *c* and *d* in the figure. Both tubes, which were developed primarily for the dot-sequential color system, eliminate parallax between the three primary images by having all three phosphors in a single plane. This is accomplished by covering the viewing screen with an orderly arrangement of dots of phosphors grouped in threes. Each group consists of one red, one green, and one blue phosphor dot, each of which is aluminized to increase the screen brightness. The dot size is small, there being 117,000 groups of three in the original model, or 351,000 dots in all on the 9- by 12-in. active face on a 16-in. metal-cone tube. Switching from one color dot to another is accomplished with the aid of a mask electrode which is perforated with 117,000 holes, or one hole for each phosphor-dot group. The exact manner of color selection is best described for the two tubes separately.

In the three-gun tube (Fig. 18-13c) each hole in the mask is so positioned with respect to its associated three-color dot group that it screens each color dot from all but one of the electron guns. That is, the angle of approach of the electron beam from each of the guns is arranged so that, in passing through a mask aperture, the beam can hit only one color dot in the group of three. This holds true over the entire tube face. Regardless of what point in the raster is being scanned, the beam from the "red gun" can only reach the red phosphor dots. The same holds true for the beam from the blue and green guns also. Thus, color selection is determined by whatever gun is emitting electrons at any given instant. Gun emission is controlled by applying keying pulses, which are derived from the color switch, in sequence to the cathodes of the three guns. Notice,

¹⁵ See, for example, "RCA's New Direct-view Tri-color Kinescope." *Radio and Television News*, 43, 6 (June 1950); or Vin Zeluff, "Tubes at Work." *Electronics*, 23, 6 (June 1950).



(c)



(d)

Fig. 18-13 (cont.) (c) R.C.A. three-gun tricolor kinescope. (Courtesy of *Radio and Television News*.) (d) R.C.A. single-gun tricolor kinescope. (Courtesy of *Radio and Television News*.)

then, that color switching does not take place in the video channel at all. The sequential video information is fed to a control grid common to all the guns; switching is accomplished in the cathode circuits. The resulting image appears in full color because the individual color dots are so small that they effectively merge together when observed from normal viewing distance.

Since the three guns cannot occupy the same physical position, some problem arises in making their beams converge properly over the entire tube face. This is overcome by changing the voltage on a fourth grid in the gun assembly as the beams move over the scanning raster. The necessary correcting voltage, or dynamic convergence voltage, is derived from the horizontal and vertical sweep systems in the receiver.

Color separation in the single-gun tricolor kinescope is accomplished in essentially the same manner as in the three-gun tube, except that the position of the single electron beam is rotated by an auxiliary deflection yoke so that it appears to originate from three different sources at different times. Each of these virtual sources corresponds to one gun in the three-gun tube, so that the rotating beam in conjunction with the perforated mask and three-dot-group phosphor screen yields an image in full color. Details of the auxiliary yoke circuit may be identified in Fig. 18-13*d*.

Notice again that no switching is used in the video channel which contains the dot-sequential color information. The actual switching is provided by the magnetic rotating yoke, which ensures that a phosphor dot of proper color is selected at any given instant. Driving currents for the beam-rotating yoke are derived from a color synchronizer so that the color dot selection is properly synchronized with the video color samples. Dynamic convergence is also applied through a small auxiliary focusing coil.¹⁶ We now have described the basic elements of several possible types of color-television systems. It remains to consider the details of some of the complete systems which have been proposed for adoption by the Federal Communications Commission.

¹⁶ At the time of writing only experimental models of the two types of tricolor kinescopes have been produced. Indications are that they might well serve as the reproducing unit in the color television system that will finally be standardized by the F.C.C.

18-12. Bandwidth and Compatibility

Two basic problems which arise in setting up a color television system warrant special attention before we consider the details of any particular system. These are concerned with bandwidth and compatibility of the color system with the commercial black-and-white system. A color television system is said to be compatible if its signal as radiated can be received by a conventional black-and-white receiver and viewed as a *monochrome* image on that receiver without modification, save that an R-F converter unit external to the receiver may be required. The possibility of the converter is allowed because at present, pending an F.C.C. ruling, it is not known whether color transmission will be authorized on the current 12 black-and-white channels or whether it will be restricted to proposed channels in the U.H.F. portion of the spectrum. When one reckons the number of commercial receivers in use at the present time—which is in the millions—the need for a compatible color system is apparent; the black-and-white receivers should not be rendered obsolescent by the new system of transmission.

In almost direct opposition to compatibility is the bandwidth required in color transmission unless resort is made to special techniques, such as multiplexing or a change in transmission standards. The reason for this may be seen from Hartley's law. This law states that if the effects of noise be neglected, the quantity of information which may be transmitted in a bandwidth Δf in a time Δt is proportional to the product $\Delta f \Delta t$. It may be seen at once that if a color image of geometric resolution equivalent to that of a black-and-white television image is transmitted, either a greater bandwidth or a longer frame interval (time to transmit a single picture) is required than in the black-and-white system. This follows because the color image contains more information than its monochrome counterpart. Basically, then, a color system will either have a bandwidth greater than 6 mc, or scanning standards other than those of the black-and-white system will be required. In discussing specific color proposals in the remainder of the chapter we shall see how these basic limitations can be overcome with the use of multiplexing.

It will be observed in the work that follows that the actual multiplying techniques used are an extension of the sampling concepts described in Chapter 1 and the section on interlaced scanning.

THE R.C.A. SIMULTANEOUS COLOR TELEVISION SYSTEM¹⁷

As early as 1939 the R.C.A. Laboratories had experimented with a simultaneous color television receiver, but found that technical limitations at that time prevented a satisfactory means of recombining the three color images at the receiver. Subsequent work on a sequential system led to use of that form by the National Broadcasting Company in 1941, but rapid advances in tube and circuit design during the war years made the earlier simultaneous system feasible, and further development work was activated on a simultaneous color system which was to satisfy the requirements of compatibility; that is the scanning standards of the black-and-white system were to be used. In essence, the R.C.A. proposal was a system like that illustrated in Fig. 18-6 where the three color images are transmitted in parallel. It is reasonably safe to say on the basis of testimony presented in the F.C.C. color television hearings in 1950 that the simultaneous system to be described is obsolete and will not be used in practice. It is worthy of study, however, because it illustrates several of the basic problems encountered in color transmission, it shows how special techniques may be applied to increase bandwidth utilization, and it serves as a reference for the other systems to be described. Our discussion will consider, first, ways and means of transmitting the three color channels, second, the pickup equipment, and third, the receiver and viewing equipment.

18-13. Color-channel Bandwidth

In analyzing the three primary images which are developed by the color-pickup equipment, it was found experimentally that the green signal contained sufficient detail and contrast so that it could be used to provide an excellent black-and-white picture; hence it was decided to make the green standards identical to those of the commercial black-and-white system in order that the design criterion

¹⁷ "Simultaneous All-Electronic Color Television." Progress Report by the RCA Laboratories Division, Princeton, N. J., *RCA Review*, VII, 4 (December 1946). R. D. Kell, "An Experimental Simultaneous Color-Television System," Part I—Introduction; G. C. Sziklai, R. C. Ballard, and A. C. Schroeder, Part II—Pickup Equipment; K. R. Wendt, G. L. Fredendall, and A. C. Schroeder, Part III—Radio Frequency and Reproducing Equipment, *Proc. IRE*, 35, 9 (September 1947).

of compatibility be satisfied. It was also found that the acuity of the eye in the blue end of the visible spectrum was limited so that a narrow bandwidth could be allocated to the blue signal without degradation of the reproduced color image. The red signal was of equal importance to the green in color reproduction so that equal bandwidths were chosen for the red and green signals. Thus to meet the requirements of color reproduction, and black-and-white reproduction from the green channel, the following transmission standards were chosen:

| | <i>Red channel</i> | <i>Blue channel</i> | <i>Green channel</i> |
|-----------------------------|--------------------|---------------------|----------------------|
| n | 525 | 525 | 525 |
| f_f , fields per sec | 60 | 60 | 60 |
| Interlace | 2:1 | 2:1 | 2:1 |
| Video bandwidth, megacycles | 4.5 | 1.3 | 4.5 |

In order to allow reception of the green channel by a black-and-white receiver the standard R.M.A. synchronizing wave form and the F-M sound program were associated with the green video signals. The bandwidths of the three simultaneous monochrome-picture signals are illustrated at *a* in Fig. 18-14.

18-14. Methods of Transmission

The pickup equipment developed for televising colored slides is of the simultaneous type which has already been described. It is apparent at once that the three signals cannot be used as is to modulate a single carrier simultaneously. Two methods of superimposing the three signals into a modulated wave are illustrated in Fig. 18-14. The obvious method is shown at *b*, where, in effect, three separate transmitters are provided, one for each color channel and operating at a different carrier frequency from the others and whose outputs are combined in a triplexer unit to feed a single transmitting antenna. While such a system is practicable, it would be expensive, requiring in effect three relatively high-power transmitters.

A second means of combining the three monochrome signals utilizes the previously mentioned carrier principle as shown at *c* in Fig. 18-14. The green video signal remains as is while the two other channels are used to amplitude-modulate subcarriers at low power level. One sideband of each of these two modulated subcarriers is suppressed by a filter and the two remaining sidebands are combined in a mixer to give a new combined video signal. The effect of the subcarrier

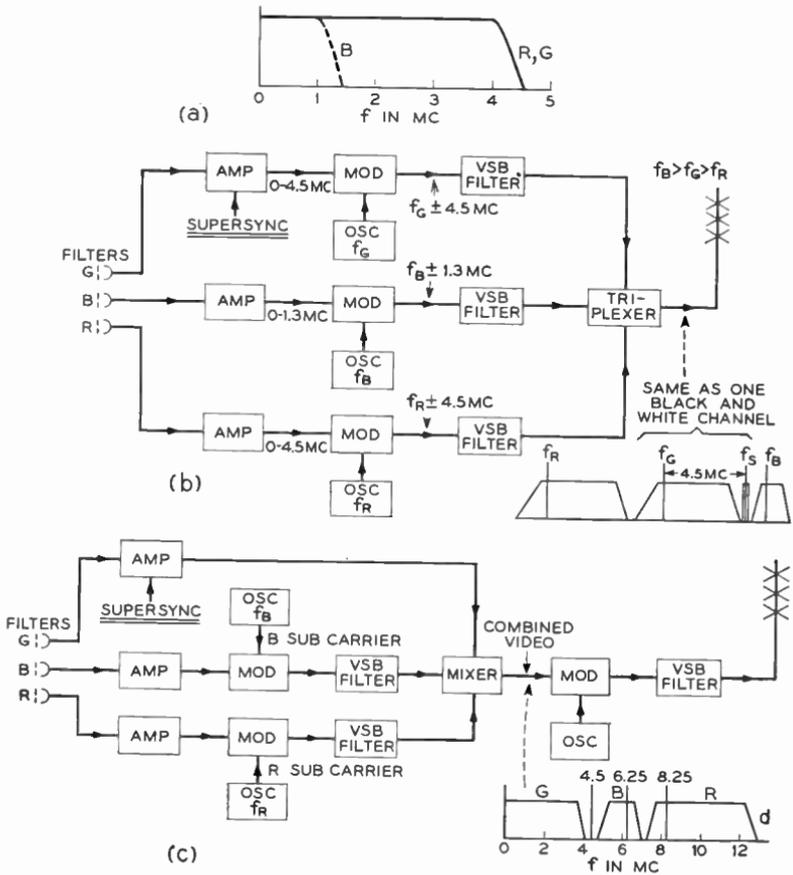


Fig. 18-14. The R.C.A. simultaneous color system. (a) Video bandwidths allowed for the three color channels. (b) The signal from each color channel modulates a separate R-F carrier, resulting in the three-carrier type of transmission. (c) The blue and red signals are shifted upwards by 6.25 and 8.25 megacycles, respectively, to form a combined video signal 13 megacycles wide, which is used to modulate a single R-F carrier. This gives a subcarrier type of radiated signal.

modulation on the red and blue signals is to shift those signals upward in frequency by an amount equal to the subcarrier frequency. The resulting combined video signal has the spectrum indicated at d in the figure. Notice that the over-all effect is to combine the three monochrome signals into a single video modulating signal which

covers a frequency range from zero to 13 megacycles with a specific portion of that spectrum allotted to each primary color. Since these are separated in frequency, they may be separated from each other in the receiver. Of the two types of modulation just described, the latter was chosen because it was based on techniques commonly used in carrier telephone and radio work and hence required less developmental work. It also had the additional advantage of allowing the use of a simpler receiver than the three-carrier system. The final signal which is radiated from the transmitting antenna contains the three primary monochrome images, all on a single common R-F carrier.

While our discussion has been confined to the system block diagram only, this restriction must be stated about certain of the component circuits in the over-all system. Wherever the three color channels are mixed or separated—that is, in any modulator, mixer, or detector—extreme care must be exercised to ensure that the circuit is linear in order that cross-modulation between the colors does not occur.

18-15. The Receiver

The major components of a receiver designed to handle the subcarrier type of radiated signal are shown in the block diagram of Fig. 18-15. The front end of the receiver is similar to that of the black-and-white receiver, except that it is designed to handle an unusually broad band of roughly 14.5 megacycles. The same bandwidth requirement is imposed on the common I-F amplifier system and the linear detector which handles all three signals. The fre-

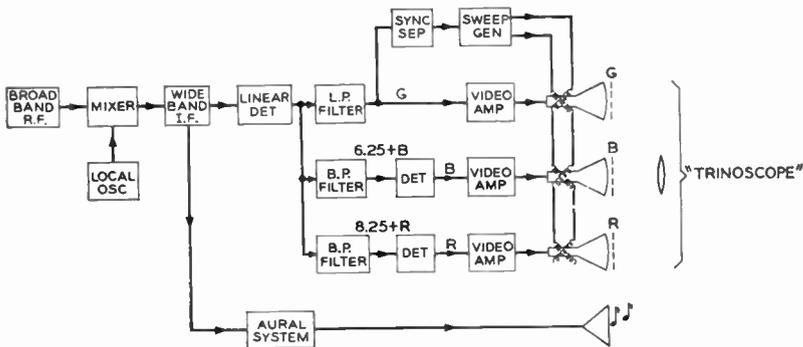


Fig. 18-15. Simultaneous color receiver for subcarrier type of signal.

quency distribution at the output of the detector is the same as that of the combined modulating signal shown at the inset of Fig. 18-14*d*. Since the three color signals occupy different portions of the spectrum, they may be separated by the use of passive filters. Two additional detectors then serve to separate the red and blue signals from their respective subcarriers of 8.25 and 6.25 megacycles. The three final signals are amplified and applied one to each of three separate kinescopes in the Trinoscope assembly. The supersync signals are separated from the green video channel by a conventional sync stripper and are applied to a common sweep-generating circuit, which feeds the three deflection yokes in parallel.

As previously described the three cathode-ray tubes of the Trinoscope assembly are arranged physically so that their face-centers form the corners of an equilateral triangle and each tube face is covered with the appropriate filter so that a monochrome image is produced there. Superposition of the three images is accomplished by the lens system illustrated in Fig. 18-10*c*. The axis of each lens is displaced from the center of its tube toward the center of the triangle by an amount sufficient to register the three images on a common projection screen. Color reproduction results from a direct addition of the three primary color images on the screen.

The means by which the color R-F signals are received by a converter and a black-and-white receiver are easy to understand. The converter is designed to tune to the green portion of the R-F signal with a 6-megacycle bandwidth. With this pass band properly centered, the green band plus the aural signals, whose carrier is 4.5 megacycles above the green carrier, are accepted and "beat-down" by the local oscillator and mixer to the V.H.F. band. The resulting signal from the converter is similar to that from a commercial black-and-white transmitter and may be applied directly to the antenna terminals of the receiver. The operation of the receiver is unaffected: frequency conversion and sound, sync, and picture separation all occur in the usual manner and the final image appears in the usual black-and-white presentation.

If one were to appraise the simultaneous color system just described, the following chief factors would be evident. Its primary advantage is that the radiated signal may be used by a black-and-white receiver in conjunction with a frequency converter, even though only a black-and-white image would be displayed. The chief dis-

advantage is the three-unit kinescope assembly required at the receiver, a unit in which exact electrical and mechanical adjustments are required to ensure correct register of the three separate primary color images. Further, the radiated signal requires a bandwidth far in excess of 6 megacycles.

18-16. Bandwidth Reduction by the Mixed Highs Technique

A reduction of roughly 25 per cent in the bandwidth required for transmitting the video signals in the simultaneous color system may be effected by utilizing the principle of "mixed highs," which is based on the experimentally determined fact that satisfactory reproduction of the color picture results if the large details of the image are reproduced in color and the fine-grain details in shades of gray.¹⁸ To accomplish this in the color television system the upper two megacycles of each color channel are filtered out from the channel, and these filtered-out "highs" are then added together to form a fourth channel of the "mixed highs," which contains the fine detail of the image. One possible layout for accomplishing this is illustrated in Fig. 18-16. The output from each camera tube is passed through

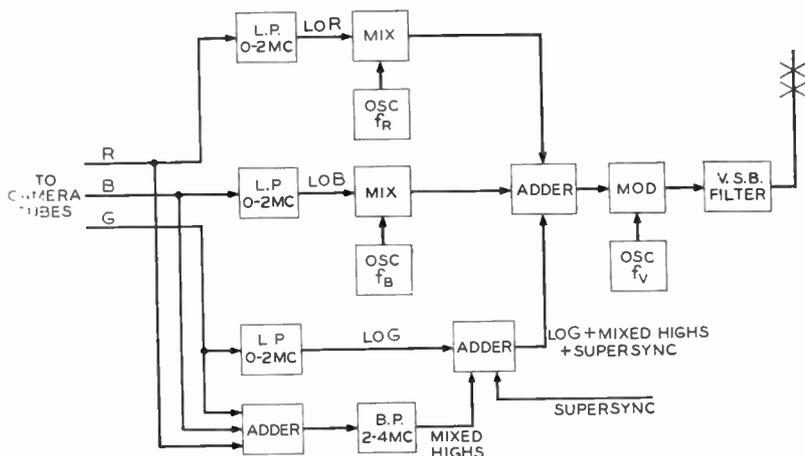


Fig. 18-16. Block diagram of a transmitter for a simultaneous color program that uses the principle of mixed highs.

¹⁸ The same technique is used in color printing, where fine details are printed in black.

a low-pass filter, having a 2-megacycle cutoff frequency. The resulting "red lows" and "blue lows" are handled in the same manner as previously described. The three camera outputs are also combined in an adding circuit and passed through a 2- to 4-megacycle band pass filter whose output, the "mixed highs," is then added to the supersync and the "green lows." It may be seen from the remainder of the diagram that the spectrum of the radiated signals consists of a band for the greens and "mixed highs," one for the "low blues," and one for the "low reds." The saving in bandwidth over the last system results from reducing the red and blue bands to only 2 megacycles each.

At the receiving end an inverse process is required: The "mixed highs" are recombined with the "low reds" and "low blues" so that each of the CRT's, say in a Trinoscope, reproduces the 0- to 2-megacycle components of one primary color plus the mixed highs.

The addition of the mixed high principle to the simultaneous process does not affect the compatibility of the system. The green channel, which includes the highs from all channels, plus the associated sound program still lies within the standard 6-megacycle bandwidth and hence may be received and reproduced as a monochrome image by a black-and-white receiver. We shall see that the mixed high principle may also be used for bandwidth reduction in other color transmission systems.

THE C.B.S. SEQUENTIAL COLOR TELEVISION SYSTEM¹⁹

The Columbia Broadcasting System entered the field of development work in color television in the early part of 1940. At that time both concepts of color transmission, sequential and simultaneous, had already been demonstrated, the former by Baird in England in 1928, and the latter by the Bell Telephone Laboratories in 1929 (and as we have seen by R.C.A. in the latter part of 1940). A survey of the field led to the belief on the part of the C.B.S. staff that the

¹⁹ P. C. Goldmark, J. N. Dyer, E. R. Piore and J. M. Hollywood, *op. cit.*; P. C. Goldmark, E. R. Piore, J. M. Hollywood, T. H. Chambers, and J. J. Reeves, "Color Television," Part II. *Proc. IRE*, **31**, 9 (September 1943); P. C. Goldmark and R. Serrell, "Color and Ultra-High Frequency Television." *Proceedings of the 1st National Electronics Conference*, October 1944; P. C. Goldmark, J. N. Dyer, E. R. Piore, and J. M. Hollywood, "Color Television." *J.A.P.*, **13**, 11 (November 1942): statement of P. C. Goldmark, *op. cit.*

sequential system of color transmission was the more feasible, and their efforts were directed toward developing a color television system of the sequential type of such quality that it would be immediately available for commercial telecasting for the general public. It was also believed that the high entertainment value of the colored images warranted the development of the color system on its own, unhampered by any restrictions of the black-and-white system. It was this philosophy that resulted in the system proposed for adoption by the F.C.C. in 1947.

