MOTION PICTURE SOUND ENGINEERING

Motion Picture Sound Engineering

A Series of Lectures
presented to the classes enrolled in the
courses in Sound Engineering given by the

RESEARCH COUNCIL
of the
ACADEMY OF MOTION PICTURE ARTS AND SCIENCES
HOLLYWOOD, CALIFORNIA

SIXTH PRINTING

NEW YORK
D. VAN NOSTRAND COMPANY, INC.
250 FOURTH AVENUE

PRINTED IN THE UNITED STATES OF AMERICA

TYPOGRAPHY BY MAGOFFIN & DOYE, HOLLYWOOD, CALIF.

To the memory

of

Irving G. Thalberg

who

inaugurated the Research Council

M/s. Lakshmi Book Store, New Delhi on 6/11/57 —

PREFACE

In the period which has elapsed since the publication of the book "Recording Sound for Motion Pictures," by the Academy of Motion Picture Arts and Sciences in the year 1931, there have been many advances in the development of equipment and technique for recording sound for the motion picture.

Most of these advances have taken place as a result of the cooperative endeavor of a large number of technical experts, and this book covering current sound recording and sound reproducing practice results from a further extension of the technical cooperative idea as exemplified by the Research Council of the Academy of Motion Picture Arts and Sciences.

Although the Academy of Motion Picture Arts and Sciences has engaged in cooperative research ever since its organization in 1927, the Research Council as it now functions was organized in 1934, having at that time eight cooperative projects in the hands of eight separate Committees. Founded upon the recognition of the value of a group judgment based upon an orderly study of all of the facts, the Council was set up to coordinate technical problems within the industry, the entire effort being directed towards getting pictures of better quality upon the screen and getting them there at the lowest net cost and with the highest net efficiency.

In the Council the motion picture industry has available a smoothly working machinery to handle industry projects involving investigation beyond the scope of any one individual studio or company and of tackling problems which can be dealt with more effectively and more economically through cooperative action rather than by the individual companies working separately and thus duplicating their development effort and expense.

Since its reorganization in 1934, the Council's activities have steadily grown until at the present time there are thirty-six technical committees operating under Council sponsorship, investigating problems of sound recording, sound reproduction, projection, laboratory practice, film preservation, photography, lighting and set acoustics.

The work of all of the various technical committees functions through the Council proper, which consists of a chairman representing

the producing companies and one technical representative for each or the studios participating in this work—these being: John Livadary, representing Columbia Pictures Corporation; Douglas Shearer, representing Metro-Goldwyn-Mayer Studios; Farciot Edouart, representing Paramount Productions; John Aalberg, RKO-Radio Pictures; E. H. Hansen, 20th Century-Fox Film Corporation; Thomas Moulton, United Artists Studios; Homer Tasker, Universal Pictures Corporation; and Nathan Levinson, Warner Brothers-First National Studios.

The Council acts as the governing body directing the work and activities of the Committees, each of which in turn operates under the direction of a chairman who is selected from the staff of one or another of the participating companies and chosen for his particular qualifications in the field within which the Committee is to function.

The members of the Council, and the chairmen and members of the various committees all donate their time and knowledge to carry on this cooperative activity, receiving no compensation other than the satisfaction of having participated in a worthwhile industry activity and of having achieved results which without their participation would have been impossible.

We believe that this book, which has resulted from approximately two years' work by the Council's Committee on Industrial Education and the group of sound course instructors and authors who have prepared materials included in it, will become a valuable addition to the technical literature on the motion picture and we offer it to the industry with a feeling of great pride.

Wm Koenig

Chairman,
Research Council
Academy of Motion Picture Arts and Sciences.

Hollywood, California.
January 3, 1938.

ACKNOWLEDGMENTS

In presenting MOTION PICTURE SOUND ENGINEERING to the industry the Research Council of the Academy of Motion Picture Arts and Sciences acknowledges the cooperation of A. P. Hill, of Electrical Research Products, Inc., who prepared the lectures given in the two courses in the Fundamentals of Sound Recording, one in the fall of 1936 and the other in the spring of 1937, and Messrs. Fred Albin, of United Artists Studios, L. E. Clark, of the RCA Manufacturing Co., and John Hilliard and Harry Kimball of Metro-Goldwyn-Mayer Studios, who prepared the lectures presented in the Advanced Course in Sound Recording given during the spring of 1937.

In addition, the Council and the industry are indebted to Electrical Research Products, Inc., the Metro-Goldwyn-Mayer Studios, the RCA Manufacturing Company, and the United Artists Studios for their helpful cooperation in making it possible for these experts to undertake the responsibility of these Courses and perform this service to the motion picture industry.

The time intervening since the close of the last classes of the courses and the issuance of this book in its final form has not been idle. Early in this period a policy of revision and expansion was decided upon, in order that MOTION PICTURE SOUND ENGINEERING might, in its final form, deal as completely and authoritatively as possible with the subjects with which it is concerned.

Several chapters not presented originally to the sound course classes were written for inclusion in the book by Wesley C. Miller and Kenneth Lambert in order that the text material might completely cover the subject of motion picture sound recording and reproducing.

Exhaustive editing and criticism of one another's work by each of the group of authors, helped to shape the book into its final form.

The Research Council's Committee on Industrial Education under the Chairmanship of Dr. J. G. Frayne, assisted by Barton Kreuzer, Dr. Burton F. Miller, William Thayer and Ralph Townsend has served to guide the instructors, both in their presentation of the lecture material to the various classes, and in the preparation of the book itself.

The Council can hardly make adequate acknowledgment to those few individuals who, in addition to the authors, have throughout the

period of writing and editing of the book given of their time, thought and energy to help in its preparation. To these few, particular thanks are due: Wesley C. Miller, Nathan Levinson, and John Hilliard.

In addition to the authors and others mentioned above, the Council is indebted to the following who have helped in one way or another in the preparation of this book: W. E. Beatty of Warner Brothers-First National Studios, Barton Kreuzer, Wallace V. Wolfe, and Michael Rettinger of the RCA Manufacturing Company, and K. F. Morgan, Don Loye, W. H. Garrison, and R. W. Wight of Electrical Research Products, Inc.

In addition, acknowledgment is made to the Acoustical Society of America, the American Institute of Electrical Engineers, the Society of Motion Picture Engineers, the Institute of Radio Engineers, the Bell System Technical Journal, the Brush Development Company, Electrical Research Products, Inc., the RCA Manufacturing Company, P. Blakiston's Son & Co., Inc., McGraw-Hill Book Company, Inc., the University of Chicago Press, John Wiley and Sons, and Dr. John F. Blackburn for their cooperation in making available illustrations used throughout this text.

In addition, the editing and proofreading of the book has benefited to a very considerable degree by the able and conscientious assistance of William F. Kelley of the Research Council staff.

Gordon S. Mitchell

Manager,
Research Council
Academy of Motion Picture Arts and Sciences

January 3, 1938.

CONTENTS

CONTENTS

CONTENTS

FOREWORD

Early in 1936 the Research Council of the Academy of Motion Picture Arts and Sciences, recognizing the great advancements in the sound field during the previous few years and realizing the value to the industry of a completely trained personnel in the studios, appointed the Council's Committee on Industrial Education, under the chairmanship of Dr. J. G. Frayne, assisted by Barton Kreuzer, Dr. Burton F. Miller, William Thayer and Ralph Townsend, to investigate the needs for vocational education within the motion picture studios.

This committee subsequently arranged and conducted a Course in the Fundamentals of Sound Recording, which was attended by 100 selected studio sound department employees, divided into two classes of 50 each.

Both of these classes, meeting twice each week, were instructed by A. P. Hill of Electrical Research Products, Inc.

Enrollments in this first course in the Fundamentals of Sound Recording were limited to 100, each studio's quota being chosen by the director of that studio's sound department. The first announcement of the course calling for applications for enrollment to be submitted through the studio sound directors, resulted in the filing of 206 applications. In consideration of this unforseen popularity of the Council's educational program, it was decided to repeat the Fundamental Course again the next fall, under the direction of Mr. Hill, to an additional group of 50 sound technicians.

Concurrent with the second of the Fundamental Courses, the Advanced Course in Sound Recording, under the direction of Messrs. Fred Albin, L. E. Clark, John Hilliard and Harry Kimball was arranged for a group of 50 studio sound engineers consisting of those who had completed the first Fundamental Course, and others having the necessary previous education and experience.

This educational program has proven to be of great value to the 200 studio sound department employees who were enrolled, each of whom is of greater value to the industry because of the added knowledge gained from the Courses.

Stenographic transcripts were made of the lectures presented in these Courses, and material included in this book has been assembled from these transcripts.

FOREWORD

Any discontinuity in the text results from the fact that the book is made up of a series of lectures which were originally prepared for presentation to the sound course classes and subsequently edited by their authors for inclusion in this work.

We believe that MOTION PICTURE SOUND ENGINEERING will answer a definite need for an authoritative treatise on sound recording and sound reproducing, and anticipate that this volume will become an important reference authority on sound in motion pictures.

Nathan Levinson

Vice-Chairman,

Research Council

Academy of Motion Picture Arts and Sciences.

Hollywood, California.
January 3, 1938.

In Appreciation

The individuals listed on the following pages, by serving on various Committees and by participating in its activities, have contributed greatly to the success of the
Research Council
of the
Academy of Motion Picture Arts and Sciences,
and through their constant interest and activity
the Council is able to continue its record of achievement.

DARRYL ZANUCK, *Chairman 1933.*
SAMUEL J. BRISKIN, *Chairman 1934-1935.*
WILLIAM KOENIG, *Chairman 1936-1937.*

John Aalberg
Phillip Abrahams
Lawrence Aicholtz
Fred Albin
Emma Alblanalp
Joseph Aller
John Arnold
J. A. Ball
Lionel Banks
H. J. Barnicoate
Thad Barrows
J. F. Barry
James Basevi
Jesse Bastian
C. N. Batsel
M. C. Batsel
Anne Bauchens
W. B. Beatty
Fred Beetson
Gerald Best
O. C. Binder
A. C. Blaney
John Boyle
John G. Bradley
Frank Brandow
W. G. Brennan
James Brocklehurst
Norbert Brodine
Oscar Brodine
H. D. Buckley
Frank Burns
G. A. Burns
W. S. Burrows
Sidney Burton
W. W. Byrne
Frank Cahill
Richard Cahoon
F. S. Campbell
Leslie I. Carey
Ganahl Carson
John Cass
Wallace Castle
O. O. Ceccarini
Merle Chamberlain
Gordon Chambers
L. E. Clark
William H. Clark
Lewis M. Clement
Harry Cohan
L. W. Conroe
Jack Cosgrove
J. Crabtree
J. I. Crabtree
Duncan Cramer
George Crane
W. T. Crespinel
Harry Cunningham
C. R. Daily
Ralph Dawson
Arthur DeBra
Harry Decker
A. de Stefano
Arthur S. Dickinson
G. L. Dimmick
Harry Dodge
William Dover
Carl Dreher
Joseph Dubray
O. L. Dupy
Elmer Dyer
Frank Dyke
Arthur Edeson
Farciot Edouart
William Eglinton
Harry Ehrreich
Ferdinand Eich
Harris Ensign
Porter Evans
Robert Fairbanks
Art Fields
Godfrey Fischer
Fred Flack
C. B. Fogel
George J. Folsey, Jr.
James Francis
Bert Frank
J. Frank, Jr.
J. G. Frayne
Alan Freedman
George Friedl, Jr.
Fred Gabourie
Fred Gage
J. J. Gain
Lewis Geib
Lola Gentry
Merritt B. Gerstad
Harvey Giffen
F. C. Gilbert
Arnold Gillespie
Keith Glennan
Bert Glennon
William Goetz
Henry Goldfarb
A. N. Goldsmith
Lloyd Goldsmith
Ellis Gray
W. E. Green
Herbert Griffin
Lorin Grignon
William Grot
George Groves
A. J. Guerin
Robert L. Gunn
G. Knox Haddow
Archie A. Hall
Ralph Hammeras
Charles W. Handley
E. H. Hansen
Franklin Hansen
W. C. Harcus
Arthur C. Hardy
Edward Harris
Sylvan Harris
Byron Haskin
F. W. Herron
Eric Heyer
Al Heiman
Henry Herzbrun
Gabriel Hess
A. P. Hill
John Hilliard
Robert Hilton
Winton Hoch
Wm. Holmes
Wm. Hornbeck
Stanley Horsley
C. F. Horstman
A. S. Howell

C. Roy Hunter
Emery Huse
Thomas Ingman
Al Inman
Lester B. Isaac
Fred Jackman
Harry Jones
Lloyd Jones
Ray June
B. B. Kahane
Wallace Kelley
E. P. Kennedy
Harry Kimball
H. G. Knox
H. F. Koenekamp
Louis Kolb
H. R. Kossman
Sid Kramer
Barton Kreuzer
William Kunzmann
Franklin LaGrande
Arthur Lamb
Kenneth Lambert
Charles B. Lang, Jr.
Grover Laube
F. T. Lay
Robert Layton
Fred Leahy
Wilson Leahy
E. D. Leishman
Herman Lentz
Michael Leshing
Nathan Levinson
Harold Lewis
C. H. Lindblom
John Livadary
Edwin C. Loeb
C L.. Lootens
Don Loye
Sidney Lund
Luigi Luraschi
Karl G. MacDonald
J. S. MacLeod
N. L. Manheim
Glenn E. Matthews
J. P. Maxfield
E. A. McClintock
Hugh McClung
R. H. McCullough
D. C. McGalliard
E B. McGreal
J. H. McNabb
C. E. K. Mees
Lewis L. Mellor
Harry Mersay
Frank Meyer
W. C. Michel
A. W. Miller
Arthur Miller
Burton F. Miller
R. P. Miller
Virgil Miller
Wesley C. Miller
William Miller
Victor Milner
J. J. Milstein
Peter Mole
Regina Molseed
Abe Montague
K. F. Morgan
Sam Morris
Charles Moskowitz
Thomas Moulton
Hollis Moyse
William Mueller
Edward Muhl
Martin Murphy
Joseph Nadel
Harry Newman
J. M. Nickolaus
N. F. Oakley
M. D. O'Brien
John Ojerholm
Emil Oster
Jack Otterson
David Palfreyman
Fred Pelton
C. C. Pettijohn
Maurice Pivar
Van Nest Polglase
George C. Pratt
Gerald F. Rackett
Elmer Raguse
W. B. Rayton
H. W. Remerscheid
Elmer Richardson
W. G. Robinson
Sid Rogell
David Ross
Harry Rubin
William Rudolph
C. C. Ryan
Loren Ryder
H. B. Santee
A. J. Savin
Gordon Sawyer
Roy Seawright
Murray Seldeen
Carl Schillinger
A. W. Schwalberg
V. B. Sease
George Seid
Douglas Shearer
H. C. Silent
Nugent H. Slaughter
W. B. Slaughter, Jr.
Theodore Soderberg
Karl Spain
Nate Spingold
E. I. Sponable
Joseph H. Spray
Herbert Starkey
Joseph Stofel
Karl Struss
John Swain
Homer Tasker
George Teague
E. W. Templen
William Thayer
William Thomas
G. A. Toepperwein
Al Tondreau
Ralph Townsend
Nick Tronolone
S. J. Twining
John Volkman
Samuel A. Waite
Stuart Wainwright
Joseph Walker
Vernon Walker
J. S. Ward
J. V. Ward
George Weltner
H. M. Wilcox
I. James Wilkinson
Ray Wilkinson
Alan Williford
R. C. Willman
Anthony G. Wise
Sidney K. Wolf
Wallace V. Wolfe
Ben Wurtzel
Frank Young
Arthur Zaugg

COMMITTEE ON INDUSTRIAL EDUCATION

J. G. FRAYNE, *Chairman*

BARTON KREUZER

BURTON F. MILLER

WILLIAM THAYER

RALPH TOWNSEND

SOUND COURSE INSTRUCTORS

(Preliminary Course)

A. P. Hill

(Advanced Course)

Fred Albin

L. E. Clark

John K. Hilliard

Harry Kimball

PART I

Chapter I

BASIS OF MOTION PICTURE SOUND

By WESLEY C. MILLER

"The world is so full of a number of things, I'm sure we should all be as happy as kings." Even though the human race fails to attain to this condition of perfection each of us has his own ideas about the world in which he lives. Consciously and otherwise we are continually absorbing from the infinite variety which surrounds us, and in turn we make our contribution through our own reflected personality. Since earliest childhood we have been sorting out and compiling impressions and associations which are limited only by our own imagination, reasoning power and environment. As the conscious mind learns to exert its main effort in directing our activities, much of this assimilation becomes involuntary and automatic. The accumulation of impressions is important — indifference to the means of accumulation results. In normal life this indifference is not necessarily a criticism. However, in the creation of artistic illusions such as the motion picture, careful analysis is required of the nature of both the voluntary and involuntary sensory impressions.

Translation of things into impressions in the consciousness of our brains and minds is through the senses. The sensations which these organs furnish to us have enabled us to build up an accumulation of experience whereby cause and effect assume the relations and associations we have learned to expect. In the natural world certain combinations of objective elements of sight and sound are familiar to us. It is only when our expectations are disturbed that we commence to wonder and to investigate. The character of a sound informs us of its probable source and time of origin. We know what to expect when we start out to identify that source and we have but to trace it back to find it.

The recording medium introduces a new element—time. The reproduced sound may no longer be traced directly to its source in point of time. Any period may elapse between the original inception and the final reproduction. The identification of source becomes a voluntary effort, and a multitude of questions arises to perplex us in the technique of the reproduction system. The motion picture craftsman desires to create in his product an illusion which plays upon the imagination of his audience to make them forget these artificial factors. By the many artifices at his command he may often transport them from their own

sphere to the entirely new surroundings which he provides for them at the screen.

1. THE RECORDING SYSTEM A ROBOT

The ear is but a part of the sense of hearing. It is the means of transmitting outside sound phenomena to the brain where the final impression is made. The brain selects that which is wanted and discards the rest so that, within reasonable limits, we hear only that which we wish to hear. This faculty of selection and concentration is very valuable to us in the atmosphere of noise which is always around us. It makes conversation possible in the midst of crowd noise at a football game. Again, in quiet surroundings, we can select from among several persons all talking together, the voice of the one we wish to hear. Physically the presence of our two ears—the binaural sense of hearing—is an invaluable aid to such concentration. Sounds reach the two ears with minute differences of time and phase and loudness which give us a means of identifying the location of the source. Perspective is established and the world of sound becomes three dimensional.

A microphone and recording system have no such discretion. They are robots which pick up everything within their range and record it to the best of their ability. They possess no inherent means of selecting only that portion which is wanted. Emphasis is not possible to them without human guidance. Moreover, the recording system is, up to the present time at least, a monaural or single-eared arrangement. Binaural reproduction is practicable experimentally and should be an advantage when perfected for general use. In any case the direction of the robot, the provision of a brain for the microphone, devolves upon the sound man. He must control its location, directional qualities and response so that the audience will hear only that which is intended for it. Reduction of undesired extraneous noise to a minimum, exaggerated perspective, control of loudness and character of sound are in his hands to provide an illusion of pseudo binaural hearing and realism on the screen.

2. NATURALNESS OF ILLUSION

Sound must naturally just fit the picture on the screen and it must seem to come from the apparent source in the picture. The philosophy of the sound illusion is complex, but fundamentally it is largely a matter of perspective and loudness. Perspective is built into the record with theatre reproduction conditions in mind. Loudness, however, is susceptible to adjustment in both the record and theatre and is a study in itself.

An audience clearly demands that more be heard from a motion picture than from the usual stage presentation, apparently recog-

nizing the voice limitations of the actors on a stage and being willing to accept the loss of certain parts of the show which cannot be heard. The motion picture audience is more difficult to please as it demands complete understanding of every word. Inasmuch as the timing of audience reactions must be predetermined and built into the picture, this demand is often hard to meet. Fortunately the correct loudness or sound volume to fit the picture is invariably greater than that which would be heard from a person at the screen, as the screen figure is usually larger. In passing, it is also of importance to note that all people do not hear equally well, which means a compromise adjustment to fit the majority.

A person raises his voice to speak more loudly. The loudness change is created by a physical change in the voice characteristics. The pitch increases, the power content becomes located higher in the frequency range and the bass elements are reduced. Somewhat the same conditions hold for any noise or for music, with the reservation that the pitch change may be either less or not present. With a reproducing system, however, loudness may be artificially adjusted over a virtually unlimited range. In such an adjustment, the entire frequency range is raised or lowered an equal amount. If the volume is improper and does not correspond to the original, the ear hears an unnatural sound from the distorted relation between the various elements. Too-loud reproductions sound over-bassed, while thinness and lack of bass result from too-low volume.

The recorder attempts to have his actors speak at a voice level appropriate to the scene and surroundings. To make this voice seem natural at the screen after its loudness is increased to fit the picture, he then adjusts the gain-frequency response of his system—equalization—to compensate for the expected difference in loudness. If the actor fails to speak in the proper tone of voice, further equalization is required to make the final reproduction sound right. If, after these corrections, the theatre plays the sound at improper volume level or loudness, the whole balance is further upset in spite of the fact that the record was properly made.

Loudness is greatly affected by reverberation. The ear hears both the sound coming directly from a source and a series of reflections in various relations of phase and attenuation. Recording technique must not only recognize reverberation but must employ it as one of the factors to enhance the illusion.. One effect of monaural listening is an apparent increase of reverberation over the binaural condition. Hence a general reduction of the normal reverberation is necessary in the record. A departure from this rule may often be desirable for certain forms of

musical recordings. In every case the original reverberation on the set must be controlled and used with reference to the additional reverberation which will be introduced by the theatre acoustic condition.

The nature of the overall gain-frequency characteristic of a complete system has always been the subject of much discussion. On first thought it would appear that distortionless reproduction would be best obtained by a system having every element capable of uniform response at all frequencies. This would be perfectly true were it not for the necessity for the adjustment just mentioned, to compensate for louder reproduced volume. There is fairly common agreement upon the nature of this adjustment, which is in general an attenuation at the low frequencies. In addition the agreement is becoming more general on the total range of frequencies which should be accommodated. About 50 to 7,500 cycles is more or less accepted as the useful range, with a definite feeling that even the upper end of this range may at times cost more in apparatus refinements and difficulties than its reproduction may be worth.

3. RADIO VS. PICTURES

The sound technique for motion pictures is quite different from that of radio. Radio music and speech require nothing of the eye. The listener may create his own imaginary picture of the setting, or more often probably gives it little or no thought. Moreover, he adjusts the volume to his immediate pleasure with no reference to the probable original volume or to any illusion of the source having been in the room. In the radio studio no limitations exist to control the microphone placement—there are no lights or camera lines to be avoided. A psychological factor is that radio is more often used as a general background with a lowered concentration demand upon the listener. The motion picture illusion, involving both sight and sound, must be more perfect and its requirements are much more complex. Aside from the need for perfect synchronism between picture and sound, the similarity of apparatus and transmission methods is recognized, but the philosophy of their use in motion picture work is different.

4. AN ENTERTAINMENT FACTORY

The motion picture industry is from every point of view one of the most peculiar in the world. Regardless of artistic desires which are its foundation, the commercial end is of manifest importance. Occasionally an artistic effort is made with no thought of pecuniary gain but it is usually the case that a profit must be made. This requirement is sometimes difficult to meet in terms of the purely artistic. The business is one of increasing demands upon the technical staff to maintain the perfection of past performances and to continually be developing

and improving the technical processes to give greater freedom to the creative genius of the producer and his staff.

Entertainment value is a measure of the success of the production. Entertainment will, however, be found to be a very complex and clever combination of all of the many elements available to the producer. The motion picture studio may very aptly be termed an entertainment factory. In many respects it is organized on the same principles as those followed in any other type of manufacturing plant. The raw material happens to be stories and situations, actors, music and effects. The embellishment and processing which this raw material gets is the technical treatment at the hands of the art director, cameraman, sound man, musician, editor and many others, guided and moulded by the hand of the director who is a master craftsman and designer with full knowledge of the capabilities and limitations of his tools.

An important corollary is that all of the various people working with the director must not only be thoroughly familiar with his desires and aims but must also have a well rounded experience and knowledge of the work of the other departments. The set designer cannot disregard acoustics even though he be tempted to use a particular form of construction for some other reason. The camera and sound people have the constant problem of the relation between camera angles and microphone location, affecting both perspective and interference between the camera lines, lighting and microphone. The musician designs his music and lyrics for the scene. Scenes are made and remade for perfection and scores of thousands of feet of film are shot. Finally the editor draws all of this material together into final finished story form and must depend upon the work of the others who have made the film to furnish him with material which has the proper mechanical and dramatic elements to combine properly into a dramatic entity of suitable length to be released.

Probably the simplest thing which the sound man does is to record sound. His familiarity with all of the capabilities and limitations of his sound system as such is the familiarity of a workman with his tools. His real creative work involves complete understanding of the overall showmanship problem and of the interlocking requirements of the several crafts. It happens that his own technical language is that of decibels, gammas and equalizers which are entire strangers to the rest of the motion picture world. It also happens that the sound man's training has usually been such that he is able to assimilate the basic knowledge of the other fields and can meet them on their own ground. The reverse is often not true. It is unfortunate that the sound man may have the reputation of being "high brow." Often he appears to ob-

struct the way of progress because he alone recognizes some of the limitations beyond which, at the moment, it is not practicable to try to go. It should be recognized that this same sound man is well able to plan the details of a scene and to produce the desired overall effect while still keeping within the limitations imposed by physical laws which no one has yet found a way to circumvent. This is not a plea for the sound man but merely an observation of experience, an observation which, when followed, usually has a beneficial result upon the completed productions.

5. A BRAIN FOR THE MICROPHONE

The sound man is the brain of the microphone. He guides the technical system and its action to produce the illusion which is desired on the screen. Without him many undesired sounds which actually occur on the set but which have no bearing upon the production might easily be recorded into the finished product. With him the desired selection and emphasis are obtained. Adjustments to compensate for the monaural listening ability of the microphone are secured by his apparatus control so that he records only that which is wanted, discarding all undesirable material. His technical knowledge of the mechanics of acoustics, equalizers, filters and other mechanisms peculiar to sound work is brought to bear. It is his job to handle and combine a series of objective factors to produce a subjective result which will entertain.

In most instances the sound man is a whole department. It is the practice, particularly in the major studios, for the control of the sound policy to be maintained within certain limits. General type of quality and characteristics, specification of acoustic and other working conditions on the set, laboratory control, maintenance of a complete electromechanical recording system—these and a multitude of other factors are coordinated and controlled by a sound department as such. The individual recordist may then go out on the firing line—the set—with a system at his command which is carefully maintained and adjusted for him and he is free to work out and to coordinate the many details which arise there. Far from stifling the individual and his efforts, this mode of operation gives him more freedom because he is relieved of the care of many technical matters which can be designed and maintained for him.

There is another member of the sound fraternity who is not usually included in a discussion of recording and reproducing technique but who plays a major part in the success or failure of the studio's sound efforts. This is the theatre projectionist, or it may be the management. Whatever the studio sound may be when reproduced under carefully controlled studio conditions, the final reproduction is in the theatre

and under its control. It is gratifying to know the degree of favorable results and intelligent handling in this essential part of the field even though much remains to be done in modernizing equipment in many theatres. This is probably the most important technical problem confronting the sound end of motion pictures at the present time. Fortunately it is being approached in a manner which indicates that improvement may be expected.

6. THE TECHNICAL SYSTEM

The basic elements of a sound recording and reproduction system are perhaps fairly well understood by even the layman. However, a brief review of the fundamental processes may not be amiss, to provide the reader with some form of outline into which he may fit the more detailed discussions as they appear.

Sound recording and reproduction involve a series of transformations of energy from one form to another. The listener should hear by means of the recording the same sound he would have heard had he been present when the original was created. If a simple reproduction system is interposed between source and listener, the latter hears the sound through the system simultaneously with its creation. The further interposition of a recording mechanism places the sound in permanent form on a recording medium so that it may be reproduced at any future time. Any such system can do no more than faithfully reproduce an original. Actually all systems have certain inherent limitations and their faithfulness is measured by the degree of refinement attained in their design and use.

Sound emanates from a source in the form of pressure variations in the air. The amplitude, frequency and phase relations of these variations are determined by the character of the sound. Their effect upon the ear produces the sensation of the particular sound. The recording and reproduction system should faithfully maintain the original pressure variation relations or the listener will not hear a counterpart of the original.

The elementary form of the sound wave is the familiar sine wave. In passing, two of its important features should be noted: First, that sound is made up of highly complex and constantly shifting combinations of frequency, relative amplitude and phase relations; second, that any wave form or group of wave forms can be shown to be the result of a combination of a number of simple sine waves of various frequencies and other determining characteristics. Not only is this true of sound but it also applies to all of the electrical, mechanical and other systems which appear in the recording process. It should be further noted in this connection that a relatively sharp distinction exists between tran-

sient and steady state conditions. These impose entirely different demands upon the technical system, even though they may both be reduced to a sine wave series.

The diaphragm of a microphone placed in the field of a sound wave will vibrate in sympathy with it, resulting in mechanical energy in the form of the diaphragm movement. Microphones of any type depend upon the equivalent of such diaphragm movement to change the resistance of a carbon button, to vary the capacity of a condenser, to move a coil or ribbon within a magnetic field or to modify the pressure upon a piezo-electric crystal. Any type of microphone thus effects a change in energy from the mechanical form of diaphragm movement to an electrical form as the output of the instrument.

In electrical form the energy variations are in such condition that any desirable amount of amplification may be provided by means of the conventional vacuum tube amplifier. Moreover, it becomes practicable in this stage to make any changes or corrections which may be desirable in relative frequency response or in other characteristics of the signal. Too much cannot be said of the importance of the amplifier and vacuum tube to sound work. As in radio and other allied fields, the vacuum tube amplifier makes the whole process practicable.

With the electrical signal amplified to a sufficient degree, the next steps are further transformations into mechanical, and finally chemical form, using light in the transition, through the medium of the modulating device which regulates the signal that goes on the film. The modulator device is in general an electrically operated mechanical shutter of which there are many types. It controls the light from a constant light source in such a manner that the amount of light falling upon a continuously moving recording film varies from instant to instant in proportion to the signal given to the modulator. The film thus receives a varying exposure which is in practically strict proportion to the original sound wave. In the ladder-like striations which appear on a variable density film record, the degree of density or film transmission change is a measure of loudness, while the length or duration of the striation is a function of frequency. With the film moving at a velocity of 90 feet per minute—18 inches per second—each cycle of a 1,000-cycle wave on the film is eighteen one-thousandths of an inch long, whereas a 6,000-cycle wave obviously is but one-sixth of this length. The variable area striation expresses relative loudness in terms of varying width on the sound track—again a variation of transmission. The striation duration or length must, of course, be the same as for variable density. Suitable chemical processing of this negative film record, that is, development and printing in much the same manner as for any

photograph, produces a positive film record capable of reproducing the signal source. This record may be reproduced at any time, and, combined on a single film with the picture to which it applies, may be used for theatre projection.

7. REPRODUCTION

Reproduction is a further series of energy transformations. The first step is optical and electrical. A fixed light source is focused to a sharp line on the film record while the latter is moving in front of a photo-electric cell. The cell thus sees varying amounts of light through the moving film record and has the important function of translating them into corresponding electrical variations. Once more the signal energy is in electrical form which can again be amplified to the desired degree and led into specialized types of the conventional loud-speaker.

The loud-speaker has an electrically operated diaphragm, coupled to the air through some form of horn or projector. The diaphragm is actuated by the amplified electrical energy from the photo-electric cell, and in vibrating mechanically produces in the air a series of pressure variations. If the process has been carefully carried out we are back where we started. These pressure variations will be accurate counterparts of those which emanated from the original source and the listener will hear sound which is a virtual copy of the original and which can be played at any required or desirable volume.

To recapitulate, the several steps in the process and the energy transformations or translations are as follows:

Recording	1.	Source	Acoustic
	2.	Air Medium	Air Pressure
	3.	Microphone Diaphragm	Mechanical
	4.	Microphone Output	Electrical
	5.	Amplification	
	6.	Modulator	Mechanical
	7.	Optical System	Light
	8.	Film—Negative Record	Photochemical
Reproduction	9.	Film—Positive Record	Photochemical
	10.	Optical System	Light
	11.	Photo-electric Cell	Electrical
	12.	Amplification	
	13.	Loud-speaker Diaphragm	Mechanical
	14.	Loud-speaker Output	Air Pressure
	15.	Ear	Acoustic

The sound man is concerned with each of these translations, their inter-relations and the technical limitations and capabilities of each.

Acoustics, frequency response, harmonic production, phase shift, resonance, power capacity, acoustic efficiency, film characteristics, optics—all of these and many more play their parts in the process and must be controlled. Electrical circuits, amplifiers, photo-electric cells, filters, equalizers, attenuation and transmission elements must continually be combined and employed most effectively for the particular end in view. The sound man must be intimately familiar with the operating principles and available types of microphones, amplifiers, modulating devices, photographic processes, loud-speakers, and the mechanical film recorders and reproducers. His success depends upon his ability to properly handle these various tools and to direct their uses in the best way to secure the result which is desired for the final product. It is the hope that the later chapters of this volume may be of benefit in securing a better understanding of these tools, and of the recording and reproduction problem as a whole and in detail.

Chapter II

THE NATURE OF SOUND

By L. E. CLARK

1. INTRODUCTION

In sound motion pictures, dialogue, sound effects and musical sequences must be picked up and recorded in some permanent manner so that later they may be reproduced in the theatre. Since the recording equipment must be acted upon by the sound, its operation depends upon the physical nature of sound, and upon how the sound behaves between the time it leaves its source and the time it arrives at the receiver. Likewise, the reproducing equipment must be the new source of sound, and what the auditor hears depends upon this physical nature of sound.

Therefore, it is proposed to discuss and describe some of the fundamental characteristics of sound so that application of these properties can be made in later chapters. To be able to do this we must understand how sound behaves; that is, how it is transmitted from place to place, how it is reflected, absorbed, refracted, and diffracted, and what effect it has upon the ear. All these things must be known so that they can be used in delivering to the auditor sound of the same quality that he would enjoy were he present at the original performance.

2. ORIGIN OF SOUND

In order to describe these fundamental characteristics of sound it is first necessary to define sound itself. A very good, easily remembered definition is that *sound is a compressional wave motion in an elastic medium.* The American Standards Association defines sound as *"an alternation in pressure, particle displacement, or particle velocity, propagated in an elastic material* or *the superposition of such propagated alternations. It is also the sensation produced through the ear by these alternations."*

Sound is generated by a vibrating body. If a bell or a tuning fork is struck, it is set into vibration; these vibrations are transferred to the surrounding medium (air in the usual case), which in turn transmits them to the receiver (ear drum, microphone, etc.). Sound

may also be generated by vibrating air columns, as in organ pipes or the human voice, and by vibrating strings. In the case of stringed instruments, the strings themselves transfer very little vibration to the air, but instead vibrate a sounding board or the body of the instrument which transfers much more vibration to the air. In order to transfer enough vibration to the air for the sound to be audible, it is necessary that a considerable area of the vibrating body be in contact with the air. In general, any vibrating body, solid or fluid, will generate sound, but some medium is necessary to transmit the sound to the listener. A bell, ringing in an evacuated bell jar, cannot be heard, although the clapper can be seen to be vibrating. As soon as air is allowed to enter the jar the bell can be heard.

3. SOUND WAVES

As the vibrating body moves forward, the air immediately in front is compressed and the pressure thus built up is relieved by the wave of compression moving outward. As the vibrating body moves back, the air is rarefied, and this wave of rarefaction moves outward behind the wave of compression. Since the vibrating body executes this motion periodically, there will be a train of alternate compressions and rarefactions travelling rapidly away from the body. If the air be examined at any given instant of time there will be found, assuming

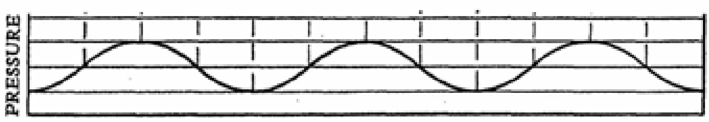


Figure 1.

a simple sine wave of vibration, a series of compressions and rarefactions equally spaced along the direction of travel. The distance between compressions, or rarefactions, or the corresponding points of any two successive waves is known as the wave length, which is designated by "λ." The number of complete vibrations that the body makes in one second is known as the frequency of vibration. Since the body is vibrating with a frequency of "f" cycles per second, it will send out f waves during a second, and as each wave has a length of λ, the sound travels over a distance of $f\lambda$ in one second. The distance traveled in one second is the velocity. Therefore, the relation between velocity, frequency and wave length is,

$$C = f\lambda \tag{1}$$

where C = velocity of sound.

This space picture of the sound wave may be expressed by the equation:

$$p = P_M \cos(-kX) = P_M \cos\left(-2\pi \frac{X}{\lambda}\right) \tag{2}$$

where p = instantaneous pressure above static atmospheric pressure

P_M = maximum instantaneous pressure, or the pressure amplitude.

X = the distance from a reference point

$k = \dfrac{2\pi}{\lambda}$

If at any one point the air be examined over a period of time, it will be found to undergo periodic compressions and rarefactions as the sound passes by. The time between successive compressions is called the periods "T," and is the reciprocal of the frequency, that is,

$$T = \frac{1}{f} \tag{3}$$

The time picture of the second wave is given by:

$$p = P_M \cos \omega t = P_M \cos 2\pi ft = P_M \cos 2\pi \frac{t}{T} \tag{4}$$

where t = time measured from a reference point

$\omega = 2\pi f$

If these two pictures are now combined, the complete space-time picture of the wave can be given by:

$$p = P_M \cos 2\pi \left(\frac{t}{T} - \frac{X}{\lambda}\right) = P_M \cos k\,(Ct - X) \tag{5}$$

This is strictly true only for plane waves, some modification being necessary for spherical and other types of waves.

The sound pressure, "p," is usually measured in dynes per square centimeter. An idea of the magnitude of a dyne can be had from the fact that a mass of one gram weighs 981 dynes. Normal static atmospheric pressure (14.7 lbs./sq. in.) is of the order of 1,013,000 dynes per square centimeter. Ordinary sound pressures range from a fraction of one dyne to several dynes per square centimeter, the threshold of a normal ear being taken as 0.000204 dynes per square centimeter at 1,000 c.p.s. Very intense sounds will have pressures of 100 or 200 dynes per square centimeter. It is seen, then, that the range of pressures in ordinary sounds is very great, of the order of 10^6 to 1, or 120 db, and yet this variation in pressure is very small compared to ordinary atmospheric pressure. The fact that these low level sounds can be heard is due to the remarkably high sensitivity of the ear.

4. VELOCITY OF SOUND

Sound travels away from its source with a velocity that depends upon the medium which is carrying the sound wave. The heavier the particles of the medium or the denser the medium, the more slowly will the wave travel, and the less compliant or more elastic the medium, the faster will the wave travel. The velocity is related to these characteristics of the medium by the equation,

$$C = \sqrt{\frac{E}{\rho}} \tag{6}$$

where $E =$ elasticity

$\rho =$ density

For air at 0° C., $E = 1,400,000$ dynes/sq. cm.

$\rho = 0.001293$ grams/c.c.

$C = 33,100$ cm./sec. or 1,080 ft./sec.

The velocity of sound in some of the more common materials is given below:

air (15° C.)	1,120 ft./sec.
water (15° C.)	4,700 ft./sec.
steel	16,500 ft./sec.
brick	12,000 ft./sec.

Because the elasticity and density vary with temperature, the velocity will also vary with temperature. The velocity of sound, however, is not dependent upon the frequency over the range of frequencies generally referred to as audible sound. It is not dependent upon the intensity except for very high intensity sounds, such as a thunderclap, in which case there is some increase in velocity.

5. REFLECTION OF SOUND

When sound traveling in one medium encounters a change in that medium or encounters another medium (such as water, colder air, or a solid object), which is large compared to the wave length of the sound, it is partially reflected back into the first medium, and partially transmitted by, and absorbed in, the second medium. The same law of reflection applies as in the case of the specular reflection of light, i.e., *the angle of incidence is equal to the angle of reflection.* In Figure 2, sound traveling in the direction AD, strikes the boundary of the two media, OO′, at an angle of incidence ADC and is partially reflected in the direction DB, at an angle of reflection CDB equal to ADC. The relative amount of energy reflected depends upon the densities of the two media, and upon the velocity of sound in them. The ratio of intensity, or

energy, of the reflected sound to that of the incident sound, is the acoustic reflectivity, "r."

The energy not reflected is transmitted by the second medium or absorbed in it. In the use of acoustical materials for sound absorption, which will be referred to later, we are interested in how much of the sound energy will be removed from the room. Since this represents all

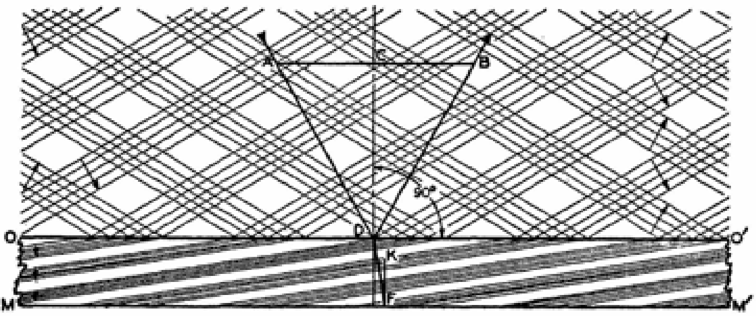


Figure 2.

the energy except that which is reflected, the acoustic absorptivity is defined as one minus the acoustic reflectivity, that is, $a = 1 - r$.

For reflection to take place, however, it is necessary that the reflecting surface be large—at least several wave lengths—in comparison with the wave length of the sound. If the surface is smaller than this, a phenomenon known as diffraction begins to become effective and it is impossible to locate the reflected sound by the laws of reflection.

6. REFRACTION OF SOUND

Referring again to Figure 2, the sound wave in the second medium travels in a direction different from that of the incident sound. This phenomenon, which is called refraction, is due to the difference in the velocity of sound in the two media. As the velocity is less in the second medium, the edge of the wave slows up when it enters this medium. The part of the wave remaining in the first medium is still traveling with its initial velocity and as a consequence there is a shift in the direction of travel from AD to DF. If C is the velocity in the medium and C′ that in the second medium, then

$$\frac{\sin \angle ADC}{\sin \angle KFD} = \frac{C}{C'} \tag{7}$$

The effect of refraction is most often observed when it is due to the existence of layers of air at different temperatures, such as cool air over a warm body of water in the evening.

7. DIFFRACTION

Diffraction is the bending of sound waves around an obstacle. If a window is opened only slightly, the entire room will be filled with any sound present on the outside. However, sunlight coming through the window, instead of spreading throughout the room, will cast a sharp shadow. If the edge of the light shadow is closely examined, diffraction would also be found, but on a much smaller scale than in the case of the sound wave.

This difference in the amount of diffraction present in the two cases is due to the different wave lengths of light and sound; sound waves being several feet in length and light waves being some twenty millionths of an inch in length.

Diffraction effects can be thought of as existing in a narrow region, only a wave length or so in extent, at the edge of the wave. Thus, when a sound wave passes by a large obstacle, such as a building, most of the energy goes straight on and only a small part is diffracted into the shadow caused by the building. However, when sound strikes a window which is open but a few inches (only a fraction of a wave length), the sound spreads out as if coming from a point source. Since most objects around us are only several wave lengths in size, we are accustomed to diffraction effects and usually do not give a second thought to the fact that we often listen to sounds whose source we cannot see.

8. INTERFERENCE

When two waves of the same frequency combine, they give rise to an effect known as interference. At points where the two waves execute their vibrations in-phase, the maximum pressure will be the sum of the pressures in the individual waves, while at points where the vibrations are 180° out-of-phase, the resultant pressure will be the difference in the two pressures. The sound will, therefore, be re-enforced at some points and destroyed at others, giving an irregular distribution of sound energy. Interference can be caused by waves coming from more than one source or by the combination of the direct and reflected sound from a single source. This commonly happens in theatres and studios, and accounts for the dead spots often encountered at certain frequencies. The interference pattern can easily be detected by establishing a pure tone in a room and then listening to the tone at various points in the room.

9. ARCHITECTURAL ACOUSTICS

The acoustics of the recording studio and theatre are of great importance because these rooms constitute two links in the chain from

the original source (through the recording and reproducing equipment), to the auditor, and any distortion originating in these rooms will be passed on to the listener. While modern science and technical advancement have made possible recording and reproducing systems that introduce negligible distortion into the sound, there is still opportunity for distortion to enter between the source and the microphone and between loud-speaker and auditor. A listener in a room seldom hears sounds that have not been changed in character in some manner en route to him. The various frequency components of the sound will be acted upon differently by the same absorbing materials in the room, the low frequencies usually being absorbed less than the high frequencies.

Some surfaces are of such size that they will reflect the high but not the low-frequency sound, the latter being diffracted around the surface. Often, therefore, a room may be very dead at some frequencies and very live at others, while at still other frequencies there may be detrimental echoes. In addition to distortion produced in this manner, the studio or theatre may be resonant at certain frequencies, or the structure may vibrate.

These are examples of distortion known as frequency distortion, in which the different frequency components are not all acted upon in like manner. There is also a type of distortion known as non-linear distortion, which is caused by failure of a system to respond in a proportionate amount for different inputs. This distortion, negligible for all but very intense sounds, is present in the air at all times because air follows the action of a gas under pressure.

As a result of the effects of serious distortion, sound usually becomes unnatural, hollow or unintelligible.

When an observer listens to a source of sound in the open, he hears only the direct sound, and as he moves away from the source the intensity of the sound falls off with the square of the distance. If, however, he listens to the sound in a room, the intensity does not decrease as rapidly as he moves away; in fact, it decreases very irregularly and at times may even increase because of the interference pattern of the enclosure. This occurs because he is receiving not only the direct sound but sound that has been once, or twice, or even many times, reflected from the surfaces of the room. To consider the matter from another point of view, when the sound energy is once introduced into an enclosed space, it remains there until used up or dissipated in some manner. Some energy is used up at each reflection and so the total is constantly decreasing.

The time required for the sound in any enclosure to die away to one-millionth of its initial intensity, or over a 60 db range, is known as the reverberation period of that enclosure, and depends upon the volume of the enclosure and upon the amount of absorbing material within it. The energy is used up by any materials that absorb sound which may be in the enclosure, and in the clothing of persons in the room—very little of it being dissipated by transmission through the air.

The sound reaching an auditor in the room consists of the direct sound, the first few beneficial reflections coming from surfaces near the source, and reverberant sound. If a reflecting surface is large and some distance from the original source, the reflected sound will arrive sufficiently later to be recognized as a distinct repetition. This is called an echo. If the reflected sound arrives 1/20 of a second later than the direct sound, there will be a noticeable echo. The reverberant sound is due to the many repeated reflections from the surfaces of the room, and is very detrimental to the perception, causing a hanging-over and general jumbling of the sound. In the case of speech, one syllable may be still reverberating in the room when the next one is spoken, making the spoken words unintelligible. The reverberant sound must be reduced if the room is to be satisfactory for speech or music, and for this reason, auditoriums, theatres and other similar places must be fitted with sound absorbing materials that reduce the reverberation time to a satisfactory value. The optimum reverberation time depends upon the volume of the room, upon the purpose for which the room is to be used, and upon the frequency of the sound. Figure 94 shows the optimum reverberation time for theatres of different size, and Figure 95 shows how this time should vary with frequency for any one room.

Moving picture auditoriums should be somewhat less reverberant than the average auditorium of the same size because the sound originates from the reproducing system, which has ample power, and it is not necessary to rely upon reverberant sound to build up the loudness. Then, too, some reverberation is recorded and this increases the effective reverberation in the theatre. Theatres should be designed so that there are no reflecting surfaces to cause echoes, and no curved surfaces to cause concentrations of sound or dead spots in the house.

In general, recording studios should be even less reverberant than theatres of the same size. This is necessary because the sound is picked up with a microphone and the effect is that of listening with one ear. This greatly increases the apparent reverberation time and the difficulty of making the direct sound stand out above the reflected sound. The ratio of direct to reverberant sound is of great importance to recording engineers because upon it depends the quality of the recorded sound. If

this ratio is too low, the sound seems reverberant, and if too high, the sound seems lifeless.

One exception to the above statement is in studio sound stages used for music recording. In this case a higher reverberation time—compared to that used for dialogue recording—is desirable to reduce the masking effect of music over dialogue.

In the presentation of a motion picture it is desirable to have the magnitude of the sound comparable to the comparative size of the actors in different scenes. That is, the audience expects to hear dialogue considerably louder when it is associated with a close-up or medium shot than in a long shot, or if the person were actually on the stage. For the same reason, sound effects must be exaggerated in like manner. This increase in loudness with respect to the original intensity is, in many cases, greater than 10 db. Because of the shift in characteristic with loudness (see Figure 4), careful control of all these factors is necessary in order that the reproduced sound seem natural, even though it may be louder than the original.

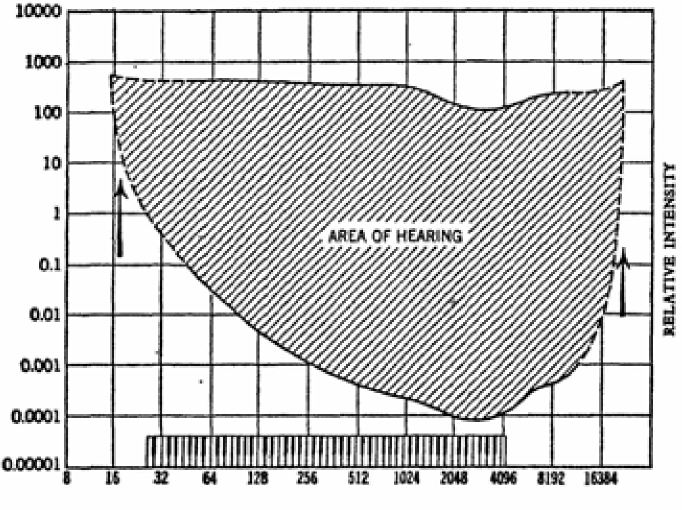
10. PERCEPTION OF SOUND

We are interested in the nature of sound and its effect upon the ear because ultimately sound is to be listened to, and what is perceived by the brain will therefore depend upon the behavior of the ear. Many extensive fundamental investigations have been carried out in this respect by Dr. Harvey Fletcher of the Bell Telephone Laboratories, with the result that our understanding of the process of hearing is very much advanced over that of a few years ago.

The three physical characteristics of sound are: Frequency, intensity and overtone content. Frequency, as previously mentioned, is the number of vibrations made in one second. The intensity is proportional to the amount of energy in the sound wave, and the overtone content is a function of the number, magnitude and phase of frequencies other than the fundamental. Most sounds are not pure tones, but are complex waves consisting of a fundamental tone (establishing the waves on the musical scale), and a number of components of higher frequencies. These latter constitute the overtones. A violin and a cornet, both playing middle C, sound different because the overtones are not the same for each instrument.

The ear does not respond uniformly to all frequencies, but is more sensitive to frequencies of about 1,000 c.p.s. (This is a form of frequency distortion, which is present whether direct or reproduced sound is under consideration.) The ear is able to respond to frequencies ranging from 20 to 20,000 c.p.s. Figure 3 shows the area of hear-

ing of an average ear. The lower line is known as the threshold of audibility and shows the minimum pressure at each frequency that



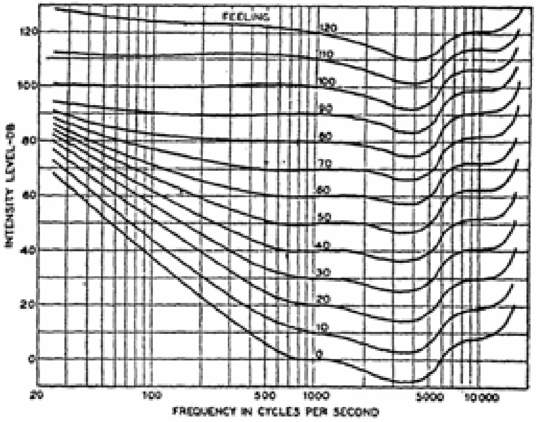
—*From "Sound," by Lemon & Schlesenger (The University of Chicago Press).*

Figure 3 — Recent data on the limits of audibility.

will cause a sensation of sound. The upper line is known as the threshold of feeling and shows the pressure, at each frequency, that will give rise to a sensation of pain. Any sound whose pressure and frequency fall within the area bounded by the lines will cause the sensation of sound.

There are three physiological characteristics of sound which approximately correspond to the three physical characteristics named above; namely, pitch, loudness, and quality. Pitch has been used interchangeably with frequency, but recent measurements indicate that the sensation of pitch also depends to a slight degree on the intensity and overtone structure. The loudness of a sound is approximately equal to the logarithm of the intensity, and is so defined for 1,000 c.p.s. Departures at other frequencies, especially the lows, are considerable and the definition should not be taken too literally for frequencies other than 1,000 c.p.s. Figure 4 shows contours of equal loudness obtained by plotting, on the hearing area of Figure 3, the locus of all points which a large number of observers indicated were of equal loudness. It will be noticed that these contours are not equally spaced but generally converge toward the low-frequency end. This is the cause of the change in quality of reproduced sound when the volume level is changed. For example: Consider a system carrying two pure tones of equal loudness, say 1,000 c.p.s. and 100 c.p.s. and let the 1,000 cycle

tone have an intensity level of + 80 db. Then, reference to Figure 4 shows that the equally loud 100 cycle tone has an intensity level



—*From a technical paper in Journal Acoustical Society of America, Volume V, No. 2, "Loudness, Its Definition, Measurement and Calculation," by Fletcher and Munson.*

Figure 4 — Loudness level contours.

of + 83 db. If the system gain is now lowered 40 db, the intensity of the 1,000 cycle tone is + 40 and that of the 100 cycle tone is + 43. Again referring to Figure 4, it is seen that the two tones are no longer of equal loudness, there being an apparent 20 db difference between the two. The balance is therefore completely destroyed, as the sensation received by the brain is entirely deficient in its low-frequency tone.

Investigations have shown that high frequencies, that is, frequencies above 1,000 c.p.s., are necessary for the intelligibility of speech. They also give brilliance to the sound. The low and bass frequencies are necessary to retain a material quality to voice and music, and in addition give "presence" to the scene on the screen.

The ear is also subject to non-linear distortion at high-intensity levels, its mechanism being unable to respond faithfully to very intense sounds. It is never possible to listen to undistorted sound as we cannot dispense with the ear between the sound and the brain. In view of the fact that the ear will respond to pressure changes of over a millionfold, and to frequencies over a range of 10 octaves, it must be considered to be a well designed mechanism.

The effect of age upon the hearing ability of the ear is shown in the table below, which indicates that, with age, the ear loses its sensi-

tivity to the high frequencies—the effect lessening as the frequency decreases.

DB Loss in Hearing with Age

Frequency	*60 to 1024 Cycles*	*2048 Cycles*	*4096 Cycles*	*8192 Cycles*
Ages 20-29 (96 ears) - - - - -	0	0	6	6
Ages 30-39 (162 ears) - - - -	0	0	16	11
Ages 40-49 (84 ears) - - - - -	0	2	18	16
Ages 50-59 (28 ears) - - - - -	0	5	30	32

It is not generally recognized that natural sound which we hear always emanates from its proper source, whereas in artificial reproduction we usually hear sounds which do not emanate from their original source. That is, in a motion picture reproduction, all sound originates from a fixed source—the speaker system. This sound includes both the sound which seems to originate from all portions of the screen, as well as sound that has its apparent source off the screen. As a consequence of this unnatural source of sound, the many factors affecting the relative loudness of the various reproduced sounds must be carefully controlled.

BIBLIOGRAPHY

(1) *American Standards Association's Tentative Standards*, Z 24.1, Z 24.2 and Z 24.3.
(2) *Sound*, F. R. Watson — John Wiley & Sons, 1935
(See for more complete reference list.)
(3) *Acoustics of Buildings*, F. R. Watson — John Wiley & Sons, 1930.
(4) *Applied Acoustics*, Olson and Massa — P. Blakiston's Son & Co., 1934.
(5) *Speech and Hearing*, Harvey Fletcher — D. Van Nostrand Co., 1929.
(6) *Vibrating Systems and Sound*, Crandall — D. Van Nostrand Co., 1927.
(7) *Architectural Acoustics*, V. O. Knudsen — John Wiley & Sons, 1932.
(8) *Loudness, Its Definition, Measurement and Calculations*, Fletcher and Munson — *Journal Acoustical Society of America*, Volume V, No. 2, p. 82.

Chapter III

TYPES OF FILM RECORDING

By L. E. CLARK *and* JOHN K. HILLIARD

1. RECORDING MEDIA

Sound recorded for motion pictures is of two general types, (1) sound-on-disc, and (2) sound-on-film. At the present time practically all commercial releases (in the United States) are on film, discs being used only within studios for sound playbacks, cueing, etc.

Because of the very limited use of disc recording at the present time, only the recording of sound on film will be discussed in this book.

2. TYPES OF FILM RECORDS

There are today two types of sound-on-film recording in general use: (1) "Variable Area," and (2) "Variable Density."

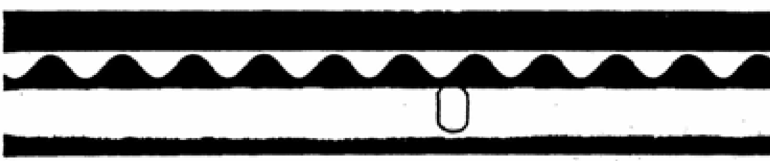
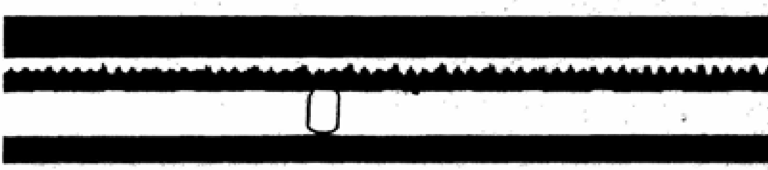


Figure 5-A — Example of a single-frequency variable area recording.



—*Courtesy RCA Manufacturing Co.*

Figure 5-B — Example of a typical variable area recording.

—*Courtesy Electrical Research Products, Inc.*

Figure 6 — Example of typical variable density recording.

In the former type, the record consists of a clear or transparent area adjoining an area of uniform density, the dividing line between the

two areas outlining the wave form of the recorded signal. The variable density record consists of transverse parallel striations of varying density—varying according to the frequency of the recorded signal.

These two types of records may be secured in a variety of ways, but the main distinction between the two is that in one type the area varies while in the other the density varies.

Figure 5-A shows an example of a single envelope variable area recording of a simple tone, while Figure 5-B is a variable area recording of a typical complicated signal. Figure 6 is an example of a variable density recording.

3. METHODS OF RECORDING

In the variable area type of recording only one method is employed for varying the exposure of the film, that is, a beam of light of constant intensity is caused to vary in its length.

In the variable density method, the slit through which the light strikes the film may be held constant and the intensity of the light source varied, or the intensity of the light may be held constant and the slit width varied.

(a) The Aeolight System

This was the first commercial sound-on-film system in use in the studios and gives a variable density record by varying the intensity of a recording light according to the signal, the light striking the film through a slit of constant width (0.001″), positioned between the lamp and the film, which passes the slit at a constant speed.

The aeo-lamp is a gaseous discharge tube, consisting of a nickel anode and a looped filament, designed to give a fairly intense beam of light at the "no-signal" level, accomplished by the use of a direct-current bias as illustrated in Figure 7. The value of the bias varies from 250 to 400 volts.

The lamp circuit shown is connected to the output of the bridging amplifier, the alternating signal component causing a variation in the bias voltage across the lamp and a consequent variation in its intensity.

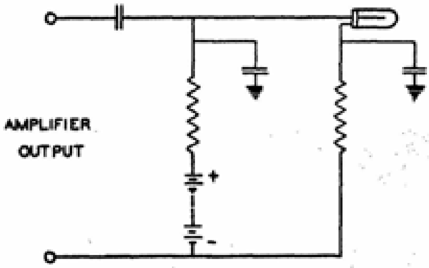


Figure 7 — Schematic circuit of a variable density recording modulator where the lamp intensity is variable.

Figure 8 shows the relation between the light intensity and the voltage. The bias value should be such that the no-signal intensity falls in the center of the straight line portion of the curve, while the maximum alternating-current variation, either plus or minus, must also be within the straight line portion, or distortion will result.

This method of producing variable density records has been superseded in the studios by the variable slit-width method.

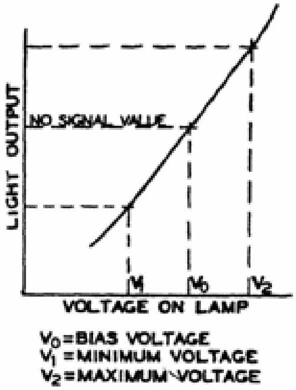


Figure 8 — Curve showing the relation between light intensity and impressed voltage of an Aeo-lamp.

(b) Variable Density Recording

Recording by the variable density method is also accomplished by means of a lamp of constant intensity, the light from which is focused by means of a lens on a slit. This slit varies in width according to the electrical intensity of the signal to be recorded, and so varies the amount of transmitted light in proportion to the electrical impulses. Such a light modulator is called a "Light Valve." The transmitted light is in turn focused upon the film by means of a second lens system.

A current carrying conductor, when placed in a magnetic field, is subjected to a force proportional to changes in the current passing through the conductor. If free to move, the conductor will move a distance proportional to the change in, and the direction governed by, the flow of current. Light valves operate on this principle.

A conductor, placed in a magnetic field and carrying a current, would tend to bow if the ends of this conductor were fixed. The center section of this conductor, if the conductor's length is great compared to its cross-section, would for all practical purposes, move parallel to itself as shown in Figure 9. In commercial application of this principle to the recording light valve, two ribbons are used, as shown in Figure 10.

The light valve ribbon is formed of two strips of duraluminum, 0.0005″ by 0.006″, spaced 0.001″ apart. The ribbon has a natural frequency of its own, depending upon its tension and length, and is consequently tuned for a frequency above the highest frequency to be recorded, which in commercial practice is always above 8,500 cycles. Tuning is accomplished in exactly the same way a violin string is tuned, by changing the tension on the ribbon. The current flows in opposite directions in each of the ribbons of the valve, resulting in an opening

and closing of the aperture between the ribbons when subjected to a signal current.

Originally, the two ribbons were mounted in the same horizontal plane. As previously mentioned, excessive modulation causes the rib-

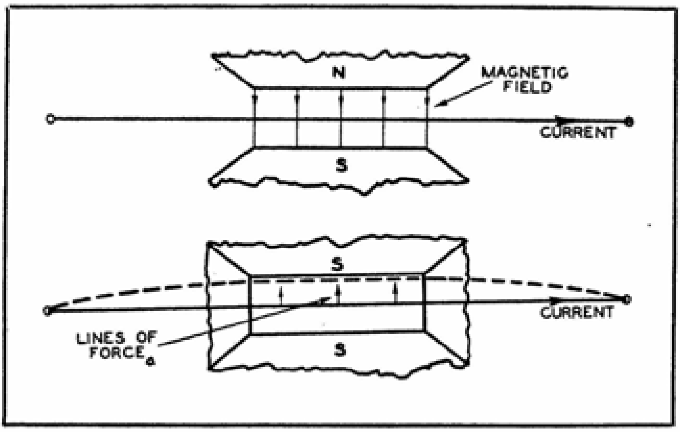


Figure 9 — Diagram showing the light valve operating principle.

bons to "clash" or strike each other. A valve so mounted is called a mono-planar valve.

In a more recently designed valve, called the bi-planar valve, the two ribbons are mounted, spaced one-half to two mils, one above the other, and vibrate in two parallel planes. This eliminates the possibility of the two ribbons striking together at periods of excessive modulation, resulting in the valve remaining in better mechanical condition, throughout its life.

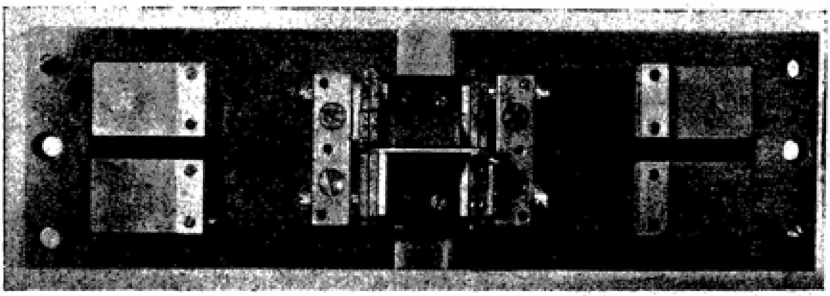


Figure 10 — Two-ribbon recording light valve.

Excellent results have been obtained with light valve recording, and the extensive use to which this method has been put in the studios

shows it to be rugged as well as efficient. However, care must be exercised in its use, and routine checks and adjustments made periodically. Azimuth adjustments must be particularly made, i.e., the horizontal plane of the valve must be perpendicular to the direction of film movement. Focal adjustments must be maintained as it is necessary to photograph the ribbon movements as accurately as possible. "Hysteresis," or the failure of the ribbons to return to their original position, may be minimized by increasing the tension, but recently designed valves are practically free of trouble from this source.

Overload and clash of the two ribbons, which result in distortion of the wave form, should be avoided. These effects occur when the ribbons modulate beyond the full or zero modulation position, which produces either a cutting off of the wave top or an actual clash of the ribbons.

Another source of possible trouble is that of "ribbon-velocity." This is due to the fact that one ribbon when opening and the other ribbon when closing, are traveling in the same direction as the film when it is passing the valve. The effective exposure recorded on the film is not only dependent upon the light flux, or intensity, but also upon the time a given point on the film requires to pass through the slit image; that is, the exposure depends upon the product of light intensity and time. At low frequencies this effect is negligible, as the velocity of the ribbon is small compared to that of the film (normally 90′ per minute), but at higher frequencies the two velocities become comparable and loss of volume as well as distortion in wave shape results.

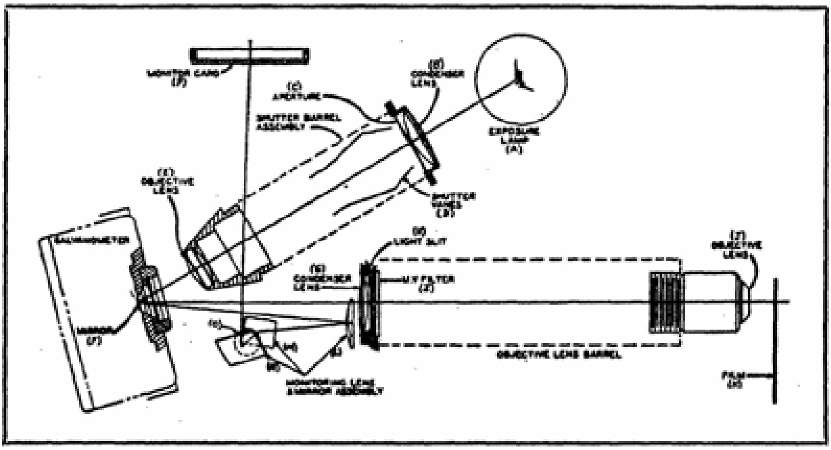
Variable density recording can also be secured by the use of a galvanometer and a penumbra in the optical system shown in Figure 11. This is the optical system normally employed for variable area recording with a galvanometer.

(c) Variable Area Recording

Fundamentally, this system consists of a concentrated lamp, an image of whose filament is focused upon a galvanometer mirror, from which the light is reflected and imaged upon a slit of constant width. The signal currents cause a rotation of the galvanometer mirror allowing either more or less of the slit to be illuminated, and thus is photographed a sound track of variable area, or width, but of constant density. The mean, or "no-signal" position, exposes one-half the sound track variations in the signal current causing a movement of the light beam either way from this mean position.

With this method a more complex optical system is necessary than for the variable density method.

Figure 11 shows such an optical system.



—*From the Journal of the S.M.P.E., September, 1937, "The RCA Recording System and Its Adaptation to Various Types of Sound Track," G. L. Dimmick.*

Figure 11 — Schematic of a variable area recorder.

The different types of tracks, discussed in the chapter on "Noise Reduction," may be produced by varying the type of aperture used. Single envelope variable area track as shown in Figure 5-B, which was the first used, employs an aperture with a rectangular opening, which at no-signal level is imaged upon the slit as shown in Figure 12 (accomplished by proper bias on the galvanometer). The incoming signal will then vary the position of the mask image upon the slit in proportion to its strength, resulting in an oscillograph record which is opaque on one side and transparent on the other in contrast to the usual oscillogram which shows only as a line on the film record.

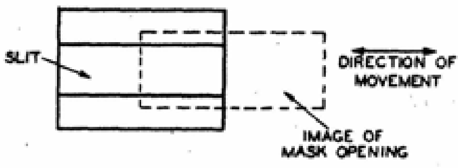
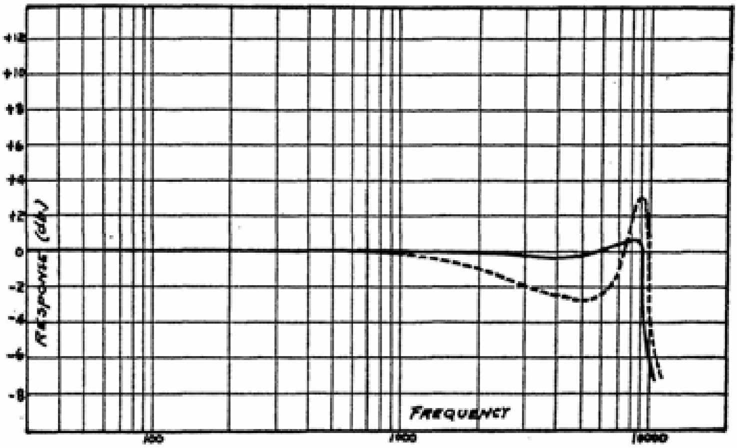


Figure 12 — Type of mask used for standard variable area recording.

The first variable area recording system employed an oscillograph consisting of a pair of fine wires to which the mirror was clamped. These wires were placed between the pole faces of a magnet and operated very similarly to the ribbons of a light valve. The complete assembly was immersed in oil to provide adequate damping. Later galvanometers are

of the moving iron type in which the armature causes the galvanometer mirror to vibrate through a simple mechanical link. These galva-



—*From the Journal of the S.M.P.E., September, 1937. "The RCA Recording System and Its Adaptation to Various Types of Sound Track," G. L. Dimmick.*

Figure 13 — Variable area recording system modulator response with frequency, with galvanometer damped (dotted curve), and with galvanometer damped and with bias condenser (full curve).

nometers are air damped, tuned to approximately 9,500 cycles, and are not critical to temperature changes.

The response with frequency of a modern galvanometer is shown in Figure 13.

The dotted line curve gives the response with no bias condenser in the circuit. The droop that occurs in the high-frequency range is caused by the inductance of the modulation winding. Earlier types of galvanometers obtained their damping by having the unit immersed in oil. Practical difficulties, however, led to the development of a damping assembly made of tungsten-loaded rubber which reduced the resonant peak at 9,000 cycles from about 12 db to about 3 db. Both the droop and peak shown in the dotted curve can be overcome and the frequency characteristic improved by placing a condenser across the bias winding to neutralize the inductive reactance. The effect of the bias capacitor upon the galvanometer response is shown by the full line curve in Figure 13.

With the variable area system certain care is required to obtain optimum results. Obviously, if the galvanometer is over-modulated,

the peaks of the signal will be cut off and distortion will result, while azimuth and focal errors, hysteresis, etc., must be minimized by correct design and adjustment of the galvanometer and lens system.

4. SLIT WIDTH

The physical width of the slit plays an important part in the quality of the resultant recorded sound, especially in the high-frequency band.

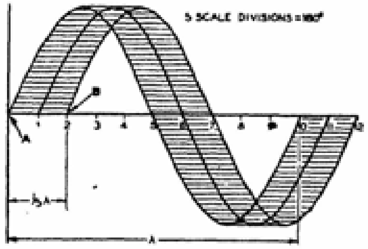


Figure 14-A.

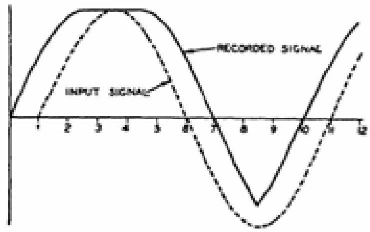


Figure 14-B.

Assume that a wave is to be recorded whose length is five times the slit width. One edge of the slit starts at the point A and the other edge at B as shown in Figure 14-A, with a resulting track as shown in Figure 14-B. It will be noticed that in one case the peak of the wave is lost and in the other accentuated—*but in both cases distorted.*

For this reason, the slit must be narrow in relation to the highest frequency to be recorded.

At 90 feet per minute the film travels 18,000 mils (1 mil = 0.001") per second, and a 9,000 cycle wave, for instance, would be two mils long. If a slit width of 1 mil is assumed, it can be seen that the

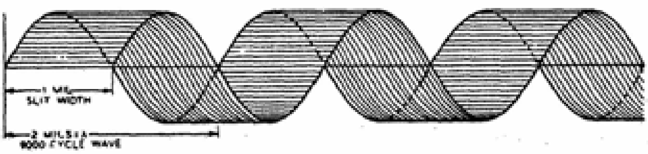


Figure 15-A.

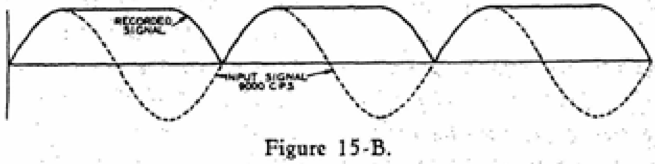


Figure 15-B.

slit width equals one-half the wave length. A study of Figure 15 shows that one-half of the wave is lost, that the other half is distorted, and the intensity changed.

Figure 15-A shows the movement of the top edge of the recording beam in relation to the long dimension of the slit, that is, in this drawing the light beam is shown moving in a vertical direction and the film in a horizontal direction toward the right. The input signal shown by the dotted line curve in Figure 15-B, is a sine wave of 9,000 c.p.s. The resultant recorded signal, shown by the full line curve in Figure 15-B, is the envelope of the area scanned by the recording beam, and shows the effect of slit width on intensity and wave form. With a slit width of 2 mils the response from a 9,000 cycle wave would be zero.

However, the lower limit of slit width is determined by other considerations, principally the difficulty of securing sufficient light to properly expose the film when very narrow slits are employed. The usual practice in the case of variable density is to employ an effective slit of ½ mil and in the case of variable area an effective slit of approximately 2/10 mil. The physical slit itself in both cases is actually considerably wider than these figures indicate, but is reduced by optical means.

5. NEW RECORDING METHODS

The constant quest by the sound departments of all the studios for a practical system of increased noise reduction which does not penalize recording methods or the quality of recorded sound, has led to refinements in the systems discussed above as well as to new recording techniques.

(a) Push-pull Recording

Sound which is to accompany motion pictures must be reproduced with the highest possible accuracy in order that the picture give the greatest possible illusion of reality. In viewing a motion picture, the audience is stimulated to a specific sound expectancy not necessarily present in the radio and phonograph.

In order to approach the ideal illusion of reality it is necessary that the recording and reproducing systems have a frequency response and volume range commensurate with the original. A high degree of linearity is essential. The frequency range should be from fifty to eight thousand cycles and the volume range from fifty to sixty db.

Reproducing systems which will meet these requirements are now available and are being installed in theatres, and the recording practice has been developed to a degree that also fulfills these requirements in the original recording form. However, it has been found necessary to limit the released volume range until such a time as the majority of theatres are equipped to handle the greater volumes.

To date, standard methods of recording on film, both variable density and variable area methods, have given a volume range of approximately forty db, at the expense of considerable distortion. In the variable density system the principal limitation has been the small linear range of density between the toe and shoulder of the characteristic curve for positive stock, plus the distortion added by the noise reduction system because of its relatively slow operating time. Push-pull recording inherently reduces these limitations by the cancellation of the internal distortions.

When sound is recorded on film, either by the variable area or variable density method, certain harmonic and extraneous components appear in the reproduced signal which were not contained in the original signal.

For the variable density recording method these consist mainly of two distortion components: one introduced by the non-linear toe and shoulder portions of the film characteristic, and the second, signal components derived from the noise reduction apparatus. For variable area recording the distortion consists of shutter or bias signals and processing effects.

Push-pull recording, like the push-pull method of operating vacuum tube amplifiers, is a method of reducing certain of these distortion effects in the electrical signal to be reproduced. The similarity between film recording and vacuum tube circuits is quite complete, the one providing a means for cancelling out certain harmonics produced in vacuum tubes because of non-linear tube characteristics, the other accomplishing the same results for film recording in connection with the above mentioned distortion effects. In fact, the possibilities for push-pull recording had their inception in the success of the vacuum tube development.

This type of recording is secured by means of two sound tracks on the film which are reproduced by a double photo-electric cell arrangement, which adds the signals from the two tracks in their proper phase relation. The two half-tracks are recorded out-of-phase on the film and the double photo-electric cell circuit is arranged to have a phase turnover, so that the signal will be reproduced in-phase (see Chapter XXII). This phase turnover is accomplished in the same manner that the outputs of two push-pull vacuum tubes are joined together by a center tapped transformer. Thus, even order harmonic content introduced by the film is cancelled out because of an out-of-phase condition in the combining transformer. As in push-pull vacuum tube circuits, it is not possible to cancel out both even and odd order harmonics.

In push-pull recording, noise reduction signals are recorded on the two half-tracks in-phase in contrast to the out-of-phase "wanted" signal. These former signals will, therefore, be cancelled out in reproduction by the photo-electric cell circuit due to its 180° phase change. Because of this cancellation feature, it is feasible, if desired, to speed up the operation of the noise reduction system (that is, allow the noise reduction signal to contain more frequency components), without damage to the reproduced signal content. This tends to decrease the amount of wave top clipping on steep wave front signals, which is a result of the sluggishness of operation of the noise reduction equipment. (See chapter on Noise Reduction.)

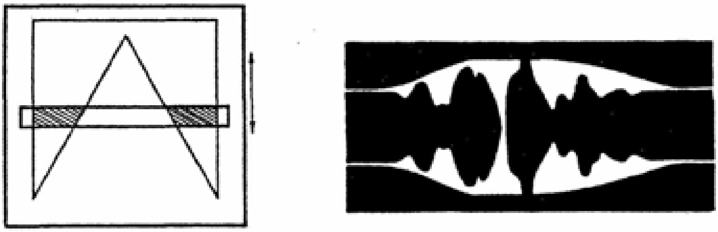


Figure 16-A — Duplex variable area.

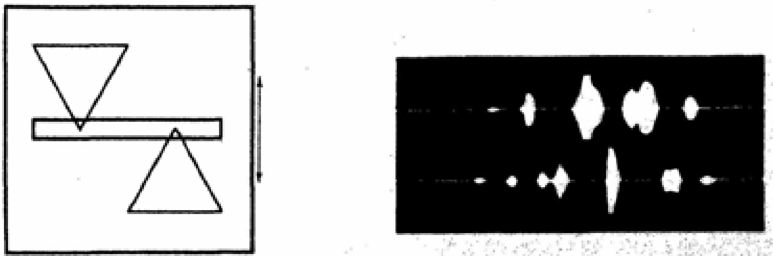


Figure 16-B — Push-pull Class B variable area.

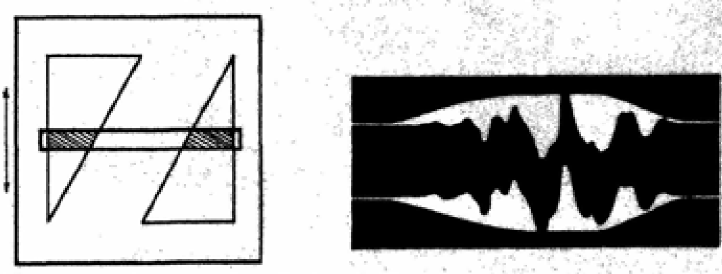


Figure 16-C — Push-pull Class A variable area.

Push-pull recording may be of three types: Class A, Class B, or Class AB—the class notation having the same significance as for amplifier operation.

In Class A recording, half the total signal energy is recorded on each half-track with the result that they are alike but out-of-phase as discussed above. (See Figure 16.)

In Class B recording, one half-track receives the positive signal energy and the other half-track receives the negative signal energy.

The adjustments necessary to produce each of these tracks are accomplished by proper external connections to the light valve in the case of variable density, and by the inclusion of the proper aperture in the variable area optical recording system.

As explained in the chapter on Noise Reduction, Class B track is not suitable, at least at the present time, for general theatre release, because of the necessity of maintaining an accurate sensitivity balance in the push-pull reproducer. Inequalities resulting from the processing of the film also add to the reproduced distortion.

The push-pull, Class A, variable area recording is accomplished by converting the optical system by change of apertures, using the same shutter in both cases. (See Figure 16.)

As the noise reduction signals are cancelled out by the reproducer circuit, increased speed of operation of the shutter, if desired, may be secured. In this case, as in reproduction from a variable density push-pull record, a considerable reduction in even harmonic distortion is obtained and distortions due to audible variations in average print transmission are cancelled out.

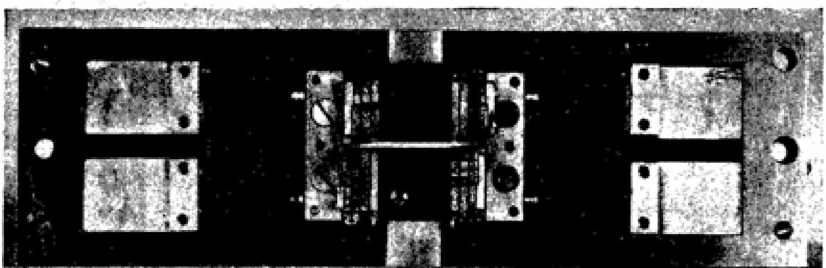


Figure 17 — Four-ribbon, push-pull light valve.

Class AB recording consists in adjusting the recording mechanism to record Class A for low volume signals and Class B whenever the signal volume exceeds a given amount. The point of separation between the Class A recording and Class B recording is determined by the setting

given to the recording mechanism. Class AB recording, up to this time, has been confined mostly to experimental work.

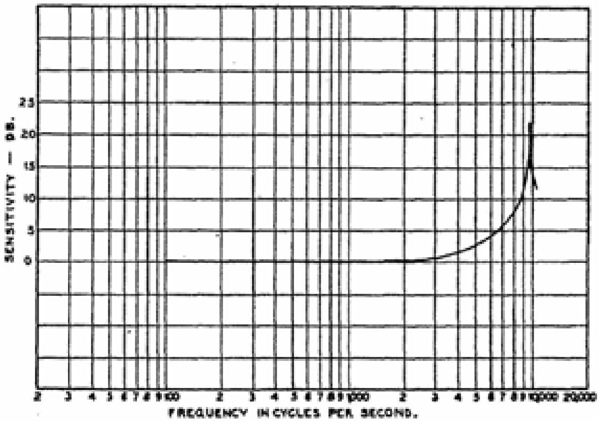


Figure 18 — Sensitivity characteristic for four-ribbon push-pull light valve (ERPI valve-RA-1061A).

For push-pull variable density recording, the light valve has been modified, as shown in Figure 17, to have two pairs of bi-planar ribbons of conventional type, each pair exposing half of the sound track, with the valve so connected that these ribbons act in push-pull. It is mechanically impossible to mount the ribbons in line with each other, so, in order to scan the track with a single slit, the image must be optically relocated on the film. This is accomplished by placing an optical flat in the path of each light beam in the form of a saw buck as shown in Figure 19. This arrangement moves the axis of the path in proportion to the angle at which the optical flat is interposed. As a result, the opposing densities on the two halves of the track are positioned horizontally across the film.

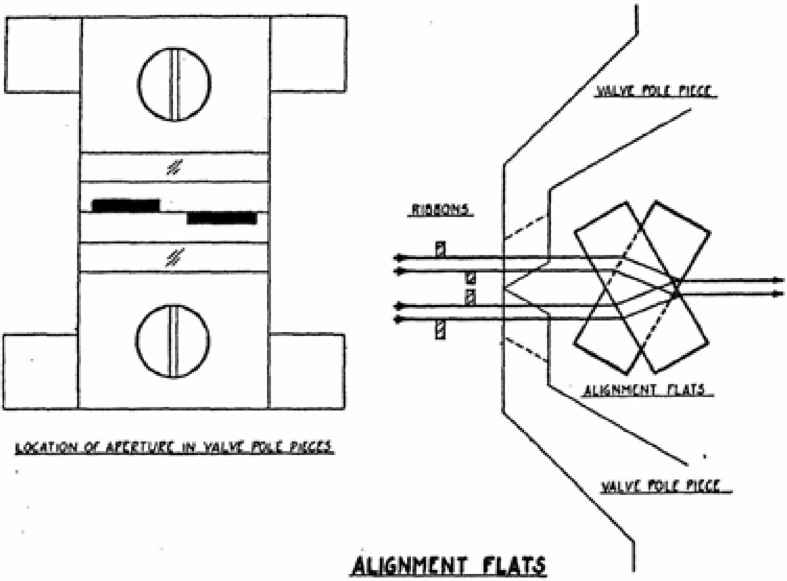
Uniform exposure, to give a uniform field across the width of the sound track, is obtained by adjusting the flat-ribbon filament lamp. These lamps, incidentally, must be carefully selected inasmuch as there is a wide variation in lamps from any one manufacturer.

If the two halves of the track are recorded in-phase, the output when reproduced on a balanced machine will be a measure of the cancellation.

In the variable area Class A push-pull system, the images of the triangular slits are so located as to be symmetrical with respect to the center of the slits, the center of the track being placed to correspond also with the center of the slits. The residual width of the track will then indicate a measure of the balance.

(b) Dimensions of Studio Push-pull Tracks

In the variable density Class A push-pull system, there are two



—*Courtesy Electrical Research Products, Inc.*

Figure 19 — Alignment flats of four-ribbon valve and associated optical system.

widths of sound track in current use in the studios: Regular (76 mils) and wide (200 mils). Regular push-pull occupies the same width as that of the original single track with the exception that there are two 35 mil tracks separated by a 6 mil septum, these tracks being scanned by the usual 84 mil slit. These dimensions are also used for push-pull movietone prints.

Wide push-pull tracks consist of two 90 mil tracks separated by a 20 mil septum. Each of these tracks is scanned by an 84 mil slit.

These three types of sound track, illustrated in Figure 20, have been adopted as Working Standards by the eight major studios. The dimensions shown in the larger sized numerals must be maintained as accurately as possible. Other values are nominal and for information only. The areas shown in black are controlled by matting the negative in the recorder—not by matting in the printing machine. This method, which limits the effective track width on the film itself, greatly reduces the effect of weaving in recorders, printers and reproducers. The volume reduction due to the slight decrease in track width in comparison to widths previously used, amounts to only a fraction of a decibel and is well justified by other advantages obtained.

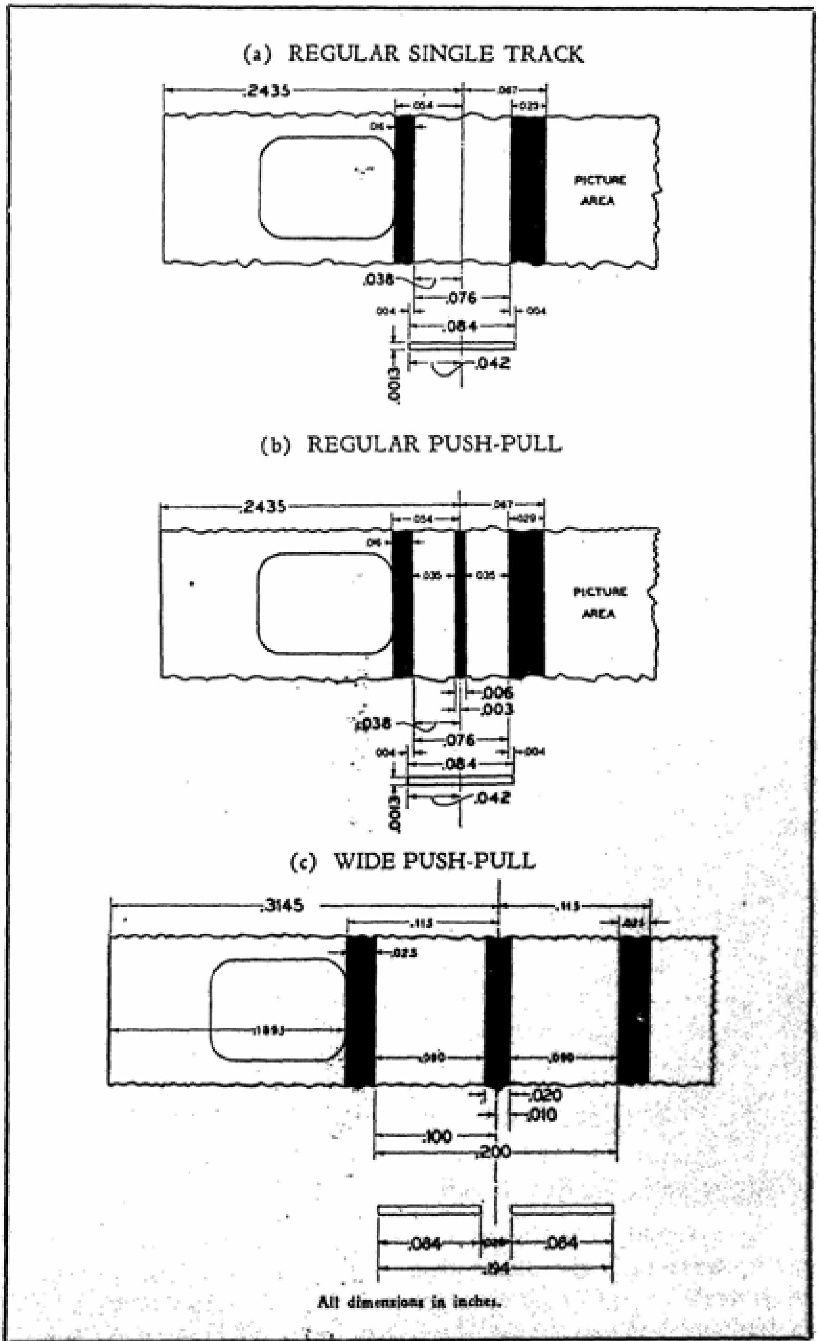


Figure 20 — Working Standards for: (a) Regular Single Track, (b) Regular Push-pull Track, and (c) Wide Push-pull Track.

For studio use the wide push-pull track has several advantages Some of these are:

1. Greater inherent fixed noise reduction is obtained, thus reducing the amount of bias current necessary for quiet operation. This decrease in bias current also reduces the breathing effect.

2. Regular and wide push-pull tracks may be intercut on the same reel and projected by a push-pull type reproducer, adjusted for the wide track. This is accomplished by printing the regular track through a wide aperture and in the area occupied by one of the two wide push-pull half-tracks. When the regular and wide tracks are intercut the two will reproduce on the same system, since the regular track utilizes only one side of the push-pull photocell arrangement as the track in the path of the other scanning slit is black. Hence, with a single adjustment the one reproducer will play both types of track with a difference only in volume efficiency when a change takes place from one track to the other.

3. Failure of one pair of ribbons still provides a standard track in case of emergency.

4. The track is separated from the sprocket holes by 25 mils compared to 16 mils in the narrow track. This increased barrier provides more isolation from 96 cycle sprocket hole modulation.

The advantages of the regular push-pull track, compared to single track, are:

(1) The same type of reproducer can be used for original as well as movietone push-pull.

(2) A four-to-one optical system can be used, which materially reduces the effective slit width on the film as well as a reduction in intermodulation provided ribbon amplitudes are limited in the case of variable density.

(3) Greater inherent noise reduction as discussed above.

(c) Squeeze or Matted Track

As push-pull reproducing systems have been installed in only a small percentage of the theatres, new techniques have been applied to single tracks, in order to secure a greater amount of noise reduction by a method which allows the use of the same type reproducer as previously used.

In the single track variable density method of recording, noise reduction by means of bias has been increased to the maximum consistent with good quality. Since this has resulted in a volume range still considered inadequate, other means of obtaining increased noise reduction must be used. As volume range is a function of the signal to noise ratio, equivalent noise reduction may be obtained by decreasing the width of the track and simultaneously increasing the percentage of modulation where the required output is considerably less than maximum.

Figure 21 — Squeeze or matted track.

A track so reduced is commercially known as a "squeeze" or "matted track."

This technique can be applied with excellent advantage to very low-volume scenes which are exemplified by whispers or subdued conversational passages, or in the use of a very low musical background to silent scenes.

One method used to obtain matted tracks was the insertion of an inverted V-type mask in the optical path, usually in fixed steps which reduced the track width to 25% or 50% of the available width.

The use of the above V-type squeeze track requires that the slit of the reproducer be evenly illuminated in order that the increase in modulation and decrease in track width exactly compensate. If this illumination is not even, matted portions will be reproduced at different levels depending upon the degree of non-uniformity of the illumination.

To overcome this difficulty, a more recent procedure is to use a W-type mat which reduces the track width from the center as well as from both outside edges (see Figure 99), and so lessens the effect of non-uniform slit-illumination.

This type equipment, which is schematically illustrated in Figure 22, also provides a continuously variable mat. The attenuator motor and the matting motor are driven by a foot controlled Selsyn motor sys-

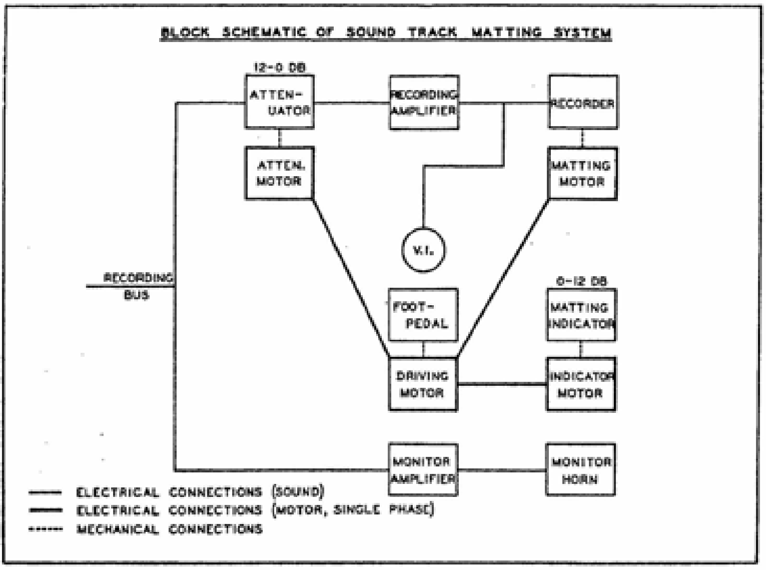


Figure 22 — Schematic diagram of equipment connections for squeeze track recording.

tem which provides the continuously variable mat, and simultaneously removes the proper amount of attenuation from the circuit to increase the modulation. The matting indicator (driven by the indicator motor through the driving motor), shows the operator the amount of matting applied, and it will be noticed that the amount of attenuation provided is exactly the inverse of the amount of matting available. From the volume indicator meter the operator is able to determine the amount of modulation upon the film and by proper control of the foot pedal, apply the necessary amount of matting which, through the simultaneously operated motor system, automatically provides the correct amount of attenuation.

Figure 21, an example of one type of "squeeze" or "matted track," shows in the upper portion of a sound track with no noise reduction applied. The center portion has a 6 db and the lower portion a 12 db noise reduction, a reduction in track width to 50% and 25%, respectively.

(d) Complementary Recording

Complementary recording provides a method, for use as an adjunct to present type noise reduction systems, to secure increased noise reduc-

tion and to partially suppress breathing effects. In addition to these benefits, inter-modulation between low and high frequencies of signals is reduced, and some overloading of the recording mechanism is avoided on abnormally high-volumed, low-frequency signals.

Modern methods of recording and reproduction do not permit much extension of the amount of noise reduction which can be secured in practice. This is due to several effects which accompany the use of biased noise reduction systems, such as breathing, distortion effects caused by the sluggishness of the bias component, and other items well known to the industry. For the method described below, these items, limiting the use of standard noise reduction systems, are partially overcome by distributing the signal load more uniformly upon the recording mechanism with respect to frequency. This is accomplished by pre-equalization and post-equalization methods of a type different than those heretofore employed for sound-on-film recording.

Figure 23 shows the circuits of two equalizers, one for use in the recording circuit to pre-equalize the recorded sound material, and the other for use with the corresponding reproducing channel to post-equalize this circuit in a complementary manner. When these two equalizers are so used, the amplitude-frequency characteristic of the reproduced sound material is unchanged because of the complementary characteristics of the two equalizers. Throughout the complete recording and reproducing channels, more than normal gain is needed to compensate for the equalizer losses, which are 14 db for the equalizers of Figure 23. Referring to the figure, it is noted that the recording equalizer has an insertion loss of 14 db at 100 c.p.s., 7 db at 1,000 c.p.s., and very little loss for the high frequencies. In general, most of the transition from 14 db loss to zero loss occurs in the frequency range from 300 to 3,000 c.p.s. The half-loss frequency of 1,000 c.p.s. is one of the design parameters of the equalizers.

A large part of the energy content of sound signals lies in the low frequency range from 200 to 500 c.p.s. Insertion of the above pre equalizer into a normal recording channel without any change in channel gain removes a large part of the load from the recording mechanism and the film, leaving a high-frequency content of approximately the same level. Due to the removal of the low-frequency load, it is now possible to increase the recording channel gain, thus securing a greater ratio of high-frequency signal to static surface noise; that is, an increase in noise reduction. Subsequent post-equalization does not destroy this increased noise reduction because of the concentration of surface noise in the upper part of the frequency spectrum. In addition to this increase

in noise reduction, a further increase is sometimes obtained due to the removal of restrictions applying to the noise reduction equipment as discussed below.

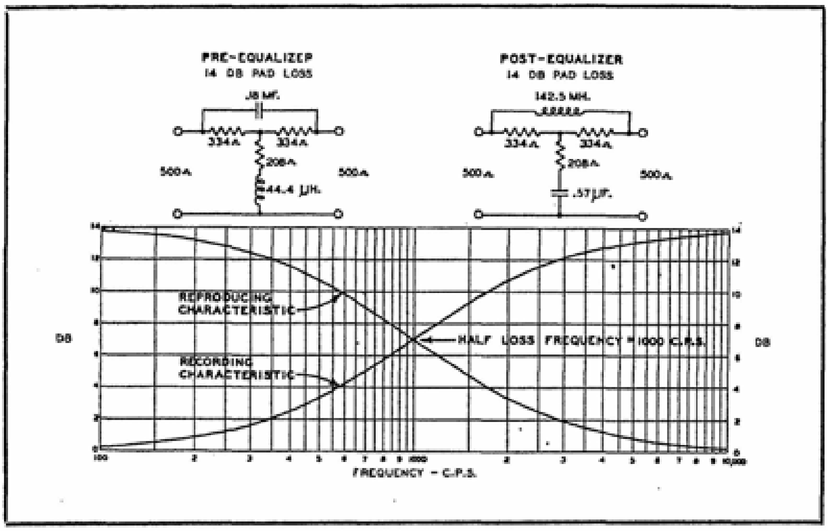


Figure 23 — Equalizer circuits used in complementary recording.

Breathing is caused by the action of the noise reduction system in placing the modulator in a position to handle the signal. Normal operation of noise reduction systems consists of a bias on the modulator or a reduction in track width, so that the minimum amount of static surface noise is present for the no-signal condition. Opening-up of the recording mechanism or track width by the rectified bias signal increases the surface noise, causing the effect known as breathing. The greater the volume of the recorded signal, the greater the breathing effect.

As already mentioned, most of the signal energy lies in the low-frequency range which means that much of the breathing is produced by the lower frequency components of the signals. Hence, use of the above type of pre-equalizer greatly decreases breathing because the recording mechanism is not modulated nearly as much by the low-frequency content. Tests have shown that breathing is practically eliminated by this method even on piano and organ tracks. It may be mentioned that in normal recording, breathing caused by the high-frequency content of signals is not nearly as objectionable as that from low-frequency components because of the low-energy content and masking effects.

Some other beneficial effects derived from this type of pre-equalized recording are: Possibility of the use, in some cases, of a greater amount

of noise reduction because breathing, which normally tends to limit the amount of noise reduction that can be used, is reduced; the reduction of spurious signal products, arising from intermodulation, because of the lower level of the low-frequency components; the reduction of bias current components for the same reason; and the reduction in the amount of wave top clipping on steep wave front signals which is normally caused by the sluggishness or time lag of the bias signal. (See chapter on Noise Reduction.)

In conclusion, it may be mentioned that this method of recording accomplishes beneficial results for two basic reasons; first, because the energy distribution of acoustic signals lies primarily in the lower part of the audible frequency spectrum and, second, because film surface noise is concentrated in the upper part of the frequency spectrum. Reducing the level of the recorded signals in the low-frequency range does not materially increase the signal to static surface noise ratio in this same range, but on the contrary permits the over-riding of surface noise in the upper-frequency range. This, in connection with the reduction in breathing, decreased intermodulation effects, and the other items outlined above, constitutes the benefits which may be derived from this method of recording.

At the present time this method is in current use on original recordings, the post equalizer being used at the time of re-recording for standard release. It is hoped that within the very near future a sufficient number of theatres will be equipped with the post equalizer in order that the full benefits of the system may be obtained by release on the movietone print.

6. PHOTOGRAPHIC REQUIREMENTS OF VARIABLE AREA AND VARIABLE DENSITY SYSTEMS

The primary requirement of any optical recording and reproducing system is that the relative amount of light falling upon the photoelectric cell be, at each instant, proportional to the amount of light passed by the exposure device in recording.

This requirement is shared by both systems, but due to the difference in the exposing devices in the two cases, a different use of the film characteristics is employed, as covered in detail in the chapter on Film Processing.

Chapter IV

NOISE REDUCTION

By FRED ALBIN

Theoretically, an ideal electrical recording and reproducing system is one which will produce an electrical output similar in every respect to the electrical input (see Figure 24). Any deviation between the two is known as distortion, the two commonest forms of which are amplitude distortion and frequency distortion. Practically, how-

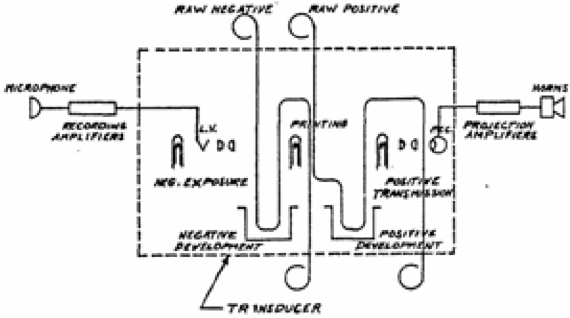


Figure 24 — Elements of sound recording and reproduction system.

ever, such a system will always present some distortion and will differ from the ideal in several respects, one of which is the introduction of "ground noise" during the recording process.

A measure of the efficacy of such a system is the volume latitude; the maximum limit of which depends upon the overload level of the system and the minimum limit of which depends upon the ground noise present on the record. Figure 25 illustrates this, showing the maximum and minimum levels and giving the volume latitude as the difference between these two levels. It is evident, therefore, that a reduction of the noise level increases the volume latitude and so enhances the value of the system.

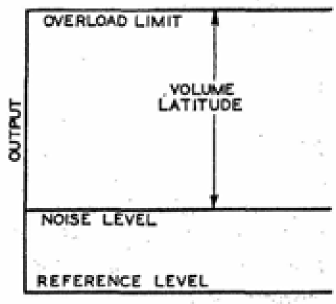


Figure 25—Diagram illustrating the volume latitude of reproducing systems.

This is further brought out by referring to Figures 92 and 96, where it can be seen that the difference in level between the two curves at any particular volume-value is the volume latitude of the system in question.

It will be noticed that the slope of the noise level curve and the minimum volume curve are equal, as any ground noise which is on the film will be amplified proportionately to the "wanted" signal contained on the film. Therefore, increasing the power of the installation will not increase the volume latitude unless the ground noise is held constant, or in other words, unless some system of noise reduction is used.

This chapter will be devoted to the manner in which noise reduction may be accomplished and to the methods of application to the variable area and variable density recording systems.

1. METHOD OF APPLYING NOISE REDUCTION TO VARIABLE AREA RECORDINGS

The variable area method will be discussed first, as the application of noise reduction is more easily visualized due to the type of track produced by this system. As explained in Chapter III, the first type of variable area record consisted of a track which had the appearance of an oscillograph record on which the area to one side of the oscillogram was transparent and the other side opaque. This is illustrated in Figure 26, where it can be seen that one side of the track is quite clear (except for

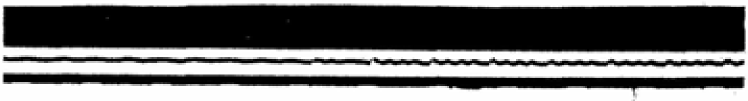


Figure 26-A — Oscillogram of a speech wave.

Figure 26-B — A single envelope variable area recording of the speech wave of Figure 26-A.

—*Courtesy RCA Manufacturing Co.*

fog grains of silver, dirt, abrasions, etc., present on and in the clear film, but which are not discernible in the cut). These foreign particles, in passing the scanning aperture, interrupt the light beam in the same manner as a modulation wave, and so produce undesirable noise in contrast to the signal itself, which is producing the wanted sound. Since these particles are located at random throughout the clear area of the track, their power of modulation is proportional to the width of the clear track area (Figure 27). Without modulation, that is, with no signal recorded on the track and without noise reduction, the widths of the

clear and opaque areas are equal. The division line between these two areas is known as the zero line or base line, and is the axis of the recorded wave during modulation. The width of the clear area determines the maximum undistorted amplitude of the modulation, that is, any

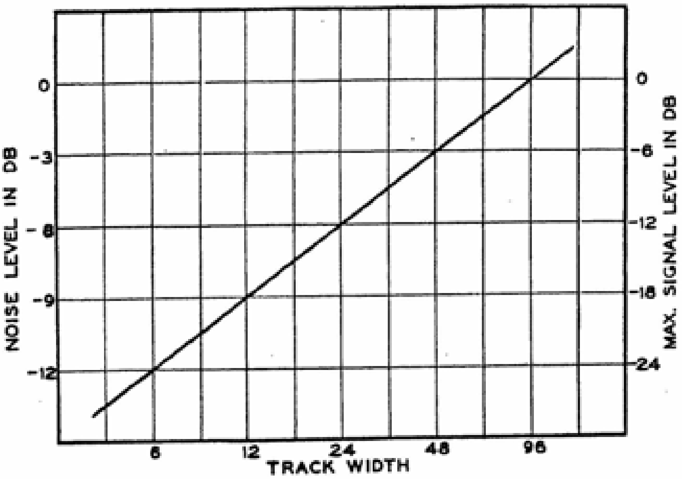


Figure 27 — Relation between track width and noise level and maximum signal level.

signal which has an amplitude greater than one-half of the track width will obviously extend beyond the track, and the peaks of the waves will be lost and harmonics introduced. At low modulation levels the clear area width is excessive and as a consequence the foreign particles mentioned above are present in a relatively large extent, which results in a serious prominence of ground noise. The obvious remedy is the reduction of the excessive width of the clear portion of the track during low modulation levels. Such a system, when applied to recording methods, is known as "noise reduction."

The first experimental plan for accomplishing noise reduction in the variable area system consisted of applying a direct-current bias to the modulator circuit. This direct-current bias caused the zero line to be shifted toward one side of the track, during no-signal periods, thus reducing the width of the clear area to some pre-determined value. During modulation periods, part of the signal energy was diverted from the modulator and used to reduce this bias current, thereby moving the zero line toward the center of the track to allow sufficient track to record the increased modulation.

However, mechanical difficulties were introduced by the weaving of the film when passing through the projector, as the modulation at low levels often moved outside the area of the scanning beam and part

of the wave was lost. Consequently, this system never was used commercially.

This difficulty was overcome by removing the bias current on the galvanometer and adopting a biased shutter which moved into the track area from the clear side of the film toward the zero line at low-modulation levels, leaving the same net width of clear track with the zero line

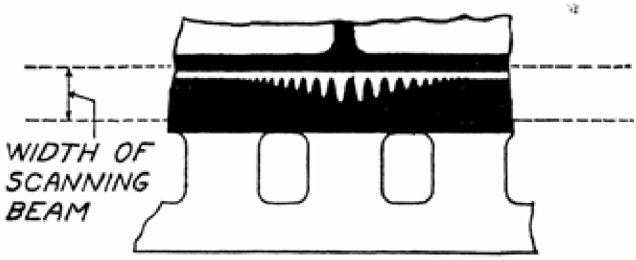


Figure 28-A — Variable area sound track with noise reduction accomplished by variable bias of galvanometer.

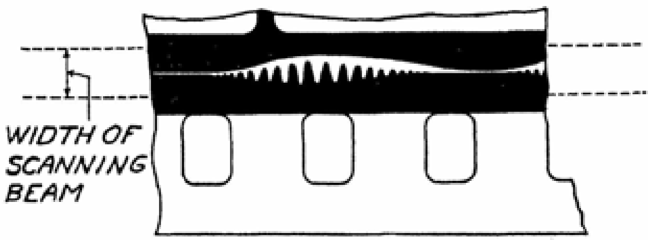


Figure 28-B — Variable area sound track with noise reduction accomplished by use of shutter.

—*From the Journal of the S.M.P.E., August, 1931. "A Shutter for Use in Reduction of Ground Noise," E. W. Kellogg and C. N. Batsel.*

still located in the center of the whole track, and avoiding the loss of modulation resulting from weave of the film. The bias current on the shutter is reduced, at high-modulation levels, by using part of the rectified signal energy to cancel this bias current. Figure 28-A illustrates an experimental sample recording made under the first plan, and Figure 28-B represents a sample recording made under the second plan. These single envelope variable area records are made with an aperture as shown in Chapter III.

Later, a different type of aperture, in the form of an isosceles triangle, was adopted to replace the original square aperture. An image of this triangle is moved in a vertical direction, and at right angles with respect to the scanning slit, as shown in Figure 29-A, resulting in a record as shown in Figure 29-B.

The image of the common area of the optical slit and triangle (shaded area in Figure 29-A) thus varies in horizontal length in accordance with the signal, and the track produced is known as a bilateral

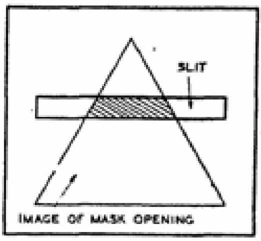


Figure 29-A — Type of aperture used in bilateral variable area recording.

—Courtesy RCA Manufacturing Co.

Figure 29-B — Sample of variable area bilateral recording.

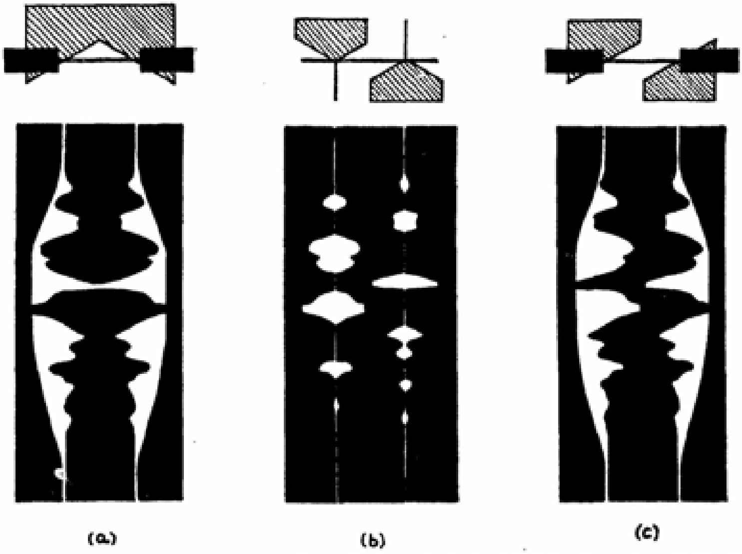
track, as it is a double envelope track identical in modulation on both sides of the zero line. The modulator may be biased to reduce the horizontal length of the image which results in a corresponding reduction of the area of the clear portion of the positive track. This method accomplishes noise reduction without the mechanical difficulties experienced with the original method.

From this followed the adoption of a double track system, which is produced by replacing the single triangular aperture with two opposing triangular apertures as shown in Figure 30-(b). An optical image of the two triangles is formed at the scanning slit after being reflected from the galvanometer mirror. The apexes of the triangular images fall on the center of the slit at the "no-signal" position, and are spaced apart half the length of the slit. The signal to be recorded causes the galvanometer to vibrate and to shift this image, according to the signal, in the direction shown. Thus, one triangular image will record the positive portion of the signal, and the other image the negative portion, which produces a track as shown in Figure 30-(b). This type track is known as "push-pull, Class B, variable area track."*

If the aperture is so designed that the mid-point of the altitude of both triangles falls on the center of the scanning slit at the "no-signal" value, Figure 30-(c), the resulting recording is known as "push-pull, Class A, variable area track."* This aperture was so modified that

*NOTE: As will be explained later in this chapter, Class B push-pull records are unsuitable for release print purposes at the present stage of advancement of reproducing equipment. For this reason, Class B push-pull recording, in both the variable area and variable density systems, is not used, and for purposes of release print nomenclature the terms "Class A" and "Class B" are dropped, and the term push-pull record always designates a push-pull Class A track.

the image falling on the scanning slit consisted of two opposing right triangles instead of the two opposing isosceles triangles, in order that noise reduction of the shutter type might be easily applied. The re-



(a) (b) (c)

—From the Journal of the S.M.P.E., September, 1937, "The RCA Recording System and Its Adaptation to Various Types of Sound Track," G. L. Dimmick.

Figure 30 — Types of variable area apertures with noise reduction shutters, with the resulting sound tracks.

sulting track is shown in Figure 30-(c), which is similar to the track of Figure 29-B, but with the two half-tracks 180° out-of-phase.

The most recently developed aperture and the one currently used in production, except where push-pull is employed, is shown in Figure 30-(a). This "M-shaped" aperture is a negative of the single track aperture of Figure 29-A and its use results in a track shown in Figure 30-(a). Noise reduction, that is, reduction in the width of the transparent portion of the track, is secured by the use of a double-shutter mask, each shutter biased toward the center of the track during periods of low modulation. Part of the signal energy is diverted from the modulator, rectified, and then used to cancel the biasing current on the shutters to allow greater track width at periods of high modulation, as illustrated in the center portion of Figure 30-(a).

In sound recording, the majority of signals are unsymmetrical about their base line,* having greater amplitudes during the compression part

* While this condition appears to be true, there is considerable difference of opinion among authorities on the subject.

of the wave. For this reason the galvanometer should be so phased with the sound wave that the compression part of the wave is recorded toward the center line, and away from the shutter side of the track. This provides a greater safety margin for high-peak waves as the total amount of available track on a bilateral record may be instantly utilized by the compression part of the signal. However, on both outside quarters of the track, the shutter, during periods in which the noise reduction is operating, masks part of the track, and high amplitude waves are apt to be cut off during the time the shutters are being removed from the sound track area. This timing action of the shutters is described in detail below.

2. METHOD OF APPLYING NOISE REDUCTION TO VARIABLE DENSITY RECORDINGS

The variable density recording system employs as a light modulator the light valve, which has a normal spacing between the two ribbons of one mil. Noise reduction may be applied to such a system by reducing the mean spacing of the ribbon to some predetermined value, resulting in a reduction of negative exposure with a consequent reduction of positive transmission. This has an effect similar to reducing the clear portion of the track of the variable area record. Part of the incoming signal is then diverted through an appropriate network to act in such a way as to increase this mean ribbon spacing as the modulation of the signal increases. This increase of mean spacing is sufficient to accommodate the increased input.

More recent designs of variable density recorders use a four-ribbon valve to expose two tracks in the area formerly occupied by a single track. The exposure of one track is controlled by one pair of ribbons and of the other track by the other pair. The modulation of each pair of ribbons may be in-phase or out-of-phase by 180°, depending upon the external connection. An in-phase relation results in a track which is the same as that produced by a two-ribbon valve with the exception that a septum divides the sound track into two parts. An external connection resulting in an out-of-phase condition produces a push-pull record.

This push-pull record may be either Class A or Class B. In the Class A type each pair of ribbons receives one-half the total signal energy, and has a fixed spacing of over one mil. The resulting track consists of two parts, one part being identical with the other, yet with the two parts 180° out-of-phase. A sample Class A push-pull, variable density recording of low frequency is shown in Figure 31-A.

In the Class B type of recording the ribbons are still connected 180° out-of-phase, but in this case one pair of ribbons receives the positive half and the other pair receives the negative half of the signal energy. This results in a type of track which is very similar in appearance to the Class A push-pull record. A sample of Class B push-pull variable density recording is shown in Figure 31-B.

While it will be shown later that the Class B system of push-pull recording is ideal from a standpoint of maximum noise reduction, it

—Courtesy Electrical Research Products, Inc.

Figure 31-A — Sample of push-pull, Class A, variable density recording.

Figure 31-B — Sample of push-pull, Class B, variable density recording.

NOTE: Figure 31-A is the usual 100 mil sound track used on release prints (actual size), while Figure 31-B is a 200 mil sound track frequently used in studio work (actual size). The 200 mil track is of course reduced to 100 mil track for release.

has, nevertheless, been superseded entirely by the Class A system because of practical difficulties. Reproduction of Class B recording requires very careful balance between the positive and negative halves of the cycles in both an amplitude and phase relation if low distortion requirements are to be fulfilled. When the recording is reproduced, the relative sensitivities of the two photo-electric cells and their associated optical and electrical circuits must be maintained exactly in balance for all frequencies within the recording range.

In practice it has been found difficult to maintain this necessary balance and the fidelity sacrificed was not offset by the improvements realized in noise reduction.

The problems encountered in the variable density system, due to the film characteristics employed by this system, add still further to the distortion by the Class B type recording. This can be seen by reference to Figure 36, where it should be noted that the curve below the

value of T_1 becomes non-linear and as this section of the curve is utilized by the Class B recording, distortion results.

3. DEFINITION OF TERMS

Before taking up the effect of noise reduction on the reproduced noise level from both variable area and variable density records, it will be necessary to establish the definition of certain terms.

In order to explain, in common terminology, the operation of noise reduction as applied to both the variable area and variable density systems, a nomenclature has been chosen, which, while not elucidative to either particular system, is descriptive when applied to both systems.

The total track width, which is the area of the film scanned by one projection scanning beam, will be termed "track." The "carrier" is the medium which, when varied in magnitude, constitutes recording. This is not to be confused with a popular specific type of noise reduction control circuit known as the "Carrier Noise Reduction System," described later. The carrier amplitude in the variable area system will be defined as the width of the clear portion of the track and in the variable density system as the mean positive projected transmission. The percentage modulation, that is, the percentage of the track covered by modulation, is the ratio of the modulation amplitude to the carrier amplitude expressed in per cent. Margin, which is the excess of the opening of the valve over the necessary amount to carry the signal (variable density), or the amount of clear track between the peak amplitude and the modulation and that portion of the track masked by the noise reduction apparatus (variable area) is expressed mathematically as

$$\text{Margin} = 20 \log \left(\frac{100}{\text{percentage modulation}}\right) \text{db} \qquad (8)$$

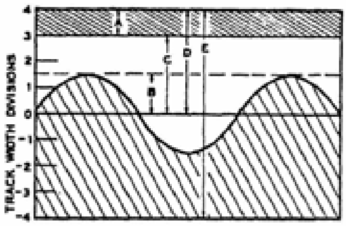
$$= 20 \log \left(\frac{\text{carrier amplitude}}{\text{modulation amplitude}}\right) \text{db} \qquad (9)$$

For example, if the carrier amplitude in the variable density system is 25% and the modulation amplitude is 12½ %, the percentage modulation is 50% and the margin is (20 log 2) 6 db.

The physical meaning of these terms is shown in Figure 32.

In practice, some value of margin in the order of 6 db is used for several reasons: First, there is a time delay in the operation of the noise reduction equipment and the margin assists the bias cancellation in its

speed, that is, it allows the noise reduction equipment a longer period of time to clear the way for a signal of increasing modulation; second, the adjustment of margin is usually made with a sine wave, which does not have as large a peak factor as do wave forms encountered in ordinary



A = Amount of noise reduction.
B = Modulation amplitude.
C = Carrier amplitude.
D = Clear portion of track with no signal and no noise reduction.
E = Track.

$$\text{Percentage Modulation} = 100\,\frac{B}{C} = 100\,\frac{1.5}{3} = 50\%$$

$$\text{Margin} = 20 \log \frac{100}{50} = 20 \log 2 = 6 \text{ db}$$

Figure 32.

recording, and since the response of the noise reduction control circuits is usually approximately proportional to the effective value of the input level, its response when predetermined by the sine wave would be insufficient to accommodate the higher modulation of a complex wave.

4. DISTRIBUTION OF NOISE IN THE RECORDING BAND

Investigations of the different sources and causes of noise have been made and the results indicate that if sufficient care is taken in the processing of the film, the greatest contribution to noise is the irregularity of the grain grouping of the photographic emulsion, as well as the irregularity in transmission characteristic of the gelatin and the base of the film itself. Again, after the film has been in use for some time it accumulates various kinds of dirt, becomes scratched, and otherwise depreciates from an ideal medium for sound recording.

Figure 33-A gives the distribution of the noise level over the frequency range used in recording, while Curve B, Figure 33-B, gives the *effective* noise level due to the non-uniform intensity curve (Curve A) of the ear. It must be noted that there is actually more energy in the low

frequencies than in the high, but the sensation level curve of the ear (Curve A) indicates that frequencies in the mid-range are actually more disturbing.

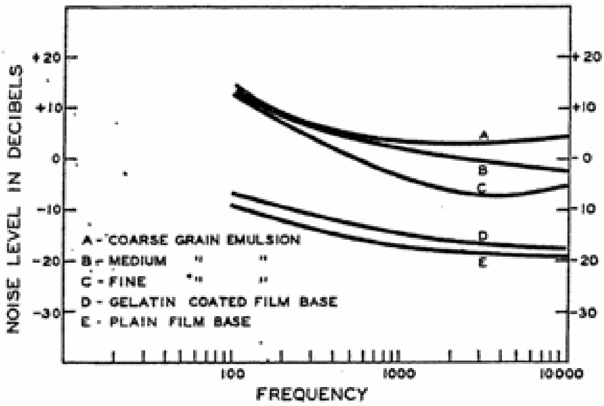


Figure 33-A — Distribution of noise level over the frequency band used in recording.

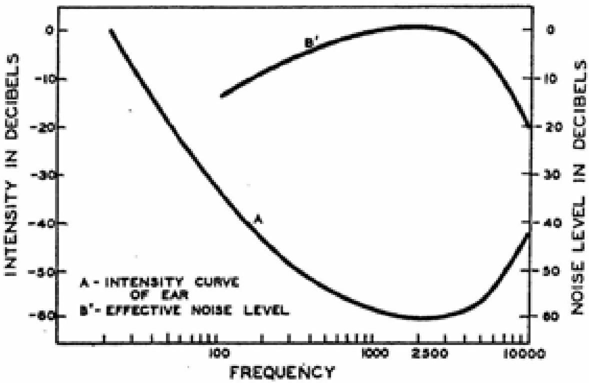


Figure 33-B — Curve A shows the sensitivity of the ear as a function of the frequency, while Curve B′ shows the effective intensity of noise produced by a film record over the frequency range used.

5. "CARRIER REDUCTION" METHOD OF NOISE REDUCTION

As previously mentioned, the most commonly used method of noise reduction is the "carrier reduction" method, commonly known as either the "bias" or "shutter" method depending upon the choice

made, wherein the carrier amplitude is reduced to a value which will accommodate the modulation amplitude with only a small margin for tolerance. Such a system is illustrated in Figure 34.

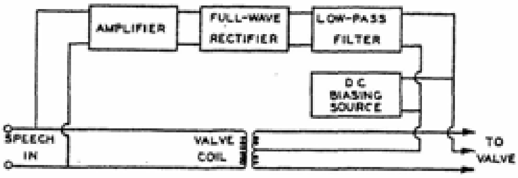


Figure 34 — Block schematic of a conventional noise reduction system.

The reproduced noise level is usually expressed in terms of amplitude and from a variable area record the reduction of amplitude is proportional to the square root of the reduction of the clear track width.

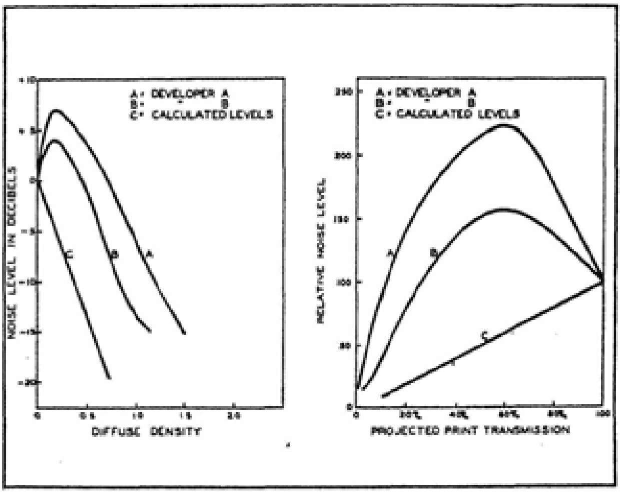


Figure 35-A — Curves showing the theoretical and measured relations between noise level and density of positive film developed in different developing solutions.

Figure 35-B — Curve showing the theoretical and measured relations between the relative noise level and the projected print transmission of positive film developed in different solutions.

For example, if a reduction of the noise level of 4 to 1 is desired, it is necessary to reduce the width of the clear portion of the track to 1/16 its initial value.

From variable density records the noise level is also proportional to the reduction of power, but in this case the amplitude of the ground noise level bears a close relation to the transmission of the film over a wide range of density, as shown in Figures 35-A and 35-B. For example, if a reduction of noise level of 4 to 1 is desired, the mean positive projection transmission must be reduced in the same ratio. This transmission is proportional to the valve spacing and consequently the desired noise reduction is accomplished by reducing the valve spacing by the desired ratio. (See Appendix, Page 527.)

In both the above cases the noise reduction is expressed in decibels as 20 log (reduction of amplitude of noise level) = 12 db.

This may be illustrated by considering the alternating-current analogy. The foreign particles present in the clear portion of the track may be considered as distributed more or less uniformly and these contribute to the reproduced noise level in a power ratio according to the amount of clear track projected. In the electrical case, if two equal alternating currents which have the same frequency and which are in-phase are impressed upon a circuit, the resultant current is equal to the sum of the two. The power in the circuit, however, is proportional to the square of the current and therefore the power has been increased four-fold by doubling the current. This condition is similar to noise reduction in variable density as any change in the transmission affects all noise sources simultaneously and the reduction of noise is proportional to the reduction of light. Therefore, the reduction of noise level is proportional to the reduction of mean positive projected transmission.

However, if the two currents mentioned above are not in-phase or are of a different frequency although of equal amplitude, the resultant current does not have the same linear relation to its components. In this case, the resultant current is equal to the square root of the sum of the squares of the component currents, or, if of equal value, to the square root of 2 times the value of either component current. Thus, the power, when the current is doubled, is also doubled. This is similar to the variable area case and shows that the reduction in noise level is proportional to the square root of the reduction of clear track width.

Therefore, to secure a 12 db reduction in the noise level, it is necessary to reduce the positive projected transmission in variable density recording to one-fourth its original value, while in the variable area system it is necessary to reduce the clear portion of the track to one-sixteenth its original value.

The following is an example of the application of this amount of noise reduction to both systems of recording:

Let it be required to reduce the noise level to one-fourth its value when noise reduction is not used. Then if the subscript "1" designates the unreduced noise level and the subscript "2" designates the noise level at one-fourth the original value:

$$NL_1 = 4\, NL_2 \qquad \text{where } NL = \text{noise level amplitude}$$

Variable Area

$$NL = K\sqrt{\text{clear track width}}$$

$$NL_1 = K\sqrt{CTW_1}$$

$$NL_2 = K\sqrt{CTW_2}$$

but $$\frac{NL_1}{NL_2} = 4 = \frac{\sqrt{CTW_1}}{\sqrt{CTW_2}}$$

and $$\frac{CTW_1}{CTW_2} = 16$$

$$CTW_2 = \frac{1}{16}\, CTW_1$$

Variable Density

$$NL = K\ (\text{positive projected transmission})$$

$$NL_1 = K\ (PPT_1)$$

$$NL_2 = K\ (PPT_2)$$

but $$\frac{NL_1}{NL_2} = 4 = \frac{PPT}{PPT}$$

and $$\frac{PPT_1}{PPT_2} = 4$$

$$PPT_2 = \tfrac{1}{4}\ PPT_1$$

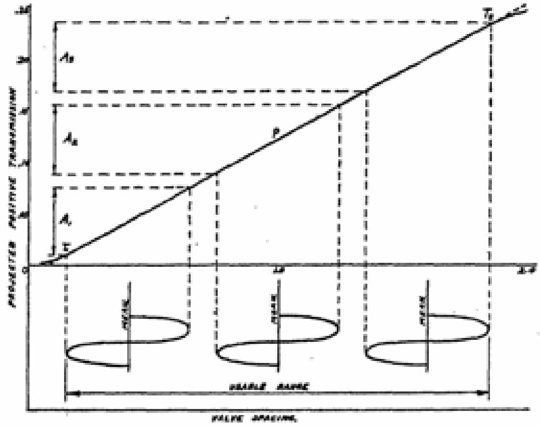
In both cases the amount of noise reduction expressed in decibels is the same, namely:

$$NR_{db} = 20 \log \left(\frac{NL_2}{NL_1}\right)$$

$$= 20 \log\ 4 = 12 \text{ db}\ .$$

Returning again to the variable density system and considering the combined steps of film processing as a whole, the positive projected transmission is proportional to the valve spacing. The signal input, which controls the variation from a predetermined mean of valve spacing, therefore causes a resulting change in the positive projected transmission, provided the relations given above are kept linear. The design of the photo-electric cell circuit serves to convert the change of transmission to an alternating electrical current whose amplitude is proportional to the change in transmission. The mean transmission, therefore, has no effect upon the output signal level which is determined solely by the above alternating current—but the mean transmission represents the source of, and determines the value of, the ground noise.

The minimum ground noise level is therefore reached when the mean transmission is reduced to the lowest allowable value. The lowest allowable value is limited by the straight line portion of the curve, which gives the relation between negative exposure and positive projected transmission. The mean exposure must be greater than this lower limit by an amount equal to the amplitude of the modulating wave, and so the minimum mean exposure is directly dependent upon the modulation level. In this respect the variable density and variable area systems of recording are similar. As shown above, the reproduced signal level, being proportional only to the modulation amplitude or change of transmission, is independent of the value of the mean transmission and, as a consequence, a reduction in this mean transmission, provided it is kept above the minimum value determined by the modulation amplitude, leads to a reduction in noise level at a constant output level.



—*Courtesy Electrical Research Products, Inc.*

Figure 36 — Curve showing the relation between positive projected transmission and the ribbon spacing of the recording valve.

Figure 36 illustrates the linear relationship between negative exposure and positive projected transmission. A constant amplitude of modulation is shown which is variable about three different mean values—it is evident that the output level in these three cases is the same.

To avoid amplitude distortion at any frequency, the output level of the recording system must be proportional to the exposure modulation amplitude. It is commonly known that the output level may be varied by changing the print transmission and that the output level is proportional to the percentage of modulation. When the print transmission is varied as a consequence of bias on the valve, the percentage modulation varies in an inverse manner, but since the output level is

proportional to the product of these two factors and as they change in an equal but inverse manner, the output level is unaffected.

When the illumination in the printing process is varied, the positive mean transmission and the positive change of transmission due to modulation are both varied, and in the same proportion. Therefore, the signal and noise levels vary together. This effect is used as a means of signal adjustment, but not as noise reduction.

However, for the above conditions to be true the exposure illumination and development conditions of both negative and positive films must be held constant. Variation of mean transmission as a result of varying illumination and development will definitely affect the reproduced level as well as the mean transmission, and as a consequence it cannot be used as a method of noise reduction.

6. "SPLIT-CHANNEL" METHOD OF NOISE REDUCTION

Another method of noise reduction illustrated in Figure 37 makes use of the frequency distribution of the film noise in the recording band. Referring again to Curve B of Figure 33-B, it is apparent that the higher frequencies represent greater effective noise level. The noise level may be reduced by the use, in the reproducer circuit, of an equalizer which causes an attenuation of these higher frequencies, but this leads to a corresponding loss of the high frequencies of the recorded sound as well as a reduction of the noise level.

There are two methods in use employing this principle of noise reduction.

In the first method, where only a single track is used, the high frequencies of the recorded sound are pre-equalized by an amount equal to the loss introduced by the post-equalizer, resulting, for all practical purposes, in the reproduction of a signal without frequency discrimination, but with a reduced noise level. (See Chapter III, Complementary Recording.)

The second system requires the use of two tracks, one for the low frequencies and the other for the high frequencies. The frequency band of the recorded sound is separated into two branches of the circuit by appropriate filters (see Figure 37), each branch operating a valve which utilizes only half the sound track. Thus the low frequencies are recorded on one-half the track and the highs on the other half. When this track is projected, a low-pass filter is inserted in the circuit from the low-frequency half of the track, effectively eliminating all noise above the cut-off frequency of the filter, and the high frequencies are reproduced

from the other half of the track in the conventional manner. This output is recombined with the output from the low-frequency side and

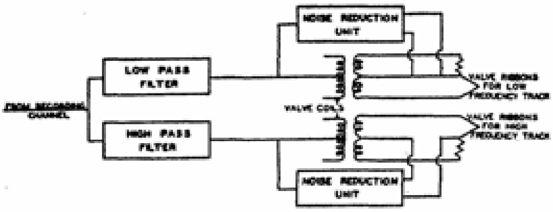


Figure 37 — Block schematic of two-channel high speed bias circuit.

fed into the same projection amplifier system. In recording this track, noise reduction of the carrier-reduction type is used only in the high-frequency circuit.

Due to the fact that the low frequencies have been eliminated from this branch, the time of operation of the noise reduction in this circuit may be accelerated, thus reducing distortion due to valve clash on the initial wave front, which is caused by too slow a removal of the valve bias when a signal of high modulation is introduced. The high frequencies in this branch of the circuit have a masking effect on the noise, which further reduces the noise level.

7. "AUTOMATIC ATTENUATION" METHOD OF NOISE REDUCTION

Another system of noise reduction which has been used with a certain degree of success in some instances, consists essentially of an automatic attenuator, sometimes known as a threshold limiting device, placed in the projection circuit to attenuate the transmission during periods in which no modulation exists. A control circuit, operated by the same energy which produces the output, actuates the attenuator and restores transmission of speech energy to a normal value during modulation periods, at which time the presence of modulation is depended upon to mask out the ground noise. This method has the added advantage of reducing noise from other sources than the film, such as extraneous noises picked up by the microphone, etc.

8. NOISE REDUCTION CIRCUITS

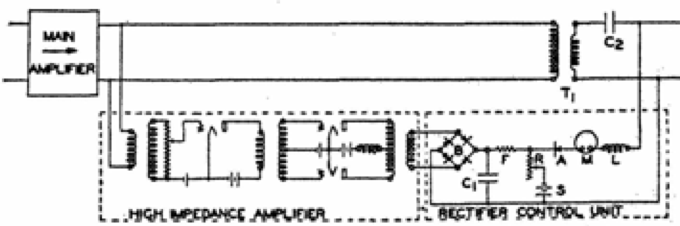
Figure 34 shows the noise reduction control apparatus for the common method of reducing noise by the reduction of carrier amplitude, accomplished either by the use of shutters or by biasing the modu-

lator. The ideal case of reduction of noise by the bias method, in which the bias follows exactly the envelope of the wave, is shown in Figure 30-(b). Any bias method would operate in this fashion if its response were in direct proportion to the input for both amplitude and phase. For this to be so, all amplifiers, rectifiers, etc., would have to have exactly linear amplitude characteristics, and all filters and other delay circuits would have to be removed from the system. This would lead to effectively zero instantaneous margin which would give the ultimate reduction in noise by the carrier reduction method.

However, in practical applications, a margin is necessary as the conditions outlined above are not fulfilled and the noise reduction control apparatus must be given sufficient time to operate and to remove the valve bias or shutter, as the case may be, when a high level of modulation requires a greater mean transmission or track width.

The noise reduction control apparatus consists of an amplifier circuit bridged on a special circuit which feeds the valve, the amplifier being coupled to a rectifier unit (see Chapter XXIII) which contains a filter. The output uni-directional potential from the filter is proportional to the amplitude (the value of which will be discussed later) of the speech energy and is used to restore the track width by removing the shutter or cancelling the bias.

Figure 38 is a schematic of a typical noise reduction system. The characteristics of this equipment are such that the direct-current output



—*From the Journal of the S.M.P.E., May, 1932. "Western Electric Noiseless Recording," H. C. Silent and J. G. Frayne.*

Figure 38 — Schematic diagram of circuits of noise reduction amplifier and control unit.

is proportional to the average value of the wave of speech input energy, and when this output is used to cancel the bias current, the net bias current is inversely proportional to the speech level.

The sensitivity of the circuit is regulated by adjusting the gain of the noise reduction amplifier and is so adjusted that when the modulator amplitude is maximum the bias current is completely cancelled. This adjustment is the margin adjustment. The rectifier prevents a reversal

of the bias current through the modulator, as it stops any further change of current when the bias current has been completely cancelled, regardless of any increase in signal level.

The modulator bias current is supplied by a battery in series with a variable resistance, this resistance being high with respect to the modulator resistance, and causing only a small bridging loss. This bias current is adjusted to the desired amount by means of the variable resistance, as indicated by the meter, this adjustment determining the amount of noise reduction applied.

Figure 39 illustrates the relation between current density and the resistance of a copper oxide rectifier. When the rectified signal voltage

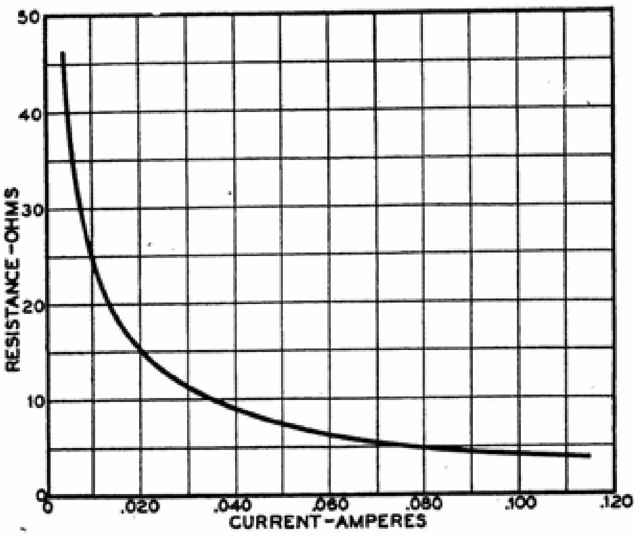


Figure 39 — Curve showing the relation between the resistance variation with current of a copper oxide rectifier.

is high, the resistance of this anti-reversing rectifier rises to a relatively high value and thus blocks any reverse valve bias current which would otherwise tend to flow. This signal rectifier has a slightly non-linear characteristic which is evident near the cut-off point.

The output of the rectifier is a uni-directional current which contains the double frequency of the original speech energy. If this component is not filtered out, it will modulate the valve, which in turn will record the same modulation on the film. Precautions must therefore be taken to prevent this modulation from entering the audible frequency range. Thus a limit to the speed of operation of the noise reduction

apparatus is established. This limit is established by that frequency which is the highest allowable below the cut-off of the system as established by the horns, filters, etc., in the circuit.

As an example, if the system will reproduce, as the lowest frequency, 20 c.p.s., and if 100% modulation is necessary for audibility at this frequency, then the maximum allowable rate of change of bias current will be approximately 1/80 or 0.0125 seconds. The necessary filtering circuit which follows the rectifier introduces a time delay in the circuit and should thus be designed for a cut-off frequency of about 20 c.p.s. as a maximum.

Due to the type of circuit and the fact that the full-wave rectifier impedance is constant, the periods for bias cancellation currents to build up and to decay are equal.

Another type of noise reduction circuit utilizes a vacuum tube rectifier in place of the copper oxide type. The filter which follows the rectifier consists of a simple condenser resistance network. The time constant of this circuit is fixed at approximately the same value as the previously described circuit.

9. TIMING CIRCUITS OF NOISE REDUCTION EQUIPMENT

The time constants of noise reduction circuits are determined in the following manner.

Consider Figure 40, which is a simplified circuit of the timing circuit of the noise reduction systems. In Figure 40

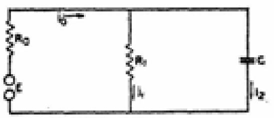


Figure 40 — Simplified circuit diagram of timing circuit of a noise reduction system.

$E =$ constant impressed voltage
$R_0 =$ resistance of generating source
$C =$ capacitance of condenser in farads
$i_0 =$ instantaneous current in R_0
$i_1 =$ instantaneous current in R_1
$i_2 =$ instantaneous current in C
$t =$ time

From the e.m.f. and current laws the following equations are obtained:

$$i_2 = \frac{dq}{dt} \tag{10}$$

$$i_1 = \frac{E - R_0 i_0}{R_1} \tag{11}$$

$$i_0 = i_1 + i_2 \tag{12}$$

From these equations may be derived an expression for i_1 in terms of t and the constants of the circuit:

$$i_1 = \frac{E}{R_1 + R_0}\left(1 - \epsilon^{-\frac{t}{RC}}\right) \tag{13}$$

Where $$R = \frac{R_0 R_1}{R_1 + R_0}$$

When $t = \infty$ then $i_1 = \dfrac{E}{R_1 + R_0}$

which is the maximum value of i_1. This is, of course, the theoretical maximum value, but practically, as the value of $\left(\frac{t}{RC}\right)$ becomes large at a very short time—as we normally measure time—after the switching operation, the value of t usually given for noise reduction systems is that value when $i_1 = 0.9$ (maximum value). (14)

This value of t is the operating time of the circuit.

When $$i_1 = 0.9 \frac{E}{R_0 + R_1} \tag{15}$$

then $$t = 2.3\, RC \tag{16}$$

With this type equipment the bias current is the last stage amplifier plate current, which is reduced by the rectified signal applied to this stage as a grid bias voltage. To avoid "over-shooting" due to excessive signal, the circuit is adjusted so that this plate current approaches the cut-off point, beyond which the amplifier stage becomes non-linear. Use is made of this fact as it enables the equipment to provide a large margin at low levels and a small margin at high levels, which decreases to zero after the cut-off point has been reached and the modulation is 100%.

The response curve, that is, the relation between bias current and relative signal input, is linear over almost the total range used. However, at the extremes of the curve, that is, when the bias current is reduced to zero or when the signal reduces to zero, the relation falls off slightly.

Another noise reduction control system of the carrier-frequency type has its bias current supplied by the rectified output of a 20,000 cycle oscillator whose level is modulated by the rectified speech energy. The 20,000 cycle circuit is the "carrier" for this particular equipment, consequently the name carrier-frequency. The output level is maximum when the signal level is zero. The response characteristics are to a

large degree controlled by the adjustment of the exciter voltage, and the bias and modulation level into the modulating tube.

Noise reduction control equipment used with a light valve requires an equalizer circuit in order that the response with frequency be made to follow the light valve resonant characteristic, which condition provides a constant margin at all frequency conditions at a given level.

The usual noise reduction systems, that is, those with a linear response between the speech input into the noise reduction system and

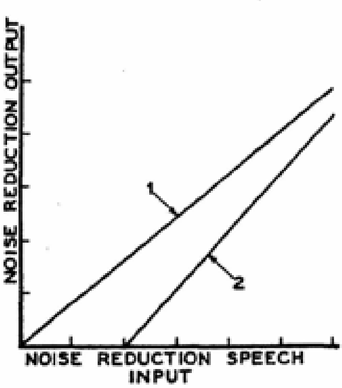


Figure 41-A — Curves showing the response characteristic of "linear response" (Curve 1), and "constant margin" (Curve 2) noise reduction systems.

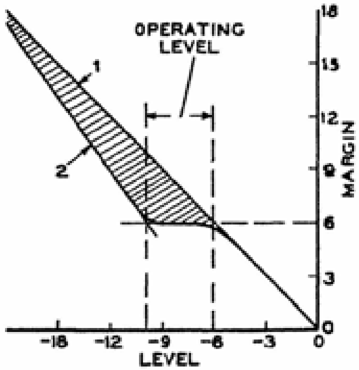


Figure 41-B — Curves showing the relation between margin and modulation level of "linear response" and "constant margin" types of noise reduction systems.

the output of the noise reduction system as described above, and as illustrated by Curve 1 of Figures 41-A and 41-B, provide too much margin in the operating region of the noise reduction system.

Another system, called the "constant margin" system, has been designed. This has a response characteristic as shown by Curve 2 of Figure 41-A. Figure 41-B illustrates the margin such a system provides at different levels, where zero level is a condition of 100% modulation. Curve 1 shows the relation of a linear noise reduction system. Curve 2 shows the margin provided at different levels by a constant margin system designed to operate at a margin of 6 db. As the modulation level is increased from 50 toward 100%, the margin approaches zero, which is the valve overload level. However, as the speech level decreases from 6 db, to say 10 db, the margin remains constant at a value of 6 db. The 10 db point (chosen arbitrarily for this drawing) is determined by the no-signal spacing of the valve ribbons and for any reduction in level below this point the margin increases in a linear manner as the modulation decreases.

Such a system further decreases the noise level for certain values of modulation. This further decrease in noise level is represented by the shaded area in Figure 41-B and represents an improvement over the linear-response noise reduction systems.

The noise reduction systems described above and those given in the chapter on recording methods, still leave a great deal to be desired from a consideration of the results obtained. In all probability a new method or a combination of old methods will eventually be worked out and used—a method which at the present time is not even anticipated.

Chapter V

RE-RECORDING AND PREPARATION FOR RELEASE

By KENNETH LAMBERT

The preparation of a picture for release involves processes which are not all of a technical nature. The philosophies of these processes, however, are tremendously important not only to the mechanical tools and operations used, but also to the very planning, writing and directing of the picture. From a strictly sound standpoint the interpretation given these philosophies may decide the choice of a recording system, the technique for using it, and the training and organization of the operating personnel.

Releasing a motion picture is about like preparing for a dinner party. You wish it to be one your friends will remember pleasantly. Many things will contribute to its success or failure. You hope your guests will come with keen expectation and will not be disappointed. Your home is always delightful, you are a good host, but you know that these particular people especially appreciate tasty, well prepared foods. Your favorite roast, its gravy flavored just to a king's taste; your particularly incomparable biscuits; and a relaxing evening to enjoy it all. Everything else is sure to be right but what if the cook fails you just this once! What if she is distracted by the children and puts too much salt in the gravy, or lets the biscuits burn? Well, at least she usually succeeds in making even the toughest meat moderately tender. You have cautioned the one who serves to be sure everything comes to the table at just the right temperature. At the last moment you go to the kitchen for one last smell, and then hope it will all arrive safely.

The producer has planned a perfect menu. His director has secured the finest foods. He has even grown most of them himself. The editor and recorder cook the meal, it is put into dishes by the laboratory, and is given to a theatre manager to serve.

It is difficult to consider the functions of editing and re-recording separately. Editing or "cutting" has been commonly considered to mean the assembly of the picture scenes and their corresponding sound in such fashion that the dramatic situations are presented forcefully in accordance with the plans of the producer. Re-recording or "dubbing"

was once the unavoidable process by which occasional necessary music or other effects were inserted into a sequence. This is no longer the case. Both the editor and the re-recorder must constantly have in mind the one common aim—how will the picture tell the story in the theatre? The editor must be more than picture-conscious, just as the re-recorder must be more than sound-conscious. Each must be a master of technique of various kinds, for motion pictures are inherently a technical product. Primarily, though, these men must be master judges of how best to put the producer's story across in a picture theatre. Their responsibility in this may exceed even the producer's, for they may be more familiar with the changing technical aspects of such presentation. They are showmen, and as such each is interested in the whole show, and not merely parts of it.

What makes up this show in a theatre? From a story standpoint, the sequences must be planned, shot and assembled into film with dramatic effectiveness, and this film must then be projected properly in the theatre. The picture must be focused sharply, with uniform, adequate illumination, on a carefully matted screen, of the most effective size for the particular theatre. The sound must be of the proper loudness, and free from distortions which might make it unintelligible, harsh, or disagreeable. The matter of effective presentation in the theatre now rests with the theatre owner and his manager and projectionist. Each must take a real interest in the mechanical performance of the theatre and all its facilities. The technique of good theatre presentation is discussed elsewhere in this book.

The process of releasing a picture divides naturally between several groups of people. The function of the laboratory is to intercut the negatives of the various picture and sound scenes in accordance with the "working" or "master" print prepared by the editor and re-recorder under the producer's supervision, and then to make the necessary release prints for use in the theatres.

The division of work between the editor and re-recorder in preparing the master print is dictated by convenience and economy. Their work is not entirely simultaneous and involves different techniques which are quite specialized but still of a nature permitting effective teamwork. Because of this, one editor can be "cutting" one picture while a re-recorder is working on the sound assembly of another picture. In this way, the equipment and personnel necessary for re-recording can be utilized most efficiently.

We may conveniently define the editing part of this process to include all of the selection and cutting together of film either for dramatic or for technical reasons. We shall only touch upon this technique.

The re-recording part then includes the assembly of the sound onto film suitable for theatre use. Of course, this is only one of the purposes for which re-recording may be employed. Re-recording is to sound what process photography is to the picture, and most of the tricks available in process photography have their equivalent in sound re-recording.

Re-recording has four general purposes: To combine sound effects, to adjust loudness or quality, to secure additional or duplicate records, and to change the kind of record, as from disc to film or from push-pull track to single track.

These basic uses and their corollaries determine the nature of a re-recording channel. A stage recording channel uses one or more transmitters on the stage to initiate the electrical energy necessary for recording either on film or on disc. A re-recording channel merely replaces these transmitters with film or disc reproducers upon which records are played. The natures of these records may be very different. The dialogue may be from a film, music from film or several discs, and sound effects either from specially synchronized films or from loops of continuously repeating sound which can be appropriately mixed-in during a scene. Fundamentally, that is the only way in which the mechanics of re-recording differ from original recording. The methods of employing stage and re-recording channels are quite different, however, which becomes strikingly apparent when the entire subject of recording and reproducing motion picture sound is examined. Let us do that briefly.

1. THE GENERAL PROBLEM

We desire sound in the theatre which will seem so natural when heard with the picture, that the observers feel they are a part of the scene, or at least that they are viewing the production through a large opening in the end of the theatre. The specific technical requirements necessary to achieve this result are discussed elsewhere in this book.

2. EARLY SOLUTIONS

In the earliest sound pictures, a frequency band from about 100 to 5,500 c.p.s. was reproduced, the width of this frequency band being limited primarily by the theatre horns. Because of the limited frequency range and the use of simple horns, phase distortion was unimportant. The volume ranges of recording media were about 20 db from surface noise to maximum modulation, but, by accepting objectionable surface noise, 30 to 35 db of recorded volume could be secured in practice. There was little variety between scenes, whole reels commonly being made with one microphone placement. Music and sound effects

were recorded at the time the action was photographed and re-recording was usually not required or employed.

As time passed, scenes became shorter, several being re-recorded from intercut film or from a number of discs, to reel-length discs for release. Some pictures were released with the sound printed on the film from intercut original sound negative. Adding effects by re-recording was still uncommon. Level changes in the theatre, if made at all, were dependent upon "Cue Sheets," furnished to the theatre by the studio and according to which the projectionist made changes in the fader or amplifier gain.

Gradually the method of adjusting volume of variable density recording by changes in track density came into use. This is a valuable attribute of this system of recording which is still used. Planned use of this technique permits soft sequences, even in re-recorded track, to be printed darker than usual with the same resulting sound volume but with lower surface noise than normal density track having low modulation. In difficult re-recording, undesired level variations may occur in an otherwise acceptable re-recorded track. In the variable density system these can be corrected in the prints, but with variable area there is, at present, no easy cure. This feature makes variable density especially useful as a release medium.

3. INFLUENCE OF NOISE REDUCTION

Noise reduction systems became available for both variable area and variable density recording. These extended the useful volume range of variable density 6 to 10 db and of variable area about 12 db. At the same time the usefulness of re-recording was becoming more apparent. Re-recording permitted sequences to be shot in short scenes, or in different sets, or even on different days. Continuous background noises could be added later, and level and quality differences between original scenes corrected. Use of noise reduction on both original and re-recorded tracks reduced the surface noise to a somewhat commercial amount in spite of the increase due to re-recording. This increase is 3 db if the noise in the original record and in the re-recording medium are equal. Because of the random frequency spectrum of noise it combines in a root-mean-square manner. That is,

$$N_{recording} = [N^2_{original} + N^2_{rerecording}]^{1/2}$$

It is apparent that if either the original or the re-recording medium had even moderately increased noise reduction, the resulting total noise would be but little more than that of the noisier medium alone.

4. MODERN SOLUTIONS

The comparative freedom in production technique and the control of the release product, resulting from a policy of complete re-recording, caused several studios to adopt such a policy soon after noise reduction became available. The increasing use of process photography caused others to re-record more and more until now most productions are completely re-recorded. Conventional film recording methods were designed for printing the sound on the film beside the picture. They are still commonly used in this manner for theatre release. They had certain defects, however, which could be reduced by operation of two such tracks in "push-pull" (see Chapter III). If this push-pull track could be used for the original recording, the noise and distortion in the re-recorded film would be little more than that to which we had previously been accustomed in original records, and we could gain the great conveniences of re-recording at little quality cost. Not only has this been accomplished, but push-pull track has been adapted to the re-recording stage as well, and many modern theatres are equipped to reproduce such track. (See Chapter X.)

When this technique is augmented by the use of the squeeze mat, volume ranges of 55 to 60 db from maximum modulation in loud parts to minimum modulation (surface noise line) in soft parts are easily obtained in conventional theatres *without changing the theatre fader setting.*

If, in addition, the theatre can be equipped with the very inexpensive equalizer for reproducing track made with the complementary recording technique, 6 to 8 db more range is possible.

It seems improbable that more than 80 db range, according to the above definition, will ever be required.

5. MODERN DISTORTION REQUIREMENTS

It is most important to remember that sound, as currently recorded for picture use, passes through two acoustical, six mechanical, three electrical, six optical, and four chemical states; and twenty-four transformations between and within these states, at least twelve of which take place in the presence of superimposed mechanical movement.

Total quality distortions of 2 to 3 % are detectable and little more is acceptable. Therefore, the total combined effect of all the partially nonlinear elements such as vacuum tubes, iron-cored coils, film characteristics, and the transformations outlined above, must be within this amount of distortion. The greatest technical difficulty in re-recording is to keep errors and distortion so small that the combination of two of

them is not objectionable, when one at a time they may not even be noticeable.

6. RE-RECORDING IS A CREATIVE PROCESS

When you hear a famous violinist in Carnegie Hall you think of him only as a great creative artist—certainly not as a mechanical technician. Yet, if you were to spend hour after hour with him during practice, it is probable that you would become quite conscious of the meticulous placing of his fingers on the strings, his bowing, and even the kind of strings used on the violin. If, during all this time he *appeared* by necessity to be engrossed with technique, it is quite possible that when you went to the concert you would still think of him as a fine technician and of his violin as a mere mechanical tool. His concert would actually be just as beautiful a creation, but your point of view would have spoiled your appreciation of it.

It is fortunate that the audience seeing a finished picture has not seen it being rehearsed over and over again in the re-recording rooms. If the re-recorder is successful, the audience is not conscious of his technique but only of the result achieved. The director and producer watching the re-recorder work out the details, however, may think him and his tools very mechanical, for in spite of what is going on inside his head (the important part of re-recording), his hands are performing a multitude of mechanical operations, and his conversations with his assistants are in terms of machines.

7. RE-RECORDING TECHNIQUE

A re-recording monitoring room is a miniature theatre or projection room, with a picture screen, horn system behind it, and a large table on which are placed all the controls the re-recorder must manipulate.

The acoustic qualities of this room are important. It must sound as nearly as possible like a representatively large theatre. In fact, it is desirable for *all* projection rooms and theatres to sound alike. Then the product made or judged in one room will be equally good in the others as well as in the releasing theatres. These projection rooms need not necessarily be the same size, if each is properly made for its size.

The basic re-recording channel may be considered to be merely a projection channel exactly as in a theatre. A track played in the re-recording room should then theoretically sound the same as in the theatre. At a convenient place in this channel, as close to the horns as practicable, a recording machine is connected. If this recording operation does not distort, then the resulting re-recorded film (from which the process gets its name), should reproduce exactly what the re-recorder heard during the re-recording operation. It is most important that this

distinction be clear; that the re-recording operation records *what the re-recorder hears* and does not *necessarily* duplicate any original record. If the re-recorder is completely satisfied with the original, and does not modify it in any way, then the re-recording will duplicate it. If different recording media are used for the re-recording operation, and each has different degrees of inherent frequency distortion, individual recorder equalizers (see Chapter XVI) must be used which will permit each record to sound as nearly as possible like the monitoring. Inasmuch as the purpose is to record exactly what goes into the reproducing horns, it is desirable that there be as little circuit between the recording machine and the horns as is practicable, and that the amplifiers which supply the energy to the horns and to the recorder be stable in order that the relative levels and quality of the two may be constant. Photoelectric cell monitoring is unnecessary, for modern modulators are quite dependable and probably have less variation than a complicated monitoring device and circuit.

There are other methods of connecting and operating re-recording channels but none seems quite as simple as the one described above.

The re-recording mixer works primarily by ear, merely using his volume indicator to show the percentage of modulation of the film under different adjustments of his channel. The volume indicator is like a speedometer—you only look at it when you are about to exceed the speed limit. In ordinary driving you keep your eyes and your mind on the road, and there is no need for the speedometer.

The re-recording mixer has a number of volume controls, one for each record he is reproducing. He can insert equalizers into one or more of these circuits if the frequency characteristics are not as desired. With certain modern arrangements of volume and equalizer controls, one operator can usually mix five or six tracks and make necessary quality adjustments.

8. RE-RECORDING FOR DUPLICATE RECORDS

The operation of duplicating an original record, or transferring it to another recording medium with its sound quality and levels unchanged, has been described above. If more than one duplicate is needed, or if more than one type of new record is required, additional recording machines are connected to the circuit through their proper equalizers. If duplication into the same recording medium is desired it can often be done by photographic "duping," thus preserving any numbers and identification marks which may be on the original film, and assuring the same percentage modulation as the original (not always easily accomplished in re-recording).

The method of determining the desired recorder equalizer characteristic is simple. A recording is made of single frequency oscillator tones covering in steps the range from 50 to 8,000 cycles. Equal energy is supplied to the modulator amplifier at each frequency. This record then is re-recorded in a channel as described above, except that there is as yet no recorder equalizer. The original frequency record and the re-recorded record are then compared by measuring the voltage developed at each frequency when the records are reproduced into a voltmeter instead of into the usual horns. An equalizer is then designed which will cause the re-recording to match the original, which is inserted in the circuit and the test repeated as a check. The re-recording should now match the original. This equalizer should be adjusted precisely as departure of as little as 1 db in this range can be discerned by ear in music and dialogue.

The re-recording operation must also be free of phase and volume distortion. Phase distortion is limited by using as simple a channel as possible, with coils having high efficiency at frequencies even below the desired 50 cycle minimum, and equalizers which do not introduce unnecessarily large phase shifts at low frequencies.

Volume and frequency distortion are checked in a similar manner. A recording is made of a single oscillator tone, reducing its level in steps from full modulation to perhaps 30 db down. This record is re-recorded and compared with the original as described above. If either record fails to reproduce the original variations of oscillator level, volume distortion exists and must be corrected. The degree of distortion in the re-recording channel is indicated by the difference between the two records.

9. RE-RECORDING TO COMBINE SOUND EFFECTS

Re-recording is the sound equivalent of process photography, and ranges from adding a bit of crowd noise or a door knock in a single scene, to construction of a complete sound accompaniment to a picture sequence originally photographed silent.

The mechanics consist simply of placing desired sounds—dialogue, music, or other effects—at the proper place in two or more tracks, reproducing the tracks simultaneously, and controlling the loudness of each so that the desired effect is heard from the horns. The composite result is recorded as described above.

10. PREPARATION OF TRACKS

This editing operation is so closely related to re-recording that it should be considered briefly.

The sound recorded at the time the picture was photographed is cut simultaneously with the picture. Sound and picture should be

synchronized within two sprocket holes, and be kept in synchronism while editing by running them over sprocket wheels mounted on a common shaft. Other tracks, if required, are run over additional sprockets on the same shaft. If the sound recorded with the scene is not entirely adequate, suitable sound effects to make it so are selected from a film library and are cut into additional tracks at the appropriate places. It is desirable that a minimum number of tracks be used, but, on the other hand, it is inconvenient to jump from dialogue to sound effects to music in one track. It is customary to keep dialogue on one track, music on another, and sound effects on as many more as necessary.

As the tracks for each picture reel are assembled, their contents are outlined on a prepared form which serves to index the structure of the reel and to transmit or record editorial and technical instructions concerning its assembly. In one studio the reverse side of this form also serves as a record of all re-recording done on the reel. The sheets representing each reel during its various previews and release are always kept together, so that the entire history of the reel is available in one place.

The selection of effects from the library and their construction into the tracks requires much imagination, ingenuity, and care. The effects must be realistic and of the proper perspective, for if false they can destroy the whole illusion of a sequence. If illusion of reality is destroyed, the value of the motion picture is destroyed.

11. COMBINING THE TRACKS

The tracks can be mixed either for a realistic effect, or for a theatrical one. Realism puts quite strict limitations on the treatment. Theatrical effects usually utilize the exaggeration of certain sounds beyond realism, or the omission of realistic sounds. The combinations here are limitless.

Studio or producer policies may determine the general treatment. Some favor realism; others lean toward fantasy. In this connection, the manner of using music is important. If music is required by the action, it is a realistic effect—if merely for creating a mood, as an underscore during a dialogue sequence, it tends to destroy realism. The nature and treatment of the picture rather automatically determine how music should be used. If the story is being told by dialogue, sound effects must be subjugated. In other places,perhaps, the effects tell the story and may dominate the dialogue. Variety in treatment is refreshing if reasonable.

12. RE-RECORDING TO ADJUST LEVELS AND QUALITY

When sound is recorded in different sets, with different channels, under different monitoring conditions, many variations of level and

quality unavoidably result. In addition, actors may not speak quite the same in consecutive picture scenes, or they may even have a cold and distorted voice quality. One set may be heavily draped, and sound dull; another with mirrors all around will reflect sound about as effectively as light. Reverberant sound will seem loud though its energy content may be small.

The finally released picture must seem so smooth that it will appear to have been photographed and recorded in one long continuous operation. Re-recording is especially useful for this purpose. Level adjustment is inherent to it. The proper level for sound is determined very definitely by the action. The sound must appear to come from the picture on the screen. Because of the enlarged picture the proper loudness is somewhat greater than in life, usually 6 to 12 db. If too soft, it will seem remote, behind the screen, and if too loud it will be grotesque. The director's and actor's interpretation of the scene cannot be changed appreciably in re-recording by making the sound loud or soft. This would merely make the scene artificially unreal.

At present the setting of absolute loudness in the theatre must be determined by the theatre manager or projectionist. The re-recording mixer tries to help determine the proper loudness by the *range* of loudness he builds into the picture. If both whispers and shouts sound natural and convincing in the theatre, the sound is probably being well projected. ***Pictures are so constructed that if the dialogue is played at the proper level all effect and musical sequences will automatically be of the desired loudness without changing the theatre amplification.***

The standard of quality is always, "Does it sound like the real thing?" Most of the public have never heard a picture actor's actual voice, but everyone knows whether his picture voice sounds human and natural.

Most of us hear with two ears. Present recording systems hear as if with only one, and adjustments of quality must be made electrically or acoustically to simulate the effect of two ears. Cover one ear and note how voices become more masked by surrounding noises. This effect is overcome in recording, partially by the use of a somewhat directional microphone which discriminates against sound coming from the back and sides of the microphone, and partially by placing the microphone a little closer to the actor than we normally should have our ears.

This closeness tends to accentuate the low and high frequencies in the record, and make it both boomy and spitty.

Losses of high frequencies inherent to recording processes tend to help correct the spittiness, but the boominess remains, which in turn is accentuated by the loudness required for good illusion in the theatre. This boominess can be reduced in a number of ways, all of them amounting to "equalization."

Equalization can have many forms, as it is merely the use of reactive networks of one kind or another to give a system a desired relative transmission ability at different frequencies. The name originated from use in telephone circuits to counteract distortions which were otherwise unavoidable. This is still its primary use in recording. Since it is a counteractive distortion when so used, it is easy to see that it must be used intelligently or undesired new distortion will result. It can naturally be used to produce a controlled degree of distortion if desired.

Equalizing networks ordinarily are made with electrical reactances, but there are acoustical and mechanical reactances which can be used if convenient. Cavities in a microphone or loud-speaker and the mechanical resonance of light valve ribbons are examples. If a system is free from volume distortion, equalization can be introduced at any convenient point in the system with equal effectiveness. Amplifier systems, however, are not entirely free from volume distortion and this must be considered carefully in the design and use of equalizers. In addition, operating convenience often dictates the point in a system at which equalizers can best be used. A small fixed amount of low-frequency attenuation is sometimes used in the stage recording channel to reduce excessive boominess. In general, it is better practice to do corrective equalizing only while re-recording. Variations in quality can only be corrected after the picture is edited. In addition, the small group of re-recorders, working under standard, consistent monitoring conditions, closely in contact with release requirements, is better able to judge what treatment is required than a large group of stage mixers working under a variety of monitor conditions not particularly representative of theatres. Rather complex and bulky equalizer equipment is sometimes required, which can only be provided conveniently in the re-recording rooms. Variable equalizers may provide means to accentuate or reduce narrow or wide bands of frequencies throughout the entire frequency range. Fixed networks may be employed to counteract or produce specific, frequently encountered, distortions. Unusual problems may require the design and construction of special networks. In general, corrective networks should have the same nominal impedance as the circuit in which they will be inserted, and, as far as possible, should be of symmetrical constant resistance structure so that they can be used in tandem in any combination which may be required.

In the choice and use of corrective equalizers, a re-recorder can only be guided by ear. The nature and amount of correction must be determined by trial.

It is important to understand that phase and volume distortions cannot be corrected by attenuation equalizers. Reverberation is a phase effect—an echo, really, which cannot be removed from a record but which can sometimes be added. Volume distortion could only be corrected by introducing an inverse distortion in most cases, an almost impossible feat.

Another type of quality change sometimes utilized is to vary the pitch of a sound, which is accomplished by operating either the reproducer or the recorder at other than normal speed. This naturally affects the duration of the sound in proportion to the speed and pitch change. However, the harmonic relationship of overtones are not upset, so moderate changes do not cause unintelligibility in dialogue or harshness in music. Such a change does modify the quality of the sound in addition to changing its pitch. A masculine voice raised an octave by this means would sound feminine.

13. SUMMARY

We are painting a picture. All of the mechanics we employ, whether acting, directing, photography or recording, are the brushes with which the thought colors are spread on the theatre screen. If one can get a better effect by smearing the paint on with his thumb than by using a more conventional tool, there is reason for doing it. Progress results from new tools developed to meet new needs.

Chapter VI

MICROPHONES

By L. E. CLARK

1. TYPES

A microphone is an electro-acoustic transducer, actuated by either the pressure produced by, or the velocity of, sound waves, which converts acoustic energy into electrical energy.

Theoretically, the current wave produced in the electric circuit should follow exactly the pattern of the pressure wave, and any differences in the two waves are the result of some deficiency in the microphone. The microphones most commonly used in studio work are, in historical order: carbon, condenser, dynamic or moving coil, ribbon or velocity, and crystal.

These can be divided into two groups: pressure operated devices, including the carbon, condenser, dynamic, and crystal; and the velocity operated device, of which the ribbon is an example. Of the first group, the carbon, condenser, and crystal are tuned; the dynamic, untuned.

The condenser, dynamic and ribbon are the most commonly used.

The output of these microphones varies from about minus 40 to minus 80 db, compared to a reference level of one volt open circuit per bar of acoustical pressure (1 bar equals 1 dyne per sq. cm.). Such a small output requires a great amount of amplification, and in some cases the first, or microphone amplifier, is actually contained within the microphone housing itself to reduce losses due to amplifier leads.

The output impedances vary from a few ohms to several thousand ohms. The high impedance instruments must have their amplifiers in the same housing to reduce losses, but the output from the low impedance microphones can be carried over a considerable distance.

2. CARBON MICROPHONE

The carbon microphone is the oldest type, operating upon the principle that the resistance of a mass of carbon granules varies with the

pressure applied. A direct current passes through the carbon, and variations in pressure cause direct-current variations which theoretically follow the pressure changes. A transformer is inserted in the microphone circuit which allows only the alternating portion of the current to be delivered to the amplifier circuit.

Commercial units of this type usually employ two such masses of carbon (the so-called double-button carbon microphone) to reduce amplitude distortion.

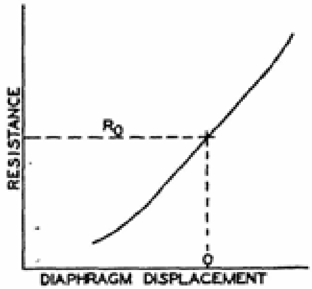


Figure 42 — Diaphragm displacement of a carbon microphone.

Figure 42 shows a graph of the resistance of the circuit plotted against the position of the diaphragm. At zero displacement the resistance has some value R_0 and as the diaphragm is displaced, distortion results as the curve is not linear. The double-button microphone attempts to correct this by superimposing an opposite curve of the same type upon the first in a push-pull circuit, which reduces not only the distortion but also the sensitivity.

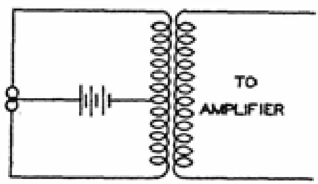


Figure 43-A — Schematic of carbon microphone circuit.

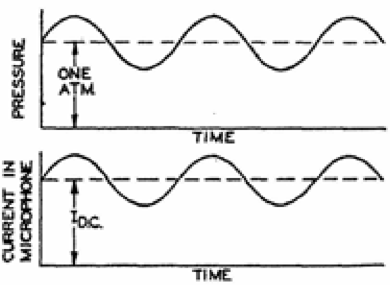


Figure 43-B — Electrical wave in microphone circuit showing the corresponding pressure wave.

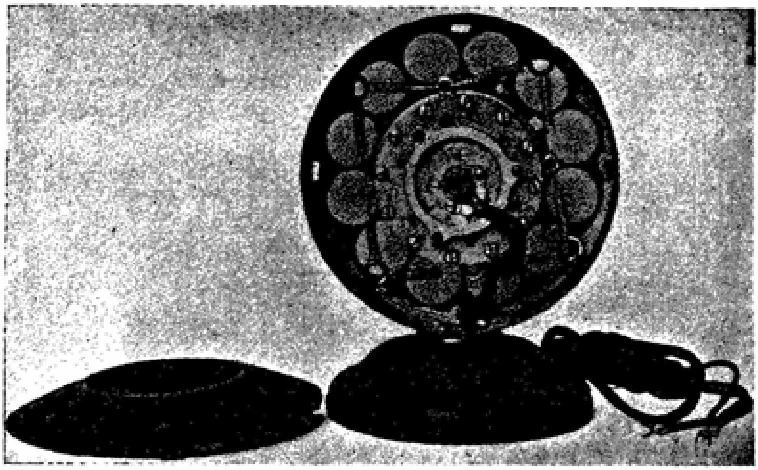
Figure 43-A shows a schematic circuit of such a microphone. Figure 43-B gives a simplified pressure wave and the corresponding electrical wave if no distortion is present. The current I delivered to the amplifier is

$$I = I_M - I_{DO} \quad (17)$$

where I_M is the total current at any instant through the microphone circuit, and I_{DO} is the no-signal or unmodulated direct current.

Figure 44 shows this microphone as commercially manufactured.

The disadvantages of this microphone are: High noise level, general instability of frequency characteristic, and tendency of the carbon granules to pack together.



—Courtesy Electrical Research Products, Inc.

Figure 44 — Carbon microphone.

The hiss or background noise is due to the great number of contacts presented by the large number of granules of carbon.

The sensitivity depends upon the condition of the carbon, and is affected by vibration, position, handling, etc., and so the microphone is very apt to lose its calibration.

This microphone is not used in the studios because of stability and noise-level requirements which it cannot meet.

3. CONDENSER MICROPHONE

The condenser microphone derives its name from the fact that its operation depends upon the variation in the capacitance between two plates. One plate, the diaphragm, is movable and the other is fixed, the two forming a variable condenser in which the distance between the plates is the variable element. A direct-current potential is applied to the two plates through a high resistance, and as the diaphragm vibrates and produces a corresponding change in capacity, voltage variations appear across the series resistor. This voltage can then be amplified, and since the resistor across

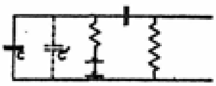
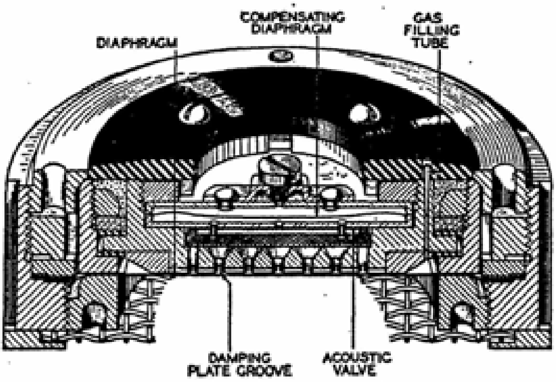


Figure 45 — Schematic diagram of a condenser microphone.

which it is developed is very high, the first amplifier tube is usually coupled through a small capacitor to the resistor.

In microphones of this type now in use the impedance is high, and the amplifier immediately associated with the microphone is in the same housing.

Figure 45 shows a schematic wiring diagram of a condenser microphone, using a resistance coupling, while Figure 46 shows a cross-section of a typical condenser microphone.



—*From "Radio Engineering," by Frederick E. Terman (McGraw-Hill Book Co., Inc.).*

Figure 46 — Cross-section of condenser microphone.

The diaphragm is duraluminum and stretched so that its resonant point is well up in the recording band. The overall diameter is only about three inches, and as a consequence the plates must be relatively close together to have an appreciable capacitance. The gas trapped in the small space between the plates acquires a stiffness and viscous damping and its action results in raising the frequency of resonance. The effect of this resonance is minimized by designing the microphone so that the resonance mentioned above occurs at the extreme upper end of the frequency range.

In order that the stiffness of the trapped air be relatively constant with frequency, the usual design divides the back plate into a number of small sections by means of crosscut grooves or concentric rings.

Equalization is provided to take care of variations in atmospheric pressure. This is shown in Figure 46 as the compensating diaphragm.

(a) Cavity Resonance

The condenser microphone is so constructed that there is a cavity or air pocket in front of the diaphragm (see Figure 46). At the fundamental frequency of vibration of this cavity, the pressure may build up to twice the value it would have if this cavity were not present.

The resonant point is usually in the neighborhood of 3,500 c.p.s., depending on the dimensions of the cavity, and has a magnitude of about 5 db at this frequency for waves approaching near the axis of the microphone.

The magnitude of the cavity resonance has been reduced in some microphones by modifying the housing so that the diaphragm is flush with the face of the microphone.

(b) Pressure-Doubling Effect

If the frequency of a sound is high and its wave length small compared to the size of the diaphragm, part of the sound wave is actually stopped, and the pressure at the diaphragm increased. In the high-frequency range, the sensitivity of the microphone may be increased as much as 6 db. The actual frequency at which this increase takes place depends upon the size of diaphragm and mounting.

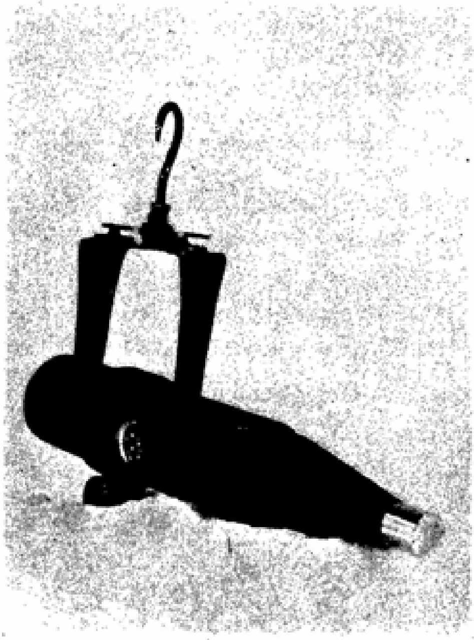


Figure 47 — Miniature condenser microphone (ERPI—D-99848 transmitter with amplier).

The effects of cavity resonance and pressure doubling on the response characteristic of a condenser microphone are shown in Figure 58. These effects are additive and give rise to the peak in the region of 3,000 cycles;

pressure-doubling contributing approximately 6 db, and cavity resonance 5 db.

The condenser microphone has a higher impedance, a lower sensitivity, and a much lower noise-level than the carbon type, and is affected by cavity, diaphragm, and air-pocket resonances (the last two being damped), and, in sizes now commercially available, by pressure-doubling. Its response characteristic is comparatively stable, and it is not affected appreciably by temperature, humidity or ordinary handling.

A smaller improved type of condenser microphone is now available commercially. The diaphragm, built flush with the face of the housing, has a diameter of approximately 0.6 of an inch.

Figure 47 shows the new type miniature condenser microphone (usually termed Baby Condenser) with its associated amplifier. The size of the microphone element itself is about eight-tenths of an inch in diameter by approximately an inch in length.

The reduction in size shifts the cavity resonant peak from the 3,000 cycle neighborhood of larger condenser microphones to the high-frequency range at approximately 10,000 cycles.

The efficiency is approximately the same, compared to the larger microphone, as the proportion of "dead" to "active" capacitance is much lower, and for this reason the voltage on the grid of the first amplifier tube is the same in both cases and no increase in noise results.

Still further reductions in size have been considered in an attempt to improve reflection and phase difference effects but the results were of negligible importance within the audible frequency range.

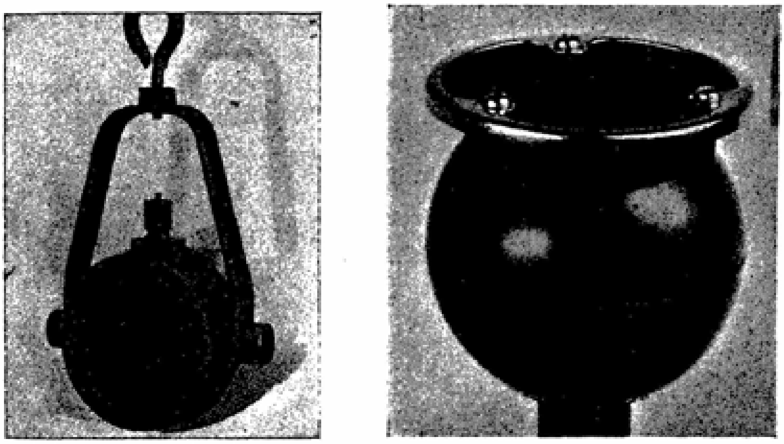
A flatter response curve is obtained from this smaller microphone (see Figure 58) due to the decrease in diameter and the elimination of the cavity.

4. DYNAMIC OR MOVING COIL MICROPHONE

This microphone operates upon the principle that a coil, when moved through a magnetic field, will have an e.m.f. induced in it. It consists mainly of a diaphragm to which a coil is rigidly fixed, the diaphragm vibrating in response to the sound waves striking its surface and causing the coil to vibrate in like manner, cutting magnetic lines of force.

By the very nature of this design, the mass of the moving parts is relatively large, and acoustic networks must be used in order to give a flatter response.

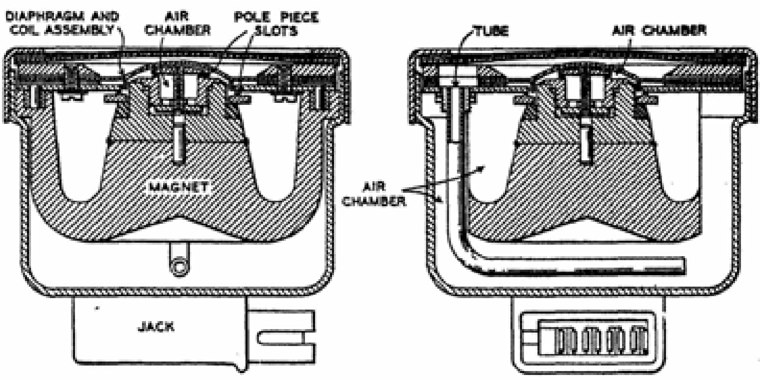
Figure 49 shows a cross-section of a dynamic microphone. It will be noticed that the diaphragm is crowned and the voice coil placed at the edge of the crown, and that there is an acoustic network coupled to the diaphragm.



—*Courtesy Electrical Research Products, Inc.*

Figure 48 — Dynamic microphones.

This microphone has a higher sensitivity than the ribbon, a low-hiss characteristic, low impedance, and resonant peaks which are



—*Courtesy Electrical Research Products, Inc.*

Figure 49 — Cross-sectional view of a dynamic microphone.

acoustically damped. The response characteristic (see Figure 58) is good, but not always uniform from one microphone to another of the same type.

5. RIBBON MICROPHONE

The ribbon microphone consists of a very light, corrugated metallic ribbon suspended under negligible tension in a magnetic field and freely accessible to air vibrations from both sides. The ribbon vibrates in the magnetic field under the action of the difference in pressure existing between the two sides and this vibration induces an e.m.f. in the ribbon. Although the generated e.m.f. of this type microphone is low, its impedance is low and, therefore, the efficiency is of the same order of magnitude as the condenser and dynamic microphones.

In order that the ribbon or diaphragm move, it must be actuated by the pressure of the sound wave. This is true for all types of microphones, including the ribbon, or velocity microphone, which is operated by the pressure difference between the front and back. Strictly speaking, it should not be called a velocity microphone, but a pressure-gradient microphone.

Although the sensitivity of this type microphone is low, its amplifier need not necessarily be in the same housing, as its impedance is also very low.

Its response characteristic is the best of all type microphones with the possible exception of the crystal, but there is a characteristic middle-frequency hum which is troublesome when the microphone is used for dialogue recording, but when used for music (its most frequent use in the studios) this hiss is not apparent.

—*Courtesy RCA Manufacturing Co.*

Figure 50 — Ribbon microphone.

It is, however, easily overloaded, and sudden sounds such as gun shots, explosions, etc., may blow the ribbon entirely out of the air gap.

Figure 50 shows a ribbon microphone.

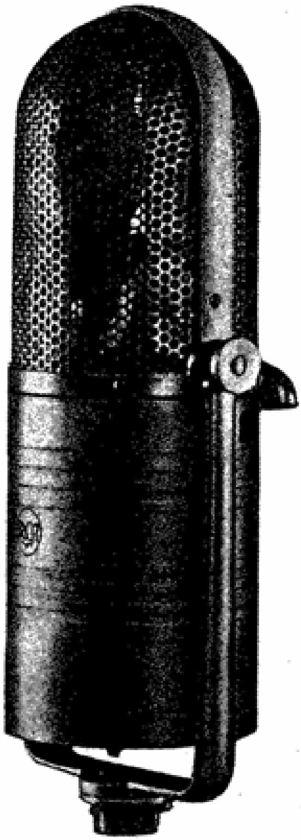
6. UNI-DIRECTIONAL MICROPHONE

The uni-directional microphone is a microphone that is partly pressure and partly pressure-gradient operated. The moving element

is again a ribbon, but it is divided into two sections. One section acts as a pressure-gradient microphone as described above. The other section is closed at the back by means of an acoustic labyrinth which removes the pressure from the rear. Thus the lower section of the ribbon is a pressure-operated device.

The output voltages of the two sections of the ribbon are in series. For sounds coming from the front, the developed voltages add, but for sounds coming from the rear, these voltages are 180° out-of-phase and, therefore, cancel. Figure 57 shows the directional characteristic of (a), a pressure-gradient microphone; (b), a pressure microphone; and (c), the combination of these to make a uni-directional microphone.

The uni-directional microphone is used where it is desired to discriminate against sounds coming from the rear of the microphone and which would be picked up by a pressure-gradient microphone. It also has a wider pick-up angle.



—Courtesy RCA Manufacturing Co.

Figure 51 — Uni-directional microphone.

7. CRYSTAL MICROPHONE

The crystal microphone is one of the latest types to be developed (Figure 53).

Its operation depends upon the principle that certain crystals have piezo-electric properties, that is, the crystal develops electrical charges on certain surfaces when subjected to mechanical stress.

The usual construction is such that the sound pressure causes the crystal to be bent, and as a consequence of the strain, voltages are produced which are proportional to the pressures of the sound waves.

The crystal microphone has a very good characteristic, provided there are no temperature changes during recording. The resonant peaks are above the audible range, and the inherent hiss is low. However, the

impedance is very high, sensitivity low, and the microphone is unstable when subjected to humidity and temperature changes.

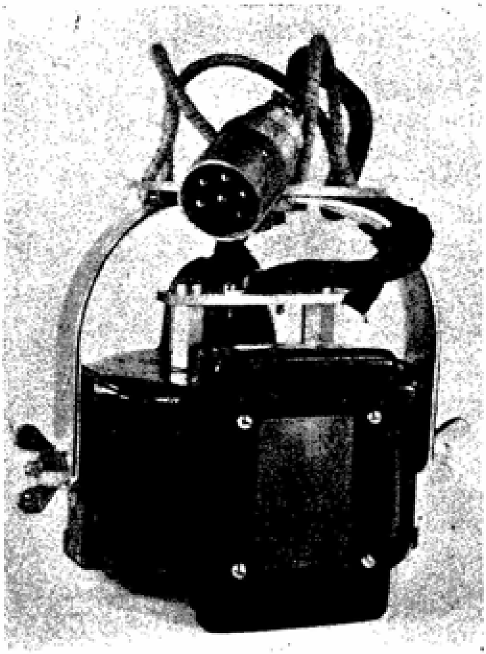
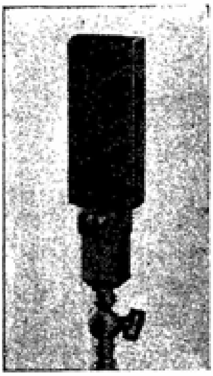


Figure 52 — Siemen's uni-directional microphone.

For these reasons, this type microphone is excellent as a test microphone, but is only used in production on musical stages where temperature and humidity conditions can be rigidly controlled.



—Courtesy Brush Development Co.
Figure 53 — Crystal microphone.

The sensitivity may be increased by the use of a diaphragm mechanically linked to the crystal. Through this lever action obtained, a greater force, for any given sound pressure, is applied to the crystal surface. This type is not used in the studios as this construction results in resonant peaks in the audible range.

The output of these microphones can be carried through a concentric cable to a separate amplifier (maximum of 50 ft.) resulting in a reduction in output level without varying the frequency characteristic.

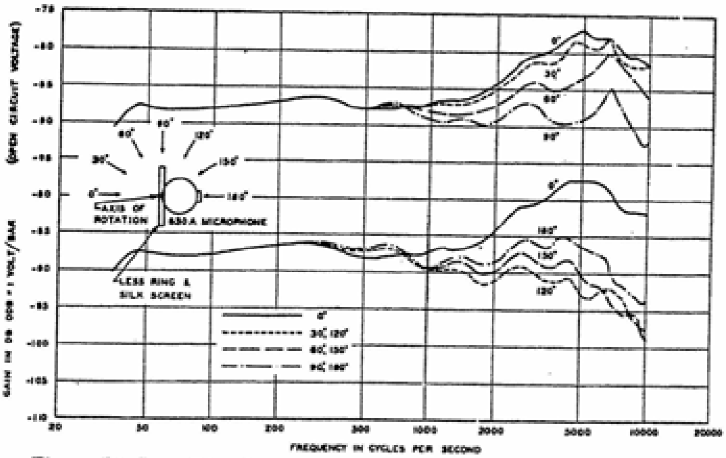


Figure 54-A — Directional response of a small dynamic microphone.

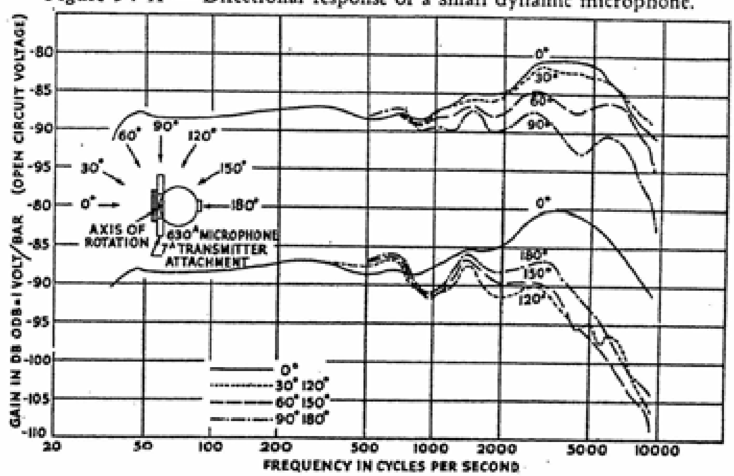


Figure 54-B — Directional response of a small dynamic microphone with baffle attachments.

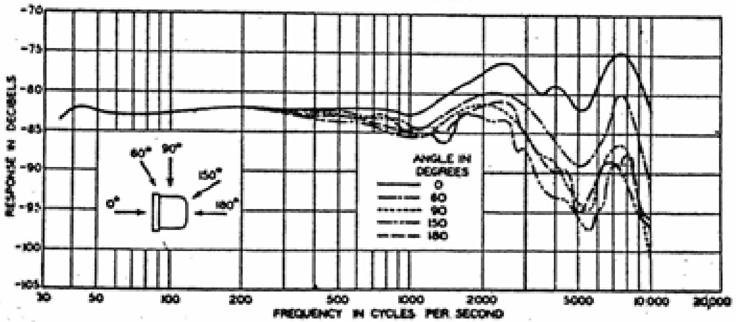


Figure 54-C — Directional response of a larger dynamic microphone.

—*Courtesy Electrical Research Products, Inc.*

8. DIRECTIONAL EFFECTS OF MICROPHONES

The directional effects of pressure-operated microphones depend upon the size of the microphone, because as the size of the diaphragm decreases, the microphone tends to act more as a "point" receiver of sound, and the effect of pressure doubling decreases rapidly.

Figure 54 shows the typical response at various angles of incidence of dynamic microphones of different sizes. Figure 54-A gives the response of a microphone, the diameter of which is approximately 2½", while Figure 54-B gives the response for this microphone when equipped with attachments to give a directional quality. Figure 54-C gives the response of a microphone with a diameter of 3¼".

The directional effects of the pressure-gradient devices depend upon the angle of incidence of the sound wave with respect to the ribbon surface. Figure 55-A shows the response of such a microphone at

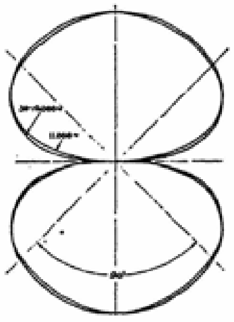


Figure 55-A.

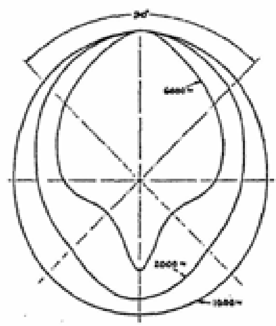


Figure 55-B.

—*Courtesy RCA Manufacturing Co.*

various angles of incidence. If the sound wave is approaching the ribbon normal to its surface, the air particles have a much greater surface to act upon, and consequently affect the ribbon to a much greater extent than if approaching from an angle on either side of the normal.

From Figure 56 it may be seen that when a sound wave is approaching from the direction indicated, the effective surface of the ribbon is no longer S but S', and the result is a weaker response from the ribbon. Figure 55-A shows that this effect is much more critical as the sound approaches nearly parallel to the plane of the ribbon.

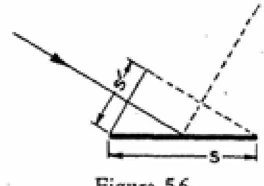
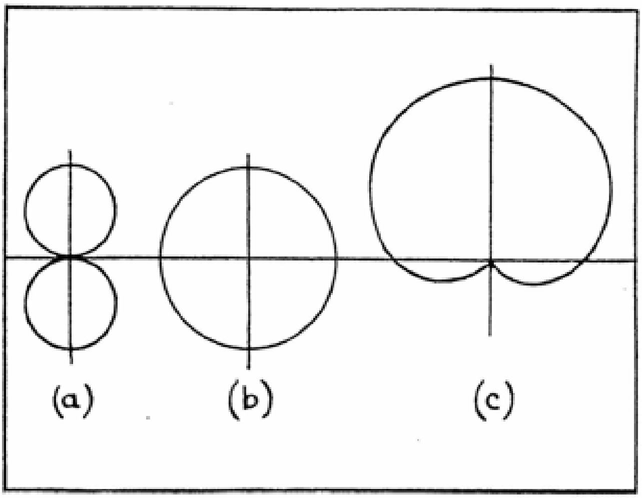


Figure 56.

Figure 55-B shows typical directional characteristics for pressure-operated microphones.

Figure 57 shows the directional effect of the uni-directional microphone. The velocity effect (a) combines with the pressure effect (b) in an additive manner on one side of the microphone axis, and in a neutralizing manner on the other side of the axis, giving the combined result as shown at (c).



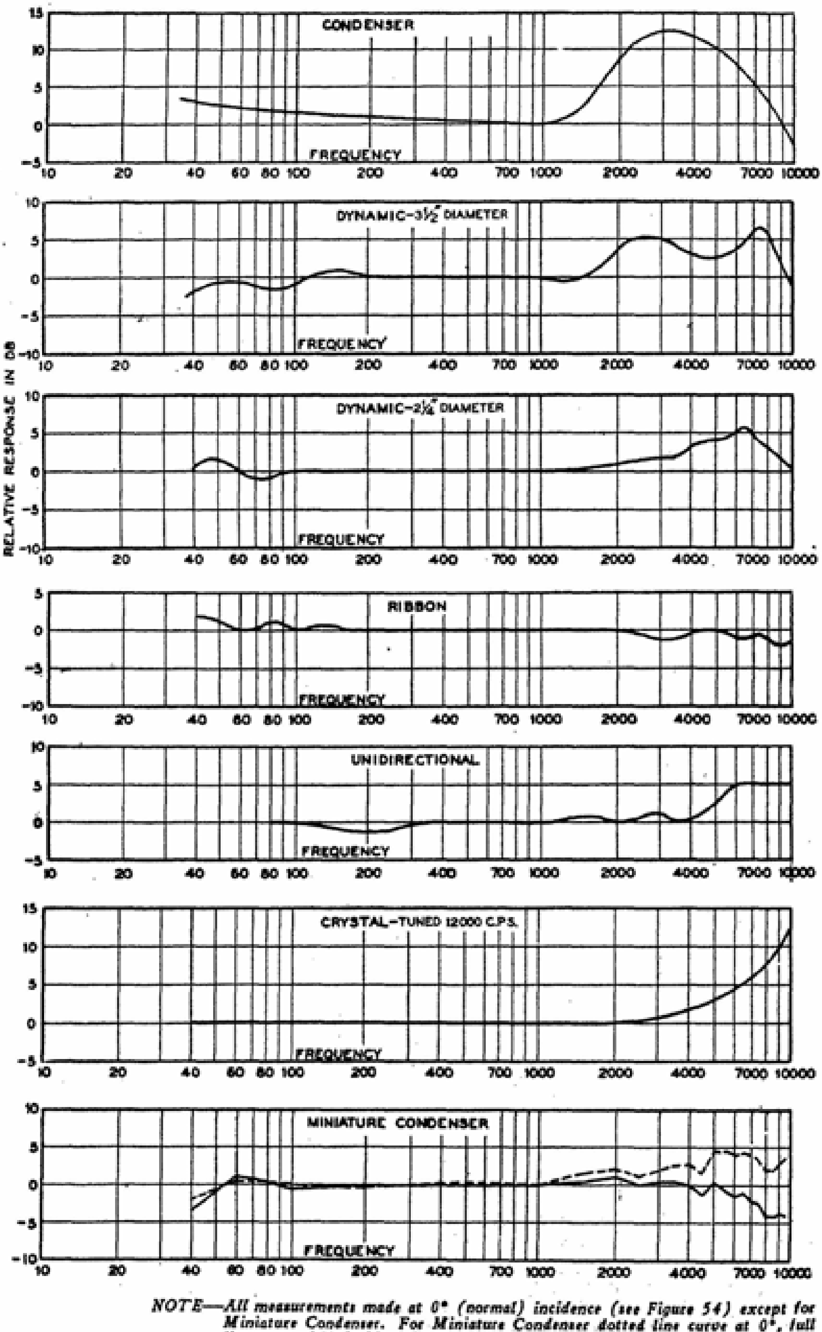
—*Courtesy RCA Manufacturing Co.*

Figure 57 — Comparison of directional characteristics of velocity (a), pressure (b), and (c), a combination of pressure- and velocity-operated devices.

The requirements for microphones for motion picture production are such that they seldom can be placed in the ideal position for sound pickup, as camera lines, microphone shadows, etc., must be taken into consideration when the microphone placement is being made. As a consequence, the microphone position is in most cases a compromise. This results in the use of microphones which are not too critical in their placement and which are fool-proof and uniform under operating conditions, even though under laboratory conditions they might not prove to be technically the best instruments.

BIBLIOGRAPHY

(1) *Applied Acoustics*, Olson and Massa.
(2) *Vibrating Systems and Sound*, Crandall.
(3) *Physical Review*, July, 1917, p. 39, E. C. Wente.
(4) *Physical Review*, May, 1922, p. 498.
(5) *Journal Acoustical Society of America*, Vol. III, No. 1, p. 44, Wente and Thomas.
(6) *Journal Acoustical Society of America*, Vol. III, No. 1, p. 56, H. F. Olson.



NOTE—All measurements made at 0° (normal) incidence (see Figure 54) except for Miniature Condenser. For Miniature Condenser dotted line curve at 0°, full line curve 90° incidence.

Figure 58 — Typical characteristics of commercially manufactured microphones.

Chapter VII

HEADPHONES AND LOUD-SPEAKERS

By L. E. CLARK *and* JOHN K. HILLIARD

Headphones and loud-speakers are electro-acoustic transducers which receive power from an electrical system, convert it to mechanical power, and then deliver it to an acoustic system. This transfer is usually made by the electrical current reacting with a magnetic field and causing an armature or coil of wire to vibrate in step with the current. The armature or coil is mechanically connected to a diaphragm which in turn vibrates the air and produces sound.

1. HEADPHONES

Headphones, as their name implies, are used against the ear, and produce variations in pressure in the small volume of air trapped in the ear between the ear drum and the diaphragm of the phone.

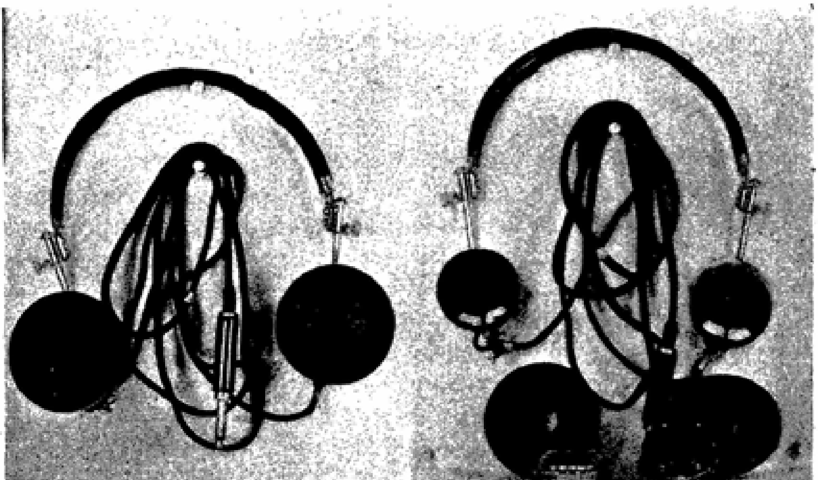


Figure 59 — Typical headset used in the studios, left view assembled, right view unassembled (ERPI No. 705).

The electro-magnetic telephone receiver, which was the earliest magnetic-type headphone, consists of an iron diaphragm vibrated by the variations in a magnetic field produced by the electric current in coils

placed over the pole tips of a permanent magnet. This receiver has a very poor frequency response characteristic, reproducing only the middle tones because of the sharp resonance of the diaphragm.

A so-called high-fidelity headphone has recently been designed which employs a moving coil, or electro-dynamic, telephone receiver, similar in construction to the moving coil microphone. The microphone and receiver cannot be used interchangeably, however, as their design requirements are different.

In the electro-dynamic receiver, the diaphragm is actuated by the force developed between a magnetic field and the electrical current flowing through the voice coil which is rigidly coupled to the diaphragm. As in the case of the microphone, the frequency response is smoothed out by an acoustic network coupled to the diaphragm.

Another type of high-fidelity headphone employs a straight conductor, or ribbon, which serves the dual purpose of carrying the actuating current and performing as a diaphragm. Its response is smoothed out by an acoustic network coupled to the ribbon.

High-fidelity headphones must be designed with a consideration of the effect of the acoustical properties of the ear, such as the air chamber in front of the ear drum, etc., and of the acoustic response of the ear, which has been discussed in Chapter II. Account must also be taken of a nominal amount of leakage which occurs between the phone unit and the ear.

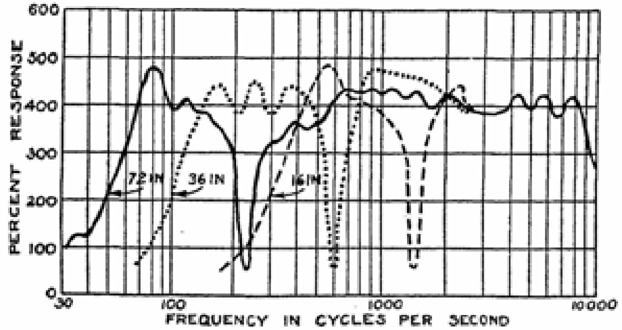
2. LOUD-SPEAKERS

Loud-speakers are electro-acoustic transducers designed to handle relatively large amounts of power, and to distribute the reproduced sound over a large area. They differ from headphones in that they must be able to produce more acoustic power and must radiate it into a different type of receiver. Fortunately, they are not subject to the same space and weight limitations which must be imposed upon headphones. In general, loud-speakers are unable to radiate low frequencies efficiently because loss of coupling occurs between the diaphragm and transmitting medium. For this reason they are usually used in conjunction with a baffle or horn.

The earliest speakers were electro-magnetically operated and were very similar to the electro-magnetic receiver. Some used a large cone for a radiating surface while others were merely powerful telephone receivers coupled to horns. They were subject to the same deficiencies in response as were the electro-magnetic receivers.

Most present day loud-speakers are of the electro-dynamic type and employ either a diaphragm or a cone. For radio receivers and other low

power applications, the speaker is usually used in conjunction with a flat baffle to increase the low-frequency radiation. As the cone vibrates, the waves sent out from the front and from the back are 180° out-of-phase, and there is a tendency for the pressure built up in front of the cone to



—From "Applied Acoustics," by H. F. Olson and Frank Massa (P. Blakiston's Son & Co., Inc.).

Figure 60 — Percentage response-frequency characteristic of 8 inch cone loud-speaker mounted in various sizes of square baffles. Measurements in free space, microphone distance equal to 10 feet.

be relieved by a flow of air into the corresponding region behind the cone instead of being radiated outward. This effect is of more importance for low frequencies where the path between the front and back of the cone is an appreciable part of the wave length. The flat baffle increases this path, prevents the air flow, and consequently increases the low-frequency radiation from the cone. Figure 60 shows the frequency response of a cone speaker in baffles of different sizes.

For motion picture theatre applications and other applications where large amounts of high quality acoustic power are required, horns or directional baffles are employed because of their much better performance and their ability to direct the power into useful areas. The horn functions much in the same manner as the flat baffle in preventing pressure equalization and at the same time results in higher efficiency and desirable directional characteristics. It can be thought of as an acoustic matching transformer that couples the relatively heavy diaphragm to the lighter air.

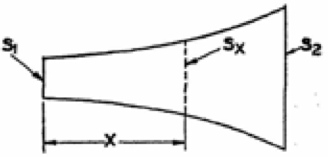


Figure 61 — Exponential horn.

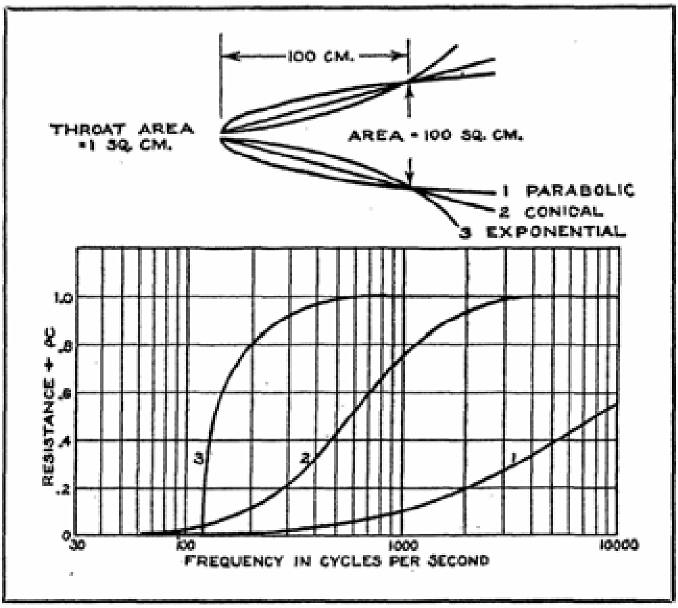
It is general practice to use horns with an exponentially increasing cross-sectional area, i.e., the area S_X at a distance, X, from the throat

is (see Figure 61),

$$S_X = S_1 \epsilon^{MX} \tag{18}$$

where S_1 = Area at throat
S_X = Area at distance X from throat
X = Distance from throat
$\epsilon = 2.718$
M = Constant which determines the taper of the horn.

The theory of horns shows that, while a horn of almost any shape will cause the impedance match to be attained at a low frequency, a horn of exponential shape will cause it to be attained at the



—*From "Applied Acoustics," by H. F. Olson and Frank Massa (P. Blakiston's Son & Co., Inc.).*

Figure 62 — Throat resistance of infinite parabolic, conical and exponential horns, having the same throat opening and the same cross-sectional area at 100 centimeters from the throat.

lowest frequency possible in any design. The impedance increases to a maximum value much more rapidly in an exponential horn. Figure 62 shows the throat resistance of different types of horns.

The resistance at the throat is given by

$$R = \frac{\rho C}{S_1}\sqrt{1 - \left(\frac{MC}{2\,\omega}\right)^2} \tag{19}$$

where ρ = density of air
C = velocity of sound
$\omega = 2\,\pi f$

For frequencies below $\omega = MC/2$, the resistance is zero, and the horn acting as a high-pass filter can radiate no power. Therefore, in the design of a practical horn, it is necessary to choose M so that the cut-off frequency of the horn is well below the lowest frequency to be radiated.

In addition, consideration should be given to the size of the mouth: a usual criterion being that its perimeter should be comparable to the wave length of the lowest frequency it is to pass. The size of the throat is governed on one hand by the largest practical size of cone or diaphragm and on the other by the restrictions imposed on the length of the horn by space requirements in the theatre, etc.

The remainder of this chapter is devoted to a description of the speaker mechanisms and horns that are used in present day high quality reproducing systems.

TWO-WAY HORN SYSTEM*

3. INTRODUCTION

The present investigation was undertaken with a two-fold purpose in mind; first, to study thoroughly the more important types of extended range loud-speaker systems in current use, and second, to develop if possible, a system which would combine practicability for theatre use with as great an improvement in quality and efficiency as could be obtained without greatly increased cost. The first objective necessarily involved an effort to learn as much as possible of the "why" as well as the "how" of the systems and individual speakers studied, while the second led to considerable investigation of certain aspects of loud-speaker design, some of which—at least in the literature of the subject—seem not to have been sufficiently emphasized in the past.

Any investigation of as wide a scope as the present one inevitably furnishes many facts not pertinent to the main issue, but useful in other fields. The main body of the paper has been written with the problem of the reproduction of sound for motion pictures ever in mind, and should be read from that viewpoint. It is felt, however, that the results referred to may form a definite contribution to other fields, such as public address work and home radio.

4. SOUND REPRODUCING SYSTEMS FOR MOTION PICTURE THEATRES

The art of modern reproduction of sound in motion picture theatres is now about eight years old. During this time there has, of course, been

* Reprint from the Technical Bulletin of the Research Council of the Academy of Motion Picture Arts and Sciences, March 3, 1936.

considerable improvement, but there has been only one major change in the standard theatre installation. This change was the adoption of the "Wide Range" and "High Fidelity" systems after 1933. The principal modifications involved were: First, a partial fulfillment of greatly needed increase in amplifier carrying capacity; second, the adoption of speaker systems which provided for the division of power between two or more groups of speakers, each operating over a limited-frequency range; and third, improvements in the sound head which reduced flutter. While these improvements considerably raised the standard of reproduction in the theatre, it was felt that the loud-speaker system still constituted the principal limitation to naturalness of reproduction. An investigation was accordingly made to determine whether a speaker system could be developed which would economically replace the present systems while providing the much needed increase in fidelity. This was found to be the case, and it is the purpose of the present paper to describe this system and the results obtained with it, and to compare it with previous systems.

Since it was not known how great a departure from a full range linear response could be tolerated for the purpose in hand, it was considered advisable to start with a system as near this as so far achieved, even though the form of apparatus available by its size and cost would prohibit its use for theatre installations. From this it was determinable how much deviation was allowable and necessary in order to obtain a commercially practical system. Such a linear system was made available, and a series of tests led to the following specifications which were found to be adequate for theatre reproduction, taking into consideration further developments in recording which may be expected within the next few years.

5. SPECIFICATIONS

Flat Overall Frequency Characteristic. The system shall not deviate by more than plus or minus 2 db, from 50 to 8000 cycles, over the entire angle of distribution within ten feet of the mouth of the horn.

High Electro-Acoustical Efficiency. It shall approach fifty per cent in order that the required amplifier capacity will not be too great.

Volume Range. The volume range shall be at least 50 db and preferably 60 db.

Reasonable Cost.

Absence of Transient Distortion and "Fuzziness." The electro-acoustical transducer shall be of such construction that it will not generate objectionable harmonics up to the peak power required, and

the phase delay between units shall be such that the sound will be equivalent to that coming from a single source.

Suitable Angular Distribution Characteristics. The sound shall be radiated through a horizontal angle as great as 110 degrees and a vertical angle of 60 degrees, with nearly uniform response at all positions.

Reasonable Compactness and Portability. Low weight.

Amplifier Capacity. The installed amplifier capacity shall be such that one acoustic watt per one thousand square feet of floor area can be delivered when the auditorium is adjusted for optimum reverberation time.

A system which will conform to, or exceed, these specifications has now been developed, and can be constructed at moderate expense.

In order to take advantage of these characteristics it has been found that when film is reproduced over a system such as this, it is necessary to keep the flutter from the sound head no greater than 0.1 of 1 per cent. Although the problem of flutter has been satisfactorily solved and heads are commercially available which will pass the 0.1 of 1 per cent flutter specifications, it should be pointed out that by far the largest majority of heads in use today will not meet this specification.

6. POWER AND FREQUENCY REQUIREMENTS

The history of the electrical reproduction of sound has been one of continual increase in amplifier carrying capacity, and in this respect, the theatre installation is no exception. Originally, output powers from 2.5 to 12 watts were considered adequate for most houses. With the advent of the later systems now in use, these powers were recommended to be increased from 3 to 6 db, depending upon the size of the house. It has been found from this investigation that it is both practical and eminently desirable to make a further increase of at least the same amount. The figure given of one acoustic watt per one thousand square feet of floor area is felt to be the minimum which will do justice to the advanced conception of reproduction of records produced with modern recording technique. It is of interest to note that this figure can be achieved allowing for considerable latitude above this point without danger of mechanical damage to the units.

The advisability of extending the frequency range of a reproducing system must be determined by balancing the gain in naturalness obtained by the extension, against the resulting increase in noise and extraneous sounds. In the present state of the recording art, a characteristic flat to 6,000 cycles is the least that will do justice to the film: an extension to 7,000 or even 8,000 cycles is advisable, and a furhter

extension is not. This is so because a further extension becomes of less and less value, due to the decreasing sensitivity of the ear and the small amount of energy in this region, and especially because above 8,000 cycles, noise, flutter and harmonics due to recording deficiencies become decidedly the limiting factor. Incidentally, since practically all recording systems include a low-pass filter with a cut-off in the neighborhood of 8,000 cycles, there is nothing on the film at high frequencies to be reproduced.

Once the high-frequency limit is chosen, the low-frequency limit is automatically fixed. It has been found that for ideal balance the product of the two cut-off frequencies must be fairly close to 400,000, so that for an 8,000 cycle upper cut-off, the lower becomes 50 cycles.

7. HIGH-FREQUENCY HORN

One of the principal limitations of present theatre installations is bad directional characteristics. The plain exponential horn has a directivity which varies with frequency; low-frequency sound is projected fairly uniformly over a wide angle, but as the frequency is increased this angle decreases rapidly until at frequencies of several thousand cycles practically all of the energy is emitted in a narrow beam. The result of this is that the reproduction becomes very "drummy" or "bassy" for that portion of the audience whose seats lie well off the axis, while the opposite is true for seats located directly on the axis. In the present system this effect is eliminated by using a radiating system for the high-frequency unit which is composed of a cluster of small exponential horns, each having a mouth opening of approximately sixty square inches. These individual units are stacked in layers to form a large horn, the mouth opening of which is spherical in shape. The principle of this high-frequency unit can best be likened to a further compacting of the typical cluster of loud-speakers, as customarily used in auditoriums and stadiums for public address systems and announcing, except that the whole array is fed from a common header and driven by two dynamic units. This type of high-frequency radiation is also a feature of the aforementioned reference system. However, the reference horn, having been developed to a very limited angle and being driven by a single mechanism, was not adaptable to theatre use, as more than one horn became necessary for full coverage. This would result in non-uniform distribution as well as complete loss of coverage for a large part of the auditorium, should one unit fail during a performance.

One of the features of the reference system is the use of a single diaphragm to reduce phase distortion. Inasmuch as theatres require parallel operation as protection in the case of failure of one unit, experiments

were made with a Y throat and two units. As a result of these experiments, it is now recognized by all concerned that any increase in phase distortion which may be introduced by the Y throat is negligible.

The diaphragms are made of duraluminum 0.002 inch thick and have an area of six square inches. The diaphragm is mounted on the

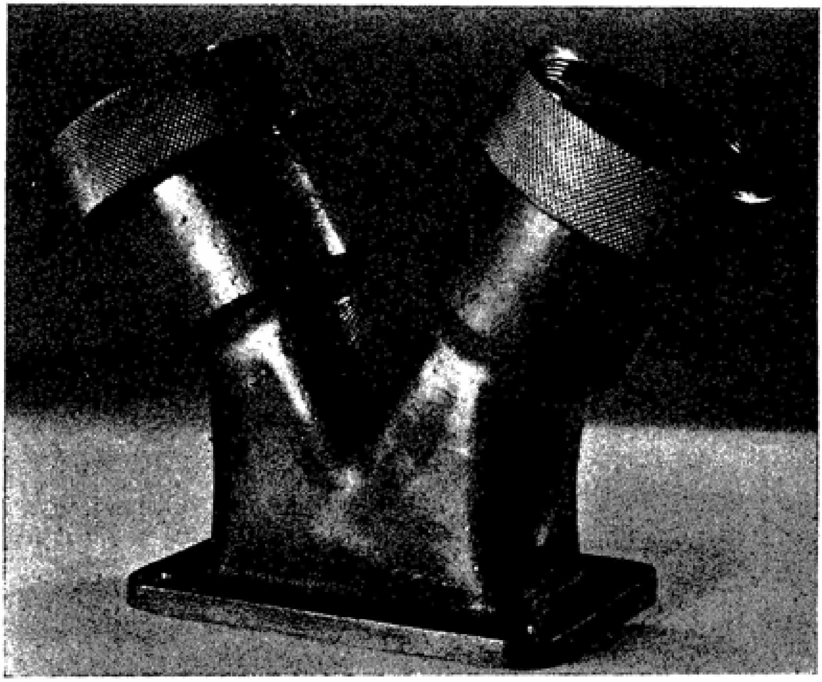


Figure 63 — The Y Throat.

back of the assembly and by the use of an annular opening, the sound that is admitted to the throat within the unit has a minimum phase distortion. (Figure 64.) This is still further reduced by having this throat exponential beginning at the annular opening, and avoids a sharp discontinuity that may exist with a tubular throat. Two units are connected by means of a Y throat to the multi-channel horn which tends to reduce the distortion of high throat pressure. The field excitation requires twenty-five watts per unit.

To obtain high efficiency energy transfer between the diaphragm and air column in an exponential horn loud-speaker, the acoustic impedance of the air must be matched with the mechanical impedance of the diaphragm. Such an impedance match is usually secured by the use of an acoustic transformer which provides a properly constricted

cross-sectional area of the sound channel between the diaphragm and the throat of the horn.

This impedance matching device provides a load for the diaphragm which may be taken as the ratio of the effective diaphragm area to the

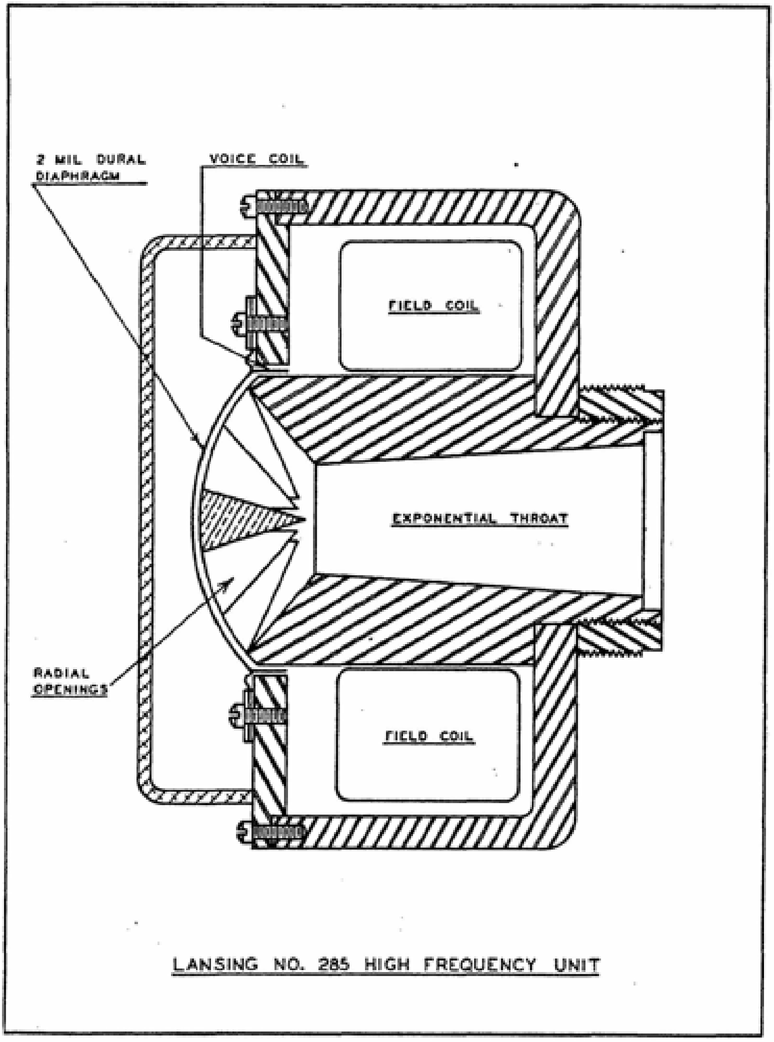


Figure 64.

cross-sectional area of the constricted sound channel. This loading factor, of such a value that it tends to damp out resonant action of the diaphragm, should provide a loading for the diaphragm which is as uniform as possible over its entire area.

The sound channel dimensions between diaphragm and horn must be of such values that waves of any frequency within the recording band which emanate simultaneously from different points of the diaphragm, will not meet at any point in the throat 180° out-of-phase—since in this case destructive interference between such waves would take place.

It is also required that any waves which originate from all points along a radial element or sector of the diaphragm, form a wave front of a desirable shape at the throat of the horn.

One method considered, to avoid destructive interference at the throat between waves which originate at the center and outside edge of the diaphragm, was to provide individual sound channels of substantially equal length from the diaphragm to the throat of the horn. This

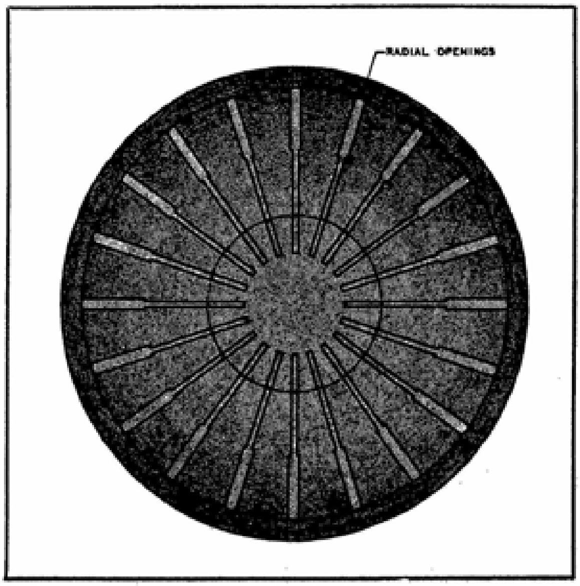


Figure 65 — Slit structure of acoustic transformer—Lansing No. 285, high-frequency unit.

resulted in the construction of acoustic transformers in which there were formed a number of annular concentric sound passages all of substantially equal length.