The basic form of this sequential color television system is illustrated in Fig. 18-17. A single pickup, communication channel, and

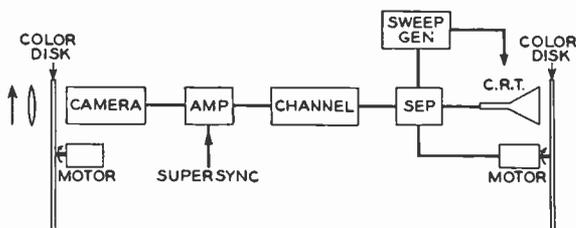


Fig. 18-17. The basic field-sequential color system. The color filters are located on rotating disks placed in front of the camera tube and the CRT.

reproducing system, all of the field-sequential type, are used. Color perception is added to the system by a rotating color disk at each end of the system. These two disks are divided into segments which are covered in a sequence of red, blue, and green color filters and are rotated in synchronism. Thus at the camera end, the camera views the televised object in a series of primary color images, that is, the camera sees first a red image, then a blue image, then a green image, and so on. These color samples of the object occur in sequence and so may be applied to the single, common communication channel. At the receiving end, the cathode-ray tube screen is viewed through a similar set of filters so that the eye sees the red, blue, and green image components in such rapid sequence that the three primaries are added by retina retentivity to give the impression of the image in full color.

The sequential color system in essence, then, is a black-and-white television system to which are added two rotating filter disks, which

serve as color analyzer and synthesizer. Since the two disks must rotate in synchronism and in phase, some form of color synchronizing signals must be provided. In the paragraphs which follow we shall consider the standards used in the sequential system and means of superimposing the color sync signals on the composite video signal.

18-17. The Color-switching Interval

The first problem to be decided in any sequential color system is the color-switching interval or portion of the scanned image which is scanned under a single primary-color filter. In essence, the question whether the filters should be switched after each element, line, field, or frame must be answered. In the U.S. color system a single image is scanned at the camera tube and at the CRT. Furthermore, the conventional 2 to 1 interlaced scan of the black-and-white system is used in conjunction with rotating color-filter disks. Subject to these limitations of the system the question may be answered in the following manner.

If the filters be changed after each element is scanned, no storage of charge in the camera tube may be utilized. The reason for this may be seen from the following considerations. If there is to be no cross-contamination of the primary colors, each element between scans must store charge under one filter only. Thus if the filters are switched after each element is scanned, no time is allowed for charge storage, and after-element switching requires the use of an instantaneous type of pickup tube. The advantages of modern camera-tube development are literally thrown away, and at best only a relatively small output signal will be developed by the camera. Another factor which would result in an even lower signal-to-noise ratio is the low transmission coefficient of the color filters, which acts to reduce the incident flux on the camera tube. For these reasons the element-by-element switching of the filters is deemed inadvisable.

If the filters be switched after each horizontal scanning line, the storage time is reduced by a factor of $2/n$ (n = number of scanning lines) as compared to the storage time available when the filters are switched after each field. We have already seen that the output of a storage tube is proportional to the storage time, thus the signal-to-noise ratio would be reduced by the same factor.

If, on the other hand, filter switching occurs after each frame, a

very high frame frequency would be required to prevent color-flicker, a problem discussed in the next section. A high frame frequency is objectionable because it raises the required video bandwidth for any given resolution. For these several reasons field-by-field switching of the color filters was considered to offer the best compromise. Another factor which favors field switching rather than line or element switching is that the color filters are rotated mechanically and considerable difficulties arise in designing the shape of the filter segments on the disk and rotating the disk at a sufficiently high rate to produce the required action.

18-18. Frame and Field Standards

Once the filter-switching rate is determined, the standards for the frame and field frequencies may be established. The scanning sequence with the filters changed after each field is illustrated at *a* in Fig. 18-18. It should be observed that six fields are required

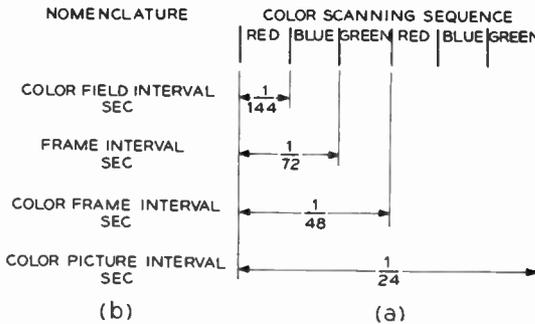


Fig. 18-18. Color-scanning sequence, intervals, and nomenclature for the field-sequential color system.

to scan the complete image in all three colors, and that any given area must be scanned four times to be covered in one color during both the odd and even lines. It is this last fact that brings about the problem of color flicker, which has been mentioned previously. To illustrate this effect, consider that a certain televised object contains a large area of a single primary hue, say red. If, then, the black-and-white field frequency of 60 fields per second be used, that red area in the reproduced image will be illuminated once every third field or at a rate of $(\frac{1}{3})(60) = 20$ times per second. That area

would appear black during the blue and green fields for it contains no components of these colors; hence, in the example the flicker rate in the red area would be 20 flicks per second, a value which lies below the critical flicker frequency. Thus to reduce color flicker, the C.B.S. system was based on a higher field frequency of 144 (as opposed to 60) fields per second. With a 2 to 1 interlaced scan the corresponding frame frequency is 72 frames per second. Since two separate scanning processes are involved, one electronic, and one mechanical for the color filters, a new set of terminology is required as shown in Fig. 18-18*b*. The color field frequency of 144 per second is the highest vertical scanning rate in the raster and is also the rate at which the color filters are switched. Two color fields comprise one frame of the raster. One color frame interval is the time required to scan through three filters, and one color picture interval is the time required to scan the entire image in all colors in both fields. It may be seen at once that the standards shown in Fig. 18-18*b* mitigate the problems of scanning motion-picture film, which runs at 24 frames per second.

Originally the black-and-white standard of 525 scanning lines was carried over into the color system; hence the required video bandwidth for the same resolution would be

$$f_v = 4.5 \left(\frac{72}{30} \right) = 10.8 \text{ megacycles} \quad (18-8)$$

As a practical matter, the bandwidth was restricted to 10 megacycles with a slight loss in horizontal resolution, which is more than compensated for by the presence of color in the final reproduced image.

Later, after the F.C.C.'s decision that transmission must be confined within a 6-megacycle bandwidth, C.B.S. lowered the horizontal scanning frequency to meet this requirement so that the final number of lines was reduced to 405. This reduction in the number of lines reduces the vertical resolution by 23 per cent, and the horizontal resolution had to be reduced by roughly 45 per cent in order to keep the signal within the 6-megacycle limit.

18-19. Color Disk Phasing and Speed Control

It has been stated that the field-sequential color transmission system consists essentially of a black-and-white television channel

to which is added two rotating color disks. A brief consideration of the problem will show that these disks must be in color-phase at all times if the proper colors are to be synthesized at the receiver; that is, when the camera tube is covered by a red filter, the cathode-ray tube at the receiver must also be covered by a red filter and so on. This requirement can best be met if the disks at both ends of the system have the same number of segments and if they rotate at the same speed. Hence we next consider means of driving all the color disks in the system at a common speed.

In the development of the C.B.S. system a number of speed-control systems have been tried. The obvious solution of using synchronous motors locked-in to the power line frequency was rejected because frequently receiving sets served by a single transmitter operate from different power systems which are not maintained in synchronism with each other. A second method which employed phonic motors was designed in which synchronism was obtained by locking the motor rotation to the vertical field frequency. Still another scheme which utilizes induction motors that are braked to the correct speed will be described. The basic circuit for the system is illustrated in Fig. 18-19. The connection between the color-

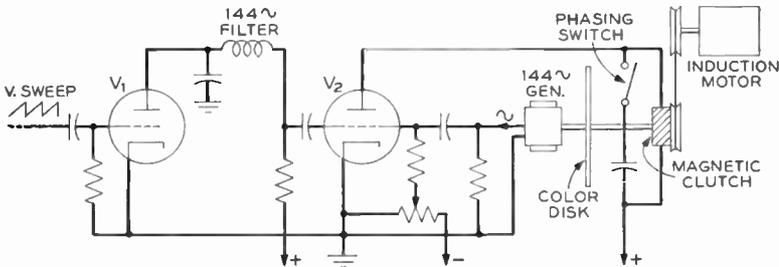


Fig. 18-19. Basic circuit for color-disk synchronization. A control current derived from the vertical sweep determines the braking action of the magnetic clutch. (After Goldmark *et al.*)

disk shaft and the drive pulley is through a magnetic clutch. Since slip between the two shafts may be controlled by varying the current through the clutch, the control circuit serves to generate the proper braking current, using the vertical-sweep voltage as a reference.

The filter in the plate circuit of V_1 delivers a 144-cycle sine wave

to the grid of V_2 , which is also fed from the output of a generator connected to the color-disk shaft. When the disk is rotating at the correct speed, the two voltages applied to V_2 will be of the same frequency and in phase. The bias on V_2 is adjusted so that rectification occurs, and pulses of current, whose average value depends upon the phase of the two voltages, flow through the brake. Filtering action is obtained with a condenser shunted across the brake winding. If for any reason the disk speed changes, the voltages on V_2 shift in phase, the brake current changes, and speed correction is obtained. Still another variation of the system replaces the magnetic clutch by a saturable reactor. The output of V_2 may then be used as the control field for the reactor, whose output determines the motor speed by governing the armature current.

It should be noticed that the control circuit only ensures proper speed of the color disk, but color ambiguity may occur. To remedy this condition a color phasing switch is wired in series with the brake filter condenser. If this switch is opened, the filtering action is removed, pulses of current, rather than a d-c current, flow through the clutch, and slipping occurs. The switch is left open until the phasing is correct. It is then closed and the speed-control circuit takes hold.

The system of disk-phasing just described may be made automatic by increasing the complexity of the over-all system. One form of automatic phasing which, in effect, replaces the manually operated phasing switch is shown in Fig. 18-20. It is apparent that automatic disk-phasing requires that the receiving equipment be able to identify some fixed point in the color sequence. To this end a color-identifying pulse may be included which appears in the supersync wave form just prior to each red field, that is, once every third field. The repetition rate of the color-identifying pulses will therefore be one-third of the color field frequency or 48 pulses per second. Since this pulse is identified with the color sequence, it may be used to disable the brake circuit until the color disk is in proper color phase. Thus in Fig. 18-20 two color-identifying pulses are compared in V_4 . The first pulse is derived from the supersync wave by the tuned filter located between V_3 and V_4 . The second 48-cycle pulse is generated by a rotating contactor which is keyed to the disk shaft and which forms the pulse by keying an E, R, C combination in the screen grid circuit of V_4 . When the two pulses

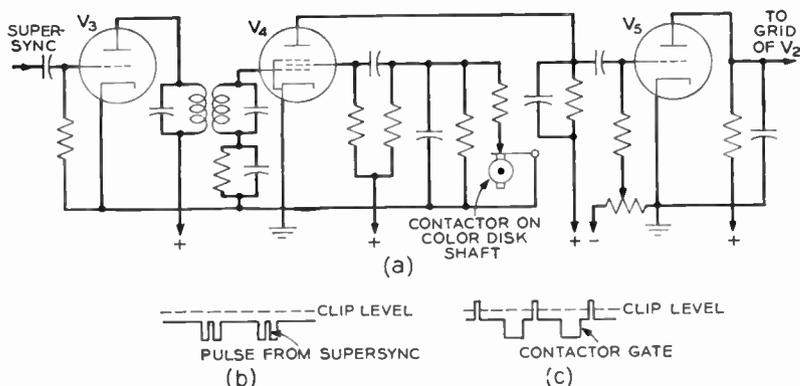


Fig. 18-20. The color disk may be phased automatically to a color identifying pulse in the supersync signal. (a) Basic form of the automatic color-phasing circuit. (b) Voltage on the plate of V_4 when the color-identifying pulse and contactor gate are in phase. (c) Voltage on the plate of V_4 when the pulse and gate are not in phase. (After Goldmark *et al.*)

are in phase as shown at *b* in the diagram no signal is applied to the brake-disabling circuit. If on the other hand, the color disk is out of phase, a pulse appears on the plate of V_2 and at the output of the clipper V_5 . This voltage is integrated and used to cut off V_2 of Fig. 18-19, and the required slipping action is obtained. Once the disk has slipped into phase, the condition shown at *b* obtains and normal braking action is maintained.

18-20. The Supersync Signal

There is nothing in the color-producing mechanism of the sequential color system which modifies the electronic sweep of the monochrome television channel, except for the frequencies of the horizontal and vertical scan components. Thus we may expect to find the same type of supersync signal that is used in commercial telecasting practice. We shall consider two modifications which may be used in the sequential color supersync wave form; first, the equalizing signals before and after the serrated vertical sync pulse may be omitted and second, some form of color-identifying pulse may be added after every third field. Omission of the equalizing pulses is justified on the grounds that the system operates perfectly well without them and, as we have seen, the color-identi-

cation pulse may be required for automatic phasing of the color disk. One general form of the supersync wave form that has been proposed by C.B.S. is shown in Fig. 18-21. It may be noticed that

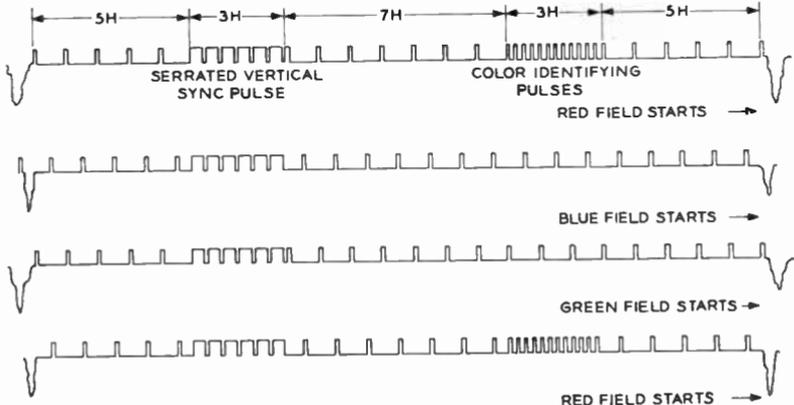


Fig. 18-21. Synchronizing wave form that may be used with the field-sequential color system. (Adapted from Goldmark *et al.*)

the color-identifying pulse is serrated in order that horizontal synchronization may be maintained for the duration of that pulse. The groups of color-identifying pulses occur at the end of every third field and hence have a repetition rate of 48 times per second.

18-21. Pickup and Receiving Equipment

The block diagram of the entire sequential color-television chain is illustrated in Fig. 18-22. The camera signal is mixed with the composite blanking and horizontal sync signals in the first amplifier. Since the three color signals are separated in time, each may be subject to a separate gain control. To accomplish this, the color mixer pulse generator furnishes three gating pulses, each of which coincides with one color. These pulses, in turn, gate in the proper gain control as each color field appears in sequence. After final mixing the composite video signal is fed to the visual transmitter. The receiver may be of the intercarrier or Dome type and conventional, except for its color disk-speed and phasing-control circuits.

Again, if an appraisal of the sequential color system is made, the

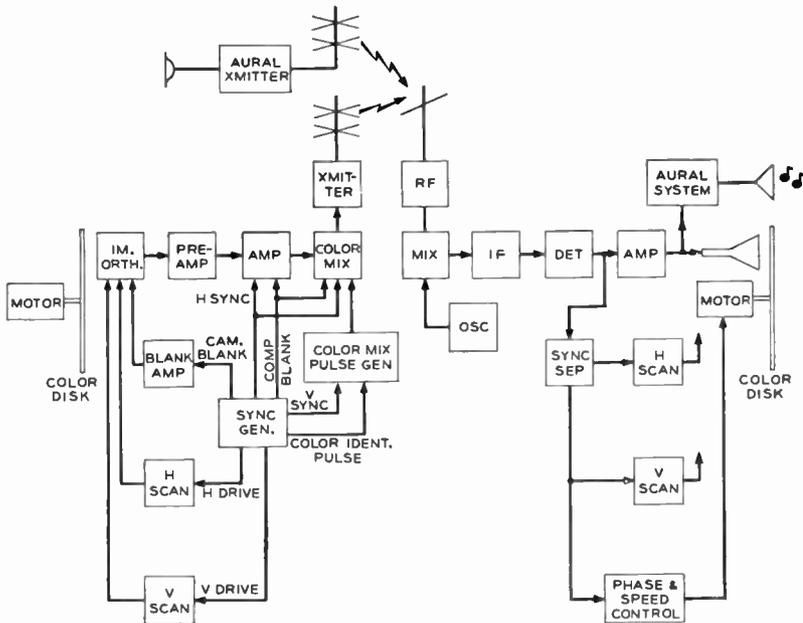


Fig. 18-22. Block diagram of the complete field-sequential color system.

following points are noted: The chief advantage of the system is that it works on live as well as film subjects and gives an excellent color picture. Its main disadvantage is that its adoption would tend to render obsolete present black-and-white receivers; since it is based on a new set of standards, no simple frequency converter would permit reception of the color signals for monochrome display on an existing black-and-white receiver; in short, it is a system which is not compatible with the commercial system already in use.

During the F.C.C. hearings in 1949 certain observers stated other disadvantages of the sequential color system. It was claimed that since the system operates on frame and field frequencies which are not integrally related to the power supply frequency, special precaution must be taken to minimize power supply hum effects at both ends of the system. At the transmitter end this was handled by operating all equipment from 144-cycle generators. It was fur-

ther pointed out that hum in the receiver upsets interlace by causing "line crawl," and that flicker caused by beating of the power supply and color sequence frequencies was objectionable. This last mentioned flicker can occur at rates as low as 12 per second and hence may be discernible at all normal levels of picture brightness. It should be stated, however, that not all observers of the demonstrations confirmed these objections about receiver performance. Still another serious disadvantage of the system is that it causes a 58% reduction in picture resolution as compared to that of the black-and-white system. This loss of picture information is the result of the lowering of the frame and line standards. It should be mentioned, however, that dot-interlace techniques, which are described in the last portion of the chapter and in the Appendix, have been utilized with the C.B.S. system and they raise the resolution to 83 per cent of that provided in black-and-white reception.

Since the C.B.S. sequential color system is not compatible with the black-and-white system, some attention has been directed to the design of adaptors or converters which will allow the color program to be received on a commercial-type monochrome receiver. Two separate problems are involved here. First, if the color program is to be reproduced in black and white only, a change-over switching arrangement is needed to change the horizontal and vertical scanning circuits from one set of standards to another. Second, if the program is to be reproduced in full color, the rotating color disk and drive motor are required.

It should be mentioned that the field-sequential color system is amenable to an all-electronic method of color reproduction at the receiving end of the system. For example, the single CRT and filter wheel may be replaced by a Trinoscope assembly or by a display similar to that of the C.T.I. system in which three images—one for each primary—appear in a vertical column on a single CRT. These three images may be covered with filters of the proper hue and merged by a lens system. While such changes eliminate the inherent objections to the color wheel assembly, they introduce new problems of image registration, and in the last mentioned proposal introduce some color parallax in the vertical direction. Even more important, the reproduction of the color image may also be effected with a tricolor kinescope of either of the types which have been described.

THE C.T.I. LINE-SEQUENTIAL COLOR SYSTEM²⁰

The third system of color television to be discussed is the line-sequential type proposed by Color Television, Inc. The basic equipment has been described and is illustrated in Fig. 18-11c. We shall confine our description, therefore, primarily to the method by which proper color interlace is produced.

18-22. Color Interlace

For the sake of complete compatibility, the C.T.I. system is based on the black-and-white standards of 525 lines and 60 fields per second. It will be remembered from our earlier discussions that three images, each corresponding to a single primary component, are focused side by side on the photocathode of an image orthicon tube. Scanning is arranged so that in one horizontal scan of the electron beam the three edge-to-edge images are covered. It should be noticed, then, that during one horizontal scan an output is produced for three lines in the image. In order to reduce confusion it is convenient to reserve the term "scan" for one sweep of the beam across the three images, and to use "line" in its usual sense of one sweep across a single image. It follows at once that in a 525-line picture, only $\frac{1}{3}(525) = 175$ scans are required. If, then, a standard supersync signal is used, the horizontal sweep generator in the camera equipment must be sensitive to every third horizontal sync pulse. Under this condition the horizontal "scan" frequency is $\frac{1}{3}(15,750) = 5250$ "scans" per second.

The interrelationship between lines and scans may be made clearer by reference to Fig. 18-23, which is drawn for a simplified 21-line system. The roman numerals refer to scans, and the arabic numbers to lines. At the extreme right in the figure the three images are superposed to show how interlacing occurs. Zero flyback time is assumed in both directions to simplify the drawing. Study of the drawing reveals an inherent difficulty in the scanning pattern just discussed. Since the total number of lines (21 in the example and 525 in the actual system) is divisible by three (the number of images scanned), no color interlace occurs. That is, each line of elements

²⁰ See, for example, J. H. Battison, *op. cit.*, or "New Directions in Color Television." *Electronics*, 22, 12 (December 1949).

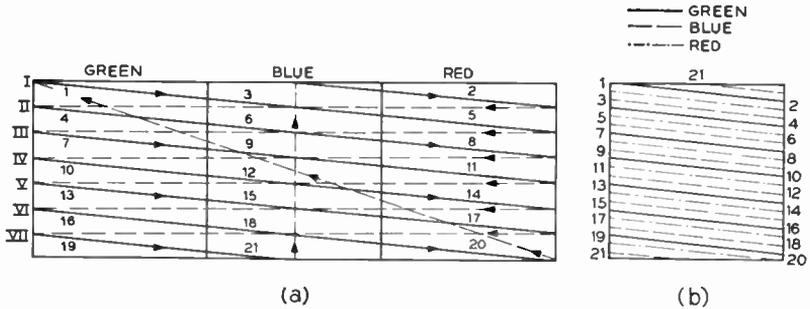


Fig. 18-23. The basic scanning pattern for the C.T.I. color system. Three side-by-side images are scanned by a single raster at one-third of the standard line frequency. In the actual system some space is allowed between adjacent images in order to leave time for horizontal flyback in black-and-white reception.

in the complete picture is always scanned in the same color. This situation is bad, because if a certain region contains only one primary, that primary would only be produced along every third line, resulting in a very coarse reproduction of the image detail in that region. In effect two-thirds of the picture information in that region is wasted.

To overcome this difficulty, C.T.I. has employed a system of commutation known as "interlaced color shift" that allows each line in the complete image to be scanned in all three primary colors.²¹ Since, as we have seen, the total number of scans and lines is divisible by three, such a commutation of colors can only be accomplished by upsetting the raster illustrated in Fig. 18-23.

In the actual system the pattern is upset in such a way that the *odd* lines only are scanned on the first three successive fields, and in the next three fields only the *even* lines are scanned. It follows, therefore, that six fields are required to scan the entire picture in all three primaries; hence, under the 60-field-per-second standard, 10 complete color pictures are transmitted each second. Table 18-1 shows the sequence of scanning and color commutation for the first four lines of the six successive fields that are required to scan one complete color frame.²² It will be observed that the usual system

²¹ Demonstrated before the F.C.C. on May 17, 1950.

²² "The Present Status of Color Television." Report of the Advisory Committee on Color Television to the Committee on Interstate and Foreign Commerce, United States Senate. A reprint of this report is published in the *Proc. IRE*, 38, 9 (September 1950).

TABLE 18-1

| <i>Field</i> | <i>Line</i> | <i>Primary</i> | <i>Field</i> | <i>Line</i> | <i>Primary</i> |
|--------------|-------------|----------------|--------------|-------------|----------------|
| | 1 | Green | | 2 | Green |
| | 3 | Blue | | 4 | Blue |
| 1 | 5 | Red | 4 | 6 | Red |
| | 7 | Green | | 8 | Green |
| | 1 | Red | | 2 | Blue |
| | 3 | Green | | 4 | Red |
| 2 | 5 | Blue | 5 | 6 | Green |
| | 7 | Red | | 8 | Blue |
| | 1 | Blue | | 2 | Red |
| | 3 | Red | | 4 | Green |
| 3 | 5 | Green | 6 | 6 | Blue |
| | 7 | Blue | | 8 | Red |

of interlace between the odd and even lines in successive fields is changed, the interlaced color shift being used in its place.

In practice the color commutation is produced by introducing a color-identifying slot in every third horizontal sync pulse during a given field interval. During the vertical blanking interval that occurs at the end of a field, the position of the slot is displaced one or two line intervals as required and remains in the new position during the next field. Still another shift occurs in the next vertical blanking interval, and so on. In the camera and color receiver the identifying notch activates a control circuit that causes the beam position to shift to a particular color image, and the required interlaced color shift takes place. More specifically, say that the color identifying notch is used to identify the green image. The notch is used to synchronize the horizontal scan that will run at one-third line frequency. Since the notch position is shifted at the end of each field, the scan begins at different points of the *vertical* sweep during successive fields. Thus the scan of the green field begins at different heights along the vertical edge of the images, and the proper color interlace is obtained.

A similar arrangement is used in the color receiver. With horizontal synchronization provided by the color-identifying notch, the scan runs at one-third normal line frequency and three edge-to-edge images are displayed on a CRT. Color is synthesized by using a filter of different hue over each image, or by utilizing a different color phosphor under each image, and by superimposing the three resulting primary images optically.

The chief disadvantages of the C.T.I. system are that it exhibits line crawl and interline flicker; it transmits only 10 color pictures

per second; and it is sensitive to electrical and optical registration problems.

The principal advantages of the C.T.I. proposal are that it is all-electronic, requiring no mechanical switching of color filters, and it is completely compatible. Since the conventional black-and-white receiver has no means of recognizing the color-identifying notches, its horizontal sweep circuits are actuated by the conventional 15,750-per-second horizontal sync pulses in the supersync, and the image is reproduced in black and white at 525 lines, 60 fields per second.

In concluding the discussion of the C.T.I. system it is interesting to point out how that system overcomes the loss of storage time which would occur if a mechanically shifted filter such as that used by C.B.S., were employed to commutate colors after each line.²³ In the present case each color image is confined to a specific region on the camera photocathode; hence any section of that photocathode may store charge for an entire frame interval less one line interval. This is true because the color information is produced simultaneously (three edge-to-edge images) but is then sampled, color by color, at line frequency by the scanning system.

THE R.C.A. DOT-SEQUENTIAL COLOR SYSTEM²⁴

The next color-transmission system to be discussed is that which uses the dot-sequential method of color sampling and color reproduction. Proposed and demonstrated to the Federal Communications Commission in 1949 by R.C.A., it is fully compatible, and further compresses the entire three-color video signal and its associated sound program into a 6-megacycle channel without change in the electronic scanning standards of the standard monochrome method of transmission. This remarkable fact is achieved by utilizing the principle of mixed highs and by combining color-sampling and time-multiplexing techniques.

²³ See section 18-17.

²⁴ A Six-megacycle Compatible High-definition Color Television System, an exhibit presented to the Federal Communications Commission by the Radio Corporation of America. See also, M. S. Kay, "Color Television?" *Radio and Television News*, 42, 6 (December 1949); or J. H. Battison, *op. cit.*

18-23. The Sampling Process

The basic equipment required at the transmitting end of the dot-sequential system is shown in Fig. 18-24. It will be observed that a simultaneous type of three-camera pickup system is used because three independent color channels, one for each of the primary hues,

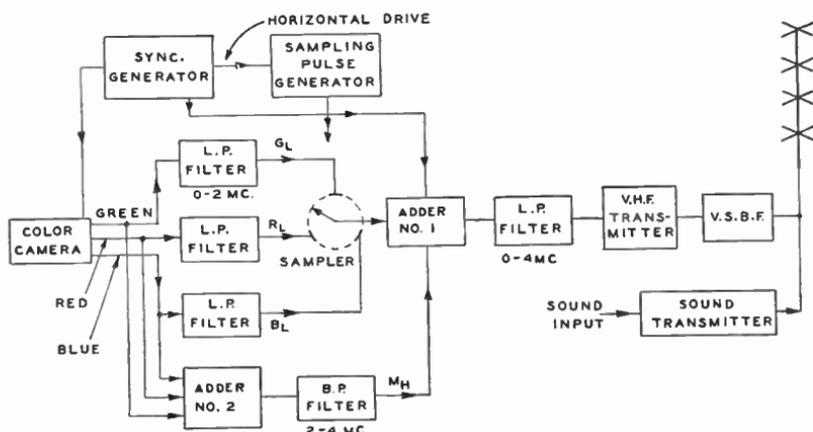


Fig. 18-24. Block diagram of the dot-sequential color transmitter.
(Courtesy of Radio Corporation of America.)

are present. The left-hand part of the diagram uses components whose functions we have already discussed. Filters are used for separating the low- and high-frequency components of each channel, 2 megacycles being the boundary frequency in each case, and the low-frequency components are sampled at a rate of 11,400,000 times per second (each color at 3,800,000 times per second) by a switch or electronic sampler. This much about the system we already know. We shall now consider the sampling process in greater detail.

The sampler is keyed by signals from the sampling-pulse generator, which, in turn, is controlled by the master sync generator. This method of interlock between the sampler and the source of the synchronizing signals ensures that the color samples are taken at the proper intervals in the scanning raster and that proper interlace of the color dots is maintained. Since the sampler and its associated components yield a compression in bandwidth, let us consider their action more closely.

Consider what happens to the low (0- to 2-megacycle) green component. Since this is sampled at a 3.8-megacycle rate, it will appear in the output of the sampler as a series of voltage pulses spaced at intervals of $1/3.8$ megacycle or $2.63 \mu\text{sec}$. The height of each pulse will be proportional to the amplitude of the green signal being scanned at the instant of sampling. Two such samples are shown in Fig. 18-24*a*. Since both are of the same height we may assume that they represent an area in the complete image which has a constant green component.

It may be shown from Fourier's theorem that these pulses consist of a d-c component plus several a-c components of frequencies which are multiples of 3.8 megacycles, the basic repetition rate. Thus, if the pulses are passed through the 0- to 4-megacycle filter of Fig. 18-24, only the d-c and fundamental components will pass through; hence the filter output (as far as the green component is concerned) will consist of a 3.8-megacycle component superimposed on a d-c component, a combination which is shown at the right in Fig. 18-25. It is important to stress that the amplitudes of the d-c and 3.8-megacycle components are proportional to the amplitudes of the green pulses from which they are derived.

The same description also applies to the red and blue channels. In each case the low-pass filter converts the sample pulses into a sine wave superimposed on a d-c component, the amplitudes of each being proportional to the amplitudes of the pulses which produce them. We must notice, however, that the color sample pulses occur in a time sequence red, blue, green, and so on and that any two adjacent pulses are displaced by $\frac{1}{3}(2.63)$, or 0.0877 , μsec . This time relationship is shown at *a*, *c*, and *e* in Fig. 18-25 for the sample pulses, and for the d-c plus sine waves at *b*, *d*, and *e*. Notice particularly that when any one of the sine waves is at a maximum, the two other waves are passing through zero. This property of the signals is basic in reconstituting the image in full color at the receiver.

Our discussion thus far has been on a color-by-color basis when actually all three signals are combined in the adder which follows the sampler. Since the circuits are linear, we can superimpose all three signals; for example, the actual sampler output consists of the samples from all three channels as shown in Fig. 18-25*g*, and since the sum of three sine waves of the same frequency is also a sine wave of the same frequency, the output of the 0- to 4-megacycle filter

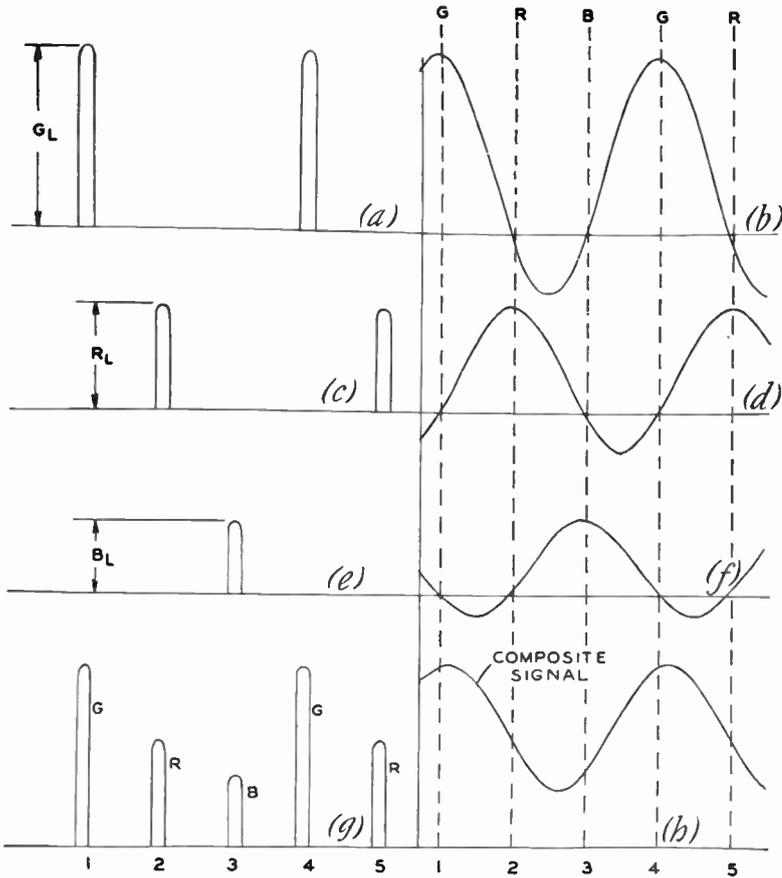


Fig. 18-25. Color sample pulses and the sine plus d-c waves that result after filtering. Notice that when any one component is at a maximum the two other components are zero. (Courtesy of Radio Corporation of America.)

consists of a single 3.8-megacycle sine wave superimposed on the sum of the three d-c components previously mentioned. This composite signal which contains information from all three color channels is shown at *h*. Observe that the original color samples may be recovered from this composite wave if it, in turn, is sampled at the intervals labeled 1 through 5 in the diagram.

Let us now move backward toward the scanning system to see

where the color samples occur in the raster. The actual spatial relationship is shown in Fig. 18-26. Notice that the primary dots (in space the dots correspond to the samples in time) occur in an orderly pattern, red, blue, green, along each line of the raster. Fur-

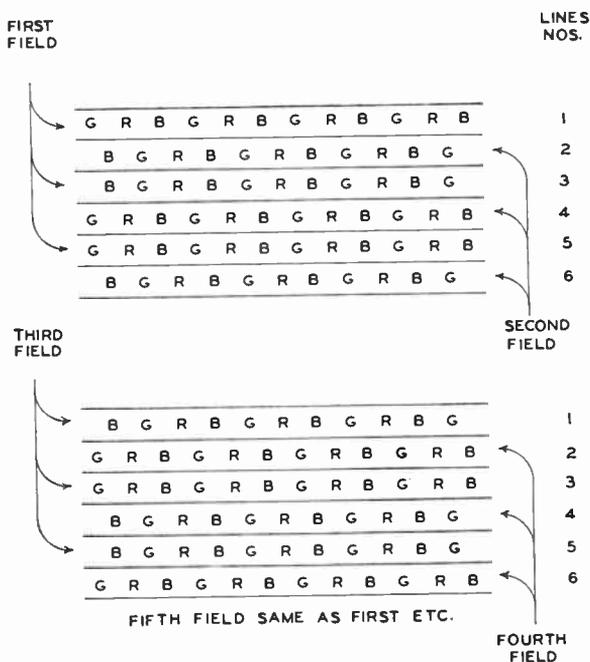


Fig. 18-26. Dot positions in the scanning raster on consecutive fields and frames. Any given dot-element is always scanned or reproduced in the same primary hue. (Courtesy of Radio Corporation of America.)

thermore adjacent dots are not contiguous, but there are holes or blank spaces between them, and in adjacent lines the positions of dots and holes are interchanged. This may be seen from the upper diagram in Fig. 18-26, which shows a portion of the raster after the first two fields have been scanned. Thus, in one frame only one-half of the entire area has been sampled; no samples have been taken corresponding to the holes. This is important. Only one-half of

the picture detail is sampled in one frame interval, even though the entire picture area has been scanned.²⁵

The remaining half of the detail in the picture is sampled in the next two fields as may be seen from the lower portion of the figure. The positions of the dots and blanks of fields 1 and 2 have interchanged in fields 3 and 4. We may now summarize the discussion to show how the sampling and multiplexing processes yield a reduction in bandwidth, so that all three colors may be transmitted in the 4-megacycle video bandwidth of the black-and-white system without loss of detail: the use of dots and blanks means that all the detail along a line is replaced by the detail in the dot spaces; thus maintaining the same bandwidth, only one-half the amount of detail is sent in one frame interval. Four, rather than two, fields are required to send the detail in the entire picture; yet the final picture contains all the detail.

Careful inspection of Fig. 18-26 shows that any one dot area is sampled for only one primary hue; hence we cannot say that every picture element is sampled for its three primary components as it is in the field-sequential system. It may be argued on the other hand that the space occupied by three color dots is so small that satisfactory color analysis is obtained.

18-24. The Receiver

The block diagram of a receiver for the R.C.A. dot-sequential color signal is shown in Fig. 18-27. The composite video signal which is fed to the sampler has the form shown in Fig. 18-25*h*. Then since the sampler is timed to take samples for the green channel at points 1 and 4, the output fed to the green channel will consist of pulses which are proportional to the original green samples at the transmitter. This is true because, as we have seen, at these intervals the composite signal amplitude is equal to that of the green pulses; the red and blue waves are passing through zero.

Following through the receiver block diagram we see that the green pulses are then filtered to produce a 3.8-megacycle sine wave

²⁵ Remember that the simultaneous type of camera delivers three channels which have all the picture detail. The rejection of one-half the detail results at the sampler. Since samples are related to the dots, it is convenient to refer everything to the dot positions in the raster.

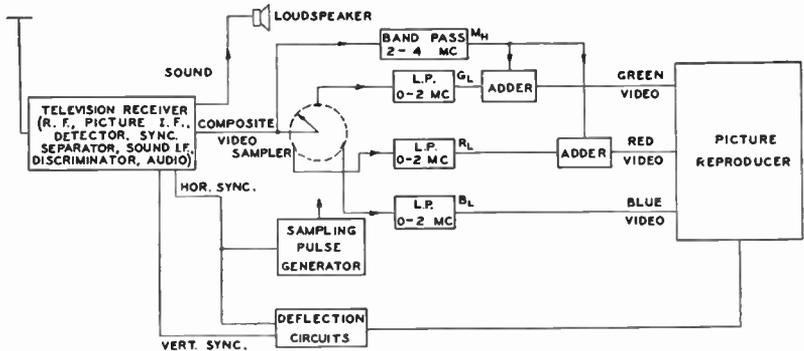


Fig. 18-27. Block diagram of the dot-sequential color receiver.
(Courtesy of Radio Corporation of America.)

superimposed on d-c (Fig. 18-28*b*) which are then fed to a CRT. The latter is driven below cutoff by negative portions of this wave, and hence, if the light-voltage characteristic of the tube is linear, a plot of green tube light output *v.* time is that shown in Fig. 18-27*c*. The dotted curves show the output for the even fields. Notice that the reproduced dot is wider than the sample. This spreading effect produces better dot interlace. The same description holds for the red and blue components. The light *v.* time curves for all three color components are shown at *c* in Fig. 18-28.

As far as reproducing assemblies are concerned, either the Trinoscope or the three-CRT dichroic mirror assemblies of Fig. 18-12 may be used. Furthermore, either of the tricolor kinescopes may be used because any one spot in the raster is always scanned and so must always be reproduced in a single color. It should be apparent that the position of the raster relative to the three-color dot groups must be maintained accurately. Otherwise, if a red dot in the original image is reproduced by a green or blue dot in the kinescope because of raster misalignment, an interchange of color primaries will result and cause improper color reproduction in the final image.

In the foregoing discussion we have neglected the effects of the mixed-high signal components in the interests of simplicity. Actually their effect on the wave forms is extremely difficult to show. They are added to the sine wave primary components and, at the receiver, are reproduced along with the d-c and 3.8-megacycle primary component in each channel. Our discussion has also been

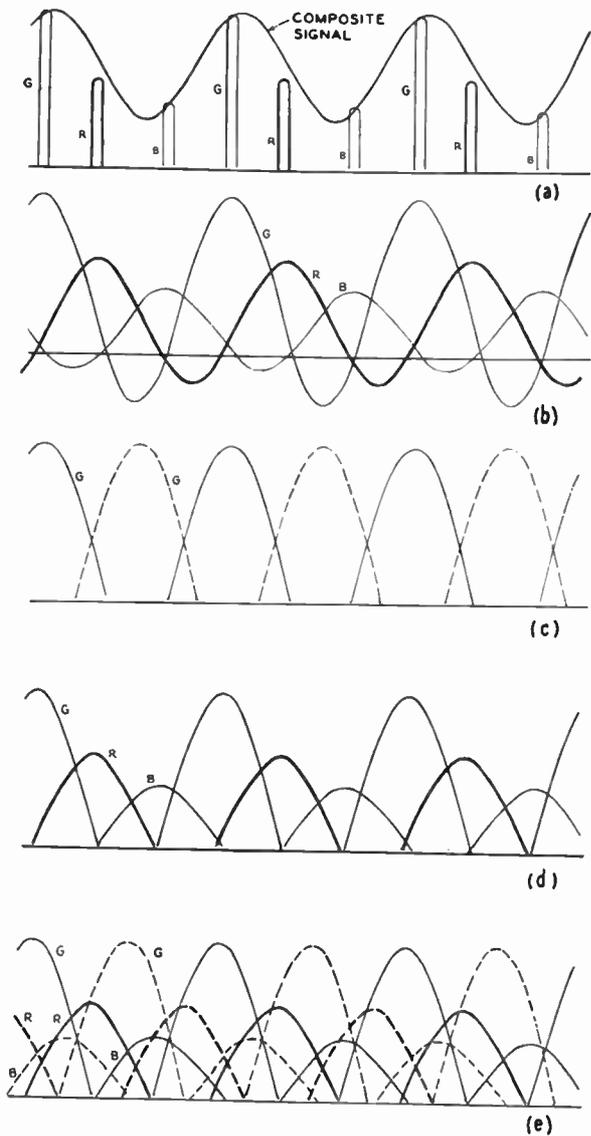


Fig. 18-28. The sampling and reproduction of primary components at the receiver. Dashed and solid lines indicate even and odd fields. (a) The composite signals and resulting color samples. (b) 3.8-megacycle plus d-c waves are derived from the sample pulses by filtering. (c) Light output of the green component along a line. (d) Light output of all three components along a line. (e) Light output of all three components along a line, for odd and even fields. (Courtesy of Radio Corporation of America.)

are zero; hence sampling at the proper intervals still produces pulses of the proper amplitude for each primary. Even though each component wave no longer has a sinusoidal shape, the frequency of the fundamental components, being determined by the sampling rate, remains at 3.8 megacycles.