Such a device proved practical and gave good performance, but was difficult and expensive to construct.

To overcome these practical difficulties, another type of acoustic transformer has been constructed. This type is illustrated in Figure

65, in a view looking into the transformer from the diaphragm. It is rugged, of relatively simple construction, and inexpensive.

The transformer provides sound channels which are disposed radially so that each opening subtends a sector of the diaphragm. Waves originating along a continuously radial sector of the diaphragm will be passed by each sound channel in the manner in which they originate A wave front which emanates from each radial element of the diaphragm advances without further subdivision into the radial sound channel to the throat of the horn, arriving at that plane in an approximately spherical form. No destructive interference is encountered, providing the maximum distance from any point of the diaphragm to the opening of the nearest radial sound channel is less than a quarter wave length of the highest frequency to be reproduced.

A loud-speaker provided with an acoustic transformer designed in accordance with these considerations is shown in Figure 64. Such a loud-speaker reproduces at a response level which is highly uniform throughout the range from 300 to 10,000 cycles.

The directional characteristics of the resulting unit are very satisfactory as found from theatre installations. It should perhaps be emphasized that lack of good distribution cannot be corrected by equalization in the electrical circuits, since for any given adjustment, the overall response is a highly varying function of position in the house. Although the characteristic can be made flat for any given position, it can not be made so for all, or even a large part of the house by this method.

8. LOW-FREQUENCY HORN

In the case of a low-frequency unit, a suitable driving mechanism was not available, and it became necessary to develop one. The unit finally adopted consisted essentially of an exponential horn with a mouth area of fifty square feet, and an axial length of forty inches, driven by four fifteen-inch dynamic units of special design. The mouth opening was extended laterally to form a flat baffle, 10′ x 12′. The paper cones are dipped with lacquer to prevent them from absorbing moisture, which would vary their response. They are connected in series-parallel to give a desirable impedance characteristic as well as to provide insurance against complete failure of the system in the event any individual unit fails. The angle of distribution is uniform through an arc of fifty degrees on each side of the axis. The use of a horn instead of a flat baffle board for low frequencies has several advantages: The efficiency is raised from ten or fifteen per cent to better than fifty per cent, which effects an enormous reduction in amplifier capacity; and undesirable radiation from the rear of the unit is considerably reduced, resulting in a decrease

to a negligible amount of the usual objectionable back stage low-frequency "hang-over." For purposes of further compactness and rigidity the low-frequency horn may advantageously be folded and in this form retains the same characteristic, if the air path length be maintained unchanged. This modification was contributed by Dr. H. F.

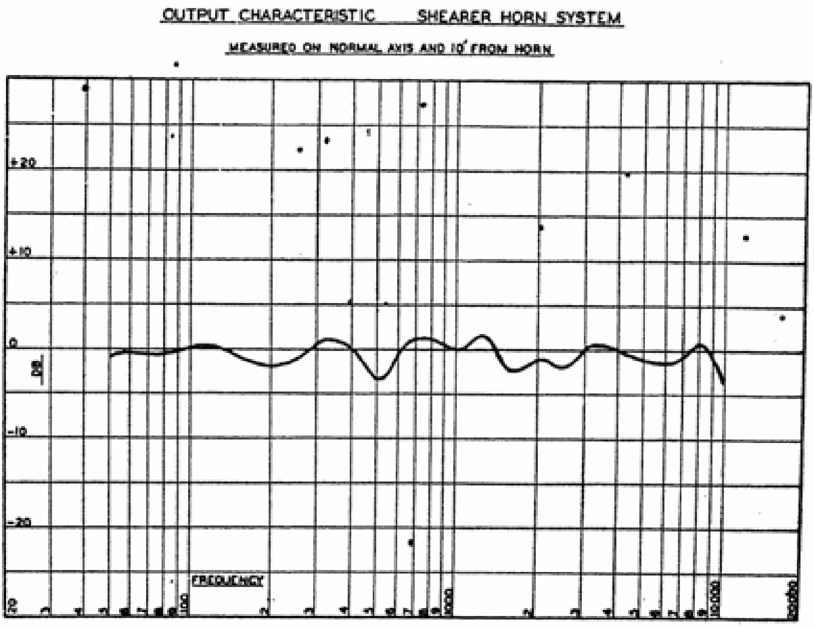


Figure 66.

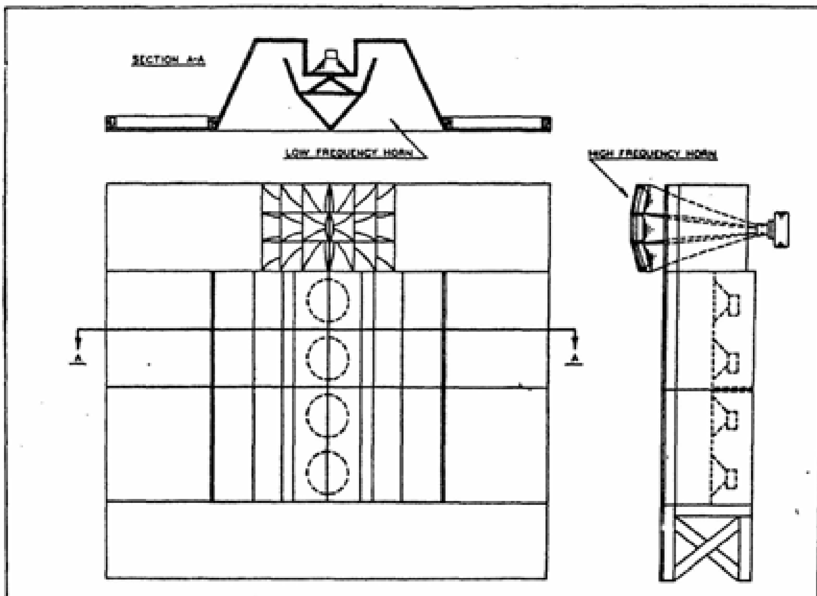
Olson of the RCA Manufacturing Co. The loading provided by the air column of the horn decreases the excursion of the diaphragms as compared to the excursion necessary to produce equivalent output from a flat baffle array, and distortion is correspondingly reduced. (Figure 66.)

With the low-frequency horn length, as specified in the design under discussion, maintained approximately equivalent to the length of the high-frequency horn, there is no time delay between the component sounds from the two horns.

9. HORN ASSEMBLY

The folded horn is assembled in sections, each section containing two driving mechanisms. They may be stacked one upon the other, depending upon the number required. Each section is adequate for an output from the amplifier of 25-30 watts for the required minimum

harmonic content. If it is desired to secure a wide lateral distribution, the sections may be placed side by side. Section AA, Figure 67, shows the construction of the horn.



SHEARER TWO WAY HORN SYSTEM
FOLDED TYPE

Figure 67.

The entire horn is assembled so that the center of the high-frequency unit is approximately 50 to 60 per cent of screen height. This position has been found by years of use to be the center of activity or "presence" on the screen and since the high-frequencies are responsible for determining the presence, the unit was so arranged. In order to keep the sound as near a point source as possible, the low-frequency horn is maintained at a position near the high-frequency horn. (Figure 68.)

The complete assembly is a unit so that it can be moved away from the screen or raised and lowered with the screen with a minimum of effort. The use of sections for the low-frequency horn allows the horn to be shipped and moved into spaces which have standard size doors.

10. DIRECTIVITY

For both the low- and high-frequency units a certain amount of directivity is required, since the best illusion is obtained if the ratio of direct to reflected sound is as high as possible. In most theatre auditoriums there should be but little energy radiated at angles greater than

about forty-five degrees from the axis, as such energy will be reflected from the walls and will consequently lessen the illusion.

There is one additional consideration with regard to directivity which should be mentioned. Dr. V. O. Knudsen has shown that at

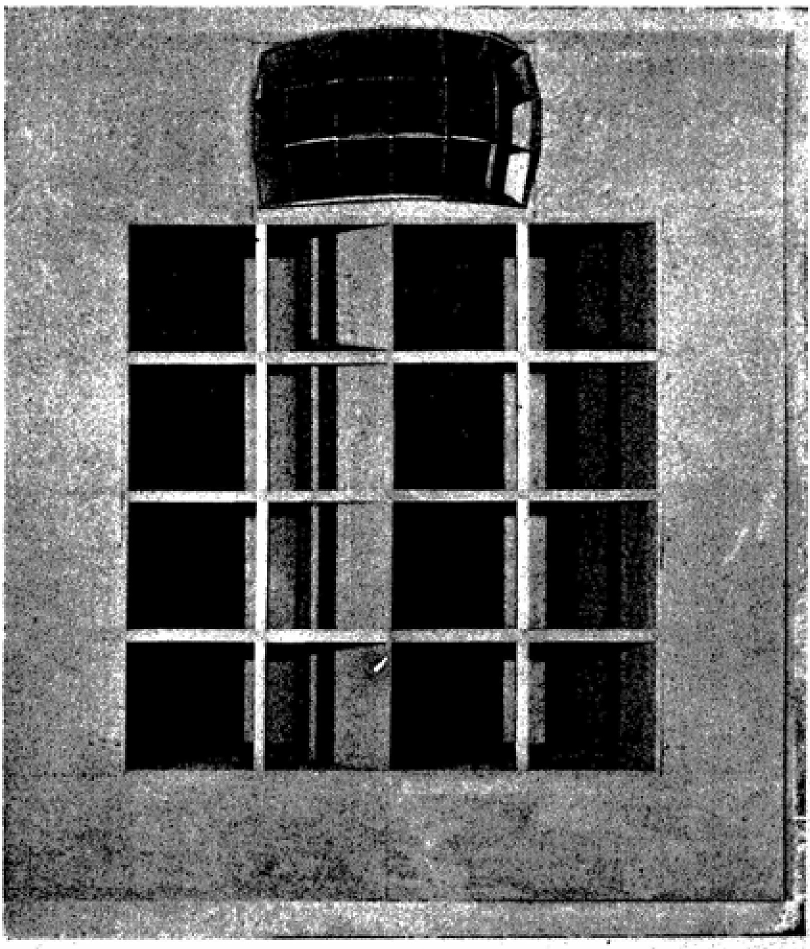


Figure 68 — The folded horn assembly.

the higher frequencies, *e.g.*, at 10,000 cycles, absorption of the atmosphere may become very serious, being as great as 0.2 db per foot under certain conditions of humidity and temperature. In large and deep houses this would result in a serious loss of high-frequencies in the rear seats. The effect can be considerably reduced by increasing the high-frequency radiations from those horns of the unit which serve these seats.

It may be done by putting a suitable amount of absorbing material in the other horns and re-equalizing to bring the overall response up to standard for the front seats. These artifices will probably not be required in most houses.

11. HARMONIC CONSIDERATIONS

A major defect of commercial loud-speakers is their large amplitude distortion. One of the striking improvements in the new system

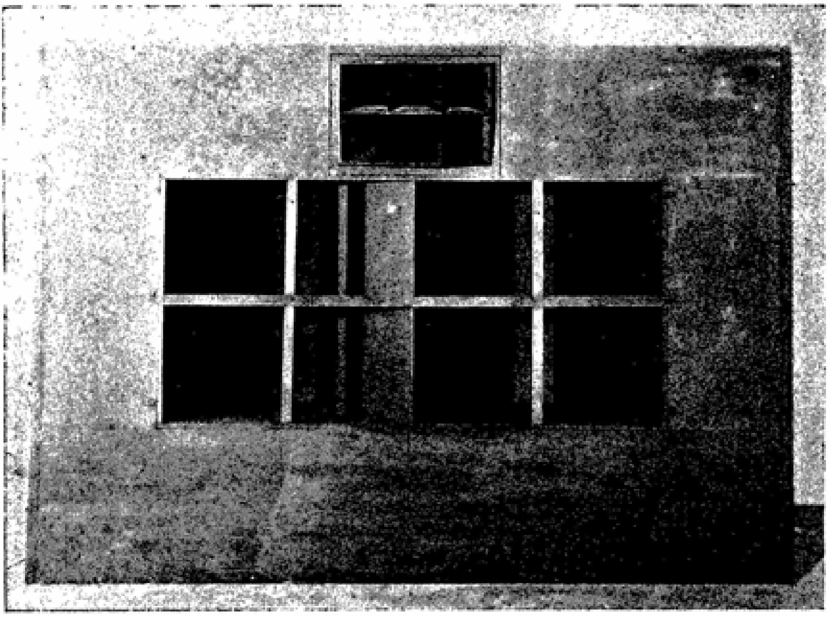


Figure 69 — Single section low-frequency folded horn with a 52 degree high-frequencies for use in studio viewing rooms and small theatres.

is its cleanness of reproduction at low frequencies. The measured harmonic content is less than four per cent at 40 cycles for 30 watts output. This is due in large part to the use of a thick and comparatively soft cone which can be driven to full excursion without break-up, and consequent harmonic production. It was found by actual listening tests that with a pure tone of forty cycles impressed, most of the cone speakers investigated gave a greater apparent loudness than the speaker finally adopted. However, when a direct comparison was made by keying the amplifier from the new unit to the unit under test, it was at once obvious that the output of the new one was fairly pure forty-cycle tone, while that of the other speakers consisted, in most cases, entirely of the second and higher harmonics. Direct measurement of the acoustic

output showed that in spite of its low apparent loudness, the fairly pure output of forty cycles was actually about 6 db higher than that of the other speakers.

This great increase in apparent loudness due to transferring part of the fundamental power into harmonics in the conventional speaker is very striking, and is undoubtedly the explanation for the alleged high efficiency of many present day speakers of all types. The loudness of the harmonics is not due to the rapid change in the sensitivity of the ear at low-frequencies which would favor the harmonics at the expense of the fundamental, since it also occurs at fairly high frequencies where the sensitivity of the ear is varying in the opposite way with frequency. With one particular pair of units tested, the effect was more striking at 1,000 to 2,000 cycles than at any other frequency. It is equally great with complex sounds, such as speech and music, although here the change in quality is somewhat less with respect to the change in apparent loudness than in the case with pure tone.

12. PHASING

Another important advantage of the new system is that it can easily be made to fulfill the requirements that the virtual sources of all the components of the reproduced sound shall coincide in the vertical plane. This condition is impossible to obtain with divided frequency range systems now in use in which the axial length of the several types of horns in a given system are widely different. In this respect, a two-unit system is much easier of adjustment than a three-way system. It might be thought that since the time delay is so small, of the order of a few milliseconds, that the effect would be inappreciable. This is true for certain types of sound such as sustained music passages, but on dialogue and especially certain types of sound effects which are of the nature of short pulses, a very objectionable distortion is usually noticeable. A striking demonstration of this fact was obtained by recording a tap dance. When this was reproduced it was found that the system with a very small time delay gave a naturalness of reproduction, while systems which had an appreciable delay reproduced the scene with far less realism. In fact, the sound did not appear to come from the screen, and, in addition, the tap was fuzzy in character with a decided echo.

This effect sounds somewhat like that of transient distortion due to the use of a filter with too sharp a cut-off, but it is actually more analogous to the echo effect often observed on long lines and with certain types of phase distortion networks.

A recent paper discusses the features of the three-way system, including some of the limitations which require special installation tech-

nique for the setting of horns, back stage draping, phasing of various horn positions, position of horns for distribution and setting of volume between horns. Familiarity with this data will assist in appreciating the principles of the present system.

It should be pointed out that the overall frequency response curve of the system should not fall off too rapidly beyond the cut-off fre-

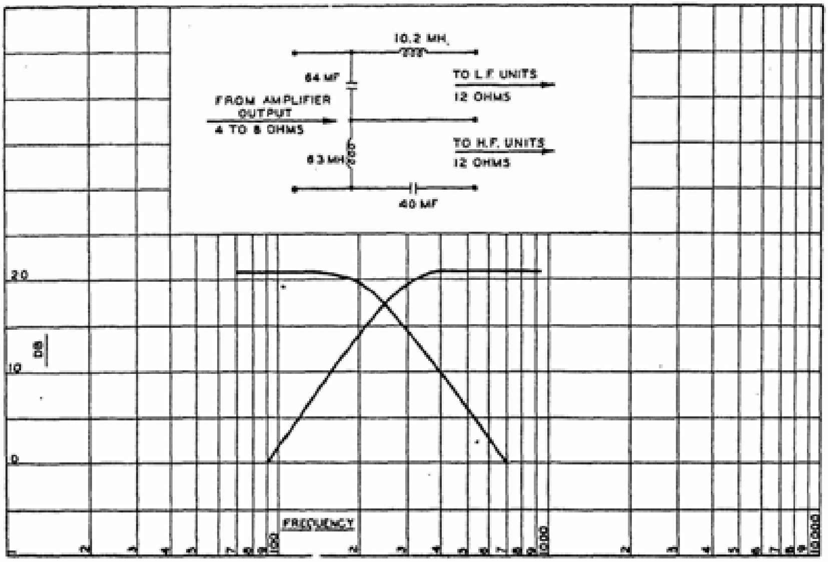


Figure 70 — Series type dividing network, Shearer Horn System.

quencies, or objectionable transient distortion will result. Probably the maximum slope that can be tolerated is of the order of 20 db per octave, or roughly, that of a single section constant-K filter.

13. DIVIDING NETWORK

The frequency chosen for the critical frequency of the dividing network is governed by several factors. If this frequency is too low, it leads to uneconomically large values of capacity in the network, and to impracticably large horns for the high-frequency unit. If too high, there is danger of running into the characteristic dip which seems to be always present in large cones, and also, it would result in dividing the prime energy of speech sounds between the two units, which is objectionable from the standpoint of good presence. If the critical frequency is chosen between 250 and 400 cycles, a good compromise results. (Figure 70.)

A dividing network was chosen which gave fairly rapid attenuation, 12 db per octave, in order to keep any appreciable low-frequency energy out of the high-frequency unit, and to minimize the effect of irregu-

larities encountered in the response curve above the designed range of the low-frequency cones. This lies somewhat above 400 cycles for an efficient low-frequency unit. Certain dividing networks in current use have attenuation curves of such gradual slope that at some frequencies the irregularities in the response curves of the speakers are actually greater than the attenuations of the networks.

The network is designed so that the reflected impedance of the horn on the amplifier is approximately 2.5 times the amplifier impedance. The loss in the network is less than 1 db in order that the full capacity of the amplifier may be utilized.

14. MEASUREMENTS

While it is recognized that indoor response measurements do not have the degree of precision that may be had in free space, they nevertheless do represent conditions under which the loud-speakers must actually be used for motion pictures. Also, for the purpose at hand, comparative measurements are sufficient and were verified by listening tests, which in the end are the final criterion. (Figure 66 shows average response.)

Irregularities in the sound pressure at the microphone due to standing wave patterns in the room are minimized by the use of a conventional warble frequency, varying plus and minus twenty-five cycles at a ten cycle rate. Tests have been run which indicate that the warble is only effective below 2,000 cycles. Above this point, the standing waves do not interfere with the correct interpretation of the response curve.

The measurements were taken in a stage 100′ x 70′ x 35′, having a reverberation time of one second at 512 cycles per second. By making these measurements indoors, tests could be made rapidly on a large number of units without interference from outside noises, due to a 60 db insulation between inside and outside, provided by the building.

The response curves were measured using a high speed level indicator capable of responding to a change in level as rapid as 300 db per second.

Douglas Shearer, head of the Metro-Goldwyn-Mayer Sound Department, brought about and directed this project. This development was engineered by the writer and contributed by Metro-Goldwyn-Mayer Studios. The cooperation of the following companies is gratefully acknowledged: Electrical Research Products, Inc.; RCA Manufacturing Co.; Lansing Manufacturing Co.; and Loew's, Inc. These companies assisted by making available test equipment the reference system, staff and theatres, which greatly facilitated the work and pro-

duced a co-ordinated result not otherwise possible. The writer also wishes to acknowledge the contribution of the Metro-Goldwyn-Mayer Sound Department, and in particular Robert L. Stephens, who has carried out the mechanical design.

15. DESIGN DATA

(a) Low-Frequency Exponential Horn

Fundamentally, the design of a low-frequency exponential horn follows the same treatment as that accorded a horn for high-frequency response. There is, however, a greater tolerance allowable in deviating from theoretically calculated values, namely: Expansion rate (governing value of cut-off frequency), mouth size, and nature of cross-section. Discontinuities which would be out of the question in high-frequency design may be permitted with little loss in a low-frequency horn. Numerous tests have borne out the above statement. A horn of folded cross-section has been chosen for general use in this system, because it permits a compactness of design not possible with a straight exponential horn. Sufficient loading has been obtained in a small space to permit the cone driving units to operate at their optimum efficiency.

For the purposes of illustrating the method of computation, a brief summary of the calculations involved in the design of a straight exponential horn will be given:

The cut-off frequency was chosen at 50 cycles per second. A 50 cycle wave has a length of 271 inches. The distance across the mouth of the horn should be at least equal to one-quarter the wave length of the lowest frequency it is desired to transmit. This value for the horn in question gives a minimum mouth size of 68 inches. The size of throat must be sufficient to accommodate four 15-inch cone speaker units. A throat size of 30 x 30 inches was chosen.

It has been found that an exponential horn whose area doubles every 12 inches will have a cut-off frequency of 64 cycles per second; one whose area doubles every 6 inches, a cut-off frequency of 128 cycles per second. From the above relationship the length for the area of the present horn to double, may be found by simple proportion:

$$\frac{64}{X} = \frac{50}{12}$$

from which $X = 15.36$ inches

From the general horn equation:

$$S_x = S_1 \epsilon^{Mx} \qquad \text{(Eq. 18)}$$

S_X = cross-sectional area at any point X
$\epsilon = 2.7183$
M = flare constant of horn
X = distance along horn axis from throat

where S_1 was chosen above as 900 square inches. M can be computed by substituting known values in the above equation:

$$1{,}800 = 900 \times 2.7183^{M15.36}$$

from which $M = 0.045$

Then the equation for the present horn becomes:

$$S_X = 900\,\epsilon^{0.045x}$$

from which the sectional area at all points X may be computed.

For a minimum mouth area of 4,624 square inches, the length is determined:

$$S_X = 900\epsilon^{0.045x}$$

$$4{,}624 = 900 \times 2.7183^{0.045x}$$

Where $X = 36\tfrac{1}{4}$ inches

It has been found, however, that while the sizes above are satisfactory from a theoretical standpoint, an increase in loading will result in a higher efficiency. An increase in length to 44 inches with a corresponding mouth size of 80 inches or 6,400 square inches has, as a result of tests, proven to be perhaps the most desirable size. The overall length, inclusive of units, then becomes approximately 55 inches. This length is considerably more than is desirable for the majority of installations.

The above analysis applies to the straight type horn rather than the folded type.

Figure 67 illustrates a horn of folded cross-section. Here it is possible to retain optimum loading conditions in a minimum space. It is, however, in this case mechanically impracticable to construct a horn of true exponential shape.

The mouth, throat size, and flare constant are determined as in the case of the straight exponential horn. Intermediate cross-sectional areas are approximated to those of a true exponential horn as closely as is feasible without involving constructional difficulties.

It has been found that the difference in response is sufficiently slight to justify this deviation from the theoretical.

(b) High-Frequency Exponential Horn

The specifications require that the overall depth or length of both low- and high-frequency assemblies does not exceed 44 inches.

This limitation of length brought about the selection of a theoretical cut-off frequency of 220 cycles per second. This value of cut-off allowed the design of a horn which fulfilled the desired requirements, such as a spread of either 90 degrees or 105 degrees with a maximum of six separate channels and a sufficient mouth size to present a reasonably small amount of discontinuity.

A brief summary of the design calculations follows:

It has been found that an exponential horn whose area doubles every 12 inches will have a cut-off frequency of 64 cycles per second; one whose area doubles every 6 inches, a cut-off frequency of 128 cycles per second. Then by simple proportion the length for the area of the present horn to double may be found:

$$\frac{64}{X} = \frac{220}{12}$$

from which

$$X = 3.5 \text{ inches}$$

From the general horn equation:

$$S_X = S_1 \epsilon^{MX} \qquad \text{(Eq. 18)}$$

where

$$S_1 = \text{area of throat (chosen as } \tfrac{1}{4} \text{ square inch)}$$

M can be computed by substituting known values in the above equation:

$$\tfrac{1}{2} = \tfrac{1}{4} \times 2.7183^{3.5M}$$

from which

$$M = 0.2$$

Then the equation for the present horn becomes:

$$S_X = \tfrac{1}{4}\ \epsilon^{0.2X}$$

From which the sectional area of the horn at all points X may be computed.

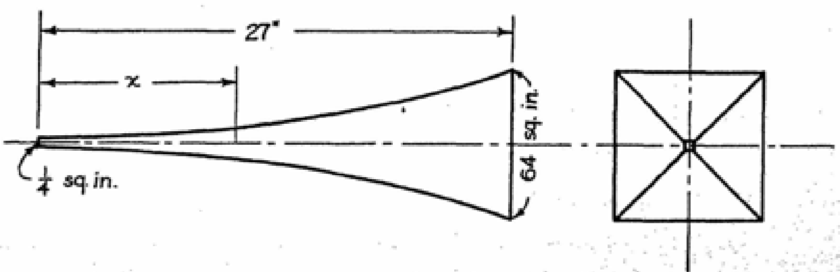


Figure 71 — A typical individual channel.

BIBLIOGRAPHY

(1) *Vibrating Systems and Sound*, Crandall.

(2) *Applied Acoustics*, Olson and Massa.

BIBLIOGRAPHY (Two-Way Horn System)

(1) *Wide Range Reproduction in Theatres*, J. P. Maxfield and C. Flannigan—*Journal Society of Motion Picture Engineers*, Vol. XXVI, No. 1—January, 1936, Page 67.

(2) *Loud-Speakers and Microphones*, Wente and Thuras—*Electrical Engineering* —January, 1934, Pages 17 to 25.

(3) *Acoustic Power Levels in Sound Picture Reproduction*, Wolfe and Sette—*Journal Acoustical Society of America*, Vol. II, No. 3—Pages 384 to 398.

(4) *The Effect of Humidity Upon the Absorption of Sound in a Room*, V. O. Knudsen—*Journal Acoustical Society of America*—July, 1931, Vol. III, No. 1, Part 1—Pages 126 to 138.

(5) *High Speed Level Recorder for Acoustic Measurements*, Wente, Bedell and Swartzel, Jr. — *Journal Acoustical Society of America* — January, 1935, Page 121.

Chapter VIII

FILM DRIVE

By WESLEY C. MILLER

Uniformity in the speed of the record and perfect synchronism between sound and picture are vitally important to the sound picture. While the subject matter in this book is largely devoted to the problems of transmission, a brief mention of the film-driving means may be desirable.

The internationally accepted standard film velocity for 35 mm. film is 90 feet per minute or 18 inches per second. Film has 16 frames to the foot, each frame having four sprocket holes so that this velocity corresponds to 24 frames or 96 sprocket holes per second. In the case of a picture, an intermittent motion device in both camera and projector causes each frame to remain stationary for a portion of its twenty-fourth of a second, and then to be pulled down to allow the next frame to take its place. Sound film, on the other hand, moves steadily past the recording or reproducing optical center, so that each three-fourths of an inch (four sprocket holes) of sound track corresponds to the accompanying picture. Inasmuch as eye and ear fail to recognize inaccuracies smaller than perhaps half a frame, the illusion is properly maintained.

In newsreel work, where speed and simplicity of operation are imperative, and where a relatively small amount of complicated editing is done, the sound record and the picture are both made on the same negative film with an offset of several frames between them to allow for the physical separation of picture and sound optical systems in the sound camera.

Two methods are used in the studio for simultaneously driving sound and picture apparatus: The synchronous system and the interlock system.

1. THE SYNCHRONOUS SYSTEM

In the studio the universal practice is to make the picture and sound on separate films which run through separate cameras and recording machines. The practical advantages are many. Negative film emulsion and processing for the best results are not the same for picture as for sound. Greater flexibility of operation of camera and recorder, increased

ability to combine and delete portions of either picture or sound, and generally better possibilities of editing are important considerations.

The use of separate films presents a driving problem which is solved by various means, depending upon the specific conditions. Camera and recorder must be geared together in some manner to maintain synchronism. As mechanical connection is not feasible, electrical means are employed. These are of two general types—synchronous and interlock, with various modifications and refinements in each.

The synchronous drive, as its name implies, employs synchronous motors to drive both camera and recorder, the power being derived from the same source. The source is usually the commercial three-phase alternating-current system or its equivalent so that frequency uniformity is excellent. Synchronous motors which have no hunting characteristics are used. The result is absolute uniformity of motor speed and constant synchronism between sound and picture. Many frequency combinations are used. For example, one of the major studios having a 50-cycle commercial supply uses a 50—60-cycle motor generator to provide 60-cycle power to drive the recorder, which happens to be designed to operate at 1,200 R.P.M., a synchronous speed obtainable from 60 cycles. At the same time, a 50—48-cycle frequency changer is used to provide 48-cycle power to operate 1,440 R.P.M. synchronous motors to drive the cameras. As both of these new frequencies are generated synchronously with the original supply, they are synchronous with each other. Other frequency combinations are used for particular purposes.

An explanation of the choice of 1,440 R.P.M. drive from 48 cycles may be of interest. The camera shutter operates 24 times per second or 1,440 times per minute. Its main shutter shaft consequently operates at a speed of 1,440 R.P.M. A motor driving it at this speed requires no extra gearing to adapt it to the camera. This not only reduces power requirements to some extent, but is extremely valuable in minimizing extraneous noise on the set, which might be occasioned by such gears. Moreover, the 1,440 R.P.M. speed was found to be a good compromise for motor size. The camera silencing bungalow or "blimp" must, of course, be made large enough to completely enclose the camera motor. Hence, size and weight penalties are incurred which are out of proportion to the motor itself, if the latter is made too large. Incidentally, camera motors may be greatly over-rated in their temperature rise, as the greatest possible length of time they may operate continuously is for the length of one one-thousand-foot roll of film—nominally about eleven minutes. There is always a period for cooling while the camera is being reloaded. In practice, a camera

motor is seldom called upon to operate continuously for the full eleven minutes.

With separate film, and with the synchronous system, a means must be provided to properly align the picture and sound track for subsequent running—"start marks" is the term employed for this means. In the simple case, the "hand-clap" serves this purpose. After both machines have attained synchronous speed an operator claps his hands in front of camera and microphone. The camera photographs his hands coming together and the sound film shows the small, easily identified striation group which was made by the sound of the clap. These are satisfactory start marks. In practice, start mark systems run all the way from the simple hand clap to fairly complicated automatic systems.

2. THE INTERLOCK SYSTEM

Another commonly used driving system employs the Selsyn type of motor and is known as the interlock system. This is virtually an electrical gear system, whereby all of the motors connected together on several separated units will start together, come up to speed at the same rate, and continue to run at identical speeds. The master gear or distributor is driven from a synchronous or other constant source so that its speed is accurately maintained. As it is interlocked with the other motor units, uniformity of their speed is assured.

As is the case of the synchronous drive, there are many combinations of distributor and motor speeds in use to meet various requirements. A common case is a 2,400 R.P.M. interlock motor for camera drive—geared down to 1,440 R.P.M. at the camera—and a 1,200 R.P.M. interlock drive for the recorder. Both motor speeds are obtainable from the same distributor by selection of suitable combinations of numbers of poles in the distributor and motor windings. Another arrangement is the use of a single synchronous motor which drives two distributors, one directly at 1,200 R.P.M. for the recording machines, and the other, through gears, at 1,440 R.P.M. which is used to drive a number of re-recording machines at 720 R.P.M.

A very useful application of the interlock system is in projection background process work. Here, camera shutter and process projection shutter must be accurately interlocked so that the camera exposure bears a definite relation to the time of projection of each frame. This relation is maintained by an interlock motor system which is again tied to the recording machine either through the same interlock distributor or through some form of synchronous drive.

Interlock drive is very useful for re-recording work. Here several film records must be reproduced simultaneously and accurate syn-

chronous relations must be maintained between these sound tracks, the picture which accompanies them, and the recording machine which records their combined output. It will readily be seen that the interlock drive provides a reliable and facile means of bringing all the machines up to speed together and of maintaining their speed relation throughout the duration of the record.

An interesting variation of the interlock system has proven very valuable for location work where light weight apparatus is required. The motors used for recorder and camera are virtually a specialized form of inverted rotary converter. Battery power drives the motors, while three-phase alternating-current connections to their windings produce a low voltage alternating current which serves to interlock the several units together. The battery sources may be separated—one to each motor unit, in which case one unit is used as the master speed control and the others are individually adjusted to it. A more satisfactory arrangement is a central battery supply, usually at the recording machine, which furnishes direct current to all of the motors. Field rheostats for each motor are located in one place and all speeds are adjusted to maintain interlock at the estimated camera and recorder loads. Final speed control and provision for starting in the interlock condition are maintained by the central operator—the recorder operator. This method of drive is becoming more and more popular for portable use to the extent that it will probably be the best ultimate means to be found for this purpose.

3. THEATRE MOTOR DRIVES

Theatre motor drives are in general of two types. One type employs a motor of peculiar electrical design such that its speed is maintained constant within narrow limits by means of electrical speed control networks. This driving means is widely used but it appears to be in the process of being superseded by another simpler method.

The newer apparatus being installed uses, for alternating-current installations, a single-phase induction motor having unusually small slip. The slip of this motor is determined for the particular projection machine load and the connecting gearing is so designed that with normal motor slip the projector drive very closely approximates synchronous speed. In other words, if synchronous speed requires a motor speed of 3,000 R.P.M. and the motor slip reduces the motor speed to say 2,900 R.P.M., then the motor will be connected to the projector through gearing which has a speed increase of the ratio of 3,000 to 2,900 built into it.

This particular arrangement has been worked out to accommodate the large majority of houses where three-phase supply is not readily

available, and is quite satisfactory. In direct-current installations, the usual practice is to provide a motor generator set to generate an alternating-current supply, and then to use the previously mentioned single-phase induction motor.

4. UNIFORMITY OF FILM MOTION

Motor drive for the film-running device is but one part of the film motion problem and is perhaps the simpler part. Given a motor running at uniform speed, it is required to couple this motor to a mechanism which will run the film itself at constant speed. As the ear is very sensitive to the results of speed variation in either the recording or the reproducing process, the uniformity of the film motion must be essentially perfect. A brief mention of some of the limitations and some of the devices used may be of value.

To begin with, the celluloid which forms the base for the film is a material which is susceptible to fairly wide variations in length due to changes in its moisture content. Its shrinkage is a known element at the various stages of the several processes, but unfortunately the limits within which shrinkage may be maintained are of relatively considerable size. Because of shrinkage variations it is imperative that the sprocket holes be used at some stage of the film pulling process to maintain synchronous speed and to control the loops which form a part of any film driving and pulling mechanism. For example, a one per cent shrinkage may readily occur and this means a change of 10 feet in length in a 1,000-foot reel. Picture mechanism definitely requires sprocket drive to ensure that the four sprocket holes corresponding to each frame are constantly pulled down by the intermittent mechanism to keep the right frame relation. In projection there is a constant 20-frame difference between picture and sound—the picture and sound apertures are just that distance apart—and this difference must be maintained throughout the whole reel to keep in synchronism.

The conventional sound film drive uses the sprocket holes to do the main film pulling and to maintain isolating loop relations, and then departs from the sprocket pull at or near the sound aperture. A roller drum coupled to a flywheel, or some similar device, maintains a high degree of approximation to absolutely uniform motion at the actual aperture. The transition between sprocket hole pull with shrinkage variation and the continuous motion at the drum, occurs in the isolating loops before and after the drum.

The design and operation problems, then, are those of insuring perfectly uniform drum action, of making the film move absolutely with the drum at the instant the former crosses the optical center, and of iso-

lating the film on the drum from any reflected sprocket tooth action from the sprockets immediately ahead of or behind the drum. The failures in these respects are noticeable in two general types of speed variation or flutter. Slow variations in film speed occasioned by motor speed changes, or more often by drum speed irregularities, appear as "wow-wows," to use the vernacular of the trade. High speed variations, which generally bear a definite relation to the 96 sprocket hole per second speed of the film, result in a broken up and harsh character in the record. These variations appearing in either the recording or reproducing process, are bad. If they occur in both processes, it is apparent that the reproduced sound will be very unpleasantly distorted.

In the early days of sound, flutter of the various kinds was a disturbing factor which was apt to occur at any moment and could not be controlled. A great deal of effort has been devoted to remedies, and present day recorder and reproducer design has minimized the possibility of these troubles occurring. Studio recorders have been commercially free from flutter for a long time. Unfortunately, the same may not be said for the many theatre applications of the older apparatus designs which are still in use. It is the hope that speed variation difficulties will be reduced to the point of being inappreciable as the apparatus now available is gradually used to replace the older installations.

Chapter IX

FILM PROCESSING

By L. E. CLARK

In sound recording, film, while acting in the nature of a delay circuit or storage device for the electrical energy of the signal, also controls the relative instantaneous amount of light reaching the photocell in the reproducing mechanism.

As previously explained, the same principles apply to the variable area and variable density methods of recording, but in the two systems different characteristics of the film are employed. It is the purpose of this chapter to explain briefly the chemistry of film processing and the application of the science of sensitometry to the commercial production of high quality sound recording.

Sensitometry, as its name implies, is very largely concerned with the measurement of sensitivity. In its modern applications, however, it embraces a much wider field and may be more completely defined as the qualitative measurement of the response of photographic material when exposed to light or other forms of radiant energy.

Motion picture sound recording on film must go through the following steps before the production is completed and ready to be shown in the theatre: The picture and sound tracks are each recorded separately on different negatives, which are then developed, and the positive, containing both the picture and the sound track, is then printed from these two negatives—this positive, called a composite print, is then developed for use in the theatre.

1. FUNDAMENTAL MEASUREMENTS

The fundamental film measurement is "transmission"; that is, the percentage of incident light which is transmitted through the film, and is expressed mathematically by

$$T \text{ (Transmission)} = 100 \times \frac{L_1}{L_0} \quad \text{(per cent)} \qquad (20)$$

where

L_0 = total incident light falling on the film and
L_1 = total amount of light passing through the film.

It can be seen that transmission, when expressed in per cent, varies from a maximum value of 100 to a minimum value of zero.

Although the measurement of this ratio is the fundamental measurement made on film, it is not usually expressed in this way but in terms of the "opacity," which is the reciprocal of the transmission, and is expressed mathematically as:

$$D \text{ (Density)} = \log O \text{ (Opacity)} = \log \frac{I}{T} \qquad (21)$$

From this it may be seen that as the transmission, expressed as a decimal, varies from one to zero, the opacity will vary from one to infinity and the density from zero to infinity.

Density is the common term used to express the measure of transmission or opacity of film.

2. DENSITY MEASURING DEVICES

Several instruments are available for measuring the ratio of the incident to transmitted light, expressed in terms of the logarithm of opacity or density. These instruments are called densitometers and are divided into one of two types, either optical or electrical, depending upon the manner in which the measurement of density is made.

(a) Electrical Densitometers

This type contains a light-sensitive cell and measures first the light without film in its path and then the light after passing through the film. The ratio of the two values is a measure of the transmission.

This type is designed to employ either a steady or an interrupted source of light depending upon the purpose for which it is intended.

(b) Optical Densitometers

This type is the most commonly used and is either polarizing or non-polarizing; the polarizing type consisting essentially of a polarizing photometer head with a suitable source of light to obtain the necessary illumination, and the non-polarizing type containing no polarizing prism.

Figure 72 shows the main constructional details of the polarizing type densitometer. S and S' are apertures through which the two light beams enter, and after being polar-

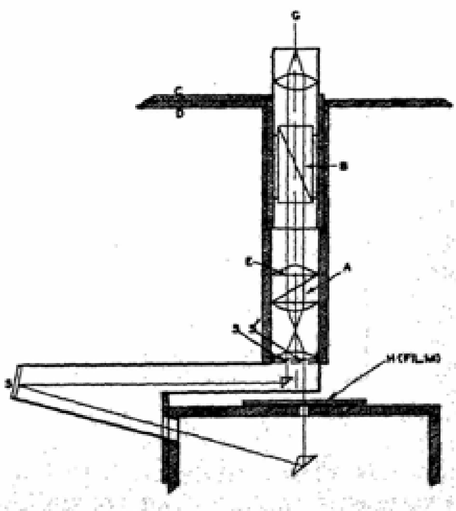


Figure 72 — Details of a polarizing type densitometer.

ized by means of the Prism A, result in two light beams whose planes of polarization are perpendicular to one another. Each beam is then directed through the analyzing prism, B, which is held in a rotating support inside the frame of the photometer head. The member D of the frame and the member C of the analyzing prism support, form a scale which serves as a means of measuring the rotation of the prism B with respect to the fixed prism A. The light field, as seen by the eye from position G, has been divided along a diameter by an image of the apex of the bi-prism E, and thus consists of two semi-circular light fields whose relative intensity is determined by the amount of rotation of B. If both light beams are uninterrupted, prism B would be rotated 45° from prism A to secure light fields of equal intensity.

If the material, the transmission of which is to be measured, is placed at H, then the beam passing through this side of the photometer head will be changed in intensity and further rotation of the prism B will be necessary before the two light fields again have the same intensity. From the new position of the scale C, the density may be directly computed.

The non-polarizing type is illustrated in Figure 73, and consists of a movable source of illumination A; two diffusing screens, B and C; four reflecting mirrors D, E, F and G; and a lens H. The photometer head is shown in detail in Figure 73-B.

The illumination on the screens B and C varies as the light source moves, and is inversely proportional to the square of the distance from

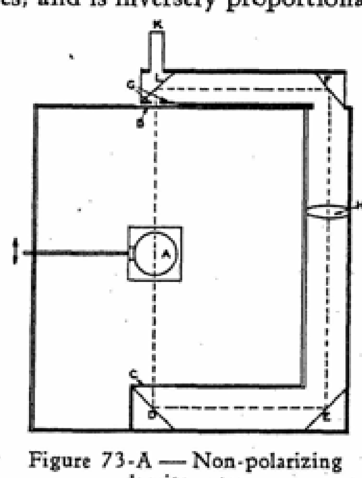


Figure 73-A — Non-polarizing densitometer.

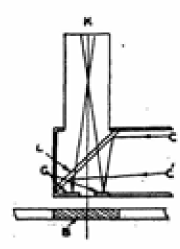


Figure 73-B — Details of photometer head used with a non-polarizing densitometer.

source to screen. As A moves toward B the intensity on B increases, according to the inverse square law, just as it decreases on C. Images of screens B and C are formed in the photometer head as shown. L is a

pane of clear glass placed at an angle of 45°, which serves to both reflect and transmit the light rays. With no absorbing material placed in the instrument, the light source A would be at some zero position depending upon the losses of light intensity in the various mirrors and lenses.

The material to be measured is then placed between the screen B and the reflecting mirror G, so that the beam passing through the screen B also passes through this material, and the eye sees a light field consisting of that part which passes through the material to be measured, surrounded by a field which comes directly from the light source. By moving A to a position where these fields are equal, the inverse square law may be applied to compute the density. A suitable scale attached to A permits the reading of these values directly.

There are, of course, many different kinds of densitometers available of the above types, each most suitable for particular uses.

3. DENSITY

Motion picture film is composed of a celluloid base approximately five-thousandths of an inch thick, carrying a thin coating of emulsion, the principal contents of which are gelatin, silver bromide and silver iodine. When the film is exposed and developed, the silver bromide, under the action of the light rays and developing solution, is converted into a layer of opaque silver, rendering film less capable of transmitting light than before exposure and development.

The maximum density obtainable depends upon the amount of silver present in the emulsion. In the case of sound positive this maximum density is usually from 3.5 to 4, but densities of this value are never required in commercial work.

4. DEVELOPMENT

If a piece of film is exposed in the light for a short period of time there will be no discernible difference in its appearance until it is placed in a proper developing solution where it will react with the solution and metallic silver will be formed on the celluloid base. To obtain this deposit of metallic silver the solution must contain a reducing agent, which is usually some organic substance which reacts with the silver bromide to form metallic silver. The reducing agent is used up in the process—and is usually a coal tar derivative broken down in certain definite ways (trade names Metol, Elon, etc.).

If the developing solution contained nothing but the reducing agent the process would be slow, as considerable time would be required for the solution to penetrate even the thin layer of gelatin and to react with the silver. For this reason the solution also contains an alkali, called an accelerator, as its purpose is to speed up the reaction. This alkali may

be sodium carbonate, caustic soda, etc. The speed of the action desired depends upon the strength of the accelerator used, and determines the graininess of the developed film.

The next substance needed is a preservative for the developer, as the solution absorbs oxygen from the air and spoils in a very short time if left standing in contact with air. This preservative is usually sodium sulphate.

One other substance, usually potassium bromide, is added which acts as a retarder or restrainer, the quantity necessary depending upon the type developer required for the particular type of film being processed.

The active agent in this developing solution is the reducing agent. All the others are added to either prolong the life of the solution or to determine the speed of its action as a developer of the film. As previously brought out, all of the silver bromide is not acted upon and it is necessary to remove this residue after developing, so the film is washed thoroughly in water and then put through a so-called fixing solution after it leaves the developing solution. This fixing solution contains "hypo" (sodium thiosulphate), its only purpose being to dissolve the silver bromide left in the gelatin.

In some cases a hardening agent is added at this point, the function of which is to harden the gelatin and to protect it against scratches, cuts, etc. After again being washed and dried, the film theoretically consists of clear cellulose covered by particles of silver, the amount of which determines the density.

5. SENSITOMETERS

Obviously, the type of solution, type of film, and amount of exposure for a given result have not yet been determined, and it can be seen that it would be of great value to be able to predetermine these points and to control results after these variables are decided upon. This is accomplished by the use of sensitometry in film processing.

In exposing, developing, and reproducing a film record, the type of film, the developing solution, the exposure, the time of development, printing, etc., are all variables, and unless carefully controlled are apt to give confusing results.

The effect of exposure on a given film which is to be developed in a standard solution for different periods of time will first be considered.

Instruments for exposing photographic material for a series of graduating and precisely known periods are known as sensitometers, and consist of a light source of known intensity and a means of produc-

ing different exposures of known relative amounts. The light source affects the results both by its intensity and spectral composition.

The exposure modulator may be one of several types, as the exposure may be varied by controlling either the intensity of light or time of exposure. One type uses a rotating sector-wheel allowing the light to pass for a smaller period of time as the distance from the axis increases; another type uses tablets of different density between the light and photographic material, thus changing in a regular manner the intensity of exposure. There are a great many varieties of these instruments, but the general operation is much the same, variations being only in the application of the following principle:

A strip of film, divided into sections, is placed in the field of illumination of a dim light of proper value. At periodic intervals of one, two, four, eight, etc., seconds, the sections are covered by moving an opaque object parallel to the length of the film. At the end of the exposure time each section has been exposed a different length of time, namely: One, two, four, eight, etc., seconds depending upon the number of times this process is repeated.

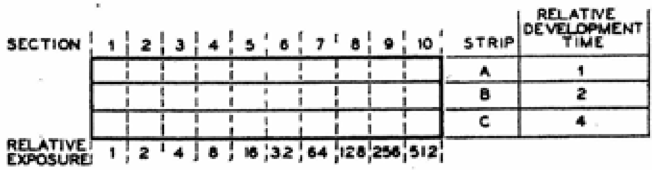


Figure 74.

This film is then cut into three lateral strips and the first developed for two minutes, the second for four minutes, and the third for eight minutes, to give a development time ratio of 1, 2, and 4. After each strip is developed, the density of each section of each strip is then read on a densitometer. Figure 74 shows the three strips of film and the relative exposure and development time for each section.

6. "H AND D" CURVES

Considering each strip separately, the densities can now be plotted with respect to the exposure, and give a curve of the form shown in Figure 75, when the exposure is plotted logarithmically. The usual method is to plot density and log of the exposure as shown by the middle horizontal scale. These curves are known as "H and D" or "density-log E" curves.

It will be noticed that each curve consists of three distinct parts, a toe, a straight line portion, and a shoulder. The toe, region 1-2, is called the region of under-exposure, as changes in exposure have little effect on density. The region 3-4, the shoulder of the curve, is the region of over-exposure, as again changes in exposure cause only small changes in density. The straight-line portion of the curve is called the region of correct exposure as in this region the density is directly proportional to the logarithm of the exposure.

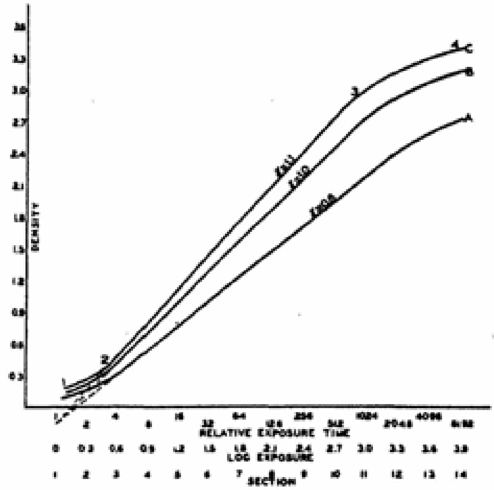


Figure 75 — Curves showing the relation between density and exposure of three strips of the same film at three different development times.

7. GAMMA

From these curves it is possible to predict results from a given film developed for a certain length of time in a certain developing solution. With these factors constant, and considering only the straight-line portion of the H and D curve, the density varies with the exposure in a regular manner. This variation, or slope of the curve, is known as the gamma (γ) of this particular curve, that is, the tangent of the angle between the H and D curve and the exposure axis is called gamma. Thus for the three curves in Figure 75, we have γ_1, γ_2, and γ_3 as shown. Gamma is a measure of the effect of relative exposures on any particular film which is left in a certain developing solution for a given length of time.

If we consider Figure 75 as H and D curves of negative film, it is now necessary to make prints from these negatives which will give another series of H and D curves, depending upon the solution used and length of developing time. Now, if a series of prints were made, each being developed a different length of time, the density of each section of each print would again vary with exposure and an H and D curve for each print could be plotted. It can thus be seen that these conditions, unless controlled, would give rise to varied and unpredictable results.

The relation between the length of developing time and the gamma is controlled by means of curves drawn for each type of film developed

in the same solution. Typical curves showing this relation are shown in Figure 76, Curve (1).

8. FOG

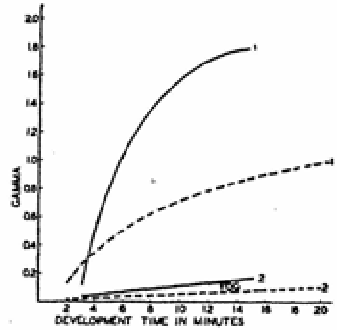


Figure 76 — Curves showing the relation between gamma and development time.

Film subjected to no exposure, but developed for any length of time, would theoretically have a density of zero. Actually, however, it has a definite density depending upon the photographic material used and the development time. As a consequence, the theoretically transparent portions of sound track actually have a certain density, which is called "fog." This is illustrated in Figure 76, where the curves marked (2) are the fog curves of the materials whose gamma-development time curves are also given in the figure.

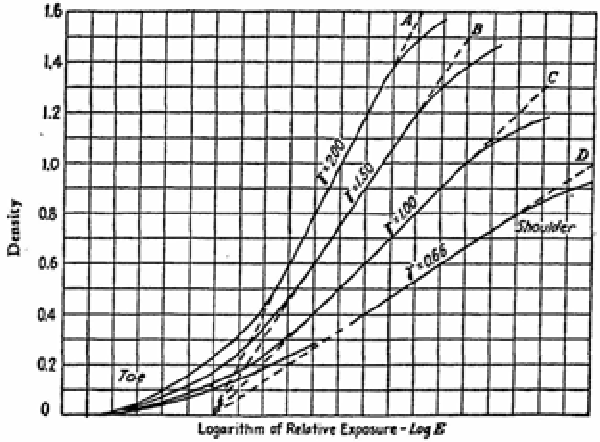
9. PHOTOGRAPHIC REQUIREMENTS OF VARIABLE AREA AND VARIABLE DENSITY RECORDS

It will be remembered that the distinction between variable area and variable density recordings is that in the former both factors of exposure, intensity and time of illumination, are constant at any point within the exposed area, while in the latter method the intensity factor is constant while the time factor varies (due to the changing width of the light-valve slit). This difference in the two types of records necessitates the use of different film characteristics for each type, even though both systems function under the same criterion, namely, that the sound record positive, printed from the negative, shall carry a distribution of density such that the intensity of the light reaching the photocell in the reproducer is directly proportional to the instantaneous pressure on the microphone diaphragm.

(a) Variable Density Requirements

From a photographic point of view, H and D curves are used in the study of emulsion characteristics and the effect of development. However, in sound track processing, where the relation between the input to the film and the output from the film are of primary importance, the overall result is considered. The basis of this relationship is still the H and D curves of the negative and positive films, but as the input of the film is measured by the negative exposure and as the positive trans-

mission is proportional to the output from the film, the most logical method of studying this relationship is between the negative exposure and the positive transmission.



—From "Recording Sound for Motion Pictures," Academy of Motion Picture Arts and Sciences (McGraw-Hill Book Co., Inc.).

Figure 77 — Typical H and D curves.

Typical H and D curves, with four different values of gamma, are given in Figure 77. The equation of the *straight-line portion* of these curves is:

$$D = \log \frac{1}{T} = \gamma \ (\log E - \log i) \qquad (22)$$

where $\log i =$ the value of $\log E$ at the intersection of the straight line portion extended, with the $\log E$ axis

$i =$ inertia of the film

Since the quantity of $\log i$ of equation (22) is a constant for a given case, then

$$T \infty E^{-\gamma} \qquad (23)$$

for both positive and negative film. Then if N and P as subscripts denote negative and positive, respectively,

$$T_N \infty E_N^{-\gamma_N}$$

and $\qquad\qquad (24)$

$$T_P \infty E_P^{-\gamma_P}$$

In printing, the exposure of the positive is controlled by the transmission of the negative so that

$$E_P \propto T_N \tag{25}$$

From this

$$T_P \propto E_N^{\gamma_N \gamma_P} = K\, E_N^{\gamma_N \gamma_P} \tag{26}$$

where K is a constant related principally to the inertia and the amount of light used in printing.

From equation (26) it can be seen that when the product, $\gamma_N \gamma_P$, equals unity, the positive transmission is directly proportional to the negative exposure. When this condition is fulfilled, provided the straight-line portion of the H and D curves are not exceeded, the true transmission ratio for the film is realized.

In practical applications in sound recording, the gamma of the exposure device (light valve), the gamma of the printing machine, and the gamma of the photo-electric cell, as well as the gamma product of the two films, affect the transmission ratio. For this reason, in practice the overall gamma varies from unity, depending upon laboratory conditions and recording equipment.

Typical examples of the various gammas are: gamma of the light valve, 0.96; gamma of the negative, 0.40; gamma of the printing machine, 0.90; gamma of the positive film, 2.00; and gamma of the reproducing equipment, 1.45.

These typical values would lead to an overall gamma of about one.

Practical experience has led to the adoption of the following standards for certain factors:

1. A fine-grained, long-scale emulsion, such as standard positive stock, seems suitable for both negative and positive sound records.

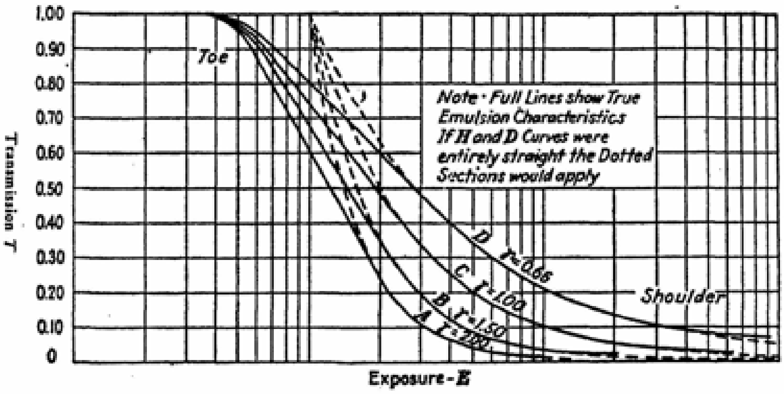
2. Positive development, and hence positive gamma are determined by the picture requirements of the positive. The reason for this is obvious, since both picture and sound track must be developed simultaneously. Commercial positive gamma is usually in the neighborhood of 2.00.

3. Negative development, that is, negative gamma, must be controlled to meet the overall requirements with fixed positive conditions. Negative gamma accordingly ranges from about 0.38 to 0.43.

4. The average positive transmission and the total range of positive transmission modulation must be such that neither excessive light nor excessive amplification is required in theatre reproduction. Accordingly, the average or unmodulated transmission of the positive is

about 0.30. (All statements of density or transmission are made with clear film as a standard, i.e., its density is taken as 0 and its transmission as 1.00.)

5. Negative exposure is theoretically so fixed that the maximum possible value is slightly below the shoulder of the H and D curve. This exposure produces an unmodulated or average negative transmission of about 0.25, depending upon the development chosen and upon practical departures from the theoretical exposure range.



—From "Recording Sound for Motion Pictures," Academy of Motion Picture Arts and Sciences (McGraw-Hill Book Co., Inc.).

Figure 78 — Typical transmission-exposure curves.

To determine the effect of a departure from an overall gamma of unity, and the practical limitations of the above, the H and D curves are drawn in a different form as follows:

The relations between density and log E as given in Figure 77, are replotted as shown in Figure 78, with transmission and exposure as the two new relative values. This may be done by use of the formula

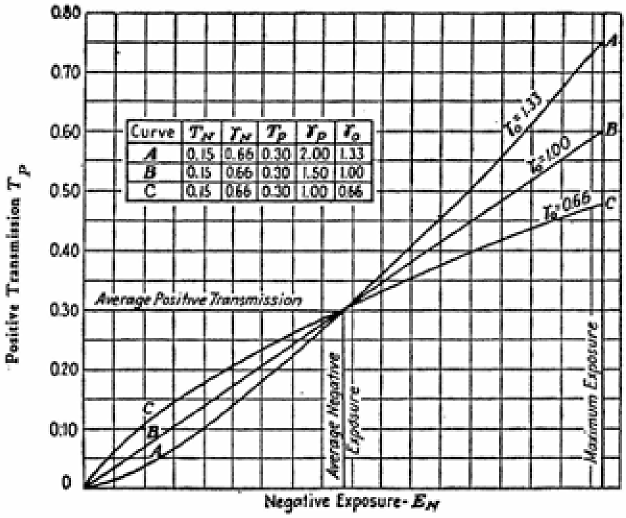
$$D = \log O = \log \frac{1}{T}$$

In Figure 78, it can be seen that the straight portions of the curves of Figure 77 tend to become curved and the curved portions straight, and that the relation between exposure and transmission is not constant.

From Figure 78, and assuming the entire H and D curve to be straight, the curves of Figure 79 can be constructed, using the values of positive and negative transmission, and the combination of gammas, as listed in the figure. These curves show that there is a definite departure from proportionality when the overall gamma is not equal to unity, even when the H and D curve is assumed to be entirely straight.

It is also evident, however, that small departures from unity can be tolerated without a severe change in quality.

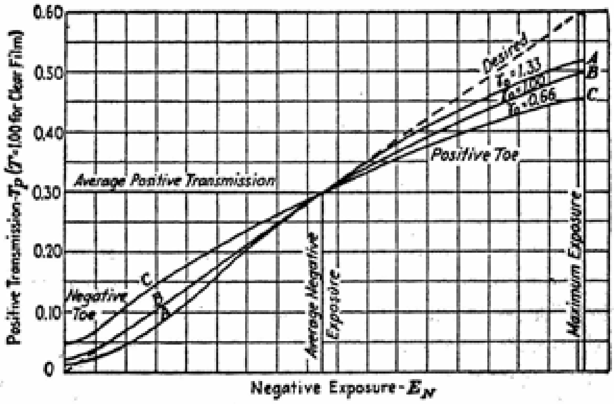
Now if the curves of Figure 79 are redrawn but with true emul-



—*From "Recording Sound for Motion Pictures," Academy of Motion Picture Arts and Sciences (McGraw-Hill Book Co., Inc.).*

Figure 79 — Curves derived from data in Figure 78.

sion characteristics, that is, considering the curved toe and shoulder portion of the H and D curves, Figure 80 results. This figure shows



—*From "Recording Sound for Motion Pictures," Academy of Motion Picture Arts and Sciences (McGraw-Hill Book Co., Inc.).*

Figure 80 — Overall relation between negative exposure and positive transmission for various overall gammas showing the effect of toe sections.

that there is a small departure from the theoretical curve in the region of low negative exposures—the result of the negative toe—and large

departures at the high-exposure end—the result of the positive toe.

Considering negative exposure in terms of modulation about the average value, with unity overall gamma, fairly high modulation without distortion is permissible on the low-exposure end, whereas on the high-exposure end distortion commences at practically zero modulation. For overall gammas higher or lower than unity, the extent of this distortion varies. It is a fact, not immediately evident from the figure, that the relative distortion becomes less as the average positive transmission is made less.

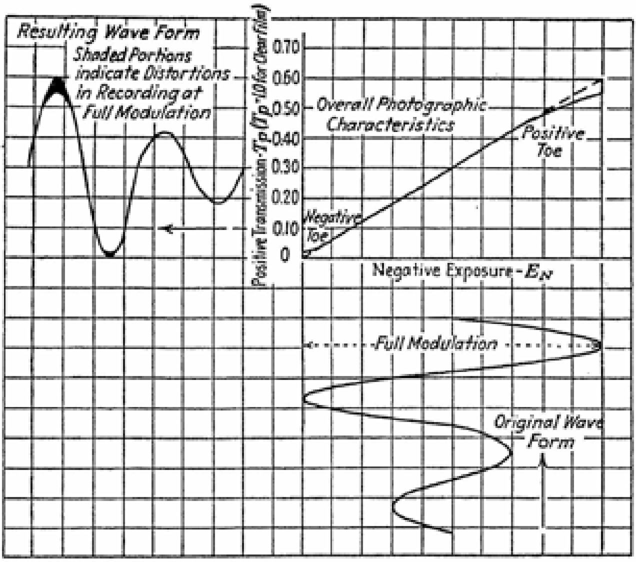
The effect of these toe sections is serious and difficult to correct. Indeed, complete correction is impossible. For low-modulation values they may be accepted, but as commercial records require maximum modulation to keep above surface noise and to maintain proper volume levels, it is often necessary to work right up to the limit of acceptable distortion—and sometimes, perhaps, beyond. The negative toe distortion is under better control than that of the positive, as the amount of use which it receives is a direct function of the control of the exposure device. The positive toe, on the other hand, is the principal source of distortion, and at the same time, is more subject to harmful variation. It seems, unfortunately, that almost every factor in the positive development process is such that it tends to exaggerate rather than to help the positive toe condition. Moreover, production negative can always be more carefully handled than the released positive, as the latter amounts to many more feet than the former, and its production must be undertaken at a reasonable cost.

Referring again to the curves of Figures 79 and 80, there is available an approximate solution to the toe distortion. Experimentally it has been found to be practicable to make relative adjustments of gamma and transmission values, principally the former, to offset one type of distortion by means of another. This fact was not contemplated in the original discussions of the theory of this method of recording.

That such adjustment is successful may be seen from the curve of Figure 81, which is representative of normal results obtained in practice. The major producing studios and the equipment companies employing the variable density method of recording, and working independently or in conjunction with one another, have arrived at this compromise and in every case the results are remarkably close.

It has been arrived at by carefully balancing the several factors involved in the annual production of several hundred million feet of annual release print and represents the best practicable compromise between the commercial and theoretical aspects. As a matter of interest, the same figure shows the effects of film-characteristic distortion upon the repro-

duction of a sine-wave negative exposure. The reproduction is excellent except at high modulation.



—From "Recording Sound for Motion Pictures," Academy of Motion Picture Arts and Sciences (McGraw-Hill Book Co., Inc.).

Figure 81 — Results in practice from relative adjustments of gammas and transmission values.

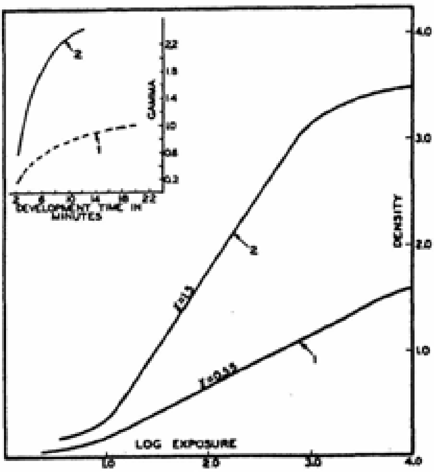
The effect on the mean transmission or no-signal density of the print on noise reduction is considered in the chapter on Noise Reduction.

(b) Variable Area Requirements

With this type of recording, the general requirements of the photographic material are different from those of variable density as the track consists of transparent and opaque areas rather than striations of varying density. For variable area, high contrast between the two regions is desirable, that is, the opaque region should be as dense as possible in contrast to the opposing region, which should be as transparent as possible, with the dividing line between the two very sharp. This calls for a development just the reverse of the variable density method.

This high contrast requirement needs an H and D curve of much sharper slope, that is, with a higher gamma, although the absolute value of this gamma is relatively unimportant. This steeper slope is

necessary in order to obtain the density contrast and image sharpness which are required for this type of track. Also, the gamma-development time curve is much shorter and higher than for variable density recording. Figure 82 gives a comparison of H and D curves and gamma-development time curves for the two types of recording.



Curves (1) for variable density.
Curves (2) for variable area.

Figure 82 — H and D curves with associated gamma-development time curves.

As previously mentioned, the dividing line between the transparent and opaque regions should be sharp, and theoretically at this boundary line there is a region of light and no light, but actually there is always stray light spilling over into the transparent region and resulting in regions of light and less light, which of course reduces the contrast.

This builds up, especially in the case of a high-frequency wave, a sort of halo which follows the wave form and produces a region of graduated density along the wave. This tends to fill up the valleys of the waves and to broaden the peaks in a manner similar to that described under "Slit-Effect" in Chapter III, and distortion and loss of high-frequency results.

To minimize this effect the recording light beam should be focused as sharply as possible.

The new ultra-violet recording, which uses a filter in the optical system to allow only the ultra-violet light to strike the film, has materially reduced this effect by keeping the exposure near the film surface.

The shape of the toe of the H and D curves for both positive and negative are highly important. The transparent portion of the print should have as low a density as possible, in contrast to a density of around 1.5 for the opaque portion. This means that the contrast on the negative must be good, and that the transparent part be of very small density, and, as a consequence, the toe of the H and D curve must be

sharp so that changes in exposure in the lower-density region cause very little change in density.

A comparison between the critical factors governing film characteristics of the two systems shows that for variable density, gamma and density control are most important, while for variable area the shape of the toe of the H and D curve and the contrast are most important.

The effect of noise reduction on variable area processing is discussed in the chapter on "Noise Reduction."

Chapter X

REPRODUCING SYSTEMS

By L. E. CLARK *and* JOHN K. HILLIARD

As brought out previously in this text, release prints for use in the theatre have both picture and sound on the same film, and the reproducing system consists of two more or less independent projectors—the picture projector and the sound projector.

We will take up the picture projector first and then explain the necessity for the two separate projector systems.

Consider first the manner in which the film is exposed and then the manner in which it is projected. The negative enters and leaves the camera at a constant speed under the action of a uniform speed sprocket, but moves through the camera at an intermittent speed under the action of a claw-like arrangement (called the intermittent movement), controlled by a cam which engages the film and brings it to a stop in a position in front of the shutter. The shutter then opens and the film is exposed. The shutter then closes and during this closure the intermittent again engages the film moving that particular frame out of the aperture and bringing the next frame into position in front of the aperture. The variation in speed during the travel of the film through the camera is taken up in two loops formed by the film, one on each side of the intermittent movement.

This entire cycle takes place at high speed, there being 24 complete exposures per second in the camera.

After the film is processed, it must be projected in exactly the same way as exposed, the film entering the picture projector at a continuous speed, passing in front of the projector aperture under the action of an intermittent movement similar to the camera mechanism, and leaving the projector again at constant speed. For obvious reasons the sound track could not be projected in this manner, as one of the primary requisites for high-quality reproduction is a constantly uniform rate of speed of the film through the sound projector.

The sound projector is therefore placed behind (in point of time) the picture projector at sufficient distance to allow the film to be again given a constant rate of speed of movement, and explains the necessity for a separate projector for the sound. This distance is usually about

20 frames and is secured by synchronizing the sound this number of frames ahead of the picture when printing the composite positive from the picture and sound track negatives.

There have been several types of sound reproducing systems in use depending upon the type of record each is designed to reproduce. However, as sound on film is now used exclusively in motion pictures, and as the newest installations are capable of reproducing only the film record, the discussion in this text will be limited to this type.

Such a reproducing system consists essentially of a sound head, an amplifier system and a horn system.

The sound head, whether for the reproduction of variable density or variable area track, is made up of a light source, an optical system, and a photo-electric cell. The light rays from the light source act as a *conveyor* of energy while the sound track acts as a *control* of this energy and determines the amplitude and frequency of current change at the photo-electric cell. From this fluctuating current, an alternating voltage is derived, and after sufficient amplification, is delivered to the horns which reproduce the sound and project it into the auditorium.

1. LIGHT SOURCE AND OPTICAL SYSTEM

There are two general scanning methods, either direct or rear scanning, both consisting of a lamp whose light rays are carried through an optical system and then either through the film and the slit, in that or in the reverse order, and then through another set of lenses to the photo-electric cell.

Direct scanning is illustrated in Figure 87. The light passes through the optical unit and is focused on the scanning slit which projects a beam .084″ × .0013″ onto the film, and the part of the beam transmitted by the film is carried to the photocell by a second system of lenses.

Rear scanning is illustrated in Figure 89. Here the light beam, after passing through the condenser lens (whose focal point is some distance in front of the film), falls directly on the film and the blob of light transmitted is focused by the objective lens upon the scanning slit (.084″ x .0013″). The transmitted light then falls upon the photocell.

2. SLIT WIDTH

The width of the scanning slit has a great effect upon the quality of the reproduced sound, both from the reproduction as well as from the recording standpoint. If the slit is made too wide, the high-frequency response is lowered and distortion is introduced, while if too narrow,

the overall response is lowered and mechanical problems of manufacture are introduced.

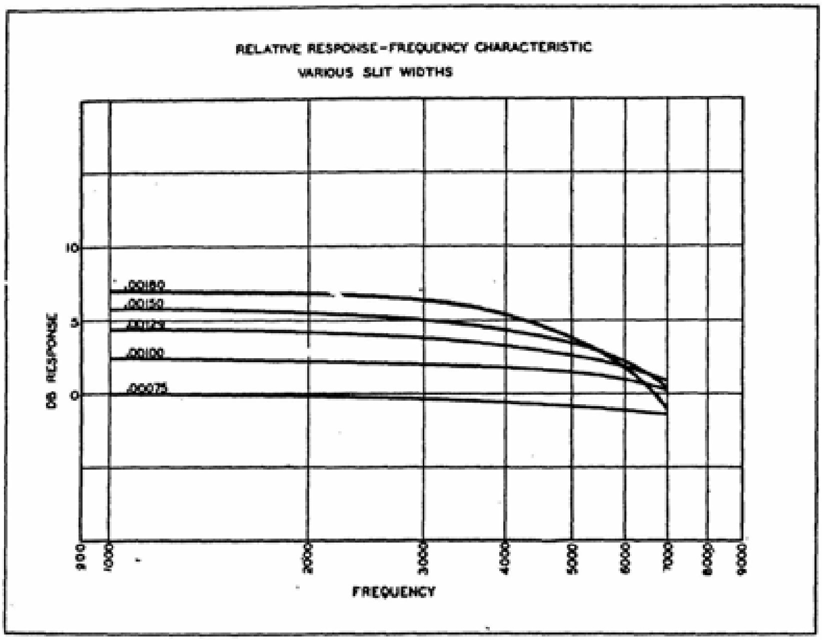


Figure 83.

Figure 83 shows the relative response from slits of different widths. The dimension of 0.0013″ was chosen as standard as giving the best compromise between the limiting factors mentioned above.

3. FILM VELOCITY

Investigation and experimentation have led to a great many improvements in the mechanical set-up of the sound reproducing system. Optical systems have been improved, vibration practically eliminated by better mechanical design, and the efficiency and response of the speaker system have been greatly improved.

Thus it was that, until very recently, one of the chief limitations to high-quality sound reproduction was the lack of constant velocity of the film as it moved past the scanning slit. Any variation in this speed of the film results in frequency modulation, which consists in a change in the relative frequencies of the signals on the film and leads to the reproduction of the signal itself as well as to the introduction of harmonics.

This is a form of distortion called flutter, and reduces the quality of the reproduced sound. However, the new-type sound heads now

available move the film at a very constant velocity and for all practical purposes, if properly adjusted, introduce no flutter.

4. PHOTO-ELECTRIC CELLS

Photo-electric cells fall under the general classification of photo-responsive devices, which are instruments responsive to radiant energy.

The term photo-electric cell, in its broadest sense, includes (1) photo-emissive cells, (2) photo-conducting cells, (3) photo-barrier cells, and (4) photo-voltaic cells. All these are similar in that they depend upon radiant energy for their action, but from each group a different reaction is secured.

In the first group, electrons are emitted from the cathode under the action of radiant energy, and collected upon the anode; in the second group, the resistance of the cell varies with the illumination; in the third group, the cell generates en e.m.f. under the action of the radiant energy; while in the fourth group, an e.m.f. is produced on one of two electrodes of certain materials when immersed in dilute electrolytes with one electrode illuminated.

Only the first group is of interest to us in this text, as this type is the only one used in motion picture sound reproduction, being far more sensitive than any of the other types in certain narrow regions of the spectrum and can be readily coupled to an amplifier system. Consequently, when the term photo-electric cell is used hereafter, this type cell is the one referred to. Photo-electric cells are responsive to radiant energy lying in or near the region of visible light, and act as a transducer of the energy, changing the radiant energy into electrical energy.

(a) Principle of Operation of a Photo-Electric Cell

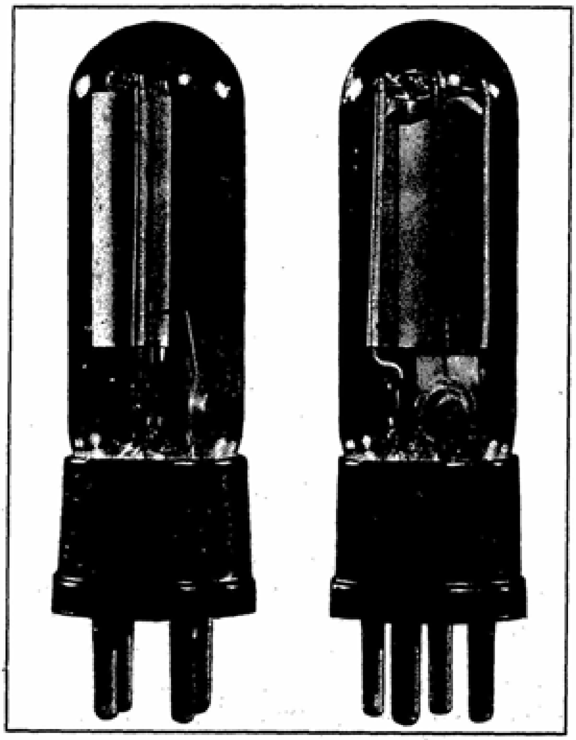
The principle of operation of photo-electric cells depends upon the fact that when radiant energy falls upon certain metals, electrons are caused to be emitted at a rate proportional to the total amount of illumination falling upon the surface. Consequently, if such a metal forms the cathode of an electric circuit, and another conductor is placed close to the metal, this second conductor will act as the anode of the circuit and a current will flow in the circuit. Thus, if properly applied, an instrument is available to change the "frozen" sound waves on the film to electrical waves, through the medium of light and the use of one of these sensitive metals.

The action of a photo-electric cell is similar to that of the vacuum tube in that electrons move through space from the cathode to the anode. However, in the vacuum tube, electrons are emitted from the cathode under the action of heat, while in a photo-electric cell they are emitted

under the action of radiant energy. The action of the vacuum tube is easily explained by the theories of physics and chemistry, but the action of the photo-electric cell is not so clearly depicted, as the same theories explain only in part the results obtained. Only at such time as the characteristics of light are more definitely known will the photo-electric phenomena be clearly understood.

(b) Photo-Electric Materials

Only a few metals emit electrons under the influence of light in any appreciable quantity. The most commonly used are the alkali metals: sodium, potassium, rubidium, and caesium, their sensitiveness decreasing in the order given. Also, each has a particular very narrow region in the spectrum where it is most sensitive, but the sensitiveness of any of these pure metals may be greatly increased by subjecting them to special treatments.



—*Courtesy RCA Manufacturing Co.*

Figure 84 — Two different views of a typical photo-electric cell used in sound reproduction.

There are two general types of photocells—those containing an inert gas and those which are evacuated.

In the latter the current consists totally of the electrons emitted from the cathode from the effect of the light falling upon the sensitive surface, but in the former the cathode-emitted electrons ionize the inert gas, and, as a consequence, when a sufficient voltage is applied the current is amplified within the cell itself.

The amount of this amplification is limited by the ionization voltage, that is, the voltage at which a visible glow discharge takes place, as this effect is self-perpetuating and injurious to the cell.

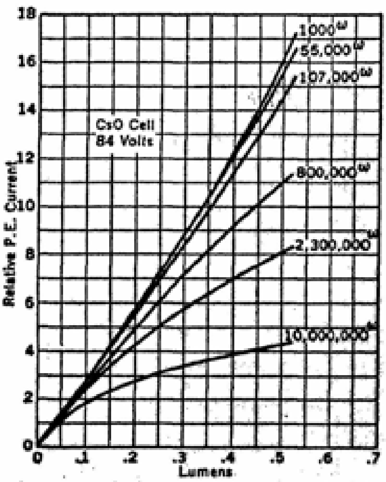
The effect of the illumination on the cell depends not only upon the material of which the cathode is made, but also upon the design of both the cathode and anode, as the photo-electric current depends upon the total illumination at any instant, and if the anode interferes with this illumination the current is reduced.

Two general types of construction are common: those with a central anode and those with a central cathode.

Figure 84 shows a central-anode cell, the type most generally used, with the cathode built in a semi-circular shape around the anode. The envelope is of glass or quartz, depending upon the metal used for the cathode and the source of illumination.

In an ideal cell, the photo-electric current is proportional to the total illumination but in practical applications of this principle the relationship is slightly non-linear because of the charging effect of the glass wall, interference of the anode, reflection effects, and other obscure phenomena not yet fully understood.

It is necessary that the impedance looking into the output of the cell be held to a low value to minimize this effect, but the impedance must also be high enough to insure protection against glow discharge. Figure 85 shows the relative current plotted against illumination at different values of terminating resistances.



—*Reprinted by permission from "Electrical Engineer's Handbook," by Pender & McIlwain, published by John Wiley & Sons, Inc.*

Figure 85 — Current-illumination relation for gas-filled photo-electric cell with various series resistances.

The photo-electric effect of any certain material depends upon the wave length of the light rays to which the cathode is exposed. In motion picture work a cathode of

caesium with a suitable light source is usually used. These cells are rated according to their output in amperes per lumen of steady incident light on the cathode with the type of light specified, or in terms of the slope of the anode current illumination curve.

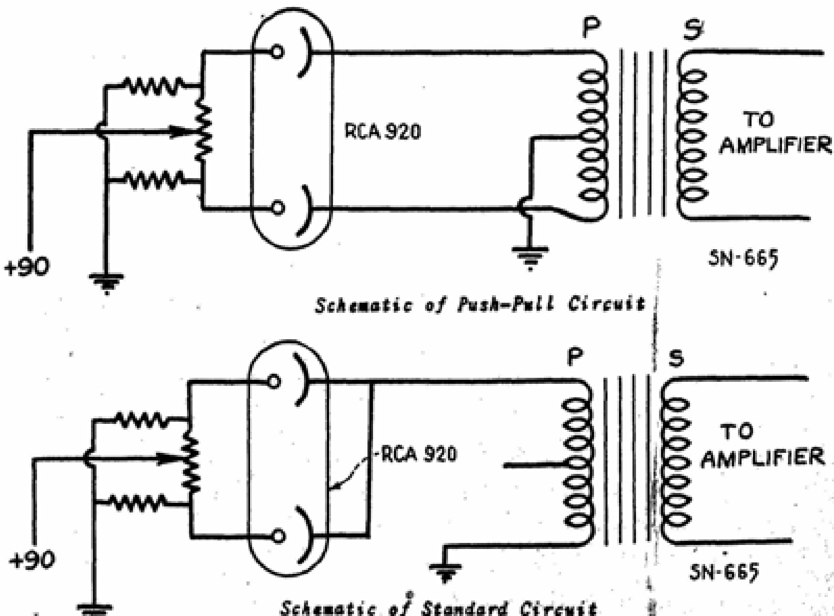
(c) Frequency Response

The frequency response of a commercial gas-filled photo-electric cell drops off in the higher-frequency range of the recording band, due to the presence of the gas in the cell and the capacitive effect between the electrodes and the circuit itself. This last effect becomes the limiting factor and determines the impedances which may be used.

As the alternating-current output of the photo-electric cell is small, the circuit must be shielded or pick-up will result. The capacitive effect of the circuit necessitates close coupling with the first amplifier (P. E. C. amplifier).

5. PUSH-PULL SOUND HEADS WITH ASSOCIATED CIRCUITS

Figure 87 shows the schematic of an RCA MI-1070 sound head, known as a direct scanning reproducer, and consisting of a push-pull photocell, a special lens and prism assembly, together with a push-pull



—*Courtesy RCA Manufacturing Co.*

Figure 86 — Schematics of RCA 920 Photocell for push-pull and single track reproduction.

photocell transformer. This combination permits the reproduction of either single track or push-pull recording.

The 920 photocell contains two anodes and two cathodes. When connected through the selector switch for single track reproduction, the cathodes are hooked up in parallel, and the photocell operates in the same manner as that of the standard cell. For push-pull reproduction, the two cathodes are separated and operate alternately through the photocell transformer. (See Figure 87.)

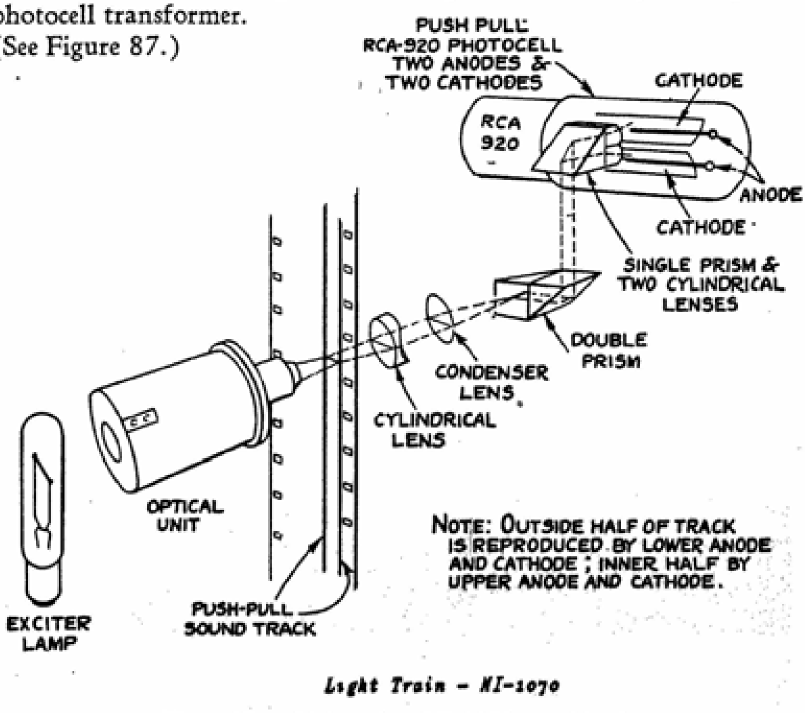


Figure 87 — Light train—MI-1070 sound head.

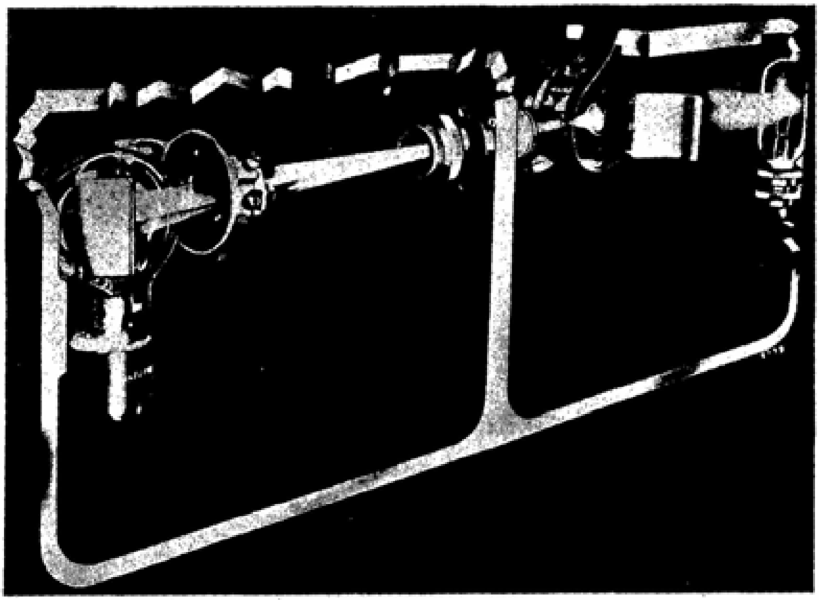
Any single track film may be used to balance this equipment for push-pull recording. When set for push-pull reproduction, the balancing potentiometer is operated to obtain minimum reproduction of the recorded sound at high volume.

Another method which is currently used in studio review room operation is to use alternating current as a source of supply for the exciter lamp and to adjust the balancing potentiometer for minimum output.

6. FLUTTER

To obtain the maximum benefit from increased volume range, flutter, due to the motion of the film at the point of scanning, must be reduced to an absolute minimum.

In the reproduction of sound recording, where a wheel with sprocket teeth is used to move the film, two types of flutter are encountered. The first is due to the irregular motion imparted to the film as it is engaged by the sprocket teeth and the second is due to the not-quite-constant speed of the sprocket wheel itself. For this reason scanners which use sprocket teeth have been eliminated in favor of drum-type equipment which practically eliminates the 96 cycle flutter. The drum also reduces low-frequency flutter when the rotary stabilizer principle is applied.*



—*Courtesy Electrical Research Products, Inc.*

Figure 88 — Western Electric sound head.

Figures 88 and 89 show another type of sound head manufactured by the Western Electric Company. This film scanning system is known as the "rear or indirect projection" type and consists essentially of an exciter lamp, condenser lens-prism assembly and objective lens, a scanning slit behind which is a collimating lens, and a photocell.

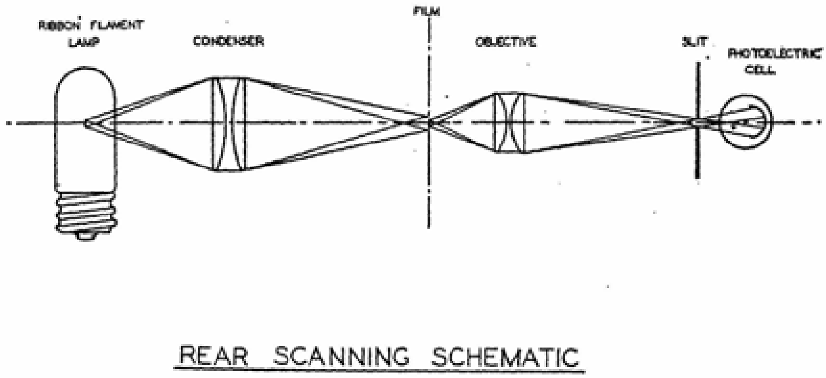
The condenser lens-prism assembly should be so adjusted that it focuses the filament image some distance in front of the film plane on the lamp side, so that the film is illuminated with a blob of light. The objective lens is adjusted to focus the track image sharply on the scanning slit, the width of which is approximately 1.3 mils.

---

* Journal of Society of Motion Picture Engineers, October, 1935, "*Technical Aspects of the High-Fidelity Reproducer,*" E. D. Cook, pages 289-312.

The drum which holds the film in place at the point of scanning is of the rotary stabilizer type similar to that used in the RCA sound head.

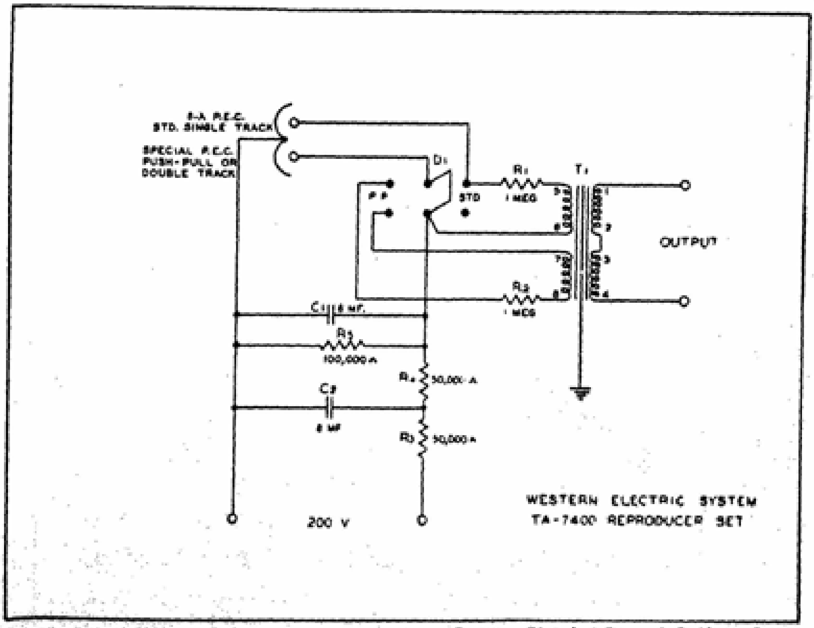
Figure 90 shows the wiring diagram of the TA-7400 reproducer set, for both push-pull and single track.



—*Courtesy Electrical Research Products, Inc.*

Figure 89 — Rear scanning schematic. Western Electric system.

The release of movietone push-pull sound track is very limited due to the small number of theatres equipped with push-pull reproducers. However, during the past year several pictures have been released with a



—*Courtesy Electrical Research Products, Inc.*

Figure 90 — Wiring diagram of TA-7400 reproducer set. Western Electric system.

limited number of push-pull copies. It is expected that within a short time enough theatres will be capable of playing push-pull so that it will be practicable to release push-pull prints on a larger scale.

7. OUTPUT POWER REQUIREMENTS FOR THEATRES

The use of "Hi-Range" release print sound tracks, now being released by some studios, requires that the theatre reproducer have suffi-

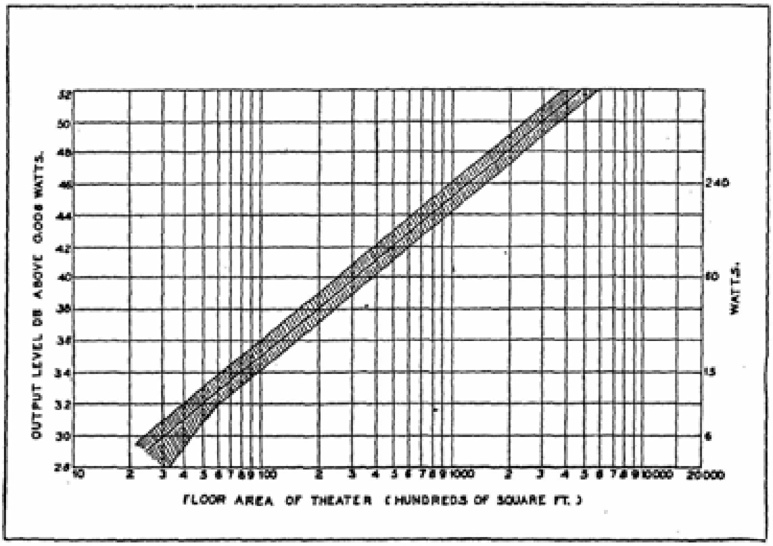


Figure 91 — Recommended amplifier output in electric watts in terms of theatre floor area in square feet.

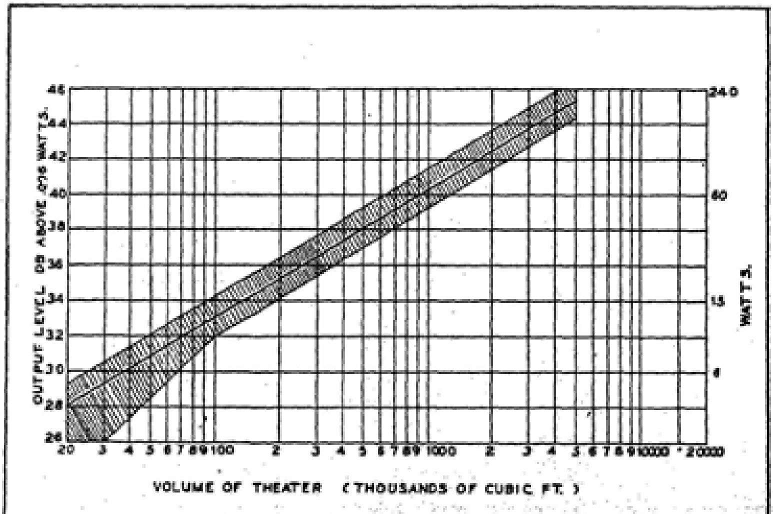


Figure 92 — Recommended amplifier output in electric watts in terms of the theatre volume in cubic feet.

cient output carrying capacity and efficiency to adequately reproduce this increased volume range without compression.

The history of the reproduction of sound has been one of continual increase in amplifier carrying capacity. Originally, output powers from two and one-half to twelve watts were considered adequate, but since then developments in recording have made it possible to utilize up to 60 db range, and it has consequently been found neces-

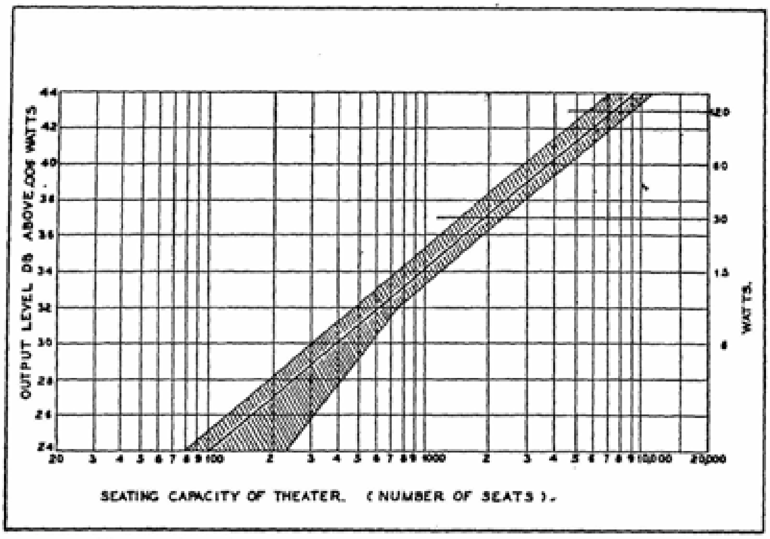
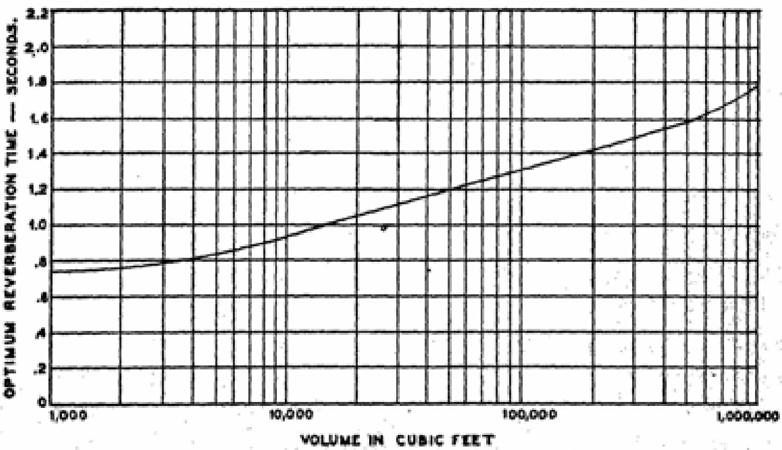


Figure 93 — Recommended amplifier output in electric watts in terms of the seating capacity of the theatre.



—*Courtesy Electrical Research Products, Inc.*

Figure 94 — Optimum reverberation times for different volumes in cubic feet for 512 cycles per second.

sary to increase the theatre equipment power-carrying capacity by large amounts.

Sound effects which incorporate screams, earthquakes, gunshots and other sound effects incident to warfare, demand sensation levels considerably higher than that which could be delivered in the past, and for this reason a maximum output level of not less than ninety sensation

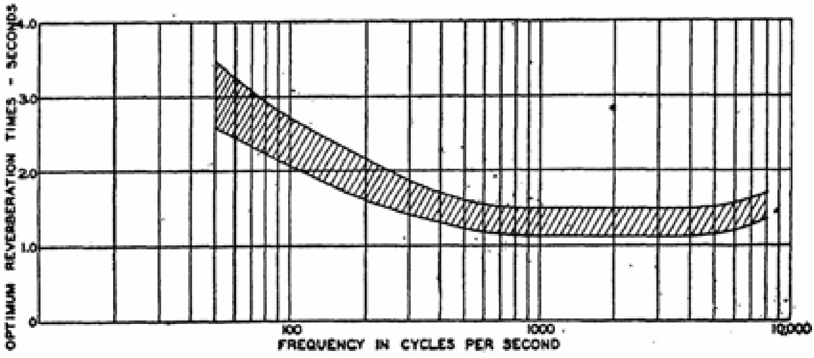


Figure 95 — Optimum reverberation times for different frequencies for a theatre of 300,000 cubic feet capacity.

units is now considered necessary, whereas in the past, amplifier carrying capacity has been limited to 80 db above the threshold of hearing.

8. VOLUME RANGE

High-quality theatre standards determine the necessity for a volume range of at least 60 db, which means that the residual noise of the

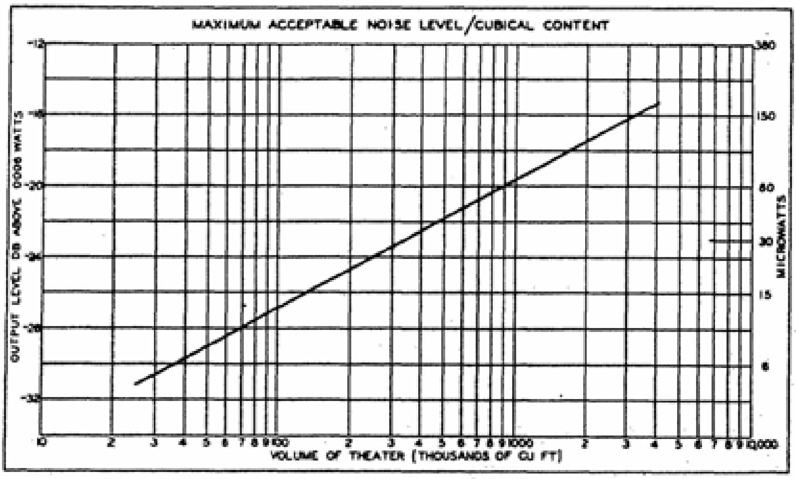


Figure 96 — Maximum acceptable noise level based on output level as recommended in Figure 92 (curve Figure 96, + 60 db = curve, Figure 92).

system, with no modulation, shall be not less than 60 db down from the maximum undistorted carrying capacity of the system. For example: A normal fifty watt system, which has a carrying capacity of plus forty db (6 milliwatt reference), should have a noise not greater than minus twenty, while smaller systems should have proportionately less noise.

Figure 96 shows the acceptable noise level for systems installed according to the recommendations given in Figure 92.

Figures 91, 92 and 93 indicate a yardstick to measure the amount of power necessary for a theatre, when either the floor space area, cubical content, or seating capacity is known. These curves, each of which is equally accurate, give the installed amplifier capacity necessary to maintain the standard required for high quality sound reproduction. Since the required power is a function of the absorption or reverberation in the theatre, deviation from these curves will be required, depending upon the permissible variation from optimum reverberation conditions.

The optimum reverberation time at 512 cycles per second is shown for auditoriums of various volumes in Figure 94. Figure 95 shows the optimum reverberation times in the recording frequency range for an auditorium of 300,000 cubic-feet volume.

One method by which this increased volume range may be procured is by the use of squeeze track on the release print sound track (see Chapter III).

Prints which have such a volume range are called "Hi-Range" prints, and the above mentioned increased amplifier power is necessary in the reproducing equipment to properly reproduce such prints.

Moreover, in reproducing a high volume print, the theatre manager and projectionist should follow the usual method of setting the fader for proper dialogue volume, which will automatically insure a proper reproduced volume level for any musical or sound effects passages in the same production. *If the volume level of the music or sound effects is reduced to a point lower than that originally intended at the time of the recording, dialogue passages will be too low for satisfactory reproduction.*

If the equipment is not functioning properly or if there is insufficient power capacity, the higher volume portions will reproduce with harshness and distortion.

The use of the higher amplifier power necessary to reproduce these prints also requires that the distribution of sound throughout the theatre be particularly uniform.

A "Hi-Range" print having a range of sound intensity of 50 db produces intensity changes which closely approximate those occurring in nature, and musical passages so recorded and subsequently reproduced with adequate power, lend the added color and naturalness necessary to insure a more complete enjoyment of the presentation.

9. STANDARD NOMENCLATURE FOR RELEASE PRINT SOUND TRACKS AND STANDARD FADER SETTING INSTRUCTION LEADER

As part of its program, the Research Council Committee on Standardization of Sound Projection Equipment Characteristics recently published papers on "Procedure for Projecting Hi-Range Prints," "Standard Nomenclature for Release Print Sound Tracks," and "Standard Fader Setting Instructions," which were distributed to every theatre in the United States, Canada and Alaska.

The Standard Nomenclature for Release Print Sound Tracks and the Standard Fader Setting Instructions as specified herein, have been formally approved by the Research Council of the Academy of Motion Picture Arts and Sciences, and adopted as standards for the motion picture industry, effective December 1, 1937.

Each of these papers is included in this book because of its timely interest.

STANDARD NOMENCLATURE for RELEASE PRINT SOUND TRACKS*

As a further step in the program of coordination between studio and theatre, the Research Council of the Academy of Motion Picture Arts and Sciences recently undertook to standardize the nomenclature for release print sound tracks, particularly as developments in sound recording equipment and technique have recently led to the appearance in the theatre field of a number of various new and different types of sound track.

---

* Reprint from the Technical Bulletin of the Research Council of the Academy of Motion Picture Arts and Sciences, November 24, 1937.

SPECIFICATIONS

The Standard Nomenclature for Release Print Sound Tracks follows, with examples of each type included in the illustrations on the following pages.

Plays in "Std." Position of Sound Head Switch

Single variable density - - - - -	Figure 97
Single variable density squeeze - - -	Figure 98
Single variable density double squeeze - -	Figure 99
Unilateral variable area - - - - -	Figure 100
Bilateral variable area - - - - -	Figure 101
Duplex variable area - - - - -	Figure 102

Plays in "P.P." Position of Sound Head Switch

Push-pull variable density - - - -	Figure 103
Push-pull variable density squeeze - -	Figure 104
Push-pull variable area - - - - -	Figure 105

Classification as to Type of Recording

Figures 97, 98, 99, 103 and 104 on the following pages, illustrate the different types of variable density sound tracks, while Figures 100, 101, 102 and 105, illustrate the various variable area tracks.

As may be seen from the illustrations, these two general types of sound track differ fundamentally in that variable density recordings, either "single" or "push-pull," consist of alternate dark and light striations extending across the width of the track and gradually merging one into the other, the sound being represented by these changes in density, while the variable area recordings consist of black and clear transparent sections lengthwise of the film, the sound being represented by the wavy dividing line between these two sections.

Classification According to Power Requirements Necessary for Undistorted Reproduction

Those tracks illustrated in Figures 97, 100, 101, 102, 103 and 105, may be reproduced on those systems having a volume range which was considered adequate up to the present time and previous to the installation of the modern improved equipment with its relatively greater amplifier power.

Figure 97 — Single variable density.

Figure 98 — Single variable density squeeze, showing transition from full-width track to squeeze track.

Figure 99 — Single variable density double squeeze, showing transition from full-width track to double squeeze track.

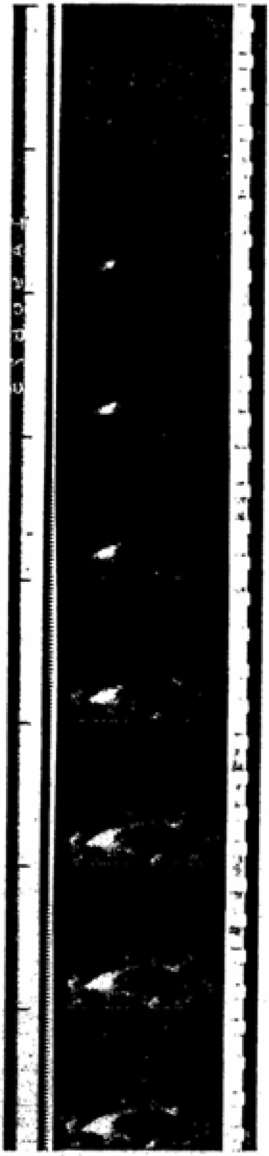


Figure 100 — Unilateral variable area.

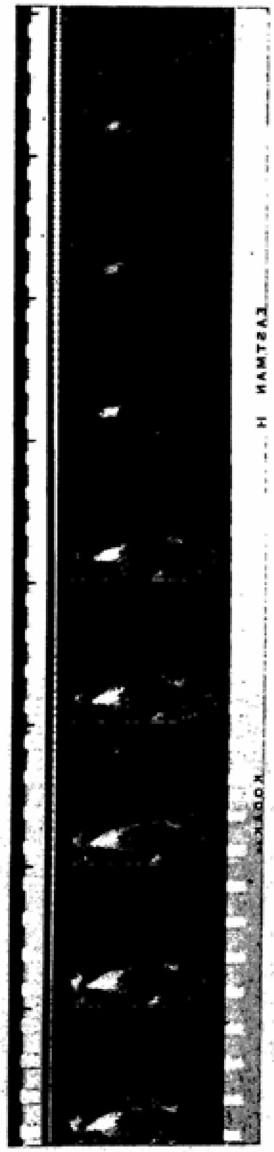


Figure 101 — Bilateral variable area.

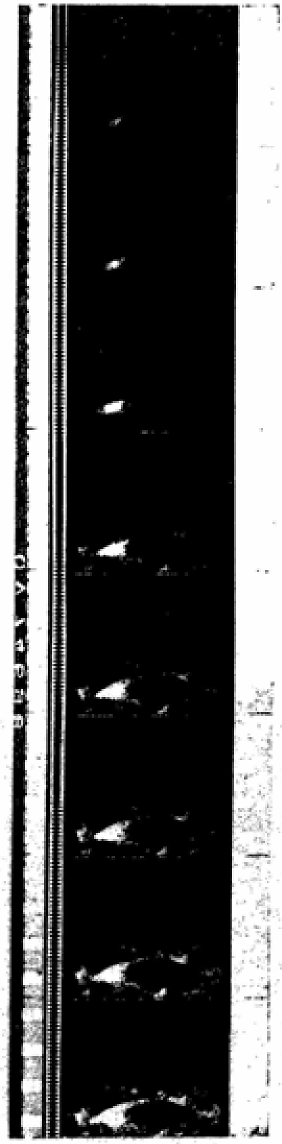


Figure 102 — Duplex variable area.

Figure 103 — Push-pull variable density.

Figure 104 — Push-pull variable density squeeze, showing full-width track, double squeeze of 6 db, and double squeeze of 12 db.

Figure 105 — Push-pull variable area.

Classification by Type of Equipment Necessary for Reproduction

"Push-pull" tracks as illustrated in Figures 103, 104 and 105, can be reproduced only on systems having a double or "push-pull" photo-cell, together with the necessary associated circuits.

Figure 104 illustrates the different amount of "squeeze," or track reduction, now being applied to variable density recordings. The upper portion of this figure shows a "push-pull" track before the application of any "squeeze," the center portion a reduction in track width of one-half, and the lower section a reduction of three-fourths, these being reductions of 6 and 12 db respectively.

STANDARD FADER SETTING INSTRUCTION LEADER*

To further aid the exhibitor in the proper handling of "Hi-Range" prints the studios, commencing about December 1, 1937, will utilize that part of the Academy Research Council Standard Release Print Leader which has been designated for use for any pertinent information to be transmitted from studio to theatre.

A portion of the specifications for the Standard Release Print Leader, indicating the location of this instructional information, is shown in Figure 106, with details of the information to be known as "Standard Fader Setting Instructions" being illustrated in Figure 107.

SPECIFICATIONS

The Standard Fader Setting Instruction Leader shall consist of 15 frames located as specified (Academy Research Council Standard Release Print Leader) in the synchronizing leader; the first frame shall designate the type of print; the second frame the type of reproducing equipment necessary to project the print; and the next nine frames the general fader setting specified in relation to an average fader setting for the particular product under consideration. The remaining frames may be used for whatever additional information the studio may wish to transmit to the theatre.

---

* Reprinted from the Technical Bulletin of the Research Council of the Academy of Motion Picture Arts and Sciences, November 24, 1937.

This instruction leader will be of assistance to the exchange in that it will facilitate the special handling required in the exchange for the various types of prints, by providing an easily noted means of identification for each type.

It should be noted that the designation "Regular" in the Standard Fader Setting Instruction Leader indicates that only one type print has been issued on the particular production under consideration. Productions with prints designated as either "Hi-Range" or "Lo-Range" will have been issued in both type prints, i.e., all productions on "Hi-Range" prints will have necessarily been issued on "Lo-Range" prints as well.

This instruction leader will also enable the projectionist to identify a print which requires a "push-pull" reproducing system as contrasted to a print requiring a "single" system.

In order to identify more plainly the "push-pull" or "single" system prints, it was decided to include both the terms "push-pull" and "single" on every leader, crossing out in the laboratory one or the other of these two to leave the appropriate term designating the type sound track on the print. The illustration of the Standard Fader Setting Instruction Leader shown in Figure 107 indicates the manner by which this was accomplished for leaders which would be included in prints containing a sound track for reproduction on a "single" system. For leaders to be included in prints containing "push-pull" tracks the word "single" would have been crossed out, leaving the word "push-pull" to indicate this type of track.

In order that the exhibitor may achieve the best results, the fader setting designated in this leader should be followed in general, inasmuch as the entire balance between the dialogue and music throughout the reel will be chosen for each designated setting.

SPECIFICATIONS FOR

Academy Research Council Standard Release Print Leader Showing Location of Standard Fader Setting Instructions

Protective Leader

Shall be either transparent or raw stock.
When the protective leader has been reduced to a length of six feet it is to be restored to a length of eight feet.

Identification Leader (Part Title)

Shall contain 24 frames in each of which is plainly printed in black letters on white background: (a) type of print, (b) reel number (Arabic numeral not less than ¼ of frame height), and (c) picture title.

Synchronizing Leader

Shall consist of 20 frames ahead of Start mark, then 12 feet, including Start mark, to picture, opaque except as specified below: In the center of the first frame there shall be printed across the picture and sound track area a white line 1/32 inch wide upon which is superimposed a diamond ¼ inch high.

The next 15 frames may be used by the studio for sensitometric or other information. If not so used this leader shall be opaque.

The Start mark shall be the 21st frame, in which is printed START (inverted) in black letters on white background. The Academy camera aperture height of .631 inch shall be used in the photography of this frame, and all others between Start mark and beginning of picture.

From the Start mark to the picture the leader shall contain frame lines which do not cross sound track area.

In the frames in which the numerals "6" and "9" appear, the words "six" and "nine" (also inverted) shall be placed immediately below the figure, to eliminate the possibility of mis-reading in the projection room due to the similarity between the inverted numerals.

Beginning 3 feet from the first frame of picture, each foot is to be plainly marked by a transparent frame containing an inverted black numeral at least ½ frame in height. Footage indicator numerals shall run consecutively from 3 to 11, inclusive. At a point exactly 20 frames ahead of the center of each footage numeral frame there shall be a diamond (white on black background) ¼ inch high by ⅜ inch wide.

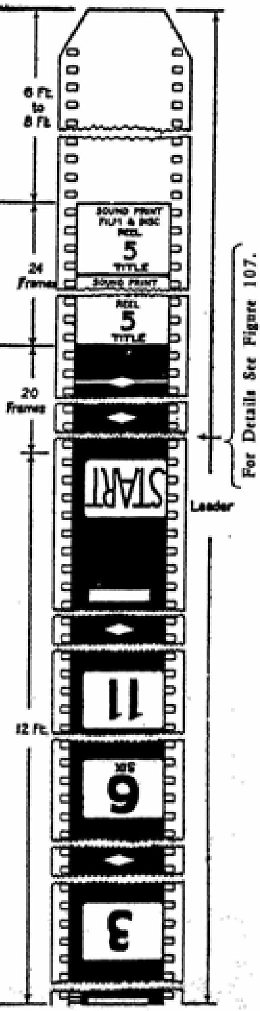


Figure 106.

Start of Picture

For specifications for motor and changeover cue location and reel end leader, see complete Academy Research Council Specifications for 35 mm. Motion Picture Release Prints in Standard 2000′ Lengths, published January 6, 1936.

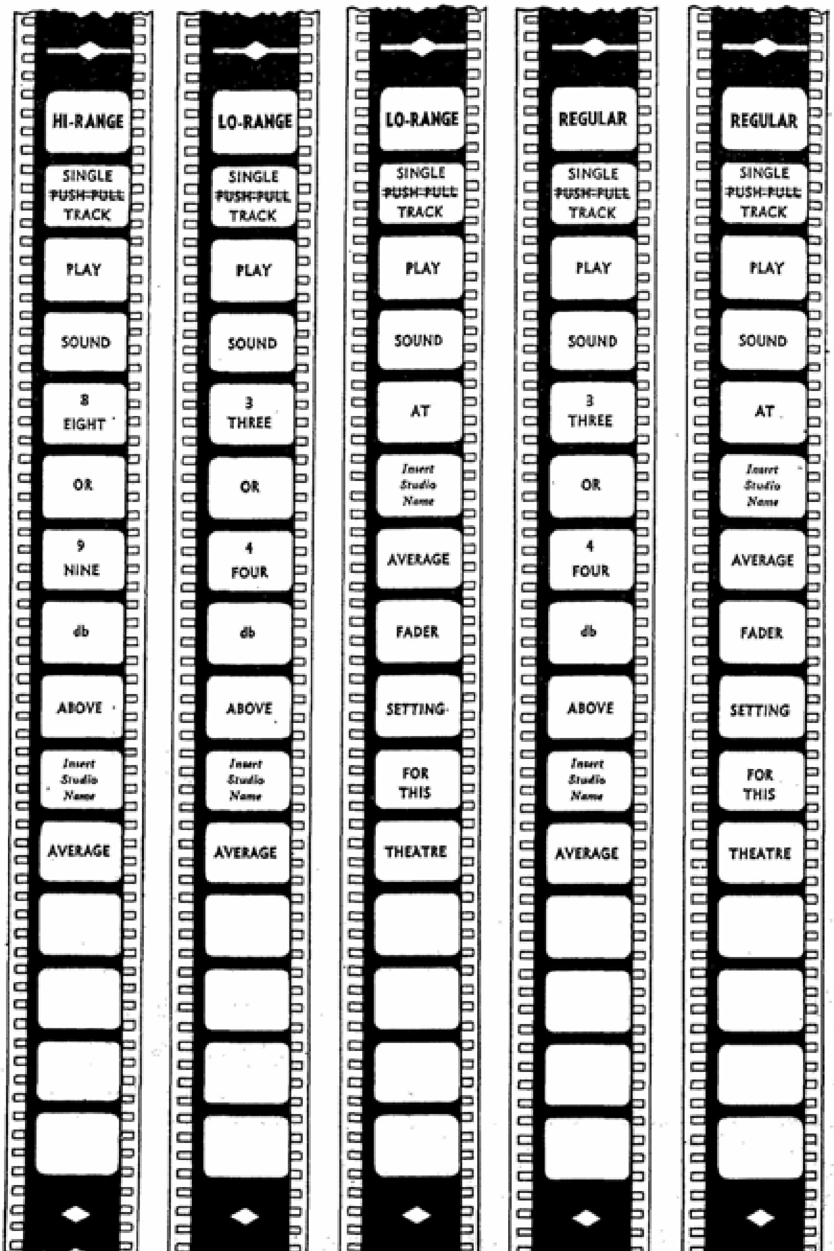


Figure 107 — Standard fader setting instruction leaders.

Chapter XI

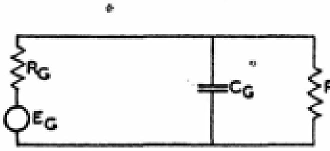
SOUND CIRCUITS

By JOHN K. HILLIARD

1. THE REASON FOR CHOOSING PARTICULAR IMPEDANCE CIRCUITS

Many students of engineering often ask why, in communication circuits, 50, 200 or 500 ohm input and output impedances are used, and why in power circuits, a voltage of 110 is used. Inasmuch as the second question is more easily answered than the first, we will dispose of it by saying that there is no technical reason for a voltage of 110, other than that such voltage just happened to be selected, and has been used commonly for want of a good reason for changing it.

As for the first question, in voice and speech circuits, the length of circuit determines the loss-frequency characteristic for any given impedance. For very high impedances such as 10,000 ohms or more, the high frequencies (5,000 to 10,000 cycles) are attenuated much more rapidly than the low frequencies, due to the fact that the reactance of the capacity of the wiring is comparable to that of the circuit involved.



Circuit 108 — Circuit illustrating the effect of capacitance at high frequencies.

Consider the circuit illustrated in Figure 108. At the lower frequencies the reactance of C_G (which represents the stray capacities of the circuit and for purposes of illustration will be assumed to be 32 $\mu\mu$f) in shunt across R_L is so high that it has practically no effect on the power dissipated in R_L and can therefore be neglected. The maximum amplification of the circuit is then

$$\frac{E_L}{E_G}=\frac{R_L}{R_G+R_L} \tag{27}$$

At higher frequencies where the values of X_C become effective in the circuit, the extent of attenuation is indicated by the ratio of actual amplification at these higher frequencies to the maximum amplification.

$$\frac{\text{Actual amplification}}{\text{Maximum amplification}} = \frac{1}{\sqrt{1+\left(\frac{R}{X_C}\right)^2}} \qquad (28)^*$$

where $R =$ the equivalent resistance of R_G and R_L in parallel.

If now we assume that R_G and R_L are each equal to 1 megohm (10^6 ohms), then $R = \frac{R_G R_L}{R_G + R_L} = 500{,}000$ ohms and, at a frequency of 10,000 cycles, $X_C = 500{,}000$ ohms.

The ratio of actual to maximum amplification

$$= \frac{1}{\sqrt{1+\left(\frac{500{,}000}{500{,}000}\right)^2}} = \frac{1}{\sqrt{2}} = 0.707$$

or an attenuation of 3 db.

At 3,333 cycles the ratio is

$$\frac{1}{\sqrt{1+\left(\frac{500{,}000}{1{,}500{,}000}\right)^2}} = \frac{1}{\sqrt{\frac{10}{9}}} = \frac{3}{\sqrt{10}} = 0.95$$

or an attenuation of 0.5 db.

If R_G and R_L were taken as 500 ohms each, the ratio of actual to maximum amplification at 10,000 cycles

$$= \frac{1}{\sqrt{1+\left(\frac{1}{2000}\right)^2}}$$

The value of $\frac{R}{X_C}$ is thus so small as to be negligible and the ratio of actual to maximum amplification is practically one.

Thus we see a given capacity is less critical when circuit impedances are low. For this reason impedances of 500 ohms or less are practical values to use.

Where inductive pick-up is to be taken into consideration it is desirable to keep the impedance as high as possible or around 500 ohms, so that the voltage is high compared with the induced voltage. This is done in modern photocell output circuits by the use of a transformer, as shown in Figure 109.

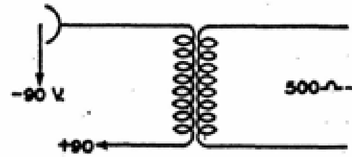


Figure 109 — Photocell output circuit coupled by a transformer.

* The derivation of equation (28) is shown on pages 178 and 179 of Terman's "*Radio Engineering*," 2nd Edition.

2. CONDITIONS FOR MAXIMUM TRANSFER OF POWER FROM ONE CIRCUIT TO ANOTHER

The maximum power will be absorbed by one network from another which is coupled to it at two terminals, when the junction impedances looking into the two networks are conjugates of each other.

It can be demonstrated that the maximum power is absorbed from a generator when the external impedance is the conjugate of the internal impedance. Let us consider the case first where Z_G and Z_L are pure resistances. In Figure 111-A, where $Z_G = R_G$ and $Z_L = R_L$

$$I_L = \frac{E_G}{R_G + R_L}$$

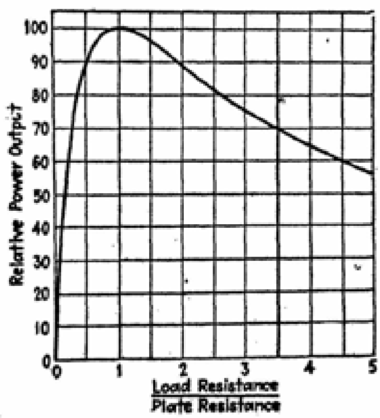
$$P_L = I_L^2 R_L = \frac{E_G^2 R_L}{(R_G + R_L)^2}$$

Differentiating the above equation and equating to zero to find the maximum power, gives

$$\frac{d P_L}{d R_L} = E_G^2 \frac{[(R_G + R_L)^2 - 2 R_L (R_G + R_L)]}{(R_G + R_L)^4} = 0$$

$$R_G^2 + 2R_G R_L + R_L^2 - 2R_G R_L - 2R_L^2 = 0$$

$$\therefore R_G = R_L \text{ (for maximum power transfer)}$$



—*From "Radio Engineering," by Frederick E. Terman (McGraw-Hill Book Co., Inc.).*

Figure 110 — Showing the relative power output at different ratios of generator and load resistances (maximum output when load resistance equals generator resistance).

If Z_G and Z_L have reactance then

$$P_L = \frac{E_G^2 R_L}{(R_G + R_L)^2 + (X_G + X_L)^2}$$

and it can be seen that as regards the reactance X, the power is a maximum when

$$X_L = -X_G \text{ (Resonance)}$$

which means that when Z_L is inductive, Z_G should be capacitive, indicating that a reactive load should be equal in impedance but negative in angle; that is, the impedances should be conjugates.

In power circuits, impedances are never matched to secure the greatest amount of power in the receiving circuit, since the efficiency under these conditions is only 50%. Power devices are rated on maximum possible load current within the safe heating range, and this safe current is considerably less than that which would flow if the circuit were matched. In recording circuits, however, where the cost of power is not a first consideration, and where currents are small so that temperature rise is negligible, considerable use is made of impedance matching to obtain maximum transfer of power from one circuit to another.

This brings us to what is known as the *reciprocity theorem*, which states that *when a source of voltage is connected across a pair of terminals of a passive four-terminal network, and an ammeter is connected across the other pair of terminals, the source of voltage and the ammeter may be interchanged without altering the reading of the ammeter.*

The effect of inserting a circuit element in a general network can be calculated by considering the current in the receiver before and after the element is inserted. The current in

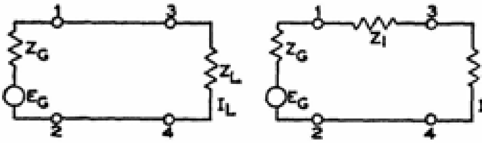


Figure 111-A Figure 111-B

the receiver when the two circuits are connected together is

$$I_L = \frac{E_G}{Z_G + Z_L} \tag{29}$$

When a series impedance Z_1 is inserted between terminals 1-3, Figure 111-B, the receiver current is

$$I'_L = \frac{E_G}{Z_G + Z_L + Z_1} \tag{30}$$

Hence, the ratio of currents before and after the insertion of the series impedance is

$$\frac{I'_L}{I_L} = \frac{Z_G + Z_L}{Z_G + Z_L + Z_1} \tag{31}$$

Likewise, the current when Z_1 is connected in shunt across the terminals 1-2, Figure 112, is

$$I'_L = \frac{E_G Z_1}{Z_G (Z_L + Z_1) + Z_L Z_1}$$

$$\text{so } \frac{I'_L}{I_L} = \frac{\dfrac{Z_1}{Z_G (Z_L + Z_1) + Z_L Z_1}}{\dfrac{1}{Z_G + Z_L}}$$

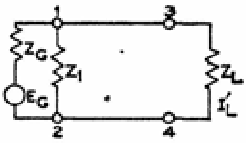


Figure 112.

$$\frac{I'_L}{I_L} = \frac{Z_1 (Z_G + Z_L)}{Z_1 (Z_G + Z_L) + Z_G Z_L} \tag{32}$$

NOTE—In the case of pure resistance, substitute R for Z in the above equations, and where Z is reactive it must be remembered that Z is a vector quantity.

Problems on the Effect of Inserting Impedances in a Network

Given: Circuit diagram Figure 111-A, where the generator and load impedances are resistances, i.e., $Z_G = R_G$ and $Z_L = R_L$.

$$R_G = R_L = 500 \text{ ohms}$$

1. What is the loss in db if a 500 ohm resistance R_1 is inserted in the circuit of Figure 111-A between terminals 1-3, as in Figure 111-B?

Let $I_L =$ the load current before insertion of the resistance.

$I'_L =$ the load current after insertion of the resistance.

Then
$$\frac{I'_L}{I_L} = \frac{R_G + R_L}{R_G + R_L + R_1} \quad \text{(Eq. 31)}$$

$$= \frac{1000}{1500} = 0.67$$

From Table VI, Page 455, an attenuation ratio of 0.67 indicates a loss of 3.5 db, or $N_{db} = 20 \log_{10} \frac{1}{0.67} = 3.5$ db.

2. What is the loss in db if R_1 of problem 1 is inserted in the circuit of Figure 111-A between terminals 1-2, as in Figure 112?

$$\frac{I'_L}{I_L} = \frac{R_1 (R_G + R_L)}{R_1 (R_G + R_L) + R_G R_L} \quad \text{(Eq. 32)}$$

$$= \frac{500\ (1000)}{500\ (1000) + 250{,}000} = \frac{500{,}000}{750{,}000} = 0.67$$

Attenuation ratio $= 0.67$

Loss $= 3.5$ db

3. *Find:* The loss in db in the above circuit if an inductance of 1.6 henries is inserted between terminals 1-3 of Figure 111-A, at

(a) 50 cycles per second

(b) 10,000 cycles per second.

Given: Circuit diagram as in Figure 111-B where $Z_G = R_G$, $Z_L = R_L$ and $Z_1 = jX$.

$$R_G = R_L = 1{,}000 \text{ ohms}; \; L = 1.6 \text{ henries}$$

(a)

$$\frac{I'_L}{I_L} = \frac{R_G + R_L}{R_G + R_L + jX} \qquad \begin{array}{r}(\text{Eq. 31, where } Z_1 = jX)\\ X = 2\,\pi fL\end{array}$$

$$\frac{I'_L}{I_L} = \frac{1000 + 1000}{\sqrt{(1000 + 1000)^2 + (500)^2}} = \frac{2000}{\sqrt{4{,}250{,}000}}$$

$$= \frac{2000}{2064} = 0.969$$

so loss in db $= 0.3$

(b) at 10,000 cycles

$$X = 2\,\pi\, 1.6 \times 10{,}000 = 100{,}000$$

so $$\frac{I'_L}{I_L} = \frac{R_G + R_L}{R_G + R_L + jX} = \frac{2000}{\sqrt{(2000)^2 + (100{,}000)^2}}$$

$$= \frac{2000}{\sqrt{10{,}004{,}000{,}000}} = \frac{2000}{100{,}020}$$

$$= 0.02$$

so loss in db $= 34$ db

NOTE: This problem illustrates the relative loss due to an inductance in series with the load (at the approximate limits of the band used in recording) at low and high frequencies.

3. A STUDY OF POWER DELIVERED BY A CONSTANT VOLTAGE GENERATOR TO A LOAD CIRCUIT

If a battery has no internal resistance, R_G, that is, if $R_G = 0$, then I_{sc} = infinity, where I_{sc} = the short circuit current. Actually, however, this is never the case, for a battery always has some effective internal resistance. This resistance may be as low as 0.001 ohms for a huge storage battery, or up to one megohm or more for a silver chloride battery, depending upon the physical and chemical make-up of the battery.

We can, however, represent a generator or a source of power as a zero resistance generator with an accompanying external resistance, thus:

$$E = E_G - I_G R_G \tag{33}$$

Where E = effective voltage across the power source with circuit closed.

E_G = open circuit voltage across the power source.

I_G = the current flowing in the closed circuit.

This is known as the equivalent generator theorem and applies to any circuit from which power can be taken. (See Figure 113.)

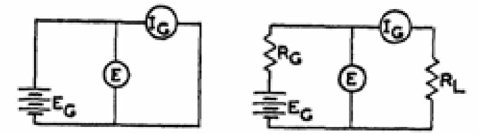


Figure 113 — Circuit diagram showing equivalent generator circuit.

If we let R_L = load resistance,

E_L = the voltage across the load,

I_L = the load current,

W_L = power consumed by the load,

$E_{oc} = E_G$

then it is found experimentally that as R_L is varied from infinity to zero, I_L varies from zero to I_{sc}, E_L varies from E_{oc} to zero and W_L ($= I_L E_L$) varies from zero to a maximum and back to zero. This maximum W_L occurs when $R_L = R_G$, at which time $E_L = \frac{1}{2} E_{oc}$; $I_L = \frac{E_{oc}}{2 R_L}$; and $W_L = W_G = \frac{E_{oc}^2}{4 R_L}$.

Here it will be noted that the condition for maximum output means $W_G = W_L$ and therefore just as much power is dissipated in the generator as in the load.

By definition, the voltage regulation of a system is

$$\text{Voltage Regulation (in per cent)} = \frac{E_{oc} - E_L}{E_L} 100 \tag{34}$$

Where

$$R_G = R_L, \text{ the voltage regulation is}$$

$$\text{V.R.} = \frac{2E_L - E_L}{E_L} 100 = 100 \text{ per cent}$$

as

$$E_{oc} = 2E_L$$

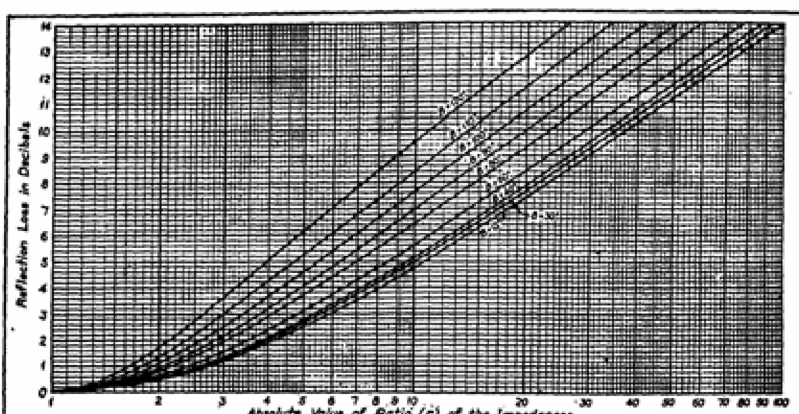
In the alternating-current case we must change our resistance to impedance, which means that we now have two independent variables instead of one, namely, the resistive and reactive components of the impedance.

In circuits which are not matched, reflection takes place at the point of mismatch. The amount of reflection will depend upon the ratio of the impedances looking in the two directions, that is

$$\text{Reflection factor} = \frac{\sqrt{4\, Z_G\, Z_L}}{Z_G + Z_L} \tag{35}$$

If $Z_G = Z_L$, the reflection factor $= 1$ (transmission factor).

The reflection in long circuits produces reversed currents, which may be out-of-phase sufficiently to nullify the incoming current, but in motion picture work this is of no appreciable value, since the time between the reflected wave and the original wave is in the order of a few microseconds, due to the short length of these circuits. However, when



The curves show the loss, in decibels, resulting from a mismatch of circuit impedances, as, for instance, connecting a 200-ohm line to a 500-ohm amplifier input.

The chart is a plot of the equation:

$$\text{Loss (in db)} = 10 \log_{10}\left[1 + \frac{(1 - r)^2}{4r\cos^2\left(\frac{\beta}{2}\right)}\right]$$

Losses for values of β not given on the chart may be found by use of this equation.

Example of the use of the chart: Assume that a circuit having an impedance of $50\angle{-40^\circ}$ is connected to a circuit with an impedance of $500\angle{130^\circ}$. Adding the angles algebraically, we find $\beta = 90^\circ$. The ratio r of the absolute values of the impedances is $500/50 = 10$. From the chart at the intersection of $\beta = 90^\circ$ and $r = 10$ is read the loss 7 db, on the left ordinate.

Figure 114 — Losses in mismatched circuits.

working with vacuum tube output circuits feeding into a loud-speaker, the impedances are matched because certain types of vacuum tubes, such as pentodes, are critical to load impedance. A loud-speaker at its resonant point when working from a pentode, illustrates this type of matching, being able to absorb a great deal more power at its resonant point since its impedance is high.

A specific example of a case where the problem of mismatch between circuits is negligible, leading to no resultant distortion, is a mixer circuit, as it is only necessary that the output circuit from the amplifier be proper, and that the input to the following tube be such a value that the frequency characteristic is maintained.

The term, characteristic impedance, Z_o, is defined as

$$Z_o = \sqrt{Z_{oc} Z_{co}} \qquad (36)$$

Hence, it is possible to find the impedance of a four-terminal network with an impedance bridge by measuring its open-circuit and closed-circuit impedances.

Example of Power Output Regulation Through Impedance Matching

1. POWER CASE:

If we assume that the open-circuit voltage across a generator is 115 volts, and that this voltage drops to 110 volts when a load drawing 30 amperes is connected to the generator, then:

$$Z_G = \frac{E_G - E_L}{I_L} = \frac{115 - 110}{30} = \frac{1}{6} = 0.167 \text{ ohm} \quad \text{(Eq. 33)}$$

If now we were to match the generator and load impedances $(Z_L = Z_G)$, then we would find that

$$I_L = \frac{E_G}{Z_G + Z_L} \qquad \text{(Eq. 29)}$$

$$\frac{115}{0.334} = 344 \text{ amperes}$$

which would burn out the equipment. It is therefore not always possible or advisable to match load and generator impedances in order to secure maximum power output.

2. COMMUNICATION CASE:

Assume $Z_G = 500$ ohms, $E_G = 20$ volts.

Then if

$Z_L = 10{,}000$ ohms	$I_L = 0.0019$ amp.	$E_L = 19.0$ volts	$P_L = 0.0361$ watts
$= 1{,}000$	$= 0.0133$	$= 13.3$	$= 0.177$
$= 500$	$= 0.020$	$= 10.0$	$= 0.20$
$= 200$	$= 0.0286$	$= 5.7$	$= 0.163$
$= 50$	$= 0.0364$	$= 1.8$	$= 0.0655$
$= 0$	$= 0.04$	$= 0$	$= 0$

FORMULAE

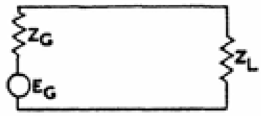


Figure 115 — Circuit diagram for Example 2.

$$I_L = \frac{E_G}{Z_G + Z_L};\ E_G = Z_G\, I_L + Z_L\, I_L;$$

$$E_L = Z_L\, I_L;$$

$$E_L = \frac{Z_L}{Z_G + Z_L}\, E_G;\ P_L = E_L\, I_L$$

4. SHIELDING

Among the important and perplexing difficulties encountered in the design of recording and reproducing circuits are pick-up and cross-talk signals. Although the amount of energy encountered is very small it may prove extremely difficult to eliminate.

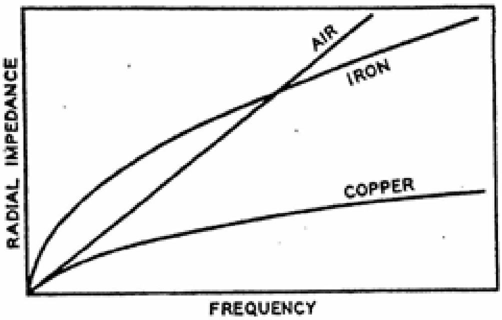
The problem of shielding is comparatively simple in principle but, in practice, radical steps are often necessary to eliminate pick-up and cross-talk.

One method of approach is based on the fact that the movement of energy at right angles to the circuits, which is the cause of pick-up, follows the same basic law as those laws which govern the desired transmission. From this point of view an electro-magnetic disturbance starts from the conductor and spreads radially outward throughout the surrounding space, constituting the first section of the radio transmission line. At the beginning and end of the shield the electrical characteristics of this line suddenly change. Outside of the shield the radio transmission line extends into space, but if the shield is correctly constructed very little energy is able to reach this section.

In their progress through metallic substances, electro-magnetic waves are attenuated at a rate depending upon the frequency, permeability, and conductivity of the metal. This attenuation is brought about by conversion of electrical energy into heat.

The effectiveness of a shield is due partly to the attenuation and partly to the reflection occurring at the boundary of the shield because of the mismatch, in radial impedance, of the shield and surrounding

insulator. As shown in Figure 116, the radial impedance of copper is very much lower than that of air. This means that while at any particular frequency there is no reflection between iron and air, there will be a



—*From the Bell System Technical Journal.*

Figure 116 — Curves showing the radial impedance of air, iron and copper at different frequencies.

large reflection between copper and air. For this reason, in thin shields the higher attenuation loss of iron may be more than offset by the greater reflection between copper and air. If a composite shield is made of alternate layers of copper and iron, the effectiveness of the shield is very much greater when the outside layers are made of copper, due to the fact that the reflection loss between copper and air is higher since the attenuation loss is independent of the arrangement of the layers.

Another interesting fact is that while non-magnetic shields become increasingly effective with increase of frequency, this is not always true with magnetic shields. At low frequencies magnetic shields are very efficient. As the frequency increases they sometimes become less effective but ultimately reach a minimum beyond which they improve again, so that at sufficiently high frequencies they are always better than non-magnetic shields. These characteristics are due to the manner in which the impedances mismatch, as previously explained.

In the case of non-magnetic shields the impedances of the shield and dielectric are always mismatched except for zero frequency by an amount which increases with the frequency as shown in Figure 116, while for magnetic shields the mismatch is large at low and high frequencies, but is small at certain intermediate points.*

5. LONGITUDINAL CURRENTS

Very often in recording or reproducing circuits, inductive interference is picked up where currents are induced into both sides of a

* Bell Laboratory Record, *"A Theory of Shielding,"* by S. A. Schelkunoff, March, 1936, page 229-232.

circuit in the same direction in parallel paths, and the return path is by some other part of the circuit (such as through the ground). If the impedance of the return path to ground is high, cross-talk may enter the circuit. Cross-talk currents often have a considerable magnitude when the communication circuit is exposed to power lines, due to irregularities in balance in the former circuit.

The most effective way of reducing the longitudinal current is to provide a short-circuit path to ground for this current or to create an open-circuit to this longitudinal current without affecting the transverse circuit.

Accurately balanced transformers may be inserted at both ends of a circuit, with a ground at the mid-point of each transformer as indicated in Figure 117.

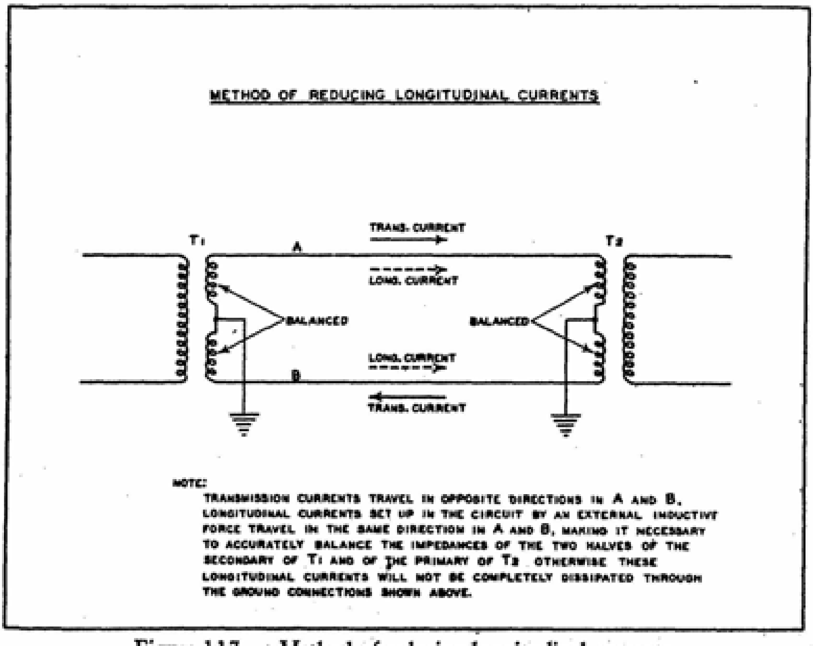


Figure 117 — Method of reducing longitudinal currents.

6. METHOD OF ACOUSTICALLY ADDING REVERBERATION TO SOUND TRACK

It is often desirable to increase the reverberation in an original sound track through the medium of recording.

During the earlier years of studio sound recording experience, this was accomplished by using staggered or offset tracks which were lined up a few frames apart and then mixed together to secure the desired effect.

Attempts were also made to use loud-speaker systems in highly reverberant echo chambers, combining in varying degrees the pick-up from them with the original sound. This system has been used to effect a different amount of "presence" in playbacks, depending upon the picture cut, to gain the proper perspective. However, when a large percentage of the sound is taken from the echo chamber, the high degree of quality required is not maintained due to the deficiencies in the loud speakers available.

Since the adoption of the current two-way loud-speaker systems, it has been found practical to re-record a sound track acoustically from an echo chamber. The distortion apparent when this track is compared back to back* with the original is a minor order effect, often not detectable at all. This is a very practical method of adding reverberation in recording without loss of the frequency characteristic of the original recording.

The necessary set-up consists of splitting the mixer into two banks, one bank for the control of those tracks which are intended to have reverberation added, the output of this bank being divided into two paths with isolation amplifiers. One path is directed into the echo chamber, the pick-up from this chamber appearing on one position of the

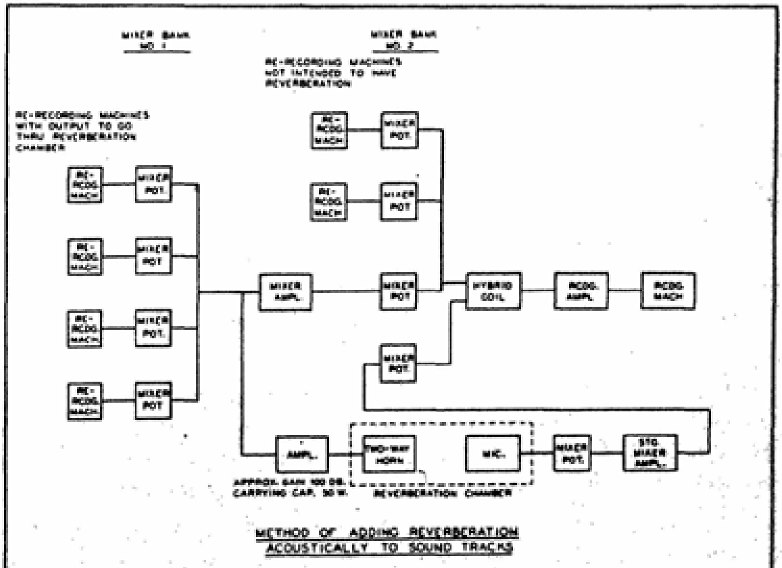


Figure 118.

*"Back to back" is a studio term for the direct comparison of two or more sound tracks run simultaneously on interlocked sound projection machines.

second mixer bank. The output of the second path of the first mixer bank appears also on the second bank. In this manner, the sound entering the echo chamber is pre-mixed and later combined with the original track to complete the desired illusion with the picture in terms of perspective. (See Figure 118.)

The echo chamber in which the horn and microphone are placed may be a room of approximately 10,000 cu. ft. volume, having a relatively high reverberation time. Studio experience indicates that a reverberation time of approximately five seconds is ample for any desired effect. This time period is obtained by lining the walls with a glazed hard surface material of considerable weight (sheet rock, or hard masonite covered with a hard paint has been found to be very practical), with the floor of painted wood or cement.

In order to obtain a flexible system of operation, so that more than one sound track may be passed through the chamber, the portion to have reverberation added must be "pre-mixed."

7. PRE-RECORDING

In order to maintain a degree of illusion of reality in music comparable to the degree of illusion of reality obtained in dialogue recording, somewhat more complex methods have been devised for recording music.

Practically all songs appearing in present day motion pictures are "pre-recorded," which means that the music is recorded before the actual filming of the picture. This is necessary in order that the continuity of the music may be maintained regardless of camera angle, which may shift in direct cuts from a long shot to medium shot or close-up and back, depending upon the dramatic requirements of the scene. Naturally, a steady flow of sound must accompany the picture regardless of the camera angle, and this steady continuation of music could not be maintained by direct recording at the time of the filming of the picture.

For "pre-recording," a play-back system has been devised, consisting of a portable disc or film reproducer, driven in step with the camera by means of either a synchronous or an interlock motor system, and used in conjunction with one or more horns located about the set.

The music previously recorded is played through this reproducing system and the artist, hearing the previously recorded song, provides the action by singing in front of the camera in tempo with the original recording. In this way a scene may be broken up into a series of short takes and need not necessarily be shot in its entirety at one time.

An automatic stop-start turn table, controlled by the director on the set, has proven a valuable tool in the use of play-backs.

The success of the pre-recording method depends upon the individual aptitude of the artist in providing well synchronized sound and picture, as well as adept editing of the sound track and picture to remove any out-of-sync condition.

A number of recordings of the same musical selection may be intercut to give the best possible rendition of the musical number. Instead of selecting a completed recording because it is the best overall average take, it is possible to intercut the best parts of a number of takes, thus reducing the effort on the part of the artist to record perfect complete takes, and considerably reducing the time required to obtain any given scene.

This technique of course requires a very high degree of precision of both artist and equipment, in order that exact pitch may be maintained between takes of the same number.

Chapter XII

MEASUREMENTS IN SOUND CIRCUITS

By JOHN K. HILLIARD

1. METHOD OF MEASURING HARMONICS

THE WEIN BRIDGE METHOD OF MEASURING HARMONICS: This bridge causes a network of resistances and condensers to completely balance out the fundamental frequency and when harmonics are introduced by an overloaded amplifier or a clashing light valve it is possible not only to hear the harmonics as they are being generated, but also to measure their amplitude.

2. LOAD CARRYING CAPACITY OF AMPLIFIERS

When the input voltage to an amplifier is increased to a point causing overload, the ouput voltage does not maintain the same wave form, and the distortion that results produces harmonics of frequencies not present in the signal being amplified. Usually in high-grade amplifiers when the output is more than 10 db down from its rated output, the percentage of harmonics is very small. For some time past it has been customary to rate the load carrying capacity of an amplifier at a value which produced 5% combined harmonics. Recently, however, this rating has been superseded by 1% total harmonics since high-grade reproducing equipment requires a much higher standard than formerly used. It has been customary to rate the amplifier in terms of the r.m.s. amplitude of the harmonics expressed as a percentage of the fundamental frequency component. The tendency will be to rate the amplifier in terms of the number of decibels difference between fundamental and harmonics. When making these harmonic measurements, it will be found desirable to check new set-ups at very low frequencies as well as at 1,000 cycles, since in most cases the carrying capacity is considerably reduced at the low frequencies because of transformers which have insufficient core material. (See Figure 119.)

3. METHODS OF MEASURING IMPEDANCE

If a source of frequencies such as an oscillator as well as either a power level indicator or a peak voltmeter and a variable resistance are

available, the resistive impedances of equipments may be measured in the field by the approximate method following. In the first method

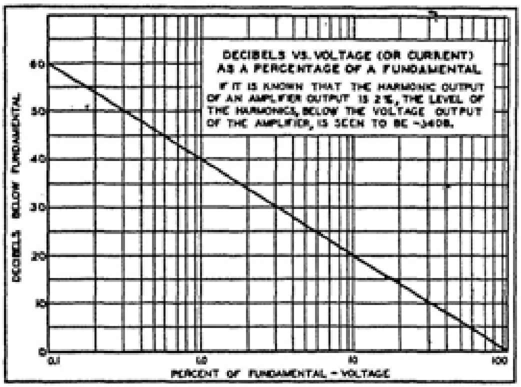


Figure 119 — Curve showing the per cent harmonics present in an amplifier.

(see Figure 120), which is of use in measuring load impedances, a calibrated variable resistance is placed between the generator and the load.

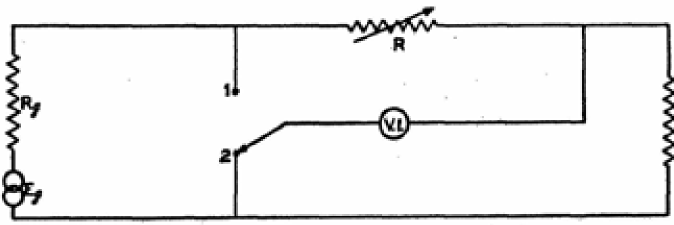


Figure 120 — Series method of measuring impedances.

The resistance is varied until, with the use of the peak voltmeter or power level indicator, the same voltage is obtained across the series re-

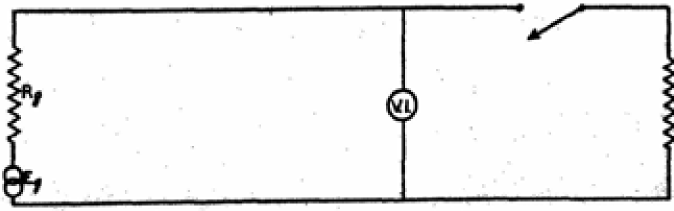


Figure 121 — Shunt method of measuring impedances.

sistance as is measured across the load. For this condition the load resistance R_L is qual to the amount indicated by the calibrated variable resistor.

The second method (see Figure 121) is useful in field work to measure the output impedance of amplifiers, oscillators, etc. It consists

of placing a voltage-indicating device across the generator terminals and measuring the output voltage with and without a known resistive load; that is, E_L and E_G respectively. For this condition $\frac{E_G}{E_L} = 1 + \frac{R_G}{R_L}$ from which R_G may be computed, since all other items of the equation are known. Where a volume indicator is used to indicate voltage, E_G and E_L may be obtained from the difference in db between the two level readings by means of the equation,

$$\text{db difference} = 20 \log \frac{E_G}{E_L}$$

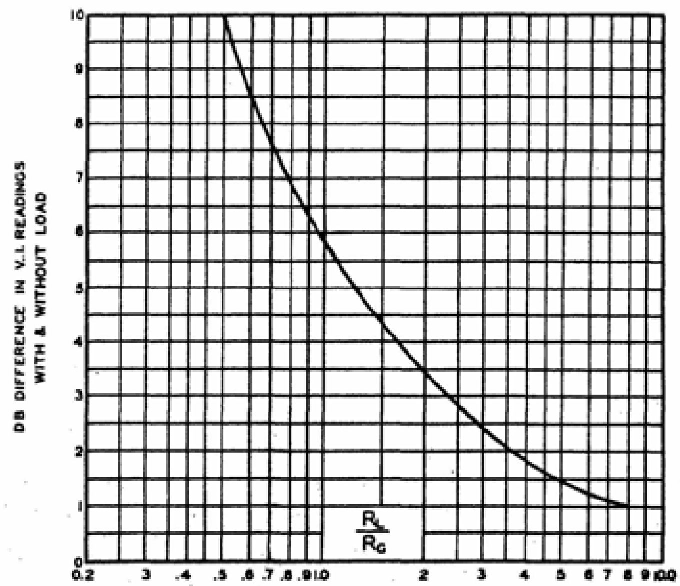


Figure 122 — Shunt method of measuring resistance impedance.

The curve of Figure 122 gives $\frac{R_L}{R_G}$ directly in terms of this difference reading in db. It is noted that where the known load impedance is equal to the generator impedance, the db difference reading is 6.0 db; that is

$$\frac{R_G}{R_L} = 1.$$

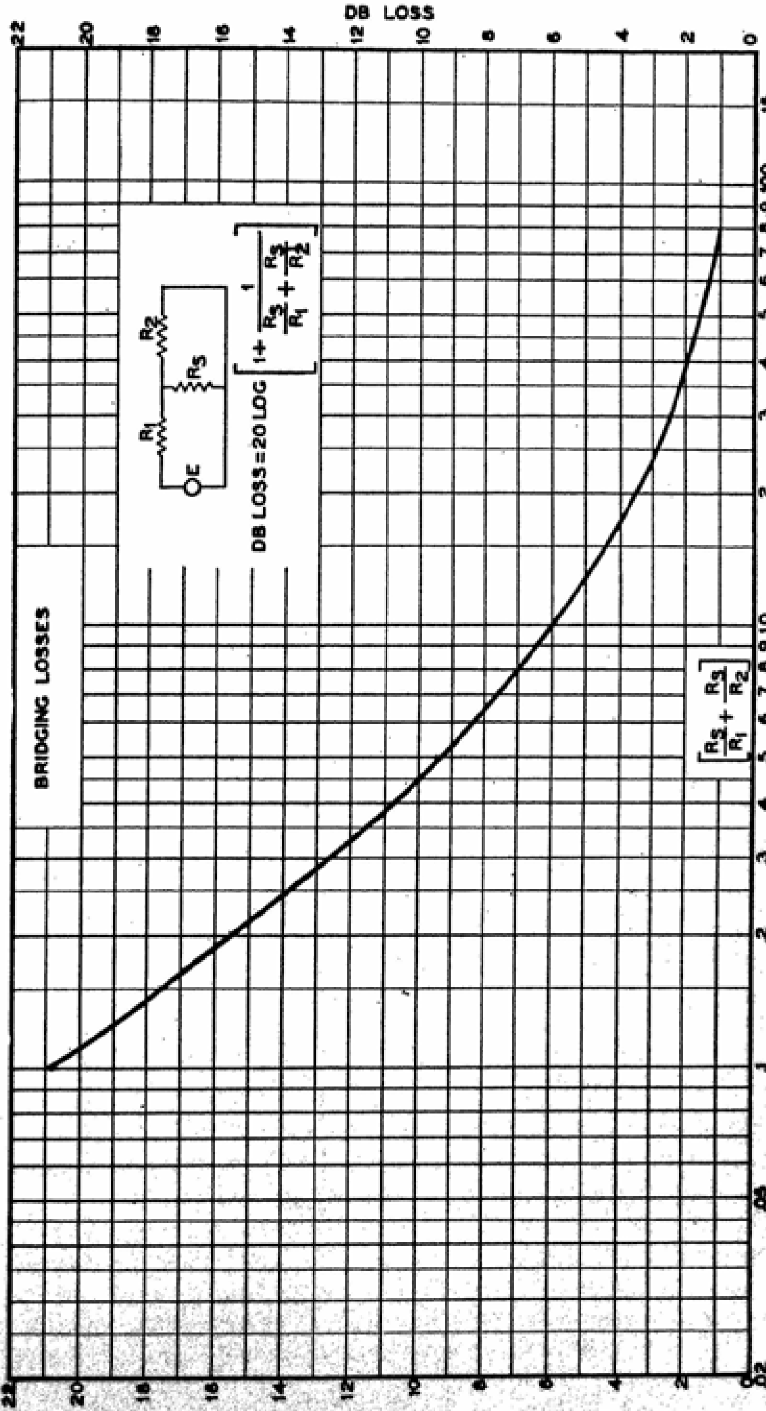

DB LOSS
BRIDGING LOSSES
E
R_1
R_S
R_2
DB LOSS = 20 LOG $\left[1+\frac{1}{\frac{R_S}{R_1}+\frac{R_S}{R_2}}\right]$
$\left[\frac{R_S}{R_1}+\frac{R_S}{R_2}\right]$

Figure 123.

GRAPHICAL SOLUTION OF BRIDGING LOSS

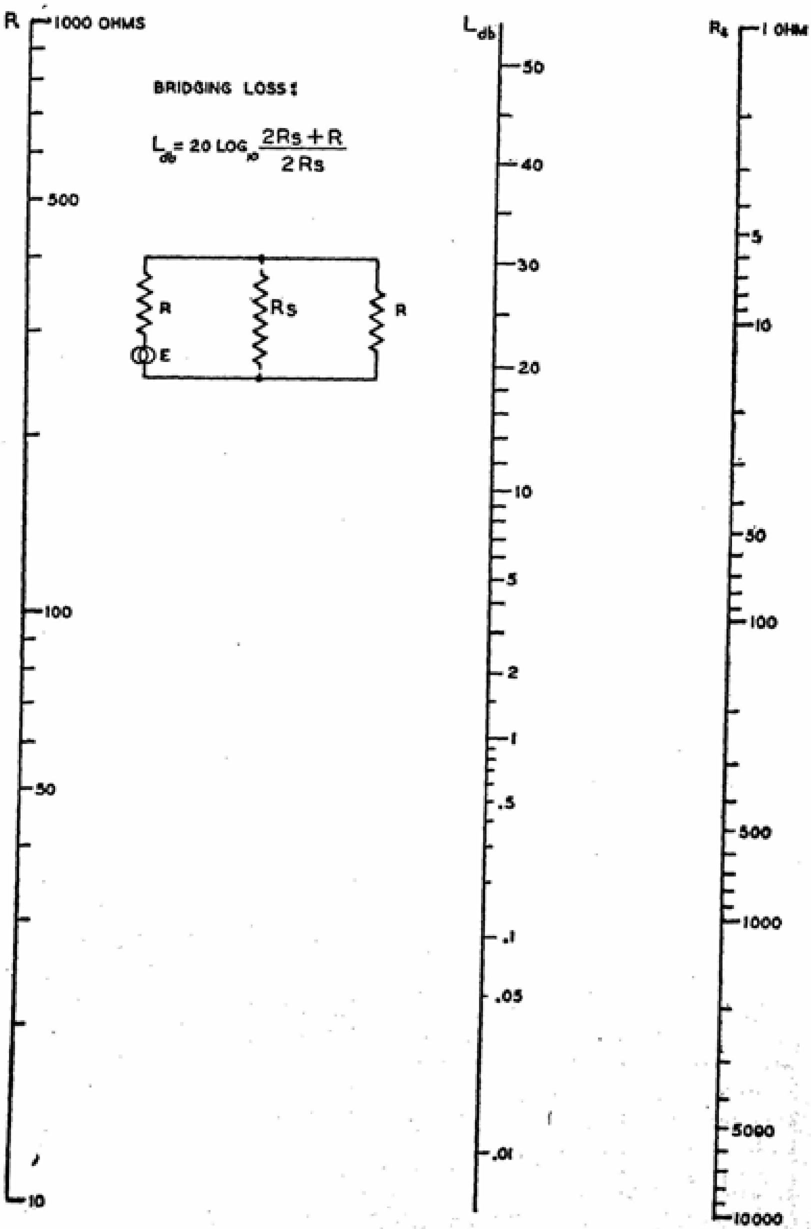


Figure 124 — Graphical solution of bridging loss.
(When R_1 and R_2 of Figure 123 are equal.)

When is it desired to measure the effect of bridging one resistive load across a circuit as shown in the bridging loss chart of Figure 123 or 124, the loss realized is equal to

$$20 \log \left[1 + \frac{1}{\frac{R_S}{R_1} + \frac{R_S}{R_2}} \right]$$

The curve of Figure 123 shows this loss plotted against the ratio

$$\left[\frac{R_S}{R_1} + \frac{R_S}{R_2} \right]$$

This curve may also be used for indicating the corrections for power level indicators when used on impedances other than those for which they are calibrated.

4. METHODS OF MEASURING DIVIDING NETWORKS

With the use of high quality, two-way loud speaker reproducing systems, which utilize dividing networks to distribute the power between the different bands, it is necessary to determine that the division of power is taking place properly. This is done by measuring their characteristics by the following method. The network is placed at the output of the amplifier from which it is to receive power, and the horn loads are removed and resistances equivalent to the speaker impedances substituted. A power level indicator is then placed across the low-frequency branch of the network, and a curve is obtained throughout the complete frequency range. The meter is then switched to the high-frequency leg and the procedure duplicated.

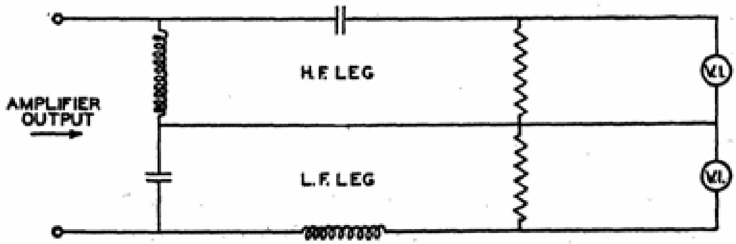


Figure 125 — Method of measuring dividing networks.

The usual half-section dividing networks now used in the field have an attenuation of 12 db per octave if the crossover point is at 300 cycles. This means that the attenuation in the low-frequency leg is 12 db down from 300 cycles at 600 cycles, and that the attenuation in the high-frequency leg is also down 12 db from 300 cycles at 150 cycles. In an efficient network of this type the minimum loss at the crossover point will be not less than 3 db and usually runs approximately 3½ db.

5. MICROPHONE MEASUREMENTS

(a) Distortion

The microphone introduces distortion into the system, depending upon the size of the microphone and the cavity resonance present in the microphone, the effect of the latter varying with the relation of the wave length of the sound to the dimensions of the cavity. Considerable research has been done in measuring these distortions, and it is a well known fact that the microphone and its enclosure should be made as small as is commercially practicable.

Due to the fact that the dynamic or moving coil type of microphone requires that the contents of the various elements of the circuit be so chosen that the magnitude of the impedances be the same at all frequencies, the microphone response is obtained by acoustic equalization. This equalization, which varies with frequency, usually takes the form of modifying the stiffness of the diaphragm, and providing vent or escapement tubes between the front and back of the diaphragm.

As a result of these various methods of equalization, it has proven difficult to manufacture these microphones commercially with a high degree of uniformity, while on the other hand, it is a comparatively easy task to so manufacture condenser microphones. Recently it has been found practicable to make small condenser microphones with a diameter of less than one inch and with gain-frequency and directional characteristics superior to those of the dynamic type. Since their size is small, distortion due to cavity and diaphragm resonance, and distortion due to field disturbance, has been greatly reduced.

In the near future it is very likely that this microphone will again come into popular use for dialogue recording.

(b) Microphone Calibration

In order to establish the characteristics of a microphone, it should be tested to determine its sensitivity, gain-frequency response characteristic, impedance-frequency characteristic, and directivity. Since sensitivity and gain-frequency characteristics can be obtained in terms of the pressure of the sound wave, the sensitivity is usually expressed in terms of the power output across a resistance load equal to that into which the microphone is designed to operate. The impedance characteristic will give a measure of the variations which arise when matching the microphone to the input of an amplifier.

A very practical method to obtain a pressure calibration of a condenser microphone, is to apply an electrostatic driving force to the

diaphragm, which can be done by mounting a fixed plate near the diaphragm, and applying an alternating voltage between the two surfaces.*

If two parallel plates are separated by a distance d cms, in air, and a statvoltage E is applied between them, the electrostatic pressure developed is

$$P = \frac{E^2}{8\pi d^2} \text{ dynes per sq. cm.} \qquad (37)$$

Now, if a sinusoidal voltage is applied where $E = e_m \sin \omega t.$, then

$$P = \frac{e_m{}^2 \sin^2 \omega t}{8\pi d^2} \text{ (instantaneous value of } P\text{)}$$

$$= e_m{}^2 \frac{(1 - \cos 2\,\omega t)^2}{16\pi d^2}$$

which shows that the pure alternating-current voltage will develop a pressure of twice the frequency.

If the same alternating-current voltage ($e_m \sin \omega t$) were applied superimposed on a fixed direct-current voltage, e_0, the pressure developed will be

$$P = \frac{(e_0 + e_m \sin \omega t)^2}{8\pi d^2} \qquad (38)$$

as E is now equal to $(e_0 + e_m \sin \omega t)$.

This equation when reduced to practical values, that is, E in volts, will be

$$P = \frac{8.85\, e_0 \times e}{d^2} 10^{-7} \qquad (39)$$

where $P =$ r.m.s. pressure in dynes per sq. cm.
$e_0 =$ direct-current polarizing potential in volts
$e =$ r.m.s. value of the alternating-current component in volts.

In the case of ribbon, uni-directional and dynamic microphones, the field calibration is the most reliable type of measurement, and can be made by placing the microphone in the field of a loud-speaker whose characteristics are known, and then measuring the output of the microphone for all frequencies impressed on the loud-speaker. Since microphones and their associated amplifiers are subjected to considerable handling when used on a motion picture production, it has been found necessary to make routine measurements at frequent intervals.

A reliable check on such equipment consists of the following pro-
A source of voltage from the measuring set is introduced by means

of a transformer, or attenuator, or combination of both, with a very

* *Applied Acoustics*, Olson and Massa, pages 225-226.

cedure:

small resistance in series with the microphone. In this way, continuity is established between the microphone and its associated amplifier which will indicate poor contact, high resistance leads or decreased gain in the microphone amplifier. In the case of the ribbon microphone, the inserted resistance should be not greater than 1/10 of an ohm, while in the dynamic type of microphone, which usually has internal impedance of approximately 30 ohms, as much as one or two ohms can be placed in series without change in characteristic.

In the making of gain-frequency responses of certain component parts of a system, it is necessary that the test frequency have a good wave form, because of the fact that when filters are connected in a circuit, their true loss will not be indicated if the test frequency contains harmonics. For example: If a high-pass filter (which normally has a sharp cut-off at 100 cycles) is being tested, and a 50 cycle frequency with harmonics in it which are not more than 10 db down at 100 cycles, is impressed on the circuit, the filter will never show more than 10 db discrimination between these two points. Also, if oscillators having a high harmonic content are used for testing light valves, it is possible to clash the valve well below the over-load point at its resonant frequency even though it is presumably being tested at a lower frequency.

6. INTERMODULATION

Since various parts of the recording circuit are not strictly linear, tests are necessary from time to time to determine the degree of non-linearity. These can be conveniently made by applying two frequencies, such as 1,000 and 1,100 cycles simultaneously. The sum and difference frequencies will then appear due to intermodulation, and a measure of this effect can be obtained by filtering out the impressed frequencies and measuring either the sum or difference frequency.

7. ATTENUATOR DESIGN TABLES

Attenuator design is not a particularly difficult subject, but involves extensive computation. In order to avoid this computation, a fairly complete table of resistive values for all of the attenuator values in use is presented.

A few words about the choice of an attenuator for a given purpose:

When working into a high impedance device such as the grid of a tube, it is usually permissible to use a potentiometer, which is the simplest form of an attenuator. Here $R_1 + R_2 = Z_1$ (see diagrams and notations below) and since this sum is constant, only one switch arm and set of contacts will be needed to make a variable attenuator.

When working between two equal impedances in a circuit, such that it is not necessary that the impedance looking back into the output

terminals of the attenuator be equal to the load impedance (probably the most common case), an *L* type attenuator may be used. Here the equation given above no longer holds, and if the attenuator is to be variable, a double set of contacts and two switch arms will be necessary. Finally, when the attenuator impedance must match the load impedance, as in mixers and certain types of amplifiers with critical input impedances, a *T* type attenuator must be used. *T* type attenuators may be either symmetrical or unsymmetrical, depending upon whether the generator and load impedances are the same. (Table I includes only symmetrical *T* type attenuators. Unsymmetrical types are discussed later.)

All of the attenuators in Table I are computed for 500 ohm circuits; attenuators for other impedances can be computed very easily from the values given by multiplying all of the resistances by the ratio of the value of the impedance under consideration to 500 ohms.

For example, suppose that a symmetrical attenuator is desired which will give a 10 db loss in a 200 ohm circuit. Two hundred divided by five hundred equals four-tenths. From the table it is found that a 10 db 500 ohm *T* attenuator has two series resistances of 260 ohms each and a shunt resistance of 253 ohms. Multiplying each of these values by 0.4, gives a new value of 104 ohms for the series arm and 141 ohms for the shunt arm.

No values have been given for *H* type attenuators since they are easily derived from the corresponding *T* attenuator. The procedure is to halve each series resistor and to place the other half in the opposite leg.

The formulae from which all of the resistances given may be computed are fairly simple, but this is not the case for attenuators to work between unequal impedances. Since we now have two variables instead of one (i.e., both attenuation and impedances ratio), any reasonably complete table would be bulky.

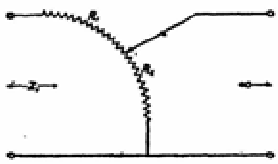
There is a minimum possible loss for which this type of attenuator can be constructed which depends upon impedance ratio. This loss varies from zero for a ratio of one, to infinity for a ratio of zero. See Chapter XXXIII for minimum loss attenuators of this type.

In conclusion, although all values have been given to three figures in the tables which follow, it is not necessary to hold to the actual values of the resistors given except in special cases where high accuracy is desired. A discrepancy of five per cent in any one resistance will cause an impedance mismatch of not more than that amount and a loss variation of only half a db.

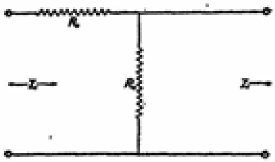
TABLE I

500 Ohm Attenuators

POTENTIOMETER

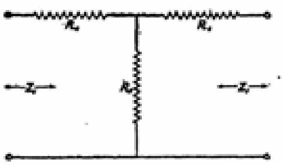


L TYPE

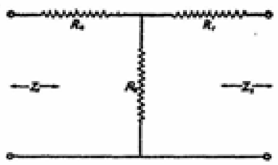


Attenuators

SYMMETRICAL



UNSYMMETRICAL



Loss, db.	R_1	R_2	R_3	R_4	R_5
1	54.5	445	4090	28.7	4340
2	103	397	1925	57.3	2160
3	146	354	1162	85.6	1420
4	185	315	855	113	1100
5	219	281	643	140	823
6	250	250	502	166	670
7	277	223	405	191	559
8	301	199	331	215	473
9	323	177	274	238	406
10	342	158	231	260	352
11	359	141	197	280	301
12	375	125	168	299	268
13	388	112	144	317	236
14	400	100	125	334	208
15	411	88.9	108	350	184
16	421	79.3	94.1	364	155
17	430	70.7	82.2	376	134
18	437	63.0	72.1	388	123
19	444	56.1	63.2	399	112
20	450	50.0	55.6	409	101
21	455	44.5	48.9	418	89.8
22	460	39.7	43.2	426	80.0
23	465	35.4	38.1	435	71.2
24	469	31.5	33.6	441	63.2
25	472	28.1	29.8	447	56.3
26	475	25.0	26.3	452	50.1
27	478	22.3	23.4	457	44.7
28	480	19.9	20.7	461	39.9
29	482	17.7	18.4	466	35.5
30	484	15.8	16.3	469	31.7
35	491	8.89	9.05	483	17.8
40	495	5.00	5.05	490	10.0
45	497	2.81	2.83	495	5.62
50	498	1.58	1.59	497	3.16
55	499	0.889	0.890	498	1.78
60	500	0.500	0.500	499	1.00

TABLE I (Continued)

Design Formulae

ATTENUATOR FORMULAE

POTENTIOMETER

$$R_1 = Z_1\ (1 - K)$$
$$R_2 = Z_1\ K$$

L ATTENUATOR

R_1 same as for potentiometer

$$R_3 = Z_1 \frac{K}{(1 - K)}$$

$$R_4 = Z_1 \frac{(1 - K)}{(1 + K)}$$

SYMMETRICAL *T* ATTENUATOR

$$R_5 = Z_1 \frac{2K}{(1 - K^2)}$$

$$R_6 = Z_1\ P_1 - P_2 \sqrt{Z_1\ Z_2}$$

UNSYMMETRICAL *T* ATTENUATOR

$$R_7 = Z_2\ P_1 - P_2 \sqrt{Z_1\ Z_2}$$

$$R_8 = P_2 \sqrt{Z_1\ Z_2}$$

The constants K, P_1 and P_2 in these formula are tabulated below. It should be noted that K is the voltage (or current) ratio corresponding to the given attenuation, and can be taken from any db-voltage ratio table or log table.

Loss, db	K	P_1	P_2
1	0.891	8.68	8.68
2	0.794	4.42	4.30
3	0.708	3.02	2.86
4	0.631	2.32	2.10
5	0.562	1.94	1.64
6	0.501	1.67	1.34
7	0.447	1.45	1.04
8	0.398	1.38	0.940
9	0.355	1.28	0.800
10	0.316	1.22	0.704
11	0.282	1.17	0.615
12	0.251	1.12	0.540
13	0.224	1.10	0.475
14	0.200	1.08	0.420
15	0.178	1.06	0.368
16	0.156	1.04	0.326
17	0.141	1.03	0.289
18	0.126	1.03	0.256
19	0.112	1.02	0.227
20	0.100	1.02	0.202
25	0.0562	1.00	0.112
30	0.0316	1.00	0.064
35	0.0178	1.00	0.036

—Data included in the above chart compiled by Dr. John F. Blackburn, Consulting Physicist, Hollywood, California.

Chapter XIII

PHASE DISTORTION

By JOHN K. HILLIARD

The composite waves of speech, music and sound effects are signals which have a wave composition quite different from those of single-frequency electrical waves. These complex compositions, which are ever changing in form at a very rapid rate (completely transient in character), are those parts of the wave train which have a bearing on the naturalness of reproduction. On the other hand, the waves show an indication of a definite steady state value after the transient part has subsided, and finally, again, we have a highly transient part of the wave beginning at the time the steady state is disturbed.

It is therefore necessary to retain all of the distinguishing qualities of a wave form to give natural reproduction of a sound. This can only be accomplished in an electro-acoustic system by making the frequency scale as wide as is commercially possible. This transient state transmission requires that a system accept steady state bands over an extremely wide frequency range.

In the reproduction of recorded material, certain distortion enters into the transmission to give imperfect results, which are due principally to three causes:

(1) The volume of the sound may be changed so as to be too loud or not loud enough.

(2) Amplitude distortion has taken place due to the fact that not all of the essential frequencies are transmitted with the same amplitude.

(3) Phase distortion is present to some extent in all recording systems. It is caused by the fact that the different frequencies composing the wave form travel with different velocities, such that their relative arrival times differ from their relative starting times. For undistorted transmission, the time of propagation of the various component frequencies must be the same.

To illustrate phase distortion; Figure 126 indicates the attenuation and delay characteristics of four band-pass filters in series. Note that the pass band is approximately 350 to 600 cycles. Now, in Figure 127, we see the oscillograms for several impulses of frequencies in the pass

and attenuation bands. Note that in each case there is a definite lag between the impulses sent and received, that the received signal in the pass band (at 480 cycles) is relatively undistorted, and that distortion

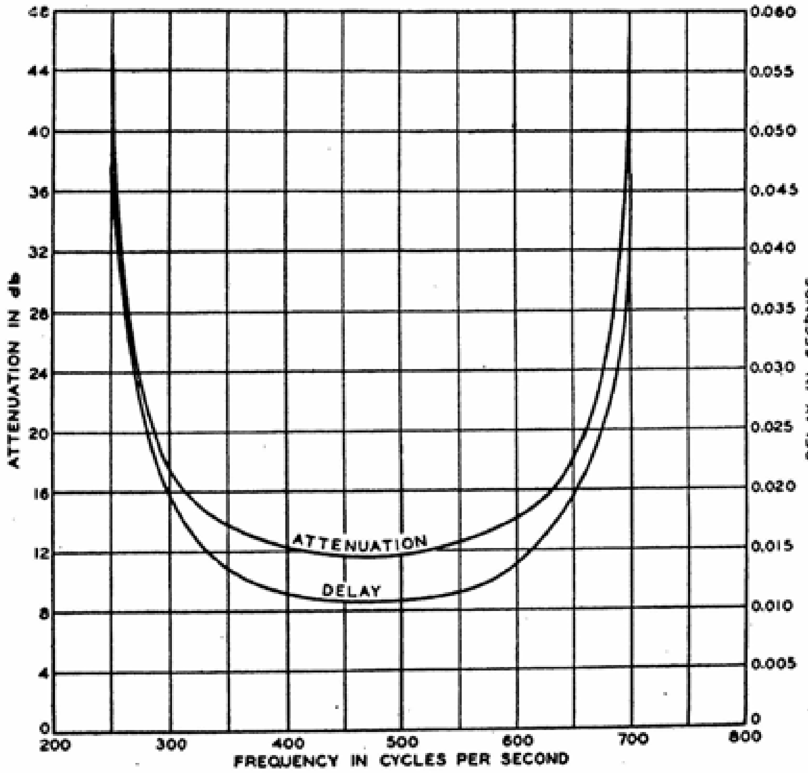


Figure 126 — Band-pass filter series — Attenuation and delay characteristics.

introduced near the edges of the pass band and in the attenuation regions is indicated by an increasing delay time in Figure 126. It will be observed that after the signal attains its steady state condition it is not distorted, and phase distortion is not present until this steady state is disturbed.

In general, we may say that the effect of phase distortion on speech is to decrease the intelligibility or to cause a loss of articulation. The effect on music is less noticeable due to the more sustained character of musical sound waves.

For transmission of transient waves without distortion, it has been shown that the time of propagation of component frequencies must be the same; this equality is obtained when the phase shift is linear with

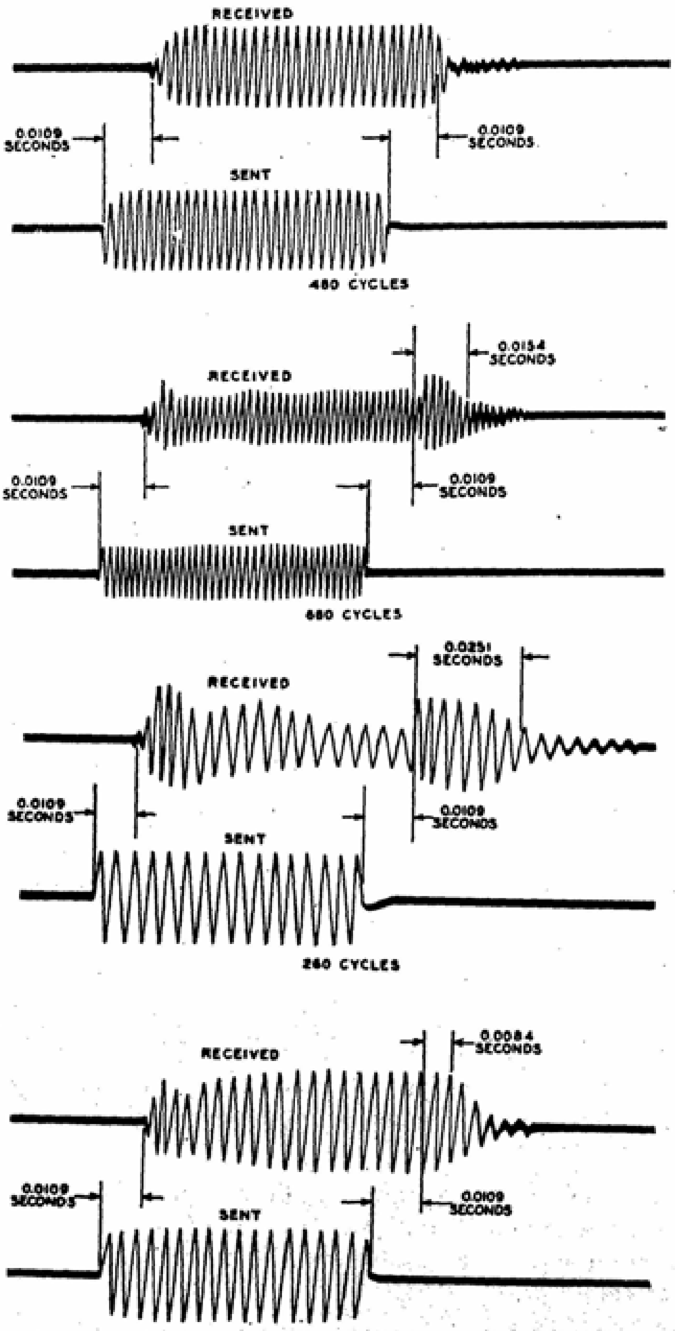


Figure 127 — Oscillograms — Frequencies of Figure 126.

frequency and equal to ($\pm n\pi$) when the frequency is zero, and n is equal to 0-1-2-3, etc. That is, when the phase shift is plotted as a function of frequency, the resulting graph should be a straight line. When a signal is applied to such a system, nothing will be received until a definite time period, after which period all waves will instantaneously assume their steady state values. The received signal will then be similar to the transmitted signal except for the *constant* phase shift introduced. This constant phase shift will not be detected by the ear. As phase changes in the component frequencies of steady state waves are not noticeable, we may assume that phase distortion occurs only in the transition periods.

Thus, a system may be linear in frequency response, as indicated by gain-frequency measurements (steady state), and yet introduce a distortion similar to non-linear systems under operating conditions.

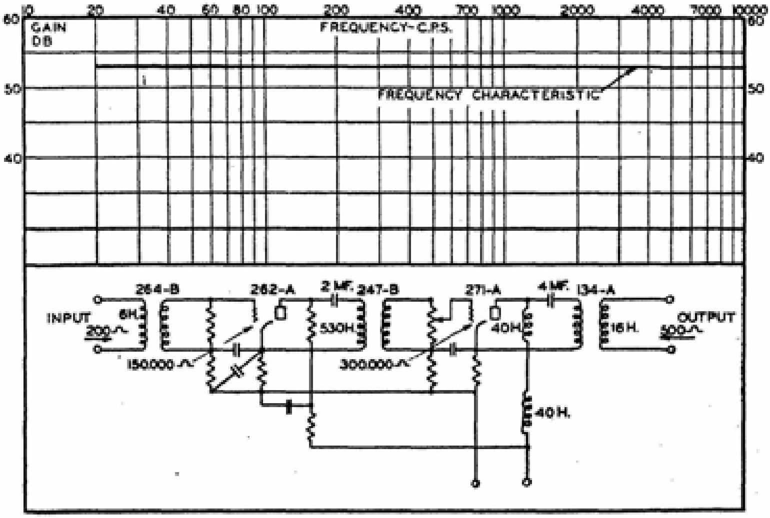
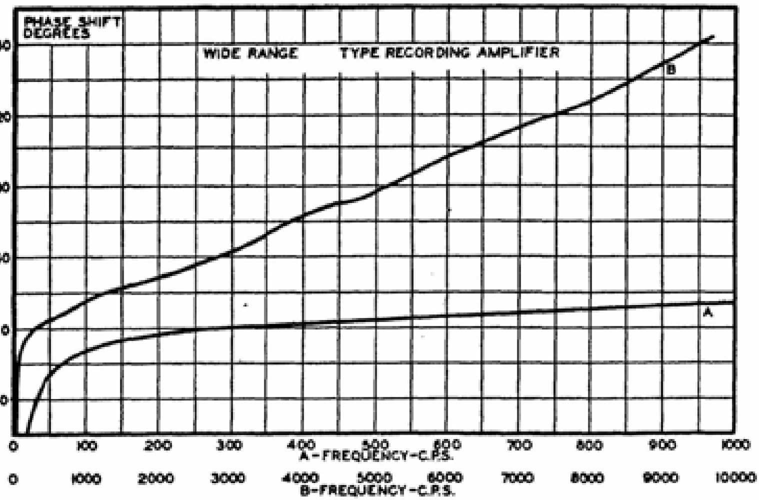


Figure 128 — Schematic diagram and frequency characteristic of a wide-range type recording amplifier.

Delay to high-frequency signals is more noticeable than to low frequencies due to the ear's dependence on the high frequencies for definition, etc. It has been determined that the maximum delay in the band from 5,000 to 8,000 c.p.s. should not exceed the 1,000 cycle delay time by more than 5 to 10 milliseconds. Also that at 50 c.p.s. the delay may be as much as 75 milliseconds more than the 1000 cycle value without noticeably affecting quality.

Both series capacity and shunt inductance delay low frequencies—the delay increasing as the values of such units are decreased. Conversely,

both series inductance and shunt capacity delay high frequencies—the delay increasing with inductance and capacity.



Upper scale for Curve A, lower scale for Curve B.

Figure 129 — Phase shift curve of Figure 128, 0 to 10,000 c.p.s.

However, we may say that in present high fidelity amplifiers, with their uniform frequency response, there is negligible delay distortion

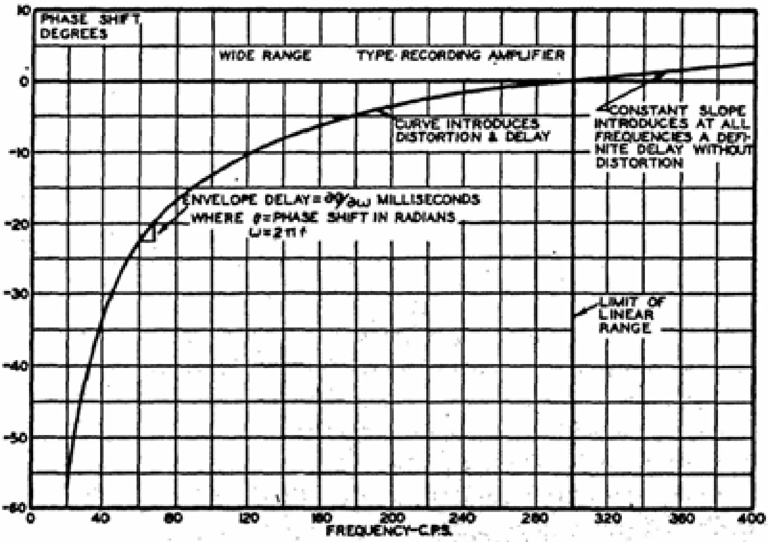
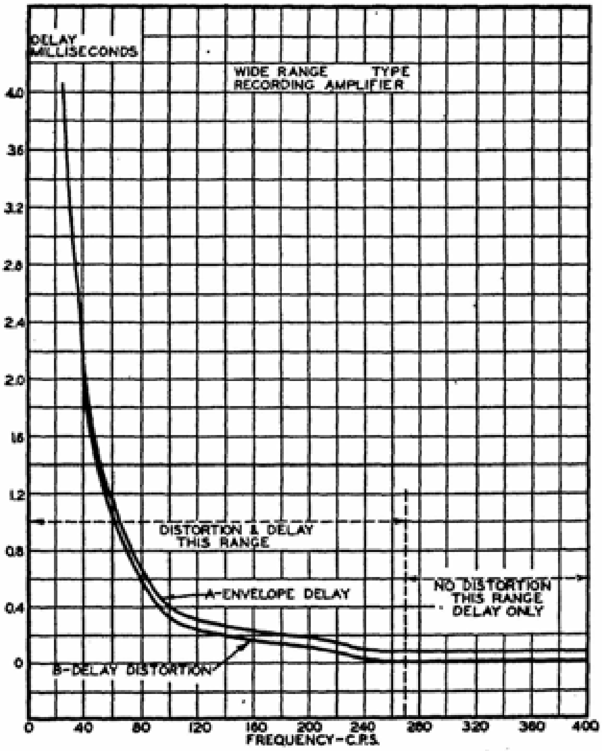


Figure 130 — Phase shift curve of Figure 128, 0 to 400 c.p.s.

in the relatively narrow frequency band employed. As an illustration we note in Figure 128 a wide range type recording amplifier with its fre-

quency characteristic. Observe the high values of shunt inductance and the large series condensers in this circuit. Now in Figure 129 the phase shift-frequency characteristic for this amplifier is seen to be relatively linear over the major portion of the frequency band. The slight irregularities or waves in the high-frequency portion are introduced by the measuring circuit, which will be described later. Observe that the devia-



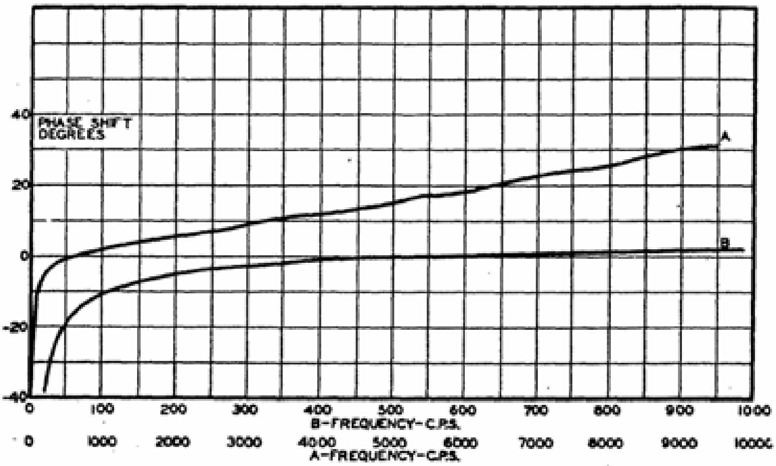
Envelope delay in seconds $= d\,\beta / d\,\omega$
where $\beta =$ phase shift in radians
$\omega = 2\,\pi\,f$

Delay distortion at any frequency = envelope delay at that frequency minus envelope delay at frequencies where phase shift-frequency relation is linear.

Figure 131 — Delay distortion curve of Figure 128 from Figure 130.

tion from linearity is confined to the lower frequencies, which low-frequency section is shown in an enlarged scale in the lower curve. A further enlargement in Figure 130, which will be used later in the calculation of distortion delay, indicates that for all practical purposes we may consider the phase shift to be linear with frequency at all frequencies above 300 c.p.s. In Figure 131 we see the actual delay distortion in this range, below 300 c.p.s. Note that the delay distortion at 50 c.p.s.

is 1.4 milliseconds, and as we have previously stated that a delay distortion of 75 milliseconds at 50 c.p.s. is not noticeable, we see that the value of 1.4 milliseconds is negligible. This figure will be discussed in detail as we proceed to the actual measurement of delay distortion. Another phase shift curve is shown in Figure 132, which is the characteristic of a standard bridging amplifier. Note that the phase shift at 20 cycles with respect to 300 cycles is 35 degrees in contrast to 58 degrees for the wide-range type amplifier of Figure 128.



Upper scale for Curve B, lower scale for Curve A.

Figure 132 — Phase shift curve of bridging amplifier.

Proceeding to the theory of phase distortion we have the following proposition: "For an applied signal the *envelope of the oscillations* in response to an e.m.f. $-E\cos\omega t$, applied at time $t = 0$, reaches 50% of its ultimate steady state value at time, $t = \frac{d\beta}{d\omega}$, and its rate of building up is inversely proportional to $\sqrt{\frac{d^2\beta}{d\omega^2}}$. The quantity $\left(\frac{d\beta}{d\omega}\right)$ is defined as the envelope delay of a system in the range where the attenuation is not a function of frequency; that is, for a distortionless system with respect to frequency characteristic this quantity is the actual delay of the signal. Envelope delay is determined from the difference in the steady state phase shift for a definite interval of frequency, i.e., the received wave is delayed by a time interval given by the slope of the phase characteristic.

$$\text{Envelope delay} = t = \frac{d\beta}{d\omega} \tag{40}$$

where t = delay in seconds
β = phase shift in radians
$\omega = 2\pi f$

Thus, referring again to Figure 130, we select a small change in frequency, f, and the corresponding change in phase shift is converted to β radians, and these two values are substituted in the formula. Solving the formula we obtain the time delay t for an intermediate frequency which we may assume to be $\left(\frac{f_1 + f_2}{2}\right)$. Proceeding in this manner we obtain the upper curve in Figure 131, which is the envelope delay. Now we note that above 300 c.p.s. the delay is constant and as it is not the actual value of delay, but the relative value from one frequency to another that is important, we subtract this constant delay from curve A, resulting in curve B, called the distortion delay curve. It should be noted that these results are in agreement with the statement that in order to have no phase distortion the phase shift must be linear with frequency (slope of phase shift curve constant) in order for the component frequencies of the signal to reach their approximate steady state condition in the same interval of time.

The major portion of phase shift in recording systems is due to the presence of filters. Delay distortion varies with the type of filter, number of sections and sharpness of cut-off; a sharp cut-off increasing the distortion in the pass band without greatly increasing the minimum delay.

We may briefly summarize the effects of different filter types as follows:

(1) In low-pass constant K filters, the phase shift increases with frequency from zero degrees at zero frequency to π radians at the cut-off frequency. The phase shift then remains constant above the cut-off frequency.

(2) In high-pass constant K filters the phase shift is constant at $(-\pi)$ radians for all frequencies below the cut-off and gradually passes from $(-\pi)$ radians to zero at infinite frequency.

In constant K type filters the phase shift is constant only in the attenuation bands and is changing with frequency over the entire transmission band. For full sections the phase shift in the attenuation band is always $(\pm n\pi)$ radians where n equals an integer. For half-sections the phase shift is one-half that of the full sections.

(3) In m-derived filters the phase shift slope in the pass band is usually less than for constant K types and the curvature is more pronounced near the edge of the attenuation band. The phase shift is not

always ($\pm \pi$) radians per section as for the constant K filter, but has this value only between the cut-off frequency and the frequency of infinite attenuation. The phase shift is zero for all other frequencies in the attenuation band.

(4) Phase shift characteristics of band-pass filters vary widely with the type of filter, but in general the amount of phase distortion is inversely proportional to the band width and is independent of the position of this band in the frequency scale. Conditions at the low-frequency end are similar to high-pass filters and at the high-frequency end similar to low-pass filters. If the straight portion of the band-pass filter phase shift curve is extended, it may intersect the phase shift axis at any point, where in low-pass filters the intersection is always at ($n\pi$) radians.

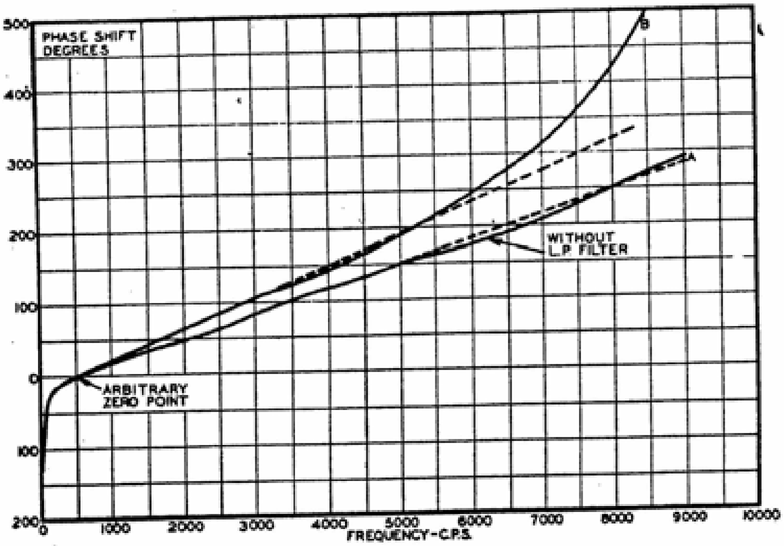


Figure 133 — Mixer + pre-amplifier + wide-range type amplifier + bridging amplifier + low-pass filter — phase shift curves.

Figure 133, Curve A, shows the phase shift characteristic of a series of amplifiers, the average value over the major portion of the frequency range being linear as indicated by the dotted line. In Curve B we see the curvature introduced by the addition of an m-derived low-pass filter of 7,500 cycles cut-off frequency. Note the deviation from linearity, also the increased slope indicating greater initial delay. In Figure 134 we see the frequency characteristic of this system, showing the sharp cut-off and rapid increase in attenuation toward the frequency of infinite attenuation.

The phase distortion data in all figures except Figures 126 and 127 were obtained from measurements made by the following method:

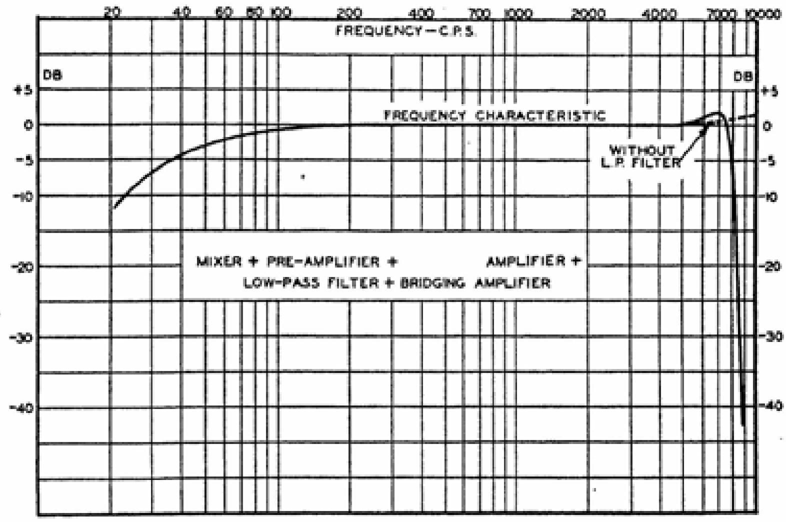


Figure 134 — Frequency characteristic of Figure 133.

Distortion delay data may be obtained from phase shift measurements, impedance measurements or direct measurements of envelope de-

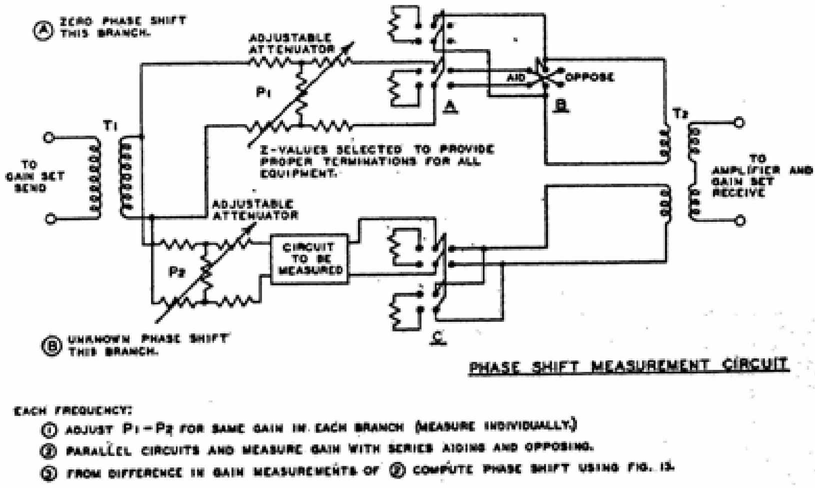


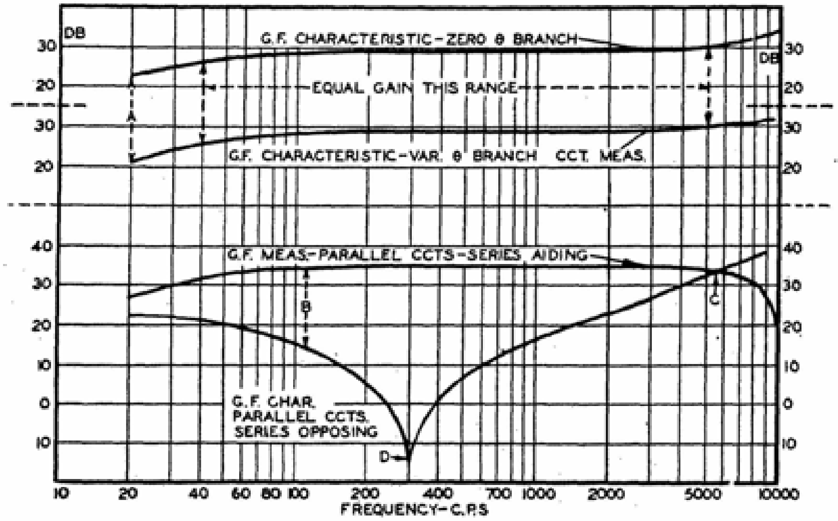
Figure 135 — Schematic diagram of phase shift measurement circuit.

lay. The phase shift method to be described is applicable to standard test equipment, will indicate small amounts of phase distortions, and is suffi-

ciently accurate for all practical purposes. This particular method of measuring phase shift is indicated schematically in Figure 135, and fundamentally consists of two parallel circuits, one with zero phase shift and the other with a varying and unknown phase shift. If, at any given frequency, we adjust the attenuators P_1 and P_2 so that the output currents at T_2 and consequently at the output of the common amplifier, are equal when the two circuits are used individually, we then have a simple condition of the vector sum and difference of the two equal currents when the circuits are paralleled and the switch B operated to series aiding or opposing. The difference between the vector sum and difference currents is indicated in db by the amount of attenuation necessary in the common circuit for equal output currents in this circuit in the "aid" and "oppose" position of the switch.

Summarized, the operations for measurement of phase shift, with reference to Figure 135, follow:

(1) Measure gain-frequency characteristic of branch B with branch A terminated.



(A) = Increased attenuation of zero β branch 1.5 db.

(B) = Difference in curves used with Figure 137 to determine phase shift.

(C) = Gain measurement the same for series aiding or opposing and phase shift equals 90°, 270°, 450°, etc., when vector sum equals vector difference.

(D) = Infinite attenuation in series opposing measurement when currents are 180° out-of-phase and phase shift equals 0°, 180°, 360°, etc.

Figure 136 — Measured curves of Figure 128 by circuit of Figure 135.

(2) Terminate circuit A and adjust attenuator P_1 so that the gain of circuit A over the flat portion of the curve is the same as the gain of

circuit B over a similar range, then proceed with gain-frequency measurement of circuit A.

(3) From these curves of A and B compute the variations in attenuation at $P_1 - P_2$, necessary to maintain equality in output currents (gain-frequency curves) in circuits A and B.

(4) Parallel circuits A and B and with switch B in "aid" position, measure gain-frequency characteristic of combination, adjusting attenuators as noted in (3), as required to maintain same transmission characteristics in both branches.

(5) Operate switch B to "oppose" and proceed with gain-frequency measurements of parallel circuits as in (4) [see note (1)].

(6) The resulting two gain-frequency curves will be similar to Figure 136, which is a measurement, made as outlined above, of the wide-range type recording amplifier of Figure 128. From these curves determine the difference in db between the aid and oppose condition at sufficient points to plot the phase shift in a degrees-frequency in c.p.s. curve using the chart in Figure 137 to convert db to degrees.

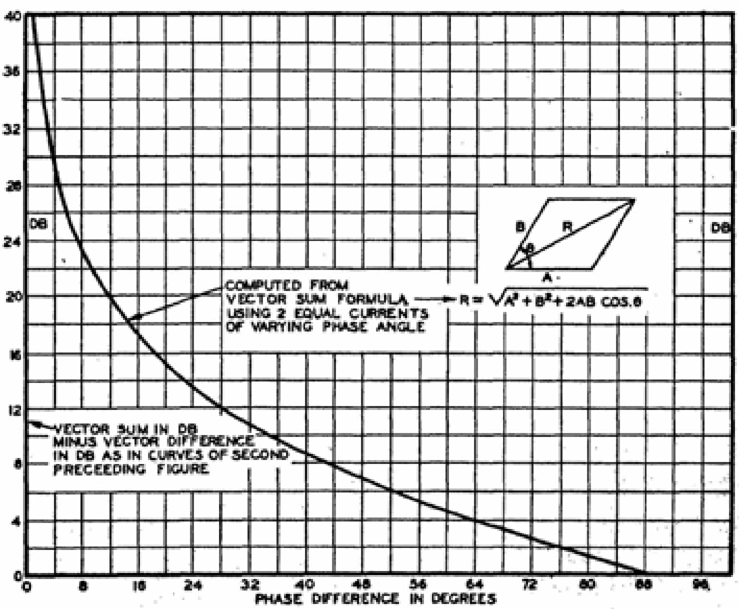


Figure 137 — Conversion chart-db/degrees for phase shift measurements.

NOTE 1:

An initial adjustment of the variable attenuators, with the circuits in the series opposing connection, and at the 0° — 180° etc., point, in

fractional db amounts, will facilitate attainment of sharp minimum values.

Again, in Figure 129 we show the resulting phase shift-frequency curve resulting from (6) in the preceding summary.

In Figure 138 note the frequency response of circuits A and B of Figure 135 when each circuit has "zero" phase shift. Note that the minimum difference in db is 44.5; referring to Figure 137, we see that this value corresponds to a phase shift of less than one degree, indicating that for practical measurements this particular circuit is satisfactory.

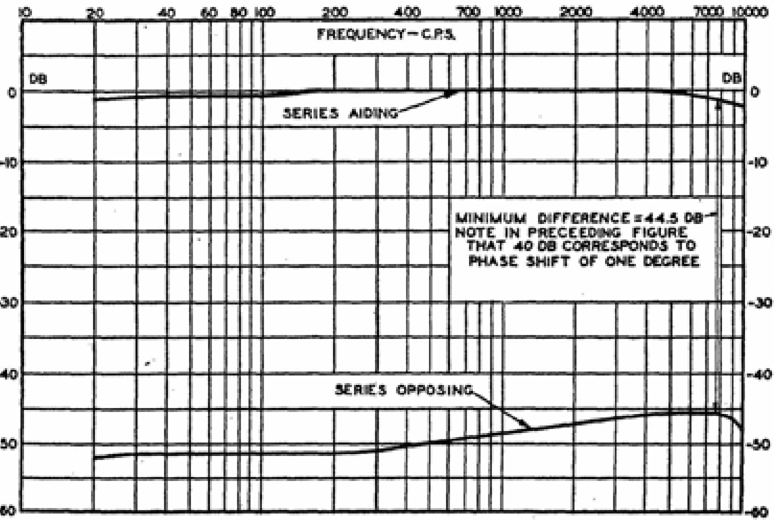


Figure 138 — Frequency response of Figure 135 with zero phase shift in each branch of parallel circuits.

While phase distortion is important in recording systems, it is not usually compensated for. If the frequency band is extended, as in television systems, or if a large number of amplifiers or filters are used in a circuit such as in a telephone system, such distortion becomes serious and must be deliberately corrected.

Correction of phase distortion is usually made by insertion of "all-pass" networks, having the property of phase shift without frequency attenuation. All-pass networks are described in detail in standard textbooks on transmission networks.

Acknowledgment is made to the *Bell System Technical Journal* for information and illustrations included in this chapter.

The above measurements were made by G. M. Sprague of Metro-Goldwyn-Mayer Studio Sound Department.

Chapter XIV

TRANSFORMERS FOR SOUND CIRCUITS

By JOHN K. HILLIARD

In power work, we are interested in changing or transforming voltages in a narrow low-frequency band from one value to another in order to effect economies in transmission, while in sound recording we are interested in transforming impedances from one value to another to obtain optimum efficiencies of transmission over a wide band of frequencies.

An impedance transforming device is known as an audio-frequency transformer, and in its ideal form modifies the magnitude of the load impedance to match it to the generator impedance. An audio-frequency transformer is similar to a power transformer except that it must operate over a wide band of frequencies instead of at a single frequency.

The ideal transformer neither stores nor dissipates energy, but in order to be ideal, a transformer would have to meet the following three requirements:

(1) Have infinite primary and secondary inductance, but a finite ratio of primary to secondary inductance.

(2) Have perfect coupling between primary and secondary.

(3) Have no resistance in primary and secondary windings.

In practice, these conditions are impossible to obtain, yet for the frequencies we use an approximation can be reached to the end that a transformer will transmit uniformly the band we require without appreciable loss:

The first condition is satisfied by using cores of extremely high permeability and a great number of turns on the coil;

The second condition is approximated by arranging the coils mechanically so that the leakage flux between turns is very small;

While the third condition is satisfied by the fact that for audio-frequencies the resistances of the windings are small compared to their impedance in the frequency ranges under consideration.

In an ideal transformer the ratio of primary to secondary voltage is equal to the ratio of the turns and hence the ratio of the currents is the

reciprocal of the voltage ratio, since there is no energy dissipated. To reflect the constants of the secondary side of the transformer to the primary, it is necessary to multiply the impedances of the primary by the ratio of secondary inductance to primary inductance or $\left(\frac{L_1}{L_2}\right)$.

For all practical purposes, with good iron core transformers the ratio of

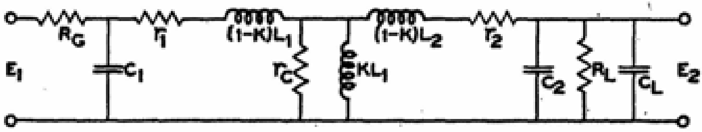
$$\frac{L_1}{L_2}=\left[\frac{N_1}{N_2}\right]^2 \tag{41}$$

also

$$\frac{E_1}{E_2}=\frac{N_1}{N_2}\text{ and }\frac{I_1}{I_2}=\frac{N_2}{N_1} \tag{42}$$

It is now possible to set up the formula for an equivalent circuit for an audio transformer, since in the practical transformer we have resistance in the coils which is the equivalent of a series loss, distributed capacities which represent shunt reactances on both primary and secondary, and losses in the magnetic circuit which look like resistance shunting the winding.

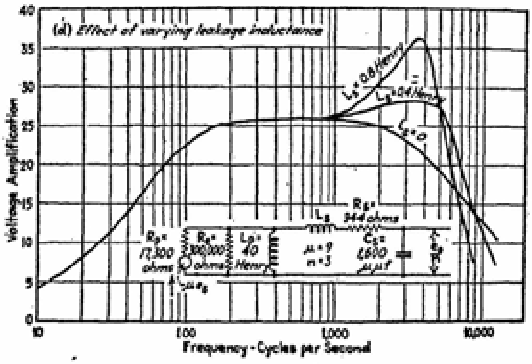
The circuit is as follows:



C_1 = distributed capacities of primary
C_2 = distributed capacities of secondary
C_L = distributed capacity of the load
K = coefficient of coupling
L_1 = inductance of primary
L_2 = inductance of secondary
r_1 = resistance of primary
r_2 = resistance of secondary
r_C = exciting current loss (core loss)
R_G = generator resistance
R_L = load resistance

Figure 139 — Equivalent circuit of audio transformer.

Since audio transformers are used to transmit a wide band of frequencies, we find that it is the extreme ends of the band which become the dominating factors. If we consider the transformer as acting between frequencies f_1 and f_2 as limits, then the mean frequency $f_M = \sqrt{f_1 f_2}$. We find that in the lower part of the band the leakage re-



—From "Radio Engineering," by Frederick E. Terman (McGraw-Hill Book Co., Inc.).

Figure 140.

actance $(1 - K)\ L_1$ (Figures 140 and 141), becomes very small as compared to the impedance of the primary. Also, the distributed capacities have a very high reactance as compared to the inductive reactance, and so the coil approximates the equivalent network of Figure 141.

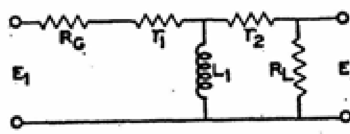


Figure 141 — Equivalent network of an audio-frequency transformer operating in the lower-frequency band.

And hence the ratio of

$$\frac{E_2}{E_1} = N \frac{R_L}{R_G + r_1 + r_2 + R_L + \dfrac{(R_G + r_1)\ (R_L + r_2)}{j\omega L_1}} \qquad (43)$$

At frequencies above the mean, the primary reactance (Figure 143) is so high that its shunting action is very small compared to the reactance of the distributed capacities, but at this point the leakage reactance must be considered as it starts to exert an appreciable effect.

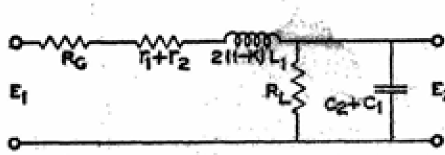
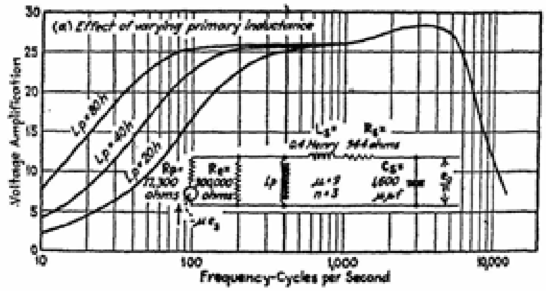


Figure 142 — Equivalent network of an audio-frequency transformer operating in the upper-frequency band.

This now gives the equivalent circuit of Figure 142.

Ordinarily an input transformer works into the grid of a vacuum tube, and the shunting effect of this high load impedance can be

neglected at the higher frequencies, so that the equivalent circuit be-



—From "Radio Engineering," by Frederick E. Terman (McGraw-Hill Book Co., Inc.).

Figure 143.

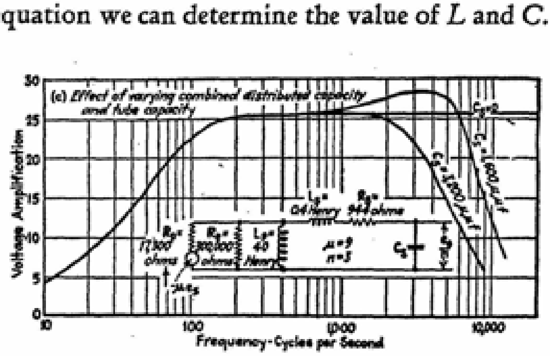
comes a series resonant circuit, and

$$\sqrt{L_e C} = \frac{1}{2\pi f_o} \qquad (44)$$

where $L_e = 2\ (1 - K)\ L_1$

$f_o =$ the resonant frequency

From this equation we can determine the value of L and C.



—From "Radio Engineering," by Frederick E. Terman (McGraw-Hill Book Co., Inc.).

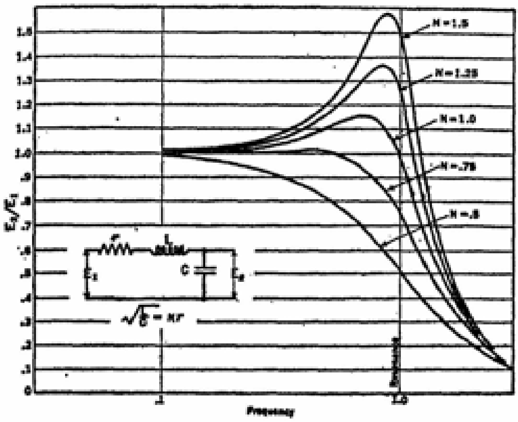
Figure 144.

If the amplification is to be uniform at high frequencies, f_o must be past the limit of the band. In input transformers it is desirable to obtain the maximum possible turn ratio, which increases both L_e and C. Hence, both quantities must be kept very low. In order to keep the amplitude at resonance low, resistance wire may be used to wind the secondary, or a very high resistance may be inserted in the circuit.

Since leakage reactance varies with N^2, and the turns ratio must be

high, there are two conflicting demands, one for many turns and the other for few.

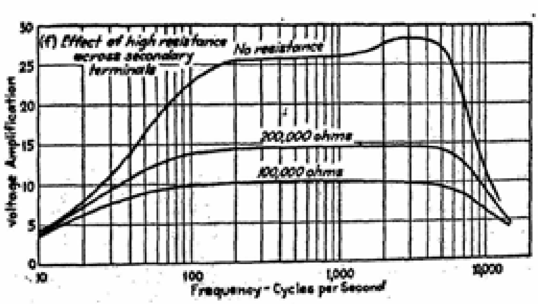
The output transformer differs from the input and inter-stage transformer in that it is a stepdown transformer, and the impedances between which it works are usually considerably less, which makes



—*Reprinted by permission from "Electrical Engineers' Handbook," by Pender and McIlwain, published by John Wiley & Sons, Inc.*

Figure 145 — Curves showing leakage resonance for different ratios of inductance, capacity, and resistance.

possible the use of a proportionately large number of turns without introducing the inherent difficulties of high leakage reactances and distributed capacity. Consequently, if weight is of no particular importance, the core material may be of the lower permeability grade (Figure 148).

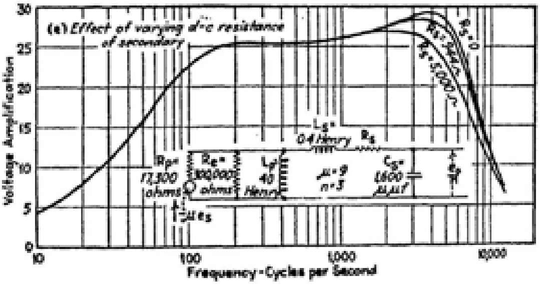


—*From "Radio Engineering," by Frederick E. Terman (McGraw-Hill Book Co., Inc.).*

Figure 146.

Improvements in frequency response during the past few years have required marked changes in transformers, some of which have

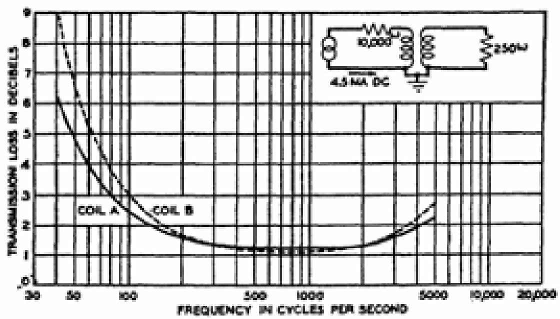
been in the application of new magnetic core materials, such as permalloy and hypernik (iron-nickel alloys having from 40-90% nickel).



—From "Radio Engineering," by Frederick E. Terman (McGraw-Hill Book Co., Inc.).

Figure 147.

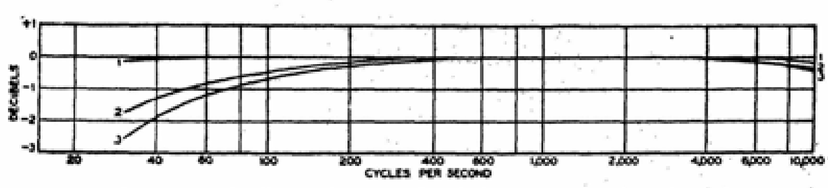
As a result, extremely high permeabilities are obtained with very small weight and loss distortion over wider frequency bands (Figure



—From the Bell System Technical Journal.

Figure 148 — Curves showing the transmission loss of (A) small output transformer, and (B) larger output transformer.

149). Reduced phase distortion is obtained since the inductance of the windings is much greater than that required for the proper trans-



Curves showing the effect of different materials on the frequency response of an audio transformer

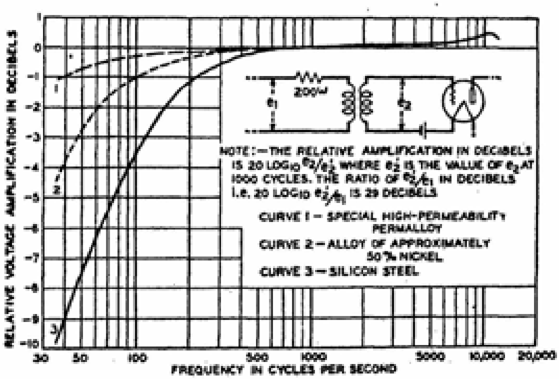
1—Alloy containing 45-50 per cent nickel 2—Experimental iron 3—Transformer steel, audio grade

—From a technical paper in Electrical Engineering, Volume 54, 1935, "Silicon Steel and Communication Equipment," by C. H. Crawford and E. J. Thomas of the General Electric Company.

Figure 149.

mission loss. Also, in most transformers, a certain amount of magnetic modulation takes place due to a magnetic non-linearity, usually indicated by harmonics of the lower frequencies occurring in the high range. This energy at the low frequencies is very great, as compared to that at very high frequencies, and for this reason, the harmonics of the low frequencies may be of the same order as the high-frequency signal.

The audio-frequency transformer differs from the power transformer in that the latter operates from a voltage source of good regulation. The audio transformer, however, operates from a voltage source of poor regulation, varying widely in amplitude and frequency range. As a result, the flux densities vary from negligible values at 10,000 cycles to approximately 5,000 lines per square centimeter at 30 cycles. Under these conditions, it is necessary to provide sufficient impedance in order that a constant voltage amplitude ratio be maintained over the operating range. High-permeability steel permits small core size, which



—*From the Bell System Technical Journal.*

Figure 150 — Showing relative effect of various core materials on the transmission characteristics of portable recording apparatus.

in turn gives lower distributed capacity and lower leakage inductance, both of which depend upon the geometry of the coil.

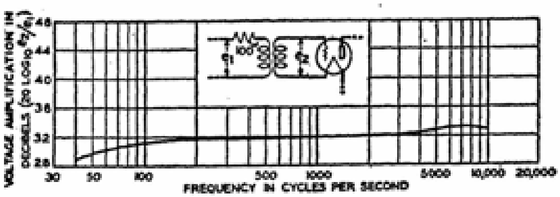
Since this is not detected in ordinary measurements, harmonic analyses must be resorted to.

The most efficient use of filters and equalizers requires constant impedance inputs or terminations. Better impedance characteristics of transformers are obtained by increasing the mutual impedance, and decreasing the leakage and capacity reactance variations.

In measuring the performance of a transformer, care must be taken to make sure that the conditions under which it is to operate are reproduced in the testing circuit, and precautions must be observed in the

tests in order that a misleading performance characteristic will not be obtained.

The permeability of magnetic core materials tends to rise rapidly from its initial values with increasing current. For this reason, if tests

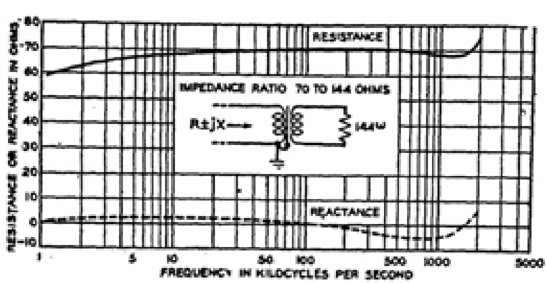


—*From the Bell System Technical Journal.*

Figure 151 — Transmission characteristic of light weight input transformer having a weight of 3½ ounces.

were made with low frequencies at considerably higher currents than are obtained under the operating condition, the measured low-frequency characteristic would indicate a much better response than that which would be obtained under service conditions.