18-25. Compatibility

Let us now check the compatibility of the dot-interlace color system. We may do this by considering what effect the composite signal of Fig. 18-25*h* has in a conventional black-and-white receiver. This wave, in conjunction with the mixed-high signal, contains all the detail required for the picture plus an additional 3.8-megacycle component, which is the result of the sampling process. The conventional receiver has no sampler or low-pass smoothing filters; hence all the detail is reproduced in monochrome on the CRT. Normally the amplitude of the 3.8-megacycle sine wave is too small to cause any trouble. Its amplitude is a function of the color value in the original image, however; hence in high-value regions of the image it will appear as a series of monochrome dots. It is claimed that these cannot be resolved at normal viewing distances and hence are not objectionable. As a matter of fact, in a white region of the picture, where the presence of dots would be most discernible, they are not present at all. This is true because all three primary components are equal in a white region; hence only a d-c and no 3.8-megacycle component is present.

Actually the compatibility of the R.C.A. system goes one step further: the dot-interlace color receiver is able to handle and reproduce a standard monochrome signal. Since the incoming signal contains all the picture detail, it is reconstituted completely, one-third of the detail being furnished by each of the three CRT's in the reproducing assembly.

Some observers of demonstrations of the dot-interlace color system have claimed that the dots in the reproduced image may not blend properly at normal viewing distances, particularly where a large area of a single primary, such as blue, is being reproduced. In such a region only one-third of the elements are active, the red and green components being zero. Further trouble might result from the beat note between the 7.6-megacycle component, resulting from sampling a single color, and a 4.5-megacycle component which

may be present because of a beat between the sound and picture carriers. It may be assumed that claims of misalignment problems in any of the three-CRT reproducing systems are eliminated by use of the tricolor kinescope.

THE C.T.I. SEGMENTAL-SEQUENTIAL SYSTEM¹

Thus far we have considered a simultaneous color system in which all three primaries are transmitted at the same time, and three systems that sample the primary information at different rates: at field frequency, at line frequency, and at a high multiple of line frequency. In the latter two systems the high color-switching rate required the use of three optical images that had to be maintained in proper register to a high degree of accuracy. Still another system has been devised, utilizing a single optical image and employing a sampling rate lying between those of the line- and dot-sequential systems. This segmental-sequential system, which samples segments of a line in the different primary colors, was described to the F.C.C. by Color Television, Inc., in August 1950, when it petitioned for a reopening of the hearings on color television. In very broad terms this proposal may be compared to the dot-sequential system in that color is sampled several times along a single scanning line; but the sampling rate is lower and is accomplished by an auxiliary optical system at the camera tube, rather than electronically. The chief advantages claimed for the system are (1) it eliminates the problem of registering three optical images, (2) it permits the use of a color receiver of relatively simple design, and (3) it is fully compatible. This segmental-sequential system will now be described.

18-26. Principle of Operation

As a starting point in considering the operation of the segmental-sequential system, let us assume that an optical image of the televised scene is focused onto the photocathode of an image orthicon tube through a color filter that consists of narrow vertical strips of the three primary colors, red, blue, and green. In the proposed system the switching from color to color is carried out at a rate equal to the 192nd harmonic of the line-scanning frequency of 15,750 per second. Thus, since an active line has a duration of approximately 53.5 μ sec, the number of filter strips required in the filter is

¹The material used on the segmental-sequential color system is used with permission of Color Television, Inc.

$$192 \left(\frac{53.5}{63.5} \right) \approx 162$$

Reference to Fig. 18-30 will show how this filter strip, in conjunction with the scanning system, operates to give a segmental-sequential output signal from the camera tube. The assumed scene corresponding to some particular line in the image is shown at *a*. It should be noticed that the transitions between regions of different color are chosen to fall in different positions relative to the filter-strip locations in order to illustrate a number of different conditions. Consider what happens in the second region corresponding to white. Since white consists of the three primaries, all three filter segments in that region allow the photocathode to be illuminated, and an output signal is developed over the entire region.

By way of contrast consider what happens in the fourth region corresponding to blue. Here only those portions of the photocathode covered by a blue filter strip are illuminated, and those covered by red or green strips remain dark; hence an output signal is developed only during the duration of each blue strip. A similar situation holds for the remaining regions, and it may be seen that the entire line is broken up into a number of color segments, each corresponding to a single primary color.

It should immediately be apparent that, unless the position of the vertical filter strips is changed relative to the scene information before the next scan of the same line in field III, each segment will always be sampled in the same color. To eliminate this condition that would, for example, reproduce the white of region 2 as three adjacent segments of blue, green, and red, the filter is displaced laterally one segment- or strip-width before the line is scanned in field III. The resulting filter-strip location and output signal are shown at *c*. Still another filter shift occurs before the line is scanned in field V, with the result shown at *d*. By an extension of this discussion it may be seen that the entire picture is scanned in segments of all three primaries in three frame intervals, thus complete color pictures are transmitted at the rate of 10 per second.

It may be surmised from the foregoing discussion that successful operation of the system so that every line segment is scanned in all three colors depends upon some satisfactory means of effectively shifting the color filter strips laterally between scans of any line in successive fields. This is accomplished by reproducing the strips on

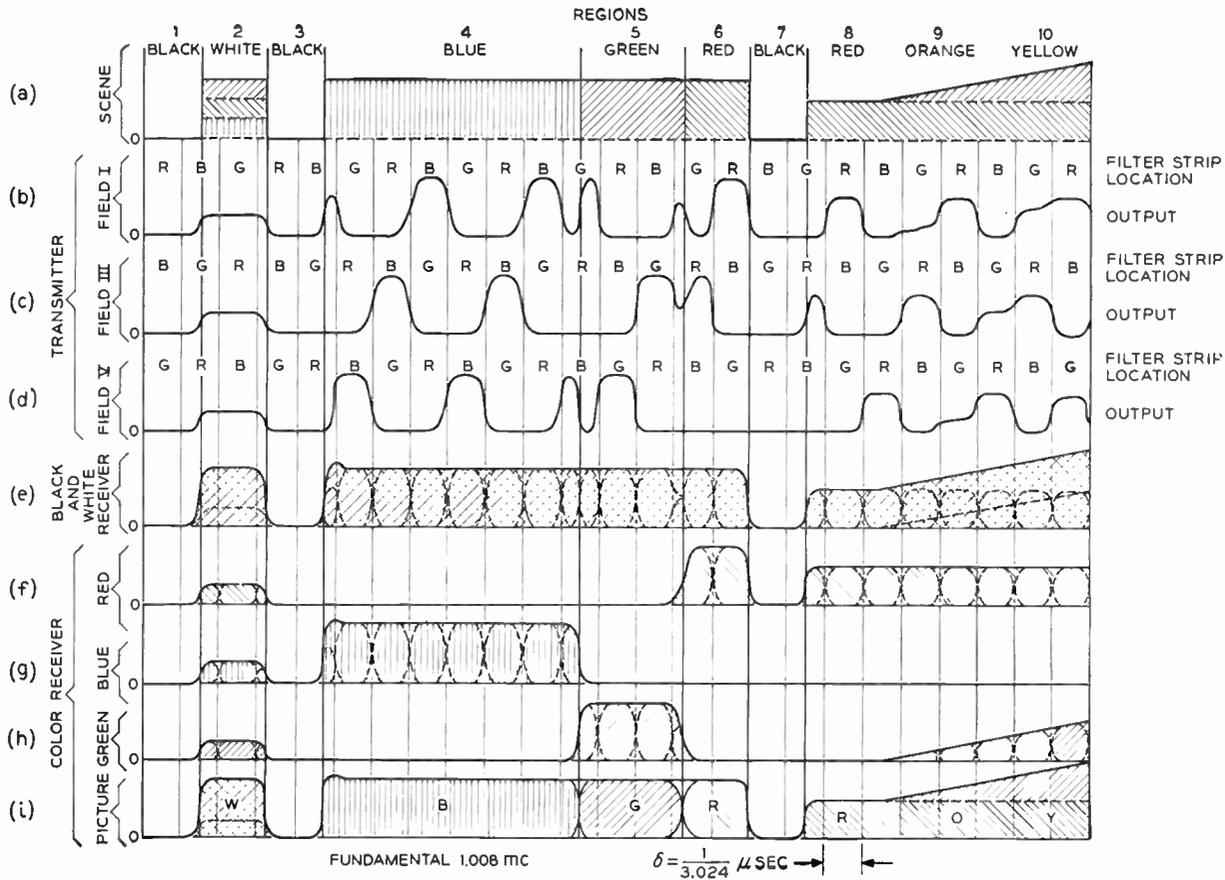


FIG. 18-30. Typical wave forms for the segmental-sequential color system. (Courtesy of Color Television, Inc.)

a loop of color motion picture film, such as Kodachrome, which is rotated in front of the camera tube by a selsyn running at proper speed.

Another problem arises from the use of the strip filter. A means must be provided to ensure that a segment of a given color, for example red, is being scanned by the camera tube at the same instant that red is being reproduced by a receiver under control of the color synchronizing signal. Precise alignment of the scanning over the entire raster on the camera photocathode is impractical because of field distortions and irreducible instabilities and nonlinearity of the scanning process; hence a means must be provided for a continuous adjustment of the scan relative to the filter. To this end a special tracking grid is located beside the color filter on the loop of Kodachrome film. This grid consists of color strips arranged in a special sequence, its image being focused onto the filter by an auxiliary lens system. This supplementary image provides a signal component in the video circuit of the camera, which component is used to compensate for beam-position deviations through a control circuit. The tracking signal is removed from the video signal to be transmitted. As the film loop rotates, the tracking and filter sections move together to shift the segment positions in successive fields.

18-27. Color Synchronization

Reproduction of the color image at the receiver in the segmental-sequential system is based on the use of a tricolor kinescope of either type described in Section 18-11. As a consequence, some form of color synchronizing signal must be transmitted along with the composite video signal so that proper gating of the kinescope guns may be accomplished. Since C.T.I. has designed the system to be fully compatible, it was decided that no changes were to be made in the standard monochrome supersync signal; hence some new means for sending the color synchronizing information had to be devised. The final method used may be understood from a more careful consideration of the camera output pulses that are shown in Fig. 18-30. As has been stated previously, the individual color segments are scanned at the 192nd harmonic of the line frequency. Thus, if we assume the output pulse from a segment to be square, it will have a width δ given by

$$\delta = \frac{1}{192(15,750)} = \frac{1}{3.024} \mu\text{sec}$$

Since the filter strips always occur in groups of three, the maximum repetition rate of the output pulses of any one primary color will be $(\frac{1}{3})(3.024) = 1.008$ mc, which corresponds to a period T of

$$T = \frac{1}{1.008} \mu\text{sec}$$

It may be seen, then, that the output for any one primary color may consist of a train of square pulses of width equal to one-third of the period T . Reference to Section 11-17 will show that when such a pulse train is analyzed into its Fourier components the following harmonics will be present: first, second, fourth, fifth, seventh . . . , but all the integral multiples of the third harmonic such as the third and the sixth will be absent. Furthermore, since the bandwidth of the system cuts off just above 4 mc, all components above the fourth harmonic will be eliminated. Thus the only components delivered to the communication channel will be 1.008, 2.016, and 4.032 mc. This bandwidth restriction accounts for the rounding of the pulse edges shown in the diagram. Notice also that the third harmonic of the color-scanning frequency is completely eliminated. As a result of this fact a color synchronizing signal of 3.024 mc may be introduced into the composite video signal. To ensure that this color sync signal will not interfere with the picture, it is introduced in opposite phase in successive scans of any line so that it cancels out in the reproduced image.

The 3.024-mc color-phasing wave rides on top of the vertical sync pulses and on alternate equalizing pulses in the supersync. In addition, bursts of a 1.008-mc component are transmitted during the last two or three lines of the vertical blanking interval to provide color phasing. The use of these waves in the color receiver is described in the next section.

It may be seen from the foregoing discussion that the color scanning rates are integrally related to the line-scanning frequency. Consequently, rigid lock-in with the sync signal generator is required and this lock-in may be accomplished by methods which we have discussed earlier. In the color sync generator a reactance-tube-controlled master oscillator is run at 378 kc, the 24th harmonic of the line frequency. This is stepped down by a 12 to 1 divider to 31.5 kc,

which may be recognized to be the master frequency of a standard sync generator. A 60-cycle component is derived from this by frequency dividers and is compared to power-line frequency. The resulting error signal is used to correct the master oscillator. The color sync signal of 3.024 mc is derived from another master oscillator under control of a reactance tube. Its frequency is corrected by deriving a 378-ke component from an 8 to 1 divider, which is compared with the first master oscillator output to develop an error signal, which in turn feeds the reactance tube. By this means, scanning and color sync frequencies are all maintained properly. The 3.024-mc frequency is also stepped down through a 3 to 1 counter to give the required 1.008-mc component. The required field-to-field phase shift of the color sync signal is obtained by tapping the wave from different sections of a delay line.

18-28. The Color Receiver

Once it has been established that the color image is to be reproduced on a tricolor kinescope of, say, the three-gun type, the operation of the receiver of the segmental-sequential system is easy to understand in general terms. As described in Section 18-11 the video signal is fed to a control grid common to all three guns in the kinescope. Separation of colors is accomplished by applying gating signals, in this case of $(1/3.024) \mu\text{sec}$ duration, to each of the cathodes in sequence; hence a gating pulse generator is required. In the proposed system this may be a relaxation oscillator, either a multivibrator or blocking oscillator, whose output is fed directly to one of the cathodes, and to a two-section delay line that gives a $(1/3.024) \mu\text{sec}$ delay per section. The output of each section feeds one of the other guns in the kinescope. By this simple means the guns are gated in proper sequence and for the proper intervals.

Since the gating must be phased properly with respect to the video signals, the frequency and phase of the relaxation oscillator are controlled from the 3.024-mc wave that is transmitted on top of the vertical sync pulses. The manner in which this control is effected will now be described. The sync component is derived from the composite video signal by a narrow band-pass filter. In order to render the control system immune to the field-to-field changes in sense of this wave that has been mentioned previously, the filtered color sync wave is fed through a frequency doubler. In the interests of simple

design and low cost this may take the form of an unfiltered full-wave rectifier using crystal diodes. The output double-frequency wave is then fed to a discriminator whose d-c output is used to control the exact frequency of the relaxation oscillator.

With only these components over and above those of a black-and-white receiver, the image may be reproduced in full color. Manual phasing of the color is required, unless an additional narrow band-pass filter peaked at 1.008 mc plus an injection circuit to control the initial phase of the relaxation oscillator are added. It may be inferred from this description that the cost differential between a segmental-sequential color receiver and its conventional monochrome counterpart will be quite low.

The manner in which the several color samples are added in the tricolor kinescope may be seen at *f* through *i* in Fig. 18-30. It will be observed that certain of the added signals exhibit a small amount of overshoot just after the leading edge. It is claimed that this transient is no larger than those frequently encountered in the commercial black-and-white system. The rounding of the leading edges is due to the approximately 4-mc bandwidth of the transmission system.

18-29. Compatibility

Since all scanning frequencies and the supersync wave form of the segmental-sequential system are identical to those of the commercial monochrome system, a conventional receiver can handle the incoming signals and will reproduce them in shades of gray. The manner in which the color samples add in this case is shown at *e* in Fig. 18-30. Since the receiver has no way of recognizing the color sync and phasing pulses, they have little or no effect on the operation of the receiver.

Inasmuch as the proposed segmental-sequential color system has not been subjected to the extensive tests of the systems that have been described previously, it was not possible, at the time of writing, to give specific data on its performance.

APPENDIX

DOT SYSTEMS OF TELEVISION TRANSMISSION^{1,2}

In Chapter 1 it was explained that certain limitations imposed by an electrical communication channel connecting the two ends of a picture transmission system required that the transmitted image be scanned and transmitted element by element. This was necessary because the two-dimensional image had to be forced into an $e(t)$ form which could be handled by the electrical system. Break-down of the picture into small elements was accomplished by causing some form of aperture to scan over the image.

In electronic scanning the effective aperture, which is furnished by the electron beam at its point of impact in the camera tube, moves continuously from element to element at such a high rate of speed that the resulting output voltage is a continuous signal. We might think of the process in this manner: At any instant, the output from the camera tube is a sample of the image at a point as shown in Fig. 1-3*b*, but since the aperture moves continuously, rather than in a step-wise manner, the several samples are merged together so that a continuous signal is produced. In the present section we wish to turn our attention to these time samples again to see if they, rather than the continuous signal, may be transmitted as the video information. If such a system is feasible, it may be termed a *dot system* of transmission. This notation arises because the amplitude of each of the time samples will be proportional to the brightness of one picture element or dot. In essence, a dot transmission system is one that sends discrete pulse samples of picture information rather than a continuous signal of picture information. Our interest in the dot transmission concept is due to the saving in bandwidth which

¹ W. Boothroyd, "Dot Systems of Color Television," Part I. *Electronics*, **22**, 12 (December 1949).

² E. M. Deloraine, "Pulse Modulation." *Proc. IRE*, **37**, 6 (June 1949).

it may afford or, conversely, in its ability to provide greater picture resolution in a given bandwidth.

A-1. The Sampling Process

From our previous studies it should be apparent that the output voltage from the camera tube is a continuous signal; hence if pulse samples are to be transmitted over the electrical system, some form of sampling device must be placed between the camera tube and the input to the communication channel. Since we are concerned here with the principles of operation of the dot system, we need not be concerned with the details of the sampler. It is sufficient for our purpose to think of it as a simple mechanical, single-pole, single-throw switch which is closed and opened successively for definite intervals. Each time it is closed, a sample of the camera output will be delivered to the remaining portions of the system.

It is an easy matter to see that each time-sample will correspond to a specific picture element or dot. The electron beam in the camera tube still scans over a conventional raster, of either the progressive or interlaced form. Then, since the vertical and horizontal velocity components of the beam are constant, any instant of time corresponds to a particular position of the scanning beam. There is a one-to-one correspondence between a time sample of the output signal and a dot position in the scanned image.

In the remaining portion of this section we shall consider two questions about the sampling process: (1) How will the pulse-sample reproduced image appear to the eye? (2) What shall be the sampling rate? We shall see that the two questions are related to each other.

In answer to the first question we note that if discrete samples of voltage spaced in time are applied to the grid of a CRT the reproduced image will consist of an array of shaded dots separated by blank spaces, that is, the image will have the general appearance of a half-tone reproduction of a photograph. For such a picture to appear satisfactory, it is only necessary that enough dots be present so that the eye cannot resolve them at normal viewing distance. If this condition is met, the blank spaces between the dots need not be filled in.

As an illustration, let us assume that the sampling rate is 8 mc. Then under commercial standards where the horizontal unblank interval is roughly $55 \mu\text{sec}$ the number of elements along a line will be

$$N_h = (8 \times 10^6)(55 \times 10^{-6}) = 440 \text{ dots/line}$$

Then, with approximately 500 active lines, the total number of dots per picture, or the figure of merit, will be

$$M = 500(440) = 220,000$$

which may be considered to be a satisfactory value.

The second question regarding the required sampling rate may be answered subjectively from the results which were just stated. A sampling rate of 8 mc may be used because it yields a satisfactory picture. A more precise answer may be found with the help of a theorem: If a function $f(t)$ contains no frequencies higher than F cps, it is completely determined by giving its ordinates at a series of points spaced $1/2F$ seconds apart.³ This means that if the sampling rate is at a frequency $2F$, the samples will be able to reconstitute any frequency component in the sampled signal up to and including F . It follows at once, then, that the 8-mc rate of our previous example will allow reproduction of all video components up to 4 mc, the top video frequency under commercial standards. We shall, therefore, consider the 8-mc sampling rate to be satisfactory.

A-2. Pulse Transmission Bandwidth

We must now consider the feasibility of sending the sample pulses over the communication channel. Actually a considerable problem is present here for, as we have previously seen,⁴ an extremely large bandwidth is required for pulse transmission. For example if we are to transmit all components of the pulses up to the fifth zero, the bandwidth must be $5/\delta$, δ being the width of the individual pulses. If we continue the previous example and assume δ to be one-half of the sampling period, the required bandwidth will be

$$\Delta f = \frac{5}{\delta} = 5(2)(8 \times 10^6) = 80 \text{ mc}$$

Clearly this is an intolerable situation: A bandwidth of 80 mc is required to transmit a video signal whose highest frequency component is 4 mc. A mathematical analysis of the sample pulses shows, however, that all of the necessary video information may be

³ See for example: C. E. Shannon, "Communication in the Presence of Noise," *Proc. IRE*, 37, 1 (January 1949).

⁴ See Section 11-17.

transmitted over a bandwidth much narrower than our figures indicate. Let us carry through this analysis.