Considerable demand has been created for portable apparatus, both for recording and testing, and as a result, we now have extremely light weight transformers of very small size. The gain-frequency characteristic, as shown in Figure 151, indicates that it is possible to construct these small transformers for specific uses with efficiencies which are comparable to, or better than, the larger type transformer.



—*From the Bell System Technical Journal.*

Figure 152 — Impedance-frequency characteristic of a well designed transformer.

Transformer Problems

1. *Given:*

Transformer step-up ratio of 1 to 5
$Z_1 = 500$ ohms
$I_1 = 0.01$ amps.
$E_1 = 10$ volts
$K =$ constant of coupling $= 1$

Find:

Z_2, I_2, and E_2

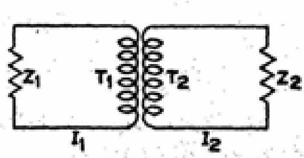


Figure 153 — Circuit diagram for transformer problems.

Circuit diagram (Figure 153):

$$\frac{Z_1}{Z_2}=\frac{N_1^2}{N_2^2};\ \frac{500}{Z_2}=\frac{1}{25};\ Z_2=500\times 25=12{,}500 \text{ ohms}$$

$$\frac{I_1}{I_2}=\frac{N_2}{N_1};\ \frac{0.01}{I_2}=\frac{5}{1};\ I_2=\frac{0.01}{5}=0.002 \text{ amps}$$

$$\frac{E_1}{E_2}=\frac{N_1}{N_2};\ \frac{10}{E_2}=\frac{1}{5};\ E_2=5\times 10=50 \text{ volts}$$

2. If the input impedance to a transformer is 200 ohms, and the load impedance is 20,000 ohms, what is the turns ratio of the transformer?

$$\frac{Z_1}{Z_2}=\frac{N_1^2}{N_2^2};\ \frac{N_1}{N_2}=T \text{ (turns ratio)}$$

$$T^2=\frac{Z_1}{Z_2}=\frac{200}{20{,}000}=\frac{1}{100}$$

$$T=\frac{1}{10}$$

3. If the plate resistance of a vacuum tube is 10,000 ohms, what must be the impedance and inductance of the primary of a transformer, coupled directly to the tube, in order that the tube output will be attenuated 3 db at 30 cycles?

$$20\log\frac{E_G}{E_L}=3 \text{ db* } (45)$$

$$\log\frac{E_G}{E_L}=0.15$$

$$\frac{E_G}{E_L}=1.413$$

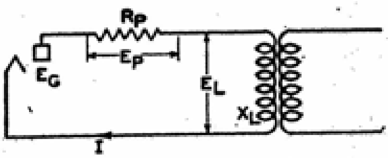


Figure 154 — Circuit diagram for problem 3.

$$\text{so}\quad \frac{E_L}{E_G}=0.707$$

Where E_L = voltage across the load

E_G = theoretical output voltage of the tube (considering the plate to filament resistance as external to the tube)

* [Formula (45) can be found directly from the Table, Chapter XXXII.]

$$\frac{E_L}{E_G} = 0.707 = \sin 45°$$

$$\therefore \quad \text{phase angle} = 45°$$

$$\tan 45° = 1 = \frac{E_L}{E_P}$$

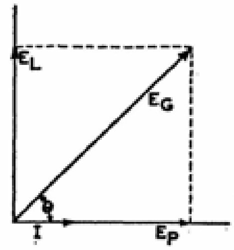


Figure 155 — Vector diagram

and $$\frac{E_L}{X_L} = \frac{E_P}{R_P} \text{ so } \frac{X_L}{R_P} = 1 \text{ and } X_L = R_P$$

$$X_L = 10{,}000 = 2\pi fL = 2\pi \times 30 \times L$$

$$L = \frac{10{,}000}{2\pi \times 30} = \frac{10{,}000}{188.4}$$

$$L = 53 \text{ henries}$$

Chapter XV

GENERAL NETWORK THEORY

By HARRY KIMBALL

1. NATURE OF NETWORKS

In the several fields of electrical engineering the term "transmission network" is used rather loosely to refer to a variety of types of electrical circuits. For instance, to the power engineer, a transmission network may mean an interconnected arrangement of transmission lines employed for the purpose of transmitting electrical power from point to point. Another use of the term, as derived from the communication field, is in reference to circuits which can be obtained from the connecting together of resistances, condensers, and inductance coils including transformers, to form equipments for use primarily as links of systems, whose function is to transmit electrically, acoustic or other signals conveying intelligence. These networks necessarily have four terminals, two for connecting to a transmission system to receive power, and two for connecting to a load to deliver power. For this reason, they are sometimes known as "four-terminal networks." The resistances, condensers, and inductance coils contained within such a network are known as the "network elements" and their electrical values are termed the "network constants."

It will be seen later that the internal structure of a network naturally divides itself into groups of electrical elements which form the branches or arms of the network. In the limiting case, these groups, may, and often do, consist of only one element. It is usually simpler in working out the mathematical relations of a network to deal with these natural groups as units rather than the individual elements. Such groups, whether consisting of one or more elements, require two accessible terminals for connecting them into the network structure. For this reason they are often known as "two-terminal networks," that is, two-terminal networks are the groups of elements from which four-terminal networks are made. A knowledge of the impedance characteristics of two-terminal networks is an important prerequisite in studying four-terminal networks. Certain fundamental information concerning these networks is therefore presented in the following work prior to the discussion of four-terminal networks.

For the networks which may be constructed in the above manner, there exist several well defined types designated in accordance with the duties they perform when inserted into a transmission facility. These networks are attenuators, attenuation equalizers, electric wave filters, and phase correctors or delay networks defined in the following material.

ATTENUATORS—An attenuator is a network whose insertion loss remains constant with change of frequency. Departure from constant attenuation occurs only when the design limitations of the component electrical elements are exceeded. The design and use of attenuators is discussed in Chapter XXXIII.

ATTENUATION EQUALIZERS—An attenuation equalizer is a network whose insertion loss varies in some desirable manner with change of frequency. Such networks usually are inserted into transmission systems to compensate for defects in the transmission-frequency characteristics of the systems. These networks are discussed in detail in the following chapter.

ELECTRIC WAVE FILTERS—An electric wave filter is a network which transmits without appreciable attenuation all frequencies within one or more frequency bands and attenuates all other frequencies. In practice, very few wave filters have more than one transmission band. Wave filters are discussed in a following chapter.

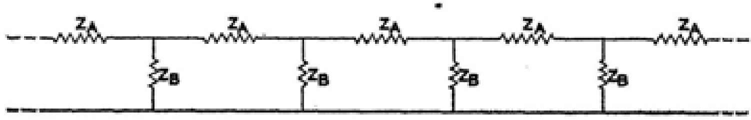


Figure 156-A — Ladder-type circuit.

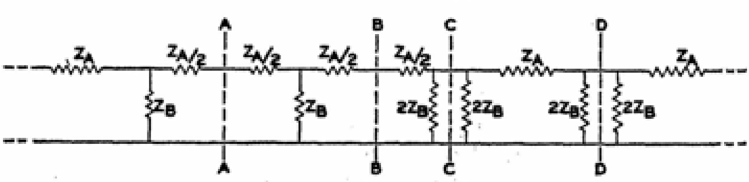
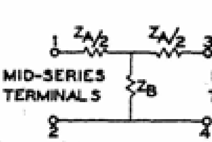
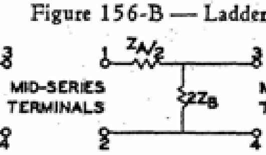


Figure 156-B — Ladder-type circuit.





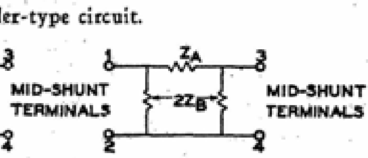


Figure 156-C — *T* type section.

Figure 156-D — *L* type section.

Figure 156-E — π type section.

Ladder-type sections.

PHASE CORRECTORS—A phase corrector is a network having the property of changing in a preassigned manner the relative phase rela-

tion of the frequencies of a given transmission band. Except for the unavoidable dissipation of the electrical elements, these networks do not attenuate the frequencies of their transmission range. Phase correctors are used with transmission systems to correct for system phase distortion similar to the manner in which attenuation equalizers correct for attenuation distortion. These networks are not within the scope of this material. (See Chapter XIII.)

It has been found that the engineering of networks is considerably facilitated by the use of distinctive unit structures operated in tandem to form configurations which, in many respects, resemble a carpenter's ladder. Many of the symbols and conceptions used in network theory thus have their origin in these "ladder-type" networks. For instance, referring to Figure 156-A, the circuit shown is a ladder-type structure having series impedances denoted by the symbol Z_A and shunt impedances denoted by Z_B. If we arrange this circuit as shown in Figure 156-B, and remove a portion of the network by cutting through two adjacent series arms at their mid-points, as shown by "AA" and "BB," we obtain the familiar *T* type section of Figure 156-C. The reason then for designating the series arms of this section by the symbol $Z_A/2$ and the shunt arm by the symbol Z_B is thus made clear. Also, the origin of the term "mid-series terminals" when referring to either of the pairs of terminals of a *T* section is evident. In a similar way, if we remove another portion of the ladder-type network by cutting through two adjacent shunt arms at their mid-points in the manner shown by "CC" and "DD," we obtain the π type section of Figure 156-E. Following the same method of designation, the series arm impedance of this section is denoted by the symbol Z_A, the shunt arms by the symbol $2\,Z_B$, and the two pairs of terminals are called "mid-shunt" terminals. Another familiar ladder-type section which can be obtained by removing a portion of the general ladder structure is the *L* type section of Figure 156-D. This section has a series arm of $Z_A/2$ and a shunt arm of $2\,Z_B$. The 1-2 terminals of the *L* section of Figure 156-D are the mid-series terminals and the 3-4 terminals are the mid-shunt terminals. *L* type sections are often referred to as half sections for the reason that two *L* sections connected together form one *T* or one π type section depending upon whether they are joined at their mid-shunt or mid-series terminals.

Transmission systems employing networks usually are arranged so that the impedances connected to the input and output terminals of the networks are resistances. That is, the impedance looking back into the system at the network input terminals is chiefly resistive in character as is also the load impedance connected to the output terminals.

The symbol R_G is used to designate the sending-end system impedance connected to a network and R_L denotes the system load impedance. Where R_G and R_L are equal they are designated in the following work by the symbol R_0. For proper performance, wave filters require that R_G and R_L be equal resistances unless the filters have impedance transforming designs. In the case of attenuation equalizers, R_G and R_L need not be equal resistances for proper performance although they are usually arranged to be equal.

For each of the types of four-terminal networks mentioned above there exists an orderly procedure of design made available as the result of a vast amount of theoretical work coupled with practical experience. These design procedures are, of course, interrelated, the methods employed for one type of network having much in common with those used for other types. It is desirable in taking up the study of networks to consider first a certain amount of "ground material" useful for any type of network. The remainder of the material in this chapter is mostly of this nature.

2. TWO-TERMINAL REACTIVE NETWORKS

Certain two-terminal reactive networks, that is, circuits not containing resistance elements, are used so frequently in the design of equalizers and filters that it is helpful to arrange the formulae for their reactances into forms especially suited to this type of work. Usually the reactance equations of such networks are expressed in terms of the electrical constants of the network elements. When the network consists of, say N elements, N electrical constants appear in the corresponding reactance equation. In equalizer and filter work it is convenient to eliminate the electrical constants from the reactance equations and use, in their place, the important critical frequencies of the networks; that is, the frequencies of resonance and anti-resonance. For any reactive network, however, the sum of the number of resonant and anti-resonant points is one less than the number of network elements. For instance, for a four-element reactive network, the resonant and anti-resonant points will number three. Hence, if these critical frequencies are used in the formulae to replace the electrical constants we are short one item for substitution purposes. To obviate this difficulty another critical frequency designated as f_S is used. The definition of f_S will be made clear as we progress. Proper selection of this frequency in connection with the equalizer work of the next chapter has made possible the simple resulting equations for the insertion losses of the networks. Following is shown the method for expressing the reactances of some simple two-terminal reactive networks in terms of these new parameters.

(a) Reactance of an Inductance Coil

The vector reactance of an inductance coil of L henries is usually expressed as follows:

$$jX = j\,2\,\pi f L. \tag{46}$$

This simple reactance network has, of course, no finite resonant or anti-resonant frequencies and hence these parameters cannot be used for substitution purposes. As a means of eliminating L from this equation, define f_S as the frequency where the reactance is, say X_S. Then we have

$$X_S = 2\,\pi f_S L$$

Solving for L gives:

$$L = \frac{X_S}{2\,\pi f_S} = \frac{X_S}{\omega_S} \text{ henries} \tag{47}$$

Equation (46) can now be written

$$jX = jX_S \frac{f}{f_S} \tag{48}$$

We have thus expressed the reactance of the inductance coil in terms of a frequency ratio f/f_S and the reactance X_S. To illustrate, suppose we wanted an inductance coil having a reactance of 500 ohms at 3000 c.p.s. From equation (47)

$$L = \frac{X_S}{\omega_S} = \frac{500}{2\pi \times 3000} = 0.0265 \text{ henries}$$

And the reactance of the coil at any frequency, f, is

$$jX = j\,500\,\frac{f}{3000}$$

(b) Reactance of a Condenser

In a similar manner the reactance of a condenser can be expressed in terms of a reactance X_S and a frequency ratio. The usual equation for the reactance of a condenser is

$$-jX = -j\,\frac{1}{2\pi f C} \tag{49}$$

At a frequency f_S let the numerical value of the reactance be X_S. Then

$$X_S = \frac{1}{2\pi f_S C}$$

Solving for C gives

$$C = \frac{1}{2\pi f_S X_S} = \frac{1}{\omega_S X_S} \text{ farads} \tag{50}$$

Using this value of C with equation (49) gives

$$-jX = -jX_S \frac{f_S}{f} \qquad (51)$$

As an example, suppose we wanted a condenser having a reactance of 500 ohms at 3000 c.p.s. Then from (50)

$$C = \frac{1,000,000}{2\pi \times 3,000 \times 500} = 0.106 \text{ microfarads}$$

and from equation (51)

$$-jX = -j\,500\,\frac{3,000}{f}$$

(c) Reactance of a Coil and Condenser in Series

In this case let Figure 157-A represent a two-terminal network consisting of an inductance L in series with a capacitance C. Figure 157-B shows symbolically the reactance characteristic of such a circuit

Figure 157-A — Inductance coil and condenser in series.

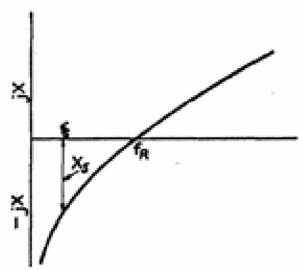


Figure 157-B — Reactance characteristic.

plotted against frequency. The symbol f_R is used to denote the resonant frequency of the coil and condenser; that is

$$f_R = \frac{1}{2\pi\sqrt{LC}} \qquad (52)$$

Again, as in the previous cases, f_S is defined as the frequency where the reactance of the circuit is numerically equal to X_S. Since this happens at two points in the frequency range, that is, once above the resonant frequency f_R and again below f_R, f_S is taken as the lower point as shown in Figure 157-B. The symbol "s" is used to denote the ratio of f_R to f_S, that is

$$s = \frac{f_R}{f_S} \text{ (greater than unity)}$$

The reactance equation for the circuit is

$$jX = j\left(\omega L - \frac{1}{\omega C}\right)$$

$$jX = j\omega_R L\left[\frac{\omega}{\omega_R} - \frac{1}{\omega\omega_R LC}\right]$$

But from equation (52) $(\omega^2_R) = \dfrac{1}{LC}$

and we have

$$jX = j\omega_R L\left[\frac{\omega}{\omega_R} - \frac{\omega_R}{\omega}\right]$$

$$jX = j\omega_R L\left[\frac{f}{f_R} - \frac{f_R}{f}\right] \tag{53}$$

by definition when $f = f_S$ then $X = -X_S$ (see Figure 157-B) which gives

$$\omega_R L = \frac{-X_S}{\dfrac{f_S}{f_R} - \dfrac{f_R}{f_S}} = \frac{X_S}{s - \dfrac{1}{s}} \tag{54}$$

Use of this value of $\omega_R L$ in connection with equation (53) gives us the form which we desire as a means of expressing the above reactance; that is,

$$jX = jX_S\frac{\dfrac{f}{f_R} - \dfrac{f_R}{f}}{s - \dfrac{1}{s}} \tag{55}$$

The values of L and C can also be expressed in terms of the new parameters in the following manner, by means of equation (54) and the relation

$$\omega_R = s\ \omega_s$$

We have

$$L = \frac{X_S}{\omega_R}\ \frac{s}{s^2 - 1} = \frac{X_S}{\omega_S}\ \frac{1}{s^2 - 1} \tag{56}$$

And when this value of L is used with equation (52) we have for C

$$C = \frac{1}{\omega_R{}^2 L} = \frac{1}{\omega_R X_S}\frac{s^2 - 1}{s} = \frac{1}{\omega_S X_S}\ \frac{s^2 - 1}{s^2} \tag{57}$$

As an illustration of the above, consider a series coil and condenser circuit having its resonant frequency at 5,000 c.p.s. and a reactive impedance of 500 ohms at 4,000 c.p.s. That is,

$f_R = 5{,}000$ c.p.s. $\quad s = 1.25$
$f_S = 4{,}000$ c.p.s. $\quad X_S = 500$ ohms.

This gives

$$L = \frac{500}{2\pi\,4{,}000}\ \frac{1}{(1.25)^2 - 1} = 0.0353 \text{ henries}$$

$$C = \frac{1{,}000{,}000}{2\pi \times 4{,}000 \times 500}\ \frac{(1.25)^2 - 1}{(1.25)^2} = 0.0287\ \mu f$$

$$jX = j\,\frac{500 \times 1.25}{(1.25)^2 - 1}\left[\frac{f}{f_R} - \frac{f_R}{f}\right] = j\,1{,}111\left[\frac{f}{f_R} - \frac{f_R}{f}\right]$$

(d) Reactance of a Coil and Condenser in Parallel

Figure 158-A represents a two-terminal network consisting of a coil and condenser in parallel and Figure 158-B shows the form of the reactance characteristic. The symbol f_R designates the frequency of

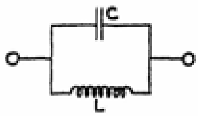


Figure 158-A — Inductance coil and condenser in parallel.

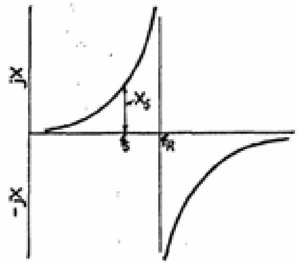


Figure 158-B — Reactance characteristic.

anti-resonance and the symbol f_S is used to indicate the point where a reactance of numerical value X_S is obtained. The reactance of this network is usually written in the form

$$jX = \frac{\dfrac{\omega L}{\omega C}}{j\left(\omega L - \dfrac{1}{\omega C}\right)}$$

which may be converted into the form

$$jX = \frac{\dfrac{\omega L}{\omega C}}{j\omega_R L\left[\dfrac{\omega}{\omega_R} - \dfrac{1}{\omega\omega_R}\,\dfrac{1}{LC}\right]}$$

At the resonant frequency, $\dfrac{1}{LC} = \omega^2{}_R$. Substituting this value of LC gives

$$jX = \frac{1}{j\omega_R C}\left[\frac{1}{\dfrac{\omega}{\omega_R} - \dfrac{\omega_R}{\omega}}\right] = \frac{1}{j\omega_R C}\left[\frac{1}{\dfrac{f}{f_R} - \dfrac{f_R}{f}}\right] \qquad (58)$$

Now by definition when $f = f_S$; $X = X_S$, which gives

$$\frac{1}{j\omega_R C} = jX_S\left[\frac{f_S}{f_R} - \frac{f_R}{f_S}\right] = jX_S\left[\frac{1}{s} - s\right]$$

$$\frac{1}{j\omega_R C} = -\,jX_S\left[s - \frac{1}{s}\right] \qquad (59)$$

Use of this value in connection with equation (58) gives the final form we desire for expressing the reactance of the parallel coil and condenser.

$$jX = -jX_S \frac{s - \frac{1}{s}}{\frac{f}{f_R} - \frac{f_R}{f}} \tag{60}$$

The values of L and C in terms of the new parameters may now be determined from (59)

$$C = \frac{1}{\omega_R X_S} \frac{1}{s - \frac{1}{s}} = \frac{1}{\omega_S X_S} \frac{1}{s^2 - 1} \tag{61}$$

And since at the resonant frequency, $\omega_R L = \dfrac{1}{\omega_R C}$ we have for L

$$L = \frac{1}{\omega^2_R C} = \frac{\omega_R X_S}{\omega^2_R}\left(s - \frac{1}{s}\right) = \frac{X_S}{\omega_R}\left(s - \frac{1}{s}\right)$$

$$L = \frac{X_S}{\omega_S} \frac{s^2 - 1}{s^2} \tag{62}$$

3. EQUIVALENT TWO-TERMINAL NETWORKS

The impedance of a two-terminal network is, of course, that obtained looking in at the terminals and takes into account the effects of all of the network elements. It is possible for certain pairs of such networks to have identical impedances at their terminals for any frequency and yet have different circuit arrangements and different element values. A pair of such networks, however, must each have the same number of elements with a definite relationship existing between corresponding elements. Two networks of these types are said to be equivalent. Not all two-terminal networks have equivalent networks. In general, however, any two-terminal network having two or more elements of the same kind has an equivalent network. By this is meant two or more resistance elements, two or more inductance elements, etc. Since equivalent networks display identical impedances at their terminals, they may be used interchangeably as parts of four-terminal networks. In this connection they are often of great commercial value for it frequently happens that one of the networks is easier and cheaper to construct than the other.

Single- and two-element networks have no equivalent circuits except those that are exactly identical, and these are not regarded as

being equivalent arrangements. For networks containing three electrical elements, important equivalent networks exist if the network has two elements of like kind. When the three elements of such a network are all different in kind, that is, a resistance, a capacitance and an inductance coil, no equivalent network exists. Two-terminal networks containing four or more elements also have equivalent networks but are not discussed in this book. However, the conditions for the equivalence of a pair of three-element networks are discussed in detail.

To illustrate the above, consider circuits A and B of Figure 159, where R_1 and R_2 are any two resistances and n is any number. The

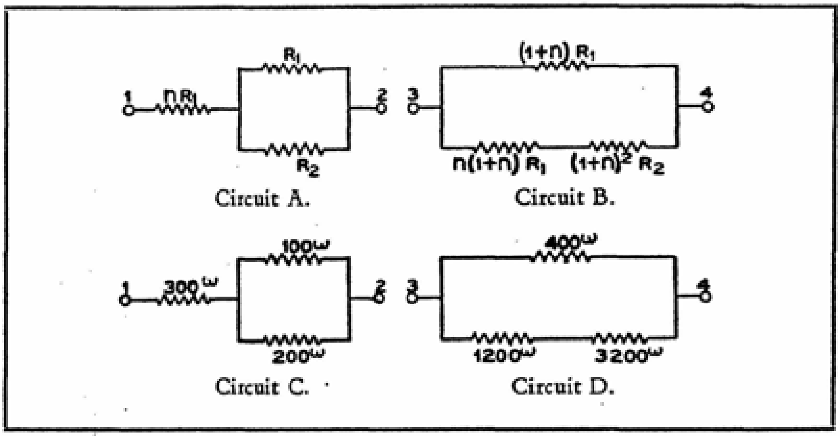


Figure 159 — Equivalent resistive networks.

resistance obtained at the terminals of circuit A of Figure 159 is

$$R_{12} = nR_1 + \frac{R_1 R_2}{R_1 + R_2} = \frac{nR^2_1 + R_1 R_2 (1 + n)}{R_1 + R_2} \tag{63}$$

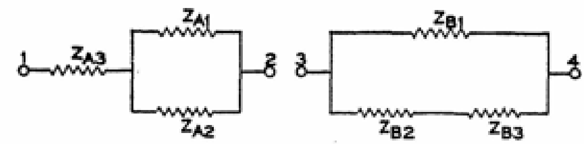
The resistance of circuit B of Figure 159 is

$$R_{34} = \frac{(1 + n)\ R_1\ [n\ (1 + n)\ R_1 + (1 + n)^2\ R_2]}{(1 + n)\ \ [R_1 + n\ R_1 + (1 + n)\ R_2]}$$

$$R_{34} = \frac{n\ R^2_1 + (1 + n)\ R_1\ R_2}{R_1 + R_2} \tag{64}$$

The resistances of the two networks are thus shown to be identical for any values of R_1, R_2, and n. Suppose, for instance, that $R_1 = 100$ ohms, $R_2 = 200$ ohms, and $n = 3$. This results in the two specific resistive circuits, C and D, of Figure 159, each of which may be shown to present a resistance of 367 ohms at its terminals.

The above illustration uses resistances for the elements of the two equivalent networks. The equivalence of a pair of similar networks



Circuit A Circuit B

IMPEDANCE RELATIONS

$$n = \frac{Z_{A3}}{Z_{A1}} = \frac{Z_{B3}}{Z_{B1}}$$

$$\frac{Z_{B1}}{Z_{A1}} = (1+n) \qquad \frac{Z_{B2}}{Z_{A2}} = (1+n)^2 \qquad \frac{Z_{B3}}{Z_{A3}} = (1+n)$$

Figure 160 — Equivalent three-element two-terminal networks.

can also be shown where impedances are used instead of resistances. Consider circuit A and circuit B of Figure 160. The impedances of these two networks are identical when their respective electrical elements are related in the manner shown by the equations given on the figure. This can be demonstrated as follows: Considering circuit A, its impedance is

$$Z_{12} = Z_{A3} + \frac{Z_{A1}\,Z_{A2}}{Z_{A1}+Z_{A2}} = \frac{Z_{A1}\,Z_{A2} + Z_{A1}\,Z_{A3} + Z_{A2}\,Z_{A3}}{Z_{A1}+Z_{A2}} \qquad (65)$$

And for circuit B we have

$$Z_{34} = \frac{Z_{B1}\,Z_{B2} + Z_{B1}\,Z_{B3}}{Z_{B1}+Z_{B2}+Z_{B3}}$$

Using the relation of elements shown in Figure 160 we may write

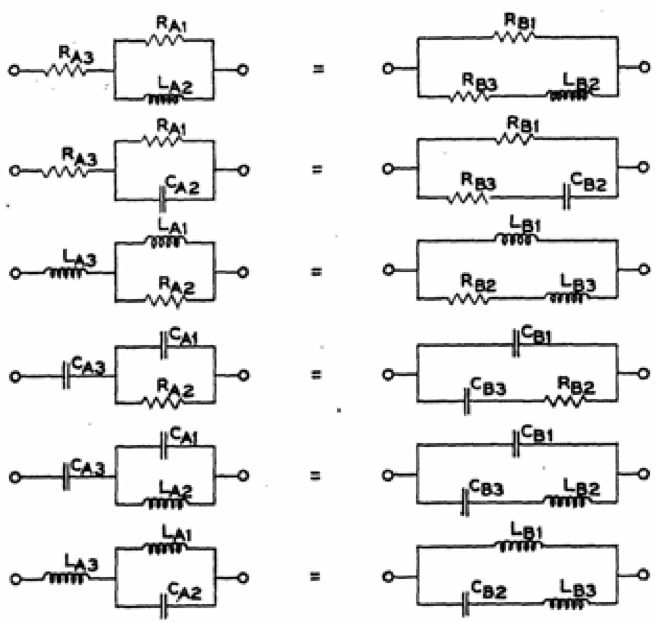
$$Z_{34} = \frac{Z_{A1}\,Z_{A2}\,(1+n)^3 + Z_{A1}\,Z_{A3}\,(1+n)^2}{Z_{A1}\,(1+n) + Z_{A3}\,(1+n) + Z_{A2}\,(1+n)^2}$$

$$Z_{34} = \frac{(1+n)\ [Z_{A1}\,Z_{A2} + nZ_{A1}\,Z_{A2} + Z_{A1}\,Z_{A3}]}{Z_{A1} + nZ_{A1} + Z_{A2}(1+n)}$$

$$Z_{34} = \frac{Z_{A1}\,Z_{A2} + Z_{A1}Z_{A3} + Z_{A2}\,Z_{A3}}{Z_{A1}+Z_{A2}}$$

$$Z_{34} = Z_{12} \qquad (66)$$

Hence, for the impedance relations shown, the two circuits A and B are equivalent. It is noted that for the practical use of these equivalent circuits the ratio n must be a real number. This is only realized when Z_{A1} and Z_{A3}, or Z_{B3} and Z_{B1} are the same kind of circuit elements; that is, both resistances, or both inductances, etc. The impedance equations of Figure 160 have been set up in the manner shown so that we can readily obtain circuit B when circuit A is known, or conversely. Figure 160 shows a number of pairs of three-element equivalent networks obtained by assigning specific element types to the general impedances of Figure 160.



$$n = \frac{R_{A3}}{R_{A1}} = \frac{R_{B3}}{R_{B1}} = \frac{L_{A3}}{L_{A1}} = \frac{L_{B3}}{L_{B1}} = \frac{C_{A1}}{C_{A3}} = \frac{C_{B1}}{C_{B3}}$$

$$(1+n) = \frac{R_{B3}}{R_{A3}} = \frac{L_{B3}}{L_{A3}} = \frac{C_{A3}}{C_{B3}} = \frac{R_{B1}}{R_{A1}} = \frac{L_{B1}}{L_{A1}} = \frac{C_{A1}}{C_{B1}}$$

$$(1+n)^2 = \frac{R_{B2}}{R_{A2}} = \frac{L_{B2}}{L_{A2}} = \frac{C_{A2}}{C_{B2}}$$

Figure 161 — Pairs of three-element equivalent networks.

4. INVERSE TWO-TERMINAL NETWORKS

The theory of inverse networks relates to an important mutual relationship which may be made to exist between the impedances of a pair of two-terminal networks when the electrical constants and configuration of one network are properly selected with respect to the other network. Consider a pair of two-terminal networks one of which presents an impedance Z_1 at its terminals, and the other having an impedance Z_2. When Z_1 and Z_2 are so related that

$$Z_1 Z_2 = R_0^2 \tag{67}$$

the networks are defined as being the reciprocal or inverse of each other with respect to R_0. In general R_0 may be a vector expression but in equalizer and filter work it is convenient to let R_0 be the resistive impedance between which the networks operate: In other words, R_0 is a real number and a known quantity in most network problems. It is noted that reciprocal or inverse networks always involve a pair of networks mutually related to each other. Complete information must be available concerning the electrical elements and circuit connections of one network before its inverse network can be found. Also, the inverse network has not been completely determined until its electrical constants and the manner in which they are connected together are known.

The inverse network corresponding to a given network whose constants and connections are known may be derived with the aid of the few pairs of simple inverse networks given in Figure 162.

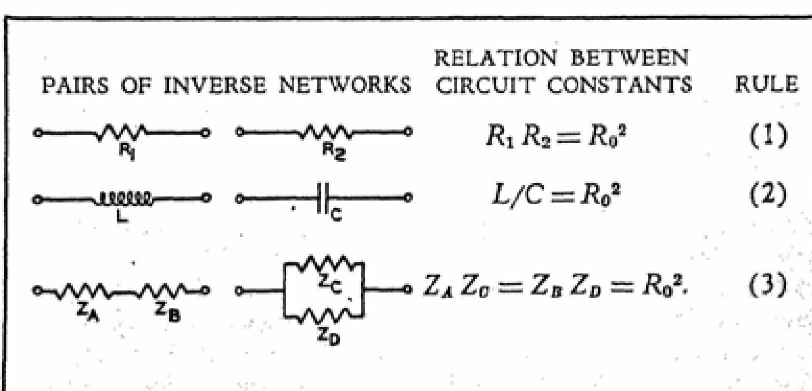


Figure 162 — Fundamental pairs of inverse networks.

The relationships shown between the electrical constants for the above pairs of inverse networks have been determined on the basis that their

corresponding impedances satisfy equation (67). It is left for the student to verify this to his own satisfaction. It is noted that:

(1) When one of a pair of inverse networks is a resistance the other network is another resistance.

(2) When one of a pair of inverse networks is an inductance the other is a capacitance, and conversely.

(3) When one of a pair of inverse networks consists of two impedances in series, the other consists of the inverse of these impedances in parallel, and conversely.

The following example illustrates the manner in which the above pairs of inverse networks may be used as fundamental rules for deriving the inverse circuits of more complex networks. Let it be required to find the inverse network corresponding to circuit A of Figure 163.

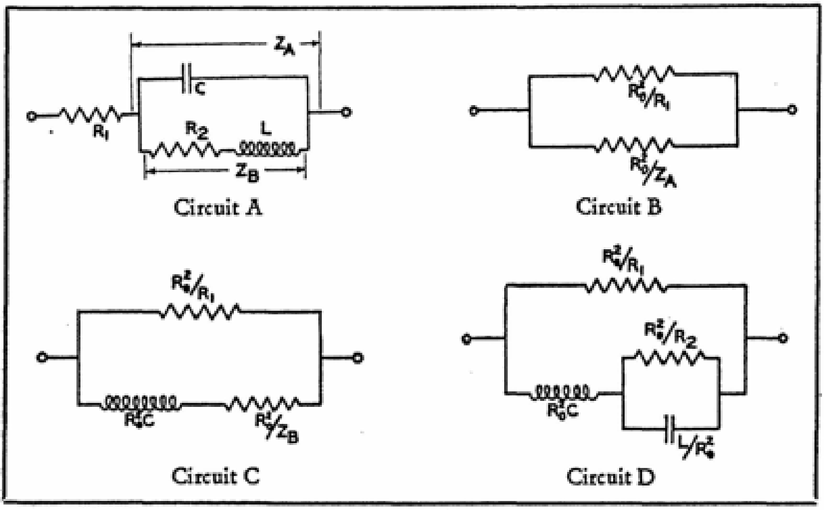
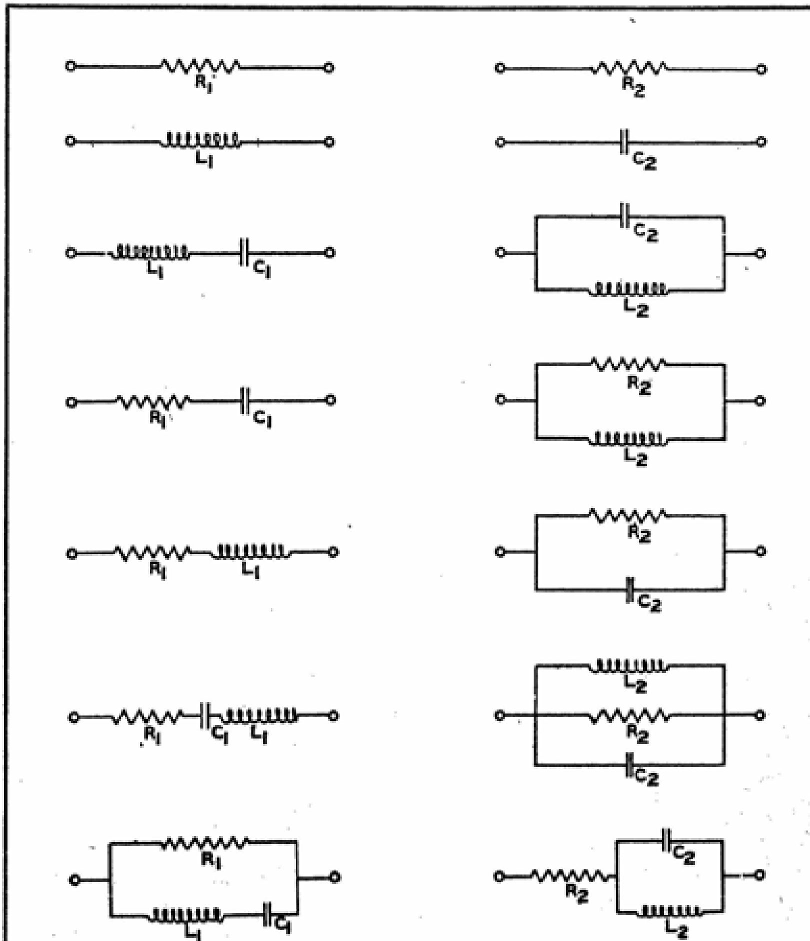


Figure 163.

Referring to Figure 163 it is seen that circuit A may be regarded as a resistance R_1 in series with an impedance Z_A. From Rule (3) the inverse of these two impedances in series is the parallel arrangement of B, where one of the impedances is the inverse of R_1, that is, R_0^2/R_1, and the other impedance is the inverse of Z_A which is R_0^2/Z_A. Now the inverse impedance R_0^2/Z_A may be simplified by regarding Z_A of circuit A as consisting of the condenser C in parallel with the impedance Z_B. Again making use of rule (3) the inverse of the impedance Z_A becomes an inductance of value $R^2{}_0C$ henries in series with the inverse of impedance Z_B which is R_0^2/Z_B. This condition is shown in circuit C. Once more a further simplification is obtained by using Rule (3) to find the inverse of Z_B which gives a resistance R_0^2/R_2 in parallel with a capacitance of L/R_0^2 farads. We are now able to write down circuit D

which is the final form for the inverse circuit corresponding to circuit A.

It is seen that the process of deriving an inverse network consists in replacing parallel circuits with series circuits and series circuits with parallel circuits and at the same time replacing all elements of the known network with their inverse elements in the inverse network. Experience in doing this makes it unnecessary to pass through the intermediate circuits as was done in the above. Figure 164 shows a number of pairs of inverse networks which will be useful for reference purposes in the equalizer work which follows.



$$R_1 R_2 = \frac{L_1}{C_2} = \frac{L_2}{C_1} = R_0^2.$$

Figure 164 — Inverse networks.

5. INSERTION LOSSES

The insertion loss of a network is a measure of its attenuation performance when operating under service conditions in a transmission facility. As already mentioned, networks usually operate between equal resistive impedances. For this condition the insertion loss is determined by the ratio of power delivered to the load before and after the network is inserted. Thus, if P_L is the power delivered to a load R_L before inserting the network and P'_L is the power after insertion of the network the insertion loss is defined as,

$$\text{I.L.} = 10 \log \frac{P_L}{P'_L} \tag{68}$$

In cases where the load resistance R_L is not equal to the sending-end impedance R_G, the power P_L is usually defined as that delivered to the load through an ideal transformer which matches the two system impedances R_G and R_L. In other words, P_L is the maximum power which the source of supply can deliver to the load. Where R_G and R_L are equal the ideal matching transformer is not required. Insertion losses are defined in this manner in order that they will be losses for the insertion

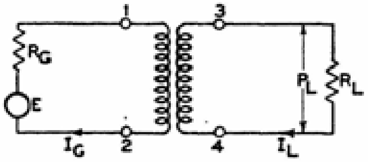


Figure 165-A — Ideal transformer matching R_G and R_L

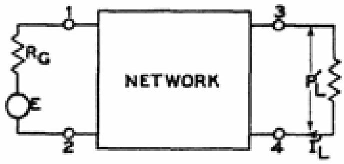


Figure 165-B.

of any network. That is, if P_L were not defined as the maximum power which could be delivered to the load, it would be possible to insert a network which under certain conditions would give an insertion gain. For some purposes this might not be objectionable but for the following wave filter and equalizer work it is desirable to exclude the possibility of an insertion loss becoming a gain.

Figure 165-A shows a cross-section of a transmission system having a sending impedance R_G and a receiving or load impedance R_L. The ideal transformer shown matches R_G with R_L so as to deliver the maximum power to R_L. For this reference circuit we have the following relations where P_G is the power sent into the transformer:

$$I_G = \frac{E}{2\,R_G}$$

$$P_G = \frac{E^2}{4\,R_G}$$

Now since the transformer is ideal it has zero dissipation loss and therefore all power sent into it is transferred on to the load. This gives

$$P_G = P_L = \frac{E^2}{4\,R_G} = I_L^2\,R_L$$

$$I_L = \sqrt{\frac{P_L}{R_L}} = \frac{E}{2\,\sqrt{R_L\,R_G}}$$

These are the reference power and reference current used in the computation of insertion losses. For the usual case where R_G and R_L are both equal resistances of some value, say R_0, the above equations reduce to the simple forms which would have been obtained had the ideal transformer not been used.

The power P'_L and the current I'_L refer to the condition where a network is inserted in the place of the ideal transformer, as illustrated in Figure 165-B. On a current basis the insertion loss formula corresponding to equation (68) becomes

$$\text{I.L.}^* = 20 \log \frac{I_L}{I'_L} \tag{69}$$

---

* NOTE: It will be noticed that there are two different notations used in this book to express insertion gains or insertion losses. Obviously, an insertion loss is a negative insertion gain, as an insertion gain is a negative insertion loss. For this reason, when using the decibel unit the loss or gain is expressed as a log function of either I'_L/I_L or I_L/I'_L. If the current I'_L (the load current after the change has been made in the circuit), is of a smaller numerical value than I_L, then an insertion loss has taken place. If the ratio I'_L/I_L had been used, a negative answer would have resulted, while if the ratio I_L/I'_L had been used, a positive answer would have resulted. These two answers would have the same numerical value but with opposite sign. In other words, either method is applicable as long as it is known whether the result is a gain or a loss.

In equalizer work where an insertion loss is always present, the ratio I_L/I'_L is used, as a positive answer results.

Chapter XVI

ATTENUATION EQUALIZERS

By HARRY KIMBALL

1. NATURE OF EQUALIZERS

As already mentioned, an attenuation equalizer is a four-terminal network whose attenuation loss, over a given frequency range, varies with frequency in some desirable manner. This means that if a number of frequencies of given amplitudes are simultaneously impressed upon the input terminals of an equalizer, the relative amplitudes of these frequencies will be changed when delivered to a load connected to the output terminals. The manner in which this change takes place is determined and can be controlled by the design of the equalizer. In sound picture work, the frequency range required extends from about 40 or 50 up to 7,500 or 8,000 cycles per second. This is called the transmission band. Any change in the relative amplitudes of the important frequency components of a signal changes the character of the reproduced sound as heard by the ear. For instance, where the upper frequencies of the transmission band are discriminated against, the signal is said to be "dull," and where they are accentuated, the signal is said to be "bright." In practice, the configurations required for the insertion loss curves of equalizers appear to vary over a wide range. Actually, many equalizer problems are but duplications of others with different values assigned to the network constants.

Equalizers may be provided with controls for varying at will the form of their insertion loss characteristics, or may be of fixed non-variable construction. Variable equalizers are used largely in re-recording work, whereas fixed equalizers are used as permanent parts of recording and reproducing equipments. The need for variable equalizers arises from the fact that in recording sound from artificial surroundings, it is, in some cases, not practicable to provide sets and pick-up conditions having the acoustical characteristics one would expect from the real scenes they represent when displayed as part of a continuous picture. The use of variable equalizers in the re-recording of a picture makes it possible to alter the amplitude relations of the frequency components of signals in a manner to create the illusion desired for the

picture, where, of course, this can be accomplished by such amplitude changes.

Except for the equalization employed in re-recording, sound systems are usually arranged to preserve, insofar as possible, the amplitude relations of the frequency components of signals. When unavoidable amplitude distortion occurs in one part of a system, fixed attenuation equalizers provide the means for making a permanent correction. In some cases attenuation distortion is deliberately inserted in one part of a system and compensated for in another part, for benefits obtained by removing the load from certain of the recording equipments as, for instance, for the the complementary method of recording discussed in Chapter III.

2. GENERAL EQUALIZER TYPES

From the great amount of work which has been done on the design theory of attenuation equalizers a number of general circuit arrangements have emerged which have proven to be the most satisfactory for general use. The network engineer does not necessarily restrict himself to the use of these few types but they do represent a large part of his kit of tools. These equalizer circuits are designated in the following manner:

(1) Series Impedance Type.
(2) Shunt Impedance Type.
(3) Full Series Type.
(4) Full Shunt Type.
(5) T Type.
(6) Bridged T Type.
(7) Lattice Type.

In general, the transmission characteristics of the circuit arrangements designated in the above manner are made to depend upon two impedances, denoted as Z_1 and Z_2, where Z_1 is usually a series impedance and Z_2 is a shunt impedance. Z_1 and Z_2 are defined as being the impedances of any pair of two-terminal networks which are inverse to each other with respect to the line impedance R_0; that is,

$$Z_1 Z_2 = R_0^2$$

In any given equalizer problem, Z_1 and Z_2 must necessarily be assigned specific circuit arrangements. The general properties of the above typical equalizer sections, however, can be determined before considering specific arrangements for Z_1 and Z_2. This is done in the work immediately following.

For each equalizer type we are principally interested in two items:

First, its insertion loss, and second, the impedance match conditions at the input and output terminals. These are determined for the above networks operating between circuit impedances of R_0 ohms at both the input and output terminals.

(a) Series Impedance Type

One of the simplest methods available for controlling, with respect to frequency, the power delivered to a load resistance from a resistive power source is to insert an impedance in series with one side of the connecting circuit in the manner shown in Figure 166-A. For this circuit the symbol, $2Z_1$, is used to designate the inserted impedance; that is, the equalizer consists of an impedance, $2 Z_1$, connected between the 1-3 terminals and with the 2-4 terminals shorted together. Without the use of the equalizer the current which the power source can deliver to the load is

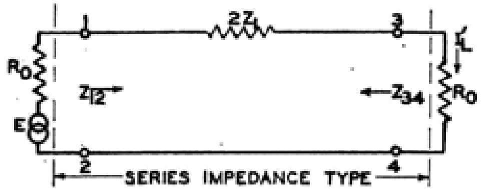


Figure 166-A — Series impedance type equalizer.

$$I_L = \frac{E}{2\,R_0}$$

Insertion of the equalizer changes this current to a value I'_L which is

$$I'_L = \frac{E}{2\,R_0 + 2\,Z_1}$$

The insertion loss by means of equation (69) then becomes

$$\text{I.L.} = 20 \log \frac{I_L}{I'_L} = 20 \log \frac{R_0 + Z_1}{R_0} \qquad (70)$$

Since the purpose of an equalizer is to secure an insertion loss which varies with frequency in some desirable manner, the impedance Z_1 must be a function of frequency to secure this effect. In other words, Z_1 cannot be entirely resistive because, if it were, the insertion loss would be constant with frequency.

The impedances looking into the network at the input and the output terminals are given by the expression

$$Z_{12} = Z_{34} = R_0 + 2\,Z_1 \qquad (71)$$

Obviously, this network does not match impedances at either the input or output terminals, and the mis-match obtained varies with frequency since Z_1 is a frequency function.

(b) Shunt Impedance Type

Another four-terminal equalizer circuit quite similar to the previous case is obtained by the use of a shunt impedance in the manner shown in Figure 166-B. For this circuit the symbol $Z_2/2$ is used to designate the shunt impedance. The load current I_L received when the equalizer is not in the circuit is again equal to $(E/2\ R_0)$, and the load current I'_L obtained when the equalizer is inserted is

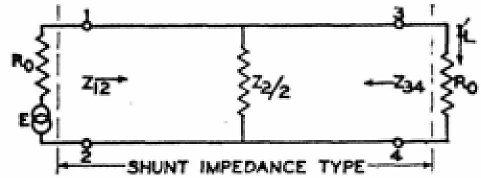


Figure 166-B — Shunt impedance type equalizer.

$$I'_L = \frac{E}{R_0 + \dfrac{R_0 Z_2}{2\,R_0 + Z_2}} \; \frac{Z_2}{2\,R_0 + Z_2} = \frac{E}{2\,R_0} \; \frac{Z_2}{(R_0 + Z_2)}$$

The insertion loss for the network then becomes

$$\text{I.L.} = 20 \log \frac{I_L}{I'_L} = 20 \log \frac{R_0 + Z_2}{Z_2} \tag{72}$$

Since the impedance Z_1 used in connection with the series impedance type circuit has been defined as being the inverse of Z_2, that is, $Z_1\,Z_2 = R_0^2$, we may also write for equation (72)

$$\text{I.L.} = 20 \log \frac{R_0 + Z_1}{R_0} \tag{73}$$

This means that the series impedance and shunt impedance type equalizers have the same insertion loss characteristics when designed with inverse impedances as their variable elements.

The impedance looking into the network either at the input terminals with the output terminals connected to R_0 ohms, or at the output terminals with the input terminals connected to R_0 ohms is,

$$Z_{12} = Z_{34} = \frac{\dfrac{R_0\,Z_2}{2}}{R_0 + \dfrac{Z_2}{2}} = \frac{R_0\,Z_2}{2\,R_0 + Z_2} \tag{74}$$

This network does not provide a match of impedances at the input and output terminals as can be seen, of course, from inspection.

(c) Full Series Constant Resistance Type

This equalizer circuit shown in Figure 166-C is quite commonly used in practice. As Figure 166-C shows, both the inverse impedances, Z_1 and Z_2, are used in this circuit. Referring to the figure, the impedance looking in at the 1-2 terminals with the 3-4 terminals connected to R_0 resistive ohms, is

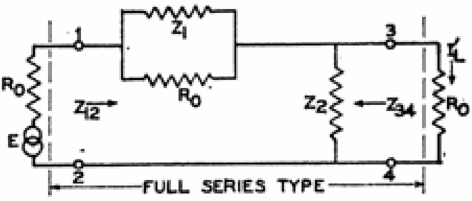


Figure 166-C — Full series constant resistance type equalizer.

$$Z_{12} = \frac{R_0 Z_1}{R_0 + Z_1} + \frac{R_0 Z_2}{R_0 + Z_2} = R_0 \frac{R_0 (Z_1 + Z_2) + 2 Z_1 Z_2}{R_0 (Z_1 + Z_2) + R_0^2 + Z_1 Z_2}$$

But by definition, $Z_1 Z_2 = R_0^2$ and we therefore have

$$Z_{12} = R_0 \tag{75}$$

The impedance Z_{34} looking into the network at the 3-4 terminals with the 1-2 terminals connected to a resistance of R_0 ohms is

$$Z_{34} = \frac{Z_2 \left[R_0 + \dfrac{R_0 Z_1}{R_0 + Z_1} \right]}{Z_2 + R_0 + \dfrac{R_0 Z_1}{R_0 + Z_1}}$$

$$Z_{34} = \frac{R_0 Z_2 (R_0 + 2 Z_1)}{(R_0 + Z_2)\ (R_0 + Z_1) + R_0 Z_1} \tag{76}$$

This impedance formula can be put in various forms by different transformations, but since it is of little importance these are omitted here. The important item in connection with the terminal impedances is that Z_{12} is a constant resistance regardless of frequency, whereas Z_{34} varies with frequency in some manner. This network then provides perfect impedance conditions at the 1-2 terminals, but not at the 3-4 terminals. The making of Z_1 and Z_2 inverse to each other makes it possible to secure this resistive impedance at the 1-2 terminals.

With regard to the insertion loss of the network the current delivered to the load resistance by the network is:

$$I'_L = \frac{E}{2 R_0} \frac{Z_2}{R_0 + Z_2}$$

Since the current without the use of the network is $I_L = \dfrac{E}{2\,R_0}$ the insertion loss becomes

$$\text{I.L.} = 20 \log \frac{R_0 + Z_2}{Z_2} = 20 \log \frac{R_0 + Z_1}{R_0} \qquad (77)$$

(d) Full Shunt Constant Resistance Type

This network shown in Figure 166-D is the inverse of the full series type network with respect to the 1-2 terminals. The terminal impedance and insertion loss formulae are derived in the same manner. It will be seen that in this case also a match of impedances is obtained at the 1-2 terminals but not at the 3-4 terminals.

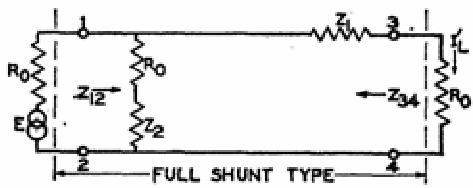


Figure 166-D — Full shunt constant resistance type equalizer.

$$Z_{12} = \frac{(R_0 + Z_2)\;(R_0 + Z_1)}{2\,R_0 + Z_1 + Z_2} = \frac{R_0^2 + Z_1 Z_2 + R_0\,(Z_1 + Z_2)}{2\,R_0 + Z_1 + Z_2}$$

But
$$Z_1\,Z_2 = R_0^2$$

$$Z_{12} = R_0 = \text{constant resistance} \qquad (78)$$

$$Z_{34} = \text{not constant with frequency}$$

$$I'_L = \frac{E}{2\,R_0}\,\frac{R_0 + Z_2}{2\,R_0 + Z_1 + Z_2} = \frac{E}{2\,R_0}\,\frac{R_0 + Z_2}{2\,R_0 + Z_2 + \dfrac{R_0^2}{Z_2}}$$

$$= \frac{E}{2\,R_0}\,\frac{Z_2\,(R_0 + Z_2)}{R_0^2 + 2\,R_0 Z_2 + Z_2^2} = \frac{E}{2\,R_0}\,\frac{Z_2\,(R_0 + Z_2)}{(R_0 + Z_2)^2}$$

$$I'_L = \frac{E}{2\,R_0}\,\frac{Z_2}{R_0 + Z_2}$$

$$I_L = \frac{E}{2\,R_0}$$

Insertion Loss =

$$20 \log \frac{I_L}{I'_L} = 20 \log \frac{R_0 + Z_1}{R_0} = 20 \log \frac{R_0 + Z_2}{Z_2} \qquad (79)$$

(e) T Type Constant Resistance Equalizer

This T type network of Figure 166-E may be analyzed by separating it into two networks in tandem, one a full series type and the other a

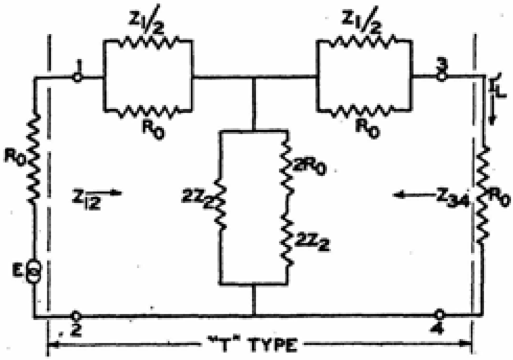


Figure 166-E — T type constant resistance type equalizer.

full shunt type as shown in Figure 166-F.

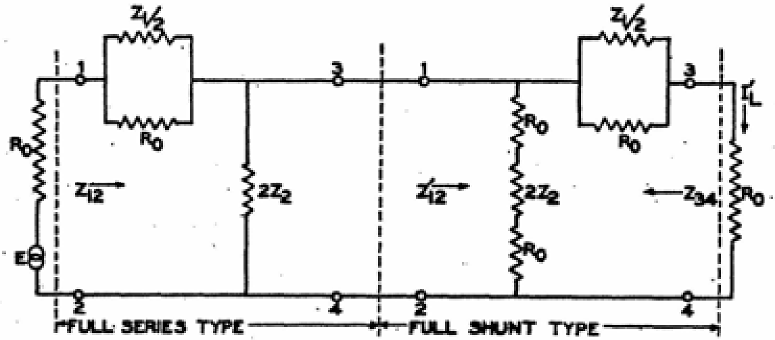


Figure 166-F — Equivalent circuit for T type equalizer of Figure 166-E.

For the full shunt type:

$$Z'_{12} = R_0 \qquad \text{[using formula (78)]}$$

$$\text{I.L.} = 20 \log \frac{2\,R_0 + 2\,Z_2}{R_0 + 2\,Z_2} \qquad \text{[using formula (79)]}$$

For the full series type:

$$Z_{12} = R_0 \qquad \text{[using formula (75)]}$$

$$\text{I.L.} = 20 \log \frac{R_0 + 2\,Z_2}{2\,Z_2} \qquad \text{[using formula (77)]}$$

From this data and because the T network is symmetrical, we have for the T network of Figure 166-E

$$Z_{12} = Z_{34} = R_0 = \text{constant resistance at both ends} \qquad (80)$$

And the insertion loss is the sum of the losses for the two component sections; that is,

$$\text{I.L.} = 20 \log \frac{2\,R_0 + 2\,Z_2}{R_0 + 2\,Z_2} + 20 \log \frac{R_0 + 2\,Z_2}{2\,Z_2}$$

$$\text{I.L.} = 20 \log \left(\frac{2\,R_0 + 2\,Z_2}{R_0 + 2\,Z_2}\right) \left(\frac{R_0 + 2\,Z_2}{2\,Z_2}\right)$$

$$\text{I.L.} = 20 \log \frac{R_0 + Z_2}{Z_2} = 20 \log \frac{R_0 + Z_1}{R_0} \qquad (81)$$

(f) Bridged T Constant Resistance Equalizer

As a means of deriving the impedance and loss characteristics of the bridged T equalizer of Figure 166-G, consider again the full shunt type equalizer arranged as in Figure 166-H.

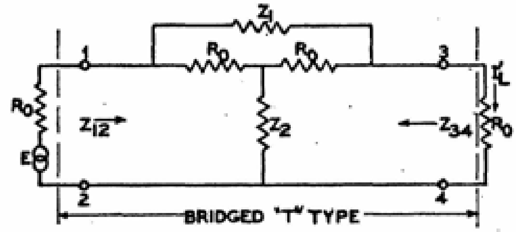


Figure 166-G — Bridged T constant resistance type equalizer.

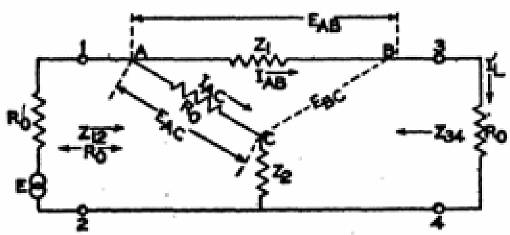


Figure 166-H — Equivalent circuit for bridged T type equalizer of Figure 166-G.

By Kirchoff's laws the voltage E_{BO} is:

$$E_{BO} = E_{AB} - E_{AO}$$

$$E_{BO} = Z_1 I_{AB} - R_0 I_{AO}$$

Since the impedance looking into the network at the 1-2 terminals is R_0, the voltage applied at these terminals is $E/2$. This gives for I_{AB} and I_{AO}:

$$I_{AB} = \frac{E}{2\,(R_0 + Z_1)}$$

$$I_{AO} = \frac{E}{2\,(R_0 + Z_2)}$$

Then the voltage E_{BC} becomes

$$E_{BC} = \frac{E}{2}\left[\frac{Z_1}{R_0 + Z_1} - \frac{R_0}{R_0 + Z_2}\right]$$

$$E_{BC} = \frac{E}{2}\left[\frac{Z_1 Z_2 - R_0^2}{(R_0 + Z_1)\ (R_0 + Z_2)}\right] \text{and as } Z_1 Z_2 = R_0^2$$

$$E_{BC} = \text{zero}$$

Since the voltage existing across the points B and C is zero, an impedance of any value may be connected between the points without disturbing the network. When a resistance of R_0 ohms is connected between the points B and C we arrive at the bridged T network of Figure 166-H. Hence, the insertion loss for the bridged T network is the same as for the full shunt circuit. Since the bridged T network is symmetrical at both ends, we know that $Z_{34} = Z_{12}$. Then we finally have for the bridged T network of Figure 166-G:

$$Z_{12} = Z_{34} = R_0 = \text{constant resistance} \tag{82}$$

and

$$\text{I.L.} = 20 \log \frac{R_0 + Z_1}{R_0} = 20 \log \frac{R_0 + Z_2}{Z_2} \tag{83}$$

(g) Bridge or Lattice Type Equalizer

Before taking up the lattice type equalizer, it will be useful to work out the current relations of the following bridge circuit when a voltage E is applied directly to one pair of the diagonal terminals and a resistance R_0 is connected across the other diagonal terminals. The impedances Z_A and Z_B shown in the figure are defined as being inverse to each other with respect to R_0; that is, $Z_A Z_B = R_0^2$. Since the configuration of the above circuit is symmetrical, it is evident that equal currents flow through each of the Z_A impedances, and similarly for the Z_B impedances. By Kirchoff's laws, the voltage drop around any closed loop is zero. Considering the closed loop 1-3-4-1,

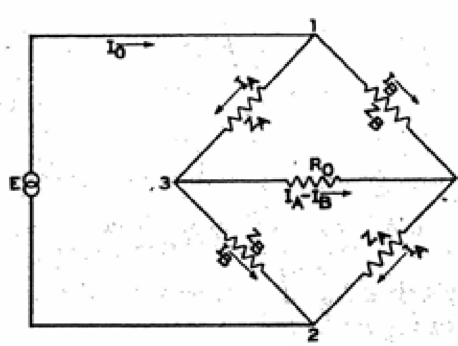


Figure 166-I — Bridge circuit.

we have,

$$I_A Z_A + (I_A - I_B) R_0 - I_B \frac{R_0^2}{Z_A} = 0$$

$$I_A (Z_A + R_0) - I_B R_0 \left(\frac{Z_A + R_0}{Z_A}\right) = 0$$

or we have

$$I_A Z_A = I_B R_0 \tag{84}$$

Now considering the loop 1-3-2-1, which includes the voltage E, we have

$$I_A Z_A + I_B \frac{R_0^2}{Z_A} = E \tag{85}$$

Substituting $I_B R_0$ for $I_A Z_A$ gives

$$I_B = \frac{E}{R_0} \frac{Z_A}{R_0 + Z_A}$$

Then from equation (84)

$$I_A = \frac{E}{R_0} \frac{R_0}{R_0 + Z_A}$$

$$I_0 = I_A + I_B = \frac{E}{R_0}\left[\frac{R_0}{R_0 + Z_A} + \frac{Z_A}{R_0 + Z_A}\right] = \frac{E}{R_0} \tag{86}$$

$$I_A - I_B = \frac{E}{R_0} \frac{R_0 - Z_A}{R_0 + Z_A} = I_0 \frac{R_0 - Z_A}{R_0 + Z_A} \tag{87}$$

It is noted from equation (86) that the impedance looking into the bridge at terminals 1-2 is a constant resistance of R_0 ohms. Also, from equation (87) the current delivered to the resistance R_0 is the current I_0 entering terminals 1-2 multiplied by the factor $\dfrac{R_0 - Z_A}{R_0 + Z_A}$

We are now in a position to consider the lattice type equalizer section of Figure 166-J. This network is the same as the above bridge circuit except that a resistance of R_0 ohms has been inserted in series with the voltage E, and the impedances Z_A and Z_B have been replaced by

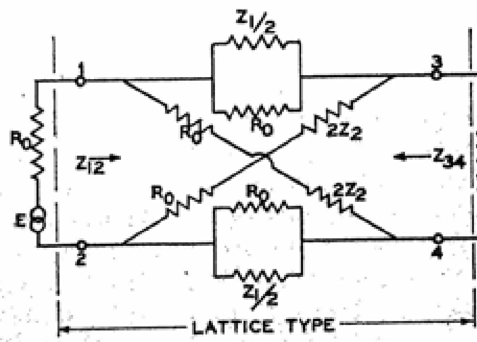
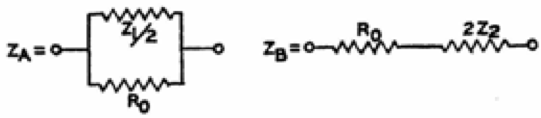


Figure 166-J — Lattice type equalizer.

the following circuits where Z_1 and Z_2 are inverse to each other:



Since the network is symmetrical about the 1-2 and 3-4 terminals we may say by means of equation (86):

$$Z_{12} = Z_{34} = R_0 \tag{88}$$

and also from equation (87)

$$I'_L = \frac{E}{2\,R_0}\,\frac{R_0 - Z_A}{R_0 + Z_A}$$

$$I'_L = \frac{E}{2\,R_0}\,\frac{R_0 - \dfrac{R_0\,Z_1}{2\,R_0 + Z_1}}{R_0 + \dfrac{R_0\,Z_1}{2\,R_0 + Z_1}} = \frac{E}{2\,(R_0 + Z_1)}$$

Since $I_L = \dfrac{E}{2\,R_0}$ we have for the insertion loss

$$\text{I.L.} = 20 \log \frac{R_0 + Z_1}{R_0} = 20 \log \frac{R_0 + Z_2}{Z_2} \tag{89}$$

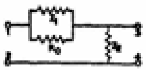
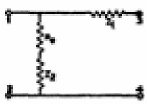
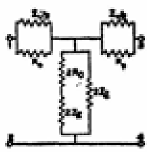
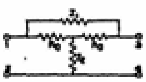
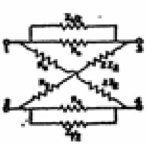
3. SUMMARY OF EQUALIZER TYPES

Figure 166 is a summary of the seven equalizer types discussed above. It is noted that the insertion loss formula, as expressed by the equation

$$\text{I.L.} = 20 \log \frac{R_0 + Z_1}{R_0} = 20 \log \frac{R_0 + Z_2}{Z_2} \tag{90}$$

is applicable to each of the equalizer types. This means that an insertion loss characteristic obtained with one of the equalizer types can be duplicated by any of the other types. The above formula also shows that the configurations of the insertion loss characteristics obtained for the equalizers of Figure 166 are determined solely by the inverse arms of the networks as represented by impedances Z_1 and Z_2. This feature makes it practicable, in a design problem, to determine the circuits of the inverse arms independent of the equalizer types with which they are to be used. In the work immediately following, the insertion loss

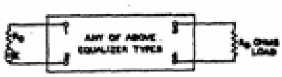
Figure 166 — Fundamental Equalizer Types

Figure	Network	Type	Z_{12}	Z_{34}	Insertion Loss
166-A		Series Imp.	Not Constant	Not Constant	$20 \log \frac{R_0 + Z_1}{R_0}$
166-B		Shunt Imp.	Not Constant	Not Constant	$20 \log \frac{R_0 + Z_2}{Z_2}$
166-C		Full Series	R_0	Not Constant	$20 \log \frac{R_0 + Z_1}{R_0}$
166-D		Full Shunt	R_0	Not Constant	$20 \log \frac{R_0 + Z_1}{R_0}$
166-E		T Type	R_0	R_0	$20 \log \frac{R_0 + Z_1}{R_0}$
166-G		Bridged T	R_0	R_0	$20 \log \frac{R_0 + Z_1}{R_0}$
166-J		Lattice	R_0	R_0	$20 \log \frac{R_0 + Z_1}{R_0}$

NOTES:

(1) $Z_1 Z_2 = R_0^2$ for all networks

(2) $20 \log \frac{R_0 + Z_1}{R_0} = 20 \log \frac{R_0 + Z_2}{Z_2}$

(3) Working Circuit = 

characteristics obtained with the use of certain common circuit arrangements for the inverse arms are discussed in detail. In doing this it is convenient to consider two types of circuits: first, those containing only reactive elements, and second, those containing both resistive and reactive elements. The purpose of this classification is made clear as we progress.

As already mentioned, where the inverse arms are entirely resistive, the seven equalizer types degenerate into attenuators. For instance, suppose we let

$$Z_1 = R_0\,(K-1) \text{ and } Z_2 = \frac{R_0}{K-1} \text{ (ohms)}$$

where K is a constant to be defined later. Use of these arms with the seven circuits of Figure 166 results in the attenuator circuits of Figure 167. The insertion loss of these attenuators is

$$\text{I.L.} = 20 \log \frac{R_0 + R_0\,(K-1)}{R_0} = 20 \log K$$

and K may now be seen to be the factor which determines the loss of the attenuators.

Another important feature in connection with the above equalizer circuits is the impedance match obtained at the 1-2 and 3-4 terminals. It is seen that two of the types shown do not give a match at either end; two produce a match of impedances at one end; and the remaining three

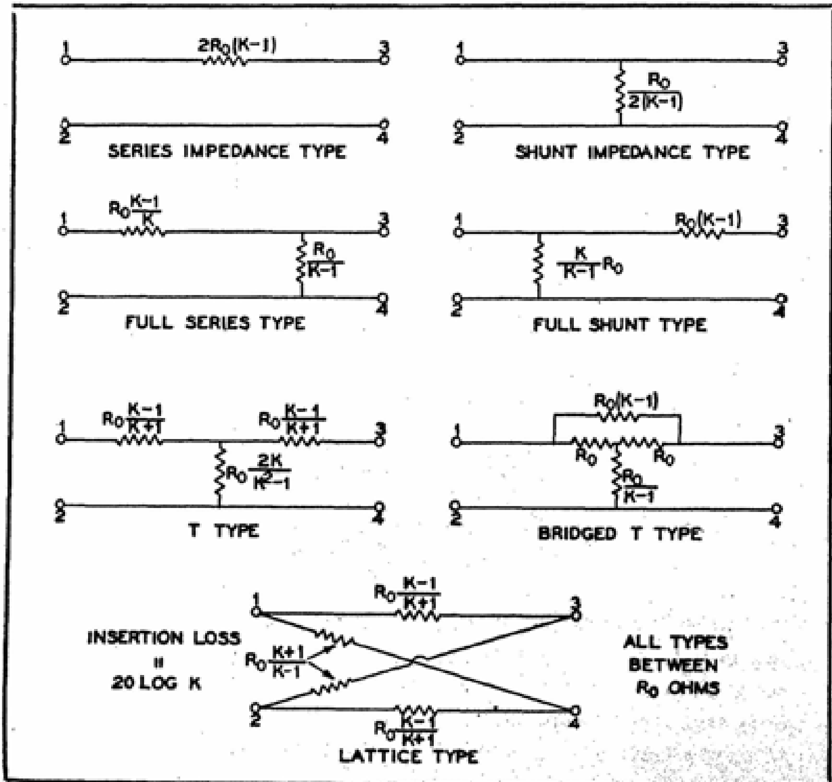


Figure 167 — Attenuator circuits.

provide a match of impedances at both ends. This dissimilarity between the networks makes one circuit more desirable than another under different conditions as discussed later.

4. INVERSE ARMS PURELY REACTIVE

When the inverse arms of equalizers consist only of reactive elements their impedances may be expressed as

$$Z_1 = jX_1 \text{ and } Z_2 = jX_2$$

where $(jX_1)\ (jX_2) = R_0^2$. Substituting these values into the general loss formula of equation (90) we have

$$\text{I.L.} = 20 \log \frac{R_0 + jX_1}{R_0}$$

The term $\left(\frac{R_0 + jX_1}{R_0}\right)$ represents the vector ratio of two currents, I_L and I'_L. A network insertion loss is proportional to the ratio of the absolute magnitudes of the two currents, which is expressed by the absolute magnitude of the vector $\left(\frac{R_0 + jX_1}{R_0}\right)$

This gives:

$$\text{I.L.} = 20 \log \sqrt{\frac{R_0^2 + X_1^2}{R_0^2}}$$

$$\text{I.L.} = 10 \log \frac{R_0^2 + X_1^2}{R_0^2} = 10 \log \frac{R_0^2 + X_2^2}{X_2^2} \tag{91}$$

Since X_1 and X_2 are pure reactances they vary with frequency between values of zero and infinity. This causes the insertion loss of equalizers employing such arms to also vary between zero and infinity over the complete frequency spectrum. It is useful to note at this time that when X_1 has a numerical value of R_0 ohms, and X_2 is also equal to R_0 ohms, the insertion loss becomes

$$\text{I.L.} = 10 \log \frac{R_0^2 + R_0^2}{R_0^2} = 10 \log 2 = 3 \text{ db} \tag{92}$$

In other words, the insertion loss passes through the three db loss point whenever the reactances X_1 and X_2 are numerically equal to R_0 ohms. This provides a useful design relation which will be used later.

Formula (91) is, of course, valid for inverse arms of any complexity. In practice, a few relatively simple pairs are sufficient for

handling a large portion of the equalizer problems encountered. In the following work, insertion loss data are given for the pairs of reactive

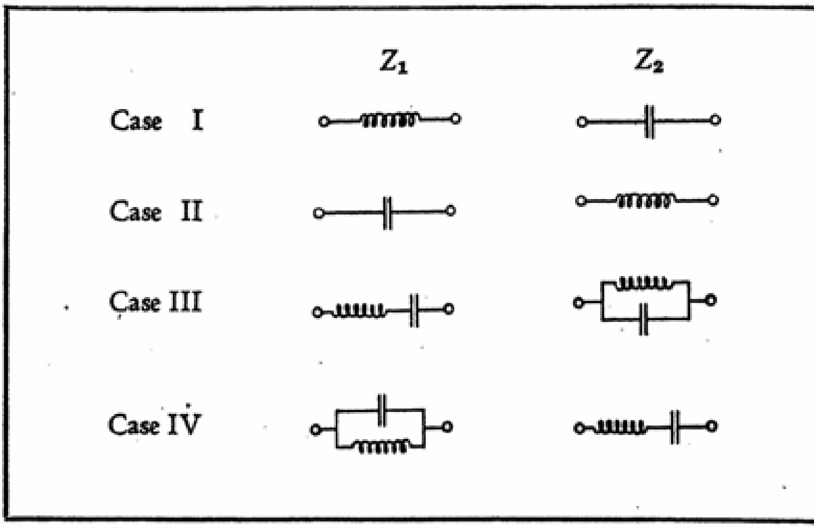


Figure 168 — Pairs of reactive inverse networks.

inverse networks of Figure 168, used in connection with any of the equalizer types of Figure 166.

(a) Case I

When $Z_1 =$ 

and $Z_2 =$ 

From inspection it can be seen that the insertion loss of equalizers employing inverse arms of this type increases with increasing frequency. For example, Figure 169 shows the full series type equalizer of Figure 166-C employing these inverse arms. Since the impedance of the shunt

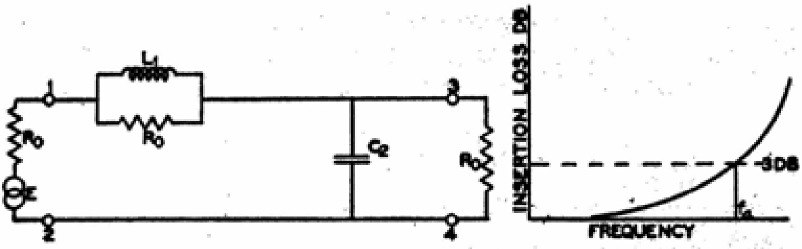


Figure 169 — A full series, Case I, equalizer with insertion loss characteristic.

condenser C_2 decreases with frequency, and that of the series inductance L_1 increases, the insertion loss becomes greater as the frequency increases.

This effect can be shown analytically in a more precise manner as follows:

From equation (48) the reactance of the inductance L_1 can be written

$$jX_1 = jX_S \frac{f}{f_S} \tag{93}$$

where X_S is the reactance at the frequency f_S. It was shown in equation (92) that where equalizers employ purely reactive inverse arms their insertion loss is three db at frequencies which makes the reactances of the inverse arms each numerically equal to R_0 ohms. Let "f_a" be the frequency where the reactances X_1 and X_2 are each R_0 ohms, or, in other words, f_a is the frequency where equalizers employing these inverse arms have three db loss. We have

$$jR_0 = jX_S \frac{f_a}{f_S} \text{ or } \frac{X_S}{f_S} = \frac{R_0}{f_a}$$

Then equation (93) can be written

$$jX_1 = jR_0 \frac{f}{f_a}$$

Using this value of X_1 in connection with equation (91) gives for the final insertion loss formula

$$\text{I.L.} = 10 \log \left[1 + \left(\frac{f}{f_a} \right)^2 \right] \tag{94}$$

Chart II shows this insertion loss formula graphically where loss in db is plotted against the frequency ratio f/f_a. As would be expected when $f/f_a = 1$ or $f = f_a$, the insertion loss is three db and the amount of loss increases with increasing frequency. In using this chart in practical problems we need to specify only the frequency f_a where we wish three db loss to occur, and read the losses for other frequencies from the chart after determining the corresponding frequency ratios.

In terms of R_0 and the design constant f_a, the values of L_1 and C_1 can be obtained by the use of equations (47) and (50). We have

$$L_1 = \frac{X_S}{\omega_S} = \frac{R_0}{\omega_a} = \frac{R_0}{2\pi f_a}$$

$$C_2 = \frac{1}{X_S\, \omega_S} = \frac{1}{R_0\, \omega_a} = \frac{1}{2\pi f_a R_0} \tag{95}$$

Column I of Chart I shows the equalizer sections obtained when this pair of inverse arms is used with the seven fundamental equalizer

types. The formulae for computing the electrical constants for each of the networks are given at the bottom of the column.

(b) Case II

When $Z_1 =$ 

and $Z_2 =$ 

This pair of inverse arms is the same as the previous pair except that Z_1 and Z_2 are interchanged. Figure 170 shows a full series type equalizer

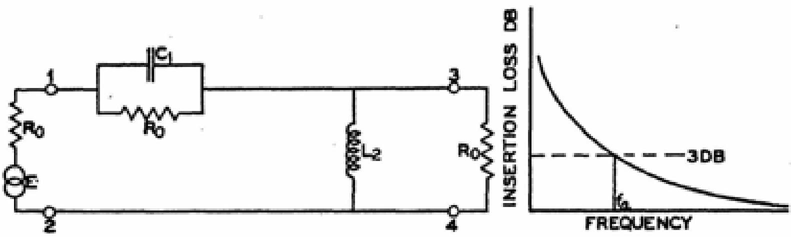


Figure 170 — A full series, Case II, equalizer with insertion loss characteristic.

employing these arms. From inspection it can be seen that the insertion loss of this equalizer should decrease with increasing frequency somewhat as indicated in the figure. The insertion loss formula expressing this characteristic is obtained in the following manner:

The reactance of the inductance L_2 is

$$jX_2 = jR_0 \frac{f}{f_a}$$

where f_a is the frequency for which the numerical value of X_2 is R_0 ohms. Substituting this value of X_2 into equation (91) we have for the insertion loss formula of equalizers employing these arms:

$$\text{I. L.} = 10 \log \left[\frac{R_0^2 + R_0^2 \left(\frac{f}{f_a}\right)^2}{R_0^2 \left(\frac{f}{f_a}\right)^2} \right]$$

$$\text{I. L.} = 10 \log \left[1 + \left(\frac{f_a}{f}\right)^2 \right] \qquad (96)$$

and the values of L_2 and C_1 in terms of R_0 and f_a as obtained from equations (48) and (50) are

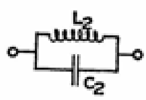
$$L_2 = \frac{X_g}{\omega_g} = \frac{R_0}{\omega_a} = \frac{R_0}{2\,\pi\, f_a}$$

$$C_1 = \frac{1}{\omega_g\, X_g} = \frac{1}{R_0\, \omega_a} = \frac{1}{2\,\pi\, f_a\, R_0} \qquad (97)$$

Chart III shows the insertion loss of equation (96) plotted against the ratio f/f_a. It is seen that in this case the insertion loss decreases with increasing frequency. Again, as in the previous case, the loss when $f = f_a$ is three db. Column II of Chart I shows the equalizer circuits obtained when this pair of inverse arms is used with the fundamental types of Figure 166, and the electrical constants for each of the circuits are given at the bottom of the column. In practical work, the frequency f_a is taken as the point where we wish three db loss to occur, and the losses at other frequencies are read from the curve after computing the corresponding frequency ratios.

(c) Case III

When $Z_1 =$ 

and $Z_2 =$ 

The general form of the insertion loss characteristic of equalizers using this pair of inverse arms is shown in Figure 171. In this figure, the circuit given is obtained by using the above inverse arms with the full

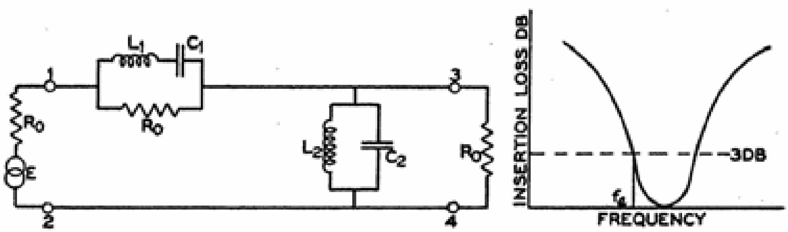


Figure 171 — A full series. Case III. equalizer with insertion loss characteristic.

series type equalizer of Figure 166-C. Since Z_1 and Z_2 are inverse to each other, L_1 and C_1 are resonant at the same frequency, f_R, that L_2 and C_2 are anti-resonant. When this occurs, the network insertion loss is zero as shown in the figure. For frequencies less or greater than f_R, the loss is finite and increases with frequency departure from f_R. At some frequency f_a, which is the frequency for which Z_1 and Z_2 are each numerically equal to R_0 ohms, the network loss is three db. The formula for precisely expressing this insertion loss characteristic is obtained as follows:

Letting $f_S = f_a$ and $X_S = R_0$ as was done in the foregoing cases, we have by means of equation (55)

$$jX_1 = jR_0 \frac{\dfrac{f}{f_R} - \dfrac{f_R}{f}}{a - \dfrac{1}{a}} \tag{98}$$

where $s = a$, and "a" is the frequency ratio f_R/f_a and by definition is greater than unity. Substitution of this value of X_1 into equation (91) gives for the insertion loss equation:

$$\text{I. L.} = 10 \log \left[1 + \frac{\left(\frac{f}{f_R} - \frac{f_R}{f} \right)^2}{\left(a - \frac{1}{a} \right)^2} \right] \tag{99}$$

This equation expresses the insertion loss of equalizers using these arms in terms of the resonant frequency f_R and the frequency f_a, where three db loss is obtained. The electrical elements, L_1, C_1, L_2 and C_2, do not appear in this insertion loss formula, but may also be expressed in terms of these design parameters. Remembering that $X_S = R_0$, $\omega_S = \omega_a$ and

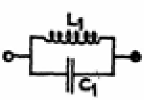
$\frac{L_1}{C_2} = \frac{L_2}{C_1} = R_0^2$ we obtain from equation (56) and (57)

$$L_1 = \frac{R_0}{\omega_a} \frac{1}{a^2 - 1} \qquad L_2 = \frac{R_0}{\omega_a} \frac{a^2 - 1}{a^2}$$

$$C_1 = \frac{1}{\omega_a R_0} \frac{a^2 - 1}{a^2} \qquad C_2 = \frac{1}{\omega_a R_0} \frac{1}{a^2 - 1} \tag{100}$$

Chart IV shows graphs of the insertion loss formula of equation (99) for a number of arbitrarily chosen values of "a". Assigning a specific value to "a" is equivalent to specifying the relation of the resonant frequency f_R to the frequency f_a where three db loss is obtained. For instance, for the curve where $a = 2$, f_a is one-half the frequency of resonance. Column III of Chart I shows the equalizers obtained when this pair of inverse arms is used with each of the network types of Figure 166. The formulae for the electrical constants of each of the circuits appear at the bottom of the column.

(d) Case IV

When $Z_1 =$ 

and $Z_2 =$ 

This pair of inverse arms is the same as the previous pair except that Z_1 and Z_2 are interchanged. Figure 172 shows a full series type equalizer employing these arms, and the general form of the associated insertion loss characteristic. It is seen that the loss is infinite at the resonant frequency f_R, and is zero at zero frequency and at an infinite frequency. At some frequency f_a, where both Z_1 and Z_2 are equal to

R_0 ohms, the loss curve passes through the three db point. As in the previous cases, f_a can be placed at any point in the frequency range be-

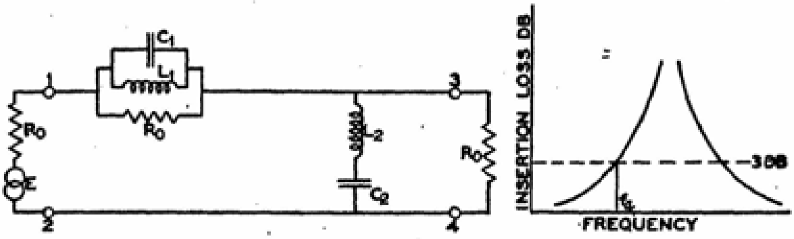


Figure 172 — A full series, Case IV, equalizer with insertion loss characteristic.

low f_R, simply by arranging Z_1 and Z_2 to be numerically equal to R_0 ohms at this frequency.

The insertion loss formula is obtained as follows: Letting $f_S = f_a$, $X_S = R_0$, and $s = a$, we have by means of equation (55)

$$jX_2 = jR_0 \frac{\frac{f}{f_R} - \frac{f_R}{f}}{a - \frac{1}{a}}$$

Substituting this value of X_2 into equation (91) gives

$$\text{I.L.} = 10 \log \left[1 + \frac{\left(a - \frac{1}{a}\right)^2}{\left(\frac{f}{f_R} - \frac{f_R}{f}\right)^2} \right] \tag{101}$$

By means of equations (56) and (57), and remembering that X_S and ω_S have been assigned the values of $X_S = R_0$ and $\omega_S = \omega_a$, we have for L_2 and C_2

$$L_2 = \frac{R_0}{\omega_a} \frac{1}{a^2 - 1} \qquad C_2 = \frac{1}{R_0 \, \omega_a} \frac{a^2 - 1}{a^2} \tag{102}$$

And by the inverse relations of $\frac{L_1}{C_2} = \frac{L_2}{C_1} = R_0^2$ we have

$$L_1 = \frac{R_0}{\omega_a} \frac{a^2 - 1}{a^2}$$

$$C_1 = \frac{1}{R_0 \, \omega_a} \frac{1}{a^2 - 1} \tag{103}$$

Again, as in the previous cases, equation (101) expresses the insertion loss of equalizers using this pair of inverse arms in terms of f_R and f_a, that is, in terms of f/f_R and $f_R/f_a = a$. Chart V shows graphs of equation (101) for different values of a. Column IV of Chart I shows the equalizers obtained when this pair of inverse arms is used with each of

the equalizer types of Figure 166. Formulae for the electrical constants are given at the bottom of the column.

5. INVERSE ARMS CONTAINING BOTH RESISTIVE AND REACTIVE ELEMENTS

Where the inverse arms of equalizers contain both resistive and reactive elements the circuit configurations considered must again be limited to those most commonly used to keep the data presented within reason. The general circuit configurations of Figure 173 will first be discussed, after which specific circuits will be taken up. The significance of the symbol "K," which determines the value of the resistance, $R_0\ (K-1)$, will be made clear later. The first pair of inverse arms of Figure 173 is included only for the purposes of generality and as shown below contributes nothing new to the design of equalizers.

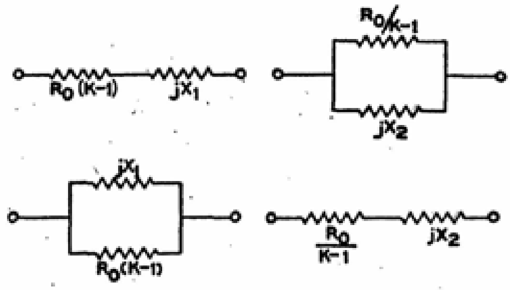


Figure 173—General pairs of inverse arms containing both resistive and reactive elements.

For this pair of arms we have

$$Z_1 = R_0\ (K-1) + jX_1$$

and when used in connection with equation (90) the insertion loss formula becomes

$$\text{I.L.} = 20 \log \frac{R_0 + R_0\ (K-1) + jX_1}{R_0}$$

$$= 20 \log \frac{R_0\, K + jX_1}{R_0}$$

$$= 20 \log K \frac{R_0 + j\dfrac{X_1}{K}}{R_0}$$

$$= 20 \log K + 20 \log \frac{R_0 + j\dfrac{X_1}{K}}{R_0} \tag{104}$$

This formula makes it clear that an equalizer using this pair of inverse arms is equivalent to a circuit consisting of an attenuator of loss equal to 20 log K, in tandem with an equalizer employing purely reactive inverse arms of values

$$Z_1 = j\frac{X_1}{K} \text{ and } Z_2 = jKX_2$$

This equivalent arrangement for a bridged T type equalizer is shown in Figure 174. Since the insertion loss performance of both of these

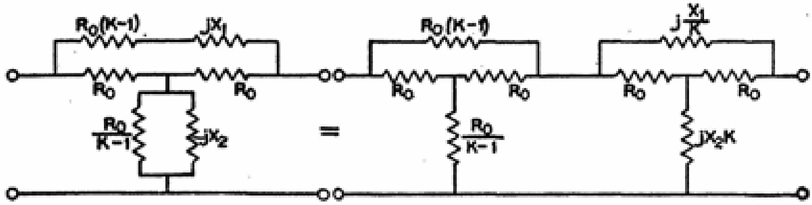


Figure 174 — Equivalent bridged T equalizer circuits.

component structures has already been discussed, equalizers employing the first pair of inverse arms of Figure 173 are not discussed further.

Referring to the second pair of inverse arms of Figure 173 we have

$$Z_1 = \frac{jR_0 X_1 (K-1)}{R_0 (K-1) + jX_1}$$

Used in connection with equation (90) this gives for the insertion loss

$$\text{I.L.} = 20 \log \frac{R_0 (K-1) + jKX_1}{R_0 (K-1) + jX_1}$$

Eliminating the vector symbol "j" by the method used in developing equation (91), this becomes

$$\text{I.L.} = 10 \log \frac{R_0^2 (K-1)^2 + K^2 X_1^2}{R_0^2 (K-1)^2 + X_1^2}$$

$$\text{I.L.} = 10 \log \left[1 + \frac{K^2 - 1}{1 + (K-1)^2 \left(\frac{R_0}{X_1}\right)^2} \right] \tag{105}$$

or expressed in terms of X_2 we have

$$\text{I.L.} = 10 \log \left[1 + \frac{K^2 - 1}{1 + (K-1)^2 \left(\frac{X_2}{R_0}\right)^2} \right] \tag{106}$$

The insertion losses obtained from equation (105) for certain specific values of X_1 are of interest at this time, namely, when $X_1 = 0$, $X_1 = \infty$ and $|X_1| = R_0 \dfrac{K-1}{\sqrt{K}}$. Successively inserting these values into equation (105) we have

$$\left[\text{I.L.}\right]_{X_1=0} = 10 \log 1 = \text{zero} \tag{107}$$

$$\left[\text{I.L.}\right]_{X_1=\infty} = 10 \log K^2 = 20 \log K \tag{108}$$

$$\left[\text{I.L.}\right]_{X_1=R_0\frac{K-1}{\sqrt{K}}} = 10 \log K \tag{109}$$

In other words, the insertion loss of equalizers employing these inverse arms varies over the range from zero to a maximum value equal to 20 log K. When the loss (20 log K) is obtained, the equalizer reactive elements are ineffective and the equalizer circuits reduce to the attenuator circuits shown in Figure 167. When $|X_1|$ takes on the special value of $R_0 \frac{K-1}{\sqrt{K}}$ the insertion loss, as shown by equation (109), is one-half its maximum value. In the following work the symbol f_b is used to denote the frequency for which $|X_1| = R_0 \frac{K-1}{\sqrt{K}}$, or what is the same thing, $|X_2| = R_0 \frac{\sqrt{K}}{K-1}$. This frequency f_b is employed as a design parameter for the following networks. It is referred to as the half-loss frequency. The frequency f_a previously used to indicate the three db loss point cannot be used where the inverse arms of equalizers contain some resistive elements for the reason that such equalizers need not have even this amount of loss

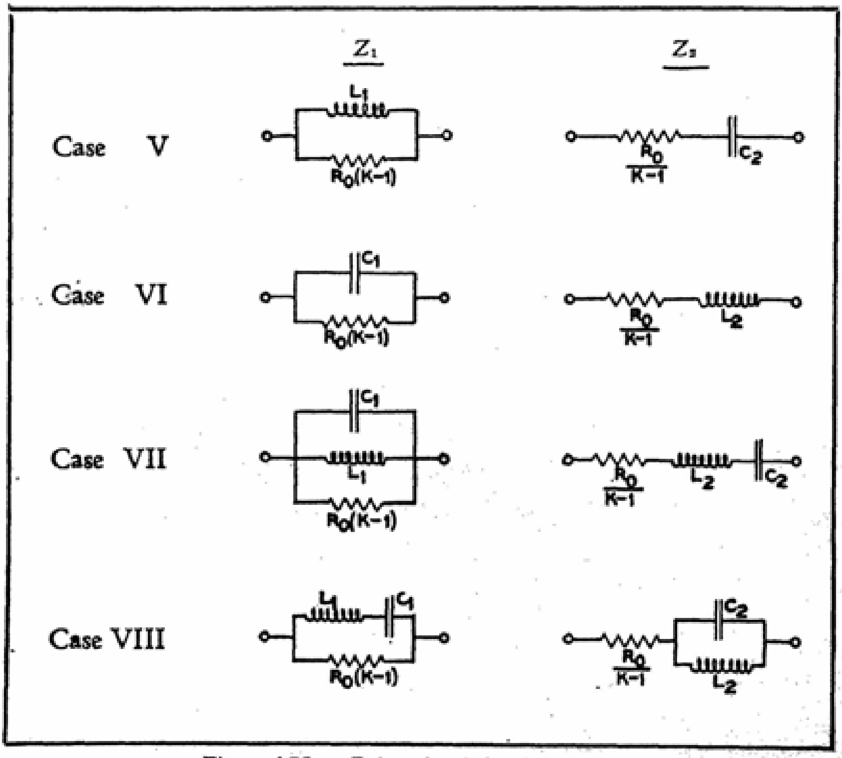
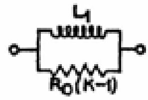


Figure 175 — Pairs of resistive and reactive inverse arms

Immediately following, insertion loss formulae are developed and chart data presented using the pairs of inverse arms of Figure 175.

(a) Case V

When $Z_1 =$ 

and $Z_2 =$ 

In this case X_2 of the Z_2 arm, is by means of equation (51),

$$X_2 = X_B \frac{f_B}{f}$$

and again letting $X_B = R_0 \dfrac{\sqrt{K}}{K-1}$ we have

$$X_2 = R_0 \frac{\sqrt{K}}{K-1} \frac{f_b}{f}$$

Substitution into equation (106) gives for the insertion loss of networks employing these inverse arms

$$\text{I. L.} = 10 \log \left[1 + \frac{K^2 - 1}{1 + K\left(\frac{f_b}{f}\right)^2} \right] \tag{110}$$

In the same manner as before, C_2 and L_1 become

$$\frac{1}{2\,\pi f_b\, C_2} = R_0 \frac{\sqrt{K}}{K-1}$$

$$C_2 = \frac{1}{2\,\pi f_b\, R_0} \frac{K-1}{\sqrt{K}} = \frac{1}{\omega_b R_0} \frac{K-1}{\sqrt{K}}$$

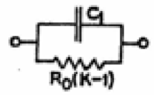
$$\tag{111}$$

$$L_1 = R_0^2\, C_2 = \frac{R_0}{\omega_b} \frac{K-1}{\sqrt{K}}$$

Chart VI shows the insertion loss formula of (110) plotted for various values of the pad loss; that is, for various values of K. Column

V of Chart I gives the equalizer sections obtained when these arms are used with the equalizer types of Figure 166.

(b) Case VI

When $Z_1 =$ 

and $Z_2 =$ 

For this pair of inverse arms we have by means of equation (48)

$$Z_2 = \frac{R_0}{K-1} + jX_B \frac{f}{f_B} \tag{112}$$

At the frequency f_b where the insertion loss of equalizers employing these arms is $10 \log K$, the value of X_B is $R_0 \frac{\sqrt{K}}{K-1}$. The reactive component X_2 of equation (112) then becomes

$$X_2 = \frac{R_0 \sqrt{K}}{K-1} \frac{f}{f_b}$$

Substitution of this value of X_2 into equation (106) gives for the insertion loss formula

$$\text{I.L.} = 10 \log \left[1 + \frac{K^2 - 1}{1 + K\left(\frac{f}{f_b}\right)^2} \right] \tag{113}$$

This insertion loss formula is now in convenient form for plotting. Chart VII shows such a family of curves plotted against the ratio f/f_b for different values of K. At zero frequency, where $f/f_b = 0$, the formula reduces to the pad loss value of $20 \log K$. Chart VII is arranged so that the pad loss varies in one db steps from two db to fourteen db, inclusive. Each curve of the charts is an insertion loss characteristic for a given value of K and plotted against the ratio f/f_b, where f_b is the frequency where the insertion loss is one-half the pad loss. The values of L_2 and C_1 in terms of R_0, K and f_b are secured as follows: At the

frequency f_b the reactance X_2, as already stated, is $R_0\frac{\sqrt{K}}{K-1}$.

Then we have

$$2\pi f_b L_2 = R_0\frac{\sqrt{K}}{K-1}$$

or

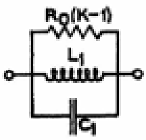
$$L_2 = \frac{R_0}{2\pi f_b}\,\frac{\sqrt{K}}{K-1} = \frac{R_0}{\omega_b}\,\frac{\sqrt{K}}{K-1} \tag{114}$$

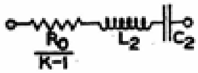
and from the inverse relationship of L_2 and C_1, we have $L_2/C_1 = R_0^2$ or

$$C_1 = \frac{L_2}{R_0^2} = \frac{1}{R_0\,\omega_b}\,\frac{\sqrt{K}}{K-1} \tag{115}$$

Column VI of Chart I shows the equalizers obtained when this pair of inverse arms is used with the equalizer types of Figure 166.

(c) Case VII

When $Z_1 =$ 

and $Z_2 =$ 

The reactive component X_2 of the above Z_2 network is by means of equation (55),

$$X_2 = X_S\frac{\dfrac{f}{f_R}-\dfrac{f_R}{f}}{s-\dfrac{1}{s}}$$

Where f_R is the frequency of resonance of L_2 and C_2. X_S is the reactance at the frequency f_S and $s = \frac{f_R}{f_S}$. Letting $X_S = R_0\frac{\sqrt{K}}{K-1}$ at the frequency f_b and letting $s = b = f_R/f_b$ we have

$$X_2 = R_0\frac{\sqrt{K}}{K-1}\,\frac{\dfrac{f}{f_R}-\dfrac{f_R}{f}}{b-\dfrac{1}{b}}$$

By means of equation (106) we obtain the following insertion loss formula for networks using these inverse arms:

$$\text{I. L.} = 10 \log \left[1 + \frac{K^2 - 1}{1 + K\left(\dfrac{\dfrac{f}{f_R} - \dfrac{f_R}{f}}{b - \dfrac{1}{b}}\right)^2} \right] \tag{116}$$

The values of L_2 and C_2 are secured from equations (56) and (57) if we remember that we have set up by definition the relations $X_s = R_0 \dfrac{\sqrt{K}}{K-1}$, $a = b$ and $f_s = f_b$. We have

$$L_2 = \frac{R_0}{\omega_b} \frac{\sqrt{K}}{K-1} \frac{1}{b^2 - 1} \qquad C_2 = \frac{1}{R_0\, \omega_b} \frac{K-1}{\sqrt{K}} \frac{b^2 - 1}{b^2} \tag{117}$$

And since by the inverse relationship between Z_1 and Z_2, $L_1 = C_2 R_0^2$ and $C_1 = L_2/R_0^2$ we have

$$L_1 = \frac{R_0}{\omega_b} \frac{K-1}{\sqrt{K}} \frac{b^2 - 1}{b^2} \qquad C_1 = \frac{1}{\omega_b R_0} \frac{\sqrt{K}}{K-1} \frac{1}{b^2 - 1} \tag{118}$$

The insertion loss formula of (116) is expressed in terms of the frequency ratio f/f_b and the parameters K and b. To chart this equation it is necessary to make curves for selected values of parameters. This has been done in Charts VIII to XVII, inclusive, where K has values which make the pad loss vary from three db to fourteen db in one db steps.

Column VII of Chart I gives the equalizers obtained when these inverse arms are used with equalizer types of Figure 166.

(d) Case VIII

When $Z_1 =$ 

and $Z_2 =$ 

Development of the insertion loss formula for equalizers using this

pair of inverse arms proceeds in the same manner as for the previous cases. We have for the reactance X_1 of the impedance Z_1

$$X_1 = X_S \frac{\frac{f}{f_R} - \frac{f_R}{f}}{s - \frac{1}{s}}$$

At the frequency f_b, $X_S = R_0 \frac{K-1}{\sqrt{K}}$ and $s = b$. This gives

$$X_1 = R_0 \frac{K-1}{\sqrt{K}} \frac{\frac{f}{f_R} - \frac{f_R}{f}}{b - \frac{1}{b}}$$

When this value of X_1 is used with equation (105) the insertion loss formula becomes

$$\text{I.L.} = 10 \log \left[1 + \frac{K^2 - 1}{1 + K \left(\dfrac{b - \frac{1}{b}}{\frac{f}{f_R} - \frac{f_R}{f}} \right)^2} \right] \tag{119}$$

By means of formulae (56) and (57), and remembering that

$$\frac{L_2}{C_1} = \frac{L_1}{C_2} = R_0^2$$

we have for the electrical constants of the inverse arms

$$L_1 = \frac{R_0}{\omega_b} \frac{K-1}{\sqrt{K}} \frac{1}{b^2 - 1} \qquad L_2 = \frac{R_0}{\omega_b} \frac{\sqrt{K}}{K-1} \frac{b^2-1}{b^2}$$

$$C_1 = \frac{1}{\omega_b R_0} \frac{\sqrt{K}}{K-1} \frac{b^2-1}{b^2} \qquad C_2 = \frac{1}{\omega_b R_0} \frac{K-1}{\sqrt{K}} \frac{1}{b^2-1}$$

Charts XVIII to XXVII, inclusive, give insertion loss characteristics for equalizers using these inverse arms as computed from formula (119). As in the previous case, each set of curves is for a particular pad loss; the range covered varying from three db to fourteen db in one db steps. Column VIII of Chart I shows the equalizer designs obtained when these inverse arms are used with the fundamental equalizer types.