To simplify the mathematics we shall consider the original camera output to consist of a single a-c component superimposed on a d-c term, that is

$$e_c(t) = E(1 + m \cos \omega t) \quad (\text{A-1})$$

We shall assume further that the sampling switch operates at a frequency f_s and is closed an interval δ seconds during each close-open cycle. If, then, a d-c voltage of 1 volt is applied to the switch input, the output voltage will be a series of pulses of frequency f_s and of width δ . This voltage may be expanded by Fourier's theorem into the form

$$e_s(t) = \delta f_s \left(1 + \sum_{n=1}^{\infty} a_n \cos n\omega_s t \right)$$

where
$$a_n = \frac{\sin n\pi\delta f_s}{n\pi\delta f_s} \quad (\text{A-2})$$

If, on the other hand, $e_c(t)$ is applied to the switch input, and if δ is small compared to the period of $e_c(t)$ so that $e_c(t)$ essentially remains constant during the sampling interval, the switch output may be expressed as

$$\begin{aligned} e_o(t) &= e_c(t)e_s(t) \\ &= E \delta f_s \left[1 + m \cos \omega t + 2 \sum_{n=1}^{\infty} a_n \cos n\omega_s t \right. \\ &\quad \left. + 2m \sum_{n=1}^{\infty} a_n \cos n\omega_s t \cos \omega t \right] \quad (\text{A-3}) \end{aligned}$$

The third term in the brackets may be expanded into a trigonometric identity and the result is

$$\begin{aligned} e_o(t) &= E \delta f_s \left\{ (1 + m \cos \omega t) + 2 \sum_{n=1}^{\infty} a_n \cos n\omega_s t \right. \\ &\quad \left. + m \sum_{n=1}^{\infty} a_n [\cos (n\omega_s + \omega)t + \cos (n\omega_s - \omega)t] \right\} \quad (\text{A-4}) \end{aligned}$$

Inspection of this equation shows that all of the camera information is contained in the first term within the braces and the other terms may be discarded by means of a low-pass filter that has a cutoff frequency f .⁵ We must notice, however, that the use of the filter imposes a restriction on the sampling frequency. If, as we have assumed in the preceding work, $f_s = 2f$, the lower sideband term ($n\omega_s - \omega$) for $n = 1$ in eq. (A-4) will have a frequency less than f , which will also pass through the filter. It follows at once that f_s must exceed $2f$ slightly if the filter output is to consist of only the original information $e_c(t)$.

Let us summarize these results: The camera output is sampled at a rate just exceeding 8 mc. The resulting train is passed through a low-pass filter having a 4-mc cutoff frequency. The filter output then is the original camera signal complete in all its components up to 4 mc, that is

$$e_{of}(t) = E' \delta f_s (1 + m \cos \omega t) \quad (\text{A-5})$$

One's immediate reaction to this summary is: Why all the bother? Why sample the signal, filter it, and end up with the original signal? The answer is that the system which we have considered points the way to an interlaced dot system that allows more detail to be transmitted over a given channel bandwidth. Let us continue to study the simple system to complete the background for the more complicated one.

The filter output is delivered over the communication channel to the receiving equipment, where it is again sampled by a switch operating synchronously with the one located at the camera. These sample pulses are applied to the grid of the reproducing CRT, which displays the image in dot form. It is interesting to notice that for the figures we have used in the illustrative examples the number of dots per inch along the screen is approximately the same in either direction, horizontal or vertical; hence the image will be reproduced with none of the usual line structure apparent.⁶

⁵ We shall assume that the filter has an ideal linear-phase characteristic. Then, since all components in the pass band are delayed by a constant interval, this delay may be omitted from the equations.

⁶ This lineless image may be reproduced in a conventional receiver if a suitable gating circuit is interposed between the video amplifiers and the CRT grid.

A-3. Dot Interlace

We may now consider the more complex system involving the principle of dot interlace, which is a further extension of the sampling concept. In Chapter 2 we saw that sampling of one-half a picture per field by means of the 2-to-1 vertical-interlaced raster yielded a 50 per cent saving in bandwidth. We shall see that, in the present case, a sampling of dots along a line will permit transmission of twice the amount of detail in a given bandwidth. This is not in contradiction to the Hartley law because the sampling process halves the picture rate. That is, under commercial line-scanning standards, only 15, rather than 30, complete pictures will be transmitted each second.

In essence the dot interlace system samples every other picture element in one frame. In the next frame the remaining elements that were not covered in the first frame are sampled. We must assume for purposes of analysis, then, that the picture remains unchanged over two successive frames. This restriction is purely of a mathematical nature and is not a real restriction in the practical application of the dot interlace technique.

Since our purpose is to demonstrate that dot interlace permits transmission of greater detail within a given bandwidth, we shall consider how the system operates at two video modulating frequencies, f and f_1 , defined by

$$f < \frac{f_s}{2} \quad (\text{A-6a})$$

and
$$\frac{f_s}{2} < f_1 < f_s \quad (\text{A-6b})$$

where f_s is the sampling frequency.

Consider, first, the analysis for the lower frequency f . For the commercial standards this will be any frequency up to 4 mc, and the sampling frequency will be 8 mc. Let the video information be given by eq. (A-1). For the first frame the operation of the sampling switch is given by eq. (A-2). In the second frame, however, the samples lie midway between those of the first frame; hence, the operation of the switch in the second frame will be expressed by:

$$\begin{aligned}
 e_{a2}(t) &= \delta f_s \left[1 + 2 \sum_{n=1}^{\infty} a_n \cos n(\omega_s t + \pi) \right] \\
 &= \delta f_s \left[1 + 2 \sum_{n=1}^{\infty} (-1)^n a_n \cos n\omega_s t \right] \quad (\text{A-7})
 \end{aligned}$$

It may be shown quite readily, that if the sampling process is applied to the video information, and the sampled output passed through a filter having a cutoff frequency of $f_s/2$, the output from either frame will be identical to the original modulating signal, except for a constant factor δf_s . This checks our previous results: A signal extending up to 4 mc may be transmitted over a 4-mc bandwidth even though it has been sampled and filtered.

The signals are sampled at the receiver again. The sampled output for the first frame is given by (A-3) times δf_s . To find the output of the sampler during the second frame we multiply the original signal (A-1) by (A-7) and the modifying factor δf_s to obtain

$$\begin{aligned}
 e_{o2}(t) &= E (\delta f_s)^2 \left[1 + m \cos \omega t + 2 \sum_{n=1}^{\infty} (-1)^n a_n \cos n\omega_s t \right. \\
 &\quad \left. + 2m \sum_{n=1}^{\infty} (-1)^n a_n \cos n\omega_s t \cos \omega t \right] \\
 &= E (\delta f_s)^2 \left\{ 1 + m \cos \omega t + 2 \sum_{n=1}^{\infty} (-1)^n a_n \cos n\omega_s t \right. \\
 &\quad \left. + m \sum_{n=1}^{\infty} (-1)^n a_n [\cos (n\omega_s + \omega)t + \cos (n\omega_s - \omega)t] \right\} \quad (\text{A-8})
 \end{aligned}$$

The two signals are then passed through a filter of cutoff frequency f_s and applied to the CRT grid. In the filter all frequencies greater than f_s are suppressed. The remaining signal components for the two successive frames are displayed on the fluorescent screen where, because of the phosphor decay time and retina retentivity, they are effectively added. Thus the final signal displayed on the tube is the sum of the filter output over two successive fields:

$$\begin{aligned}
 e_{\theta 1} + e_{\theta 2} &= E (\delta f_s)^2 [1 + m \cos \omega t + 2a_1 \cos \omega_s t + ma_1 \cos (\omega_s - \omega)t] \\
 &\quad + E (\delta f_s)^2 [1 + m \cos \omega t - 2a_1 \cos \omega_s t - ma_1 \cos (\omega_s - \omega)t] \\
 &= 2E (\delta f_s)^2 (1 + m \cos \omega t) \tag{A-9}
 \end{aligned}$$

This shows that the original modulation corresponding to a d-c component with a superimposed component of frequency, f , that is less than one-half of the sampling frequency, is reproduced on the CRT unmodified except by a constant factor $2(\delta f_s)^2$. This completes the analysis for the low-frequency component.

We next consider the analysis for a high frequency f_1 , defined by eq. (A-6b). In terms of our previous numerical example, f_1 represents any of those frequencies which lie between 4 and 8 mc that would not be transmitted in the conventional television system. In all the work that follows, numerical subscripts will be used to indicate in which frame a given equation applies. The camera signal which is assumed to be the same in both frames is given by

$$e_c(t) = E (1 + m \cos \omega_1 t) \tag{A-10}$$

Then the output of the sampler during frame 1 will be

$$\begin{aligned}
 e_{o1}(t) &= E \delta f_s \left\{ 1 + m \cos \omega_1 t + 2 \sum_{n=1}^{\infty} a_n \cos n\omega_s t \right. \\
 &\quad \left. + m \sum_{n=1}^{\infty} a_n [\cos (n\omega_s + \omega_1)t + \cos (n\omega_s - \omega_1)t] \right\} \tag{A-11}
 \end{aligned}$$

and during frame 2 it will be [by multiplying eq. (A-10) by (A-7)]

$$\begin{aligned}
 e_{o2}(t) &= E \delta f_s \left\{ 1 + m \cos \omega_1 t + 2 \sum_{n=1}^{\infty} (-1)^n a_n \cos n\omega_s t \right. \\
 &\quad \left. + m \sum_{n=1}^{\infty} (-1)^n a_n [\cos (n\omega_s t + \omega_1)t + \cos (n\omega_s - \omega_1)t] \right\} \tag{A-12}
 \end{aligned}$$

As we have seen before, the sampler output is passed through a low-pass filter having a cutoff frequency of $f_s/2$. Now since $f_s/2 < f_1 < f_s$ we notice that

$$\left. \begin{aligned} (n\omega_s + \omega_1) &> \frac{\omega_s}{2} & \text{for } n &\geq 1 \\ (n\omega_s - \omega_1) &< \frac{\omega_s}{2} & \text{for } n &= 1 \\ (n\omega_s - \omega_1) &> \frac{\omega_s}{2} & \text{for } n &> 1 \end{aligned} \right\} \quad (\text{A-13})$$

Therefore the filter output which is delivered to the communication channel in each case will be

$$e_{c1}(t) = E \delta f_s [1 + ma_1 \cos(\omega_s - \omega_1)t] \quad (\text{A-14})$$

and
$$e_{c2}(t) = E \delta f_s [1 - ma_1 \cos(\omega_s - \omega_1)t] \quad (\text{A-15})$$

The last two equations show the basis of the dot-interlace system. They show that *those signal components between 4 and 8 mc are also transmitted over a 4-mc bandwidth* in the form of a lower sideband corresponding to a sub-carrier of 8 mc, the sampling frequency. We must continue the analysis to see if these high-frequency components are restored properly at the receiver CRT. In the receiver the incoming signal is sampled again at the same rate and for the same interval that it was at the camera. Thus for field 1 we have

$$\begin{aligned} e_{10}(t) &= e_{s1}(t)e_{c1}(t) \\ &= E (\delta f_s)^2 \left[1 + ma_1 \cos(\omega_s - \omega_1)t + 2 \sum_{n=1}^{\infty} a_n \cos n\omega_s t \right. \\ &\quad \left. + 2ma_1 \sum_{n=1}^{\infty} a_n \cos n\omega_s t \cos(\omega_s - \omega_1)t \right] \\ &= E (\delta f_s)^2 \left\{ 1 + ma_1 \cos(\omega_s - \omega_1)t + 2 \sum_{n=1}^{\infty} a_n \cos n\omega_s t \right. \\ &\quad \left. + ma_1 \sum_{n=1}^{\infty} a_n [\cos(\overline{n+1}\omega_s - \omega_1)t \right. \\ &\quad \left. + \cos(\overline{n-1}\omega_s + \omega_1)t] \right\} \quad (\text{A-16}) \end{aligned}$$

and for the second field

$$\begin{aligned}
 e_{20}(t) &= e_{s2}(t)e_{c2}(t) \\
 &= E (\delta f_s)^2 \left[1 - ma_1 \cos (\omega_s - \omega_1)t + 2 \sum_{n=1}^{\infty} (-1)^n a_n \cos n\omega_s t \right. \\
 &\quad \left. - 2ma_1 \sum_{n=1}^{\infty} (-1)^n a_n \cos n\omega_s t \cos (\omega_s - \omega_1)t \right] \\
 &= E (\delta f_s)^2 \left\{ 1 - ma_1 \cos (\omega_s - \omega_1)t + 2 \sum_{n=1}^{\infty} (-1)^n a_n \cos n\omega_s t \right. \\
 &\quad - ma_1 \sum_{n=1}^{\infty} (-1)^n a_n [\cos (\overline{n+1}\omega_s - \omega_1)t \\
 &\quad \left. + \cos (\overline{n-1}\omega_s + \omega_1)t] \right\} \quad (\text{A-17})
 \end{aligned}$$

As we have seen previously, these sampled outputs are now passed through a filter having an f_s cutoff and added on the CRT. Thus the final signal on the tube grid is

$$\begin{aligned}
 e_{\rho 1} + e_{\rho 2} &= E (\delta f_s)^2 [1 + ma_1 \cos (\omega_s - \omega_1)t \\
 &\quad + 2a_1 \cos \omega_s t + ma_1^2 \cos \omega_1 t] \\
 &\quad + E (\delta f_s)^2 [1 - ma_1 \cos (\omega_s - \omega_1)t \\
 &\quad - 2a_1 \cos \omega_s t + ma_1^2 \cos \omega_1 t] \\
 &= 2E (\delta f_s)^2 (1 + ma_1^2 \cos \omega_1 t) \quad (\text{A-18})
 \end{aligned}$$

The only factor that affects the original modulation is a_1^2 . We originally specified, however, that δ is a small fraction of the sampling period $1/f_s$; therefore a_1 may be seen to be nearly 1 from eq. (A-2). Comparison of eqs. (A-9) and (A-18) shows that the dot interlace treats all components up to 8 mc in the same manner. We have also seen from eqs. (A-14) and (A-15) that these components are transmitted over a 4-mc bandwidth. This, then, is the advantage of the dot interlace system of picture transmission. As compared to the conventional system it permits transmission of twice the amount of detail in the same bandwidth. The time of transmission for a complete frame is doubled.

It should be mentioned in conclusion that a transmitted signal which is generated by a dot-interlace scanning system is amenable

to reception and reproduction by a conventional television set provided, of course, that the standard number of lines and frame frequency are used at the camera end of the system. It is for this reason that there is hope in some quarters that the transmission standards in this country will be modified to include dot interlace. Such a change will permit new receivers that incorporate a sampler and filter to reproduce a picture of higher resolution without rendering obsolete any receivers of the present design.

A-4. Dot Interlace and Multiplex⁷

The advantages of the dot interlace technique are probably most apparent when dot interlace is combined with a multiplexing system. We have already seen how this combination is effected in a color television system that was discussed in sections 18-22 and 18-23; hence we shall review only the outline of the method. In tricolor transmission three signals, one for each of the primary colors, are developed. The sampling switch is arranged to sample the three color channels in time. This is the multiplexing part of the system. Furthermore, the sampling rate is so high that only one picture element at a time is sampled in each channel. Thus the two methods are combined to permit the transmission of samples of three color channels over a 4-mc bandwidth and to make possible the reproduction of a picture whose video components extend up to 8 mc.

⁷ W. Boothroyd, "Dot Systems of Color Television," Part II. *Electronics*, **23**, 1 (January 1950).

PROBLEMS

2-1. The deflection circuits in a television receiver are adjusted so that the entire circular screen of the cathode-ray tube is used for reproducing the picture. Calculate what percentage of the image is lost in this type of presentation.

2-2. The deflection circuits in a television receiver are adjusted so that the width of the picture is equal to the diameter of the circular cathode-ray tube screen. Calculate what percentage of the image is lost in this type of presentation.

2-3. A television system requires a 4-mc bandwidth to transmit a certain image at a picture repetition frequency of 30 per second. A comparable facsimile system using the same number of lines transmits the same picture in 10 minutes. What bandwidth is required?

2-4. It is proposed that the horizontal scan in a television system be sinusoidal rather than saw-tooth in order to reduce the problem of flyback. Criticize the proposal from the point of view of picture reproduction.

2-5. Draw a sketch showing the effect on the scanning lines of a 60-cycle ripple component in the vertical deflection system. A 2:1 interlaced scan is used. The frame frequency is 30 per second. How will the ripple component affect the reproduction of detail in the reproduced image?

2-6. Explain why adjacent lines overlap near the edges of the pattern when the bidirectional scan of Fig. 2-6*b* is used. What difficulty in picture reproduction does this overlapping cause?

2-7. It is proposed that a saving in bandwidth without loss of resolution results if vertical scanning lines are used instead of horizontal lines. (See Fig. 2-7.) The proposal is based on the statement that in either case the number of lines is directly proportional to the picture dimension normal to the line direction. Criticize the proposal.

2-8a. Sketch the positions of successive odd-field lines due to 60-cycle modulation of a 24-frame-per-second picture.

b. In a 10-in. wide picture using 500 active lines what will be the maximum peak-to-peak motion of a scanning line in successive odd fields? Assume a horizontal ripple deflection of 0.25 of the spacing between adjacent lines.

3-1. A rectangular scanning raster of $\frac{4}{3}$ aspect ratio is to be produced on the face of a 7JP4 cathode-ray tube operating at an accelerating potential

of 5500 volts. Calculate the peak values of horizontal and vertical deflecting voltages required to make the raster as large as possible.

3-2. Calculate the peak values of the horizontal and vertical deflecting voltages to produce a picture of the type described in problem 2-2 on a 10HP4 cathode-ray tube. The accelerating potential is 6000 volts. If the transmitted picture is of standard aspect ratio, what will be the height of the raster on the 10HP4?

3-3. Discuss the effect of moving the focusing coil axially along the neck of a cathode-ray tube.

3-4. A certain deflection yoke used in conjunction with a 3-in. monitoring cathode-ray tube operating at 800 volts accelerating potential is found experimentally to have the following deflection sensitivities:

Vertical coil: 18 ma/in.

Horizontal coil: 78 ma/in.

What peak values of deflection current are required to produce a picture of standard aspect ratio and a diagonal of 3 in. on the monitor operating at 1000 volts for greater picture brightness?

3-5. The normal height of the image reproduced on a type 14BP4 rectangular-face cathode-ray tube is $8\frac{1}{2}$ in. A deflection yoke 5.5 cm long is centered on the neck of the tube at a point $10\frac{1}{2}$ inches from the fluorescent screen. The effective width of the deflecting field is 5.5 cm; the accelerating potential is 12 kv. Calculate the ampere-turns required to produce normal picture height. Assume the deflection field is uniform.

3-6. A television receiver uses a 15DP4 cathode-ray tube which is of the bent-gun type. The magnetic field for the ion trap is furnished by two permanent magnets held to the neck of the tube by spring clips. It is found that no raster appears on the screen. In what manner might this trouble be due to improper placement of the magnet assembly? Suggest probable remedies.

3-7. A type 16YP4 cathode-ray tube requires a peak-to-peak angular deflection of 70° to produce a scanning line of normal width. The deflection yoke is to be centered on a point 11 in. from the tube face which may be assumed flat. Using a point-by-point method, calculate the wave form of deflection current required to give constant scanning velocity along a line.

3-8. A 3:1 interlaced scanning raster is to be considered. Draw the lines for three consecutive fields and the horizontal and vertical deflection components. Assume a 10-line system with zero flyback in both directions. Is there any advantage in having the total number of lines an odd integer in this case? List the principal advantages and disadvantages of the system.

3-9. Draw the block diagram of the timing unit required to provide proper synchronization in a 3:1 interlaced scan. Label the frequency of

each oscillator and frequency-divider output. Assume 520 lines and 30 frames per second.

3-10a. Under current transmission standards the horizontal blanking interval is $0.16H$. What maximum value of horizontal flyback ratio at the receiver will permit the retrace to be completed within the blanking interval?

b. The vertical blanking interval has a duration of $0.05V$. What maximum value of vertical flyback ratio may be used?

3-11. Assuming an ideal checkerboard pattern to be transmitted, derive an expression for the rate at which picture elements are being transmitted in terms of number of lines, aspect ratio, blanking ratios, frame frequency, and utilization ratio.

3-12. Beginning from eq. (3-20) derive an expression for the deflection sensitivity, D/I , in terms of the inductance of a deflection coil.

4-1. Derive eq. (4-7) without using the superposition theorem.

4-2. In a 525-line, 2 : 1 interlaced, 60 field/sec scan, the sweep generator is designed for a maximum departure from linearity of 2%. If the vertical blanking ratio is 1/19, what is the maximum displacement of any line from its normal position in an 8 by 10-in. raster? Assume that the deflection amplifiers are linear. Will the crowding of lines occur at the top or the bottom of the picture?

4-3a. Design a saw-tooth generator using a 6SN7 discharge tube to meet the following specifications: Linearity within $\pm 1\%$, supply voltage 300 volts, charging interval 16 millisecc, flyback ratio of 1/19.

b. The saw-tooth output is to be amplified and delivered to a 7JP4 cathode-ray tube. What gain is required in the amplifier to produce a $3\frac{1}{2}$ -in. deflection?

4-4. The deflection factor of a 7EP4 cathode-ray tube is 95 volts/in. The output of a saw-tooth generator is fed through a push-pull voltage amplifier to produce a 5.6-in. sweep on the 7EP4. The constants of the saw-tooth generator are $C = 500 \mu\mu\text{f}$, $R = 1 \text{ Meg}$, $E_{bb} = 250 \text{ v}$. The synchronizing signal frequency is 15.75 kc and the flyback ratio is 1/10.

a. Calculate the required gain per tube in the amplifier.

b. Calculate the maximum departure from linearity in the saw-tooth signal. Assume no distortion in the amplifier.

4-5. A thyratron-controlled saw-tooth generator is to be used to provide the vertical scan in an unsynchronized helter-skelter scan system. The nominal vertical scanning frequency is 60 per second. An 884 is to be used.

a. Design the sweep generator to provide linearity within $\pm 1\%$. Specify the thyratron grid voltage required for a supply voltage of 250 volts. The tube drop of the 884 is 16 volts and its control characteristic is given by:

| | | | | |
|--------|-----|------|------|-------|
| e_b | 300 | 170 | 92.5 | volts |
| $-e_c$ | .30 | 17.5 | 10 | volts |

b. What nominal blanking interval should be used? Neglect ionization and deionization times.

4-6a. Design a free-running multivibrator to run at 60 cycles. The cut-off interval of one section is to be 19 times that of the other. A 6SN7 is to be used and a plate supply voltage of 300 volts is available.

b. Calculate and plot the voltage wave forms appearing on both grids and both plates.

4-7. What is the per cent error in the frequency of the multivibrator in problem 4-6 as calculated by the order-of-magnitude equation (4-42)?

4-8. The multivibrator of problem 4-6 is synchronized by positive pulses delivered from an infinite impedance source to the appropriate grid. If the synchronizing pulses occur at 15-millisecond intervals, what must be their minimum amplitude for synchronization to occur? State the assumptions made in arriving at your answer.

4-9a. Design a multivibrator to operate at 5 kc with $\tau_2/\tau_1 = 2$. Use a 6SN7, $E_{bb} = 250$, and equal plate resistors of 100,000 ohms.

b. Design a sync input system for the multivibrator. The sync frequency is 5 kc. Do not change τ_2 . Redesign the multivibrator if necessary so that it operates with the proper duty cycle when synchronized.

4-10. A certain application requires that a multivibrator deliver a square wave with $\tau_1 = \tau_2$. It is suggested that the addition of a resistance of proper value in series with each of the condensers C_1 and C_2 will improve the wave form. Discuss this suggestion and derive the necessary design equations.

4-11. Evaluate the circuit constants for the circuit of Fig. 4-19 to meet the following requirements:

$$E_{bb} = 300 \text{ volts.}$$

$$\tau_c = 60 \text{ } \mu\text{sec} \quad \tau_d = 8 \text{ } \mu\text{sec} \quad \tau_e = 0.3 R_5 C_5$$

6SN7's are used throughout.

4-12. Derive the design equations for the saw-tooth generating circuits of Fig. 4-22b and c. Be sure to account for the effect of the sync injection circuits.

4-13. A kit for building a television receiver has two blocking oscillator transformers:

Transformer A

Primary to secondary turns ratio
= 1:2.

Secondary inductance = .02 henry.

Transformer B

Primary to secondary turns ratio
= 1:4.

Primary inductance = 1.2 henry.

Which unit should be used in the horizontal deflection system? Explain.

4-14. The circuit of Fig. 4-25 is used to develop the vertical sweep signal in a television receiver. Assume a reasonable cutoff-interval for the tube and calculate typical values for R_s , C_s , C_p , and R_p .

4-15. Measurements performed on the vertical deflection coils of a yoke yield the following data:

$$L = 43 \text{ mh.} \quad Q = 4.5 \text{ at 1 kc.}$$

It is further determined that this coil in conjunction with a 7-in. CRT operating at rated voltage gives a deflection factor of 17.85 ma/in. If this coil is driven by a triode-connected 6F6 through a 2:1 ideal transformer, calculate the shape and magnitude of the grid voltage required for the 6F6 to produce a 4-in. vertical deflection on the CRT.

4-16. Design a trapezoidal generator to drive the circuit of problem 4-15. A 6SN7 is to be used with a supply voltage of 250 volts. Neglect flyback effects.

4-17. Modify the design equations for a trapezoidal generator to take into account shunt damping resistance across the deflection coils.

4-18. If shunt resistance damping is used in the 30-frame-per-second deflection system which is designed in section 4-12, what average power would be dissipated in the damping resistor? Is it necessary to recalculate the peak value of yoke current to determine the answer?

4-19. The horizontal deflection coils and output transformer of a television receiver have a natural resonant frequency of 75 kc which represents approximately the highest figure obtainable in current practice. If resistance damping is used with these components, will flyback occur within the 10- μ sec horizontal blanking interval specified by U. S. standards? What other features of resistance damping would make its use in this application undesirable?

4-20. Explain the operation of the voltage tripler shown in Fig. 4-52. Remember a pulse, not a sine wave, is applied. Consider the operation during the pulse and inter-pulse periods.

4-21. Verify eq. (4-83).

4-22. Explain why a given value of C_s has a greater effect on the flyback time when used in the circuit of Fig. 4-44a than in the circuit of Fig. 4-44b.

4-23. Explain why a change in the value of R_d in Fig. 4-50 affects the linearity of only the first half of the scan.

4-24. In Fig. 4-50 what portion of the sweep is affected by adjustment of L_1 ? Explain briefly.

4-25. Derive an expression for the flyback time of a deflection system which employs shunt resistance damping.

5-1. In the facsimile system of Fig. 1-7 one scan over the entire picture occurs over a spiral path. Can the output signal be represented by a Fourier series? If so, is a one- or two-dimensional expansion required?

5-2. A rectangular aperture has a uniform response I . Its length is $2c$ in the ξ direction, and its width is $2d$ in the η direction. Calculate and plot the aperture admittance.

5-3. Repeat problem 5-2 for a circular aperture of constant response.

5-4. It may be demonstrated mathematically that the most objectionable extraneous component in a scanned image may be eliminated by using a rectangular aperture of uniform response such that

$$|\mu| = 1 \quad \text{for} \quad k = 0, l = 0.$$

Discuss how this aperture eliminates the component on a physical basis.

5-5. The effective aperture on a CRT is circular and may be assumed to leave a $\cos^2 \theta$ variation of intensity across the width of a scanned line. If the spot radius is adjusted so that adjacent lines overlap each other a distance equal to the spot radius, a flat field is obtained. Demonstrate this analytically. What objection would arise to this overlap as far as vertical resolution is concerned?

5-6. From physical considerations why is it logical that B_k' is proportional to V_p , the vertical scanning interval? (See section 5-11.)

6-1. Derive an equation relating the foot-candle to the meter-candle.

6-2. A certain scene which is being televised has an average illuminance of 80 foot-candles. Assuming the average reflection coefficient to be 0.8, calculate the average illuminance of a photocathode on which the scene is focused by an $f/4.5$ lens. State the assumptions made in calculating your answer.

6-3. Derive an expression for the illuminance required on the cathode of the phototube used with a flying-spot scanner to produce a given signal-to-noise ratio. The shunt capacitance across R_o is not negligible.

6-4. An image dissector has a 3 by 4-in. photocathode. a. Calculate the aperture size required to give equal vertical and horizontal resolutions and a figure of merit of 100,000.

b. For what number of active scanning lines should the associated scanning system be designed?

c. If the dissector has a three-stage electron multiplier, each with secondary emission ratio 3, calculate the illuminance in foot-candles required at the photocathode to produce an output current of 45×10^{-9} amp. The luminous sensitivity of the photocathode is $20 \mu\text{amp/lumen}$.

d. Calculate the shot noise present in the output over a 4.5-mc half-power bandwidth.

6-5. When the output from an image dissector is viewed on a monitor tube, it is found that the entire image is twisted through a slight angle relative to the horizontal scanning lines in the monitor. What is the most probable cause of this effect on the image dissector? Explain your answer.

6-6. Does the output of an image dissector have a fixed black level? Explain.

6-7. A commercial type iconoscope operates into a coupling resistance of 10,000 ohms which resistor is at room temperature. Calculate the required mosaic illuminance to give a signal-to-noise ratio of 20 : 1 within a 4.5-mc half-power bandwidth.

6-8. Demonstrate that a parabolic shading voltage may be developed by integrating a saw-tooth wave of proper frequency. Assume the average value of the saw-tooth to be zero.

6-9. Draw the circuit diagram of a balanced modulator to be used for keystone correction in an iconoscope. Show the wave forms of all applied voltages. Is there any need for balancing out a spurious field frequency component in the output?

6-10. Discuss the possibility of using magnetic rather than electrostatic deflection for the vertical deflection in the orthiconoscope tube.

6-11. Explain why saturated photoemission is not possible in the iconoscope due to space charge limitation.

6-12. Why does no problem of space charge limitation occur at the photocathode in the orthicon tube?

6-13. Is it possible for the return beam in the orthiconoscope to retrace the path of the incident beam? Explain.

6-14. Derive an expression for the maximum useful electron multiplier multiplication, m , corresponding to a given beam current in the image orthicon.

7-1. Discuss the adaptability of transformer-coupled and resistance-coupled voltage amplifiers to broadband video amplification.

7-2. Discuss the relative merits of triodes and pentodes for use in video amplifiers.

7-3. Derive Kùpfmùller's rule for τ_r defined as the rise time between the 10% and 90% response values for the ideal low-pass amplifier.

7-4. Explain why the figure of merit defined in (7-71) is not equal to the gain bandwidth product of an amplifier using a four-terminal coupling network.

7-5. The delay characteristic rather than phase shift is generally plotted for a video amplifier. Derive a relationship for τ in terms of θ_2 and f_2 .

7-6. The condition of maximal flatness in a video amplifier uses the largest degree of compensation without having the amplification-vs.-frequency curve go through a peak. It is satisfied if

$$\frac{d^2A}{df^2} = \frac{dA}{df} = 0$$

By performing the necessary differentiation, show that this condition is the same as Case II in section 7-9.

7-7. Calculate and plot the steady-state amplitude and delay characteristics for an m -derived, shunt-peaked amplifier for the two conditions:

$$A. \quad R_1 = \frac{0.295 \times 10^6}{f_c C_s}; \quad L = 0.414 \times 10^{-6} C_s R_1^2; \quad C_1 = 0.352 C_s.$$

$$B. \quad R_1 = \frac{0.255 \times 10^6}{f_c C_s}; \quad L = 0.296 \times 10^{-6} C_s R_1^2; \quad C_1 = 0.125 C_s.$$

where f_c = highest frequency to be amplified, and C_1 is the shunt capacitance across L . From your results state the advantage of each design condition.

7-8. A video amplifier is being tested by the square-wave method. It is observed that the response exhibits excessive overshoot. The value of compensating inductance may be changed by moving the powdered iron core of the inductor. Should the core be moved into or out of the coil to remedy the overshoot? Explain.

7-9. A resistance-coupled amplifier using a 6AC7 is found to have an upper half-power frequency of 1 mc when a plate load resistance of 6000 ohms is used. The amplifier is to be converted for video operation. Specify the value of load resistance and shunt compensating inductance to be used to give a satisfactory transient response and a 4-mc bandwidth.

7-10. D. A. Bell (*Wireless Engineer*, July 1943) has suggested that the transient response of the ideal amplifier may be calculated by considering unit function to be the limiting value as T approaches infinity of one cycle of a 50-50 square pulse, average value zero, and of period T . Under this assumption the amplifier output is

$$e_o(t) = \frac{2A}{\pi} \int_0^{\omega_s} \frac{\sin \omega(t - \tau_d)}{\omega} d\omega, \quad t > 0$$

$$e_o(t) = 0 \text{ for } t < 0$$

Plot this response and show why it is nearer to physical reality than the response shown in Fig. 7-8b.

7-11. Verify eq. (7-103).

7-12. Design a 4-mc, 3-stage video amplifier suitable for television use. All stages are to be identical. Use type 6AC7 tubes, and simple shunt compensation.

7-13. Repeat problem 7-12 but use series compensation. State the relative merits of the two types of compensation for this design problem.

7-14. Calculate and plot the steady-state and transient performance of the amplifier shown in Fig. 7-25.

7-15. Using the steady-state curves of Figs. 7-10 and 7-12, compare in a qualitative fashion the transient response of a Freeman-Schantz compensated amplifier and a shunt-compensated amplifier of $K = 0.5$.

7-16a. Plot A_h/A_m vs. y for the range $0.5 \leq y \leq 1.5$ for a series-compensated amplifier.

b. Using this curve as a basis, compare the transient response of the series-compensated amplifier with that of the shunt-compensated amplifier with $K = 0.44$.

7-17. Construct a table giving the following values for a shunt-compensated amplifier with $K = .5$, a series-compensated amplifier, and a series-shunt compensated amplifier: shunt L , series L , R_1 , A_m/A_m (uncompensated), A_{f2}/A_m , and $\Delta\tau$, the departure from constant delay at f_2 .

7-18. By the use of Thevenin's theorem, reduce the equivalent high-band circuit of a video feedback doublet (Fig. 7-29d) to a generator $g_{m1}E_{\phi}$, shunted by an impedance, feeding a parallel load of C_{s2} and R_1 . On the basis of this equivalent circuit, explain how high-band compensation takes place.

7-19a. Design a shunt-compensated video amplifier, $K = 0.5$, for a top video frequency of 4 mc. Use a 6AC7 and assume a C_s of $24 \mu\mu\text{f}$.

b. Calculate and plot the amplitude and delay characteristics in the mid- and high-frequency bands.

7-20a. Compensate the amplifier of problem 7-19 so that the lower half-power frequency is 40 cycles. Neglect degenerative effects of the cathode and screen grid.

b. Calculate and plot the low-band response, taking into account cathode and screen effects.

7-21. A d.c. restorer employing a 1-V tube has the following values.

$$R = 1 \text{ Meg}, \quad R_g = 5 \text{ k-ohm}, \quad R_d = 4 \text{ k-ohm}, \quad C = .1 \mu\text{f}.$$

Calculate the bias developed for an applied voltage of 30 volts peak, of width $5 \mu\text{sec}$, and of repetition rate 15.75 kc.

7-22. Design a cathode follower to match a 52-ohm coaxial cable. Use a blocking condenser to isolate d.c. from the cable. Be sure to adjust for proper bias on the tube.

8-1. Analyze the paraphase amplifier of Fig. 8-6.

8-2. Is it possible in the paraphase amplifier of Fig. 8-6 to compensate for the difference in grid voltages by adjustment of the two plate load re-

sistances? If it is, derive an expression relating these load resistances so that both output voltages have the same magnitude.

8-3. Explain how the use of push-pull deflection in the CRT permits a fair degree of non-linearity in sweep voltages. Would this be of particular advantage in the Type I closed system? Explain.

8-4. Explain how the horizontal sweep circuits in the Type III closed system are protected from the vertical sync pulse.

8-5. It is proposed to design a simple closed television system utilizing a minimum number of parts. Two commercial cathode-ray oscilloscopes are available, one of which is to be used as a flying-spot scanner, and the other as a receiver. Sweep voltages are to be generated by the two sweep circuits internal to the oscilloscopes. Draw a diagram of the complete system, giving circuit values for all components not in the oscilloscopes. State what modifications, if any, need be made in the oscilloscopes and what provisions, if any, need be made for synchronization. Give some thought to delay problems arising from the location of the two sweep generators, one at each end of the system. To what form of subject matter is transmission over this system limited?

10-1. Flatness of field may be achieved by adjusting the viewing distance from the screen because of the limited visual acuity of the eye. Derive an expression for viewing distance, in terms of picture height and the number of active lines, which just allows adjacent scanned lines to be resolved. Assume that the eye can just resolve to points separated by 1 minute of arc.

10-2. Criticize the use of a checkerboard pattern as a subject for checking resolution in a television system. It is recommended that the criticism be based on the problems of reproducing a repetitive function or a pulse.

10-3. Discuss the reasons why the horizontal-resolution curves of Fig. 10-7 resemble the transient-response curve of an amplifier.

10-4. Derive the relationship between the arbitrary units of V and H in Fig. 10-8 and the units of the same quantities in the preceding figures.

10-5. Make a calibration chart for lines of resolution and effective bandwidth for the wedges shown in the test pattern of Fig. 10-9.

10-6. Explain how "smearing" of the station call letters in a received test pattern indicates poor transient response in the system.

10-7. Criticize the test patterns of Figs. 10-9 and 10-10 for their ability to show if the CRT spot is defocused as it moves away from the center of the screen.

10-8. Explain why receiver defects are more noticeable when test pattern, rather than normal program material, is being received.

11-1a. Derive the output voltage equations for a differentiating circuit and for an integrating circuit, each using a resistor-inductance combination.

11-1b. Compare the desirability of these circuits with their RC counterparts.

11-2. Discuss the separation of horizontal and vertical sync pulses from the following point of view: differentiating circuits are sensitive to steepness and timing of leading edges; integrating circuits are sensitive to pulse width. Justify this manner of considering differentiation and integration.

11-3. It is found that a vertical sweep circuit which uses the circuit shown in Fig. 11-5 slips out of synchronization when a low-amplitude sync pulse is applied. Explain the reason for this. Explain the need for holding the sync amplitude within fixed limits in order to maintain synchronization.

11-4. Design a 5 : 1 counter circuit using an 884 thyatron as the discharge device. The input is a 50-50 square wave of 5000 cps. Use the following constants:

$$E_a = 250 \text{ volts, } E_{bb} = 250 \text{ volts, } \Delta E_n \geq 2 \text{ volts.}$$

See problem 4-5 for the thyatron control characteristic.

11-5. What objections are there to using a thyatron discharge device in a television sync generator counting circuit? What other form of discharge device would be more satisfactory? Why?

11-6. Derive the equations for the shunt admittance presented by a reactance tube. Assume that the impedance of the RC branch and r_p are both large and that ωRC is small compared to 1.

11-7. A reactance tube circuit is to be used to control the frequency of a tuned-plate oscillator. The reactance tube is a 6SJ7 whose g_m - e_c characteristic for a plate voltage of 250 volts and a screen voltage of 100 volts is given by

$$g_m = (3000 + 500 e_c) \mu\text{mho} \quad \text{for } -5 \leq e_c \leq -1 \text{ volt.}$$

$$R = 1 \text{ Meg} \quad C = 300 \mu\mu\text{f}$$

The self-bias on the 6SJ7 is to be -3 volts. The oscillator tank condenser is $250 \mu\mu\text{f}$.

a. What value of tank inductance is required to produce an oscillator frequency of 31.5 Kc when no external control voltage is applied to the 6SJ7?

b. Draw the circuit diagram giving all circuit values.

c. What range of frequency may be covered by applying a d-c control voltage from -2 to $+2$ volts to the 6SJ7 grid?

11-8. The comparator circuit of Fig. 11-19 feeds the reactance tube of problem 11-7, but the phase of the 60-cycle sine wave is reversed from that shown at d in the figure. Will the control voltage developed across C_3 be of the correct polarity to correct the oscillator frequency? Explain. What changes could be made in the reactance tube circuit to remedy this condition? Explain.

11-9. Two 6SJ7's are connected in a voltage-adding circuit with a plate load resistance of 30,000 ohms. The plate supply voltage is 300 volts. Both screens are operated at 100 volts and the bias on each tube is -3 volts.

- a. Over what range of input voltage will the operation be satisfactory?
- b. Two square waves, one of twice the duration of the other, are applied to the grids. Each has a peak-to-peak amplitude of 1 volt. Plot the output wave form.
- c. Check your result using eq. (11-35).

11-10. An adding circuit consists of a single pentode with grid-to-cathode resistor R_g , plate-to-grid feedback resistor R_f , and plate-to-cathode load resistance R_L . Two input signals which are to be added are applied to the circuit in the following manner: each voltage in series with a resistor R is connected between grid and cathode.

- a. Show that if $R_f = R$, $R_g \gg R$, and the gain of the stage without feedback is large, the magnitude of the output voltage is equal to the sum of the two input voltages.
- b. Compare the relative merits of this adder and that described in the text.

11-11. Discuss the use of a pentagrid mixer tube as a voltage adder.

11-12a. Is it possible to delay a square wave, of width $20 \mu\text{sec}$ and repetition frequency 15,750 per second, $10 \mu\text{sec}$ per section without appreciable distortion in a constant- k artificial line? $R_0 = 50$ ohms. Explain.

- b. What effect would the distortion have on the output wave form?

11-13. Design a constant- k delay line to provide an overall delay of at least $2 \mu\text{sec}$ with a delay per section of $0.5 \mu\text{sec}$. The nominal impedance is to be 52 ohms and the applied square wave has a frequency of 15,750 pulses per second and a pulse width of $20 \mu\text{sec}$. Use m -derived terminating half-sections.

11-14a. Design a phase-shift network for operation at 15.75 ke. A center-tapped, 6.3-volt filament transformer rated at 10 amperes is available.

- b. If a double-pole, double-throw switch is mounted on the potentiometer in the phase shifter, how may it be used to increase the range of phase shift?

11-15. The horizontal sync pulse from a sync generator is being checked on an oscilloscope employing a 15,750-cycle sine-wave sweep. With the pulse centered on a 10-cm base line, the width of the pulse at the 10% amplitude points measures 2.81 cm.

- a. Does the pulse width lie within the tolerance limits specified by the F.C.C.?
- b. By what per cent does the pulse width deviate from the nominal value of $0.08 H$?

12-1. By extending eq. (12-29) show that vestigial sideband transmission introduces phase modulation of the carrier. This type of distortion causes trouble in receivers of the intercarrier type. What may be done at the transmitter to minimize this distortion?