CHART I

(a)

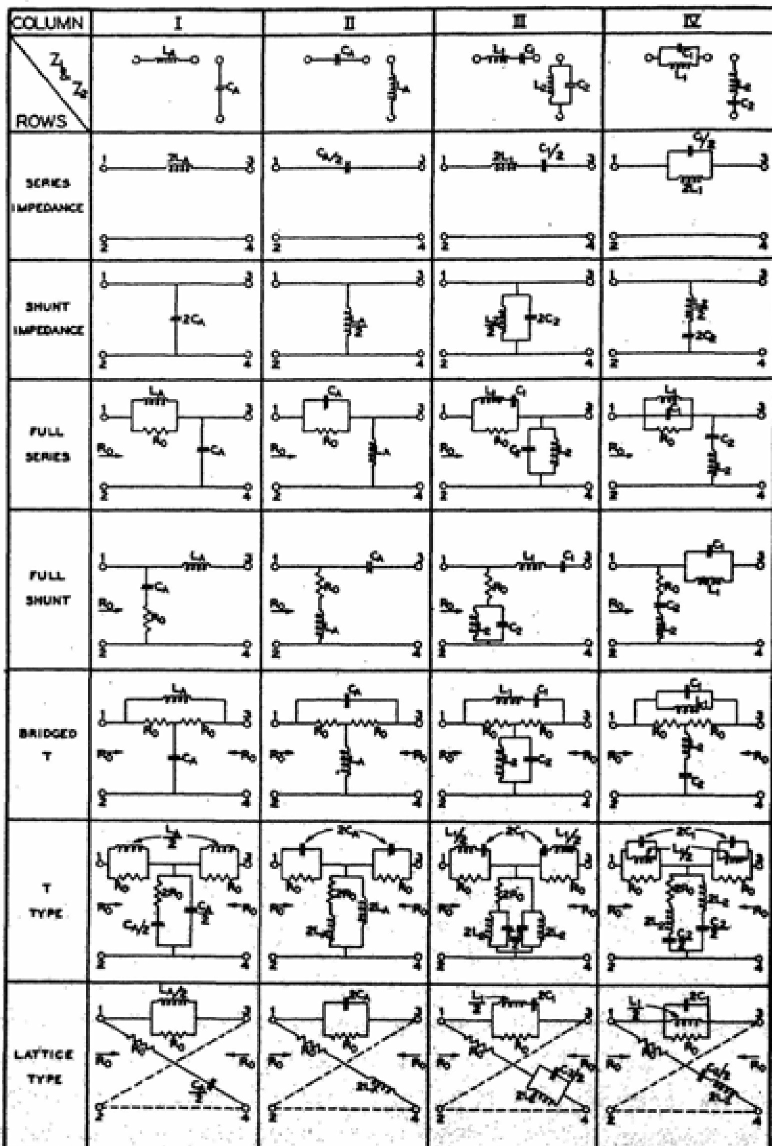

COLUMN
I
II
III
IV
ROWS
SERIES IMPEDANCE
SHUNT IMPEDANCE
FULL SERIES
FULL SHUNT
BRIDGED T
T TYPE
LATTICE TYPE

CHART I

(b)

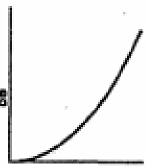
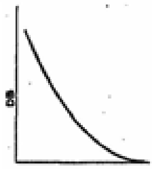
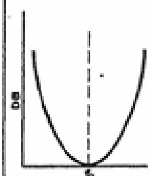
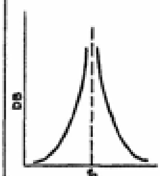
COLUMN	I	II	III	IV
INSERTION LOSS CHARACTERISTIC	 FREQUENCY	 FREQUENCY	 FREQUENCY	 FREQUENCY
REFER TO	CHART II	CHART III	CHART IX	CHART X
CURRENT RATIO $\left\lvert \left[Z_1 / Z_1' \right]^2 \right\rvert$	$1+\left[\frac{f}{f_A}\right]^2$	$1+\left[\frac{f_A}{f}\right]^2$	$1+\left[\frac{\frac{f}{f_R}-\frac{f_R}{f}}{\alpha-\frac{1}{\alpha}}\right]^2$	$1+\left[\frac{\alpha-\frac{1}{\alpha}}{\frac{f}{f_R}-\frac{f_R}{f}}\right]^2$
DESIGN FORMULAE	$L_A=\frac{R_0}{2\pi f_A}=\frac{R_0}{\omega_A}$ $C_A=\frac{1}{2\pi f_A R_0}=\frac{1}{\omega_A R_0}$ $f_A=\frac{1}{2\pi\sqrt{L_A C_A}}$ $R_0=\sqrt{\frac{L_A}{C_A}}$		$L_1=L_A\frac{1}{\alpha^2-1}$ $L_2=L_A\frac{\alpha^2-1}{\alpha^2}$ $C_1=C_A\frac{\alpha^2-1}{\alpha^2}$ $C_2=C_A\frac{1}{\alpha^2-1}$	$L_1=L_A\frac{\alpha^2-1}{\alpha^2}$ $L_2=L_A\frac{1}{\alpha^2-1}$ $C_1=C_A\frac{1}{\alpha^2-1}$ $C_2=C_A\frac{\alpha^2-1}{\alpha^2}$
NOTES	f_R=RESONANT FREQUENCY OF Z_1 & Z_2 ARMS f_A=FREQUENCY OF 3 DB INSERTION LOSS f=ANY FREQUENCY $\alpha=\frac{f_R}{f_A}$=DEFINED AS GREATER THAN UNITY R_0=EQUALIZER RESISTANCE INSERTION LOSS=10 LOG $\left\lvert\left[\frac{Z_1}{Z_1'}\right]^2\right\rvert$ L=INDUCTANCE IN HENRIES C=CAPACITANCE IN FARADS			

CHART I

(c)

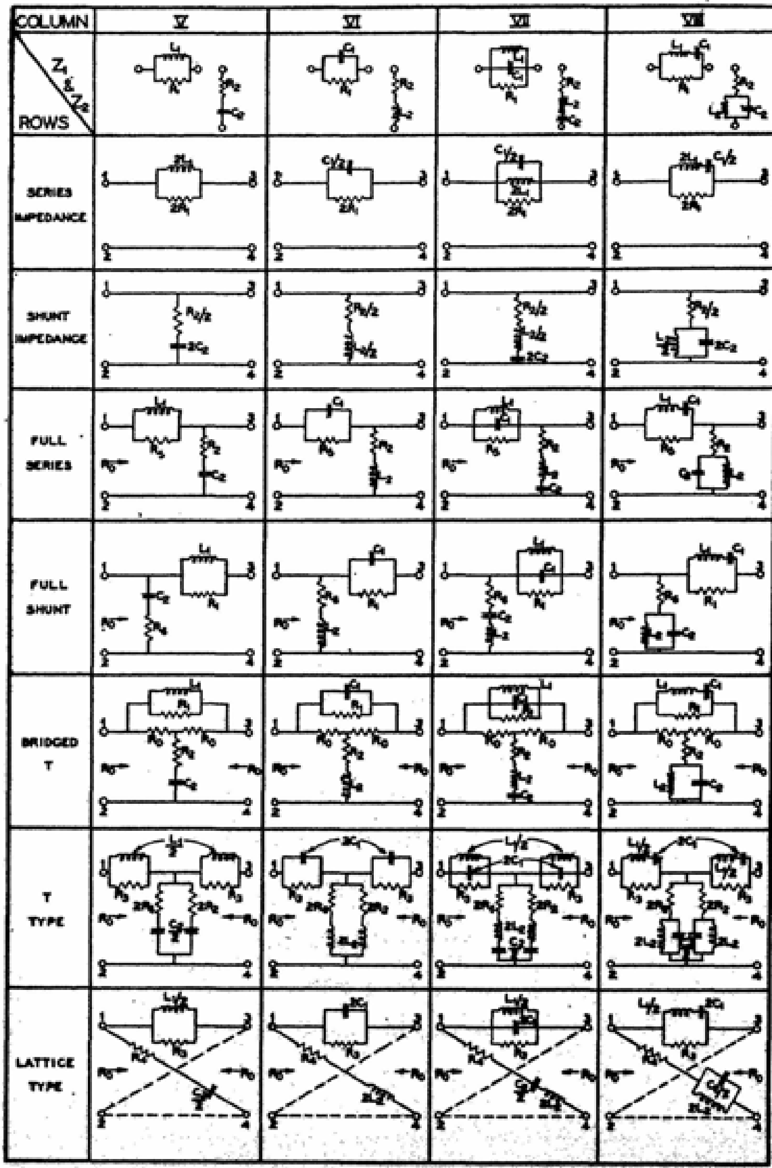

COLUMN
V
VI
VII
VIII
ROWS
SERIES IMPEDANCE
SHUNT IMPEDANCE
FULL SERIES
FULL SHUNT
BRIDGED T
T TYPE
LATTICE TYPE

CHART I

(d)

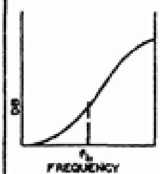
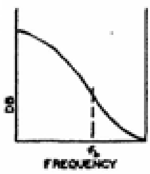
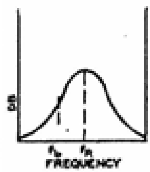
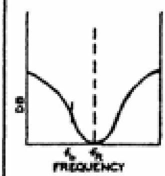
COLUMN	V	VI	VII	VIII
INSERTION LOSS CHARACTERISTIC				
REFER TO	CHART XI	CHART XII	CHARTS [illegible] TO [illegible]	CHARTS [illegible] TO [illegible]
CURRENT RATIO $\left\|\left[I_1/I_2\right]^2\right\|$	$1+\frac{K^2-1}{1+K\left(\frac{f_b}{f}\right)^2}$	$1+\frac{K^2-1}{1+K\left(\frac{f}{f_b}\right)^2}$	$1+\frac{K^2-1}{1+K\left[\frac{\frac{f}{f_R}-\frac{f_R}{f}}{b-\frac{1}{b}}\right]^2}$	$1+\frac{K^2-1}{1+K\left[\frac{b-\frac{1}{b}}{\frac{f}{f_R}-\frac{f_R}{f}}\right]^2}$
DESIGN FORMULAE	$L_1=L_B\frac{K-1}{\sqrt{K}}$ $L_2=L_B\frac{\sqrt{K}}{K-1}$ $C_1=C_B\frac{\sqrt{K}}{K-1}$ $C_2=C_B\frac{K-1}{\sqrt{K}}$		$L_1=L_B\frac{K-1}{\sqrt{K}}\frac{b^2-1}{b^2}$ $L_2=L_B\frac{\sqrt{K}}{K-1}\frac{1}{b^2-1}$ $C_1=C_B\frac{\sqrt{K}}{K-1}\frac{1}{b^2-1}$ $C_2=C_B\frac{K-1}{\sqrt{K}}\frac{b^2-1}{b^2}$	$L_1=L_B\frac{K-1}{\sqrt{K}}\frac{1}{b^2-1}$ $L_2=L_B\frac{\sqrt{K}}{K-1}\frac{b^2-1}{b^2}$ $C_1=C_B\frac{\sqrt{K}}{K-1}\frac{b^2-1}{b^2}$ $C_2=C_B\frac{K-1}{\sqrt{K}}\frac{1}{b^2-1}$
	FOR ALL NETWORKS $R_0=\sqrt{\frac{L_B}{C_B}}$ $R_3=R_0\frac{K-1}{K+1}$ $L_B=\frac{R_0}{2\pi f_b}=\frac{R_0}{\omega_b}$	$R_1=R_0(K-1)$ $R_4=R_0\frac{K+1}{K-1}$ $C_B=\frac{1}{2\pi f_b R_0}=\frac{1}{\omega_b R_0}$	$R_2=R_0\frac{1}{K-1}$ $R_5=R_0\frac{K-1}{K}$ $f_R=\frac{1}{2\pi\sqrt{L_B C_B}}$	$R_6=R_0\frac{K}{K-1}$
NOTES	f_R=RESONANT FREQUENCY OF Z_1 & Z_2 ARMS f_b=FREQUENCY OF ONE-HALF PAD LOSS f=ANY FREQUENCY $b=\frac{f_R}{f_b}$=DEFINED AS GREATER THAN UNITY INSERTION LOSS=10 LOG $\left\|\left[\frac{I_1}{I_2}\right]^2\right\|$ PAD LOSS=MAXIMUM LOSS=20 LOG K L=INDUCTANCE IN HENRIES C=CAPACITANCE IN FARADS R_0=EQUALIZER RESISTANCE			

CHART II

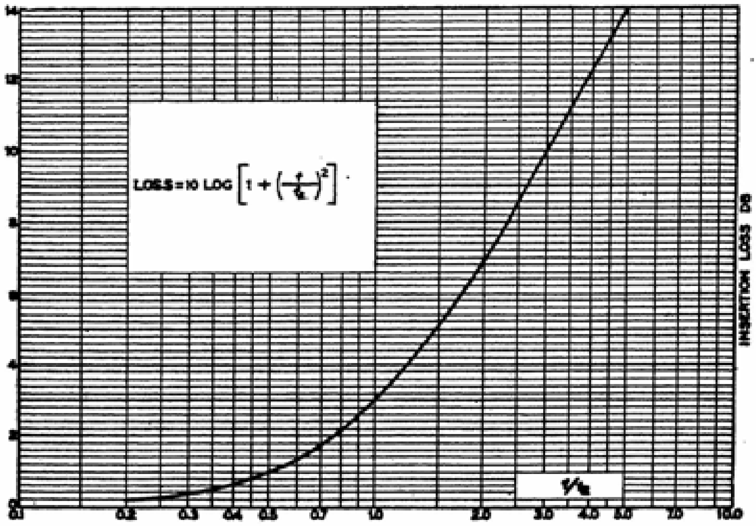


CHART III

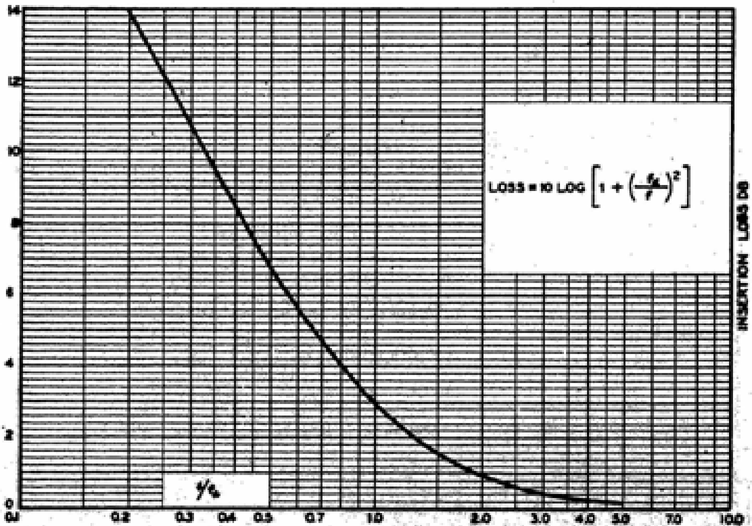


CHART IV

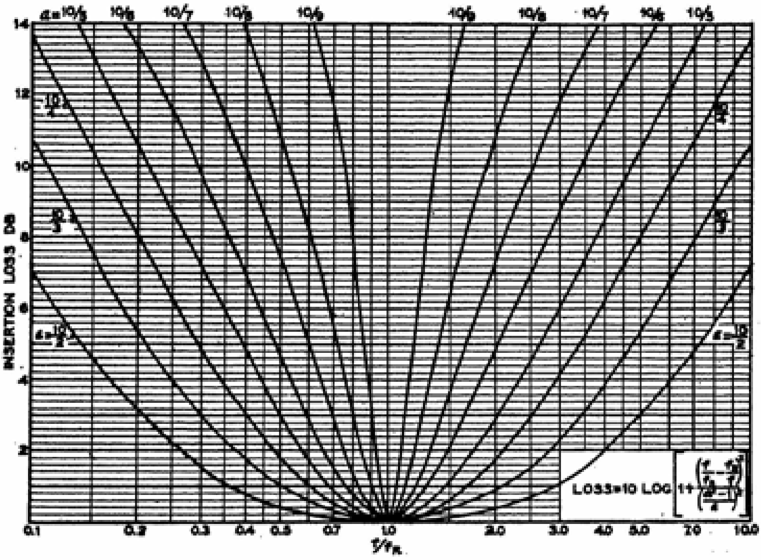


CHART V

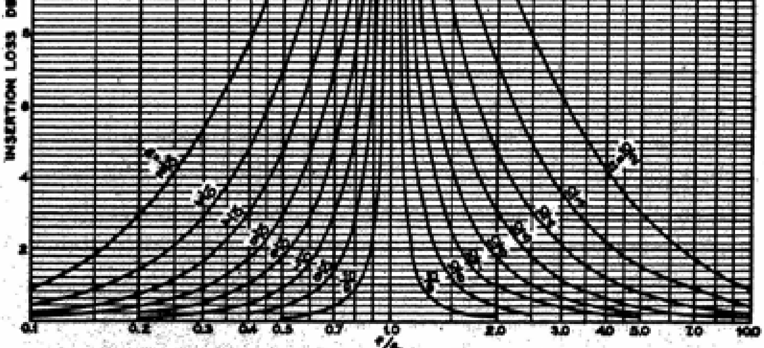
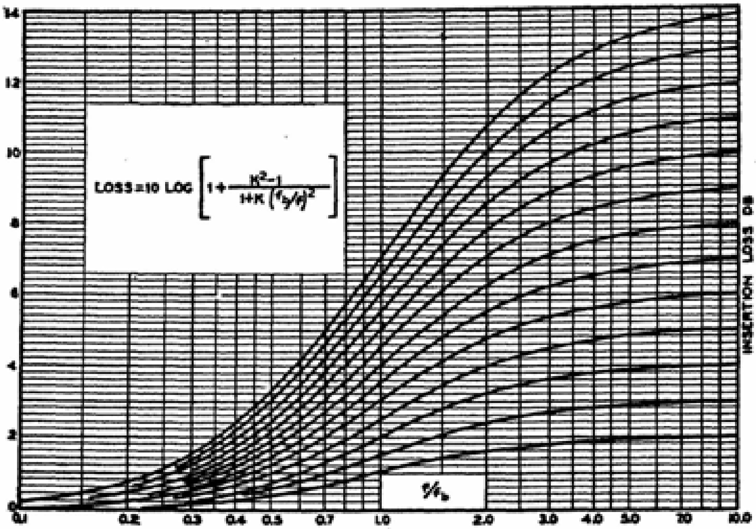
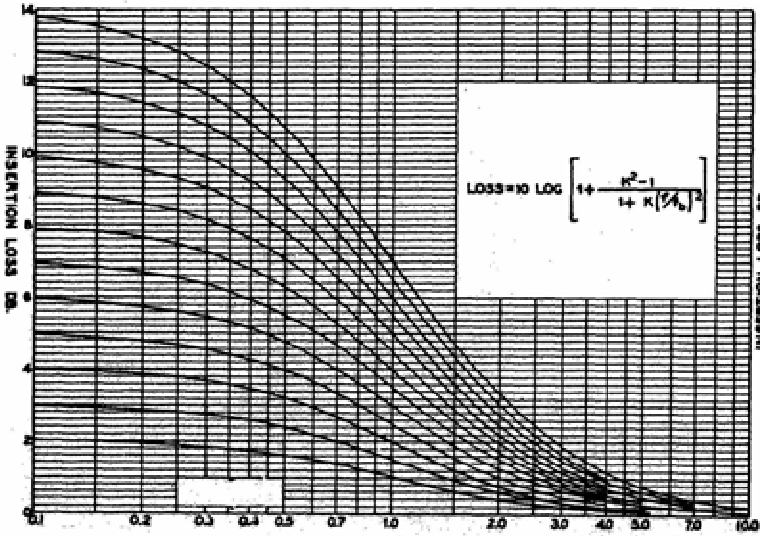


CHART VI



f/f_b

CHART VII



f/f_b

CHART VIII

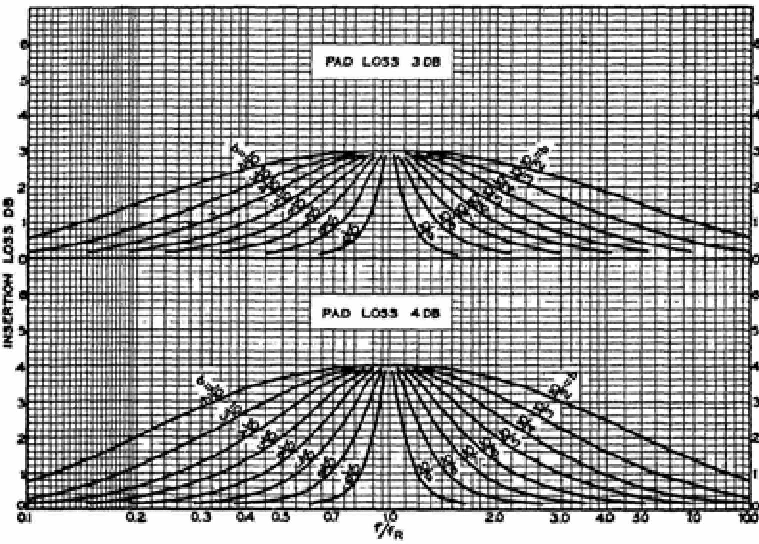


CHART IX

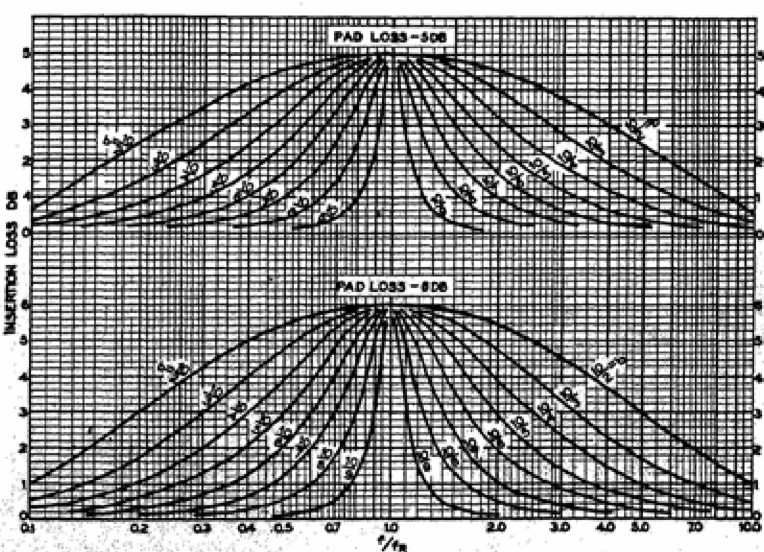


CHART X

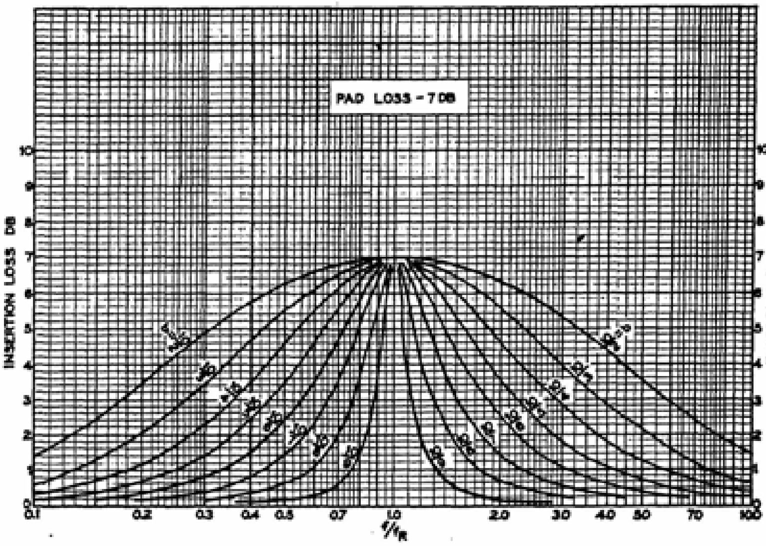


CHART XI

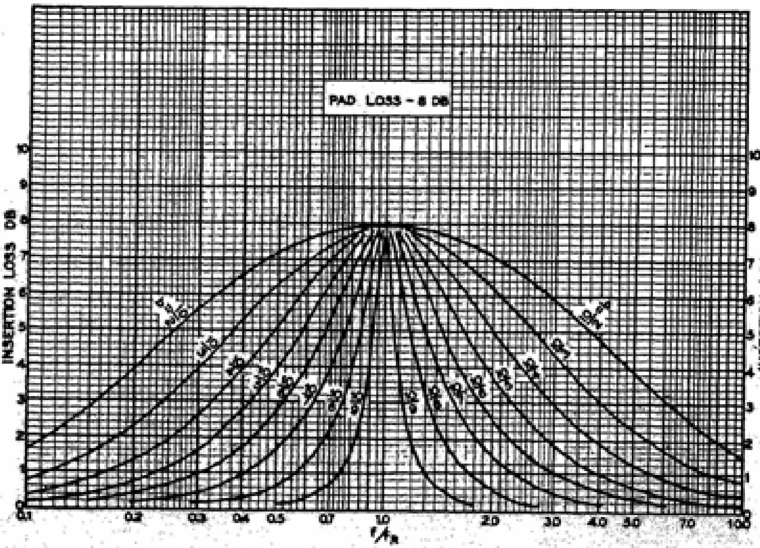


CHART XII

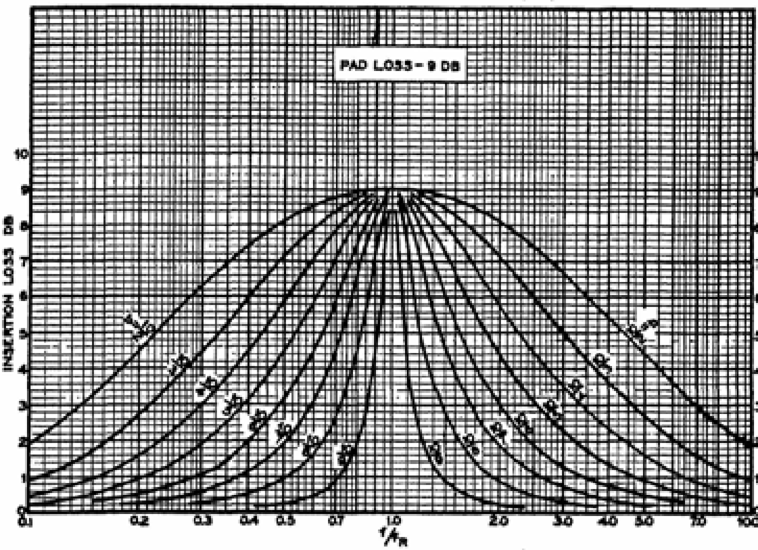


CHART XIII

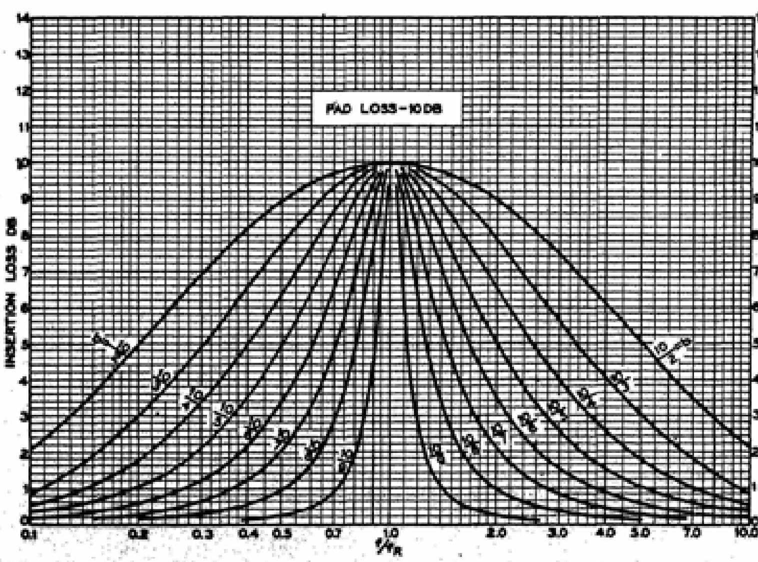


CHART XIV

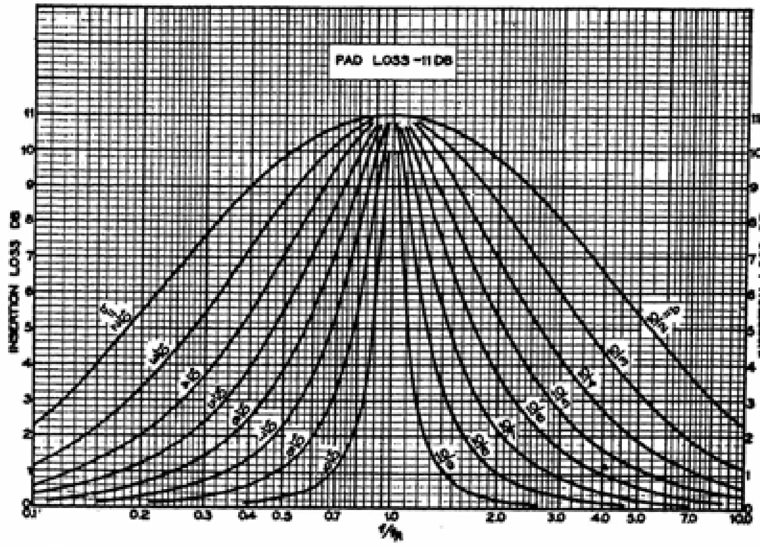


CHART XV

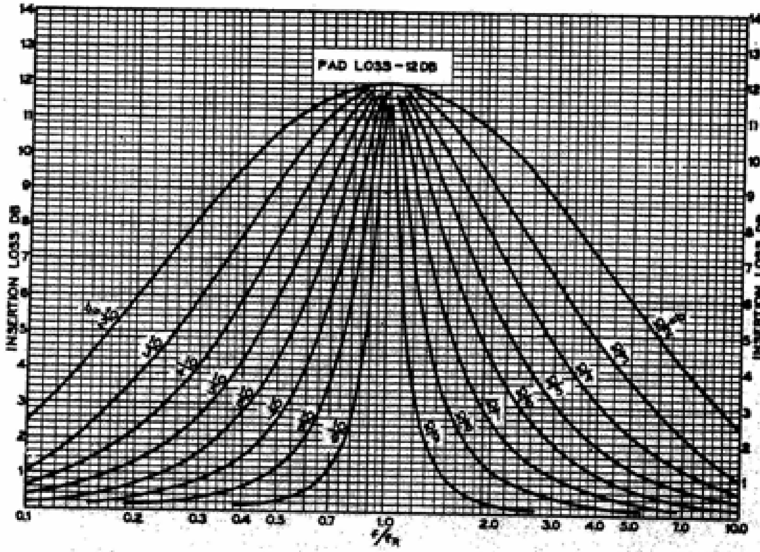


CHART XVI

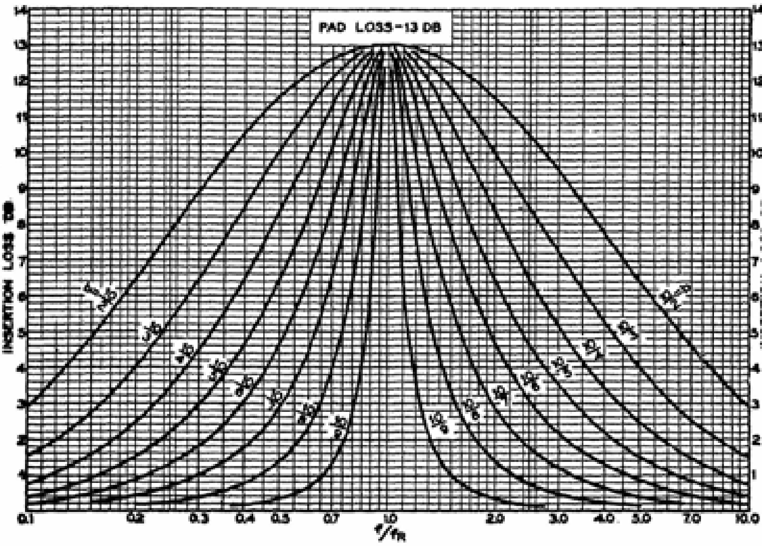


CHART XVII

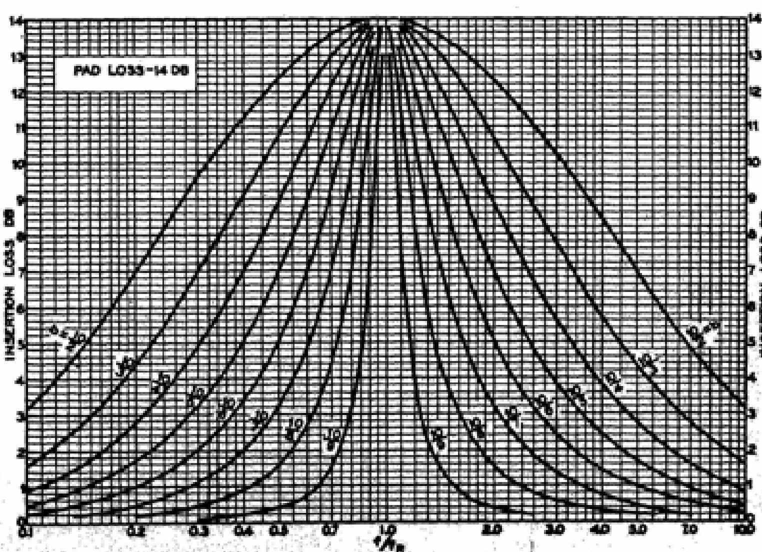


CHART XVIII

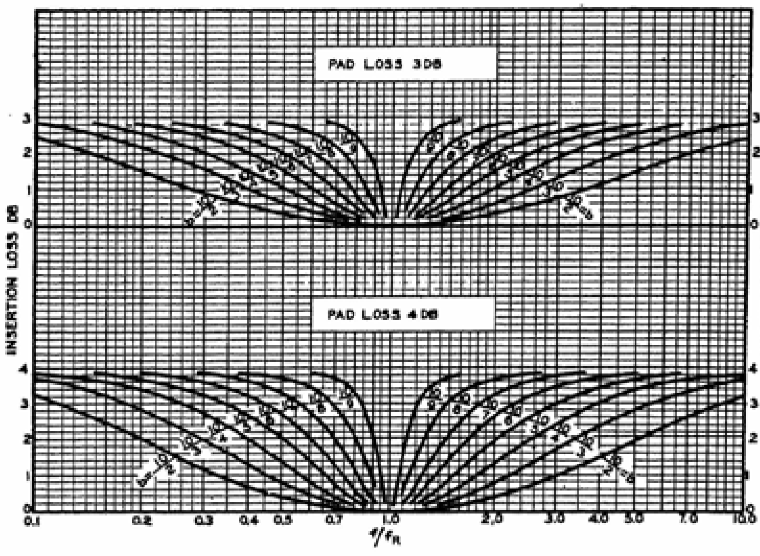


CHART XIX

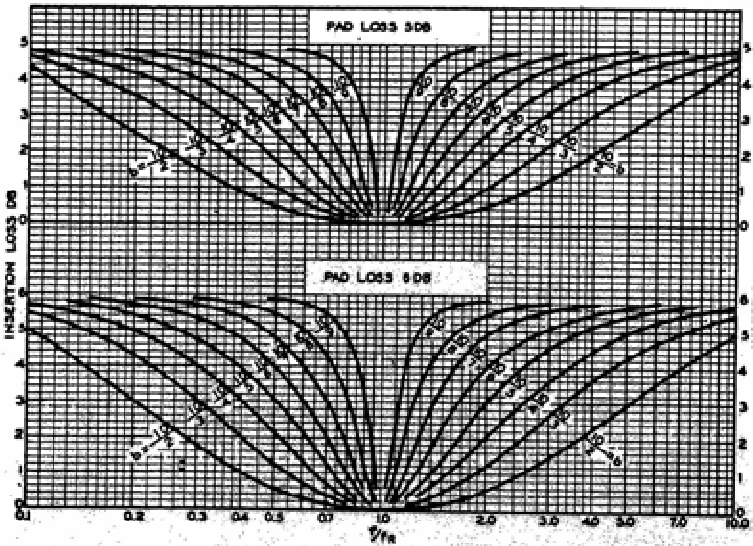


CHART XX

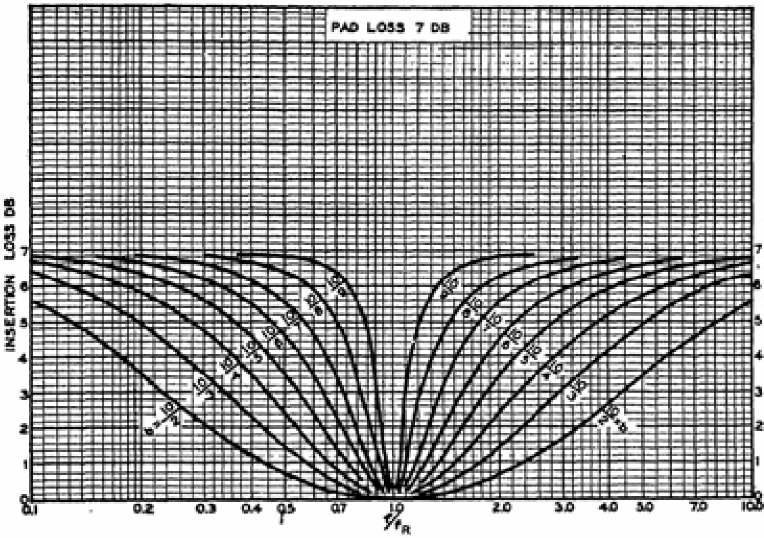


CHART XXI

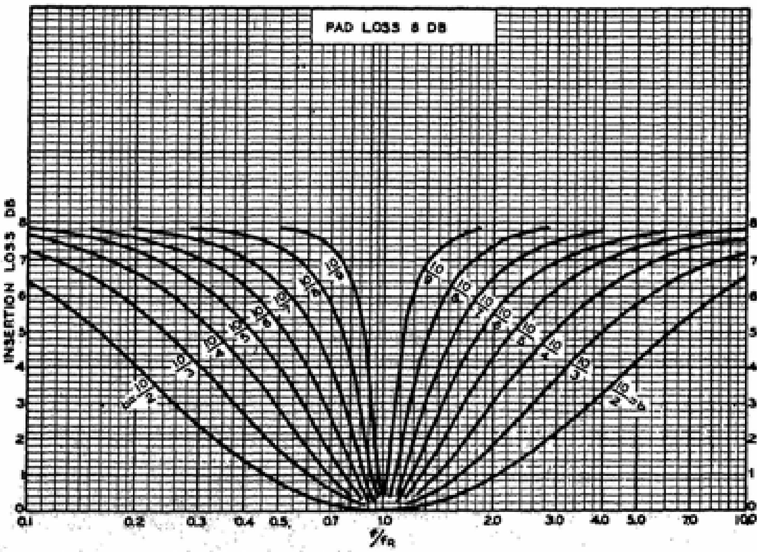


CHART XXII

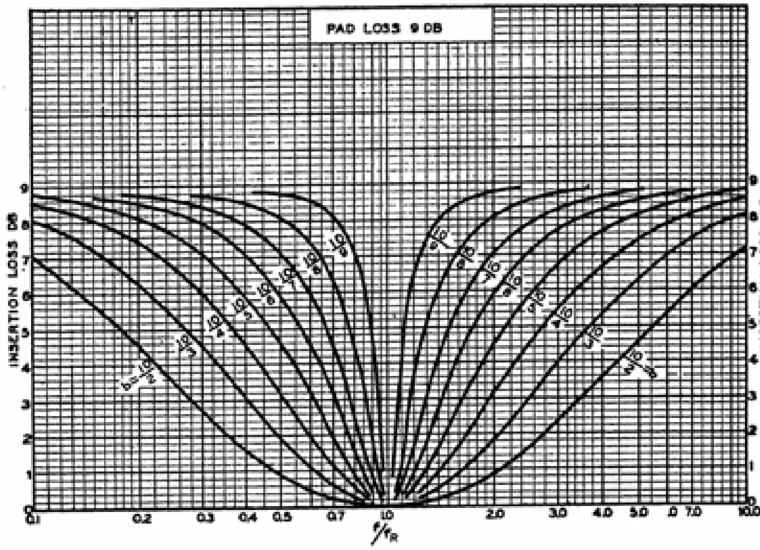


CHART XXIII

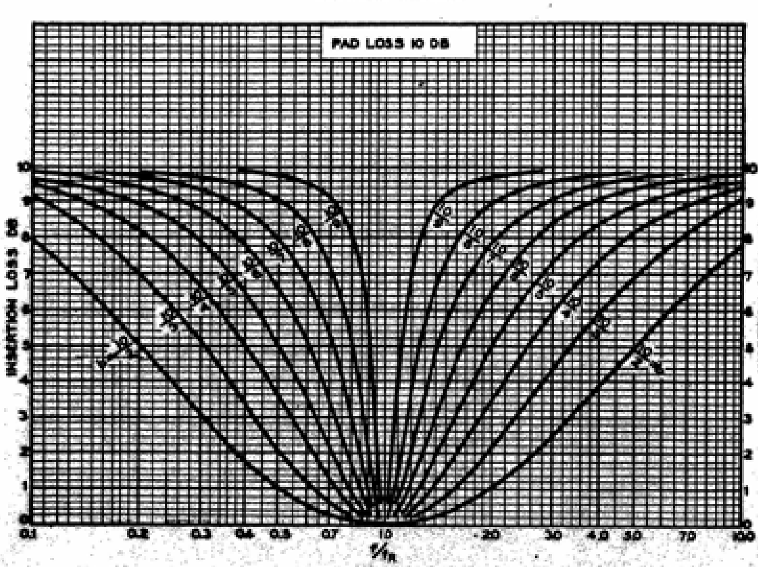


CHART XXIV

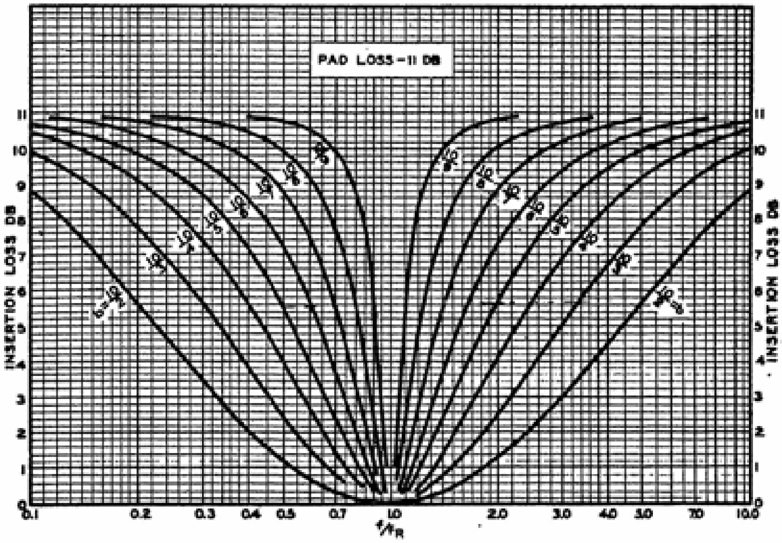


CHART XXV

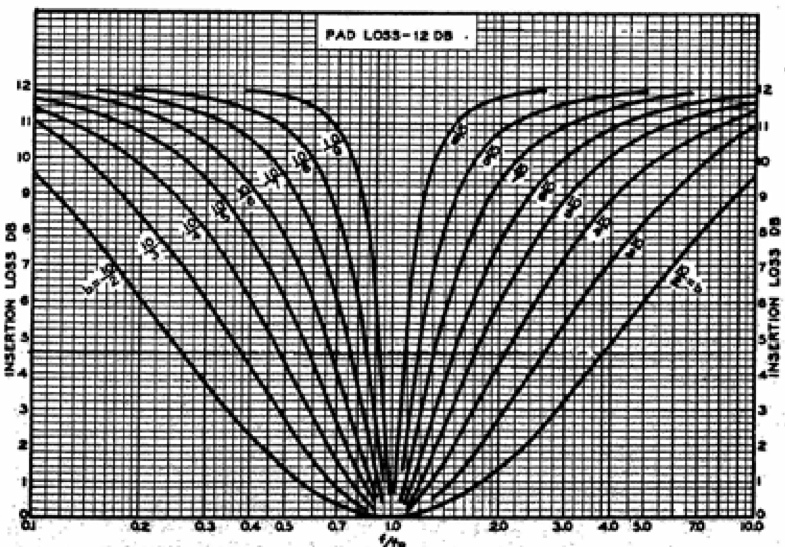


CHART XXVI

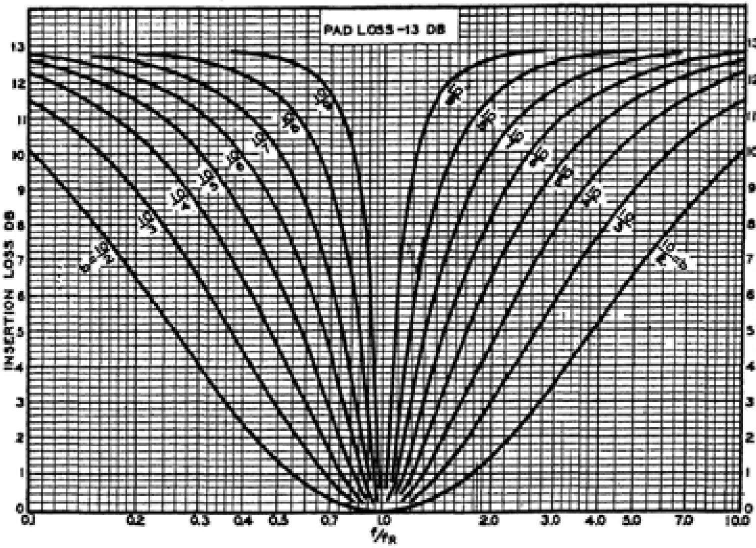
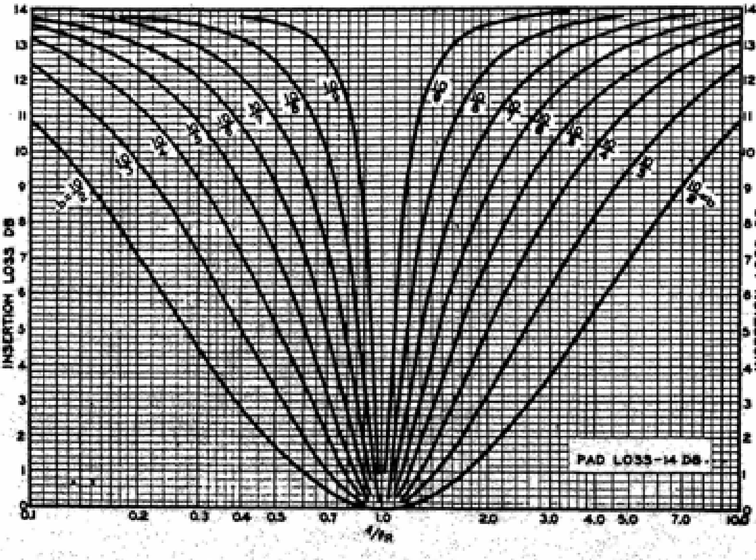


CHART XXVII



Chapter XVII

EQUALIZER DESIGN

By HARRY KIMBALL

The preceding chapter has made available design formulae and charts for the engineering of a number of useful equalizer types. The material is arranged in a manner permitting its use, in most instances, without reference to the underlying theoretical bases, although for best results one should be familiar with the mathematical development work. In using the charts it is necessary to be familiar with the symbols employed in presenting the data in order to formulate network requirements to the best advantage and to carry on the subsequent design work. With this in mind the different symbols used for the charts are summarized below:

$20 \log K =$ Pad or attenuator insertion loss of equalizers where this feature is included in the design. When the pad loss is known the value of K may be computed therefrom.

$f =$ Any frequency (c.p.s.).

$f_R =$ Frequency of resonance or anti-resonance of the Z_1 and Z_2 arms (c.p.s.).

$f_a =$ Frequency of 3.0 db insertion loss (c.p.s.).

$f_b =$ Frequency where one-half the pad loss of an equalizer is obtained (c.p.s.).

$a = \dfrac{f_R}{f_a} =$ Always greater than unity.

$b = \dfrac{f_R}{f_b} =$ Always greater than unity.

$R_0 =$ Connecting system resistance.

In the work following, ten examples are given to show the use of the charts. These examples have been selected to illustrate different features encountered in practice, and in some cases the designs given are those actually used for certain equalizers of sound systems. The primary purpose of the examples is to familiarize the reader with the detail work involved in using the charts to arrive at equalizer designs which

meet system requirements, and in estimating the equalizer insertion loss. In practice, the determination of the electrical elements of an equalizer is usually the last item on the work program, since, when the design parameters to give the required equalization characteristic are known, computing the electrical elements is merely a matter of working through some formulae.

As already mentioned, Chart I has eight columns, each of which contains equalizer types having identical insertion losses. In order to facilitate the following work, the equalizers of Column I, for instance, are referred to as Type I equalizers and so on for the remainder of the columns.

Example 1. Design and show the insertion loss characteristic for a bridged T, Type I equalizer, having 3 db insertion loss at 4,000 c.p.s. and working in a 500 ohm circuit.

From these data we have

$$R_0 = 500 \text{ ohms} \qquad f_a = 4{,}000 \text{ c.p.s.}$$

Reference to Column I of Chart I gives the electrical constants of the equalizer as

$$L_A = \frac{R_0}{2\pi f_a} = \frac{500}{2\pi \times 4{,}000} = 0.0199 \text{ henries}$$

$$C_A = \frac{10^6}{2\pi f_a R_0} = \frac{1{,}000{,}000}{2\pi \times 4{,}000 \times 500} = 0.0795 \text{ microfarads}$$

Using these electrical constants with the bridged T, Type I equalizer of Chart I, we secure the equalizer circuit of Figure 176. The insertion loss characteristic shown is obtained by arbitrarily selecting a number of frequencies, computing the frequency ratios f/f_a and reading the corresponding insertion losses from Chart II as follows:

Frequency (c.p.s.)	f/f_a	*Insertion Loss (db)*
1,000	0.25	0.2
2,000	0.5	1.0
3,000	0.75	2.1
4,000	1.0	3.0
5,000	1.25	4.0
6,000	1.5	5.1
7,000	1.75	6.0
10,000	2.5	8.5

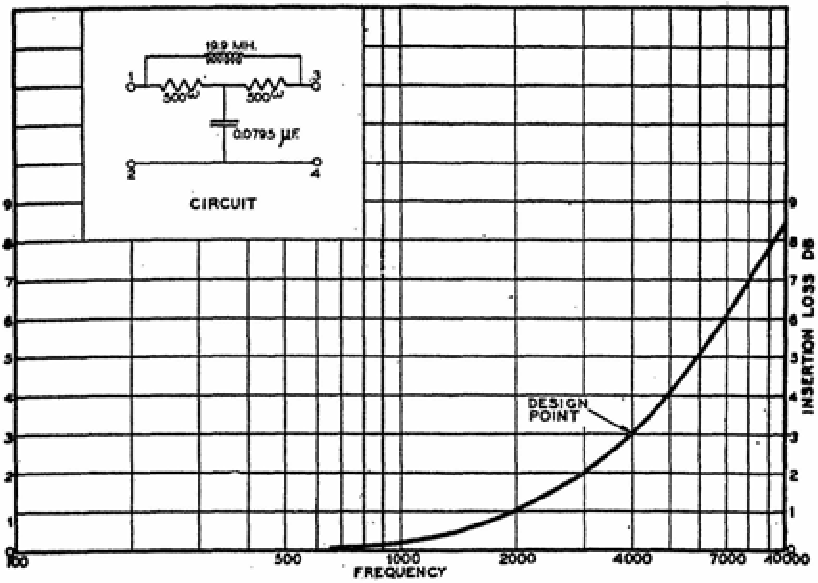


Figure 176 — Circuit and insertion loss characteristic of Example 1.

Example 2. Design and show the insertion loss characteristic for a full-shunt, Type I equalizer having 6 db insertion loss at 5,000 c.p.s., and working in a 200 ohm circuit.

Reference to Chart II shows that 6 db insertion loss gives, $\dfrac{f}{f_a} = 1.7$.

We then have

$$\frac{f}{f_a} = \frac{5{,}000}{f_a} = 1.7$$

That is $f_a = \dfrac{5{,}000}{1.7} = 2{,}940$ c.p.s. and $R_0 = 200$ ohms. From Column I of Chart I we may now compute

$$L_A = \frac{R_0}{2\,\pi\, f_a} = \frac{200}{2\,\pi \times 2{,}940} = 0.0108 \text{ henries}$$

$$C_A = \frac{10^6}{2\,\pi\, f_a R_0} = \frac{1{,}000{,}000}{2\,\pi \times 2{,}940 \times 200} = 0.270 \text{ microfarads}$$

The insertion loss at specific frequencies is secured by selecting frequencies and computing as follows:

Frequency (c.p.s.)	f/f_a	*I.L. from Chart II*
1,000	0.34	0.4
2,000	0.68	1.6
3,000	1.02	3.1
4,000	1.36	4.5
5,000	1.70	6.0
7,000	2.38	8.4
10,000	3.4	11.0

Figure 177 shows the full-shunt equalizer for these values of the electrical elements and the corresponding insertion loss characteristic.

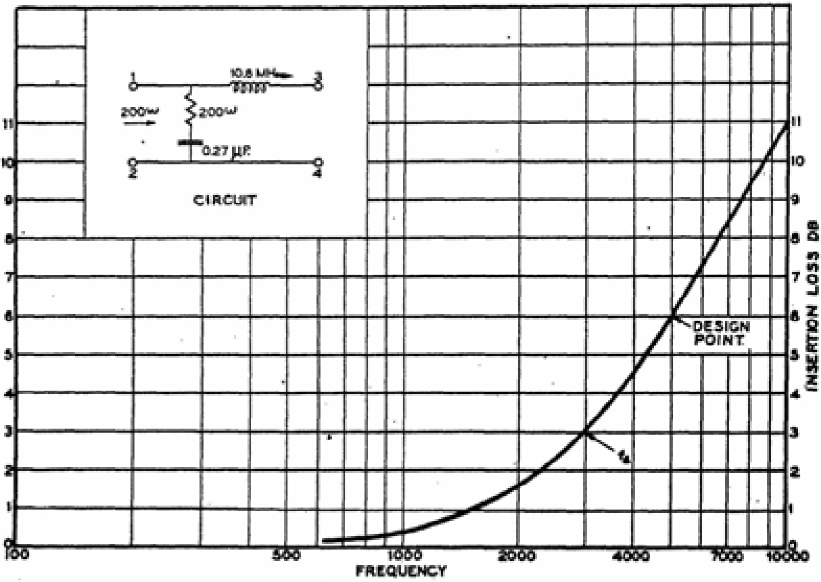


Figure 177 — Circuit and insertion loss characteristic of Example 2.

Example 3. Design and show the insertion loss characteristic for a bridged T, Type II equalizer having 7 db insertion loss at 50 c.p.s. when operating in a 125 ohm circuit.

Referring to Chart III, which applies to the Type II equalizers, we have for 7 db loss, $\frac{f}{f_a} = 0.5$. Since, for this condition $f = 50$ c.p.s., the value of f_a becomes

$$f_a = \frac{50}{0.5} = 100 \text{ c.p.s. and } R_0 = 125 \text{ ohms}$$

Then, from Column II of Chart I, the equalizer electrical elements are

$$L_A = \frac{R_0}{2\pi f_a} = \frac{125}{2\pi \times 100} = 0.199 \text{ henries}$$

$$C_A = \frac{10^6}{2\pi f_a R_0} = \frac{1{,}000{,}000}{2\pi \times 100 \times 125} = 12.7 \text{ microfarads}$$

Arbitrarily selecting frequencies, computing the corresponding frequency ratios, and using Chart III to determine the associated insertion losses, gives the data below:

Frequency (c.p.s.)	f/f_a	*Insertion Loss*
30	0.3	10.8
40	0.4	8.6
50	0.5	7.0
60	0.6	5.8
80	0.8	4.0
100	1.0	3.0
150	1.5	1.6
200	2.0	1.0
400	4.0	0.2

Figure 178 shows the bridged T equalizer obtained by using these electrical elements and its insertion loss characteristic.

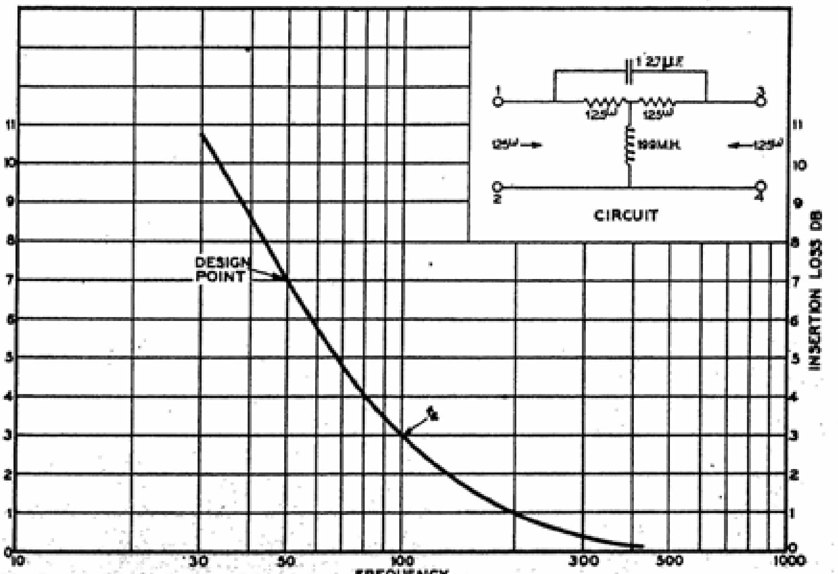


Figure 178 — Circuit and insertion loss characteristic of Example 3.

Example 4. Design and show the insertion loss data for a Type IV equalizer for use in a 500 ohm circuit, having 3 db insertion loss at 5,400 c.p.s. and with the Z_1 and Z_2 arms resonant at 9,000 c.p.s.

From these data we have

$f_R = 9{,}000$ c.p.s. $\qquad a = \frac{9{,}000}{5{,}400} = 1.667 = \frac{10}{6}$

$f_a = 5{,}400$ c.p.s. $\qquad \frac{a^2 - 1}{a^2} = \frac{1.78}{2.78} = 0.64$

$$\frac{1}{a^2 - 1} = \frac{1}{1.78} = 0.56$$

Referring to the Type IV equalizers of Chart I, the electrical constants are computed as follows:

$$L_A = \frac{R_0}{2\pi f_a} = \frac{500}{2\pi \times 5{,}400} = 0.0147 \text{ henries} = 14.7 \text{ m.h.}$$

$$C_A = \frac{10^6}{2\pi f_a R_0} = \frac{1{,}000{,}000}{2\pi \times 5{,}400 \times 500} = 0.059\ \mu\text{f.}$$

$$L_1 = L_A \frac{a^2 - 1}{a^2} = 14.7 \times 0.64 = 9.41 \text{ m.h.}$$

$$L_2 = L_A \frac{1}{a^2 - 1} = 14.7 \times 0.56 = 8.25 \text{ m.h.}$$

$$C_1 = C_A \frac{1}{a^2 - 1} = 0.059 \times 0.56 = 0.033\ \mu\text{f.}$$

$$C_2 = C_A \frac{a^2 - 1}{a^2} = 0.059 \times 0.64 = 0.038\ \mu\text{f.}$$

Insertion loss data may now be read from Chart V, on the curve designated as $a = \frac{10}{6}$. Selecting a number of frequencies and computing ratios for $f_R = 9{,}000$ gives the following insertion loss data:

Frequency	f/f_R	*Insertion Loss (db)*
2,000	0.22	0.2
3,000	0.33	0.6
4,000	0.45	1.4
5,000	0.55	2.3
6,000	0.67	4.2
7,000	0.78	7.4
8,000	0.89	13.4
9,000	1.00	Infinite

Figure 179 shows this equalizer and its insertion loss characteristic.

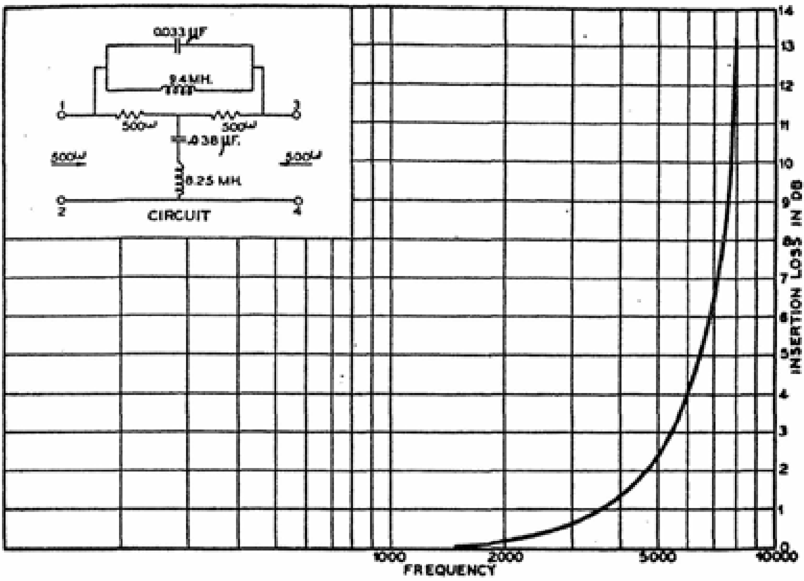


Figure 179 — Circuit and insertion loss characteristic of Example 4.

Example 5. Design and determine the insertion loss characteristic for a bridged T, Type IV equalizer working in a 200 ohm circuit and having a 4 db loss at 5,000 c.p.s., and an 8 db loss at 8,000 c.p.s.

Here again Chart V shows the insertion loss data applicable to this type of equalizer. In this example, however, the data given is not in a form permitting the immediate selection of any one of the insertion loss curves. Before the equalizer can be designed means must be found for determining any two of the three quantities f_a, f_R and a. This could be done by setting up and solving a pair of simultaneous equations based upon the loss data given, but this method is laborious and not entirely practical. The following graphical method is more satisfactory in most cases.

Prepare the following table where a is allowed to successively take on the values given in Chart V, that is, 10/2, 10/3 and so on. Using the 4 db loss point which occurs at a frequency of 5,000 c.p.s., prepare a second column showing the value of the ratio $\left(\frac{5{,}000}{f_R}\right)$, as read from the curves for each value of a. From this column compute a third column showing the ratio $\left(\frac{7{,}000}{f_R}\right) = 1.4\left(\frac{5{,}000}{f_R}\right)$. Now read the inser-

tion losses from the different "a" curves for each value of $\left(\frac{7,000}{f_R}\right)$. The particular value of $\left(\frac{7,000}{f_R}\right)$ which gives close to 8 db insertion loss is the correct one to use.

a	$\frac{5,000}{f_R}$ (4 db loss)	$\frac{7,000}{f_R}$	Insertion Loss at 7,000 c.p.s.	f_R	f_a
10/2	0.24	0.336	6.2		
10/3	0.36	0.50	7.0		
10/4	0.46	0.65	8.2	11,000	4,400
10/5	0.56	0.79	9.8		
10/6	0.66	0.925			
10/7	0.75	1.05			
10/8	0.84				
10/9	0.93				

Referring to the above table, a value of a equal to 10/4 gives 8.2 db at 7,000 c.p.s. The correct value of a is therefore a little larger than 10/4. For this problem, however, we will neglect this small increase. The value of f_R, corresponding to $a = 10/4$, is $f_R = 11,000$ c.p.s. and f_a is 4,400 c.p.s. In actual practice it is not always necessary to prepare the above table as the different steps involved can be done mentally.

With these data the design work for the equalizer can be completed.

We have

$$a = \frac{10}{4} \qquad \frac{a^2 - 1}{a^2} = 0.84 \qquad \frac{1}{a^2 - 1} = 0.19$$

and from Column IV of Chart I

$$L_A = \frac{R_0}{2\pi f_a} = \frac{200}{2\pi \times 4,400} = 0.00725 \text{ henries} = 7.25 \text{ m.h.}$$

$$C_A = \frac{10^6}{2\pi f_a R_0} = \frac{1,000,000}{2\pi \times 4,400 \times 200} = 0.181\ \mu\text{f.}$$

$$L_1 = L_A \frac{a^2 - 1}{a^2} = 7.25 \times 0.84 = 6.1 \text{ m.h.}$$

$$L_2 = L_A \frac{1}{a^2 - 1} = 7.25 \times 0.19 = 1.37 \text{ m.h.}$$

$$C_1 = C_A \frac{1}{a^2 - 1} = 0.181 \times 0.19 = 0.0344\ \mu\text{f.}$$

$$C_2 = C_A \frac{a^2 - 1}{a^2} = 0.181 \times 0.84 = 0.152\ \mu\text{f.}$$

Data for plotting the insertion loss curve of this equalizer are obtained by preparing the following table from Chart V for the curve $a = 10/4$:

Frequency	f/f_R	*Insertion Loss*
1,000	0.091	0.2
2,000	0.182	0.6
3,000	0.273	1.4
4,000	0.364	2.4
5,000	0.455	4.0
6,000	0.546	5.6
7,000	0.637	8.0
8,000	0.73	11.0
9,000	0.82	14.0

Figure 180 shows this equalizer and its insertion loss characteristic.

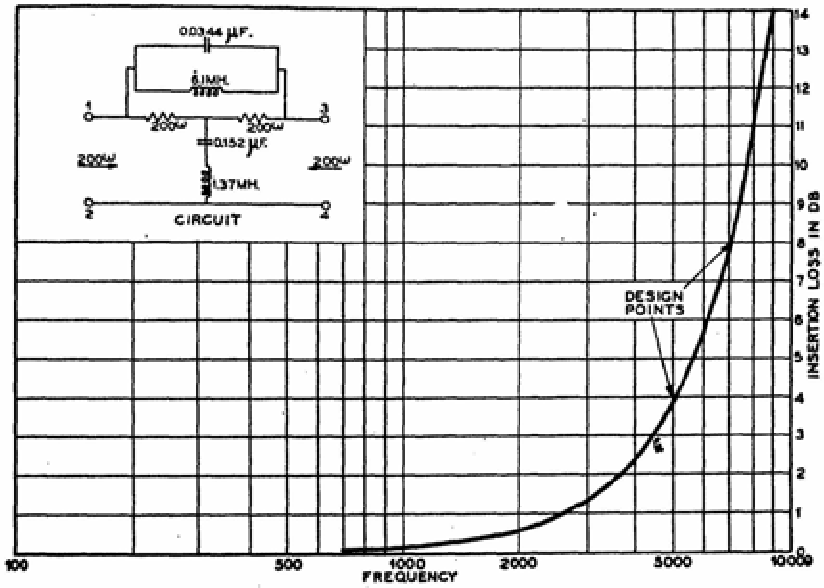


Figure 180 — Circuit and insertion loss characteristic of Example 5.

Example 6. Design a series impedance, Type IV equalizer for use with a 125 ohm circuit and having 4 db insertion loss at 100 c.p.s., and 8 db loss at 60 c.p.s.

For this case, f_R must be of lower frequency than either of the frequencies specified above in order to secure the loss characteristic desired. This means that in the frequency range of interest f/f_R is greater than

one. Referring to Chart V and using the 4 db loss point to prepare the following work table, as in the previous example, we obtain the following results:

a	(4 db Loss) $\frac{100}{f_R}$	$\frac{60}{f_R}$	I.L. at 60 c.p.s.	f_R	f_a
10/2	4.2	2.5	8.0	24	4.8
10/3	2.8	1.68	10.0		

In this case the first value of "a" chosen gave 8.0 db loss at 60 c.p.s., which is the result required. It is noted that the corresponding value of f_a (where 3 db insertion loss is obtained) appears to be very low in the frequency range. As a matter of fact, 3 db loss is obtained for two frequencies; that is, 4.8 c.p.s. and 120 c.p.s. The frequency f_a, however, was defined as being lower in frequency than f_R in order to have "a" always greater than unity. For computation work, therefore, the lower value of f_a must be used although this frequency is of no further interest. We therefore have

$$f_R = 24 \text{ c.p.s.} \qquad \frac{a^2-1}{a^2} = 0.96$$

$$f_a = 4.8 \text{ c.p.s.}$$

$$a = 5 \qquad \frac{1}{a^2-1} = 0.0416$$

From Column IV of Chart I

$$L_A = \frac{R_0}{\omega_a} = \frac{125}{2\pi \times 4.8} = 4.15 \text{ henries}$$

$$C_A = \frac{10^6}{\omega_a R_0} = \frac{1{,}000{,}000}{2\pi \times 4.8 \times 125} = 265\ \mu\text{f.}$$

$$L_1 = L_A \frac{a^2-1}{a^2} = 4.15 \times 0.96 = 4.0 \text{ henries}$$

$$C_1 = C_A \frac{1}{a^2-1} = 265 \times 0.0416 = 11.0\ \mu\text{f.}$$

Figure 181 shows the circuit of the series impedance, Type IV equalizer, which uses these constants and the corresponding insertion loss characteristic obtained in the same manner as for the previous examples. In constructing this equalizer, care should be taken to use an inductance coil having, in the region of 100 c.p.s., an effective resistance which is low compared to the circuit impedance of 125 ohms.

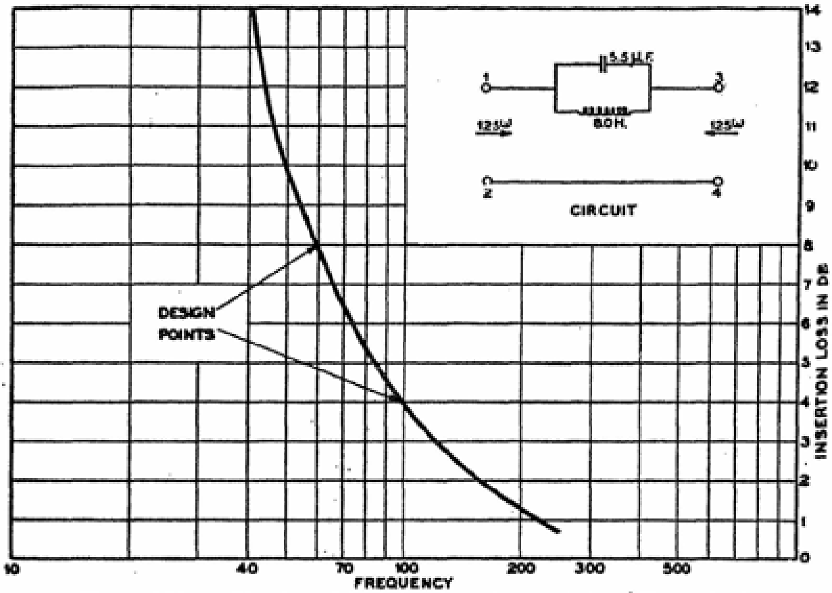


Figure 181 — Circuit and insertion loss characteristic of Example 6.

Example 7. Determine the insertion loss characteristic and proper working conditions for the equalizer given in the circuit diagram of Figure 182.

This is a bridged T, Type I equalizer and its design formulae are as indicated in Column I of Chart I. We have

$$R_0 = \sqrt{\frac{L_A}{C_A}} = \sqrt{\frac{0.018}{0.150}} \times 10^8 = 346 \text{ ohms}$$

$$f_a = \frac{1}{2\pi\sqrt{L_A C_A}} = \frac{159.2}{\sqrt{0.018 \times 0.15}} = 3{,}060 \text{ c.p.s.}$$

The equalizer, then, should work in a 346 ohm circuit and has 3 db loss at 3,060 c.p.s. The complete insertion loss curve, obtained in the manner described in Example 1, is shown in Figure 182.

f	f/f_a	*I.L. from Chart II*
1,000	0.327	0.4
2,000	0.654	1.6
3,000	0.98	2.95
4,000	1.31	4.2
5,000	1.63	5.6
7,000	2.28	7.8
9,000	2.94	9.8

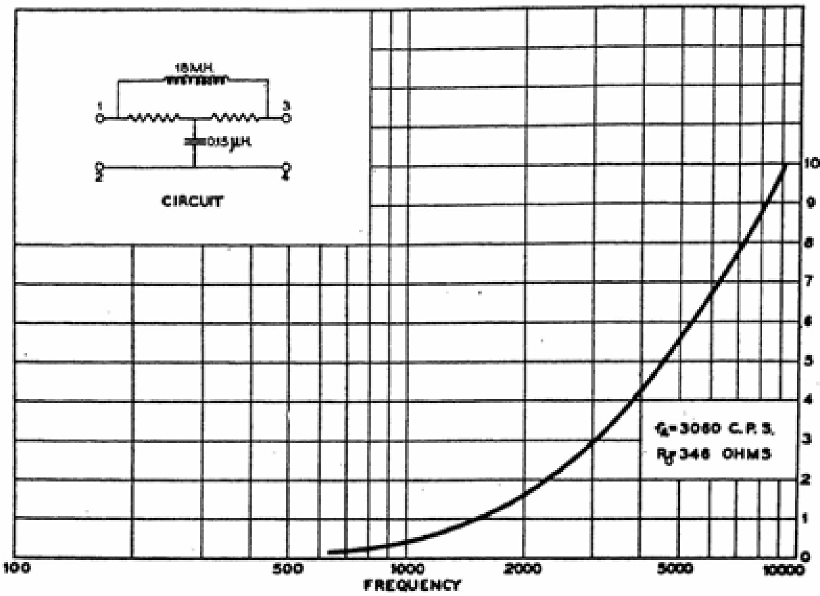


Figure 182 — Circuit of Example 7 with computed insertion loss characteristic.

Example 8.

(a) Design and show the insertion loss characteristic for a bridged T, Type V equalizer, for use in a 500 ohm circuit, having a pad loss of 14 db and a half-loss frequency of 1,000 c.p.s.

(b) Design and show the insertion loss characteristic for a bridged T, Type VI equalizer for use in a 500 ohm circuit, having a pad loss of 14 db and a half-loss frequency of 1,000 c.p.s.

(c) Show the insertion loss characteristic for the two equalizers when connected in tandem.

The data for the two equalizers may be summarized as follows:

$R_0 = 500$ ohms $\qquad \dfrac{K-1}{\sqrt{K}} = 1.79$

$f_h = 1.000$ c.p.s.

$K = 5.01$ where $20 \log K = 14$ db $\qquad \dfrac{\sqrt{K}}{K-1} = 0.559$

For both equalizers we have from Chart I

$$L_B = \frac{R_0}{2\pi f_h} = \frac{500}{2\pi \times 1{,}000} = 0.0795 \text{ henries} = 79.5 \text{ m.h.}$$

$$C_B = \frac{10^6}{2\,\pi\, f_b R_0} = \frac{1{,}000{,}000}{2\,\pi \times 1{,}000 \times 500} = 0.318\ \mu\text{f.}$$

$$R_1 = R_0\ (K - 1) = 500 \times 4.01 = 2{,}000 \text{ ohms}$$

$$R_2 = \frac{R_0}{K-1} = \frac{500}{4.01} = 125 \text{ ohms}$$

For the Type V equalizer we have from Chart I

$$L_1 = L_B \frac{K-1}{\sqrt{K}} = 79.5 \times 1.79 = 142.5 \text{ m.h.}$$

$$C_2 = C_B \frac{K-1}{\sqrt{K}} = 0.318 \times 1.79 = 0.57\ \mu\text{f.}$$

For the Type VI equalizer we have from Chart I

$$L_2 = L_B \frac{\sqrt{K}}{K-1} = 79.5 \times 0.559 = 44.4 \text{ m.h.}$$

$$C_1 = C_B \frac{\sqrt{K}}{K-1} = 0.318 \times 0.559 = 0.178\ \mu\text{f.}$$

Figure 183 shows the circuits of the two equalizers using the above values for the electrical elements, the insertion loss characteristics for each of the equalizers operating alone, and for tandem operation as determined by preparing the table below by means of Charts VI and VII. It is noted that the two equalizers are complementary to each other; that is, their insertion loss when operated in tandem is constant with frequency and equal to the equalizer pad loss.

Frequency (c.p.s.)	f/f_b	*I.L. of Type V Equalizer from Chart VI*	*I.L. of Type VI Equalizer from Chart VII*	*Tandem Loss*
100	0.1	0.2	13.8	14
200	0.2	0.8	13.2	14
300	0.3	1.6	12.4	14
400	0.4	2.4	11.6	14
600	0.6	4.2	9.8	14
800	0.8	5.7	8.3	14
1,000	1.0	7.0	7.0	14
1,500	1.5	9.3	4.7	14
2,000	2.0	10.7	3.3	14
3,000	3.0	12.1	1.9	14
5,000	5.0	13.2	0.8	14
7,000	7.0	13.6	0.4	14

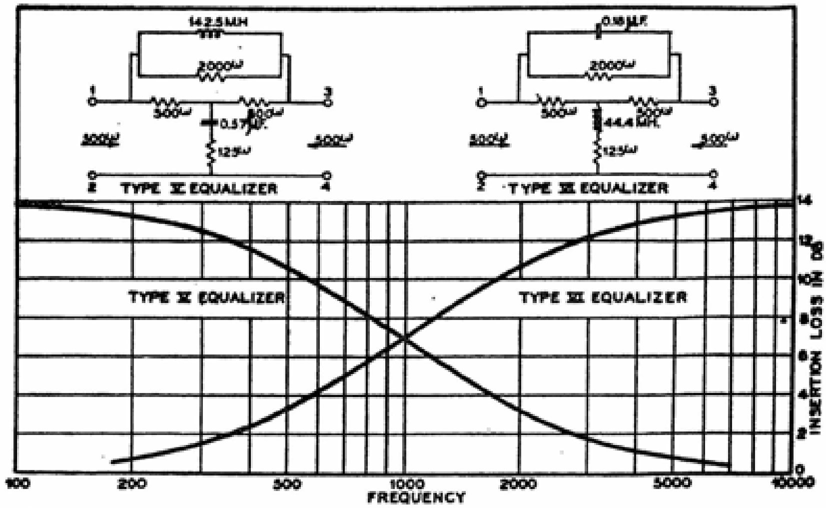


Figure 183 — Circuits of equalizers of Example 8 with computed insertion characteristic loss of each equalizer and the insertion loss characteristic of the two operating in tandem.

Example 9. Design a bridged T, Type VII equalizer for use in a 200 ohm circuit, having its maximum insertion loss of 4 db at 3,000 c.p.s. and its half-loss frequency at 1,500 c.p.s.

We have

$R_0 = 200$ ohms $\qquad f_b = 1{,}500$ c.p.s.

$f_R = 3{,}000$ c.p.s. $\qquad b = 2 = 10/5$

$$K = 1.59 \text{ where } 20 \log K = 4 \text{ db}$$

Referring to Column VII of Chart I, the following terms are needed for computing the electrical elements:

$$\frac{K-1}{\sqrt{K}} = \frac{1.59-1}{1.26} = 0.469 \qquad \frac{\sqrt{K}}{K-1} = \frac{1.26}{0.59} = 2.14$$

$$\frac{b^2-1}{b^2} = 0.75 \qquad \frac{1}{b^2-1} = 0.334$$

$$L_B = \frac{R_0}{\omega_b} = \frac{200}{2\pi \times 1{,}500} = 0.0212 \text{ henries} = 21.2 \text{ m.h.}$$

$$C_B = \frac{10^6}{\omega_b R_0} = \frac{1{,}000{,}000}{2\pi \times 1{,}500 \times 200} = 0.53\ \mu\text{f.}$$

The electrical elements are

$$L_1 = L_B \frac{K-1}{\sqrt{K}} \frac{b^2-1}{b^2} = 21.2 \times 0.469 \times 0.75 = 7.45 \text{ m.h.}$$

$$L_2 = L_B \frac{\sqrt{K}}{K-1} \frac{1}{b^2-1} = 21.2 \times 2.14 \times 0.334 = 15.15 \text{ m.h.}$$

$$C_1 = C_B \frac{\sqrt{K}}{K-1} \frac{1}{b^2-1} = 0.53 \times 2.14 \times 0.334 = 0.378 \ \mu\text{f.}$$

$$C_2 = C_B \frac{K-1}{\sqrt{K}} \frac{b^2-1}{b^2} = 0.53 \times 0.469 \times 0.75 = 0.187 \ \mu\text{f.}$$

$$R_1 = R_0 \ (K-1) = 200 \times 0.59 = 118 \text{ ohms}$$

$$R_2 = \frac{R_0}{K-1} = \frac{200}{0.59} = 340 \text{ ohms}$$

Figure 184 shows the bridged T, Type VII equalizer, using these element values and the corresponding insertion loss characteristic when operating in a 200 ohm circuit. The loss data were obtained from Chart VIII for a value of $a = 10/5$ and for a pad loss of 4 db.

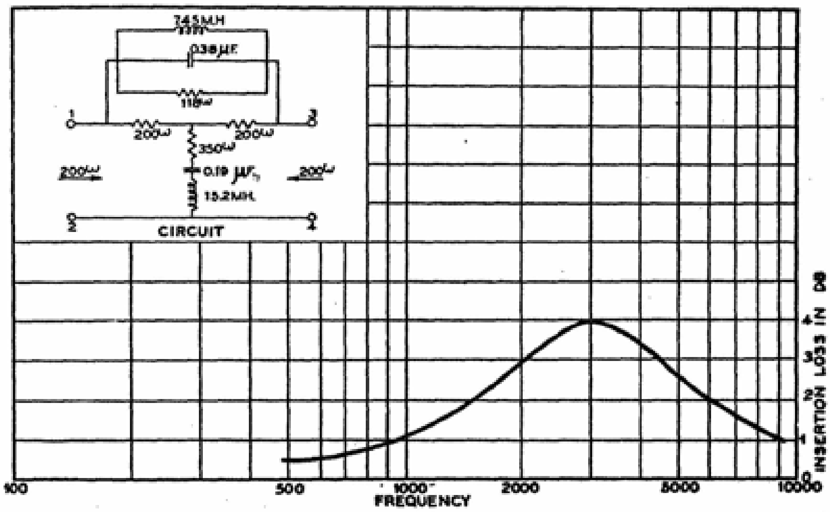


Figure 184 — Circuit and insertion loss characteristic of Example 9.

Example 10. Design a shunt impedance, Type VIII equalizer having an effective gain of 6 db at 1,000 c.p.s., half this gain at 700 c.p.s., and for use in a 200 ohm circuit.

From these data we have

$f_R = 1{,}000$ c.p.s. $\qquad b = 10/7$

$f_b = 700$ c.p.s. $\qquad 20 \log K = 6$

The following related data are needed for computing the electrical elements:

$$\frac{K-1}{\sqrt{K}}=\frac{1}{1.414}=0.707 \qquad \frac{b^2-1}{b^2}=0.51$$

$$\frac{\sqrt{K}}{K-1}=1.414 \qquad \frac{1}{b^2-1}=0.96$$

$$L_B=\frac{R_0}{2\,\pi\,f_b}=\frac{200}{2\,\pi\times 700}=0.0455 \text{ henries}$$

$$C_B=\frac{10^6}{2\,\pi\,f_bR_0}=\frac{1{,}000{,}000}{2\,\pi\times 700\times 200}=1.135\ \mu\text{f.}$$

From Column VIII of Chart I, we have for the electrical elements

$$R_2=\frac{R_0}{K-1}=\frac{200}{1}=200 \text{ ohms}$$

$$L_2=L_B\frac{\sqrt{K}}{K-1}\ \frac{b^2-1}{b^2}=45.5\times 1.414\times 0.51=32.8 \text{ m.h.}$$

$$C_2=C_B\frac{K-1}{\sqrt{K}}\ \frac{1}{b^2-1}=1.14\times 0.707\times 0.96=0.775\ \mu\text{f.}$$

Figure 185 shows the shunt impedance, Type VIII equalizer to correspond to these electrical elements. The insertion loss characteristic shown was obtained from Chart XIX for a pad loss of 6 db and for $f_R=1{,}000$ c.p.s. It is noted that when this equalizer is inserted into a circuit an effective gain of 6 db at 1,000 c.p.s. results.

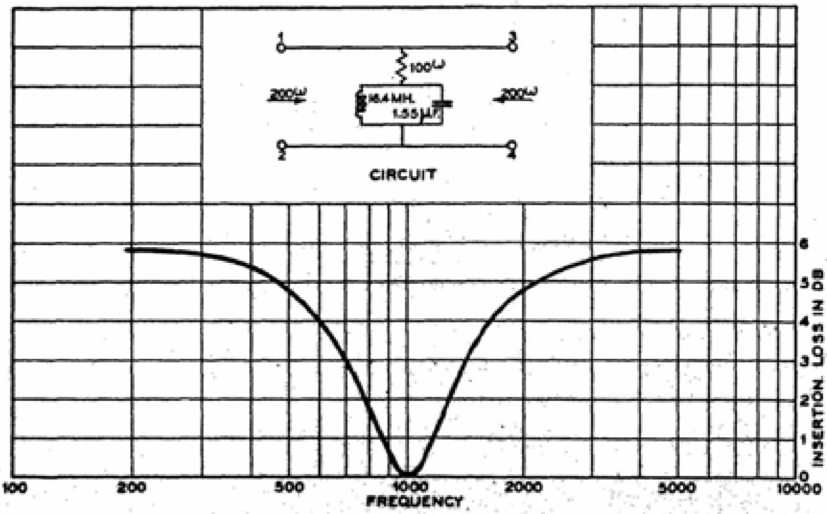


Figure 185 — Circuit and insertion loss characteristic of Example 10.

Chapter XVIII

WAVE FILTER THEORY

By HARRY KIMBALL

1. WAVE FILTERS

Electric wave filters, like attenuation equalizers, are four-terminal networks having a pair of input terminals and a pair of output terminals. Between the input and output terminals is an orderly array of electrical elements arranged so as to produce a specified insertion-loss characteristic when connected between the proper terminal impedances. Unlike equalizers, resistive elements are excluded from wave filters, that is, only inductive and capacitive elements are used to interconnect the input and output terminals. The reason for the exclusion of resistive conductors is contained in the purpose of a filter, that is, to transmit without appreciable loss all frequencies of the transmission band and to attenuate by a prescribed amount frequencies lying outside this band. If resistive elements were used in conventional type filters, attenuation loss would result within the transmission band.

Electric wave filters usually consist of a number of filter sections or unit four-terminal networks connected in tandem on a matched impedance basis to form the complete filter. It is not necessary that a filter consist of more than one section, but usually the transmission characteristic desired is such as to require the use of multiple sections. In this respect filters are different from equalizers, where in a great many cases, the desired equalization curves may be secured without tandem operation of sections.

Although wave filters transmit the frequencies of their transmission band without appreciable attenuation loss, they do shift the relative phases of all the frequencies. This is an inherent feature of filters that cannot be avoided, although in some cases means are available for controlling the phase shift characteristic so as to minimize its effect on transmission. In many transmission systems, the effect of phase shift is not of sufficient magnitude to require correction, while in certain types of systems corrective means must be employed. In the work which follows, phase shift within filters is considered only on a descriptive basis.

For design and nomenclature purposes, wave filters are classified into four types in accordance with the character of their attenuation characteristic. These types are:

1. Low-Pass Filters.
2. High-Pass Filters.
3. Band-Pass Filters.
4. Band-Elimination Filters.

For low-pass filters, the passing band includes the frequency range from zero frequency to some other finite frequency. For high-pass filters, the passing band covers the range from an infinite frequency down to some other frequency. A band-pass filter transmits a definite band of frequencies, and attenuates frequencies lying outside of the band. A band-elimination filter transmits all frequencies except a band of frequencies which are attenuated. Band-elimination filters, although forming an integral part of classified filter theory, are seldom used for the reason that there is very little commercial need for filters having their type of attenuation characteristic. In the following work, low-pass and high-pass filters are the primary subjects of discussion. The figure below illustrates the transmission and attenuation ranges for the different types of filters.

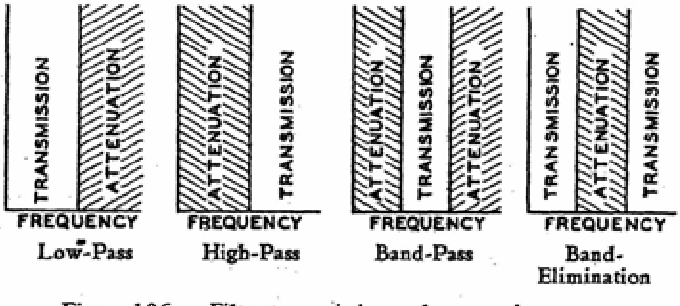


Figure 186 — Filter transmission and suppression ranges.

2. IMAGE IMPEDANCES — FOUR-TERMINAL NETWORKS

In wave filter work, the use of certain parameters known as image impedances is of great value. Consider, for the moment, the *L*-type network of Figure 187, having a series arm of 300 ohms resistance and a shunt arm of 500 ohms resistance. Let it be required to find two other resistances (image impedances) denoted by the symbols Z_I and Z'_I, which, when connected to the 1-2 and 3-4 terminals respectively, of the network, will simultaneously match impedances at the network

terminals. This match of impedances is obtained in this case when Z_I is 490 ohms and Z'_I is 306 ohms. To illustrate, the resistance looking

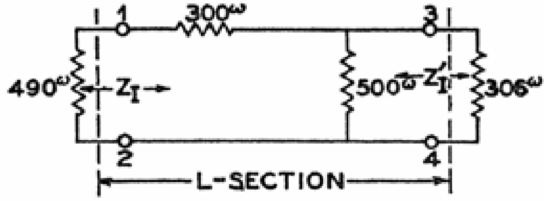


Figure 187 — *L*-type network terminated in image impedances.

into the network of Figure 187 at the 1-2 terminals with 306 ohms connected to the 3-4 terminals is

$$Z_{12} = 300 + \frac{500 \times 306}{806} = 490 \text{ ohms}$$

and that at the 3-4 terminals with 490 ohms connected to the 1-2 terminals is

$$Z_{34} = \frac{500\ (300 + 490)}{500 + 300 + 490} = 306 \text{ ohms}$$

For any network only one pair of impedances is capable of producing the desired impedance match conditions simultaneously at the input and output terminals. Thus, for Figure 187, 409 ohms and 306 ohms are the only pair of resistances which will simultaneously match impedances at the 1-2 and 3-4 terminals. Z_I and Z'_I are known as "image impedances" for the reason that the impedances looking in both directions from the terminals in question are "images" of each other. Also, Z_I is known as the mid-series image impedance because it is associated with the mid-series terminals of the network, and Z'_I is termed the mid-shunt image impedance since it relates to the mid-shunt terminals.

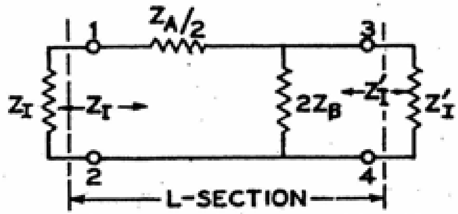


Figure 188 — Image impedances of *L* sections.

General formulae are not difficult to derive for expressing the image impedances of networks in terms of the network electrical constants. Figure 188 shows an *L* section having a series arm of impedance $Z_A/2$

and a shunt arm of impedance $2Z_B$. Image impedances of values Z_I and Z'_I are shown connected to the 1-2 and 3-4 terminals respectively. It is desired to express Z_I and Z'_I in terms of Z_A and Z_B. To do this, we have available the conditions that the impedance looking into the 1-2 terminals of the network with a load of Z'_I connected to the 3-4 terminals, must be equal to Z_I, and similarly, the network impedance at the 3-4 terminals must be equal to Z'_I with the load of Z_I connected to the 1-2 terminals.

We have at the 1-2 terminals

$$Z_I = \frac{Z_A}{2} + \frac{2Z_B Z'_I}{2Z_B + Z'_I}$$

and at the 3-4 terminals

$$Z'_I = \frac{2Z_B\left(\frac{Z_A}{2} + Z_I\right)}{2Z_B + \frac{Z_A}{2} + Z_I}$$

Simultaneous solution of these two equations for Z_I and Z'_I gives

$$Z_I = \sqrt{Z_A Z_B\left(1 + \frac{Z_A}{4Z_B}\right)} \quad \text{Mid-series image impedance.} \tag{120}$$

$$Z'_I = \sqrt{\frac{Z_A Z_B}{1 + \frac{Z_A}{4Z_B}}} \quad \text{Mid-shunt image impedance.} \tag{121}$$

These equations show that the image impedances of a network are determined by the connections and electrical constants of the network. As an application of the formulae, we have for Figure 187, $Z_A/2 = 300$ ohms and $2Z_B = 500$ ohms. Using these values of Z_A and Z_B with the above equations, we secure the image impedance values already mentioned.

$$Z_I = \sqrt{600 \times 250\left(1 + \frac{600}{1000}\right)} = 490 \text{ ohms}$$

$$Z'_I = \sqrt{\frac{600 \times 250}{1 + \frac{600}{1000}}} = 306 \text{ ohms}$$

Equations (120) and (121), as derived, have been shown to apply only to L-type sections. It may be shown, however, either mathematically or by other processes, that they apply also to symmetrical T and π sections. For instance, Figure 189 shows two L sections joined at their mid-shunt terminals to form a T section. The second L section of

Figure 189 is connected to a load impedance Z_I at its 1-2 or mid-series terminals. Then its impedance at the 3-4 terminals is Z'_I as discussed

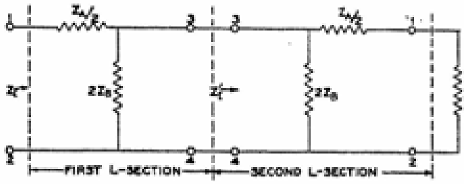
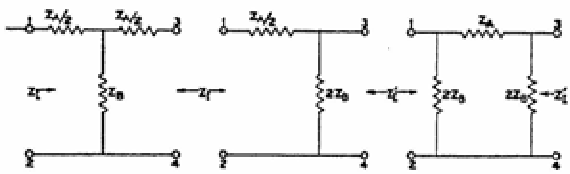


Figure 189 — Image impedances of T sections

above. This supplies a load of Z'_I for the first L network. For this load condition the impedance looking into the 1-2 terminals of the first L section is Z_I as already proven. Then the T section formed by the joining of the two L sections has equal image impedances of Z_I at both of its sets of terminals. In a similar manner it can be shown that when two L sections are joined at their mid-series terminals to form a π section, its image impedances are both alike and equal to Z'_I. Figure 190 summarizes these image impedance characteristics for L, T and π sections.



$$Z_I = \sqrt{Z_A Z_B \left(1 + \frac{Z_A}{4Z_B}\right)} \qquad Z'_I = \sqrt{\frac{Z_A Z_B}{1 + \frac{Z_A}{4Z_B}}}$$

Attenuation loss T and π sections

$$= 20 \log \frac{\sqrt{1 + \frac{Z_A}{4Z_B}} + \sqrt{\frac{Z_A}{4Z_B}}}{\sqrt{1 + \frac{Z_A}{4Z_B}} - \sqrt{\frac{Z_A}{4Z_B}}}$$

Attenuation loss L section $= \frac{1}{2}$ of T or π sections.

Figure 190 — Image impedance and attenuation losses of general L, T, and π filter sections.

In the design of filters by conventional methods it is customary to join L, T and π type sections in tandem in sufficient numbers and types to secure the insertion loss characteristic desired. It will be shown later that such sections should always be joined at terminals having like

image impedances in order to avoid internal reflection effects. Certain basic-type sections are available which permit the designing of filters in this manner with the maximum of flexibility. The general sections listed in Figure 191 are sufficient for practically all high-pass, low-pass and symmetrical band-pass filter designs, which include a considerable part of these filters.

Referring to Figure 191, the impedances Z_1 and Z_2 are those of any pair of reactive two-terminal networks. These impedances do not include resistive elements for the reason already explained—that filters in order to freely transmit throughout their transmission bands cannot dissipate energy internally. Z_1 and Z_2 are also defined as being inverse to each other with respect to R_0; that is $Z_1Z_2 = R_0^2$. The circuit configurations assigned to Z_1 and Z_2 determine the type of filters secured from the sections of Figure 191. For instance, if Z_1 is the impedance of an inductance coil, all of the sections become low-pass filter sections, whereas if Z_1 is the impedance of a condenser, high-pass filter sections are obtained. The symbol "m" of Figure 191 is a design parameter having any value between zero and unity. The image impedance and insertion loss characteristics of the filter sections are both controlled by the value assigned to m in a manner which will be discussed later.

The filter sections of Figure 191-A are called basic or constant K filter sections; those of Figure 191-B are called series m-derived sections, and those of Figure 191-C are termed shunt m-derived sections. The constant K terminology originates from the practice of some authors of using the letter "K" to indicate the constancy of the product Z_1 and Z_2; that is, $Z_1Z_2 = K$. For the m-derived sections it may be mentioned that by mathematical processes the sections of Figures 191-B and 191-C can be derived from the basic sections of 191-A, although this method is not used here.

Comparing the basic-type filter sections of Figure 191-A to the general sections of Figure 190, we have by comparison, $Z_1 = Z_A$ and $Z_2 = Z_B$. Then, from equations (120) and (121), the image impedance formulae for the basic sections are given by the equations

$$Z_I = \sqrt{(Z_1Z_2)\left(1+\frac{Z_1}{4Z_2}\right)} = R_0\sqrt{1+\frac{Z_1}{4Z_2}} \qquad (122)$$

$$Z'_I = \sqrt{\frac{Z_1Z_2}{1+\frac{Z_1}{4Z_2}}} = \frac{R_0}{\sqrt{1+\frac{Z_1}{4Z_2}}} \qquad (123)$$

Again, if the series m-derived sections of Figure 191-B are compared with the general sections of Figure 190, we secure

$$Z_A = m\,Z_1$$

$$Z_B = \frac{Z_2}{m} + \frac{1-m^2}{4m} Z_1 = \frac{Z_2}{m}\left[1 + (1-m^2)\frac{Z_1}{4Z_2}\right]$$

These may be rearranged to give the relations

$$Z_A Z_B = Z_1 Z_2\left[1 + (1-m^2)\frac{Z_1}{4Z_2}\right]$$

$$= R_0^2\left[1 + (1-m^2)\frac{Z_1}{4Z_2}\right] \qquad (124)$$

$$\frac{Z_A}{4Z_B} = \frac{m^2 \dfrac{Z_1}{4Z_2}}{1 + (1-m^2)\dfrac{Z_1}{4Z_2}} \qquad (125)$$

$$1 + \frac{Z_A}{4Z_B} = \frac{1 + \dfrac{Z_1}{4Z_2}}{1 + (1-m^2)\dfrac{Z_1}{4Z_2}} \qquad (126)$$

Use of these relations in connection with equations (120) and (121) gives for the image impedances of the series m-derived sections:

$$\begin{array}{l}\text{Mid-series}\\ \text{image impedance}\\ \text{of series}\\ m\text{-derived sections.}\end{array} = R_0\sqrt{1 + \frac{Z_1}{4Z_2}} = Z_I = \begin{array}{l}\text{Same as the}\\ \text{corresponding}\\ \text{image impedance}\\ \text{of basic sections.}\end{array} \qquad (127)$$

$$\left.\begin{array}{ll}\text{Mid-shunt image} & = R_0 \dfrac{\left[1 + (1-m^2)\dfrac{Z_1}{4Z_2}\right]}{\sqrt{1 + \dfrac{Z_1}{4Z_2}}} \\ \text{impedance of series} & \\ m\text{-derived sections.} & = Z'_I\left[1 + (1-m^2)\dfrac{Z_1}{4Z_2}\right]\end{array}\right\} \begin{array}{l}\text{Denoted}\\ \text{hereinafter}\\ \text{by the}\\ \text{Symbol } Z'_{Im}\end{array} \qquad (128)$$

Once more, when the shunt m-derived sections of Figure 191-C are compared with general circuits of Figure 190 we have,

$$Z_A = \frac{\dfrac{4m^2}{1-m^2} Z_1 Z_2}{mZ_1 + \dfrac{4m}{1-m^2} Z_2} = \frac{mZ_1}{1 + (1-m^2)\dfrac{Z_1}{4Z_2}}$$

$$Z_B = \frac{Z_2}{m}$$

By obvious transformations we secure the relations

$$Z_A Z_B = \frac{Z_1 Z_2}{1 + (1 - m^2)\dfrac{Z_1}{4Z_2}} = \frac{R_0^2}{1 + (1 - m^2)\dfrac{Z_1}{4Z_2}} \qquad (129)$$

$$\frac{Z_A}{4Z_B} = \frac{m^2 \dfrac{Z_1}{4Z_2}}{1 + (1 - m^2)\dfrac{Z_1}{4Z_2}} \qquad (130)$$

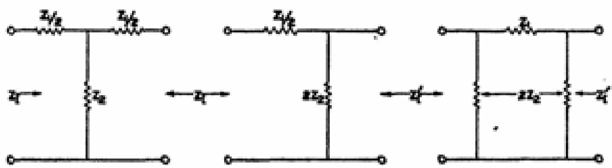
$$1 + \frac{Z_A}{4Z_B} = \frac{1 + \dfrac{Z_1}{4Z_2}}{1 + (1 - m^2)\dfrac{Z_1}{4Z_2}} \qquad (131)$$

Use of these relations with equations (120) and (121) gives for the image impedance formulae of the shunt m-derived sections

$$\left.\begin{aligned} \text{Mid-series image impedance of shunt } m\text{-derived sections.} &= \frac{R_0\sqrt{1 + \dfrac{Z_1}{4Z_2}}}{1 + (1 - m^2)\dfrac{Z_1}{4Z_2}} \\ &= \frac{Z_I}{1 + (1 - m^2)\dfrac{Z_1}{4Z_2}} \end{aligned}\right\} \text{Denoted hereinafter by the Symbol } Z_{Im} \qquad (132)$$

$$\text{Mid-shunt image impedance of shunt } m\text{-derived sections.} = \frac{R_0}{\sqrt{1 + \dfrac{Z_1}{4Z_2}}} = Z'_I = \text{Same as the corresponding image impedance of basic sections.} \qquad (133)$$

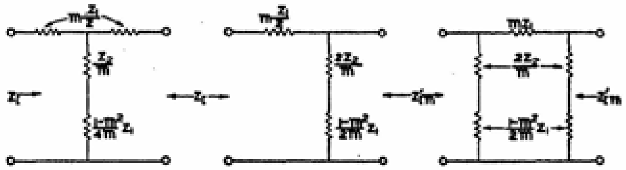
For the series m-derived sections of Figure 191-B, it is observed that the mid-series image impedance is not dependent upon the value of m but is identical with the corresponding impedance of the basic sections. This feature enables the designer to join in tandem the basic and series-derived type sections at their mid-series terminals without incurring internal reflection losses. Also the mid-shunt image impedance of the shunt m-derived sections is independent of m, and is identical with the corresponding image impedance of the basic sections, permitting the joining in tandem of these sections at their mid-shunt terminals. Series-derived and shunt-derived sections cannot be connected in tandem on a matched image impedance basis, either at their mid-series or mid-shunt terminals. Figure 192 shows a number of illustrative combinations of



$$Z_I = R_0\sqrt{1+\frac{Z_1}{4Z_2}} \qquad Z'_I = \frac{R_0}{\sqrt{1+\frac{Z_1}{4Z_2}}} \qquad Z_1 Z_2 = R_0^2$$

$$\text{Attenuation loss for basic sections} = 20\log\frac{\sqrt{1+\frac{Z_1}{4Z_2}}+\sqrt{\frac{Z_1}{4Z_2}}}{\sqrt{1+\frac{Z_1}{4Z_2}}-\sqrt{\frac{Z_1}{4Z_2}}}$$

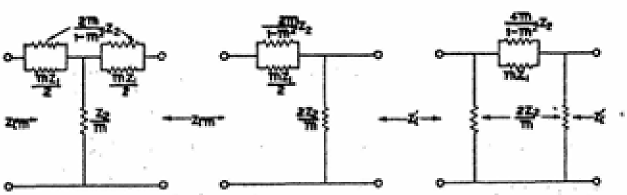
Figure 191-A — Basic filter sections.



$$Z'_{Im} = Z'_I\left[1+(1-m^2)\frac{Z_1}{4Z_2}\right]$$

$$\text{Attenuation loss for } m\text{-derived sections} = 20\log\frac{\sqrt{1+\frac{Z_1}{4Z_2}}+m\sqrt{\frac{Z_1}{4Z_2}}}{\sqrt{1+\frac{Z_1}{4Z_2}}-m\sqrt{\frac{Z_1}{4Z_2}}}$$

Figure 191-B — Series *m*-derived filter sections.



$$Z_{Im} = \frac{Z_I}{1+(1-m^2)\frac{Z_1}{4Z_2}}$$

Attenuation loss = attenuation loss of Figure 191-B.

Figure 191-C — Shunt *m*-derived filter sections.

Figure 191 — Some fundamental filter sections.

the sections of Figure 191 connected in tandem on a matched image impedance basis.

Each of the image impedance equations given above contains the term $\frac{Z_1}{4Z_2}$ which varies with frequency in a manner determined by Z_1 and Z_2. Hence, the image impedances of each of the sections of Figure 191 vary in some manner with frequency. This does not affect the impedance match obtained between the sections of a filter since the image impedances of any two joining sections are balanced against each other. However, at the input and output terminals the filter image impedances are balanced against the connecting system impedances, which are usually constant resistances, with the result that it is never possible to match impedances at the terminals of a filter. Means are available, however, usually at a certain amount of additional expense, to reduce the terminal mis-match as much as is needed. This point will be considered later for low- and high-pass filters from an image impedance standpoint. A filter, then, may be considered a number of filter sections in tandem, having ideal impedance match conditions between the sections, but a certain amount of mis-match at the input and output terminals.

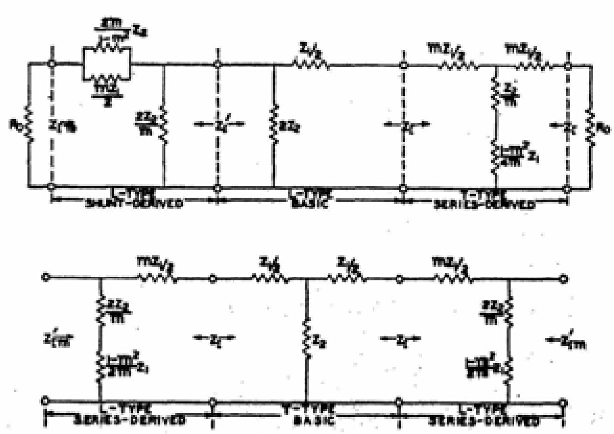


Figure 192 — Tandem operation of filter sections.

3. ATTENUATION LOSSES

When a filter section is working between connected loads which match its image impedances its insertion loss is given the special name of attenuation loss. The attenuation loss of a number of filter sections joined in tandem on a matched image impedance basis is simply the sum of the attenuation losses of the individual sections. Although, in practical work, it is never possible to arrange a transmission system to pro-

vide image impedance connecting loads at the filter terminals, the attenuation losses of filters and filter sections are important factors in determining the complete filter insertion loss.

As a basis for arriving at the attenuation losses of the filter sections of Figure 191, consider the L section of Figure 188 working between a sending-end impedance Z_I and a load impedance Z'_I as defined in equations (120) and (121). This provides a match of impedances at the input and output terminals of the section. For this condition, a current I'_L is delivered to the load by a voltage E, inserted in series with the sending-end impedance Z_I. The maximum current I_L which the voltage E can deliver to the load without the use of the filter, but with an ideal matching transformer used instead, as shown in Figure 165-A, is

$$I_L = \frac{E}{2\sqrt{Z_I Z'_I}}$$

We have for I'_L

$$I'_L = \frac{E}{2Z_I}\,\frac{2Z_B}{Z'_I + 2Z_B}$$

The current ratio becomes

$$\frac{I_L}{I'_L} = \sqrt{\frac{Z_I}{Z'_I}}\left(\frac{2Z_B + Z'_I}{2Z_B}\right) = \sqrt{\frac{Z_I}{Z'_I}} + \sqrt{\frac{Z_I Z'_I}{4Z_B^2}}$$

From equations (121) and (120) we have the relations:

$$\sqrt{\frac{Z_I}{Z'_I}} = \sqrt{1 + \frac{Z_A}{4Z_B}} \text{ and } \sqrt{Z_I Z'_I} = \sqrt{Z_A Z_B}$$

From this, the current ratio of the L section is

$$\frac{I_L}{I'_L} = \sqrt{\frac{Z_A}{4Z_B}} + \sqrt{1 + \frac{Z_A}{4Z_B}} \qquad (134)$$

and the attenuation loss is

$$\text{Attenuation Loss } L \text{ Section} = 20 \log\left[\sqrt{\frac{Z_A}{4Z_B}} + \sqrt{1 + \frac{Z_A}{4Z_B}}\right]$$

The attenuation loss for two L sections in tandem, that is, for a T section or a π section, is twice this attenuation loss.

$$\begin{matrix} T \text{ and } \pi \text{ Sections} \\ \text{Attenuation Loss} \end{matrix} = 20 \log\left[\sqrt{\frac{Z_A}{4Z_B}} + \sqrt{1 + \frac{Z_A}{4Z_B}}\right]^2 \qquad (135)$$

In filter work, attenuation data in graphical form are usually compiled for the T and π sections, and the attenuation loss of L sections is taken as one-half of such data. For this reason, equation (135) is more commonly used than (134). Equation (135) may be expressed in other forms by use of the relation

$$\left[\sqrt{1+\frac{Z_A}{4Z_B}}+\sqrt{\frac{Z_A}{4Z_B}}\right]\left[\sqrt{1+\frac{Z_A}{4Z_B}}-\sqrt{\frac{Z_A}{4Z_B}}\right]=1$$

We have

Attenuation loss T and π sections.

$$= 20\log\frac{I_L}{I'_L} = 20\log\left[\frac{\sqrt{1+\frac{Z_A}{4Z_B}}+\sqrt{\frac{Z_A}{4Z_B}}}{\sqrt{1+\frac{Z_A}{4Z_B}}-\sqrt{\frac{Z_A}{4Z_B}}}\right] \qquad (136)$$

$$= 20\log\left[\frac{1}{\sqrt{1+\frac{Z_A}{4Z_B}}-\sqrt{\frac{Z_A}{4Z_B}}}\right]^2 \qquad (137)$$

These formulae are noted in Figure 190 for reference purposes.

The attenuation loss formulae for the general wave filter sections of Figure 191 can be developed with the aid of the material shown in Figure 190. Comparing the basic filter sections of Figure 191-A with the general sections of Figure 190, we have

$$Z_A = Z_1 \text{ and } Z_B = Z_2$$

and the attenuation loss for full T and π basic sections becomes

Attenuation loss for T and π sections of Figure 191-A.

$$= 20\log\left[\frac{\sqrt{1+\frac{Z_1}{4Z_2}}+\sqrt{\frac{Z_1}{4Z_2}}}{\sqrt{1+\frac{Z_1}{4Z_2}}-\sqrt{\frac{Z_1}{4Z_2}}}\right] \qquad (138)$$

The attenuation losses of the series m-derived T and π sections of Figure 191-B are identical with those of the shunt m-derived sections of Figure 191-C. We have, by comparing Figures 191-B and C with Figure 190, and by use of relations expressed in equations (125), (126), (130) and (131):

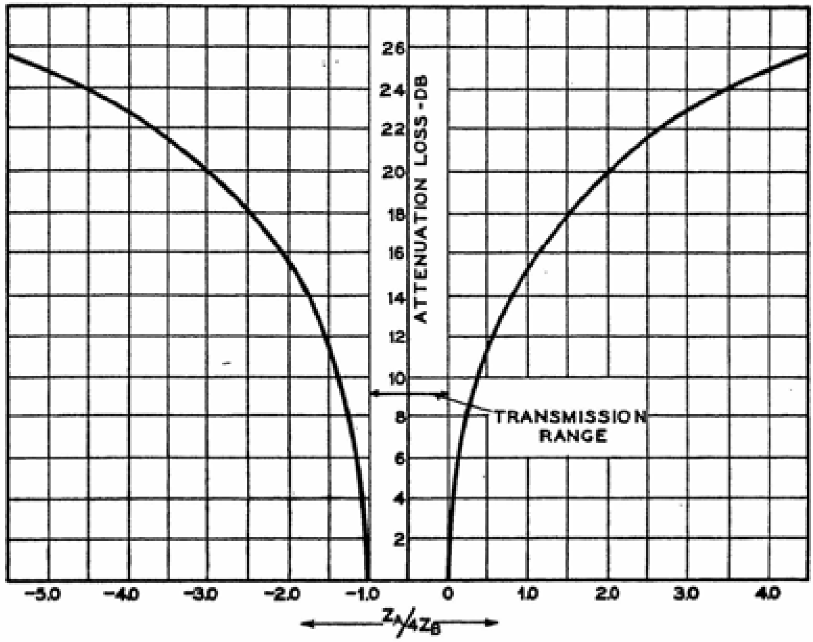
Attenuation loss of T and π, series and shunt m-derived sections.

$$= 20\log\frac{\sqrt{1+\frac{Z_1}{4Z_2}}+m\sqrt{\frac{Z_1}{4Z_2}}}{\sqrt{1+\frac{Z_1}{4Z_2}}-m\sqrt{\frac{Z_1}{4Z_2}}} \qquad (139)$$

4. TRANSMISSION AND ATTENUATION RANGES

Equation (136) expresses the attenuation loss of the general T and π sections of Figure 190 in terms of the impedance ratio $\left(\frac{Z_A}{4Z_B}\right)$ where for filters, both Z_A and Z_B are pure reactances. The ratio of any two react

ances is a real number either positive or negative in value. That is, the ratio $\left(\frac{Z_A}{4Z_B}\right)$ varies between minus infinity and plus infinity for Z_A and Z_B having any reactive values. Figure 193 shows a graph of attenuation loss



$$\text{Attenuation loss} = 20 \log \frac{\sqrt{1+\frac{Z_A}{4Z_B}}+\sqrt{\frac{Z_A}{4Z_B}}}{\sqrt{1+\frac{Z_A}{4Z_B}}-\sqrt{\frac{Z_A}{4Z_B}}}$$

where Z_A and Z_B = Pure reactances.

Figure 193 — Attenuation loss curve when Z_A and Z_B are pure reactances.

as computed from equation (136) for $\left(\frac{Z_A}{4Z_B}\right)$ covering the ratio range from -5.5 to $+4.5$. It is seen that varying amounts of loss are secured for all values of $\left(\frac{Z_A}{4Z_B}\right)$ except when this ratio has values between zero and minus one. In other words, for $\left(\frac{Z_A}{4Z_B}\right)$ having a value between zero and minus one, the filters freely transmit, whereas for other values attenua-

tion is secured. The critical conditions separating the transmission from the suppression ranges are,

$$\frac{Z_A}{4Z_B} = 0$$

and

$$\frac{Z_A}{4Z_B} = -1 \tag{140}$$

These are general equations applying to any filter sections having full-series and full-shunt reactive impedances of Z_A and Z_B. For the basic sections of Figure 191-A we have, $Z_A = Z_1$; $Z_B = Z_2$; and

$$\frac{Z_A}{4Z_B} = \frac{Z_1}{4Z_2}$$

Then the critical conditions defining the band limits for these sections are

$$\frac{Z_1}{4Z_2} = 0$$

and

$$\frac{Z_1}{4Z_2} = -1 \tag{141}$$

For the series and shunt m-derived sections of Figures 191-B and 191-C we have from equations (125) or (130)

$$\frac{Z_A}{4Z_B} = \frac{m^2 \dfrac{Z_1}{4Z_2}}{1 + (1 - m^2)\dfrac{Z_1}{4Z_2}} \tag{142}$$

Equating this value to zero gives for one band-limiting condition

$$\frac{Z_1}{4Z_2} = 0 \tag{143}$$

and equating $\left(\frac{Z_A}{4Z_B}\right)$ equal to minus one gives for the other band-limiting condition:

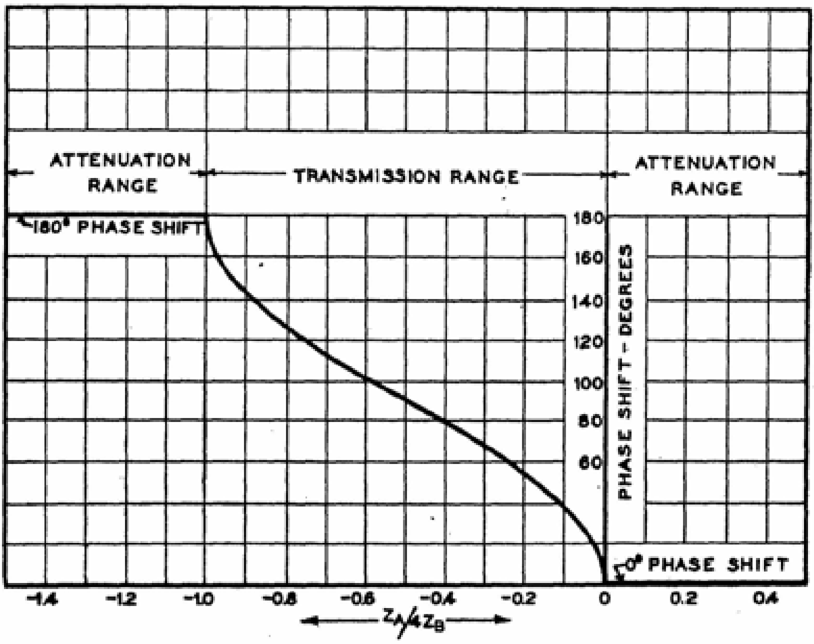
$$\frac{m^2 Z_1}{4Z_2} = -1\left[1 + (1 - m^2)\frac{Z_1}{4Z_2}\right]$$

which gives

$$\frac{Z_1}{4Z_2} = -1 \tag{144}$$

Since equations (143) and (144) defining the band-limiting conditions for m-derived sections are the same as equation (141) for basic sections, it is demonstrated that all of the sections of Figure 191 have the same band-limits for the same values of Z_1 and Z_2.

Although the current ratio of $\left(\frac{I_L}{I'_L}\right)$ for the basic and derived sections of Figure 191 remains constant in magnitude at a value of unity throughout the transmission range, thus resulting in zero attenuation loss in this range, a considerable shift in phase between the two currents I_L and I'_L is experienced. The amount of this phase shift for any value of $\left(\frac{Z_A}{4Z_B}\right)$ may be obtained by evaluating equation (136) for arbitrarily chosen values of $\left(\frac{Z_A}{4Z_B}\right)$. Figure 194 is a graph of the phase shift between the unattenuated load current I_L and the attenuated load current



$$\text{Formula: } \frac{I_L}{I'_L} = \frac{\sqrt{1+\frac{Z_A}{4Z_B}}+\sqrt{\frac{Z_A}{4Z_B}}}{\sqrt{1+\frac{Z_A}{4Z_B}}-\sqrt{\frac{Z_A}{4Z_B}}}$$

Figure 194 — Phase characteristic of T and π sections computed from formula.

I'_L. It is noted that for all positive values of $\left(\frac{Z_A}{4Z_B}\right)$ the phase shift is zero: that is, I_L and I'_L are in-phase. Again, for values of $\left(\frac{Z_A}{4Z_B}\right)$ which are more negative than minus one, the phase shift is 180 degrees. In

other words, in the attenuation ranges the phase shift is constant, being zero degrees for $\left(\frac{Z_A}{4Z_B}\right)$ greater than zero, and 180 degrees for $\left(\frac{Z_A}{4Z_B}\right)$ less than minus one. In the transmission range the phase angle changes from zero degrees for $\left(\frac{Z_A}{4Z_B}\right)$ equal to 0, to 180 degrees for $\left(\frac{Z_A}{4Z_B}\right)$ equal to minus one.

5. INSERTION LOSSES

The previous work has provided formulae for estimating the insertion losses of filters working between connected loads which supply a match of impedances at the input and output terminals, and for this condition the insertion loss is given the special term of attenuation loss. When filters operate between other impedances, as they ordinarily do in practice, the complete insertion loss is secured by adding correction losses to the attenuation losses to take account of the mis-match of impedances at the terminals. The magnitude of the correction losses, of course, decreases with improved impedance match conditions. In dealing with these correction losses they are usually separated into two parts; that is, Terminal Losses, and Interaction Losses. From a wave theory standpoint this division is in harmony with the events which take place within the filter.

To develop the complete insertion-loss formula of a filter working between equal sending- and receiving-end impedances of R_0 ohms, consider the circuit of Figure 195. Here the filter shown in block schematic

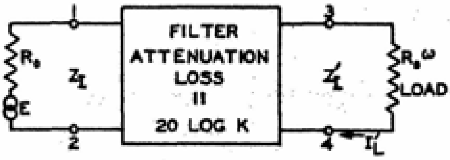


Figure 195 — Insertion losses.

form is assumed to have an image impedance of Z_I at one end and Z'_I at the other end. The attenuation loss, when operating between image impedances, is 20 log K where K is the vector current ratio for this condition. The image impedances and K are assumed to be known quantities arrived at by the methods previously given. Referring to Figure 195 the current I_L delivered to the load without the use of the equalizer is

$$I_L = \frac{E}{2R_0}$$

It can be shown* that the load current I'_L delivered to R_0 through the filter is

$$I'_L = \frac{2\,E\,\sqrt{Z_I Z'_I}}{K\,(Z_I + R_0)\;(Z'_I + R_0)}\;\frac{1}{1 - \left(\dfrac{Z'_I - R_0}{Z'_I + R_0}\,\dfrac{Z_I - R_0}{Z_I + R_0}\,\dfrac{1}{K^2}\right)}$$

Then the insertion loss of the filter can be arranged as follows:

$$\text{I.L.} = 20 \log \frac{I_L}{I'_L}$$

$$\text{I.L.} = 20 \log K \qquad \text{(Attenuation loss)}$$

$$+\, 20 \log \frac{R_0 + Z_I}{2\sqrt{R_0 Z_I}} \qquad \text{(Terminal loss at one end)}$$

$$+\, 20 \log \frac{R_0 + Z'_I}{2\sqrt{R_0 Z'_I}} \qquad \text{(Terminal loss at other end)}$$

$$+\, 20 \log \left[1 - \left(\frac{1}{K^2}\,\frac{Z_I - R_0}{Z_I + R_0}\,\frac{Z'_I - R_0}{Z'_I + R_0}\right)\right] \qquad \text{(Interaction loss)}$$

For this general insertion loss formula the first term is the attenuation loss for the filter. The second and third terms are the terminal losses, and the fourth term is known as the interaction loss. It is noted that the terminal and interaction losses are zero when $Z_I = Z'_I = R_0$. In other words, where a match of impedances is obtained at both ends of the filter, the filter insertion loss is the same as its attenuation loss. It is only for mis-match conditions at the terminals that terminal and interaction losses need to be used as correction factors.

The two terminal losses applying to a filter are in no way related to each other as indicated by the above formulae. For either end of a filter, the terminal losses are determined by the match of impedances obtained at that end. In connection with the discussion given later of low-pass and high-pass filters, it will be shown that the frequency characteristics of the image impedances for these filter sections is most likely to depart from a constant resistance of R_0 ohms in the neighborhood of the cut-off frequencies of the sections. In other words, for conventionally designed filters the terminal losses are usually very small (close to zero) for frequencies of the transmission band remote from the cut-off frequency, but become larger in the vicinity of the cut-off frequency. In the attenuation range the terminal losses are usually appreciable, but, except in precision designs, can be neglected where a safe margin of attenuation is provided by the filter attenuation losses.

---

* K. S. Johnson, *"Transmission Circuits for Telephonic Communication,"* Chapter XI.

The interaction loss of a filter, as given by the above formula, is physically due to multiple reflections of energy back and forth through the filter structure. Two factors of the filter design jointly determine the magnitude of the reflected components: One, the impedance match at each of the filter terminals; and two, the filter attenuation loss as expressed by K. The first of these factors determines the magnitude of the reflected components at the filter terminals, whereas the second factor acts to attenuate the components as they traverse from end to end of the filter. For example, where a match of impedances is obtained at either of the pair of filter terminals, the interaction loss is zero as borne out by the above formula. Again, in the attenuation range of a filter, the interaction loss is usually negligible because the reflected components are suppressed in traversing the filter; that is, K is large. Interaction losses, then, like terminal losses, are of greatest importance in the cut-off range of filters because in this range the terminal mis-match of impedance is greatest and the filter attenuation loss is not sufficiently large to suppress the reflected components.

Chapter XIX

LOW-PASS AND HIGH-PASS FILTERS

By HARRY KIMBALL

The previous chapter has made available the general network sections of Figure 191, and has shown the methods used for connecting the sections in tandem to form composite structures. All of the sections are expressed in terms of the pair of impedances, Z_1 and Z_2, which for filter work have thus far been defined as relating to any pair of inverse two-terminal reactive networks. It has been shown that the attenuation characteristics for the sections depend upon the impedance ratio $\left(\frac{Z_1}{4Z_2}\right)$, and that a range of free transmission is obtained for all values of this ratio lying between zero and minus one. It may further be said that because of the manner in which Z_1 and Z_2 are defined, it is impossible to select a pair of inverse networks for which a free transmission range is not secured, when used with the filter sections of Figure 191. The actual placement of the transmission band for these sections, with respect to the frequency spectrum is, of course, determined by the circuit configurations assigned to the Z_1 and Z_2 arms. For instance, it is shown below that, when the Z_1 arm is an inductance and the Z_2 arm is a capacitance, all of the sections of Figure 191 are low-pass filters. Again, if the Z_1 arm is an inductance and capacitance in series, and the Z_2 arm is the corresponding inverse circuit, band-pass filter sections are obtained. We have, then, but to assign to Z_1 and Z_2, different circuit configurations to secure different types of filter sections. In the following work this is done for low-pass and high-pass filters, and formulae and data relating to these filter sections are worked out from the general data already presented.

LOW-PASS FILTERS

1. LOW-PASS FILTER SECTIONS

A comprehensive group of simple and very practical low-pass filter sections is obtained from the general sections of Figure 191 when the Z_1 arm is made an inductance and the Z_2 arm the corresponding inverse capacitance. Denoting the inductance by the symbol L_0 and the capacitance by C_0 we have,

$$Z_1 = j\,2\pi fL_0 = j\,\omega\,L_0$$

$$Z_2 = -j\frac{1}{2\pi f\,C_0} = -j\frac{1}{\omega\,C_0} \tag{145}$$

Using these values of Z_1 and Z_2 with the filter sections of Figure 191, we obtain the low-pass filter sections of Figure 196.

2. TRANSMISSION AND ATTENUATION BANDS

For the general sections of Figure 191, it has been shown that the band limiting conditions are given by the equations

$$\frac{Z_1}{4Z_2} = (-1) \text{ and } \frac{Z_1}{4Z_2} = 0$$

We have by means of (145)

$$\frac{Z_1}{4Z_2} = -\pi^2 f^2 L_0 C_0 \tag{146}$$

Making use of the first condition gives

$$-\pi^2 f^2 L_0 C_0 = -1$$

Let f_c denote the value of f for which this equation is valid, that is

$$f_c = \frac{1}{\pi\sqrt{L_0 C_0}} \tag{147}$$

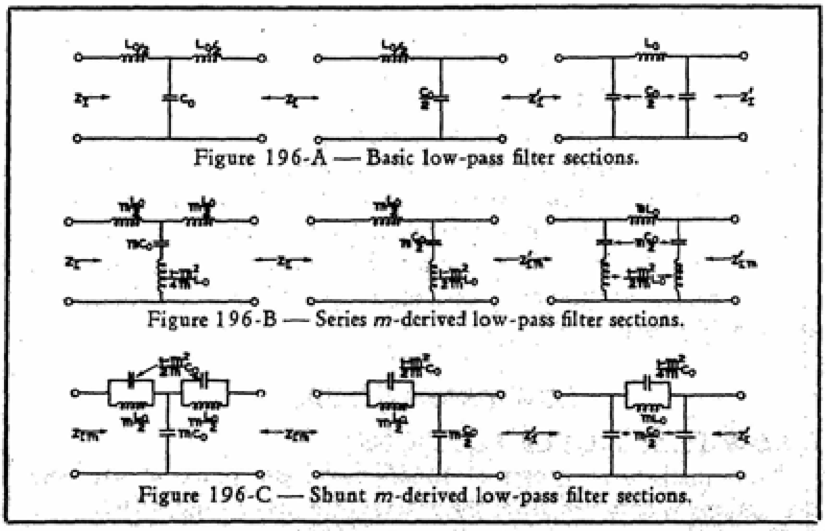


Figure 196-A — Basic low-pass filter sections.

Figure 196-B — Series m-derived low-pass filter sections.

Figure 196-C — Shunt m-derived low-pass filter sections.

Figure 196 — Low-pass filter sections.

This frequency, f_c, separates the transmission range from the suppression range, and is known as the cut-off frequency of these low-pass filter

sections. Another cut-off frequency of $f = 0$ may be obtained by using the other band limiting condition, namely: $\frac{Z_1}{4Z_2} = 0$, but this frequency is not usually referred to as a cut-off point. The low-pass filter sections of Figure 196 then freely transmit from a frequency of zero cycles per second up to f_c cycles per second, and attenuate for frequencies higher than f_c.

3. ELECTRICAL CONSTANTS

With the values of R_0 and f_c available as known design data, the constants L_0 and C_0 can be determined from equation (147), aided by the additional information that L_0 and C_0 are inverse to each other: that is, $L_0/C_0 = R_0^2$. Solving these equations simultaneously, gives for the values of L_0 and C_0 in terms of R_0 and f_c

$$L_0 = \frac{R_0}{\pi f_c}$$

$$C_0 = \frac{1}{\pi f_c R_0} \qquad (148)$$

4. IMAGE IMPEDANCES

We may now write

$$Z_1 = j\,2\,R_0\frac{f}{f_c}$$

$$Z_2 = -j\frac{R_0}{2}\frac{f_c}{f}$$

$$\frac{Z_1}{4Z_2} = -\left(\frac{f}{f_c}\right)^2 \qquad (149)$$

$$1 + \frac{Z_1}{4Z_2} = 1 - \left(\frac{f}{f_c}\right)^2 \qquad (150)$$

The image impedances Z_I, Z'_I, Z_{Im} and Z'_{Im}, as given in equations (122), (123), (132) and (128) respectively, may now be expressed in terms of R_0 and f_c as follows:

$$Z_I = R_0\sqrt{1-\left(\frac{f}{f_c}\right)^2} = \text{Mid-series image impedance of basic type and series } m\text{-derived low-pass filter sections.} \qquad (151)$$

$$Z'_I = R_0\frac{1}{\sqrt{1-\left(\frac{f}{f_c}\right)^2}} = \text{Mid-shunt image impedances of basic type and shunt } m\text{-derived low-pass filter sections.} \qquad (152)$$

$$Z_{Im} = \frac{R_0\sqrt{1-\left(\frac{f}{f_c}\right)^2}}{1-(1-m^2)\left(\frac{f}{f_c}\right)^2} = \text{Mid-series image impedance of shunt } m\text{-derived low-pass filter sections.} \quad (153)$$

$$Z'_{Im} = R_0 \frac{\left[1-(1-m^2)\left(\frac{f}{f_c}\right)^2\right]}{\sqrt{1-\left(\frac{f}{f_c}\right)^2}} = \text{Mid-shunt image impedance of series } m\text{-derived low-pass filter sections.} \quad (154)$$

It is noted that the image impedances for the derived type networks, as expressed by formulae (153) and (154), break down or reduce to the corresponding formulae for the basic type networks when "m" has a value of one. For instance, if the value $m = 1$ is used in connection with equation (153) the formula reduces to that of equation (151), and similarly, formula (154) reduces to formula (152). In other words, the basic networks may be considered as special cases of the derived networks.

Since the above formulae are expressed in terms of the frequency ratio $\left(\frac{f}{f_c}\right)$ it is convenient to make graphs of these image impedance functions for various values of $\left(\frac{f}{f_c}\right)$ plotted along the horizontal axis. For the vertical axis the ratios $\left(\frac{Z_{Im}}{R_0}\right)$ or $\left(\frac{R_0}{Z'_{Im}}\right)$ are used in order to make the graphs more useful. This method of plotting may be seen more clearly when equations (153) and (154) are expressed as follows:

$$\frac{Z_{Im}}{R_0} = \frac{R_0}{Z'_{Im}} = \frac{\sqrt{1-\left(\frac{f}{f_c}\right)^2}}{1-(1-m^2)\left(\frac{f}{f_c}\right)^2} \quad (155)$$

Figure 197 shows a graph of equation (155) for various values of m, and for values of $\left(\frac{f}{f_c}\right)$ within the transmission band of the above low-pass filter sections; that is, when $\left(\frac{f}{f_c}\right)$ varies over the range from zero to one. It is noted that the image impedance characteristics are resistive throughout the transmission range. The curve for $m = 1$ is

the image impedance characteristic for the basic-type sections as explained above. One item worth special attention is the form of the different image impedance curves for certain values of "m". For instance, when $m = 0.6$, the image impedance curve is approximately

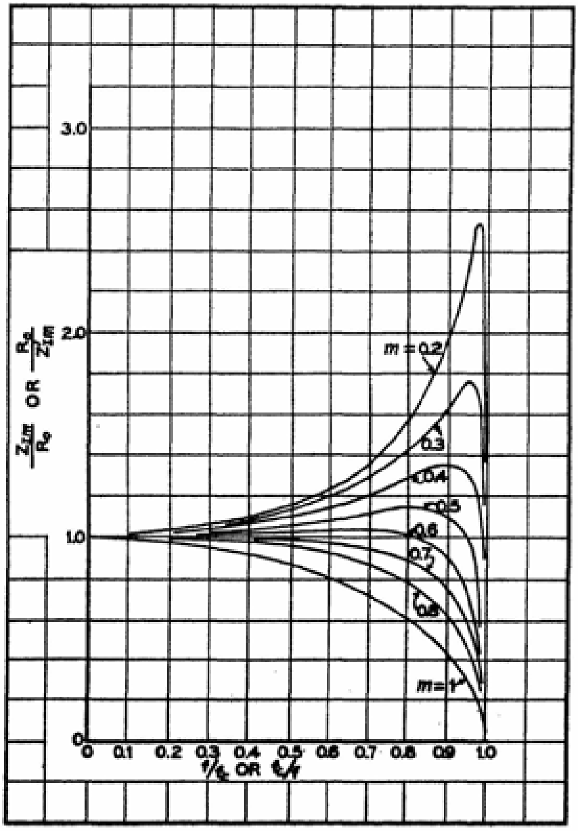


Figure 197 — Image impedance characteristics for low- and high-pass filter sections within the transmission band for various values of m.

constant up to about eighty-five per cent of the cut-off frequency. This means that derived sections employing a value of $m = 0.6$ have image impedance characteristics Z_{Im} and Z'_{Im} which are approximately constant at R_0 ohms over the greater part of the transmission band. Consequently, such sections are often used at the ends of filters to reduce terminal and interaction losses throughout the transmission band.

5. ATTENUATION CHARACTERISTICS

Attenuation characteristics for the low-pass filter sections of Figure 196 may be obtained from the general formula of equation (139) as

applied to these sections. Using the value of $\left(\frac{Z_1}{4Z_2}\right)$ from equation (149) with equation (139) gives

$$\text{Attenuation loss for low-pass } T \text{ and } \pi \text{ sections} = 20 \log \frac{\sqrt{1-\left(\frac{f}{f_c}\right)^2} + m\sqrt{-\left(\frac{f}{f_c}\right)^2}}{\sqrt{1-\left(\frac{f}{f_c}\right)^2} - m\sqrt{-\left(\frac{f}{f_c}\right)^2}} \tag{156}$$

In the attenuation range of low-pass filters where f is greater than f_c, the term $\left[1-\left(\frac{f}{f_c}\right)^2\right]$ is always negative. For graphing purposes it is desirable to express equation (156) in a manner to eliminate this negative condition.

$$\text{Attenuation loss for low-pass } T \text{ and } \pi \text{ sections} = 20 \log \frac{\sqrt{\left(\frac{f}{f_c}\right)^2 - 1} + m\frac{f}{f_c}}{\sqrt{\left(\frac{f}{f_c}\right)^2 - 1} - m\frac{f}{f_c}} \tag{157}$$

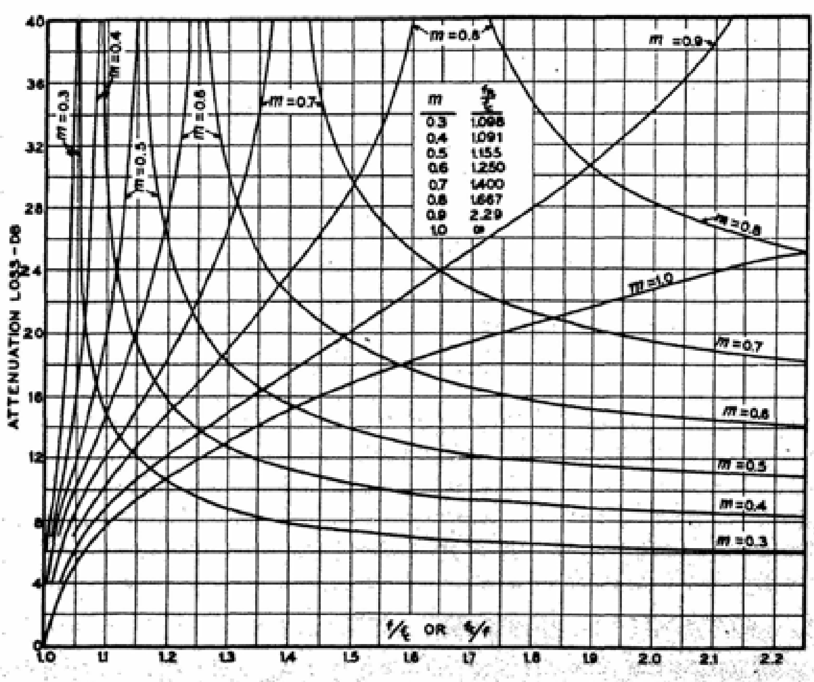


Figure 198 — Attenuation losses for low- and high-pass filters.

Figure 198 shows attenuation loss curves computed from this formula for a number of arbitrarily chosen values of m and plotted against the ratio $\left(\frac{f}{f_c}\right)$. The characteristics apply to all of the T and π sections of Figure 196 and may be used for L sections except that in this case, the insertion loss is one-half the curve values. It is noted that equation (139) reduces to (138) when m has the special value of $m = 1$. Then in Figure 198, the characteristic designated as $m = 1$, applies to basic T and π low-pass filter sections, whereas the remainder of the curves are for derived sections.

Referring to Figure 198, it is noted that the loss characteristics for derived sections have points of infinite attenuation (peak points). These occur at values of $\left(\frac{f}{f_c}\right)$ where the Z_A or Z_B arms of the sections are resonant or anti-resonant, as the case may be. For the T type, series m-derived section of Figure 196, the shunt arm is resonant when

$$j\,2\,\pi f \frac{1 - m^2}{4m} L_0 = j \frac{1}{2\pi f\, m\, C_0}$$

Let $f = f_R$ be the frequency for which this equation is valid. Solving we have

$$f_R{}^2 = \frac{1}{\pi^2\,(1 - m^2)\,L_0\,C_0}$$

Using the values of L_0 and C_0 given in (148) we have

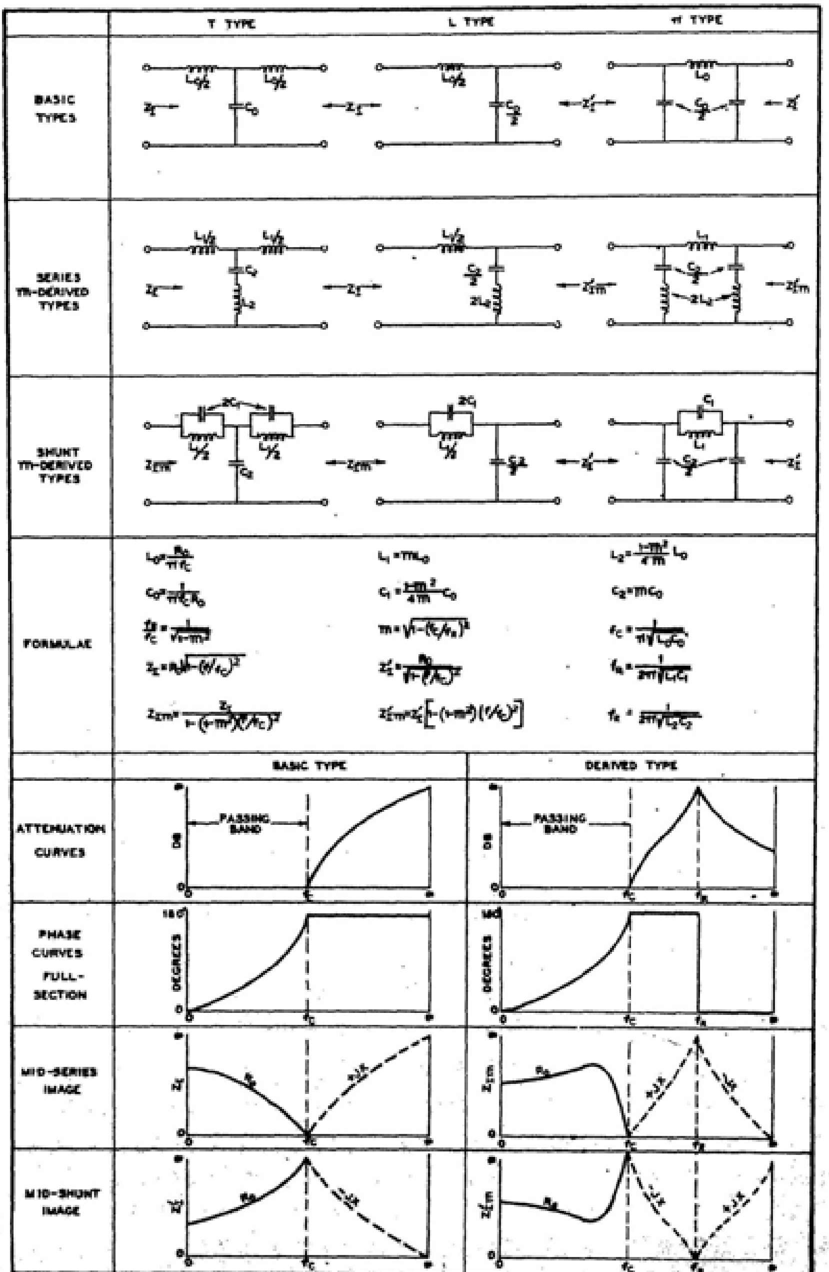
$$f_R = \frac{f_c}{\sqrt{1 - m^2}} \tag{158}$$

This equation states that when $m = 1$ the resonant frequency is infinite, and when $m = 0$, the resonant frequency is equal to the cut-off frequency f_c. In other words, the peak point for low-pass filter sections may be placed anywhere between the cut-off frequency f_c, and an infinite frequency by giving to m a value between zero and one. A basic section may be regarded as a derived section having $m = 1$ and $f_R =$ infinity.

6. SUMMARY OF LOW-PASS FILTER DATA

Chart XXVIII shows the low-pass filter sections of Figure 196 and the associated design data discussed above.

CHART XXVIII



Low-pass Filter Data

HIGH-PASS FILTERS

7. COMMON BASIC- AND DERIVED-TYPE SECTIONS

In the same manner as for the low-pass filters just discussed, a group of useful high-pass filter sections is obtained from the foregoing ladder-type structures when the series arm of reactance Z_1 is a capacitance and the shunt arm of reactance Z_2 is an inductance. Where L_0 indicates the inductance and C_0 the capacitance, this gives

$$Z_1 = -j\frac{1}{2\pi fC_0}$$

$$Z_2 = j2\pi fL_0 \tag{159}$$

When these elements are used in connection with the filter sections of Figure 191, we obtain the following high-pass filter sections of Figure 199.

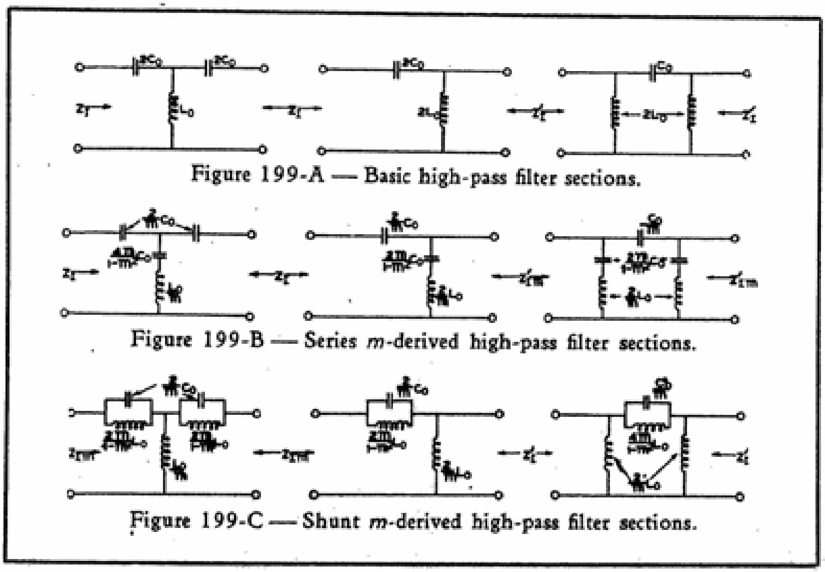


Figure 199-A — Basic high-pass filter sections.

Figure 199-B — Series m-derived high-pass filter sections.

Figure 199-C — Shunt m-derived high-pass filter sections.

Figure 199 — High-pass filter sections.

8. TRANSMISSION AND ATTENUATION RANGES

Again making use of the band limiting conditions of equation (141) in connection with the values of Z_1 and Z_2 shown above, we have

$$\frac{Z_1}{4Z_2} = \frac{1}{4}\left[-j\frac{1}{2\pi fC_0}\right]\left[\frac{1}{j2\pi fL_0}\right] = -\frac{1}{16\pi^2 f^2 L_0 C_0} = 0$$

The only value of f for which this equation is valid is when f is infinite, or in other words, one of the cut-off frequencies for the high-pass filter sections of Figure 199 is at an infinite frequency. This cut-off frequency for high-pass sections, like the zero cut-off frequency for low-pass sections, is not usually referred to as a cut-off point for the reason that it cannot be incorporated into design formulae. Now, from the other band limiting condition of (141) we may write:

$$\frac{Z_1}{4Z_2} = -\frac{1}{16\,\pi^2 f^2 L_0 C_0} = -1$$

If f_c is used to denote the frequency for which this equation is valid, we have

$$f_c = \frac{1}{4\,\pi\sqrt{L_0 C_0}} \qquad (160)$$

This frequency f_c denotes the lower limit of the transmission band of the above high-pass filter sections, that is, the transmission band extends from the frequency f_c to infinity. On this basis the attenuation range includes all frequencies between zero and f_c cycles per second.

9. ELECTRICAL CONSTANTS

Simultaneous solution of (160) in connection with the inverse relation existing between L_0 and C_0, namely

$$\frac{L_0}{C_0} = R_0^2$$

makes it possible to express L_0 and C_0 in terms of R_0 and f_c. We have

$$L_0 = \frac{R_0}{4\,\pi f_c}$$

$$C_0 = \frac{1}{4\,\pi f_c R_0} \qquad (161)$$

For any high-pass filter, therefore, when f_c and R_0 are known quantities, the values of the primary constants L_0 and C_0 may be computed from these formulae. The remaining properties of the high-pass filter sections can now be evaluated in terms of R_0, f_c and m.

10. IMAGE IMPEDANCES

Using these above values of L_0 and C_0, the following relations between Z_1 and Z_2 may be set up:

$$Z_1 = -j\frac{1}{2\pi f C_0} = -j\,2\,R_0\frac{f_c}{f}$$

$$Z_2 = j\,2\,\pi f\,L_0 = j\frac{R_0}{2}\frac{f}{f_c}$$

$$\frac{Z_1}{4Z_2} = -\left(\frac{f_c}{f}\right)^2$$

$$1 + \frac{Z_1}{4Z_2} = 1 - \left(\frac{f_c}{f}\right)^2 \tag{162}$$

By means of these relations, the image impedances Z_I, Z'_I, Z_{Im}, and Z'_{Im}, as given in equations (122), (123), (132) and (128) respectively, can be expressed in terms of R_0, f_c, and m as follows:

$$Z_I = R_0\sqrt{1-\left(\frac{f_c}{f}\right)^2} = \text{Mid-series image impedance of basic type and series } m\text{-derived high-pass filter sections.} \tag{163}$$

$$Z'_I = R_0\frac{1}{\sqrt{1-\left(\frac{f_c}{f}\right)^2}} = \text{Mid-shunt image impedance of basic type and shunt } m\text{-derived high-pass filter sections.} \tag{164}$$

$$Z_{Im} = \frac{R_0\sqrt{1-\left(\frac{f_c}{f}\right)^2}}{1-(1-m^2)\left(\frac{f_c}{f}\right)^2} = \text{Mid-series image impedance of shunt } m\text{-derived high-pass filter sections.} \tag{165}$$

$$Z'_{Im} = R_0\frac{1-(1-m^2)\left(\frac{f_c}{f}\right)^2}{\sqrt{1-\left(\frac{f_c}{f}\right)^2}} = \text{Mid-shunt image impedance of series } m\text{-derived high-pass filter sections.} \tag{166}$$

It is again noted that when the value of $m = 1$ is used with the derived image impedances of (165) and (166) the formulae break down to give formulae (163) and (164), which relate to the basic-type sections. For plotting purposes, formulae (165) and (166) may be written,

$$\frac{Z_{Im}}{R_0} = \frac{R_0}{Z'_{Im}} = \frac{\sqrt{1-\left(\frac{f_c}{f}\right)^2}}{1-(1-m^2)\left(\frac{f_c}{f}\right)^2} \tag{167}$$

From this formula the amount the image impedances Z_{Im} and Z'_{Im} depart from R_0 ohms may be plotted against the ratio $\left(\frac{f_c}{f}\right)$ for various values of m. It is not necessary to make any additional curves to show the graph of formula (167) as the curves of Figure 197 which were plotted to show the image impedances of low-pass filters are applicable simply by designating the horizontal as of $\left(\frac{f_c}{f}\right)$ instead of $\left(\frac{f}{f_c}\right)$.

11. ATTENUATION CHARACTERISTICS

Again, as for low-pass filters, the attenuation characteristics for the above high-pass filter sections may be obtained by means of the general attenuation formula of (139)

That is,

$$\text{Attenuation loss} = 20 \log \left[\frac{\sqrt{1+\frac{Z_1}{4Z_2}}+m\sqrt{\frac{Z_1}{4Z_2}}}{\sqrt{1+\frac{Z_1}{4Z_2}}-m\sqrt{\frac{Z_1}{4Z_2}}}\right]$$

For high-pass filter sections where $\frac{Z_1}{4Z_2} = -\left(\frac{f_c}{f}\right)^2$ this reduces to

$$\text{Attenuation loss} = 20 \log \left[\frac{\sqrt{1-\left(\frac{f_c}{f}\right)^2}+m\sqrt{-\left(\frac{f_c}{f}\right)^2}}{\sqrt{1-\left(\frac{f_c}{f}\right)^2}-m\sqrt{-\left(\frac{f_c}{f}\right)^2}}\right] \tag{168}$$

Now, in the attenuation range of high-pass filters, that is, in the frequency range from zero cycles per second to f_c cycles per second the ratio $\left(\frac{f_c}{f}\right)$ is always greater than unity, and hence the factor $\left[1-\left(\frac{f_c}{f}\right)^2\right]$ is always negative in this range. It is therefore desirable to rearrange (168) as follows:

$$\text{Attenuation loss} = 20 \log \left[\frac{\sqrt{\left(\frac{f_c}{f}\right)^2 - 1} + m\frac{f_c}{f}}{\sqrt{\left(\frac{f_c}{f}\right)^2 - 1} - m\frac{f_c}{f}} \right] \qquad (169)$$

This formula for high-pass filter sections is the same as that for low-pass filters except that it is expressed in terms of the ratio $\left(\frac{f_c}{f}\right)$ instead of $\left(\frac{f}{f_c}\right)$. Consequently, the curves of Figure 198 may also be used for high-pass filters as well as low-pass filters simply by using ratios $\left(\frac{f_c}{f}\right)$ instead of $\left(\frac{f}{f_c}\right)$ along the horizontal axis.

Infinite loss points occur in the attenuation curves whenever the ratio $\left[\frac{\sqrt{\left(\frac{f_c}{f}\right)^2 - 1} + m\left(\frac{f_c}{f}\right)}{\sqrt{\left(\frac{f_c}{f}\right)^2 - 1} - m\left(\frac{f_c}{f}\right)} \right]$ is infinite, or what is the same thing, when the denominator of the expression is zero. Thus

$$\sqrt{\left(\frac{f_c}{f}\right)^2 - 1} - m\left(\frac{f_c}{f}\right) = 0$$

Denoting the value of f for which this equation is valid by f_R, we have,

$$\frac{f_c}{f_R} = \frac{1}{\sqrt{1 - m^2}} \quad \text{or} \quad \frac{f_R}{f_c} = \sqrt{1 - m^2}$$

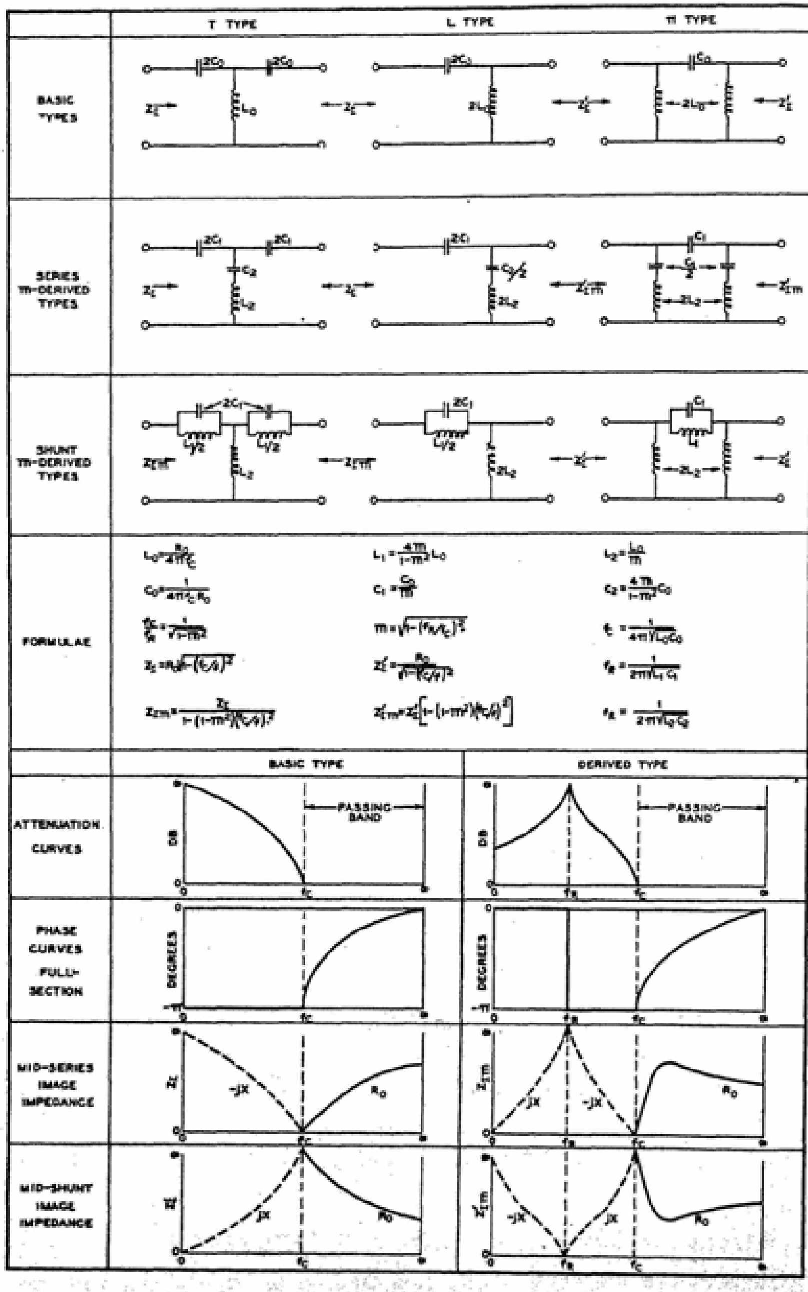
When, for a derived filter section, the value of $\left(\frac{f_R}{f_c}\right)$ is known, m may be determined from the formula

$$m = \sqrt{1 - \left(\frac{f_R}{f_c}\right)^2} \qquad (170)$$

12. SUMMARY OF HIGH-PASS FILTER DATA

Chart XXIX shows the high-pass sections of Figure 199 and the associated design data developed above.

CHART XXIX



High-pass Filter Data

13. SOME DESIGN CONSIDERATIONS

The foregoing work has supplied data for designing the common types of low-pass and high-pass filters. To review, it has been mentioned that where a filter consists of several sections, the sections are joined together on a matched image impedance basis. Charts XXVIII and XXIX, showing the various low-pass and high-pass filter sections, indicate the terminals of the sections that have like image impedance. It is only necessary to join these sections together at such like terminals to secure a properly connected composite filter.

The total attenuation loss for a filter consisting of several sections is simply the sum of the attenuation losses of the individual sections. For example, suppose a low-pass filter consists of two sections; one, a derived section of $m = 0.6$, and the other a basic-type section ($m = 1$). Referring to Figure 198, the attenuation loss for the two sections is obtained by adding the curve for $m = 0.6$ and the curve for $m = 1.0$. In arranging the sections in a filter to meet given attenuation requirements, advantage should be taken of the types of characteristics obtained from both the basic and derived sections. For instance, referring to Figure 198, it is noted that basic-type sections are useful to secure loss at frequencies remote from the cut-off, whereas derived sections give maximum attenuation at their peak points, the frequency location of which may be placed by the value assigned to m. Full sections should not be used where half-sections (L sections) will suffice. Filters, then, are built up of basic and derived sections with the peak points of the derived sections placed at strategic points to meet the attenuation characteristic desired.

In connection with attenuation losses it is to be remembered that the insertion loss of a filter is equal to its attenuation loss plus two correction losses, namely terminal and interaction losses. Because of the difficulty of presenting and using graphical data for estimating terminal and interaction losses, such data are not given in this book. Except in the vicinity of the cut-off frequency, however, a knowledge of the attenuation characteristic of a filter is sufficiently accurate for most purposes. In the cut-off range, reference may be made to filter design books when it is desired to predict precisely the transmission characteristic. Experience with the characteristics of typical filters, how ever, often makes this unnecessary.

The data of Figure 197 show that the image impedance characteristic for derived sections having a value of m equal to about 0.6 is approximately a constant resistance over a large part of the transmission band. These sections, then, make good end sections for filters where it is desired to reduce terminal and interaction losses by matching the filter

impedances with the connecting system loads. It is often necessary to use such sections for terminations even though their attenuation characteristic may not be used to the best advantage in building up the filter to meet desired requirements. Of course, if this is the case, half-sections are employed to avoid the unnecessary wasting of elements.

14. SERIES AND PARALLEL OPERATION OF LOW-PASS AND HIGH-PASS FILTERS

In certain electrical circuits, as for instance, dividing networks for loud-speaker systems, it is sometimes necessary to divide the frequency band into two parts by means of low-pass and high-pass filters operated in series or in parallel.

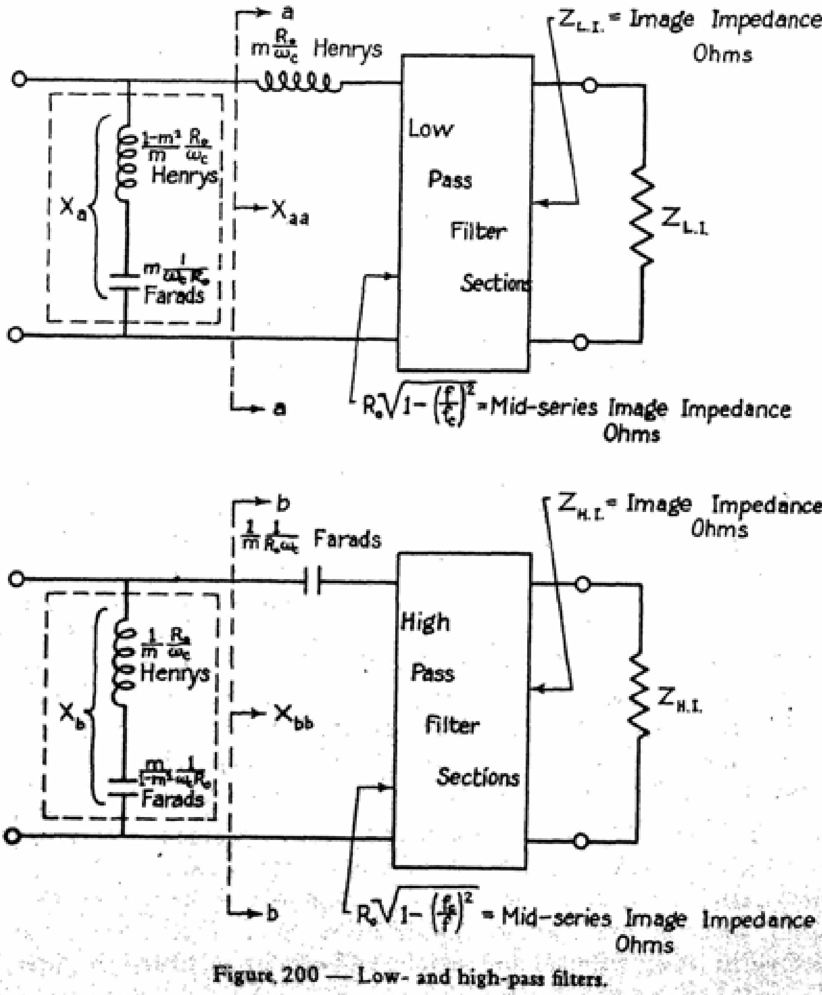


Figure 200 — Low- and high-pass filters.

In arranging a low-pass and a high-pass filter for series or parallel operation, it is only necessary to employ special design methods in connection with the first half-filter sections of each of the filters, the remaining sections being designed in accordance with conventional filter practice. For the first half-filter sections, m-derived types are used, the filters having mid-shunt terminations for parallel operation and mid-series terminations for series operation. Figure 200 shows a low-pass filter and a high-pass filter, each of which has its first half filter section shown in detail, and the remaining filter sections in block schematic form. It is recognized that the half sections shown are each mid-shunt terminated m-derived types, the constant "m" determining the place in the frequency range where infinite attenuation is theoretically obtained and also controlling the frequency configuration of the mid-shunt image impedances. This pair of filters when used individually and when the filter sections contained in the block schematics are designed according to conventional filter practice, should have reasonably good attenuation characteristics.

In showing the changes which it is necessary to make in the terminating sections of the filters of Figure 200 to make parallel operation possible, it is interesting to first consider some of the reactive impedances of the filters at different points. In order to be specific, it is assumed that each of the filters is designed for a cut-off frequency of f_o cycles per second, and that the low-pass filter has an image impedance of R_0 resistive ohms at zero frequency, and the high-pass filter has the same image impedance at an infinite frequency.

Figure 201 shows the reactive impedances listed below for the filters (these reactive characteristics were calculated for a value of m equal to 0.6, other curves for other values of m not being included because of the confusion they would cause in the drawing):

(1) Impedance characteristic in the passing band of the coil and condenser shunted across the input of the low-pass filter designated in Figure 200 as X_a.

(2) Impedance characteristic in the passing band of the coil and condenser shunted across the input of the high-pass filter designated in Figure 200 as X_b.

(3) Image impedance characteristic in the attenuation range of the low-pass filter at the point *aa* of the filter and designated as X_{aa}.

(4) Image impedance characteristic in the attenuation range of the high-pass filter at the point *bb* of the filter and designated as X_{bb}.

It is seen that the reactive impedances of curve $\frac{X_a}{R_0}$ and curve $\frac{X_{bb}}{R_0}$ are very nearly alike over the complete passing band of the low-pass filter, and likewise the same is true of curve $\frac{X_{aa}}{R_0}$ and curve $\frac{X_b}{R_0}$ over the passing band of the high-pass filter. This leads to the conception of

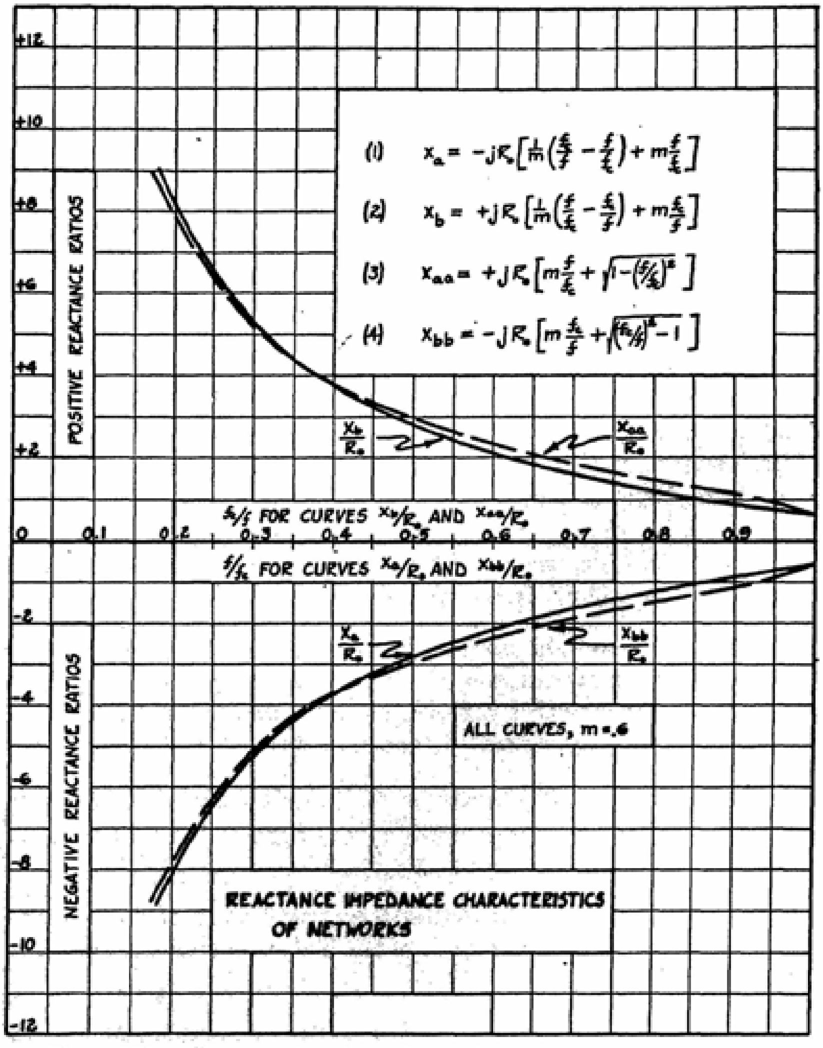


Figure 201 — Reactance impedance characteristic of networks in Figure 200.

placing the two filters in parallel by omitting the shunt coils and condensers at the input ends of the filters as in Figure 202-A. In other words, in the passing band of the low-pass filter, the filter elements of the high-

pass filter take the place of the shunt coil and condenser. Conversely, in the passing band of the high-pass filter, the filter elements of the low-pass filter take the place of the shunt coil and condenser omitted from the high-pass filter. Then in Figure 202-A the low-pass unit in its transmitting range should exhibit practically the same attenuation and impedance characteristics as the low-pass filter of Figure 200 and likewise for the high-pass unit in its passing range. In the attenuation ranges

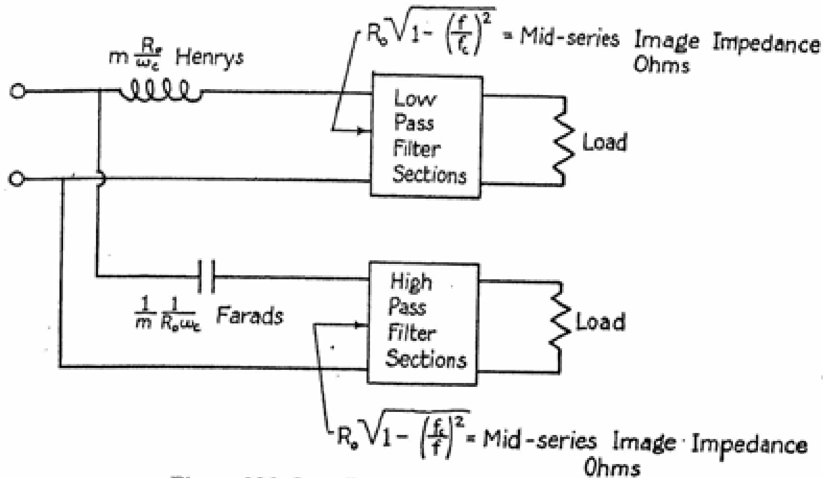


Figure 202-A — Parallel-type dividing network.

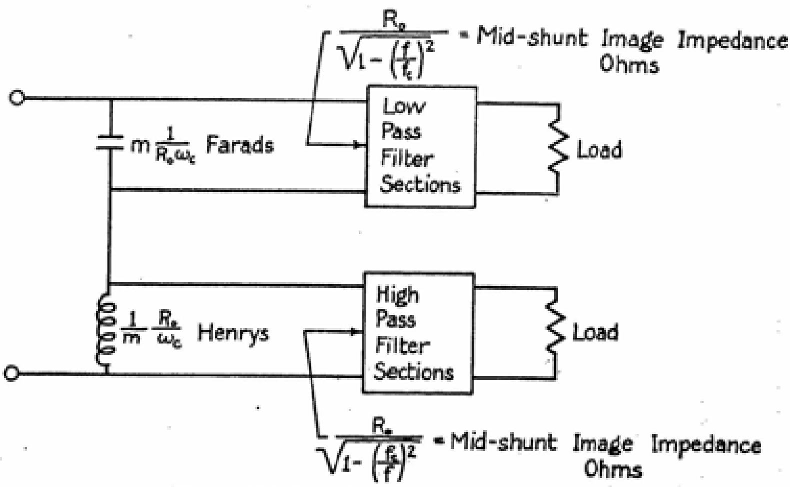


Figure 202-B — Series-type dividing network.

of the filters, the networks of Figures 200 and 202-A will not be respectively the same, but any lack of attenuation which the network of Figure 202-A has lost by paralleling can always be made up by adding more

sections in the block schematics. Figure 202-A then gives a method for designing the end half-sections of a low-pass filter and a high-pass filter, where they are to be operated in parallel.

The same procedure as above could be gone through for arriving at the design of a low-pass filter and a high-pass filter which are to be operated in series. This, however, is not necessary as the series type network can be arrived at by using the principle of inverse networks. The network of Figure 202-B is the inverse of the network of Figure 202-A and represents the design procedure employed for placing a low-pass filter and a high-pass filtet in series. For both the series and parallel methods a value of $m = 0.6$ is usually employed, since the networks are inverse to each other.

Chapter XX

DIVIDING NETWORKS FOR LOUD-SPEAKER SYSTEMS

By HARRY KIMBALL

In the design of linear sound reproducing equipment where it is desired to faithfully reproduce tones from about 50 cycles per second to about 8,000 cycles per second, it is common practice to divide the frequency range into two or more parts and provide one or more loud-speakers for each of these frequency ranges. The speakers employed for the different bands are, of course, differently designed, each speaker being particularly suitable for its own band. Since it is not possible to design speakers which will faithfully and efficiently reproduce frequencies in one preassigned band, and sharply attenuate frequencies outside of the band, it is necessary to supply an electrical network between the final power amplifiers and the speakers to deliver the correct frequency band to each of the sets of loud-speakers. These networks have acquired the name of "dividing networks."

In practice, loud-speaker systems may be of the two-way or three-way types. Because of the predominance of two-way systems, only networks for use with these systems are discussed in this text.

For the two-way system the speakers handling the lower frequencies are termed the low-frequency speakers or low-range speakers. In like manner, the speakers which reproduce higher frequencies are called the upper-frequency speakers or upper-range speakers. For each of the two frequency bands one speaker unit or a number of speakers are arranged in series-parallel combinations to secure the proper combined load.

Dividing networks are not usually of the sharp cut-off type, that is, they are not arranged to uniformly transmit frequencies of a given band and then sharply attenuate all other frequencies. Rather, they transmit the band frequencies almost uniformly and gradually slope off, thereby giving a certain amount of over-lap between the assigned frequency ranges. While theoretically it may seem desirable to arrange dividing networks to cut-off sharply, from a commercial standpoint the sharpness of cut-off is necessarily a compromise between expense and effectiveness. For well designed speaker systems, the rate of change of attenuation should at least be sufficient to suppress objectionable irregularities

in the response of one horn in its transmitting range because of sound coming from the other horn in its suppression range. From an analysis of a large number of speaker systems it appears that dividing networks should provide at least 10 to 12 db attenuation one octave away from their cut-offs. In considering networks having greater rates of change of attenuation it should be remembered that increased attenuation is accompanied by increased losses in the transmitting ranges, which, for high powered systems at least, is to be avoided. Costs also may mount up unreasonably if a large amount of filtering is employed. For these reasons, and considering the magnitude of the irregularities which one speaker produces in the transmitting range of the other, it appears that few dividing networks should employ more attenuation than about 18 db per octave.

In a two-way system, the frequency where both sets of speakers receive equal amounts of energy is called the cross-over point. In other words, the cross-over point is the point of separation between the two bands of frequencies. In developing speaker systems a trial cross-over point is usually arbitrarily selected, keeping in mind the characteristics of the upper- and lower-range speakers which are to be used, costs and other items, and later moved one way or the other if found unsatisfactory when the system is operated as a whole.

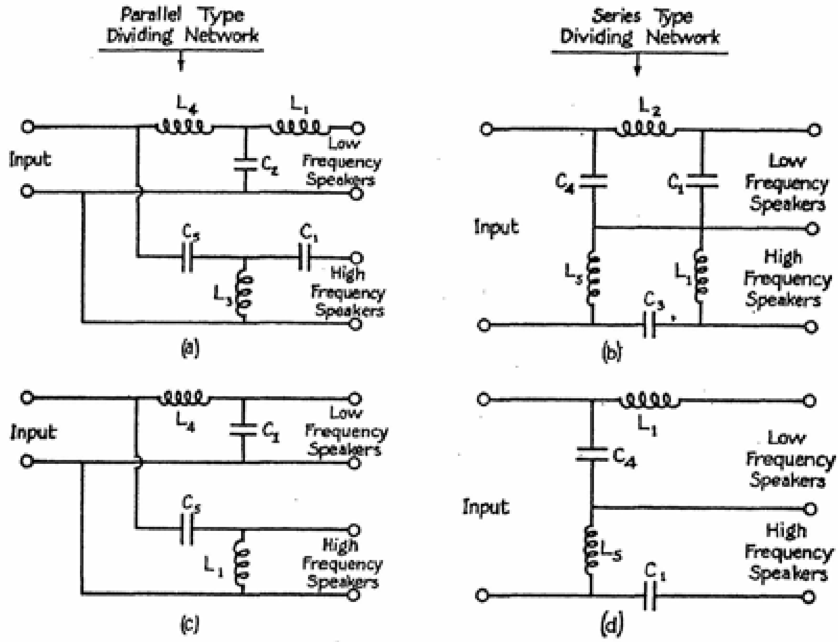
A two-way dividing network consists of a low-pass filter and a high-pass filter designed to operate from a common source at their input terminals. Two methods are in general use for connecting the filters in series or in parallel at their input terminals; namely, (1) the filter method discussed in Chapter XIX; and (2) the constant resistance method taken up later in this chapter. Each of these methods is capable of giving good results. The filter method is the more commonly used—probably because it is better known—and is somewhat more flexible in design.

1. SERIES-PARALLEL FILTER-TYPE DIVIDING NETWORKS

Chapter XIX has provided methods for designing the first half-sections of low-pass and high-pass filters for series and parallel operation. Referring to Figure 202, it is necessary for dividing network purposes to determine the number of filter sections required for the block schematic portions of the networks.

As already mentioned, the number of these sections depends upon the rate of attenuation desired in the suppressed frequency ranges, or in other words, how much frequency overlap of the low-frequency range and the high-frequency range is to be permitted. Where it is

desired to secure an attenuation rate of change of about 18 db per octave, one filter section for each of the block schematics of Figure 202 is sufficient, and where a rate of change of about 12 db per octave is satisfactory, half-sections will suffice.



$$L_1 = \frac{R_0}{\omega_c} \text{ henries} \qquad C_1 = \frac{1}{\omega_c R_0} \text{ farads}$$

$$L_2 = \frac{2R_0}{\omega_c} \text{ henries} \qquad C_2 = \frac{2}{\omega_c R_0} \text{ farads}$$

$$L_3 = \frac{R_0}{2\omega_c} \text{ henries} \qquad C_3 = \frac{1}{2\omega_c R_0} \text{ farads}$$

$$L_4 = (1 + m)\frac{R_0}{\omega_c} \text{ henries} \qquad C_4 = (1 + m)\frac{1}{\omega_c R_0} \text{ farads}$$

$$L_5 = \frac{1}{1 + m}\frac{R_0}{\omega_c} \text{ henries} \qquad C_5 = \frac{1}{1 + m}\frac{1}{\omega_c R_0} \text{ farads}$$

Figure 203 — Dividing network designs.

Figure 203 shows four dividing network circuits, two of which are parallel types and the other two, series types. Network (a) is secured by the use of basic-type T sections for the schematic portions of the parallel type circuit of Figure 202, a low-pass section being used for the low-fre-

quency side and a high-pass section for the upper-frequency side. Network (b) is obtained by the use of π type low-pass and high-pass sections with the series type circuit of Figure 202. Networks (c) and (d) are secured in the same manner, except that L-type low-pass and high-pass filter sections are used. The formulae given in Figure 203 provide means of computing the electrical constants of the networks for any value of R_0 and f_c.

In Figure 204 the values of the electrical constants for the elements of the different networks of Figure 203 have been tabulated for a cut-off frequency of f_c c.p.s., and a value of R_0 equal to 10 ohms. For other

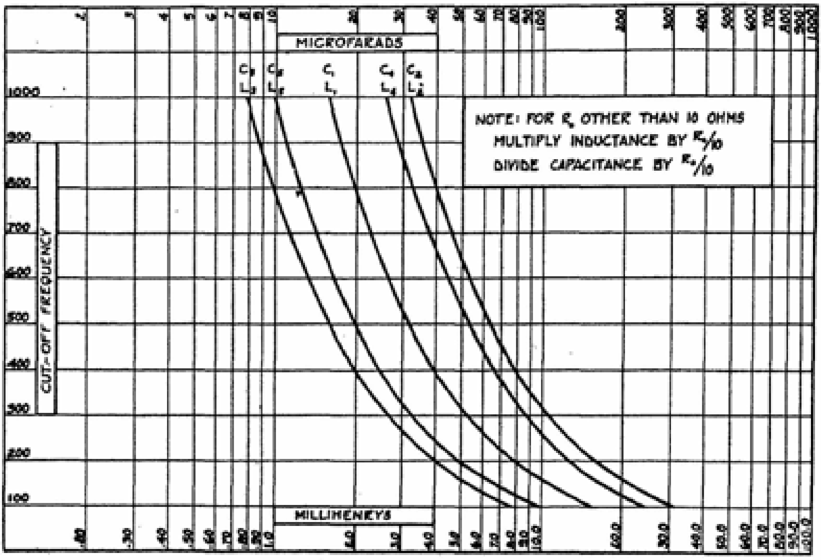


Figure 204 — Electrical constant for Figure 203 networks, $m = 0.6$ and $R_0 = 10$ ohms.

values of R_0 the values of the coil inductances increase directly with increase in the value of R_0, and capacitances decrease with increase in the value of R_0.

For computing these data the value of m was taken as 0.6 as this value was shown in Chapter XIX to be the best for the series and parallel operation of filters.

2. ATTENUATION AND PHASE CHARACTERISTICS

The attenuation characteristics for the different dividing networks of Figure 203 when operating between resistances of R_0 ohms are shown

in Figure 205. In preparing the curves it was assumed that the coils and condensers were non-resistive. In practice, where the component electrical elements contain some resistance, the curves will be slightly different, especially in the cross-over region. This, in some cases, has the effect of shifting the cross-over point slightly away from the theoretical cut-off point.

The amount of attenuation which the networks exhibit in their passing bands is especially important for high powered sound systems.

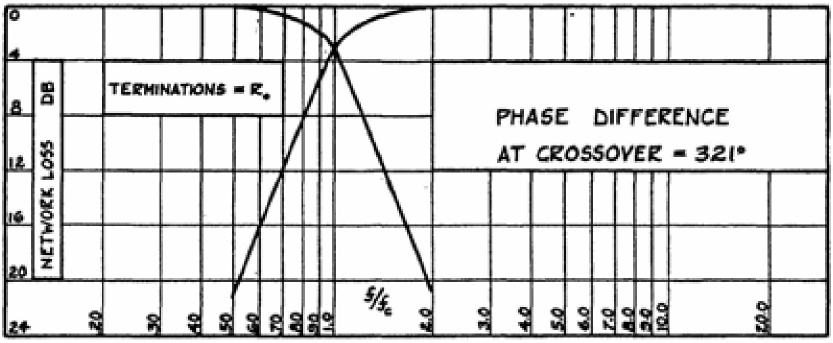


Figure 205-A — Attenuation characteristics of networks (a) and (b), Figure 203.

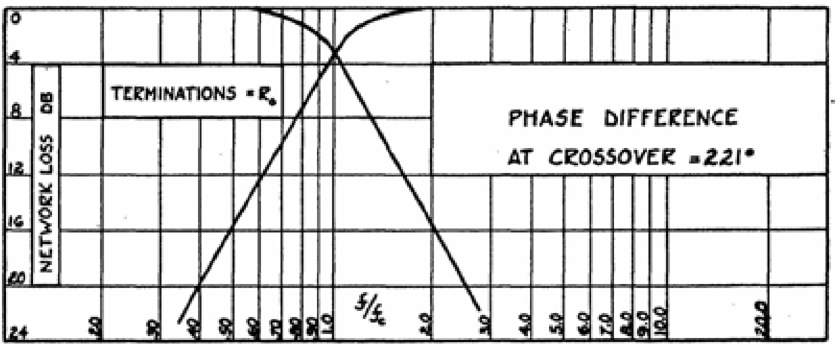


Figure 205-B — Attenuation characteristics of networks (c) and (d), Figure 203.

For instance, in a 100-watt system one db loss in the dividing network means about 25 watts loss of power. This, it is easy to realize, is an important loss, because the final power amplifiers must supply this amount of power in addition to that needed to produce the desired sound energy. Where care is taken to make use of low resistance coils for the dividing network, this loss can be reduced to about 0.5 db. In arriving at the most desirable relative location for the high-range and low-range speakers to achieve this effect, it is useful to have available the

phase shift of the dividing networks at the cross-over points. These phase shifts are noted in Figure 205.

3. IMPEDANCE CHARACTERISTICS

The impedances obtained at the input terminals of the dividing networks of Figure 203 vary, of course, from network to network. In general, the impedances of any of the corresponding series and parallel types are inverse to each other. This follows from the manner in which the series-type networks were derived from the parallel-type networks. In order to obtain image impedance characteristics at the input terminals of the networks, it would be necessary to use image impedance matching loads at the output terminals as shown in Figure 200. In practice, where the speakers are arranged in series-parallel combinations to give load resistances as closely as possibly to R_0 ohms at the high- and low-range output terminals, the input impedances will not adhere strictly to image impedance characteristics. However, for networks (a) and (b) which contain a fair amount of masking because of the intervening filter sections, this change should not be great in the filter passing bands.

In connection with networks (c) and (d) which provide a lesser amount of masking, it is instructive to view the input impedances obtained for various values of "m" and for speaker load resistances of R_0 ohms. Figures 206-A and 206-B give the resistive component and the reactive component, respectively, of the input impedance for network (d) of Figure 203. It is observed that some improvement in the sending-end impedance might be obtained by using $m = 0.45$ instead of $m = 0.6$. The change in attenuation curve which a design of this nature produces, however, is very small and for this reason design data for such a network were not included in Figure 204.

For the designs of Figure 203, some have mid-series images impedance characteristics at the sets of output terminals and others are mid-shunt terminated. A feature of the mid-series termination is that the image impedance in the passing band is a resistance of R_0 ohms at frequencies remote from the cut-off, and gradually reduces, theoretically, to zero ohms at the cut-off. The mid-shunt termination, on the other hand, is R_0 ohms at points remote from the cut-off and gradually increases to infinity ohms at the cut-off. That is, these two image impedance characteristics are inverse to each other. Now it so happens that low-frequency dynamic speakers have a mechanical anti-resonance point at frequencies of 100 c.p.s. or lower. This results in the low-frequency speakers having a relatively high input impedance at the lower frequencies which gradually reduces to the nominal impedances as the frequency increases. Hence in some cases more uniform overall attenuation

results have been secured in connection with low-frequency dynamic speakers by using dividing networks having internal mid-series image impedances at the low-frequency horn terminals. For instance, network

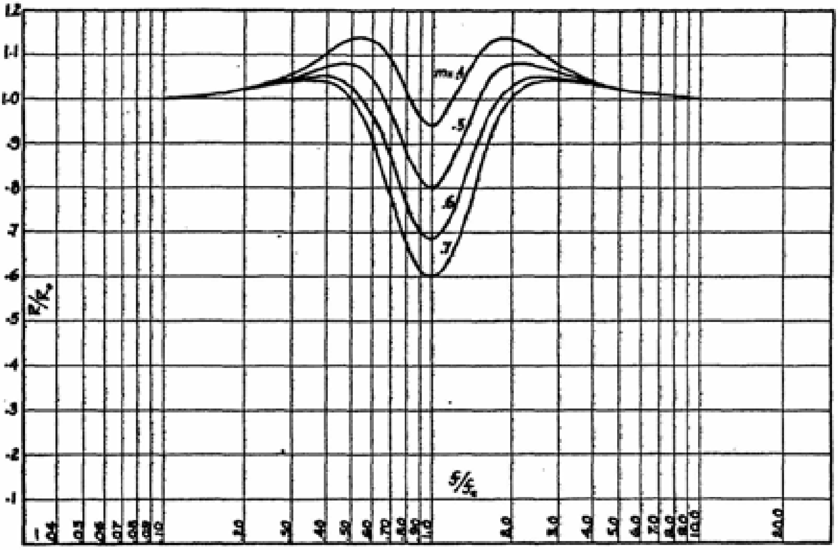


Figure 206-A — Resistive component of impedance at input terminals of network (d), Figure 203.

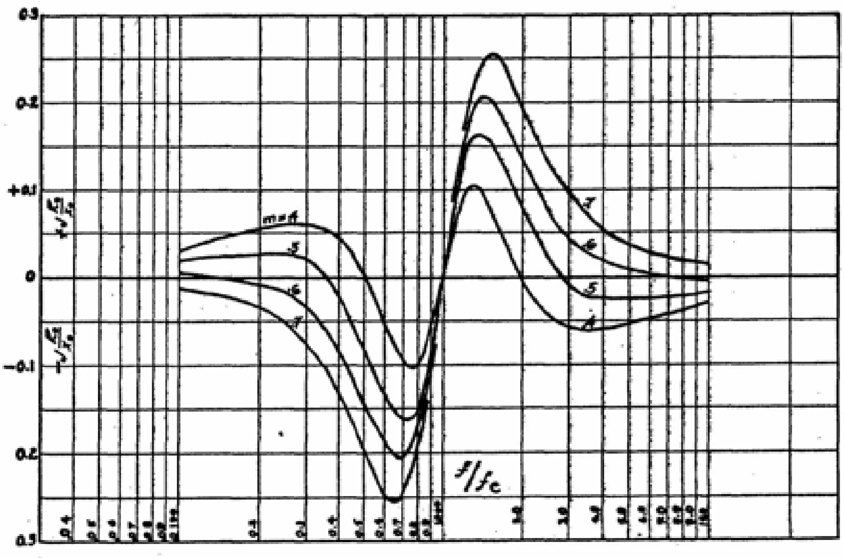


Figure 206-B — Reactive component of impedance at input terminals of network (d), Figure 203.

(d) of Figure 203 in actual operation gave better results than network (c) because of better impedance matching conditions for the low-frequency units.

4. CONSTANT RESISTANCE DIVIDING NETWORKS

A constant resistance dividing network is one whose resistance at the input terminals is constant with change of frequency for the proper resistive loads connected to the two pairs of output terminals. The circuits of Figure 207 may be used as a basis for the design of such networks. Referring to this figure, Z_1 and Z_2 in all of the circuits are specified as being inverse to each other with respect to the resistance R_0, that is $Z_1 Z_2 = R_0^2$.

For Circuit A, Figure 207, the impedance presented at the 1-2 terminals is

$$Z_{12} = \frac{R_0 Z_1}{R_0 + Z_1} + \frac{R_0 Z_2}{R_0 + Z_2}$$

$$Z_{12} = R_0 \left[\frac{Z_1}{R_0 + Z_1} + \frac{R_0}{R_0 + Z_1} \right]$$

$$Z_{12} = R_0 \text{ ohms} \tag{171}$$

For Circuit B, Figure 207, we have in a similar manner

$$Z_{12} = \frac{\frac{Z_1}{\sqrt{2}} \left[R_0 + \frac{Z_2}{\sqrt{2}} \right]}{R_0 + \frac{Z_1}{\sqrt{2}} + \frac{Z_2}{\sqrt{2}}} + \frac{\frac{Z_2}{\sqrt{2}} \left[R_0 + \frac{Z_1}{\sqrt{2}} \right]}{R_0 + \frac{Z_1}{\sqrt{2}} + \frac{Z_2}{\sqrt{2}}}$$

$$Z_{12} = \frac{R_0 \frac{Z_1}{\sqrt{2}} + R_0 \frac{Z_2}{\sqrt{2}} + Z_1 Z_2}{R_0 + \frac{Z_2}{\sqrt{2}} + \frac{Z_1}{\sqrt{2}}} = R_0 \text{ ohms} \tag{172}$$

From these equations we learn that the impedances looking into the 1-2 terminals of the Circuits A and B of Figure 207 are resistances of R_0 ohms. Another pair of circuits also presenting constant resistance impedances at their terminals may be obtained by taking the inverse of Circuits A and B, of Figure 207, with respect to R_0. Doing this we secure C of Figure 207 as the inverse of Circuit A, and Circuit D as the inverse of Circuit B.

Now let Z_1, the reactive impedance, be an inductance of L_0 henries having a numerical value of R_0 ohms at some frequency f_a. In a similar manner let the Z_2 circuit be the corresponding inverse capacitance C_0 of

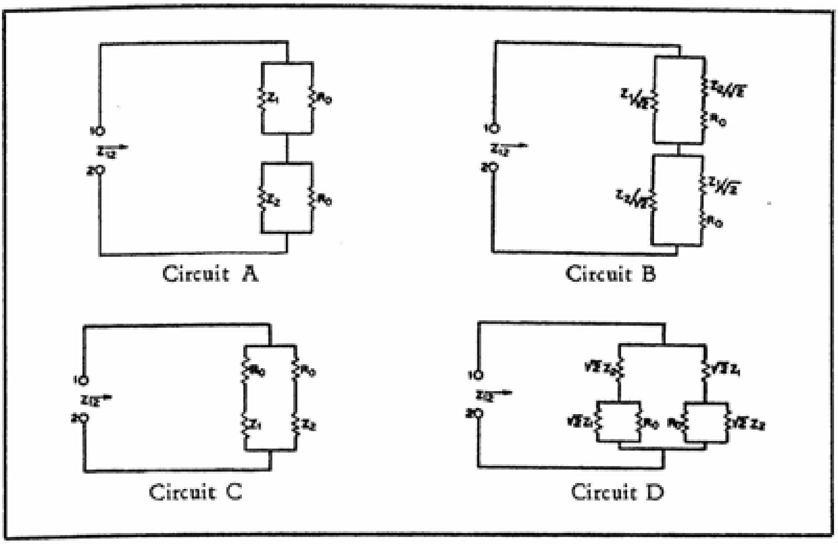


Figure 207 — Constant resistance circuits.

such a value that Z_2 is also numerically equal to R_0 at the frequency f_a. We have

$$Z_1 = j\,2\pi f\,L_0 = j\,R_0\frac{f}{f_a} \tag{173}$$

$$Z_2 = -j\frac{1}{2\pi f\,C_0} = -j\,R_0\frac{f_a}{f} \tag{174}$$

$$L_0 = \frac{R_0}{2\pi f_a} \qquad\qquad C_0 = \frac{1}{2\pi f_a R_0}$$

We may also write

$$\frac{Z_1}{\sqrt{2}} = j\,2\pi f\left(\frac{L_0}{\sqrt{2}}\right) = j\,2\pi f\,L_1 \text{ where } L_1 = \frac{L_0}{\sqrt{2}} \tag{175}$$

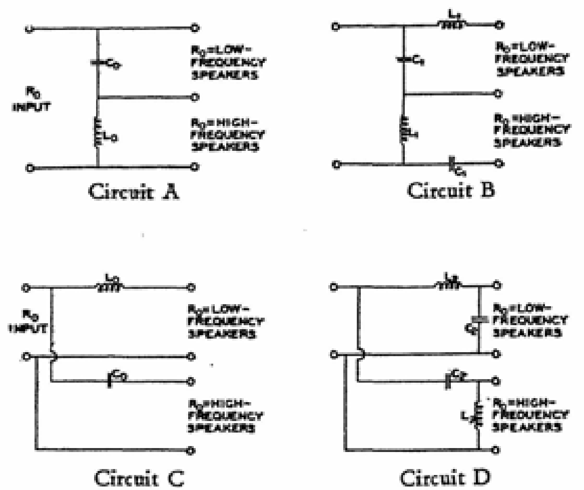
$$\sqrt{2}\,Z_1 = j\,2\pi f\,(\sqrt{2}L_0) = j\,2\pi f\,L_2 \text{ where } L_2 = \sqrt{2}\,L_0 \tag{176}$$

$$\frac{Z_2}{\sqrt{2}} = -j\frac{1}{2\pi f\,(\sqrt{2}\,C_0)} = -j\frac{1}{2\pi f\,C_1} \text{ where } C_1 = \sqrt{2}\,C_0 \tag{177}$$

$$\sqrt{2}\,Z_2 = -j\frac{1}{2\pi f\left(\frac{C_0}{\sqrt{2}}\right)} = -j\frac{1}{2\pi f\,C_2} \text{ where } C_2 = \frac{C_0}{\sqrt{2}} \tag{178}$$

Using these element values for the Z_1 and Z_2 impedances of Figure 207, and rearranging the circuits somewhat, we obtain the circuits of

Figure 208 where the resistances R_0 are now shown as loud-speaker loads. It will be recognized that these circuits are in forms suitable for



$$L_0 = \frac{R_0}{2\pi f_a} \qquad L_1 = \frac{L_0}{\sqrt{2}} \qquad L_2 = \sqrt{2}\, L_0$$

$$C_0 = \frac{1}{2\pi f_a R_0} \qquad C_1 = \sqrt{2}\, C_0 \qquad C_2 = \frac{C_0}{\sqrt{2}}$$

f_a = Cross-over frequency of network
Inductances in henries.
Capacitances in farads.

Figure 208 — Constant resistance dividing networks.

dividing networks, and, of course, the impedance looking into the networks with the proper speaker loads is R_0 ohms. It now remains to determine the transmission loss to the low- and high-speaker units for each of the circuits.

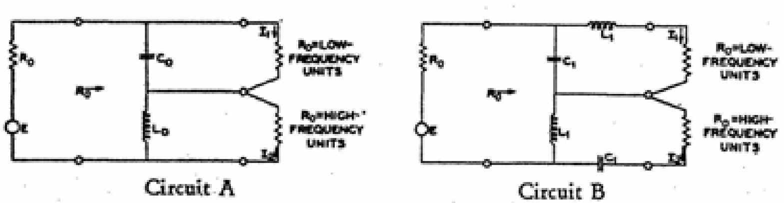


Figure 209 — Constant resistance networks operating between impedances of R_0 ohms.

Circuit A, Figure 209, shows the dividing network of Circuit A, Figure 208, working between resistances of R_0 ohms at all pairs of terminals. The currents I_1 and I_2 delivered by the source of supply to the low-frequency units and to the high-frequency units, respectively, are as follows:

$$I_1 = \frac{E}{2R_0} \frac{\frac{1}{j\omega C_0}}{R_0 + \frac{1}{j\omega C_0}} \qquad I_2 = \frac{E}{2R_0} \frac{j\omega L_0}{R_0 + j\omega L_0} \quad (179)$$

By means of equations (173) and (174) these can be written

$$I_1 = \frac{E}{2\,R_0} \frac{-j\,R_0 \frac{f_a}{f}}{R_0 - j\,R_0 \frac{f_a}{f}} = \frac{E}{2\,R_0} \frac{1}{1 + j\frac{f}{f_a}}$$

$$I_1 = \frac{E}{2\,R_0} \frac{1}{\sqrt{1 + \left(\frac{f}{f_a}\right)^2}} \angle\phi \quad (180)$$

$$I_2 = \frac{E}{2\,R_0} \frac{j\,R_0 \frac{f}{f_a}}{R_0 + j\,R_0 \frac{f}{f_a}} = \frac{E}{2\,R_0} \frac{1}{1 - j\frac{f_a}{f}}$$

$$I_2 = \frac{E}{2\,R_0} \frac{1}{\sqrt{1 + \left(\frac{f_a}{f}\right)^2}} \angle\phi + 90^\circ \quad (181)$$

Since we are not interested in the angle ϕ, we will not determine its value here.

The current I_0 delivered to either the low or the high units, when separately connected to the source of supply without the use of the network, is

$$I_0 = \frac{E}{2\,R_0}$$

Then the insertion loss equations for the low-frequency side and for the high-frequency side of the network of Circuit A, Figure 209, are

obtained from the current ratios $\left(\frac{I_0}{I_1}\right)$ and $\left(\frac{I_0}{I_2}\right)$ respectively. We have

Insertion loss for low-frequency filter of Circuit A, Figure 209.

$$\begin{cases} = 20 \log \sqrt{1 + \left(\frac{f}{f_a}\right)^2} \\ = 10 \log \left[1 + \left(\frac{f}{f_a}\right)^2\right] \end{cases} \qquad (182)$$

Insertion loss for high-frequency filter of Circuit A, Figure 209.

$$\begin{cases} = 20 \log \sqrt{1 + \left(\frac{f_a}{f}\right)^2} \\ = 10 \log \left[1 + \left(\frac{f_a}{f}\right)^2\right] \end{cases} \qquad (183)$$

In the same manner the insertion loss equation for the network of Circuit B, Figure 208, can be obtained. Circuit B, Figure 209, shows this network connected between resistances of R_0 ohms at all pairs of terminals. The currents I_1 and I_2 are obtained as follows:

$$I_1 = \frac{E}{2\,R_0} \frac{\frac{1}{j\omega C_1}}{R_0 + j\omega L_1 + \frac{1}{j\omega C_1}}$$

$$I_1 = \frac{E}{2\,R_0} \frac{-j\frac{R_0 f_a}{\sqrt{2} f}}{R_0 + j\frac{R_0 f}{\sqrt{2} f_a} - j\frac{R_0 f_a}{\sqrt{2} f}}$$

$$I_1 = \frac{E}{2\,R_0} \frac{-j\frac{f_a}{f}}{\sqrt{2} + j\left(\frac{f}{f_a} - \frac{f_a}{f}\right)}$$

$$I_1 = \frac{E}{2\,R_0} \sqrt{\frac{1}{1 + \left(\frac{f}{f_a}\right)^4}} \quad \underline{/\phi - 90^\circ} \qquad (184)$$

The current I_2 is

$$I_2 = \frac{E}{2\,R_0}\,\frac{j\,\omega\,L_1}{R_0 + j\,\omega\,L_1 + \dfrac{1}{j\,\omega\,C_1}}$$

$$I_2 = \frac{E}{2\,R_0}\sqrt{\frac{1}{1+\left(\dfrac{f_a}{f}\right)^4}}\;\underline{\angle \phi + 90^\circ} \qquad (185)$$

The insertion loss equations for the low-frequency side and the high-frequency side of the network of Circuit B, Figure 208, are determined from the ratios $\left(\dfrac{I_0}{I_1}\right)$ and $\left(\dfrac{I_0}{I_2}\right)$. This gives

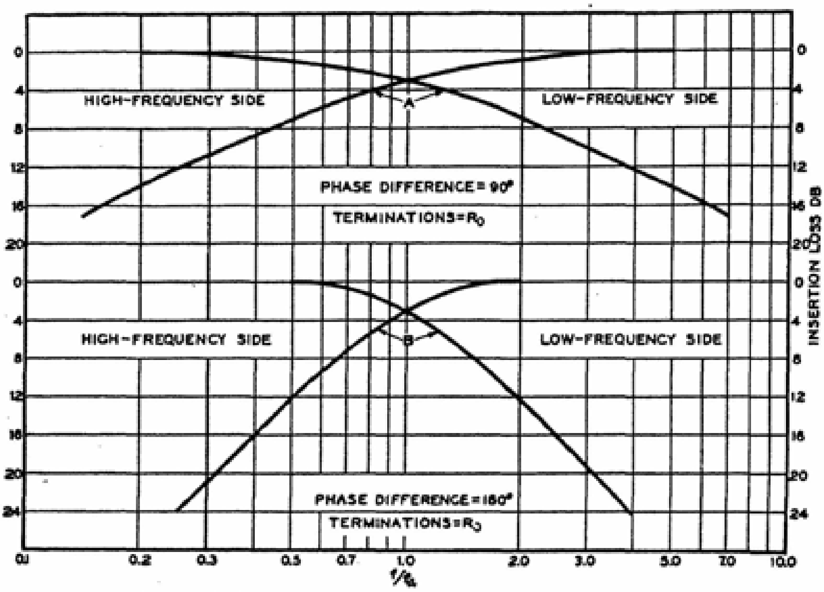
$$\begin{array}{l} \text{Insertion loss for} \\ \text{low-frequency channel} \\ \text{of Circuit B, Figure 209.} \end{array} \left\{ \begin{aligned} &= 20 \log \sqrt{1+\left(\frac{f}{f_a}\right)^4} \\ &= 10 \log \left[1+\left(\frac{f}{f_a}\right)^4\right] \end{aligned} \right. \qquad (186)$$

$$\begin{array}{l} \text{Insertion loss for} \\ \text{high-frequency channel} \\ \text{of Circuit B, Figure 209.} \end{array} \left\{ \begin{aligned} &= 20 \log \sqrt{1+\left(\frac{f_a}{f}\right)^4} \\ &= 10 \log \left[1+\left(\frac{f_a}{f}\right)^4\right] \end{aligned} \right. \qquad (187)$$

The above insertion loss formulae, as derived, apply to networks A and B of Figure 208. Because of the inverse relationship existing between networks A and C and also between networks B and D it can be shown that equations (182) and (183) apply to circuit C as well as to A and equations (186) and (187) apply to circuit D as well as to B.

Figure 210 is a graph of the insertion losses for the networks of Figure 208 plotted against the ratio $\left(\dfrac{f}{f_a}\right)$. It is noted that the cross-over point occurs at the frequency $f = f_a$ where the insertion loss is 3 db and hence the frequency f_a is the cross-over frequency of the networks. The insertion loss characteristics of curves A of Figure 210 apply to the

networks A and C, Figure 208. The insertion loss change per octave is seen to be relatively small and for this reason these networks are seldom used in practice. The insertion loss characteristics of Curves B,



Curves A for networks of Circuits A and C, Figure 208.
Curves B for networks of Circuits B and D, Figure 208.

Figure 210 — Constant resistance dividing networks insertion loss characteristics.

Figure 210, apply to the networks B and D, Figure 208. These characteristics are not quite as good as those of Figure 205-B which apply to the filter-type dividing networks (c) and (d) of Figure 203. Comparison of the two characteristics shows that the filter-type networks are slightly better both in the transmission and in the attenuation ranges. For some purposes it might be construed that the constant resistance feature of the networks C and D, Figure 208, might be an advantage, but the value of this is doubtful when one remembers that the connecting speaker loads may depart considerably from the ideal loads. From a manufacturing standpoint it is simpler to construct the constant resistance networks than the corresponding filter-type networks because of the fact that the inductance coils are alike as are the condensers.

Referring to equations (180) and (181) applying to networks A and C of Figure 208, it is seen that the currents I_1 and I_2 are 90 degrees out-of-phase with each other for any frequency. Also, equations (184) and (185) indicate that the phase relation for

networks B and D of Figure 208 is 180 degrees for any frequency. These data are given in Figure 210 for reference purposes in phasing the loud-speaker units.

5. CONSTRUCTIONAL FEATURES

The methods employed for assembling and wiring a dividing network may vary considerably depending upon system requirements.. In regard to the filter coils, as already mentioned, it is desirable to give considerable thought to their type and to their effective resistance. Where iron or some form of steel is used for the coil cores, modulation results if the coils are overloaded. Also, if the effective resistance is unduly large, excessive loss in the filter bands will be obtained. It has been found that large size air core coils solve the problem about as effectively as any other method.

Chapter XXI

VACUUM TUBES

By FRED ALBIN

1. TWO-ELEMENT VACUUM TUBE

There are numerous types of vacuum tubes available for a great number of diverse purposes.

The more common type, and the one usually used in amplifiers, is the hot cathode high-vacuum tube. As explained in Chapter XXXVIII, this tube depends for its operation upon the flow of electrons in the evacuated space between the cathode and the anode. This flow of electrons constitutes an electric current.

The electrons are negative particles of electricity, boiled from the interior of, and evaporated from the surface of the metal comprising the cathode—this taking place as a result of the high temperature of the metal. The rate at which the electrons are emitted from the surface is a function of the temperature and depends upon the material comprising the cathode. The expression for the emission rate in amperes per square centimeter of radiating surface is known as Richardson's equation:

$$I_s = A\sqrt{T}\,\epsilon^{-B/T} \tag{188}$$

where $I_s =$ radiation current in amperes per sq. cm.

$A =$ a constant of the emitting surface

$B =$ another constant involving the work function (in volts) for the cathode material.

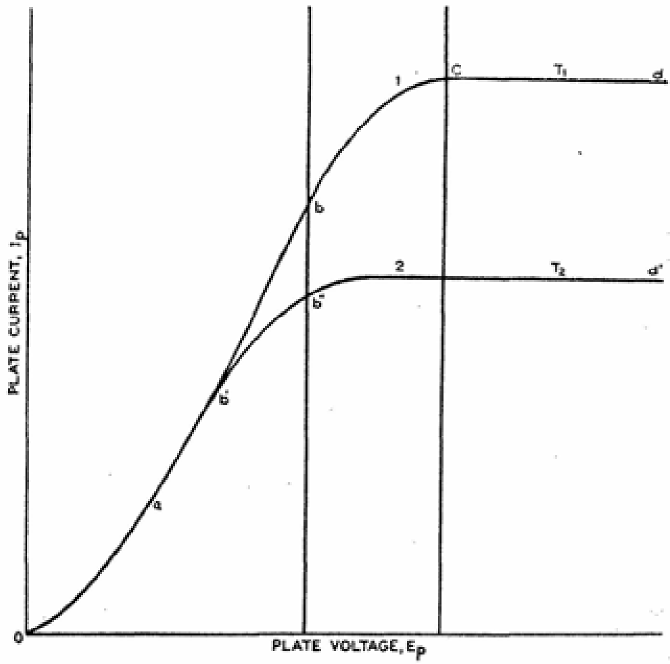
$T =$ temperature

The materials used for cathodes are: Pure tungsten metal, tungsten impregnated with thorium, and tungsten coated with barium and strontium oxides. From the standpoint of efficiency of the cathode (measured in terms of milliamperes of emitted current per watt of cathode heating energy), the pure tungsten is the least efficient, and is used only in rectifier and amplifier tubes with high anode voltages. Impregnating the tungsten with thorium greatly increases the efficiency of the cathode, enabling it to emit a sufficient quantity of electrons at a lower temperature, thus reducing the loss of energy by heat radiation and prolonging the life period of the tube. A cathode of barium and

strontium oxide mixture is the most efficient, and is usually used for small tubes. The efficiency of oxide-coated cathodes over pure tungsten is about 35 to 1.

Cathodes may consist of either a wire through which the heating current is passed, or a metallic sleeve surrounding, but electrically insulated from, the heating element. The latter design is used when the heating current is to be alternating.

As the negative electrons leave the cathode, the remaining charge is positive, and the wandering electrons eventually return to the attractively charged cathode unless there is an anode in the vicinity. If there is an anode within the evacuated enclosure, which is connected electrically to the cathode through an external circuit, the current flows through this connection to the cathode and neutralizes the charge. For proper operation of the vacuum tube as an amplifier or rectifier, the cathode must emit a copious supply of electrons. The rate at which they flow to the cathode is determined by the electric field strength in the



—From Journal of the S.M.P.E., February, 1935. "Simple Theory of the Three-Electrode Vacuum Tube," H. A. Pidgeon.

Figure 211 — Plate current-plate voltage characteristics of a two-element vacuum tube.

vicinity of the cathode. The field is produced by the charge in the anode (determined by the potential applied externally and usually known as

the plate potential), and the space charges produced by the pressure of negative electrons in the space. The relation between plate current and voltage is

$$I_P = K E_P^{3/2} \tag{189}$$

where I_P = plate current
E_P = plate voltage
and K = a constant (within certain limits) depending upon the size and proportions of the tube.

Figure 211 is a plot of the relation given in equation (189).

T_1 and T_2 represent curves with the cathode at normal and reduced temperatures, respectively. The current for any voltages below b' is normal and not affected by small changes of temperature. However, age has an effect similar to that of reduction of temperature, and the temperature variation tolerance is thereby consumed. The test for an expired tube, therefore, is a reduction of a cathode temperature by a reduction of current arbitrarily set at 10%. A considerable reduction of current with voltage indicates an expired tube.

A vacuum tube which contains only a cathode and a plate (anode) is known as a "Diode," and is used as a rectifier. Considering current flow from negative to positive (which is opposite to the usual convention, but more logical in discussions of vacuum tubes—see Footnote, Page 443), the current can only flow from cathode to plate, and the tube is therefore a rectifier. It lacks the fidelity of an ideal rectifier as the current is a (3/2) power function of the applied voltage. Improvement in linear response, if desired, is usually accomplished by adding series resistance. This makes the total resistance of the circuit large compared to the variational resistance of the diode, thus reducing the distortion.

2. TRIODE AMPLIFIER TUBE

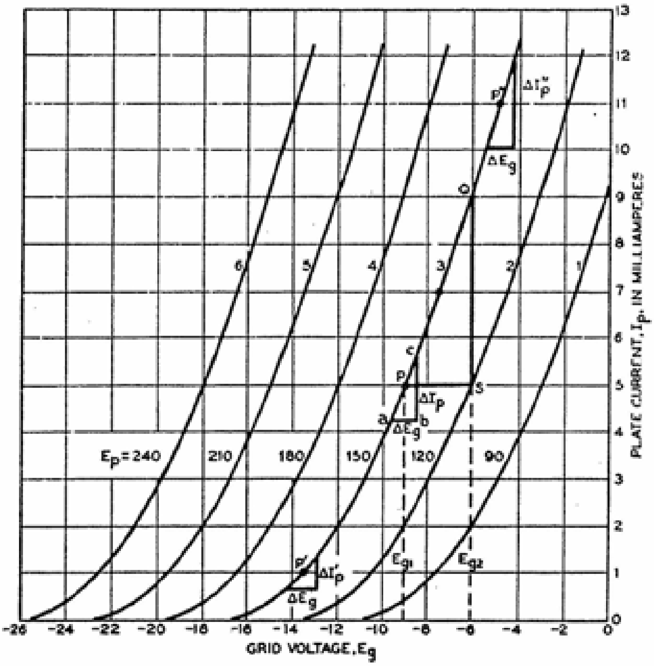
With the addition of a third electrode the tube is called a "Triod. (see Figure 355), and forms the simplest amplifier tube. The third electrode is a grid, usually composed of a spiral wire wound around the cathode in the space between the cathode and the plate, but electrically insulated from both. The relative positions of the three electrodes, together with the relative spacing between the turns of the spiral, have a direct effect upon the voltage amplification factor of the tube.

Both the grid and plate are effective in producing a field potenti.. around the cathode, which affects the current flow. The effect by the grid is greater than that of the plate by the quantity μ, the voltage ampli-

fication factor. With a potential E_g on the grid, the plate current I_P is given by the following expression:

$$I_P = K\ (E_P + \mu\, E_g)^{3/2} \tag{190}$$

Although K and the exponent (3/2) are given as constants, the variation is small in the operating region of plate current, but large when the plate current approaches either zero or the saturation region.



—*From Journal of the S.M.P.E., February, 1935. "Simple Theory of the Three-Electrode Vacuum Tube," H. A. Pidgeon.*

Figure 212 — Plate current-grid voltage characteristics of a triode vacuum tube.

Figure 212 is a family of curves obtained by plotting equation (190), with various values of E_P and E_g. If the constants did not vary as mentioned, the curves would have exactly the same shape.

3. VOLTAGE AMPLIFICATION FACTOR

The relative effectiveness of the grid and plate voltages in changing the plate current is the voltage amplification factor μ. The expression for μ in terms of changes of grid and plate voltage is:

$$\mu = -\frac{\Delta\, E_P}{\Delta\, E_g} \tag{191}$$

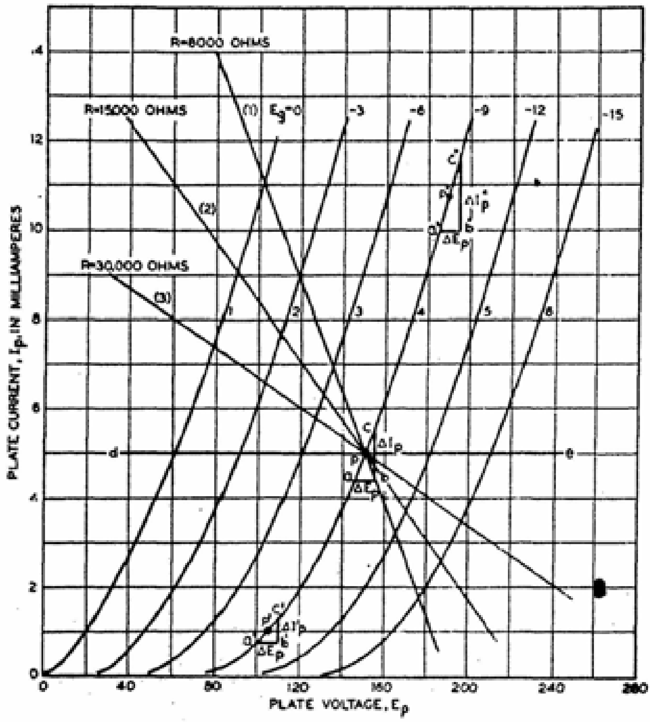
Reference to the point P on the 150 volt curve of Figure 212, gives the plate current as 5 milliamperes, and the grid bias as minus 9 volts. A

decrease of plate voltage to 120 volts offsets the plate current change due to a change of grid bias to minus 6 volts. The μ of the tube is then

$$\mu = -\frac{150-120}{(-9)-(-6)} = -\frac{30}{-3} = 10$$

An important constant of a tube is the dynamic plate resistance or plate impedance, which differs from the static value of plate resistance (the simple ratio of plate voltage and current).

Figure 213 is a plot of plate current versus plate voltage with various values of grid voltages held constant. The point P corresponds to the



—*From Journal of the S.M.P.E., February, 1935. "Simple Theory of the Three-Electrode Vacuum Tube," H. A. Pidgeon.*

Figure 213 — Plate current-plate voltage characteristics of a triode.

operating point P of Figure 213, representing the same values of plate current and voltage and grid voltage.

If the plate voltage is caused to vary by an amount ΔE_p between the two points a and b, the operating point moves along the curve from

a to c and the resulting plate current change, ΔI_P, is indicated by the distance b to c.

The ratio of ΔE_P to ΔI_P is the plate impedance R_P,

$$R_P = \frac{\Delta E_P}{\Delta I_P} \tag{192}$$

that is, the plate impedance at any point P is the reciprocal of the slope of the curve of plate current versus plate voltage at that point. At P, Figure 213,

$$R_P = \frac{10}{1.25 \times 10^{-3}} = 8000 \text{ ohms.}$$

The "figure of merit" of an amplifier tube is the mutual conductance, which is a measure of the rate of change of plate current with variations of grid voltage. Referring again to Figure 212, a change of ΔE_g grid volts, which is indicated by the distance P to S, results in a change in plate current ΔI_P from S to Q.

The mutual conductance is then

$$S_m = \frac{\Delta I_P}{\Delta E_g} = \frac{9-5}{(-9.125)-(-6)} 10^{-3} = \frac{4 \times 10^{-3}}{3.125} = 1250 \text{ microhms}$$

The three parameters μ, R_P, and S_m, are related

$$S_m = \frac{\mu}{R_P} = G_m \tag{193}$$

The mutual conductance value is about the same for all tubes of the same size, so that tubes of high μ also have high plate impedance.

An an amplifier, the vacuum tube must have a load impedance across which a drop in potential results from the flow of plate current. The net plate voltage is the difference between the constant supply voltage and the output signal voltage.

Figure 214 is a repetition of Figure 212 with dynamic characteristics indicated for three values of non-reactive load impedance, 8,000 ohms, 15,000 ohms and 30,000 ohms. These curves are numbered 1, 2 and 3, respectively.

The load line may be defined as the locus of points determining the plate current of the tube as the grid potential varies, with a given load resistance in the plate circuit. The slope of the load line is the "Dynamic Mutual Conductance." It may be determined in terms of other constants of the circuit by the use of the small triangles $PP'b$ and PP_2b in Figure 214. Point P represents a plate potential $E_P = 150$ volts

with grid potential of minus 9 volts, and plate current $I_P = 5$ milliamperes. The load resistance is $R_L = 15{,}000$ ohms; the load voltage,

$$E_L = I_P R_L = 0.005 \times 15{,}000 = 75 \text{ volts}$$

The plate supply potential,

$$E_B = E_P + E_L = 150 + 75 = 225 \text{ volts}$$

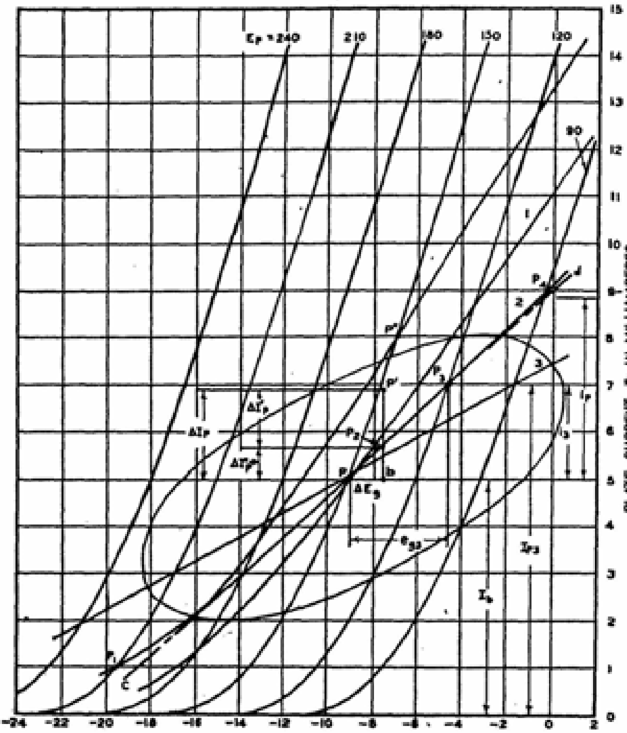


Figure 214 — Dynamic or load characteristics of a triode obtained with load resistances of 8,000, 15,000 and 30,000 ohms.

If the grid potential is changed by $\Delta E_g = +1.5$ to $E_g = -7.5$ volts, and E_P remains at 150 volts, the plate current increase would be ΔI_P to P' at 7 milliamperes, according to the expression for mutual conductance.

$$\Delta I_P = S_m (\Delta E_g) = \frac{\mu}{R_P} (\Delta E_g) \tag{194}$$

However, due to the increased plate current through R_L, the plate voltage would be less than 150 volts by ΔE_P, and the current would be less than the value at P' by ΔI_P, and a new value will be located by P_2.

From the definition of plate resistance, when E_g is constant, the reduction of plate current and voltage are related by

$$\Delta I'_P = \frac{\Delta E_P}{R_P} \qquad (195)$$

The drop in plate voltage would be caused by the increased current through the load resister

$$\Delta E_P = (\Delta I_P'') R_L \qquad (196)$$

where $\Delta I_P''$ is the net change of plate current at the reduced plate potential. From equations (195) and (196)

$$\Delta I'_P = \frac{R_L}{R_P} \Delta I_P'' \qquad (197)$$

Substituting the equivalents of ΔI_P and $\Delta I'_P$

$$\Delta I_P'' = \Delta I_P - \Delta I_P' = \frac{\mu}{R_P} \Delta E_g - \frac{R_L}{R_P} \Delta I_P''$$

or

$$\Delta I_P'' = \frac{\mu}{R_P + R_L} \Delta E_g \qquad (198)$$

The definition of mutual conductance being the rate of change of plate current with variation of grid potential, the dynamic mutual conductance of the tube with load is:

$$S_m = \frac{\Delta I_P''}{\Delta E_g} = \frac{\mu}{R_P + R_L} \qquad (199)$$

This is the equation for the straight line *cPd* which is tangent to the curve at *P*. The departure of the curve from this line represents distortion, or lack of linear relation between the input (E_g) and the output (I_P approximately equal to E_L).

The difference between the straight line and the curve is negligible around point *P*, and at the higher values the ratio of departure to the total magnitude is small, so that the percentage distortion is small. However, at the lower end of the curve, for current values below 2 milliamperes, the absolute and relative values of distortion are large. For this reason, the low limit of the useful range should be established at about 2 milliamperes. The high limit is usually established at the value where E_g equals zero, at which point distortion is introduced as the grid current causes a drop in grid potential. The quiescent point for minimum distortion is set by adjusting the grid bias potential midway between the two limiting values. The maximum undistorted power is that output power which is the result of an input wave with ampli-

tude equal to the grid bias potentials or to the difference between the quiescent point and either limiting value.

Figure 214 illustrates the dynamic characteristic when the load is reactive (as are many loud-speakers). It is obvious that the amplitude of the grid input potential must be less than allowable with a non-reactive load in order not to exceed the limiting values. Thus, for maximum power to be obtained from an amplifier, the power factor of the load should be as close to unity as possible.

4. MAXIMUM OUTPUT POWER

The maximum transference of power is possible when the load impedance equals the plate impedance, in which case one-half of the power is dissipated in each impedance. It is assumed, of course, that the load impedance is non-reactive. When the input potential is the maximum allowable, e_g (max.), the potential at the load terminals is

$$E_L = \frac{\mu R_L}{R_P + R_L} e_g \text{ (max)} \tag{200}$$

The power output is

$$W_0 = \frac{E_L^2}{R_L} = \left[\frac{\mu R_L}{R_P + R_L} e_g \text{ (max)}\right]^2 \frac{1}{R_L}$$

$$= \left[\frac{\mu}{R_P + R_L} e_g \text{ (max)}\right]^2 R_L$$

$$= \text{watts instantaneous power}$$

$$\text{Where } R_L = R_P, \text{ then } W_0 = \frac{\mu^2}{4 R_P} e_g^2 \text{ (max) peak watts.} \tag{201}$$

The average or root-mean-square values of the output will be much less, depending upon the form factor. For this reason, when an amplifier is used to amplify speech or music with complex wave forms, a large margin is allowed between the output level as indicated by an average power reading volume indicator and the rated capacity of the amplifier.

5. SCREEN-GRID TUBES

The addition of another electrode in the vacuum tube, in the form of a screen which completely surrounds the plate, has given this tube the name "Screen-grid tube." The screen-grid is usually operated at a positive potential about half the magnitude of the plate potential. The space current flow is a function of all three potentials.

$$I_P + I_{sg} = K\left(E_g + \frac{E_{sg}}{\mu_{sg}} + \frac{E_P}{\mu_P}\right)^{3/2} \tag{202}$$

where μ_{sg} and μ_P are factors determining the effectiveness of E_{sg} and E_P in controlling the electrostatic field as compared with the effectiveness of E_g. By design, the plate is quite effectively shielded, and the space current is to that extent independent of E_P. Therefore, the expression for space current reduces to

$$I_P + I_{sg} = K\left(E_g + \frac{E_{sg}}{\mu_{sg}}\right)^{3/2} \tag{203}$$

The family of curves in Figure 215 indicates the independence of the plate current with respect to the plate potential. Furthermore, with variation of I_P by some other variable (such as the resistance in the circuit), I_{sg} varies in an opposite and approximately equal amount, keeping the total space current almost constant.

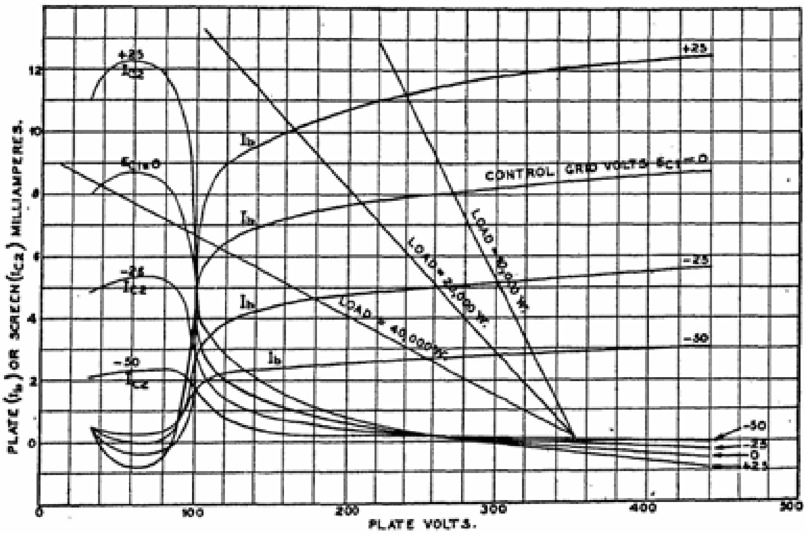


Figure 215 — Average plate characteristics of a screen-grid tube (Screen voltage = 100 volts).

When, due to input potential to the grid, the plate potential is caused to vary to a value less than the screen potential, secondary emission from the plate to the screen causes an actual reversal of plate current and results in the severe kink in the characteristic curves as illustrated in Figure 215. The region below the kink is unsuitable for amplifier purposes. The advantage of the screen-grid tube is much more pronounced in the amplification of radio frequencies, by virtue of the fact that the screen shield around the plate prevents coupling with the

grid with a consequent reduction of input impedance. It is also a good audio-frequency amplifier when used as a voltage amplifier.

The determination of the constants μ, R_P and S_m is practically identical to the method described for the triode. The screen potential is maintained constant when determining the constants involving the control grid and the plate.

Again referring to Figure 215, load lines are indicated for various values of load impedance. The values of grid potential are in equal increments and should intercept equal increments of the length of the load lines in order to represent linear response.

It is apparent that load resistance greater than 20,000 ohms cannot be used without non-linearity where the grid potential lines converge at the kink. The screen-grid tube differs from the triode in that the load impedance is lower than the plate impedance. The screen-grid tube cannot be used efficiently as a power amplifier because of the limited undistorted output obtainable.

The addition of a cathode grid or suppressor grid between the screen grid and the plate prevents the flow of any secondary emission currents

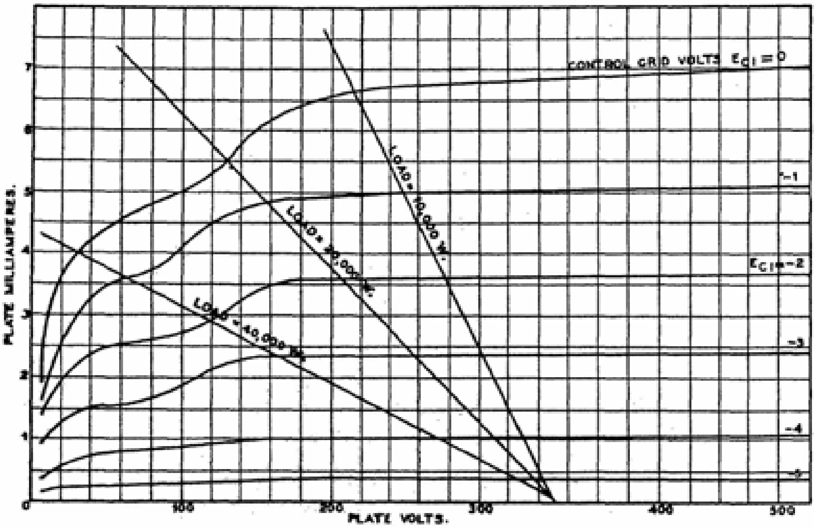


Figure 216 — Average plate characteristics of a pentode tube with suppressor grid.

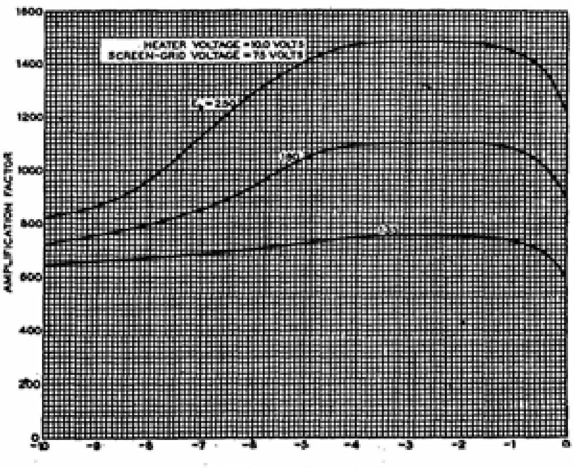
from plate to screen grid. The suppressor grid is usually operated at cathode potential, although for special purposes it may be at some other potential. In some amplifiers employing reverse feedback, the feedback

circuit actuates the suppressor grid and provides control equally as well as the control grid, but independent of it.

Figure 216 illustrates characteristic curves of a pentode, using the suppressor grid. The improvement in overcoming the kink is obvious. Load lines indicate limitations of load impedance, and the variation in lengths of increments of the load lines between equal grid potential values is again an indication of distortion. The determination of μ, R_P and S_m is the same as for the screen-grid tube.

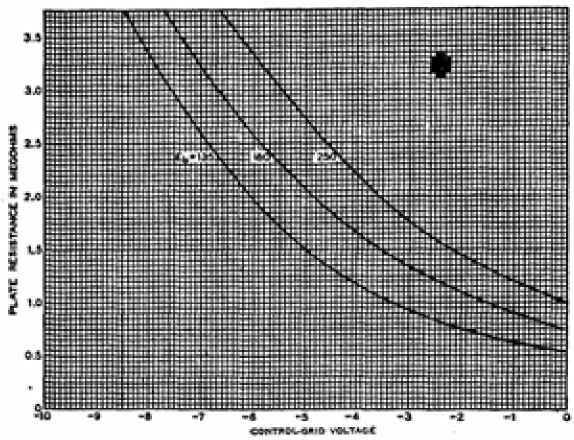
The pentode as a power amplifier has the advantage over the triode of a much greater amplification factor. The distortion is always high, and for that reason the pentode is decidedly inferior unless reverse feedback is used to reduce distortion. The high internal impedance makes it a poor driver for reactive loads, or for Class B amplifiers, where good regulation is necessary.