12-2. Draw the rotating vector diagrams for the following systems of transmission: (a) double sideband, (b) single sideband, (c) vestigial sideband, assuming the lower sideband amplitude is one-half that of the upper sideband. Briefly discuss the problem of phase distortion of the carrier for each of the three systems.

12-3. The data for eq. (12-64) may also be derived by making the slopes of impedance vs. frequency equal for the lumped-constant resonant circuit and its transmission-line equivalent. Derive the necessary equations and verify the data.

12-4. Verify, by plotting, that the case $n = 1$ in eq. (12-64) gives the closest agreement between the series resonant circuit and its transmission-line equivalent.

12-5a. Design a type A filter using coaxial line elements for channel No. 7. The characteristic impedance is to be 52 ohms. The inner diameter of the largest tubing available is 3 in.

b. If the breakdown voltage between conductors is 75 kv per in. what is the maximum voltage which may be applied to the filter input?

12-6. The delay between a video transmitter operating on channel No. 13 and its m -derived vestigial sideband filter is to be .01 μ sec or less in order to minimize the effect of signals reflected from the filter to the transmitter. What maximum length of lossless coaxial cable may be used to connect the two units? Carry out the calculations at the video carrier frequency.

12-7. Design an m -derived vestigial sideband filter for channel No. 5. The inner diameter of the largest tubing available is 3 in. The input impedance is to be 72 ohms. Calculate and show all dimensions.

12-8. It is desired to build a Balun of the form shown in Fig. 12-23 for use with a transmitter operating on channel No. 5. The materials on hand are such that the ratio of the outer, quarter-wave sleeve to middle conductor diameter is 5 : 1. The input impedance between terminal b and ground is to be the same at the two ends of the transmitted band, 76 mc and 81.5 mc respectively.

a. Show that the sleeve length should be $\lambda/4$ at the arithmetic, rather than the geometric, mean of the band.

b. What will be the impedance between terminal b and ground at the ends of the band?

c. Would a Balun using this ratio of tube diameters give better performance on channel No. 13 than on channel No. 5? Explain.

13-1. Verify eq. (13-18).

13-2. Sketch typical wave forms for each of the monitoring positions shown in Fig. 13-13. Assume the monitor sweep frequency to be 15.75 kc in each case.

13-3. Sketch the wave form of the composite video signal as it would appear on a monitor that has a sweep operating at frame frequency.

13-4. Compare the expressions for power gain of grounded-plate and grounded-grid amplifiers operating in Class A. Neglect interelectrode capacitances. Compare these results for the amplifiers operating in Class C.

13-5. Compare the relative merits of the synchronized clamping circuit and the diode clamper. Discuss the feasibility of designing a synchronized clamper employing a single triode driven from a synchronizing source. State in what respects its operation would be inferior to that of the circuit of Fig. 13-14.

13-6a. Design a 300-ohm constant-resistance network.

b. It is determined that the high-wattage resistor used in the network has a shunt capacitance of $8 \mu\mu\text{f}$. Which form of the network could be used to the best advantage? Explain.

c. In general will compensation be required if this network is used as the plate load of a video amplifier? Explain.

13-7. Design the parallel wire line plate load for the push-pull output stage of a video transmitter which is to be operated on channel No. 12. The total output capacitance is $7 \mu\mu\text{f}$. Neglect the length of the plate leads within the tubes.

13-8. Design a broadband Balun for use in channel No. 5. The input coaxial lead has a characteristic impedance of 52 ohms. Choose constants so that the line-to-ground impedance is maximized.

14-1. A 72-ohm receiving antenna is connected to 30 feet of 72-ohm coaxial cable. Calculate the available noise power available at the output end of the cable which has an attenuation of 2.4 db. per hundred feet.

14-2. Calculate the available power gain for the grounded-plate, grounded-grid, and grounded-cathode stages. Neglect capacitances and grid leak resistances. From these results, which of the three connections would you choose for input stage of a receiver? Does this choice confirm the results of Chapter 14?

14-3. Verify eqs. (14-38) through (14-42).

14-4. Calculate the available power gain and noise figure for a grounded-cathode amplifier stage employing a 6J6 and the following constants:

$$6J6: \quad I_b = 8.5 \text{ ma}, \quad g_m = 5.3 \times 10^{-3} \text{ mho}, \quad r_p = 7100 \text{ ohms.}$$

$$R_A = 300 \text{ ohms}, \quad R_i = 300 \text{ ohms}, \quad R_L = 1000 \text{ ohms.}$$

Assume that conjugate impedance matches are maintained at all junctions.

14-5. Repeat problem 14-4 but replace the 6J6 by a 6AG5 which has the following constants:

$$I_b = 5.5 \text{ ma}, \quad I_{c2} = 1.6 \text{ ma}, \quad g_m = 4.75 \times 10^{-3} \text{ mho.}$$

$$r_p = 300,000 \text{ ohms}$$

14-6. Repeat problem 14-4 for the grounded-grid and grounded-plate connections.

14-7. Repeat problem 14-5 for the grounded-grid and grounded-plate connections.

14-8a. Discuss the possibility of lowering the noise figure of a receiver by deliberately mismatching the antenna transmission line at the receiver input by means of a lossless transformer.

b. Why is such a procedure generally of little use in television work?

14-9. A single-ended R-F stage is to be designed. Compare the relative merits of the 6J6 triode and the 6AG5 pentode for use in the stage.

14-10. Discuss reasons for designing the R-F stage to be as narrow-band as possible. What compromises have to be made because of transient response requirements?

14-11. Discuss possible problems due to Miller effect in triode R-F stages.

14-12. In what television bands is image interference from the commercial FM band (88 to 108 mc) most pronounced in a television receiver which employs a 26.4-mc visual I.F. frequency?

14-13. Explain how FM interference can produce black bars on the CRT of a television receiver.

14-14. A television receiver employs a 26.4-mc visual I.F. frequency. To what frequency should the plate load of the converter be turned? Stagger tuning is not used.

14-15a. Design a series-type trap to reject the 21.25-mc component in a video I.F. amplifier system.

b. Design a shunt-type trap employing two condensers to reject the same frequency.

c. Derive an expression for the degree of reject in decibels of the 21.25-mc component relative to the 25.75 component.

14-16. Explain why an AFC circuit is not necessary to provide noise immunity for the vertical sweep system.

15-1. Using the approximation $\omega + \omega_o \approx 2\omega$, show that eq. (15-1) reduces to the arithmetic symmetry form. (See Footnote 1, Chapter 15.)

15-2. Derive an expression for the half-power bandwidth of a high- Q single-tuned stage in terms of the circuit constants.

15-3. Three synchronously tuned, cascaded amplifier stages are to have an overall bandwidth of 4 mc. The center frequency is 25 mc. 6AK5's are used.

$$g_m = 5000 \text{ micromho}, \quad r_p = 340,000 \text{ ohms}, \quad C = 12 \mu\text{mf.}$$

- a. Design the component stages.
- b. Calculate the over-all gain of the amplifier.
- c. The plate supply voltage for each stage is 120 volts. What maximum input voltage may be used with the amplifier?

15-4. Discuss the practical troubles which might arise in applying eqs. (15-1) and (15-4) to a tuned amplifier that uses a triode.

15-5. What gain may be realized from a single-tuned amplifier stage employing a 6AK5 and having a half-power bandwidth of 3.5 mc? Assume that the following stage also employs a 6AK5.

15-6. A four-stage video I.F. amplifier consists of two groups of staggered pairs, the groups being synchronously tuned. The over-all bandwidth is 4.5 mc and the center frequency is 24 mc.

- a. Design the four stages using 6AK5's.
- b. Calculate and plot the over-all gain as a function of frequency.
- c. What video I.F. should be used with this amplifier? Explain.
- d. Design appropriate traps to give the response curve the proper shape.

15-7a. Design a staggered quadruple employing 6AK5's to have the same selectivity function as the amplifier in problem 15-6.

b. What advantages are provided by the staggered quadruple amplifier over the other design with regard to stability, gain, and alignment?

15-8. Figure 12-8b shows the ideal shape of the video I.F. pass characteristic. For what integral value of n does the selectivity function of eq. (15-38) most nearly approach the ideal? Assume that trap circuits may be used to help shape the curve.

16-1. Derive eq. (16-4a).

16-2. Explain the bending of a wave in terms of velocity if the refractive index of the atmosphere varies with altitude.

16-3. A television receiving antenna is located 15 miles from a transmitting antenna which radiates a peak signal of 12 kw in channel No. 6. The center of the transmitting antenna is 320 feet above the ground. At what height will the receiving antenna be in a field of 600 microvolts per meter?

16-4. A ghost is received on a television receiver. Discuss how you might tell if it is due to mismatch in the antenna-transmission line-receiver system or to a multiple-path transmission effect. Assume that the transmission line is 40 feet long.

16-5a. Calculate the lengths required in the antenna system of Fig. 16-8.

b. If the receiver to be used has a nominal input resistance of 300 ohms, suggest two methods of overcoming the mismatch problem at the receiver input.

16-6. Design a 300-ohm Yagi array for channel No. 13.

16-7. A receiver is located 4 miles from Station A at a relative azimuth of 0° . Stations B and C are located approximately 45 miles away at a relative azimuth of 10° . The channel assignments are as follows:

| <i>Station</i> | <i>Channel</i> |
|----------------|----------------|
| A | 11 |
| B | 4 |
| C | 7 |

a. Specify the length and orientation of a folded dipole to give a good compromise on reception from the three stations.

b. Recommend one form of array to better the reception. Give reasons for your choice.

16-8. A receiver is located approximately midway between two transmitters of equal power operating in channels No. 5 and No. 6. The airline distance between the transmitters is approximately 100 miles. Recommend a receiving antenna array which will allow either station to be received with a minimum of adjacent channel interference.

16-9. It is found that the input stage of a receiver located close to a transmitter is being overloaded. How should the situation be remedied? Assume the transmission-line and receiver input impedance are both 300 ohms.

16-10. Discuss the conditions under which a rhombic antenna may be used to advantage in television reception.

16-11. Draw the block diagram of a slave unit to be used with a common-video master distribution system.

16-12. An apartment house has a master antenna and distribution system which is designed to feed six 300-ohm-input receivers. No amplifiers are used. Each set is bridged across the main 300-ohm feeder through two 750-ohm resistors, one on each side of the line.

a. Sketch the distribution system.

b. Neglecting cable losses, calculate the signal required at the antenna to deliver a 50- μ volt signal to each receiver.

c. Calculate the isolation in decibels between any two receivers. Neglect cable losses.

d. What maximum local oscillator voltage may be tolerated at the input terminals of any receiver so that the signal-to-oscillator voltage ratio at any other receiver is at least 10 : 1? Assume a 50- μ volt signal voltage.

18-1. Will equal intensities of the three primaries shown in Fig. 18-5 reproduce white? If not, what relative intensities are required?

18-2. Using Fig. 18-4, determine the relative intensities of the three primaries

$$X = 0.1, Y = 0.11; \quad X = 0.17, Y = 0.8; \quad X = 0.6, Y = 0.39;$$

to synthesize the color $X = 0.3, Y = 0.4$.

18-3. Explain what primaries should be used in the Skiatron system of color reproduction.

18-4. Consider the possibility of using time multiplex with the Skiatron color system.

18-5. Discuss the feasibility of using the color switching scheme of Fig. 18-10*b* with a switching rate higher than the field frequency. Consider the problem for camera tubes of both the instantaneous and storage types.

18-6. Tabulate the relative merits and disadvantages of the several color synthesizing systems described in section 18-11.

18-7. Discuss means for utilizing a tricolor kinescope in a field-sequential color system. Draw a block diagram of a receiver utilizing the three-color tube.

18-8a. Draw block diagrams of a simultaneous color system which uses the principle of "mixed highs."

b. Sketch channel diagrams for the system indicating reasonable bandwidths for each signal component.

18-9. In the C.B.S. field-sequential color system, color switching after each line is not feasible due to the reduction in storage time (see section 18-17). Why is this not a limiting factor in the C.T.I. color system?

18-10. At what speed must the filter disk of Fig. 18-10*b* rotate in order to conform to the C.B.S. color standards?

18-11. Explain how a 12-cycle-per-second flicker may be present in a C.B.S. color receiver.

18-12. Explain how a C.B.S. color receiver would have to be modified to receive a conventional black-and-white program.

18-13. Explain how the spacing between the three images appearing on the face of a camera tube in the C.T.I. system may be calculated to maintain the normal horizontal flyback relationships required for compatibility.

18-14. Devise an electronic system for color switching in the R.C.A. color system. It is suggested that suitably derived signals be used for gating three channels.

18-15. Summarize the means by which dot interlace increases the resolution of a picture transmitted over a fixed bandwidth.

18-16. Compare the reproduction of a dot-interlace color picture by a tricolor kinescope and by a Trinoscope assembly.

AUTHOR INDEX

- Albert, A. L., 42
Alter, R. S., 479
Applegarth, A. R., 414
Arguimbau, L. B., 278
- Ballard, R. C., 207, 724
Barco, A. A., 232
Barrett, R. E., 354
Bartelink, E. H. B., 587
Battison, J. H., 713
Baum, R. F., 629
Beam, R. E., 497
Bedford, A. V., 20, 32, 243, 302, 311,
312, 381, 412, 685
Beers, G. L., 685
Beers, Y., 549
Bereskin, A. B., 289
Bingley, F. J., 701
Blewett, J. P., 433
Boothroyd, W., 761, 771
Bowie, R. M., 62
Bradburd, E., 479
Brainerd, J. G., 659
Bronwell, A. B., 497, 718
Brown, G. H., 467, 515
Brown, W., 439
Burnett, C. E., 186
- Carter, P. S., 653
Cawein, M., 147
Chambers, T. H., 730
Clark, E. L., 601, 615
Clarke, J. G., 397
Cohen, R. M., 562
Cornell, J. A., 600
Cruft Electronics Staff, 114, 295
- de Gier, J., 615
Deloraine, E. M., 761
De Vore, H. B., 192
Dimmick, G. L., 705
Dishall, M., 629
Di Toro, M. J., 277
- Dome, R. B., 604
Downes, L. C., 688
Dupuy, J., 196
Dyer, J. N., 14, 169, 701
- Easton, A., 301, 416
Eddy, W. C., 196
Engstrom, E. W., 21, 24, 44, 374, 685
Epstein, D. W., 8, 44, 64, 113, 188, 215
Erde, B., 687
Evenitt, W. L., 493, 649
- Farnsworth, P. T., 208
Fink, D. G., 17, 23, 84, 365, 389, 403,
484, 489, 651
Finneburgh, L., II., 661
Flory, L. E., 199
Foster, D. E., 309, 591
Frankel, S., 504
Fredendall, G. L., 302, 311, 312, 381,
446, 724
Freeman, R. L., 297
Friend, A. W., 117, 146
Fubini, E. G., 529
- Gadsden, C. P., 559
George, R. H., 41
Glauber, J. J., 504
Goldberg, H., 549
Goldman, S., 277, 301
Goldmark, P. C., 14, 28, 169, 535, 701,
702, 730
Goodman, B., 668
Goodman, M. M., 354
Grammer, G., 668
Gray, F., 150
Greenlee, L. E., 669
Greenstein, P., 626
Griffin, D. A., 661
Guillemín, E. A., 137, 314, 430
- Hallmark, C. E., 418
Hamilton, G. E., 531

- Handel, R. R., 254
 Harding, C. F., 41
 Harris, W. A., 561
 Hart, W. L., 417
 Heim, H. J., 41
 Herold, E. W., 562
 Hodgman, C. D., 704
 Hollywood, J. M., 14, 701, 730
 Huggins, W. H., 629
- Iams, H., 192, 245, 248
- Jacobs, D. H., 617
 Janes, R. B., 254
 Johnson, E. O., 566
 Johnson, J. B., 201
 Johnson, R. E., 254
- Kahnke, J., 587
 Kallman, H. E., 432, 679
 Kamen, L., 681
 Kay, M. S., 615, 744
 Kell, R. D., 20, 24, 32, 44, 216, 243,
 381, 446, 724
 Kimball, C. N., 288
 Kline, M. B., 341
 Koek, W. E., 88
 Kolster, F. A., 674
 Kusonose, Y., 146
- Landis, D. O., 615
 Larsen, M. J., 327, 586
 Law, H. B., 254
 Lee, H. W., 12, 41, 621
 Liebscher, A., 115
 Lord, G. T., 571
- Macnee, A. B., 559
 Maloff, I. G., 8, 44, 113, 215, 221
 Masters, R. W., 541
 Mautner, L., 287, 334
 Mautner, R. S., 613
 Meagher, J. R., 657
 Merrill, L. L., 586
 Mertz, P., 150
 Millman, J., 88
 Mills, B. Y., 341
 Milwitt, W., 439
 M.I.T. Electrical Engineering Staff, 46
 M.I.T. Radar School Staff, 121, 543,
 667
 Montfort, R. A., 440
- Moon, P., 188, 693
 Morgan, J. M., 342
 Morrison, J. F., 543
 Morton, G. A., 54, 181, 196, 199, 245,
 381
 Moskowitz, S., 286
 Mulligan, J. H., Jr., 334
 Mumford, W. W., 544
- Noe, J. D., 325
 North, D. O., 203
- Odessy, P. H., 416
 Olsen, R. K., 531
- Paine, R. C., 483
 Palmer, R. C., 287
 Peddie, W., 693
 Pensak, L., 64
 Piore, E. R., 14, 701, 730
- Racker, J., 286, 479
 Radio Research Laboratory Staff, 540
 Rankin, J. A., 309, 591
 Reese, H. M., 694
 Reeves, J. J., 730
 Reich, H. J., 422
 Ridenour, L. R., 621
 Rinia, H., 615
 Roberts, W. van B., 663
 Robinson, D. M., 41
 Rose, A., 248, 254
 Rosenthal, A. H., 621
 Rubel, J. H., 433
 Ryder, J. D., 193, 539
- Sanabria, U. A., 30
 Sanders, R. W., 12, 210, 343
 Schade, O. H., 613
 Sehantz, J. D., 297
 Schlesinger, K., 61, 119, 601
 Schoenfeld, E., 439
 Schottky, W., 200
 Schroeder, A. C., 207, 603, 724
 Schultz, J. L., 507
 Seal, P. M., 312
 Seeley, S., 88
 Seeley, S. W., 288, 604
 Seeley, W. J., 135
 Serrell, R., 14, 657, 730
 Shannon, C. E., 763
 Shenk, E. R., 93

- Silver, M., 414, 543
Smith, J. P., 412
Sobel, A. D., 561
Soller, T., 54, 115
Somers, F. J., 440
Spielman, S. C., 132
Starner, C. J., 499
Starr, M. A., 54, 115
Stewart, J. A., 568
Sutro, P. J., 529
Sziklai, G. C., 24, 207, 216, 724
- Terman, F. E., 273, 319, 467, 493, 532,
568
Trainer, M. A., 20, 32, 243
Trevor, B., 653
- Valley, G. E., Jr., 54, 115, 287, 333
van Alphen, P. M., 615
Van Dyck, 17
- Von Ardenne, M., 350
- Wallenstein, J. P., 504
Wallman, H., 287, 333, 559, 629
Ware, P., 567
Watters, R. L., 587
Webb, R. C., 342
Weimer, P. K., 254
Wendt, K. R., 603, 724
Wheeler, H. A., 277
Wiggin, J. F., 688
Woodward, O. M., Jr., 670
Wright, A., 565, 599
- Yunker, E. L., 543
- Zaharis, G., 439
Zeidler, H. M., 325
Zeluff, V., 204, 342, 719
Zworykin, V. K., 32, 41, 44, 54, 181,
196, 199, 216, 245, 381, 615

SUBJECT INDEX

- A.G.C. 601
 - keyed 603
- Ampere turns, magnetic deflection 59
- Amplification 266
 - relative 271
- Amplifier:
 - balanced 596
 - cascaded, synchronously tuned 627
 - cascode 559
 - clamping 328
 - reduction of grid swing 332
 - deflection:
 - balanced 349
 - bandwidth 350
 - figure of merit 571
 - gain-bandwidth factor 626, 628
 - gain-bandwidth product 288
 - grounded-grid 499
 - class A 501
 - class C 501
 - high frequency compensation, summary 309
 - ideal:
 - delay time 284
 - low-pass 284
 - overshoot 285
 - ring 285
 - I.F. 581, 609, 647
 - rise time 285
 - Küpfmüller's rule 286
 - line 507
 - low frequency compensation 316
 - bias circuit 322
 - compensating network 319
 - coupling circuit 323
 - summary 324
 - modulated 520
 - paraphase 349
 - pentriode 325
 - preamplifier:
 - iconoscope 232
 - image orthicon 263
- Amplifier (*Cont.*):
 - single-tuned 622
 - bandwidth 624
 - gain-bandwidth factor 626
 - selectivity function 626
 - square wave testing:
 - high frequency 309
 - low frequency 325
 - stabilizing 507
 - stagger tuned Ch. 15, 622 f
 - transient response:
 - calculation of, by square wave 312
 - low frequency 325
 - video:
 - cascaded 287
 - compensation:
 - Freeman-Schantz 297
 - shunt 289
 - simple 292
 - transient response 300
 - vector diagrams 299
 - direct coupling 515
 - distortion:
 - amplitude 267
 - delay 267
 - frequency 267
 - non-linear 267
 - phase 267
 - feedback 334
 - mid-band 334
 - low-band 335
 - high-band 336
 - figure of merit 288
 - high level 511
 - m*-derived shunt peaking 302
 - parallel tubes in 511
 - resistance coupled 269
 - mid-band 270
 - high-band 271
 - low-band 273
 - transient response 279

- Amplifier (*Cont.*):
 video (*Cont.*):
 series compensation 303
 transient response 306
 series-shunt compensation 307
 transient response 307
 steady-state requirements 265,
 267, 269, 275
 transient response 275
- Antenna:
 batwing 541
 broadbanding 538
 crossed dipoles 533
 crossover network 661
 dipole 532, 658
 broadbanding 660
 folded 663
 directivity 532
 fan 669
 lightning protection 665
 master system
 video distribution 677
 R-F distribution 679
 parasitic 666
 radiation pattern 532
 receiving Ch. 16, 648 ff
 requirements 657
 rotatable 674
 spikes 670
 stacked 535, 671
 superturnstile 541
 transmitting requirements 531
 turnstile 535
 unipole 664
 V 669, 671
 Yagi 668
- Aperture 37
 numerical, of a lens 193
- Aperture distortion 163
 of circular aperture 176
 one-dimensional 165
 two-dimensional 174
 with electron beams 181
- Aspect ratio, checking 392
- Azimutrol 675
- Back porch (Fig. 11-11) 408, 507
- Balun 527
 broad-band 528
 narrow-band 483
- Bandwidth:
 approximate for television transmis-
 sion 18
 checking 394
- Bazooka (*see* Balun)
- Bias lighting 238
- Black-negative signal 187
- Blacker-than-black 356
- Blanking 411
 camera 370
 composite signal 346
 interval 70
 line 370
 ratio 70
- Blocking oscillator 112
 controlled 600
 design values 114
 hold control 115
 synchronization 115
- Brightness (*see* Luminance)
- Brightness control 352
- Camera tubes Ch. 6, 185 ff
 color response 198
 iconoscope 216
 image dissector 208
 image iconoscope 246
 image orthicon 254
 monoscope 185
 orthiconoscope 248
 static image generator 43, 185
- Candle 188
- Capacitance, of resistor 513
- Carrier-difference reception 604 (*see*
 also Receiver, intercar-
 rier)
- Cascode amplifier 559
- Cathode follower:
 available power gain 558
 impedance matching 338
 input admittance 233
 noise figure 558
 voltage gain 560
- Cathode-ray tube:
 aluminized screen 64
 crossover point 47
 decay characteristic Fig. 2-10, 28
 decay time 8
 deflection factor 52
 deflection sensitivity 51
 typical values 52

- Cathode-ray tube (*Cont.*):
 electron gun 47
 phosphor, decay time 8
 tricolor 719
 velocity-controlled 41
- Channels, R-F 371
- Chromoscope 718
- Clamping 241, 328, 352
 in sweep generator 89
 keyed 507
 synchronized 507
- Closed systems Ch. 8, 342 ff
 defined 343
 Purdue 43
 type I 343
 type II 352
 type III 354
- Coaxial line:
 characteristic impedance 471
 reactance 471
- Color:
 analyzing systems 709
 brilliance 695
 chroma 695
 complementary 705
 eye response 694
 filter 704
 dichroic mirror 705
 hue 695
 matching 695
 primaries 697
 reproducing systems 715
 chromoscope 718
 Telechrome tube 718
 tri-color kinescope 719
 Trinoscope 715
 saturation 695
 synthesis:
 additive 706
 subtractive 706
 transmission systems 713
 triangle 697
 trichromatic coefficients 699
- Color television:
 basic components (*see* Color)
 C.B.S. field-sequential 730 ff
 compatibility 723
 C.T.I. line-sequential 741 ff
 C.T.I. segmental-sequential 754 ff
 mixed highs 729
 R.C.A. dot-sequential 744 ff
- Color television (*Cont.*):
 R.C.A. simultaneous 724 ff
 Skiatron system 707
- Communication channel, requirements
 of 7
- Comparator circuit 422
 adjustment 424
 in receiver 598
- Compatibility 723
- Composite video signal 355
 amplitude components, 396
 in closed system 356
 wave-form components 396
- Constant-*R* network 467, 513
- Contrast:
 checking 392
 control 352
- Counter circuit 414
 analysis 415
 discharge circuits 419
- Critical damping 134
- Critical flicker frequency, for CRT,
 Fig 2-9, 27
- Crossover network:
 antenna 661
 constant-*R* 467, 513
- Damper tube (*see* Reaction scanning)
- Damping:
 shunt-resistance 129
 resistance 133
 advantages 136
 diode 139
 advantages 142
- Dark spot:
 iconoscope 226
 correction 227
 image orthicon 262
- d.c. insertion 239
- Deflection:
 crossed *E* and *H* fields 250
 crossed *H* fields 252
 electrostatic 48 ff
 balanced 349
 compared to magnetic 61
 flared plates 50
 positioning 596
 magnetic 55 ff
 damping (*see* Damping)
 reaction scanning 142, 596

- Deflection (*Cont.*):
 - magnetic (*Cont.*):
 - receiver circuits 591
 - yoke construction 58 ff
 - types 44
- Degeneration, reduction of capacitance 233
- Delay network 429
 - artificial line 430
 - multivibrator 433
- Dichroic mirror 705
- Differentiating circuit 357
 - analysis 397
- Diode line 502
- Diplexer 529
- Dipole (*see* Antenna)
- Directional coupler 544
- Directivity pattern (*see* Antennas)
 - dipole 659
- Discharge tube:
 - thyatron 87
 - vacuum 89
 - plate resistance 89
- Discriminator (*see* Comparator)
- Dissipation factor 625
- Dome system 604 (*see also* Receiver, intercarrier)
- Dot generator 443
- Dot interlace 744, 766, 771
- Duality principle 137
- Dynode 211
- Efficiency, storage:
 - orthicon 253
 - iconoscope 224
- Electron:
 - charge 45
 - lens:
 - electrostatic 47
 - magnetic 52
 - mass 45
 - multiplier 210, 259
 - volts 195
- Elemental area 4, 219
- Emission, in iconoscope 220
- Equalization 407
 - pulses 410
- Equilibrium voltage:
 - high-velocity scanning 222
 - low-velocity scanning 248
- Eye:
 - retina 21
- Eye (*Cont.*):
 - visual acuity 21
- Facsimile:
 - compared to television 7
 - Times system 8 ff
- Ferry-Porter law 25
- Field:
 - definition 31
 - even 72
 - frequency 31
 - odd 72
 - period, interlaced scan 76
 - strength 650
- Figure of merit:
 - picture 4, 219
 - minimum 5
 - tube 288, 625
- Filter:
 - color 704
 - constant- k 430
 - distributed constant 472, 476
 - low-pass 430
 - m -derived 432, 479
 - receiver attenuation 461, 462
 - tolerance 464
 - transmitter attenuation 460, 462
 - vestigial sideband
 - crossover 467
 - m -derived 479
 - notching 477
 - off-center tuned 466
 - type A 467
 - type B 477
- Flicker 25
 - critical frequency 25
 - interline 31
 - overall 31
- Flux:
 - radiant 188
 - luminous 188
- Flyback:
 - in magnetic deflection 128
 - minimum 128
 - interval 22
 - power supply 147
 - voltage multiplier 149
 - ratio 23
 - horizontal 70
 - vertical 69
 - resonance 139

- Flying spot scanning:
 electronic 204, 710
 mechanical 39
- Focusing:
 electrostatic 45
 control 48
 magnetic:
 short coil 52
 construction 53
 control 54
 long coil 213
- Fourier integral 283
- Fourier series:
 exponential form 164
 in scanning analysis 151
 one-dimensional 151
 two-dimensional 152
 components 154
 in transient response 312
 of saw tooth 350
- Frame 24
 frequency 18, 24
 motion picture 25, 683
 period 68
- Frequency:
 bands, FCC designation 648
 channels, television 371
 critical flicker 27
 crossover 469
 constant- R network 514
 cutoff 292
 field 75, 76
 frame 18, 24
 geometric mean 586
 half-power:
 lower 272
 upper 274
 I.F.:
 aural 579
 visual 579
 line 75, 76
 negative 284
 normalized 272, 291, 625
 power line 35
 top video 20, 292
- Frequency divider 345
- Frequency modulation, effect on picture 582
- Gain 266
- Gain-bandwidth factor:
 single-tuned amplifier 626
- Gain-bandwidth factor (*Cont.*):
 staggered n -uple 628
 staggered pair 628
- Gain-bandwidth product 288
- Gamma 253
- Ghost, location 655
- Gibbs' phenomenon 314
- Grounded-grid amplifier:
 class A 501
 available power gain 558
 noise figure 558
 voltage gain 560
 class C 501
 neutralization 499
- Hartley law 723
- Iconoscope, 41, 44, 216 ff
 bias lighting 238
 dark spot 227
 correction 227, 237
 d-e insertion 239
 efficiency, storage 224
 electron bombardment 220
 equilibrium potential 222
 electrostatic type 343
 equivalent circuit 224
 keystone correction 241
 illuminance levels 245
 mosaic 217
 output current, theoretical 219
 output voltage 220
 preamplifier 232
 storage 216, 218
- I-F amplifier 581, 609, 647
- I-F frequency:
 aural 579
 visual 579
- Illuminance 189 (*see* Illumination)
- Illumination 189
 apparent 27
 motion pictures 28
 photocathode 192
- Image dissector 44, 208 ff
 output current 209
 magnetic focusing 213
 signal to noise ratio 213
 multiplier 210
 noise 212
- Image iconoscope 44, 246
- Image intensification 246, 255

- Image orthiconoscope 44, 254 ff
 characteristics 260
 electron multiplier 259
 noise 260
 image intensifier 255
 mesh 256
 mosaic 257
 preamplifier 263
 shading 262, 264
 target 257
- Impedance matching, cathode follower 338
- Impulse generators 92
 blocking oscillator 112
 multivibrator 93 ff
- Inductuner 566
- Integrating circuit 357
 analysis 399
- Intercarrier reception (*see* Receiver)
- Interference, co-channel 654
- Interlace (*see* Scanning)
 color 741
 dot 744, 766, 771
- Ion spot 62
- Ion trap 65
- Keystone correction 241
 generator 243
- Kinescope, tricolor 719
- Küpfmüller's rule 286
- Lambert's law 189
- Lens:
 electron:
 electrostatic 47
 magnetic 52
 optical:
 equation for illuminance 193
 numerical aperture 193
- Line:
 active 70
 frequency 75, 76
 inactive 70, 412
 interval 72
 width 72
- Linearity:
 control 145
 departure from 82
 testing 392
- Lock and key 13
- Lumen 188
- Luminance 190
- Luminosity curve 188
- Luminous emittance 191
- Luminous intensity 188
- Magnetic deflection 116 ff (*see also* Damping)
 direct-coupled 116
 grid voltage for 119
 minimum flyback time 137, 777
 (Prob. 4-19)
 transformer coupled 117
 trapezoidal generator 121
 design procedure 126
- Mangin mirror 617
- Millilambert 192
- Mixed highs 729
- Mixer:
 frequency 571
 voltage adder 347, 348, 360
- Modulation:
 amplitude 446
 grid 494
 plate 492
- Modulation index 447
- Monitor, driven 370
- Monoscope 185
- Mosaic 217
 equilibrium potential 222
 two-sided 257
- Motion pictures:
 flicker reduction 29
 frame frequency 25
 illumination levels 28
 projector:
 16 mm. 687
 shutterless 688
 35 mm. 685
 standards 683
 televising Ch. 17, 682 ff
- Multivibrator 93 ff
 analysis 93
 circuits 111
 delay 433
 design curves Fig. 4-16, 101
 design procedure 100
 synchronization 108, 110
- Neon tube:
 picture reproducer 38
- Neutralization:
 grounded-grid stage 499
 R-F stage, receiver 559

- Nipkow disk 38
- Noise:
- available power 550
 - available power gain 551
 - bandwidth 551
 - figure 552
 - generator, equivalent circuits 549
 - image orthicon 260
 - flying spot scanner 206
 - Johnson 201
 - multiplier 212
 - shot 199
 - thermal agitation 201
 - tube:
 - equivalent resistance
 - pentode 204
 - triode 203
 - flicker 203
 - partition 203
 - shot 202
- Noise figure:
- cathode follower 558
 - grounded-cathode stage 558
 - grounded-grid stage 558
 - minimum 560
- Open system, defined 352
- Orthiconoscope 44, 248
- characteristics 252
 - deflection system 250
 - low-velocity scanning 248
- Oscillator:
- Colpitts 581
 - crystal 498
 - Hartley 581
 - receiver 577
 - stability requirement 577
 - resonant line 496
 - Ultraudion 581
- Overlap, effect on line structure 378
- Overshoot 285, 287
- Pairing 378
- Parallel operation of tubes 511
- Partition effect (*see* Noise)
- Pedestal 241, 321
- Perveance 146
- Phase shift 266
- relative 271
 - network 15
 - analysis 443
- Phosphor, decay time 8
- Photoemission:
- color response 195
 - Einstein's equation 194
 - empirical laws 193
 - threshold frequency 194
 - work function 194
- Photometry:
- illuminance 189
 - inverse square law 190
 - Lambert's law 189
 - luminance 190
 - luminosity curve 188
 - luminous emittance 191
 - luminous flux 188
 - luminous intensity 188
 - radiant flux 188
 - units of 192
- Phototube, output voltage 198
- Pickup:
- requirements 6 (*see also* Camera tubes)
 - flying spot 204
 - noise in 206
 - phosphor decay time 207
- Picture 24
- frequency 24
 - to reconstitute motion 8
 - repetition rate 24
 - shape:
 - circular 14
 - rectangular 17
 - square 17
 - transmission Ch. 1, 1 ff
 - parallel 5
 - sequential 6
- Pitch 377
- Plate load, of modulated amplifier 521
- parallel line 524
- Power:
- in negative transmission 486
 - in positive transmission 487
- Power gain:
- available:
 - cathode follower 558
 - grounded-cathode stage 556
 - grounded-grid stage 558
 - grounded-grid amplifier:
 - class A 501
 - class C 504
- Power supply, high voltage 147
- Projection system:
- reflective 617

- Projection system (*Cont.*):
 refractive 615
 Skiatron 620
- Propagation, V.H.F. 649
 anomalous 654
 field strength 650
 ghosts 655
 line-of-sight range 650
- Protelgram 619
- Pulse cross 443
- Pulsewidth measurement:
 sawtooth sweep 440
 sinusoidal sweep 441
- Purdue system 41
- Q 523, 623
- Random scan 354
- Raster 67
- Rayleigh-Carson theorem 658
- R-C circuit:
 charge 78, 79
 discharge 78
 time constant 79
- Reactance curves:
 open-circuited line 471
 short-circuited line 471
- Reactance tube 421
- Reaction scanning 142, 596
 damper tube 144
 diode type 146
 linearity control 145
 voltage boosting 147
- Receiver Ch. 14, 546 ff
 A.F.C., 579
 horizontal sweep 597
 A.G.C. 601
 bandwidth 562
 deflection circuits 591
 A.F.C. 597
 front end 372, 561
 continuous tuning 566
 functional diagram 548
 I-F amplifier 580
 characteristic 584
 double-tuned 584
 stagger-tuned 587
 trap circuits 587
 intercarrier 604
 circuit 607
 principle 605
 minimum noise figure 560
- Receiver (*Cont.*):
 mixer 571
 oscillator 578
 projection 615
 R-F stage 561
 neutralization 559
 tuning 562
- Reproducing system, requirements of 6
- Resolution:
 effect of line overlap 376
 factors governing 374
 figure of merit 374
 horizontal 383
 checking 394
 over-all 388
 testing 392
 vertical 381
 checking 392
- Ring 285
- Ripple, effect on scanning 32 ff
- Rise time 285
- R.M.A. signal Fig 11-11, 408
- Roc bird 24
- Sampling, concept of 5
- Saw-tooth voltage, generator 80, 115
 departure from linearity 82, 84
 trigger tubes 87
 conversion ratio 83, 84
 constant current 86
- Saw-tooth wave, spectrum 350
- Scan:
 horizontal 23
 interval 22
 vertical 23
- Scanning:
 defined 7
 effect on picture reproduction Ch. 5,
 150 ff
 electronic 44 ff
 development 41
 linearity, testing 392
 mechanical:
 facsimile 10
 flying spot 40
 mirror 40
 Nipkow disk 38
 pitch 377
 trace 14
- Scanning generators Ch. 4, 75 ff
- Scanning geometry:
 bidirectional 21

- Scanning geometry (*Cont.*):
 direction 21
 effect of ripple 21
 helter skelter 71
 interlaced 30, 72
 even-line 76
 odd-line 76
 need for 389
 progressive 30, 67
 random 71, 354
 speed 21
 spiral 15
 standards 371
- Scanning lines:
 intensity 376
 optimum number Ch. 10, 374 ff
 pairing 378
- Scophony system 11, 41
- Secondary emission ratio 211
 aluminum 186
 carbon 186
 nickel 222
- Serration 405
- Shading:
 iconoscope 227, 237
 generator 230
 image orthicon 262, 264
- Shot noise:
 emission limited 199
 voltage limited 199, 203
- Sidebands 447
- Signal-to-noise ratio:
 iconoscope 233
 image dissector 213
 minimum usable 199
 receiver 552
- Skiatron 620
 for color reproduction 707
- Snell's law 46
- Spectrum:
 saw-tooth wave 350
 scanned image:
 interlaced scan 160
 progressive scan 157
 square pulse 429, 764
- Square wave:
 spectrum 429, 764
 testing:
 high response 309
 low response 325
- Stagger tuning Ch. 15, 622 ff
 staggered n -uple 638
- Stagger tuning (*Cont.*):
 staggered n -uple (*Cont.*):
 arithmetic form 645
 asymptotic forms 642
 gain-bandwidth factor 628
 selectivity function 628, 641
 staggered pair 629
 gain-bandwidth factor 636
 selectivity function 629
- Standards:
 committees for 366
 intercarrier reception 613
 picture Ch. 2, 13 ff
 receiver I.F. 579
 synchronizing signal 408
 transmission 370
- Standing wave ratio:
 defined 539
 measurement 543
- Static image generator:
 monoscope 185
 Purdue project 43
- Storage principle 216, 218
- Supersync:
 blacker than black 356
 closed system 356
- Sweep generator:
 saw-tooth 80, 115
 spiral 15
 trapezoidal 121
- Sync stretcher 488, 510
- Sync stripping 360
- Synchronization Ch. 11, 395 ff
 blocking oscillator 115
 multivibrator 108, 110
 sweep generator 92
- Synchronizing signal:
 composite 408
 development 402
 dimensions 408
 equalizing pulse 410
 horizontal 402
 separation:
 horizontal and vertical 356, 401
 sync and video 360
 serration 405
 testing:
 dot generator 443
 pulse cross 443
 saw-tooth sweep 440
 sinusoidal sweep 441
 vertical 403

- Sync generator 345, 412 ff
 comparator circuit 422
 counter circuit 414
 delay network 429
 reactance tube 421
 requirements 412
 operation 434
 voltage adder 425
- Telecasting standards:
 R-F 371
 scanning 371
- Telecasting system:
 commercial Ch. 9, 365
 block diagram 368
- Telechrome tube 718
- Television, definition 1
- Test pattern 390, 391
 testing band width 394
- Thermal agitation (*see* Noise)
- Thyratron:
 deionization time 88
 trigger tube 87
 with feedback 87, 88
- Tilt 282, 326
- Time constant, R-C circuit 79
- Timing unit, interlaced scan 75
- Transconductance, conversion 576
- Transient response:
 calculation 312
 ideal amplifier 283
 ideal receiver 384
 R-C amplifier 279
 series peaking 306
 series-shunt peaking 307
 shunt peaking 301
 testing 309, 325
- Transmission:
 dot, Appendix, 761 ff
 double sideband 447
 negative 484
 positive 485
 receiver attenuation 461
 single sideband 451
 transmitter attenuation 460
 vestigial sideband 454
 filtering 465 ff
- Transmitter Ch. 13, 484 ff
 carrier frequency 648
 components:
- Transmitter (*Cont.*):
 components (*Cont.*):
 antenna 531
 coupling loop 527
 diplexer 529
 filter (*see* Filter)
 modulating section 519
 oscillator:
 crystal 498
 resonant line 496
 plate load 521
 R-F section 496
 transmission lines 527
 video section 506
 Du Mont 490
 General Electric 489
 modulation:
 grid 494
 plate 492
 monitoring 507
 requirements 489
 R.C.A. 491
- Trap circuits:
 absorption 590
 degenerative 589
 series 587
 shunt 587
- Trinoscope 715
- Ultraudion 580
- Unblank interval 70
- Unipole 664
- Unit function 277
- Utilization coefficient 19
- Utilization ratio 382
- Vestigial sideband transmission Ch. 12,
 446
- Voltage adder, analysis:
 graphical 427
 linear 426
- Voltage booster 147
- Voltage multiplier (*see* Flyback power
 supply)
- Wavelength, in picture Fourier compo-
 nents 156
- Work function 194
- White-positive output 253
- Young-Helmholtz theory 694