A modification of the construction of the control grid is used to produce what is known as a variable-mu or remote cut-off tube. The modification consists of winding the grid in a spiral or cylindrical coil with variable spacing between turns. Omitting several turns in the grid also accomplishes the same results to some degree. While this type tube was originally intended for radio receiver use, it has found a valuable place in amplifiers where it is required to vary the gain electrically. Figures 217, 218 and 219 illustrate respectively the effect upon the μ, R_P and S_m of varying the control grid bias potential. Obviously, the amplitude of the input signal potential must be low as compared to



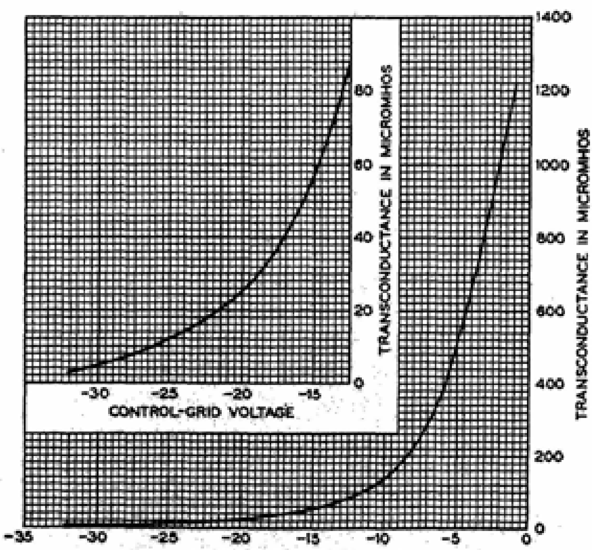
—*Courtesy Electrical Research Products, Inc.*

Figure 217 — Illustration showing the effect upon the amplification factor of variations in control-grid bias potential of a variable-mu suppressor-grid pentode tube.



—*Courtesy Electrical Research Products, Inc*

Figure 218 — Illustration showing the effect upon the plate resistance of variations in control-grid potential of a variable-mu suppressor-grid pentode tube.



—*Courtesy Electrical Research Products, Inc.*

Figure 219 — Illustration showing the effect upon the mutual conductance of variations in the control-grid potential of a variable-mu suppressor-grid pentode tube.

the grid bias: First, to avoid distortion due to the non-linear response curve; and second, to allow a control of gain. Generally, two tubes are used in a push-pull circuit. The input disturbance caused by grid bias potential control is cancelled by the opposing phase relations of the two tubes. Furthermore, even harmonic distortion is cancelled.

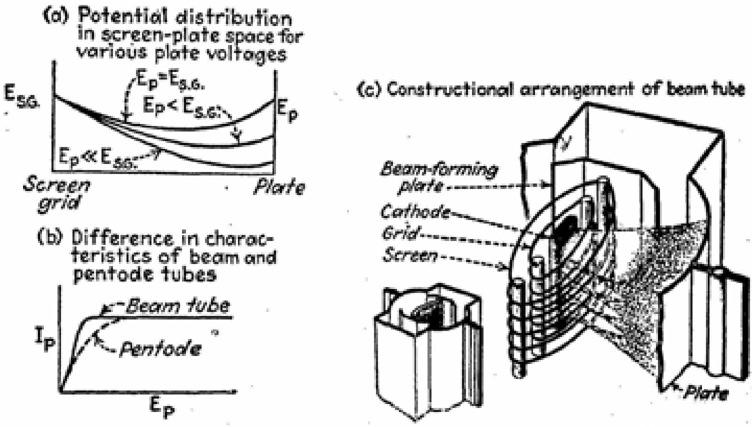
The load impedance is generally a pure resistance of low value as compared with the tube impedance, and the variation of plate impedance causes no variation other than gain, which variation is desired.

6. MISCELLANEOUS TUBES

A tetrode or pentode tube may be connected to serve as a space-charge grid tube. The grid nearest the cathode is the space charge grid, so named because the positive potential neutralizes the field due to the density of electrons in the space. Between this grid and the next, which is the control grid with negative potential, the field charge becomes zero, and a virtual cathode is thus formed. The virtual cathode is relatively close to the cathode, resulting in a tube with a large amplification factor. The plate impedance is also comparatively low so that the voltage amplification may be quite large. The chief fault of such a tube is the large current required by the space-charge grid which absorbs a large portion of the total emission, limiting the current for useful output power.

(a) Beam Power Tube

A modification in the physical design of a screen-grid tube has resulted in a tube for power amplification known as a "Beam Power Tube." The screen-grid wires are in line with wires of the control grid. A set of baffle plates on two sides of the cathode and grid assembly



—From "Radio Engineering," by Frederick E. Terman (McGraw-Hill Book Co., Inc.).

Figure 220 — Constructional details of a beam tube with comparative characteristics.

is electrically connected to the cathode, and into the plate converges the field of electrons, which would otherwise stray. (Figure 220 illustrates the construction).

The screen-grid current is lower than that of ordinary power pentodes by virtue of being in the shadow cast by the control grid. Any

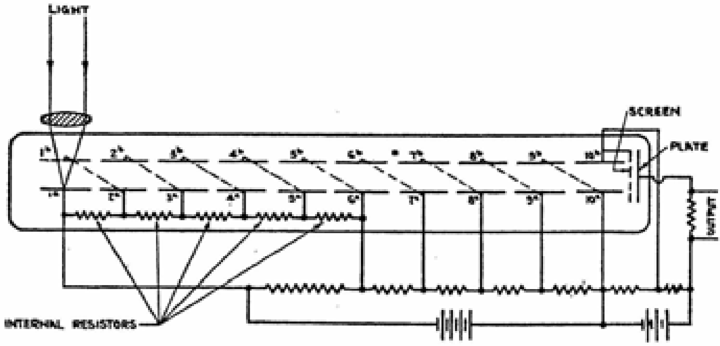
tendency for secondary emission is repelled by the very force of the concentrated primary electrons, and a suppressor grid is not needed.

The results and characteristics are similar to the common power pentode except for the efficiency, which is higher.

(b) Electron Multiplier Tube

The electron multiplier tube is a vacuum tube employing the secondary emission of electrons as the method of amplification. Figure 221 illustrates a schematic plan of the tube construction.

The source of primary electrons may be a hot cathode, but usually is a photo-active surface illuminated by a light source which is the source of energy it is desired to amplify. The primary electrons are attracted to the next electrode by the positive potential. This second



—From the proceedings of the Institute of Radio Engineers, March, 1936, "The Secondary Image Multiplier, a New Electronic Device," V. K. Zworykin, G. A. Morton and L. Malter.

Figure 221 — Constructional details of one type of electron multiplier tube.

electrode is coated with a photo-active surface such as caesium oxide, and while it acts as an anode for the first electrode, it emits secondary electrons in multiplied quantities to the third electrode. This process is repeated for six or more stages, making possible voltage amplification in the order of one million.

The advantages over ordinary tubes are: First, effectively six stages of amplification are accomplished with extreme efficiency all within the one envelope; second, input impedance is extremely high; and third, the frequency range possible far surpasses the ordinary array of cascaded tubes. For television reception, amplification of minute energy at high frequency is made possible by such a tube.

It is necessary to supply a relatively high potential to the electrodes. The polarity of each electrode is positive with respect to the one preceding it, which requires a total voltage equal to the sum of voltages required for each stage. However, no cathode power is required. The noise level is low as the Schott effect is avoided.

Chapter XXII

AMPLIFIER CIRCUITS

By FRED ALBIN

As explained in Chapter XXXIX, the amplifying property of a vacuum tube depends upon the control by the grid (which is actuated by relatively small voltages) of relatively large amounts of power flowing in the anode circuit of the tube. The method of classifying and coupling amplifiers is also explained, and it is the purpose of this chapter to explain these couplings and their effects in greater detail as well as their application to the pentode tube in addition to the triode tube.

1. EQUIVALENT CIRCUIT OF AN AMPLIFIER

When calculating the constants or gain of an amplifier it is not necessary to include the unidirectional voltages and currents present in the circuit due to the various batteries necessary to maintain the correct tube-element potentials. The actual currents and voltages existing in the various circuits during operation of the amplifier are those which are due to the alternating-current components (the result of the alternating signal input), plus the currents and voltages existing in the amplifier at the "no-signal" input—provided the plate resistance and the amplification factor are constant over the variations caused by the signals considered.

Analysis of the performance of an amplifier is usually accomplished by setting up the so-called equivalent circuit of the amplifier. This equivalent circuit gives almost exact results as the plate resistance and amplification factor of the tubes are practically constant. However, for large values of signal voltage the error involved becomes appreciable and the equivalent circuit must be modified.

The basic circuits of triode and pentode tubes acting as amplifiers are shown in Figures 222-A and 222-B, where Z_L is the load impedance; that is, the impedance into which the tube delivers energy.

Figure 223-A gives the equivalent circuit of Figure 222-A. This equivalent circuit considers that the vacuum tube is a generator of constant voltage equal to the actual signal voltage (e_s) impressed, multi-

plied by the amplification factor μ. The resistance between the filament

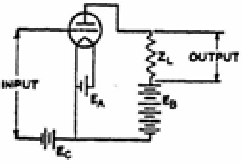


Figure 222-A — Basic triode circuit.

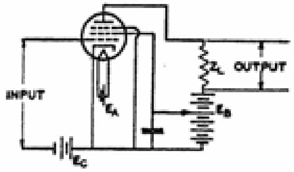


Figure 222-B — Basic pentode circuit.

and plate is given as an external resistance, R_p, and may be considered as the internal resistance of the generator (μe_s).

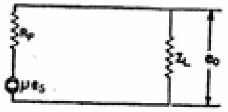


Figure 223-A — Equivalent circuit of Figure 222-A.

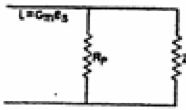


Figure 223-B — Equivalent circuit of Figure 222-B.

From Figure 223-A, and from a consideration of the information and equations given on Pages 512 and 513, it can be seen that the amplification per stage is given by the ratio e_0/e_s; that is, the ratio of the output to the input signal voltages in terms of the tube and circuit constants:

$$\frac{e_0}{e_s} = \mu \frac{Z_L}{Z_L + R_p} \tag{204}$$

This equation is used in the calculation of all amplifier circuits, but it must be remembered that the impedance Z_L is the total load into which the tube is working, while the output voltage, e_0, is the voltage delivered to the grid of the next stage of the amplifier or to the sound reproducing equipment, or to any receiving circuit.

Figure 223-B gives the "constant current" form of equivalent circuit which is more applicable to pentode tube circuits. This circuit considers that a current of constant value equal to

$$i = G_m e_s \tag{205}$$

is delivered to the load impedance. This equation follows from Figure 223-B and the formulae given in Chapter XXXIX. These will be reviewed at this time to make this more clear.

From Chapter XXXIX:

$$i_p = \mu e_s \frac{1}{R_p + Z_L} \tag{206}$$

Multiplying the right side of equation (206) by $\left(\frac{R_p}{R_p}\right)$ gives

$$i_p = \frac{\mu}{R_p} e_s \frac{R_p}{R_p + Z_L} \tag{207}$$

Multiplying equation (207) by Z_L gives

$$i_p Z_L = e_0 = e_s G_m \frac{R_p Z_L}{R_p + Z_L}. \tag{208}$$

This equation will reduce down to the same form as (204) and is equivalent to a generator supplying a current equal to $(G_m e_s)$ flowing from the source through an impedance formed by the plate resistance and the load impedance in parallel. Equation (205) is given in this form as it is a more applicable tool for pentode circuits.

The reason for the various types of coupling for amplifiers is explained in Chapter XXXIX and the fundamental information and circuits given in this reference material should be reviewed before reading the rest of this chapter.

2. THEVENIM'S THEOREM

Thevenim's Theorem states that *any linear network (circuit constants independent of voltage and current), containing one or more sources of voltage and having two terminals, may be replaced at these two terminals by an equivalent generator* E *having an internal impedance* Z_{12}, *where* E *is the internal voltage of, and* Z_{12} *is the impedance across the two terminals of, the original network, with* Z_{12} *measured when all sources of voltage in the network are short-circuited.* Consider Figure 224. By applying this theorem to Figure 224-A, the impedance Z_{12}, with the generator (μe_s) short circuited, is

$$Z_{12} = \frac{R_p R_c R_{GL}}{R_p R_c + R_p R_{GL} + R_c R_{GL}} \tag{209}$$

We then have the circuit of Figure 224-B with Z_{12} as shown and the generator voltage as yet unknown. By making the equivalent circuit generator voltage equal to $(e_s G_m Z_{12})$ the same current will be delivered to Z_L as in the actual circuit where the generator voltage is (μe_s) and the generator resistance is R_{12}. This theorem is given at this time

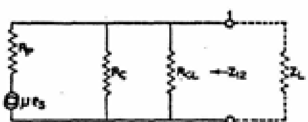


Figure 224-A — Equivalent amplifier circuit.

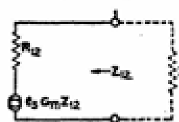


Figure 224-B — Equivalent circuit of amplifier after applying Thevenim's Theorem.

as its use facilitates the calculation of the equivalent circuits of amplifier circuits just as it simplifies many network problems.

3. DIRECT-RESISTANCE COUPLING

The simplest form of amplifier coupling is accomplished by the use of a resistance. Amplifiers employing this coupling are commonly termed direct-current amplifiers. Such a coupling, using neither inductance nor capacitance, introduces no reactance into the circuit and is used where it is necessary to amplify extremely low frequencies. It is not strictly a direct-current amplifier but is so-called because the periodic change in voltage is of such a low frequency that it approaches a direct current which has a frequency of zero. If series reactance were introduced in the circuit at these low frequencies it would seriously affect the amplification.

Figures 225-A and 225-B illustrate typical direct-resistance coupled amplifiers using triode and pentode tubes. Figure 225-C gives the equivalent circuit.

The amplification in the low- and middle-frequency range is given by the equation

$$\frac{e_0}{e_s} = \mu \frac{R_c}{R_c + R_p} \text{ (low - and middle-frequency range)} \qquad (210)$$

which follows directly from equation (204) where $Z_L = R_c$, as the shunting impedance of the capacities of the tube and wiring C_g, com-

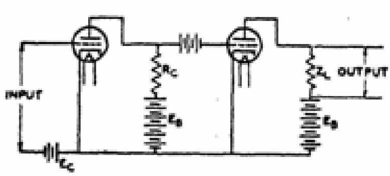


Figure 225-A — Direct-resistance coupled amplifier with triode tubes.

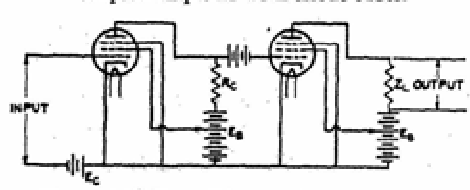


Figure 225-B — Direct-resistance coupled amplifier with pentode tubes.

Figure 225-C — Equivalent circuit of Figure 225-A.

R_p = Plate resistance of tube

R_c = Coupling resistance

C_g = Combined tube and stray wiring capacities.

Figure 225 — Direct-resistance coupled amplifier.

pared to the shunting resistance R_c, is practically an open circuit and may be neglected. In well designed amplifiers this assumption introduces an error of less than one per cent.

From equation (210) it is evident that the amplification at the low and middle frequencies is independent of frequency; that is, all the terms in the formula are constants, provided the plate resistance and the amplification factor remain constant over the variations in signal considered. The amplification is maximum throughout this range.

In the high-frequency range; the shunting effects of the capacities of the tube and stray wiring become appreciable and must be considered. From Figure 225-C and Thevenim's Theorem, the high-frequency amplification is

$$\frac{e_0}{e_s} = \frac{\mu}{R_p}\,\frac{R_c\,R_p}{R_c+R_p}\,\frac{X_g}{\sqrt{X_g{}^2+\left(\dfrac{R_c\,R_p}{R_c+R_p}\right)^2}} \text{ (high-frequency range)} \tag{211}$$

where $$X_g = \frac{1}{2\pi f\, C_g}$$

This reduces to

$$\frac{e_0}{e_s} = G_m\, R_{eq}\,\frac{1}{\sqrt{1+\left(\dfrac{R_{eq}}{X_g}\right)^2}} \tag{212}$$

where $$G_m = \frac{\mu}{R_p}$$

and $$R_{eq} = \frac{R_c\,R_p}{R_c+R_p}$$

Dividing equation (212) by equation (210) gives the ratio of the amplification at high frequencies to the maximum amplification; that is, the amplification at the low and middle frequencies. This results in a ratio

$$\frac{\text{Amplification at high frequencies}}{\text{Maximum amplification}} = \frac{1}{\sqrt{1+\left(\dfrac{R_{eq}}{X_g}\right)^2}} \tag{213}$$

which shows that the extent to which the amplification falls off at the high frequencies is determined by the ratio of the equivalent resistance of the circuit to the capacitance of the circuit. Plotting equations (210) and (213) with the relative amplification on the vertical scale and the ratio of $\left(\dfrac{X_g}{R_{eq}}\right)$ on the horizontal scale, gives a general gain-frequency characteristic curve as shown in Figure 226 where the frequency

(f_a), at which the capacity of the circuit begins to take effect, depend- upon the capacity of the tube and the arrangement and wiring of th circuit, and where the slope of the curve beyond the frequency f_a de- pends upon the ratio $\left(\frac{R_{eq}}{X_g}\right)$.

It will be noted that the cut-off of this curve is gradual, and so f_a may be considered as shown in the figure, although theoretically it would actually fall at a slightly lower point in the frequency scale.

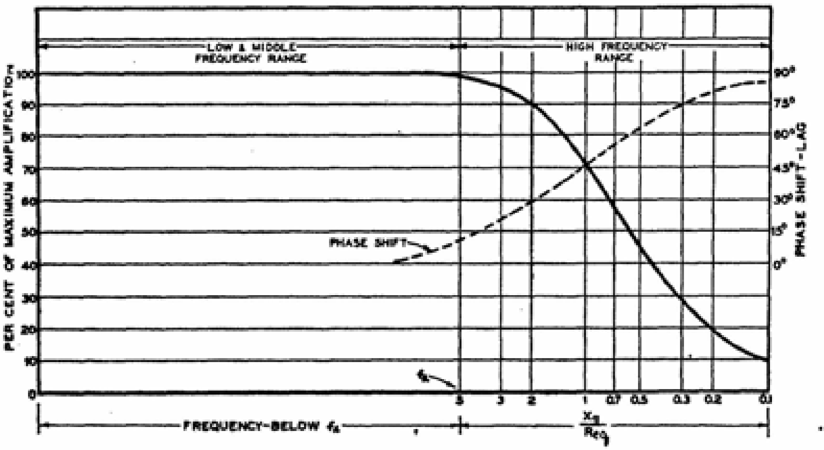


Figure 226 — Gain-frequency characteristic (relative amplification-frequency) curve of a direct-resistance coupled amplifier.

From formula (210) and Figure 226, it is evident that the amplification is constant from zero frequency to that frequency at which the capacity of the circuit becomes appreciable and it is this fact that leads to the use of direct-resistance coupled amplifiers. Their most common use is in noise reduction, photo-electric cell densitometer, and volume indicator circuits. The principal objection to such amplifiers is that they require a greater "B" battery voltage than any other type of coupling. Also, a drift in impedance of any one of the tubes in the amplifier will change the output of the last stage.

4. RESISTANCE-CONDENSER COUPLING

The next step in coupling is to insert a condenser between the grid of two adjacent tubes in the amplifier circuit, as well as a grid leak resistance in parallel with the coupling resistance, the purposes of which are to provide a coupler to deliver a voltage to the next tube and at the same time to prevent the "B" battery voltage of one tube from being applied to the grid of the succeeding tube. Such a coupling is shown in Figure 361 with equivalent circuits given in Figure 362.

The actual and equivalent pentode circuits are given in Figures 227.

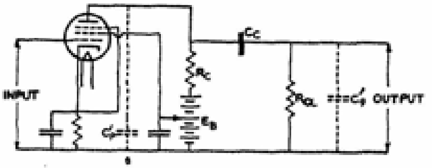


Figure 227-A — Resistance-condenser coupled amplifier with pentode tube.

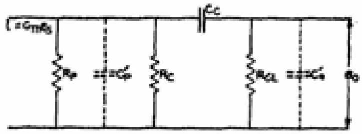


Figure 227-B — Equivalent circuit of Figure 227-A.

In the middle range of frequencies (equivalent circuit Figure 362-D), the effect of the stray capacities and the series condenser are negligible and the amplification is given as

$$\frac{e_0}{e_s} = \mu \frac{R_L}{R_L + R_p} \text{ (middle-frequency range)} \qquad (214)$$

which follows from equation (204) where

$$Z_L = R_L = \frac{R_c R_{GL}}{R_c + R_{GL}}$$

And for the pentode tube, from Figure 227-B, again neglecting the capacitances of the circuit, the amplification in the middle frenquencies is

$$\frac{e_0}{e_s} = G_m \frac{R_L R_p}{R_L + R_p} \text{ (middle-frequency range)} \qquad (215)$$

From equation (214) it is evident that in this particular frequency range the amplification depends upon the amplification factor of the tube and the resistance in the circuit. At other frequencies where the capacities must be taken into effect the relative gain will evidently be less. For this reason the amplification as given by equation (214) is the maximum amplification obtainable from such an amplifier.

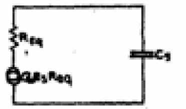


Figure 228-A — Simplified equivalent circuit of Figure 362-A applicable at high frequencies.

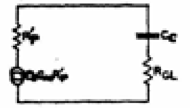


Figure 228-B — Simplified equivalent circuit of Figure 362-A applicable at low frequencies.

In the high-frequency range the shunting effect of the stray capacities must be considered. The equivalent circuit is shown in Figure 362-E. From a consideration of Figure 224 and Thevenim's Theorem, this equivalent circuit may be replaced by the circuit of

Figure 228-A, where $R_{eq} = Z_{12}$ as given in Equation (209). The solution of Figure 228-A gives the high-frequency amplification as

$$\frac{e_0}{e_s} = G_m R_{eq} \frac{1}{\sqrt{1+\left(\frac{R_{eq}}{X_g}\right)^2}} \text{ (high-frequency range)} \quad (216)$$

Dividing Equation (216) by Equation (214) gives the ratio of the amplification in the high-frequency and middle-frequency range:

$$\frac{\text{Amplification at high frequencies}}{\text{Maximum amplification}} = \frac{1}{\sqrt{1+\left(\frac{R_{eq}}{X_g}\right)^2}} \quad (217)$$

where $$X_g = \frac{1}{2\pi f C_g}$$

and $$R_{eq} = \frac{R_p R_c + R_p R_{GL} + R_c R_{GL}}{R_p R_c R_{GL}}$$

This shows that, as in the direct-resistance coupled amplifiers, the amount the amplification falls off at the high frequencies depends upon the ratio of the equivalent resistance to the reactance of the stray capacities of the circuit.

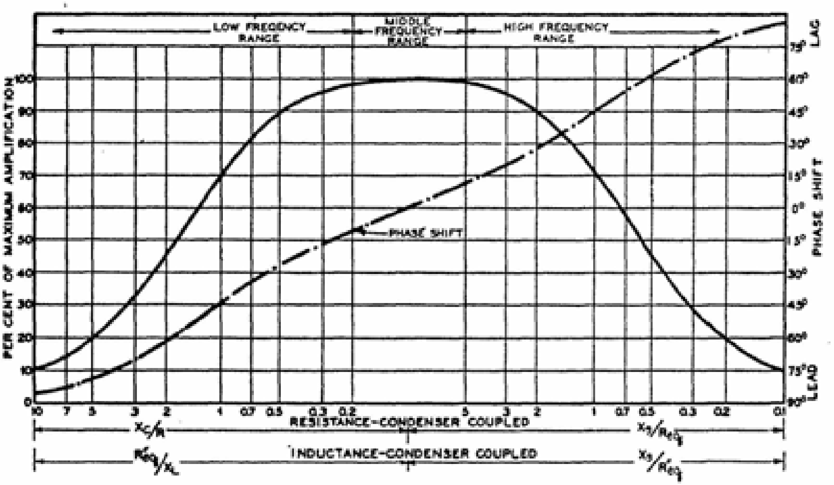
At the low frequencies, the shunting effect of the stray capacities may be neglected, but the series reactance of the coupling condenser, as shown in Figure 362-B, becomes appreciable and must be considered. By the use of Thevenim's Theorem, the equivalent circuit may be replaced by the circuit shown in Figure 228-B and a manipulation of this circuit gives the ratio of the amplifications at the low- and middle-range frequencies as

$$\frac{\text{Amplification at low frequencies}}{\text{Maximum amplification}} = \frac{1}{\sqrt{1+\left(\frac{X_c}{R}\right)^2}} \quad (218)$$

where $$X_c = \frac{1}{2\pi f C_c}$$

and $$R = R_{GL} + R'_p = R_{GL} + \frac{R_p R_c}{R_p + R_c}$$

From a consideration of formulae (214), (217) and (218), the general amplification curve of a resistance-condenser coupled amplifier may be calculated and drawn as shown in Figure 229.



It will be noted that these two amplifier curves, when plotted as general amplification curves, are identical. In application to particular cases they will not be the same, as different values of reactance and equivalent resistance will apply.

Figure 229 — Gain-frequency characteristic curve of resistance-condenser and impedance-condenser coupled amplifiers.

The maximum amplification of both the direct-resistance and resistance-condenser coupled amplifiers is seen to depend directly upon the amplification factor of the tube, as well as the ratio between the resistances in the circuit. As the coupling resistance is increased it would seem that the amplification would also increase. However, an increase of this resistance causes a greater voltage drop in the coupling resistance and makes necessary the use of a higher voltage "B" battery to keep the potential between cathode and anode at a constant value as the coupling resistance is increased. This in turn lowers the efficiency of the tube and counteracts the benefits derived from increasing the coupling resistance. Figure 365 shows this graphically, as it may be seen that the curve showing the relation between amplification and the coupling resistance has a critical value of coupling resistance beyond which there is little change in the amplification. With pentode tubes, an increase in the coupling resistance necessitates a change in the control-grid and screen-grid potentials or a virtual cathode may be formed in front of the suppressor grid and the tube will no longer act in the desired manner.

5. IMPEDANCE-CONDENSER COUPLING

An impedance-condenser coupling employs a shunt inductance coil and a series condenser, as shown in Figure 363, just as a resistance-condenser coupling uses a shunt resistance and a series condenser. The

principal advantage of the inductance type over the resistance type lies in the fact that a lower "B" battery voltage is used as the direct-current

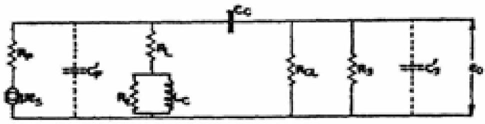


Figure 230-A — Exact equivalent circuit of impedance-condenser coupled amplifier.

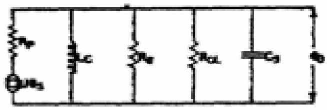


Figure 230-B — Practical equivalent circuit of impedance-condenser coupled amplifier.

voltage drop is lower through the coupling impedance than through the coupling resistance. The initial cost, however, is greater and the coupling impedance must be shielded against pick-up.

Figure 230-A gives the exact equivalent circuit of such an amplifier while Figure 230-B gives the practical equivalent circuit.

The amplification in the middle range of frequencies where L_c, C_c and C_g may be neglected, is

$$\frac{e_0}{e_s} = \mu \frac{R'_{eq}}{R'_{eq} + R_p} \text{ (middle-frequency range)} \qquad (219)$$

where $R'_{eq} = \dfrac{R_e R_{GL}}{R_e + R_{GL}}$

and $\quad R_e$ = a resistance giving equivalent loss to the eddy current loss

from which it may be seen that the amplification in the middle range of frequencies is relatively high and practically constant. The amplification in this range is the maximum amplification, as the shunting reactance of L_c is appreciable at low frequencies, while the shunting effect of the tube and stray wiring capacities cause the amplification to drop off in the high-frequency range.

The amplification at the low frequencies is

$$\frac{e_0}{e_s} = \mu \frac{R'_{eq}}{R'_{eq} + R_p} \frac{1}{\sqrt{1 + \left(\frac{R''_{eq}}{X_L}\right)^2}} \text{ (low-frequency range)} \qquad (220)$$

where $$R''_{eq} = \frac{R_p R_{GL} R_e}{R_p + R_{GL} + R_e}$$

and $$X_L = 2\pi f L_C$$

The amplification in the high-frequency range is given by

$$\frac{e_0}{e_s} = \mu \frac{R'_{eq}}{R'_{eq} + R_p} \frac{1}{\sqrt{1 + \left(\frac{R''_{eq}}{X_g}\right)^2}} \text{ (high-frequency range)} \quad (221)$$

where $$X_g = \frac{1}{2\pi fC_g}$$

Dividing Equations (220) and (221) by (219) gives the ratios between the amplification at low and high frequencies and the maximum amplification:

$$\frac{\text{Amplification at low frequencies}}{\text{Maximum amplification}} = \frac{1}{\sqrt{1 + \left(\frac{R''_{eq}}{X_L}\right)^2}} \quad (222)$$

$$\frac{\text{Amplification at high frequencies}}{\text{Maximum amplification}} = \frac{1}{\sqrt{1 + \left(\frac{R''_{eq}}{X_g}\right)^2}} \quad (223)$$

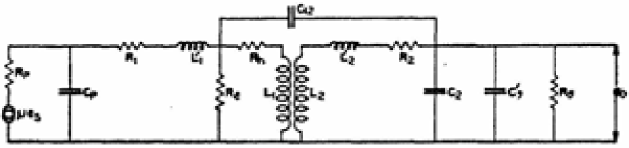
From equations (219), (222), and (223) the general amplification curve of an impedance-condenser coupled amplifier may be plotted as shown in Figure 229. In this case, as with resistance-condenser coupling, the inductance may be increased only to a certain value to cause an added increase in amplification. A comparison of Figures 365 and 366 shows that the amplification-coupling inductance curve is even more critical than the curve of Figure 365, having a much steeper slope in the region of lower values, as well as a sharper cut-off.

In modern amplifiers in which the plate voltage is supplied from a source of alternating-current power through a transformer and rectifier, the advantage of lower "B" battery voltage values is no longer a major consideration and does not offset the larger initial cost, so that resistance coupled amplifiers are usually preferred over the inductance coupled.

6. TRANSFORMER COUPLING

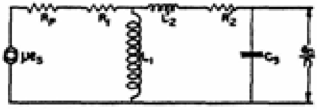
In a transformer coupled amplifier, the coupling consists of a transformer placed between the input source and the grid of the first tube, and between the plate of the first tube and the grid of the next, as shown in Figures 369, 370, and 371. The transformer not only replaces the coupling resistance or inductance, but makes unnecessary the use of a coupling condenser or grid leak resistance.

Figure 231-A gives the exact equivalent circuit of a transformer coupled amplifier and Figure 231-B gives the practical equivalent circuit



R_1 = resistance of primary
R_2 = resistance of secondary
L_1 = inductance of primary
L_2 = inductance of secondary
L'_1 = leakage inductance of primary
L'_2 = leakage inductance of secondary
R_e = eddy current loss
R_h = hysteresis loss
R_P = plate resistance of tube
R_G = input resistance of next tube
C_{12} = capacitance between primary and secondary
C_p = capacitance of tube and distributed capacitances of primary
C'_g = capacitance of next tube
C_2 = distributed capacitance of secondary

Figure 231-A — Exact equivalent circuit of a transformer coupled amplifier.



$$L'_2 = \frac{L_2}{n^2} + L'_1 \qquad R'_2 = \frac{R^2}{n^2} \qquad n = \frac{T_2}{T_1}$$

$$C_g = n^2\,(C_2 + C'_g) + C_{12}\,(n \pm 1)^2$$

T_1 = number of turns on the primary
T_2 = number of turns on the secondary

Figure 231-B — Practical equivalent circuit of a transformer coupled amplifier.

Figure 231-C — Equivalent circuit of transformer coupled amplifier applicable in the low-frequency range.

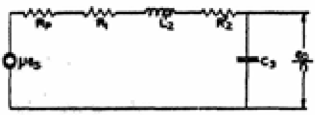


Figure 231-D — Equivalent circuit of transformer coupled amplifier applicable in the high-frequency range.

Figure 231 — Exact and equivalent circuits of a transformer coupled amplifier.

with the transformer ratio reduced to unity.

In the middle-frequency range the amplification is given by

$$\frac{e_0}{e_s} = \mu\,n \text{ (middle-frequency range)} \qquad (224)$$

where $$n = \frac{T_2}{T_1}$$

Formula (224) follows from Figure 231-B, by assuming R_1 L_1, R'_2 and C_g as negligible, which may be done without appreciable error. The inductance of the primary, and the capacitance of C_g, are so high that for all practical purposes these circuits are open. The eddy current loss is relatively low and need not be considered.

In the low-frequency range the reactance of the primary introduces an appreciable loss, and the amplification, where R_1, R'_2 and C_g may still be neglected, is

$$\frac{e_0}{e_s} = \mu\, n \frac{1}{\sqrt{1 + \left(\frac{R_p + R_1}{X_L}\right)^2}} \qquad \text{(low-frequency range) (225)}$$

where $$X_L = 2\pi fL_1$$

From equation (225) it can be seen that the extent to which the amplification falls off at the low frequencies depends upon the ratio of the plate and primary resistances to the inductive reactance of the primary. Although the reactance of the primary, as given in equation (225) is determined from the alternating component of the current through the primary, it must be remembered, in practice, where the core of the transformer is made of magnetic material, that the inductance will vary with the amount of direct current flowing in the circuit.

At the high frequencies the inductance of the primary has even less effect than at the middle frequencies and so may be neglected, but the leakage inductance and the shunting capacity must be taken into consideration. From Figure 231-D, the amplification at the higher frequencies is

$$\frac{e_0}{e_s} = \mu\, n \frac{X_g}{\sqrt{(R_p + R_1 + R'_2)^2 + (X'_2 - X_g)^2}} \qquad (226)$$

(high-frequency range)

where $$X'_2 = 2\pi fL'_2$$

$$X_g = \frac{1}{2\pi fC_g}$$

From a consideration of Figure 231-D it can be seen that at some frequency f_R [see equation (44)], the leakage inductance L'_2 and the shunting capacity C_g are in series resonance, resulting in an appreciable resonant peak in the amplification at this frequency, f_R.

At resonance in a series circuit containing capacitance and inductance, the voltage across the coil and condenser, being nearly 180°

out-of-phase, is much greater than the applied voltage (see Chapter XXXVII). The reactance

$$X_R = j\,2\pi fL = \frac{j}{2\pi fC} \qquad (227)$$

and the current in the circuit, as the capacitive and inductance reactances are equal and opposite (see Figures 343 to 346), is

$$I_R = \frac{E_0}{R} \text{ where } E_0 = \text{applied voltage}$$

Then the reactive voltage is

$$I_R X_L = \frac{E_0}{R}\,j\,\omega L = \frac{E_0}{R}\,\frac{j}{\omega C} \text{ (at resonance)} \qquad (228)$$

The ratio $\left(\frac{X_L}{R}\right)$ or $\left(\frac{X_c}{R}\right)$ of Equation (228) is termed the Q of the circuit, that is,

$$E_L = E_0 Q$$

$$\text{(at resonance)} \qquad (229)$$

$$\text{or } Q = \frac{E_L}{E_0}$$

Equation (226) therefore reduces to

$$\frac{e_0}{e_s} = \mu\,n\,\frac{1}{\sqrt{\frac{1}{Q_R{}^2}}} = \mu\,n\,Q_R \text{ (at resonance)} \qquad (230)$$

where $$Q_R = \frac{X_L}{R_p + R_1 + R'_2} = \frac{X_C}{R_p + R_1 + R'_2}$$

and $Q_R = Q$ of the circuit at resonant frequency, f_R.

Equation (230) shows the effect this resonant point has upon the amplification as, although the amplification in the middle frequencies is (μn), the amplification at frequency f_R, is $(\mu n Q_R)$, or the middle frequency amplification multiplied by the factor Q_R. At frequencies above and below f_R, but still in the high-frequency range, Equation (226) may be expressed in terms of Q_R, and the ratio f/f_R as

$$\frac{e_0}{e_s} = \mu\,n\,\frac{1}{\sqrt{\left(\frac{f}{f_R Q_R}\right)^2 + \left[\left(\frac{f}{f_R}\right)^2 - 1\right]^2}} \qquad (231)$$

Plotting Equations (224), (225), and (231) gives a gain-frequency characteristic curve as shown in Figure 232, for three different values of Q_R. The effect of different values of Q_R upon the high-frequency amplification is shown in Figure 232.

Consideration of Figure 232 shows that for good low-frequency response the inductance of the primary must be high, and the plate resistance of the tube low. In the high-frequency range, the value of Q_R is about 0.75 for linear response to the sloping off point of the curve.

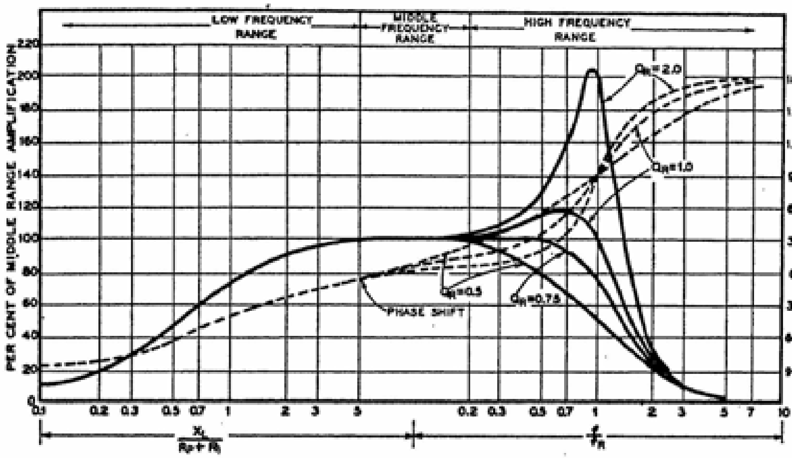


Figure 232 — Gain-frequency characteristic curve of a transformer-coupled amplifier at different values of Q_R.

For a discussion of transformer characteristics and for typical gain-frequency characteristics of transformer coupled amplifiers, see Chapter XIV.

The similarity of Figures 226 and 229 and the lower-frequency range of Figure 232 with the attenuation curves of low- and high-pass filter sections as shown in Charts XXVIII and XXIX, Chapter XIX, should be noted, as the equivalent circuits in particular frequency ranges of these amplifiers are equivalent to low- or high-pass filters depending upon the frequency range considered. Also, the similarity of the amplification curves of a transformer-coupled amplifier as given in Figure 232 should be compared with Figure 197, as the latter shows the relative effect changes in impedance have on the cut-off point of the filter section as well as on the relative amplification in this frequency range to which this amplifier may be compared. This similarity between the gain-frequency curves of these amplifiers and low- and high-pass filter sections, if noted, should lead to a greater familiarity with the formulae, as derived in this chapter, and to a greater flexibility in their use.

7. VOLTAGE AND POWER AMPLIFIERS

In a multi-stage amplifier, that is, one consisting of two or more tubes in cascade, it is a function of all amplification stages except the last to deliver the largest possible voltage to the next stage. The duty of the last stage, called the power amplifier, is to deliver the greatest amount of power to the load.

The voltage amplification of a single-stage amplifier is expressed by the equation:

$$V.A. = \mu \frac{Z_2}{R_P + Z_2} \sqrt{\frac{Z_1}{Z_2}} \qquad (232)$$

Where μ = amplification factor of the tube
R_P = plate impedance
Z_1 = input circuit impedance
Z_2 = plate circuit load impedance

From this expression, it may be seen that the voltage amplification is directly proportional to μ, the voltage amplification factor of the tube. Because of this, tubes with a high μ are used for the first stages of an amplifier. For tubes of equal values of mutual conductance, the ratio of the value of the plate impedance to the value of mu is approximately constant. In general, then, tubes with a high amplification factor have a high internal plate impedance. The loss of high-frequency energy due to capacitive effects in the high impedance circuit is a limiting factor to the gain per stage of amplification, as may be seen from Figures 230-A and 231-A.

In motion picture work, the first stage of an amplifier is often coupled to sources of very low power such as a microphone or a photo-electric cell. If the source has a high internal impedance, the usual method is to couple it directly to the grid circuit of the first stage, using only isolating condensers to avoid conduction of bias and polarizing potentials. Noteworthy examples of high impedance sources are condenser and crystal microphones, and photo-electric cells. In the case of low impedance sources such as dynamic or ribbon microphones or transmission lines, a step-up transformer is used to couple the source to the grid circuit of the first amplifier stage.

Referring again to equation (232), it may be seen that the voltage amplification is proportional to the square root of the impedance of the circuit connected to the grid. The loss of high-frequency energy due to capacitive effects in the high impedance circuit is the limiting factor to the gain accomplished by increasing the impedance. In practice, impedances as high as 200 megohms are used when the coupling

is direct. The practical limit to the impedance in the case of transformer coupling is about one-half megohm. Such a coupling reduces to an equivalent circuit as shown in Figure 231-A where R_p is the impedance looking back into the line and the signal voltage, that is, the voltage to be amplified, is given as (μe_s). This generator circuit is usually of low impedance compared to the grid circuit of the first tube, and consequently a transformer coupling greatly increases the efficiency by stepping up the grid potential by transformer action.

The characteristics of input, interstage, and output transformers are given in Chapter XIV.

The function of an interstage transformer is to couple the plate circuit of one amplification stage to the grid circuit of the next stage. The primary winding is designed to operate with the plate circuit as an input impedance. The secondary winding has the same limitations as the input transformer. When an amplifier stage is to be transformer coupled to the following stage, it is more practical to use a tube with average μ and average plate impedance, than to use a tube with high plate impedance, which would be impractical to match with the primary winding.

In the power stage, where the primary requisite is to deliver the maximum amount of power to the load, tubes with low plate impedances and co-existently low mu are used. The practical limit to this low value of plate impedance, which is in the neighborhood of 800 ohms, is determined by constructional problems.

Transformer coupling of plate circuit to load is generally used. The use of resistance coupling would result in a serious loss of energy. A transformer not only provides a path of low direct-current resistance for the plate circuit, but also acts as an impedance transforming device as explained in Chapter XIV, thus retaining the efficiency of the energy transmission. For maximum energy transference, the transformer has windings with a ratio of impedances of primary to secondary equal to the ratio of impedances of the tube and the load on the amplifier. When this transformer secondary circuit is terminated by the load, the iterative impedance of the primary circuit is a load on the tube which equals the tube impedance, this condition providing for the greatest delivery of power. However, when harmonic generation is considered, it is necessary to provide a load on the tube of at least twice the tube impedance, in which case the ratio of impedances of the transformer windings is equal to the ratio of twice the tube impedance to the load impedance.

8. REGENERATION

In commercial multi-stage amplifiers, for economical and practical reasons, the direct-current supply necessary to maintain the several tube-element potentials is secured from a rectified power source. The fact that this source is common to all stages of the amplifier, leads to the transfer of energy other than signal energy between stages. This is termed regeneration and affects the over-all gain of the amplifier when compared to the ideal amplifier which has separate batteries for each tube. In the latter case, the overall amplification is the total of the individual stages, but in the former case, the amplification usually differs from that of the ideal amplifier.

Regeneration takes place between all stages in the amplifier, due to the difference in energy level in each stage, but the greatest instance of regeneration is between the first and last stages, due to the relatively greater difference in energy levels. This spurious energy may be transferred at various phase relations with the signal voltage, and if in-phase and of sufficient energy, will lead to a building up of the input and to a "singing" or oscillating amplifier. If, however, the spurious energy transferred back to the first stage is out-of-phase with the signal voltage, the effect may lead to an overall improvement in the operating conditions of the amplifier, provided the feedback and signal voltages remain out-of-phase, as explained under feedback. For these reasons, if regeneration is present, its phase in relation to the signal voltage must be under control, or the regeneration must be reduced in magnitude in comparison to the magnitude of the signal voltage.

Although the principal source of magnetic coupling in audio-frequency amplifiers is between the input and output transformers, magnetic coupling of any kind is to be avoided by shielding, proper arrangement of the wiring, and proper relative positioning of the transformers.

As the most common source of regeneration is the inclusion of the same impedance in the plate circuits of two different stages, and as regeneration is more prevalent in transformer coupled amplifiers due to the type of impedance and the larger plate currents for a given plate voltage, a transformer coupled amplifier having a common impedance in the first and power stage will be considered. Figure 233-A shows such a circuit in which the plate currents of each stage flow through the common impedance, Z_c, resulting in a change in the amplification, depending upon the phase relation between the respective currents.

The principal effect of Z_c is to change the amplification of the first stage of the amplifier and this effect becomes important only when the

voltage transfer from the power stage becomes of a magnitude comparable to the signal voltage. The circuit in Figure 233-A may be re-

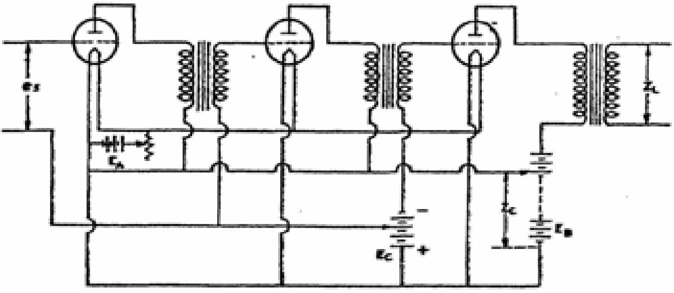


Figure 233-A — Actual circuit of an amplifier with a common impedance in the first and last stages.

placed by the equivalent circuit of Figure 233-B, where the amplification of the intermediate stages is lumped and shown in schematic form.

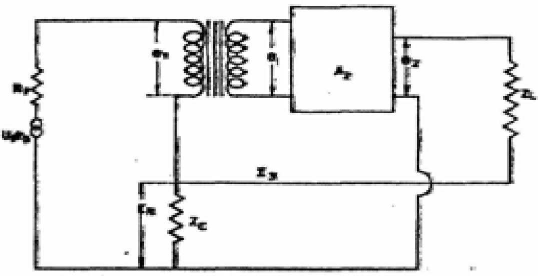


Figure 233-B — Equivalent circuit of Figure 233-A.

The gain of the group stages will be designated as A_2, and the first interstage transformer assumed to have a ratio of unity.

In Figure 233-B, the voltage e_2, acting through the grid of the final stage causes a voltage drop in Z_c by means of the plate current of the last stage. This voltage, E_R, acting in the plate circuit of the first tube, may increase or decrease the signal voltage depending upon the vector relation between the two voltages. The term, "negative regeneration," is usually used when E_R is negative with respect to $(\mu_1 e_s)$. The total voltage acting in series with the first stage plate circuit is

$$E = \mu_1 e_s + E_R \tag{233}$$

where e_s = signal voltage

μ_1 = amplification factor of first tube

E_R = voltage drop in common impedance due to final stage plate current.

The voltage, E, is the new signal voltage after regeneration has taken place. From this,

$$\frac{e_1}{E} = \frac{A_1}{\mu_1} \tag{234}$$

where e_1 = output of first stage
A_1 = amplification of first stage without regeneration.

Then, from the figure, it can be seen that the common impedance voltage,

$$E_R = I_3 Z_c = e_2 K Z_c \tag{235}$$

where K represents a factor determined by the tube and resistance constants of the final stage, which would determine the plate current I_3.

Substituting the value of E_R given in equation (235) into equation (233) gives

$$E = \mu_1 e_s + e_2 K Z_c \tag{236}$$

and solving for E in (234)

$$E = \frac{e_1 \mu_1}{A_1} = \mu_1 e_s + e_2 K Z_c \tag{237}$$

so $$e_1 = \frac{A_1}{\mu_1} (\mu_1 e_s + e_2 K Z_c)$$

but $$e_2 = A_2 e_1$$

$$e_1 = A_1 \left(e_s + \frac{A_2 K Z_c e_1}{\mu_1} \right)$$

$$\frac{e_1}{e_s} = \frac{A_1}{1 - \dfrac{A_1 A_2 K Z_c}{\mu_1}} \tag{238}$$

which shows that the amplification is altered by the factor

$$\frac{1}{1 - \dfrac{A_1 A_2 K Z_c}{\mu_1}}$$

by the presence of regeneration, and is negligible when the term $\left(\dfrac{A_1 A_2 K Z_c}{\mu_1} \right)$ is small compared to unity.

Regeneration is decreased by keeping the value of Z_c small. This is often accomplished by shunting a condenser across the common impedance where the reactance of the condenser is small compared to the by-passed impedance.

Where the plate voltages are secured from a filtered, rectified power source, the common impedance is the reactance of the shunting condenser in the filter of the rectifier. In this case, a filter is used to isolate the common impedance from the input stages. The latter filter consists of a series resistance or inductance and a shunt condenser, the relative values being of such a nature as to make the series impedance high in comparison to the shunt impedance at all frequencies at which the filter is to be effective. One or more of these filter sections may be used, depending upon the amount of filtering necessary, but the result is to introduce less voltage into the input plate circuit by shunting the greater part of the current from the output stage.

NEGATIVE FEEDBACK AMPLIFIERS

Recent developments in amplifier design have led to the deliberate introduction of a certain amount of negative regeneration to improve the overall operating characteristics of the amplifier. The theory of feedback will be considered, and then the effects upon the amplifier characteristics given.

10. FEEDBACK THEORY

The entire amplifier system may be represented in block schematic as shown in Figure 234 where the overall gain of the amplifier is represented by "A." Part of the output is diverted through an attenuator "T" and fed into the input of the amplifier with the signal voltage e_s.

The relation between the input and output voltages is then given by the equation

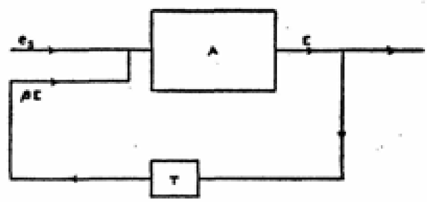


Figure 234.

$$E = A\ (e_s + \beta E) \tag{239}$$

where $\beta =$ that part of the output fed back into the input.

$e_s =$ input signal voltage

$E =$ output voltage

Solving equation (239) for the amplification factor gives

$$\frac{E}{e_s} = \frac{A}{1 - \beta A} \tag{240}$$

where

$A\,\beta =$ feedback factor.

Equation (240) may be restated thus:

$$\frac{E}{e_s} = A\left(\frac{1}{1 - \beta A}\right) \tag{241}$$

which shows that the amplification, as a result of the introduction of feedback, has been modified by the factor $\left(\frac{1}{1 - \beta A}\right)$

Equation (240), by multiplying the numerator and denominator by (βA) and properly arranging terms, may be also expressed as:

$$\frac{E}{e_s} = -\frac{1}{\beta}\left(\frac{1}{1 - \frac{1}{\beta A}}\right) = -\frac{1}{\beta} \tag{242}$$

where (βA) is large compared to 1.

From equation (242) it may be seen that if the feedback factor is large compared to unity, the amplification is practically independent of the amplification of the amplifier itself, but is inversely proportional to the quantity β.

11. STABILITY CONTROL

Although the above theory is relatively simple, practical applications are more involved because of the very special control required of phase shifts in the amplifier and feedback circuits. This control must be exercised not only in the required frequency range but also in a very wide frequency band above and below the useful range. Unless this control is maintained, oscillation or singing will take place, rendering the amplifier unstable.

A and β as given in formula (240) may be either numerical or vectorial terms. As may be seen from a consideration of the amplification curves previously given for the various types of amplifiers, A is almost invariably a vectorial term, as its magnitude varies with frequency. Figure 229 gives the amplification and phase shift curves for feedback amplifiers with resistance-condenser, or resistance-impedance coupling, and Figure 232 gives the amplification and phase shift curves of an amplifier with a transformer coupling. It will be noted that the phase shift is greatest at frequencies where the amplification is low, except in a transformer coupled amplifier, which gives large values of phase shift near the series resonant peak of the amplification curve. As a general rule, an amplifier which does not have a flat frequency characteristic invariably has phase shift, and conversely, an amplifier

with a flat frequency characteristic has very little phase shift. The first part of this rule holds more generally than the second part.

In formula (240), β is numerically equivalent to the phase relation of the output and input, that is, if the gain of the system is to be increased, β is a positive quantity, while if the gain is to be decreased, β is a negative quantity.

As long as the total phase shift of $A\beta$ is less than 180° throughout the total frequency range, from zero to infinity, the feedback voltage remains negative with respect to the signal voltage and the operation of the amplifier will be stable. In practical cases, however, outside of the transmission band, the total shift often exceeds 180°, which is allowable if the gain of the amplifier decreases sufficiently.

12. EFFECT OF NEGATIVE FEEDBACK ON THE PERFORMANCE OF AN AMPLIFIER

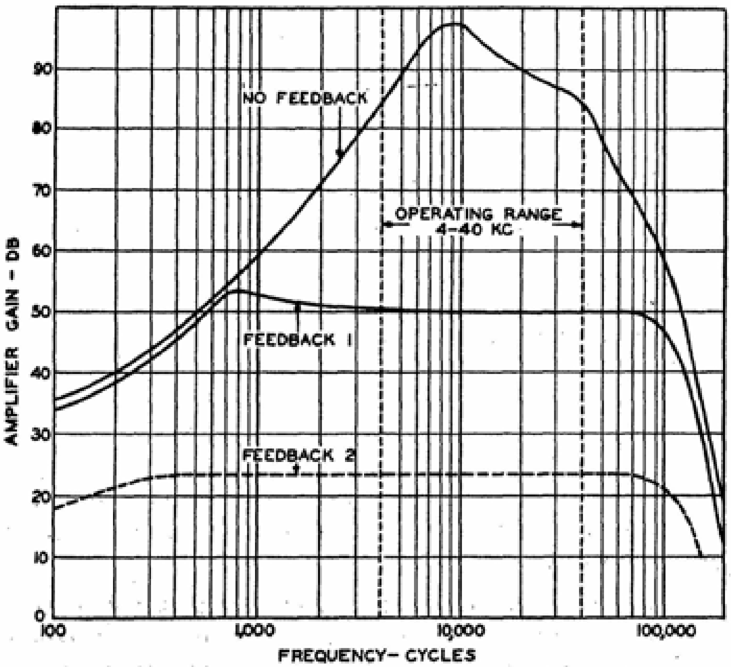


Figure 235 — Gain-frequency characteristics of an amplifier with and without feedback

(a) Effect on Gain-frequency Characteristic

A consideration of equation (241) shows that the gain of an ampli-

fier will be reduced by the factor, $\left(\frac{1}{1-A\beta}\right)$, provided the phase relation of the feedback voltage is proper with respect to the signal voltage. Where relatively large amounts of negative feedback are used, the amplification is reduced still further and becomes independent of the amplification of the amplifier. This may be seen from equation (242), where the amplification depends solely on the amount of feedback present.

This loss in gain is overcome by designing the amplifier with a higher gain than is necessary, and then by feeding back part of the output into the input. Thus, by throwing away the excess gain in a controlled manner, an amplifier with a more constant gain characteristic results.

Figure 235 shows the gain-frequency characteristics of an amplifier without feedback, and with two different amounts of negative feedback. It is to be noted that the gain is very constant throughout the operating range when feedback is employed.

(b) Effect of Negative Feedback on Distortion

In well designed multi-stage amplifiers, the output or power stage overloads first. The preceding stages have a margin of about 10 decibels before overload occurs. For this reason, any distortion present usually takes place in the plate circuit of the last stage, and this distortion, when feedback is used, re-enters the amplifier out-of-phase with the voltage causing the original distortion, thereby tending to cancel out the distortion.

If the symbol "*d*," represents the distortion voltage present in the output voltage *E*, without feedback, then with feedback, the distortion voltage *D* is:

$$D = \frac{d}{1 - A\beta} \tag{243}$$

As the voltage *d* re-enters the amplifier and appears at the output as $A\beta d$, out-of-phase with the voltage causing the distortion, it tends to cancel this distortion.

Equation (243) assumes that the undistorted output is the same with and without feedback. This is approximate and not entirely justified, as the comparison is made between the amount of distortion present with, and the amount present without, feedback. However, practical measurements agree reasonably well with equation (243). Figure 236 illustrates the improvement in loss of distortion in relation to the improvement in gain, and the deviation between the three curves is the amount of error involved.

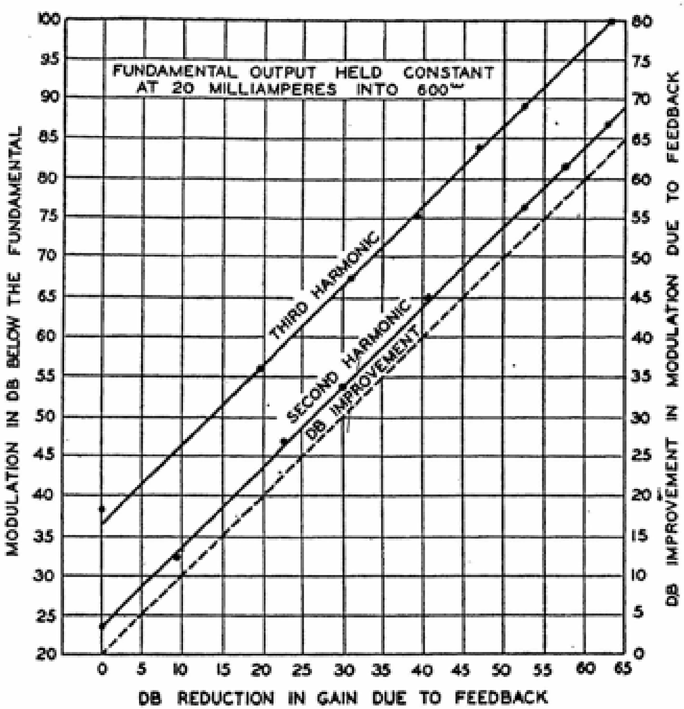


Figure 236 — Typical curves showing the reduction of harmonics with feedback.

(c) Effect on Noise

Noises introduced into the amplifier tend to be cancelled out in the same manner that distortion is reduced, provided the noise appears beyond the grid circuit of the first stage. Any noises appearing in the grid circuit of this stage, such as thermal agitation in the tube circuit, will not be reduced as they appear as part of the signal voltage. The amount the noise is reduced depends upon the place in the amplifier in which the noise is introduced, that is,

$$N = \frac{a}{a_0} \frac{n}{1 - A\beta} \tag{244}$$

where N = noise output voltage with feedback
n = noise output voltage without feedback
a_0 = total amplification
a = amount of amplification received by the noise voltage

From equation (244) it may be seen that the reduction in noise is greatest when the noise is introduced in the power supply circuits of the last stage and decreases when the noise is introduced in the lower level stages.

(d) Effect on Equalization

As the gain-frequency curve of an amplifier in which negative feedback is present depends upon the amount of feedback employed, as well as upon the type of feedback coupler used, such a circuit presents a useful tool where equalization is desired. Desirable equalization can be accomplished by designing the characteristic of the feedback circuit to be the inverse of the gain characteristic desired in the amplifier.

(e) Effect on Phase Shift, Delay, and Delay Distortion

The effect on the phase shift and delay is illustrated in Figure 237, where curves are given showing these characteristics with and without feedback.

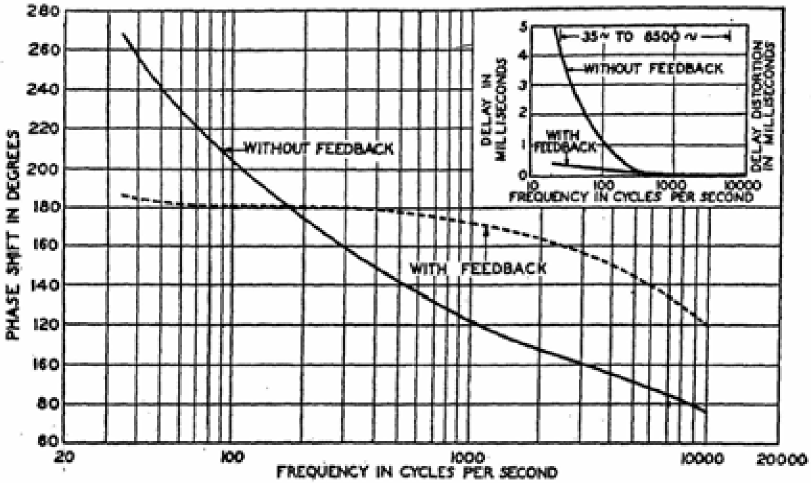


Figure 237 — Phase shift, delay and delay distortion, with and without feedback for a single tube voice frequency amplifier.

In general, where the feedback is intentionally introduced into an amplifier and where the phase of the feedback is properly controlled, the result is an amplifier which has a more constant gain with frequency, a greater freedom from distortion, and greater stability. The improvement depends upon the characteristics of the amplifier as they exist without feedback.

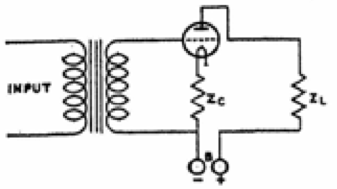
Acknowledgment is made to H. S. Black, and the *Bell System Technical Journal*, for information, nomenclature, and illustrations used in Sections 9 to 12, inclusive.

13. FEEDBACK AMPLIFIERS

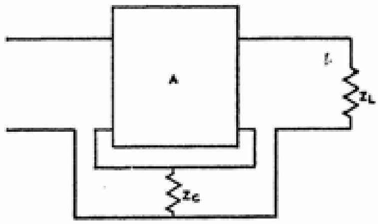
The simplest form of reverse feedback circuit uses a resistor in the cathode circuit which is common to the plate and grid circuits.

Figure 238-A illustrates the circuit of a single stage amplifier with reverse feedback, and Figure 238-B illustrates the equivalent circuit.

Here, the signal voltage drop caused by Z_c, which is included in the plate circuit, is in series with, and opposite to, the input signal.

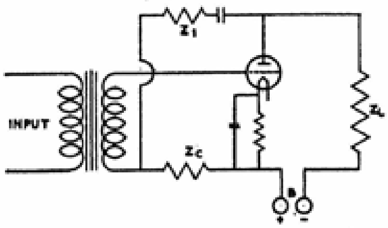


Actual Circuit
Figure 238-A.

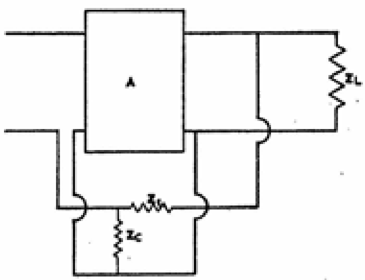


Equivalent Circuit
Figure 238-B.

Feedback voltage obtained from an impedance in series with the load, as in this case, tends to maintain the load current (rather than the load potential), constant with respect to input potential. A feedback circuit, with feedback potential obtained by a shunt, or bridging circuit, tends to maintain the load potential in proportion to the input potentials (see Figures 239-A and 239-B).



Actual Circuit
Figure 239-A.

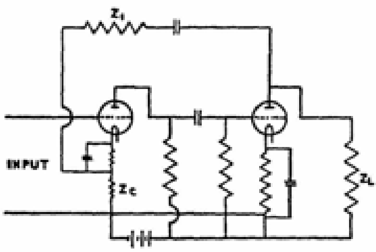


Equivalent Circuit
Figure 239-B.

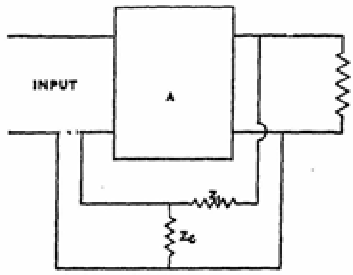
The phase of the feedback circuit is the same for three or any higher odd numbers of stages. For two or other even numbers of stages, the phase is reversed. Figures 240-A and 240-B are schematics of two-stage amplifiers. It may be observed by comparing Figure 240-B with Figure 239-B, that the feedback phase is reversed at the input of the amplifier.

There are as many ramifications of the feedback circuit as there are different amplifier circuits. For simplicity, interstage transformers are avoided. If they are used, they are so designed that the magnetic leakage is kept low to avoid large phase shifts.

Feedback is more effective when the amplifier in the loop has about three stages. The sacrifice of output power in a one-stage amplifier, such as in Figure 239, required to provide the feedback voltage offsets



Actual Circuit
Figure 240-A.



Equivalent Circuit
Figure 240-B.

any merit of the feedback action. However, in a three-stage amplifier, the power absorbed is negligible.

A simple rule in the design of a three-stage amplifier, to allow the application of reverse feedback without encountering oscillation, is to limit the frequency response of the first two stages to the frequency band required. Thus, the attenuation outside of the transmission band increases rapidly, and is sufficiently great to avoid oscillation when the phase-shift reaches 180°. The final stage, however, should have no frequency discrimination within the transmission band. Any loss in the output circuit lowers the available output power, and this loss cannot be recovered by feedback.

The tubes used in multi-stage amplifiers employing feedback are often pentode tubes, taking advantage of the large voltage gain characteristics without concern for harmonic distortion.

It may be observed that the Beta circuit has the configuration of an L network, and the attenuation may be calculated on the basis of any pad, taking due consideration of impedance terminations and accompanying losses.

14. PUSH-PULL AMPLIFIERS

A push-pull amplifier circuit, as illustrated in Figure 241-A has the merit of being free from regeneration through common power supply circuits.

The secondary of the input transformer is divided into two equal halves by a mid-tap, which is connected to the cathodes of both tubes. The opposite ends of the secondary are separately connected to the

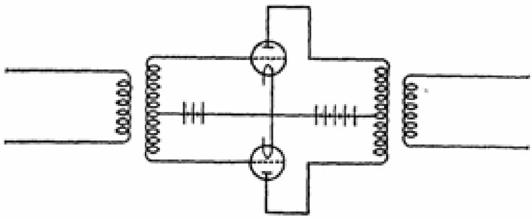


Figure 241-A — Single stage, push-pull amplifier circuit.

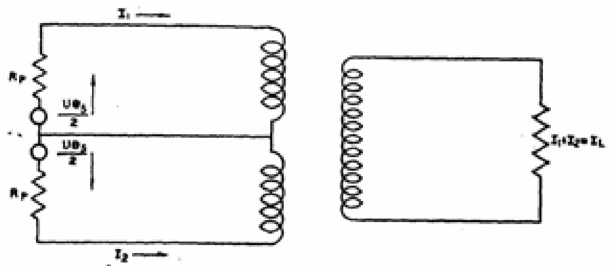


Figure 241-B — Equivalent circuit of Figure 241-A.

grids of the tubes. Thus, an incoming signal entering the primary produces potentials on the two grids which are 180° out-of-phase.

The equivalent circuit is shown in Figure 241-B with the two voltages $\left(\frac{\mu_1 e_s}{2}\right)$ and $\left(\frac{\mu_2 e_s}{2}\right)$, 180° out-of-phase. This leads to a plate current in one half of the circuit which is equal, but opposite in direction, to the current in the other half. Consequently, there is no current flowing in the mid-tap of the primary of the output transformer. The output transformer serves to re-add the output of the two tubes in-phase.

To suppress any tendency for a signal current component to flow in the mid-tapped circuit due to differences in tubes, dissymmetry of the transformers, etc., a series inductance is usually inserted in this circuit. Thus, the circuit may include an impedance common to a preceding stage without introducing regeneration, as no signal current flows in the common impedance.

Another advantage is that the direct-current components of the two plate circuits cancel each other, leaving only the signal component as a

magnetizing force on the core. The consequent reduction of flux density in the core improves the performance of the transformer.

If the two tubes are matched, then the positive half-cycle of the incoming wave receives the same amplification by one tube as the negative half does by the other tube, even under overload conditions. Thus, the symmetry of the wave is preserved, and no even harmonics are generated by the amplifier. Any even harmonics present in the output circuit would cause the output wave to be dissymmetrical, and the harmonic energy would have to flow through the mid-tapped circuit. The series inductance in the circuit is effective in impeding the flow of even harmonic distortion energy. Thus, greater undistorted power may be obtained from a push-pull amplifier than from a half-wave amplifier with the same total tube capacity.

Chapter XXIII

RECTIFIERS

By FRED ALBIN

A rectifier in an electrical circuit is a device which has either non-linear or unilateral conductance, or both, and whose usual function is to produce a unidirectional current from an alternating current. It may also be used to generate harmonics from a pure sine wave. In fact, harmonics are produced only by a non-linear conductance, taken in a broad sense.

1. APPLICATION OF RECTIFIERS

Rectifiers are used in the motion picture industry for many purposes, some of which are:

1. Noise-reduction control circuits
2. Volume indicators
3. Detectors or demodulators
4. Exciter lamp power supply
5. Amplifier tube filament power supply
6. Amplifier anode power supply
7. Horn field power supply
8. Motor speed control circuits
9. Relay and miscellaneous direct-current appliance power supply.

This list includes loads with a wide range of impedance and power requirements.

2. TYPES OF RECTIFIERS

To accommodate these load requirements, many types of rectifiers are used. Economically, certain types are best for a limited range of load impedance and power requirements. The following list of types of rectifiers is roughly in the order of impedance, although there is considerable overlap in the impedance ranges of any two types:

1. High-vacuum hot cathode
2. High-vacuum hot cathode with grid
3. Gas filled hot cathode
4. Gas filled hot cathode with grid
5. Gas filled cold cathode
6. Contact (copper oxide, etc.)

For example, the high-vacuum type is used in circuits with impedance of 500 ohms to greater than 50 megohms, while the contact type is used in circuits ranging from a fraction of an ohm to several thousand ohms. The high-vacuum type is best suited for rectification of voltages in excess of several thousand, while either the gas filled hot cathode type or the contact type is well suited for low voltage and high current, such as exciter lamp power supply.

The use of a grid in the high-vacuum tube provides amplification and rectification simultaneously, the result of virtually combining an amplifier tube with a rectifier. The use of a grid in the gas tube allows the control of a high voltage by a lower control voltage. This circuit is used as a regulated rectifier when, for example, constant output voltage under conditions of varying load current is required.

A description of the construction of a high-vacuum hot cathode rectifier was given in the beginning of the chapter on vacuum tubes. As described there, the tube consists of a cathode surrounded by the anode, and the ensemble enclosed in a highly evacuated envelope with electrical connections brought out through insulated bushings. The cathode may consist of a wire heated directly by a current through it, or it may consist of a metal sleeve enclosing a heating coil, in which case the cathode is heated by thermal conductivity, but with no electrical connection from cathode to heater within the envelope. This tube is commonly known as a "Diode."

Another type of anode may be in two electrically separated halves, or there may be a pair of cathodes and anodes within the one envelope. The name "Duo-diode" or "full-wave" rectifier is used to identify this type.

The area of the cathode emitting surface is directly proportional to the current rating of the anode circuit of the tube. The spacing between the cathode and the anode is directly proportional to the voltage of the circuit being rectified. The internal impedance (cathode to anode) of the tube is also directly dependent upon the spacing of these elements. With these design factors in view, rectifier tubes are made in a large range of sizes and proportions to fill the wide range of load requirements.

The characteristic curve of a typical rectifier is repeated here:

$$I_A = K\,(\epsilon + E_A)^{\frac{3}{2}} \qquad (245)$$

By inspection of Figure 242, it may be seen that the tube fulfills the requirements of a rectifier in that it is both a non-linear and a uni-

lateral conductance. It is strictly unilateral in that no current flows in the reverse direction in any case. It falls short of the requirements of an ideal rectifier in having internal impedance, and being non-linear at both low- and high-current regions. Departure from linearity in the low-current region may be seen by comparison to the dotted curve in the figure.

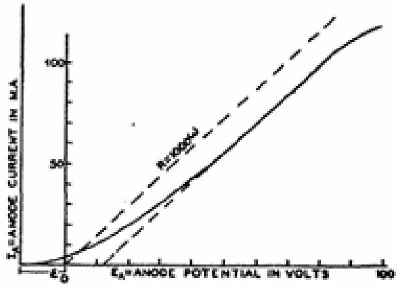
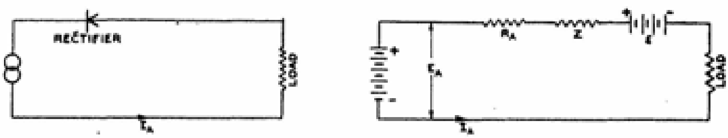


Figure 242 — Characteristic curve of a typical rectifier.

An equivalent circuit for any positive value of anode voltage is illustrated in Figure 243.



Actual Circuit

Equivalent Circuit

E_A = Anode potential
R_A = Characteristic impedance
Z = Variable impedance—function of I_A
ϵ = Emission potential

Figure 243 — Equivalent rectifier circuit for positive anode voltage.

In many applications the inefficiency due to the internal impedance is not a major consideration. When it is, such as in power supply circuits where the loss of power is high, the tube with a low value of impedance is chosen. Usually, the non-linear response is not serious. However, in the case of precision instruments employing rectifiers, and in the case of rectifiers used as detectors, the distortion must be kept at a minimum. This is accomplished by choosing a tube with current capacity in the correct order of magnitude, and with a characteristic impedance which is small compared to the value of load resistance. In this way, the variational resistance is made to be a small portion of the total resistance of the circuit, thereby reducing the proportional variation. The constant resistance values have no effect on the linearity, although the efficiency of the circuit is impaired by mismatch of impedance of the generator and load circuits.

The introduction of gas into the envelope of the rectifier greatly reduces the internal resistance, thus improving the anode efficiency and the voltage regulation, but with a sacrifice of maximum inverse peak anode voltage.

The behavior of the electronic action with presence of gas is somewhat more involved. If the potential exceeds the ionization potential of the gas, electrons are knocked from the neutral gas molecules by collision. The positive ions move slowly toward the cathode, and since they are in approximately the same quantity as negative electrons emitted from the cathode, their presence in the vicinity of the cathode reduces the space charge to nearly zero. The normal flow of negative electrons from cathode to anode and the flow of positive ions in the reverse direction, is caused by a relatively low voltage which is practically constant for all values of current. This voltage is somewhat greater than the ionization potential of the gas. For mercury vapor in equilibrium with mercury, the cathode to anode potential is about 15 volts. The current must be limited by means of an external resistance (so that at no instant does the voltage exceed 22 volts) to protect the cathode from permanent damage by positive ion bombardment.

Another common type of rectifier employs Argon gas. These tubes are commonly used as rectifiers operating at voltages under 100 volts, to supply power to theatre exciter lamps, horn field windings, etc.

3. THREE-ELECTRODE RECTIFIERS

Three-electrode vacuum tubes, or triodes, are often used as rectifiers for volume indicators, vacuum tube voltmeters, detectors, etc. They may be operated either as "grid rectifiers" or "plate rectifiers."

The plate rectifier circuit is similar in operation to that of a diode rectifier preceded by a triode amplifier. In this circuit, the grid and

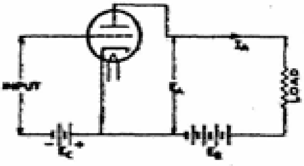


Figure 244-A — Three-electrode rectifier.

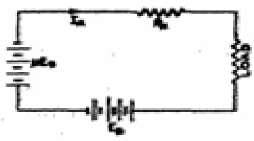


Figure 244-B — Equivalent circuit of three-electrode rectifier.

anode voltages are so chosen that the anode current is practically at the cut-off point. The incoming signal potential is applied to the grid. As in an amplifier, and differing from the diode circuit, the impedance of the load on the generator is infinite. The anode current flows during the positive half cycle only. The current involves the "B" battery voltage and the signal voltage multiplied by the voltage amplification factor of the tube, that is,

$$I_A = K\,(E_A + \epsilon + \mu E_g)^{\frac{3}{2}} \tag{246}$$

Grid rectification depends upon the non-linear conductance of the grid-cathode space. In this case, there is no grid bias potential, and the "B" battery voltage is adjusted so that the anode current is normal with no input signal.

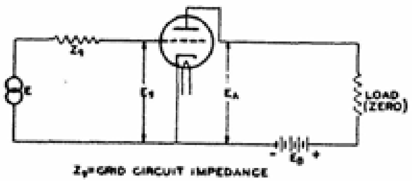


Figure 245 — Circuit of grid rectifier without grid bias potential.

The characteristic grid voltage-grid current curve must be available to determine the parameter of this circuit. (See Figure 246.)

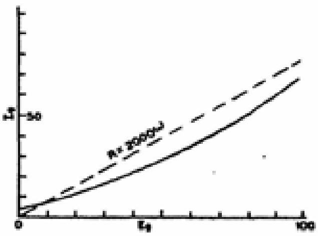


Figure 246 — Grid voltage-grid current characteristic curve.

The slope of this curve at any point is the grid resistance at that point:

$$R_g = \frac{\Delta E_g}{\Delta I_g} \tag{247}$$

The applied grid voltage is the generator voltage less the loss in the grid circuit impedance due to the grid current through it:

$$E_g = E - Z_g I_g \tag{248}$$

During the negative half of the cycle, the grid current is zero and

$$E_g = E \tag{249}$$

During the positive half of the cycle, the grid current is finite, and the grid voltage is reduced to $(E - Z_g I_g)$.

The anode current for the negative half cycle is

$$I_A = K\,(E_A - \epsilon - \mu E)^{\frac{3}{2}} \tag{250}$$

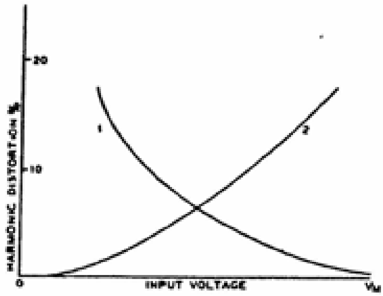
and for the positive half cycle is

$$I_A = K\,[E_A - \epsilon - \mu\,(E - Z_g I_g)]^{\frac{3}{2}} \tag{251}$$

Thus, a non-linear current flows in the anode circuit, which fulfills the requirements of a rectifier.

The grid circuit impedance is usually a capacitance with leakage.

The principal difference in the two methods of rectification is in the distortion versus voltage characteristics as illustrated in Figure 247. The anode rectification plan has large distortion at low amplitudes while the reverse is true for the grid rectification plan.



Curve 1 = Anode rectification.
Curve 2 = Grid rectification.
V_M = Maximum voltage.

Figure 247 — Distortion-input voltage characteristic curve.

The contact rectifier employs the phenomenon that different resistance is presented to the flow of current in opposite directions at the junction of two dissimilar conductors. The commonly used conductors are cuprous oxide on copper. Contact with the oxide is made by a washer of lead tightly pressed against the oxide surface. This element forms a rectifier. Combinations of these may be used to change the voltage or current capacity. A characteristic curve of one element is given in Figure 248.

This device is decidedly non-linear, but, if used as single elements, not unidirectional. When two or more elements are combined in a full-wave form, the back current is cancelled out, and the net current in the forward direction is equal to the large current of one-half of the circuit less the small back current of the other half of the circuit. Full-wave rectifier circuits may consist of the center-tapped or bridge type, which will be described later.

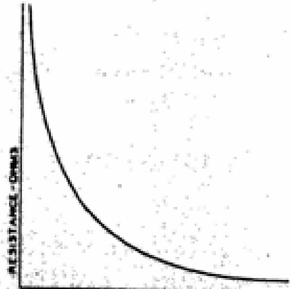


Figure 248 — Characteristic resistance-current curve of a single element contact rectifier.

The large variation of resistance with current density as illustrated in Figure 248 must be minimized if the rectifier is to be used, in conjunction with a meter, for measuring purposes. The common method of improving the operation is by the addition of series and shunt resistances in the rectifier circuit, and will be described later in the discussion of volume indicators.

The resistance-temperature coefficient of copper oxide is large, and as the rectifying properties of the rectifier element depend directly upon the resistance variation, some discrepancy is introduced by temperature variation when the rectifier is used with a measuring instrument. The output of any rectifier with an alternating-current input, is a complex current containing a unidirectional component and an alternating component of a nature depending upon the form of the impressed current and the characteristic of the rectifier. For example, an ideal half-wave rectifier with a sinusoidal input and a resistance load has an output wave with both a direct- and an alternating-current component. The direct-current component has a magnitude equal to one-half the average value of the sine wave. The alternating component is a complex wave with components of frequencies equal to the impressed alternating-current frequency, and with all even and odd harmonic components. The amplitude of the harmonics diminishes as the order of harmonics increases. On the other hand, a full-wave rectifier has a direct-current output equal to the average value of the impressed sine wave. The alternating complex wave contains only odd harmonics, starting with the third, of the impressed wave.

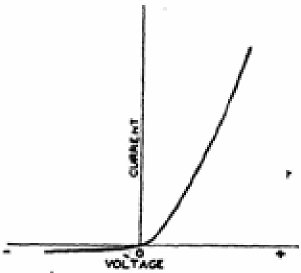


Figure 249 — Voltage-current curve of Figure 248.

If the input wave is complex, the output direct-current component is still equal to the average value of the input wave form, and the alternating component contains frequencies in harmonic relation to the components of the input wave.

In most cases it is desirable to remove the alternating components from the output current, which is done by means of one or more sections of the several types of low-pass filters. The type chosen depends upon the requirements. The chapter on filters describes the characteristics of the several types illustrated in Figure 250.

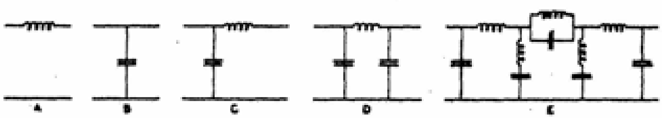


Figure 250 — Several low-pass filter sections.

If the filter is provided with a mid-shunt termination as in (C) and the rectifier impedance is low compared to the load impedance, the potential of the direct-current component increases and approaches the

peak value of the impressed alternating wave, rather than the average value.

Use is made of this action in certain apparatus such as volume indicator circuits, noise-reduction control circuits, etc.

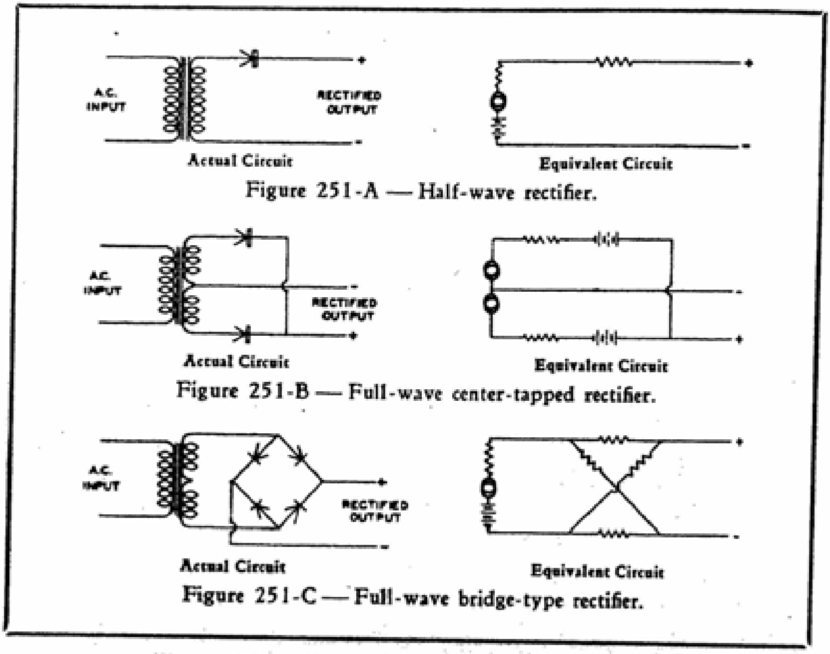


Figure 251 — Actual and equivalent rectifier circuits.

Rectifiers are not ideal from the standpoint of linear conductance. For this reason, the output wave form and even the direct-current component of the output is a complex function of the input wave form, and for use as a measuring instrument, the non-linearity must be reduced to a low value.

Chapter XXIV

VOLUME INDICATORS

By FRED ALBIN

1. TYPES OF VOLUME INDICATORS

In sound recording and sound reproducing systems, the sound intensity (pressure amplitude of the sound wave) is directly proportional to the voltage amplitude in the electrical circuit. Therefore, an indicator which responds to the electrical amplitude is also an indicator of the sound level. Such an instrument is called a volume or power level indicator and is a voltage operated electrical device which has essentially the same function as an ordinary voltmeter used in everyday electrical measurements.

These volume indicators have a scale which is usually calibrated in decibels, which scale is referred to some arbitrary power value called the reference level. The indicator is calibrated in shunt with a definite value of non-reactive impedance and its subsequent use requires match-

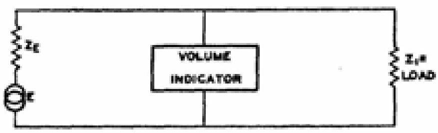


Figure 252 — Method of inserting volume indicator into circuit to be measured.

ing with this identical impedance value or a correction factor must be applied to account for the difference between the impedance used in the calibration and the impedance of the circuit under test. (See Figure 252.)

For example, the reference power level usually used in the motion picture industry is 6 milliwatts (0.006 watts) and the common circuit impedance is 500 ohms.*

*NOTE: In this case the impedance of both the generator and load is 500 ohms, and the impedance of the power level indicator is relatively infinitely high, that is, in the order of 5000 ohms, which results in only a small bridging loss when the indicator is connected to the 500 ohm circuit.

The potential applied to the power level indicator under operating conditions is

$$E = \sqrt{WZ} \qquad (252)$$

where E = potential applied to indicator circuit in volts
W = power in watts
Z = non-reactive impedance of load

If at any instant the power is 6 milliwatts (which is the reference level used) the indicator reads "zero level" and the voltage applied at the terminals of the indicator is

$$E = \sqrt{0.006 \times 500} = \sqrt{3} = 1.73 \text{ volts}$$

All other power level values as expressed in decibels are determined from the formula:

$$\text{Power level in db} = 10 \log_{10} \frac{P_1}{P_0} \text{ (See Page 520)} \qquad (253)$$

where $P_0 = 0.006$ watts
P_1 = power level reading

The correction factor to be applied when the load impedance differs from the calibration impedance, is given as

$$\text{Correction factor in db} = 10 \log_{10} \frac{Z_0}{Z_1} \qquad (254)$$

where Z_0 = calibration impedance (500 ohms)
Z_1 = load impedance

This correction factor should be added to the reading of the indicator.

The nature of the speech energy to be measured by the volume indicator is very complex as it contains frequency components ranging from 40 to 10,000 cycles per second as well as transient characteristics with large decrements.

The requirements of the volume indicator greatly exceed the usual requirements of a voltmeter in that the response must be uniform over this wide frequency range and the motion must be damped so that the indicator responds to the rapidly changing amplitude of the wave. For

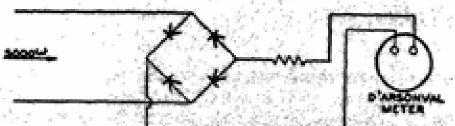


Figure 253 — Volume indicator with copper oxide rectifier.

these reasons, the large majority of volume indicators use a movement employing the d'Arsonval principle, which movement is preceded by a rectifier which converts the alternating speech current to a unidirectional current. (See Figure 253.) Such a combination can have a full scale sensitivity of 3.46 volts and a resistance of 5,000 ohms without using vacuum tubes in the circuit. When either a vacuum tube rectifier of the

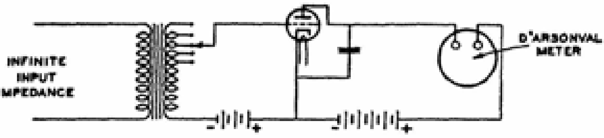


Figure 254 — Volume indicator with grid rectifier.

three-electrode type or a pre-amplifier is used the sensitivity may be substantially increased and at the same time the bridging impedance increased to practically infinity. (See Figure 254.)

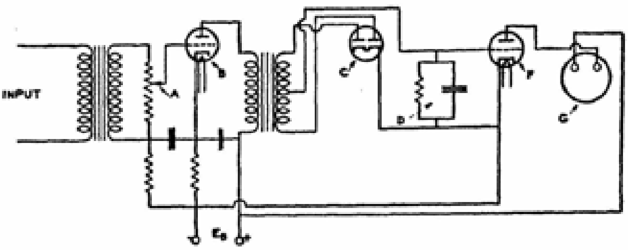
The assumption of an ideal rectifier and uniform current scale meter as an indicator, gives readings which are proportional to the average value of the complex wave. The "work values" of the wave, however, are the root-mean-square (r.m.s.) values. The energy of sound in air is also proportional to the r.m.s. value. However, the energy which overloads amplifiers, modulators, etc., is the peak or maximum value of the wave. It is therefore this value in which the sound engineer is interested when controlling overload.

The relations between the three values of the wave, that is, the average, the peak, and the r.m.s. values, are dependent upon the component frequencies and their phase relations. Studio experience has taught the sound engineer to allow for the large difference between the r.m.s. value, as read by the indicator, and the probable peak value, by providing a safety margin.

In sound recording, in order to maintain a high ratio between signal and noise level, it is imperative to record the signal level as high as permissible without risking modulator overload. To aid the sound mixer (who controls the amount of modulation recorded) in determining the peak value of the signal, a "peak-reading" volume indicator is used. In this type indicator a rectifier with a capacitive reactance load, as described in the chapter on rectifiers, is employed. The response is approximately proportional to the peak value of the speech wave. Figure 255 shows a simplified circuit of such a power level indicator.

A cathode ray oscilloscope, another type of volume indicator, has the advantage of indicating peak values of any complex wave with

components which are in excess of 10,000 c.p.s. It has the unique merit of being practically without inertia for these frequencies, the greatest



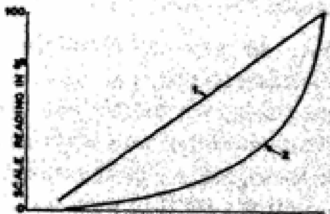
A = Sensitivity control
B = Linear speech frequency amplifier
C = Full-wave rectifier
D = Timing circuit
E = Linear direct-current amplifier
F = Indicating meter—reverse direction

Figure 255 — Circuit of peak reading volume indicator with linear response.

limiting factor to the speed of its response being the retentivity of the observer's eye. The amplitude of the displacement of the sweeping beam (oscillogram) is proportional to the voltage output of the testing circuit to which the oscilloscope deflecting plates are connected.

2. SCALE LENGTH OF VOLUME INDICATORS

Sound recording levels cover a range of about 30 decibels, which represents an amplitude ratio of approximately 33 to 1. On a linear scale the low level would cause an indication of only 3% of that caused by the modulation of the 30 db above this low level. As it is only the high-level values in which we are usually interested, that is, in levels between 90% and 100% of the highest level, indicators for low levels are relatively unimportant as far as studio practice is concerned. As a



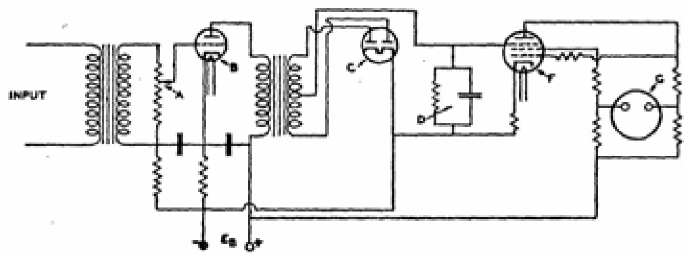
Curve 1 = linear decibel scale
Curve 2 = linear amplitude scale

Figure 256 — Response curves of two types of indicators showing the greater sensitivity at low input signal levels for logarithmic response.

consequence, a scale which is much more desirable than the linear scale is one on which equal increments represent uniform decibel steps of level; that is, the response of the indicator should be proportional to the log of the voltage applied to its terminals. The range of such an indicator may be increased many times over one employing the linear scale. (See Figure 256.)

One method of providing this exponential response is to shape the pole pieces of the galvanometer movement in such a way that an increase of current in a logarithmic manner causes a linear deflection. This, however, appreciably decreases the damping, especially at large scale readings.

Another method is in the addition of an amplifier having an exponential response. Such an amplifier must have an output which is



A = Sensitivity control
B = Linear speech-frequency amplifier
C = Full-wave rectifier
D = Timing circuit
F = Direct-current amplifier with exponential response
G = Indicating meter.

Figure 257 — Circuit of peak reading volume indicator with linear decibel scale.

proportional to the log of the input voltage over the desired range in level. Figure 257 illustrates the position of such an exponential amplifier in the volume level indicator circuit.

Another type of volume indicator in use, which responds to the peak values of complex waves over a very wide difference of level, consists of a horizontal row of 15 Neon lamps arranged adjacent to a suitable scale. Each lamp is associated with an individual amplifier and attenuator, with all amplifiers connected to the common input circuit. The overall sensitivity of each lamp and circuit diminishes in 3 db steps from one end of the scale to the other, so that the lamp on one end glows at the lowest level, while the lamp on the other end requires a level of 45 db above the first level before it glows. Thus the progress of the illumination from the low level side of the scale is an indication

of the power level. The fact that the Neon lamps flash over, or strike, at a certain peak potential makes this instrument a peak responding device.

Most volume indicator circuits are equipped with calibrated attenuators to extend their useful range. The bridging loss over all values of attenuation level and frequency must be constant to avoid introducing frequency or amplitude distortion in the circuit under test.

In general, all types of indicators which employ rectifiers, etc., and which consequently depart from the ideal, are calibrated empirically and so have a scale which automatically eliminates errors in reading due to distortion characteristics of any equipment in the circuit.

PART II

Chapter XXV

ELEMENTARY CONSIDERATIONS

By A. P. HILL

1. CONSTITUTION OF MATTER

Matter may be classified under two headings: (a) Elements; and (b) Compounds.

A compound is composed of two or more elements; an example being water, which is composed of hydrogen and oxygen. An element, on the other hand, does not consist of any matter as ordinarily interpreted other than that which it represents.

If an analysis is made of a compound such as water, it shows that the smallest particle of water that can exist as such is composed of two minute particles of hydrogen and one of oxygen. These three particles combine to form what is known as one molecule of water, a molecule being the name given to the smallest particle of a compound that can exist as such.

An analysis of an element such as hydrogen shows that there is no material other than hydrogen in its makeup; and the smallest particle of an element that can enter into any chemical combination is called an atom. The chemical symbol for water — H_2O — indicates that one molecule of water consists of two atoms of hydrogen and one of oxygen. Sulphuric acid — the chemical symbol for which is H_2SO_4 — indicates that one molecule of this substance is composed of two atoms of hydrogen, one of sulphur, and four of oxygen.

At one time the atom was considered to be the ultimate particle of matter. In fact, the word "atom" means something that cannot be cut. In more recent years, however, it has been discovered that the atom itself is a complex structure and its construction has been likened to that of our solar system which has the sun as its nucleus, and surrounding it various planets, each revolving around the sun in its own orbit. Similarly, an atom consists of a nucleus and minute particles surrounding it; these particles have been named "electrons." All atoms

consist of similar particles each having a nucleus and varying numbers of electrons surrounding it.

Thus it can be shown that all matter consists fundamentally of the same elemental particles, and just as with a quantity of bricks, a simple or complex structure can be erected, so with the same elemental particles, simple and complex elements exist, some necessitating the use of a minimum number of particles, others being much more complex. An example of the latter is a molecule of cane sugar which is composed of 12 atoms of carbon, 22 of hydrogen, and 11 of oxygen.

2. ATOMIC WEIGHTS

Table II is a list of all known elements arranged in order of their respective weights. Included in the list is their chemical symbol and the atomic weight of each which expresses how many times heavier each atom is than the atom of the lightest substance, i.e., of hydrogen. As a matter of convenience, chemistry does not refer the atomic weights to hydrogen directly, but rather to the 16th part of the weight of an atom of oxygen. The discrepancy occurring by this method is negligible. It will be noted that there are two blank spaces in this list of elements, namely 85 and 87. Elements which will eventually occupy these two places have not yet been discovered. Nevertheless, many of their properties are already known.

Several years ago the second place in this list of elements was vacant. Finally its presence was discovered, by means of spectrum analysis, in the sun. It was consequently named after the sun — "helium" — and later was discovered on the earth and is now used, as is well known, in the inflation of dirigibles.

3. SIZES OF ATOMS AND ELECTRONS

As described above, an atom of any of the 92 elements consists of a nucleus with one or more electrons in orbital motion around it.

So far as is known, the nucleus consists of both positive and negative electricity*, and normally in any atom the amounts of positive and negative electricity are equal, thus resulting in an atom that has no definite electrical charge.

---

*NOTE: Because of the purpose for which this paper is intended, no mention is made of the presence in the nuclei of deutrons, positive electrons, etc. This information can be obtained from the *Bell System Technical Journal* for July, 1933, and succeeding numbers, under "*Contemporary Advances in Physics — The Nucleus,*" by Karl K. Darrow.

Remembering that a compound consists of two or more molecules, each molecule consisting of two or more atoms, and each atom consisting of a nucleus with one or more electrons in motion around it, it may be of interest to state the relative dimensions of some of these particles. The dimensions of an atom of hydrogen (the lightest known element) are:

Diameter	$= 1.1 \times 10^{-8}$* centimeters
Mass	$= 1.66 \times 10^{-24}$ grams

The dimensions of an electron are:

Diameter	$= 3.2 \times 10^{-13}$ centimeters
Mass	$= 8.9 \times 10^{-28}$ grams
Electrical charge	$= 1.59 \times 10^{-19}$ coulombs

The above dimensions give very little practical idea of the minuteness of the particles with which we are dealing. Various descriptions have from time to time been given to illustrate these microscopic values, one of which is that if a baseball were magnified to approximately the size of the earth, an electron magnified the same number of times would be about the size of the head of a small pin.

TABLE II

The Natural Sequence of the Elements

ATOMIC NUMBER	SYMBOL	NAME	ATOMIC WEIGHT
1	H	Hydrogen	1.0077
2	He	Helium	4.00
3	Li	Lithium	6.939
4	Be	Beryllium	9.02
5	B	Boron	10.82
6	C	Carbon	12.000
7	N	Nitrogen	14.008
8	O	Oxygen	16.000
9	F	Fluorine	19.00
10	Ne	Neon	20.2
11	Na	Sodium	22.997
12	Mg	Magnesium	24.32
13	Al	Aluminum	26.96
14	Si	Silicon	28.06
15	P	Phosphorus	31.024

* NOTE: $10^{-8} = \frac{1}{100,000,000} = 0.000,000,01.$

TABLE II (Continued)

ATOMIC NUMBER	SYMBOL	NAME	ATOMIC WEIGHT
16	S	Sulphur	32.065
17	Cl	Chlorine	35.458
18	A	Argon	39.91
19	K	Potassium	39.095
20	Ca	Calcium	40.07
21	Sc	Scandium	45.10
22	Ti	Titanium	47.9
23	V	Vanadium	50.96
24	Cr	Chromium	52.01
25	Mn	Manganese	54.93
26	Fe	Iron	55.84
27	Co	Cobalt	58.97
28	Ni	Nickel	58.69
29	Cu	Copper	63.57
30	Zn	Zinc	65.38
31	Ga	Gallium	69.72
32	Ge	Germanium	72.38
33	As	Arsenic	74.96
34	Se	Selenium	79.2
35	Br	Bromine	79.916
36	Kr	Krypton	82.9
37	Rb	Rubidium	85.44
38	Sr	Strontium	87.62
39	Y	Yttrium	89.0
40	Zr	Zirconium	91.0
41	Nb	Niobium	93.1
42	Mo	Molybdenum	96.0
43	Ma	Masurium	
44	Ru	Ruthenium	101.7
45	Rh	Rhodium	102.91
46	Pd	Palladium	106.7
47	Ag	Silver	107.880
48	Cd	Cadmium	112.41
49	In	Indium	114.8
50	Sn	Tin	118.70
51	Sb	Antimony	121.77
52	Te	Tellurium	127.5
53	I	Iodine	126.932
54	X	Xenon	130.2

TABLE II (Continued)

ATOMIC NUMBER	SYMBOL	NAME	ATOMIC WEIGHT
55	Cs	Caesium	132.81
56	Ba	Barium	137.37
57	La	Lanthanum	138.91
58	Ce	Cerium	140.25
59	Pr	Praseodymium	140.92
60	Nd	Neodymium	144.27
61	Il	Illinium	
62	Sa	Samarium	150.43
63	Eu	Europium	152.0
64	Gd	Gadolinium	157.26
65	Tb	Terbium	159.2
66	Dy	Dysprosium	162.52
67	Ho	Holmium	163.4
68	Er	Erbium	167.7
69	Tm	Thulium	169.4
70	Yb	Ytterbium	173.6
71	Lu	Lutecium	175.0
72	Hf	Hafnium	178.6
73	Ta	Tantalum	181.5
74	W	Tungsten	184.0
75	Re	Rhenium	
76	Os	Osmium	190.8
77	Ir	Iridium	193.1
78	Pt	Platinum	195.23
79	Au	Gold	197.2
80	Hg	Mercury	200.61
81	Tl	Thallium	204.4
82	Pb	Lead	207.20
83	Bi	Bismuth	209.00
84	Po	Polonium	(210)
85			
86	Rn	Radon (Ra emanation)	222
87			
88	Ra	Radium	225.95
89	Ac	Actinium	(226)
90	Th	Thorium	232.15
91	Pa	Protoactinium	(230)
92	U	Uranium	238.17

Chapter XXVI

STATIC ELECTRICITY

By A. P. HILL

Electricity, as ordinarily understood, may be described under two headings: (a) Static, and (b) Dynamic electricity. Before discussing these it will be advisable, however, to give some consideration to the media in which electrical charges move.

1. CONDUCTORS AND INSULATORS

All materials may be classified as either conductors or insulators. There is, however, no fundamental difference between the conductor and the insulator, the latter describing merely a poor conductor.

From our knowledge of the electron described in the previous chapter, we see that each electron has a definite electrical charge, namely: 1.59×10^{-19} coulombs. A number of years ago, before it was definitely known what constituted an electrical current, it was assumed that it consisted of a flow of some kind from the positive to the negative pole of the generating source. This is now known to be incorrect, a current of electricity consisting simply of a drift of electrons along a conductor from the negative to the positive pole of the generator. In all text-books written up to the present time, the old classical theory describing a current flow as taking place from the positive to the negative pole of the source of supply is used. Throughout the following discussions (i.e., the remainder of this book), however, a current flow of electricity will always be considered as being a movement of electrons from the negative to the positive pole of the supply source, in the external circuit.

In some substances the nucleus of the atom exerts a strong retentive force on the electrons surrounding it. It is consequently difficult to tear these electrons loose from one atom and pass them on to the next in line in a material. If this condition exists, the material is called an insulator. If, on the other hand, the retentive force of the nucleus on the electron is relatively weak it is comparatively easy to move these electrons from one atom to the other in a direction from the negative to the positive pole of the generator. Such a substance is called a conductor. Silver, copper and certain other metals are examples of

good conductors. Glass, ebonite, etc., are poor conductors, and therefore, good insulators. Table III, at the end of this chapter, gives a list of some of the better known conductors and insulators in the order of their conducting ability. Consequently, at the head of the list will be found those substances that are usually considered to be conductors, and at the bottom those that are usually considered to be insulators.

2. SOME ELECTRICAL UNITS

An analogy for the flow of electrical current is that of the flow of water through a pipe. Figure 258 shows such a system.

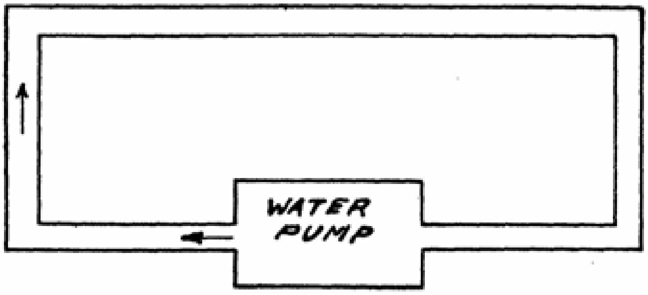


Figure 258 — Mechanical analogy for current flow.

The flow of water taking place is usually said to consist of a certain number of gallons per minute. In the electrical case we have a corresponding unit of quantity, the name of which is the "coulomb".

Since, as previously mentioned, the charge of an electron is 1.59×10^{-19} coulombs, a coulomb of electricity consists merely of 6.3×10^{18} electrons. In a similar manner, a gallon of water could be shown to consist of a certain number of minute particles or drops. The practical units in these two cases, however, are the gallon and coulomb.

As in the case of the water analogy, the rate of flow is expressed in gallons per minute, so in the electrical case the rate of flow can be expressed in coulombs per second. Another name has been given to this unit: That of "ampere," which is a rate of flow of one coulomb per second.

Since a current flow of electricity consists of a drift motion of electrons through a conductor, it is necessary to apply an electrical pressure to the circuit in order to start the electrons in motion. The unit of electrical pressure is the volt, and it is described as that pressure which will maintain a current flow of one ampere against a resistance of one ohm.

The unit of resistance is the ohm; it is the amount of resistance offered by a column of mercury 106.3 centimeters high and weighing 14.4521 grams at 0° C.

3. DISPLACEMENT CURRENTS

In a water system, as shown in Figure 259, if a rubber diaphragm is stretched across the pipe as shown at *a* and a pressure be applied in a direction shown by the arrow, the diaphragm would be displaced due to the difference in pressure exerted on its two sides.

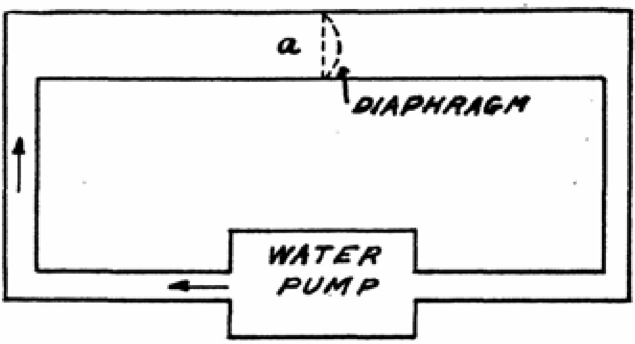


Figure 259 — Analogy for charge on a condenser.

The water in this case would merely be displaced and there would not be a steady flow unless the diaphragm were broken. Similarly in the electrical case, if an electrical potential is applied to the circuit shown in Figure 260, the electrons in the circuit would be displaced and a large number of them piled up on the plate *a*, thus giving it a negative static charge as indicated.

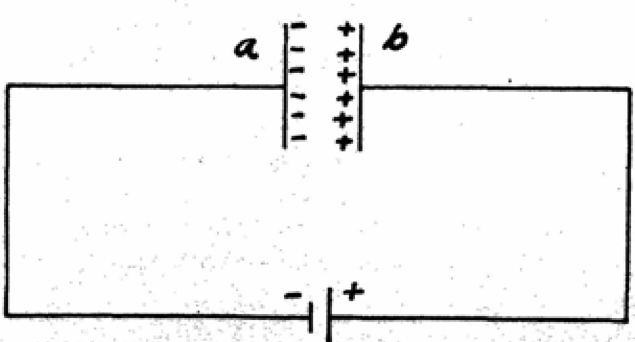


Figure 260 — Charge on a condenser.

In this case the insulating medium between the two conducting plates *a* and *b* acts as the rubber diaphragm in the previous analogy and prevents the electrons from flowing steadily around the circuit. Such an arrangement of two conducting plates separated by an insulator is called a condenser.

4. CAPACITY OF A CONDENSER

The action taking place in an electrical condenser may be shown to bear a considerable resemblance to that of a rubber balloon. If air is forced into the balloon at a pressure of one lb. per sq. inch, the balloon will be distended and, depending upon the elasticity and thickness of the material of which the balloon is composed and its initial size, a certain number of cubic inches of air will be forced into it. Now, if the pressure be increased to two lbs. per sq. inch, more air will be forced in, and if the pressure be raised too high, the sides of the balloon will eventually burst. It is obvious from this description that the capacity of such a balloon cannot be stated without specifying the pressure that is to be applied. A convenient method of expressing the capacity would be the number of cubic inches of air the balloon would hold upon application of a pressure of one lb. per sq. inch. Similarly in the case of the electrical condenser, the number of electrons that can be forced onto a condenser plate will depend upon the size of the plate, the type of insulating material separating the two plates, and its thickness.

Some insulating materials conduct electrostatic lines of force better than others. The basis of comparison of the different materials is taken as that of air at atmospheric pressure and the ability of a substance to conduct these lines of force as compared with air is called the dielectric constant, or specific inductive capacity. In Table IV, at the end of this chapter, is shown a list of materials and their specific inductive capacities.

A condenser of unit capacity is one in which, when a pressure of one volt is applied, one coulomb (6.3×10^{18}) electrons is forced into it. The unit of capacity is termed the farad.

The above statement expressed mathematically is:

$$C = \frac{Q}{E} \text{ or } Q = CE$$

where $C =$ capacitance in farads

$Q =$ quantity of electricity in coulombs

$E =$ pressure in volts

From the above description it will be realized that the capacity of a parallel plate condenser is determined by (1) the area of the plates, (2) the number of plates, (3) the dielectric constant, and (4) the thickness of the dielectric. The formula expressing the capacity of a parallel plate condenser is as follows:

$$C = \frac{8.84\ KA\ (n-1)}{10^8 \times l}$$

in which C = capacitance in microfarads
K = specific inductive capacity
A = area of one side of one plate in square centimeters
n = number of plates
l = thickness of the dielectric in centimeters.

NOTE: For the capacitance of other types of condensers see Bureau of Standards circular No. 74, *Radio Instruments and Measurements*.

TABLE III

Resistivity and Temperature Coefficient

MATERIAL	RESISTIVITY IN OHMS PER MIL-FOOT AT 20° C.	TEMPERATURE COEFFICIENT PER DEGREE C. PER OHM AT 0° C.
Silver	9.89 to 11.2	0.00435
Copper—		
Pure	10.15	0.00447
Annealed International Standard	10.4	0.00427
Hard Drawn	10.7	0.00413
Lead	12.1 to 12.9	0.00426
Aluminum	17.0	0.00423
Steel, hard	28.5	0.00165
Tungsten	35.8	0.00423
Zinc	36.4 to 39.6	0.00435
Brass	39	0.00102
Platinum	58.5 to 100	0.00412
Tin	63 to 75.5	0.00405
Iron, commercial	66.4 to 81.4	0.00618
Nickel	72.3 to 84.4	0.00682
Steel, rail	83.5 to 130	———
Tantalum	93.8	0.00285
Steel, soft	105	0.00448
Gold	133 to 136	0.0040
Platinum-iridium	148	0.00123

TABLE III (Continued)

MATERIAL	RESISTIVITY IN OHMS PER MIL-FOOT AT 20° C.	TEMPERATURE COEFFICIENT PER DEGREE C. PER OHM AT 0° C.
German Silver	181	0.000363
Manganin	250 to 450	0.000301
Antimony	251	0.00423
Monel Metal	257	0.00206
IaIa	295	0.000005
Constantin	300	0.000005
Advance Metal	326	0.000018
Superior	521	0.000823
Mercury	580	0.000713
Iron, hard cast	592	————
Nichrome	670	0.000444
Nichrome II	670	0.00016
Bismuth	718	0.00505
Calorite	720	————

TABLE IV

Specific Inductive Capacities

Dielectric Constants (Average Values)

MATERIAL	DIELECTRIC CONSTANT	MATERIAL	DIELECTRIC CONSTANT
Air	1	Glass, hard crown	7.0
Paraffin	2.1	Glass, flint	9.9
Petroleum	2.1	Alcohol (0° C.), amyl	17.4
Benzene	2.3	Ammonia	22
Ebonite	2.7	Acetone	26.6
Shellac	3.1	Alcohol (0° C.), ethyl	28.4
Gutta-percha	4.1	Alcohol (0° C.), methyl	35
Mica	5.8	Glycerine	56.2
Glass, lead	6.6	Water (pure)	81

Dielectric Strengths (Average Values)

MATERIAL	DIELECTRIC STRENGTH IN VOLTS PER CENTIMETER	MATERIAL	DIELECTRIC STRENGTH IN VOLTS PER CENTIMETER
Air	30,000	Gutta-percha	120,000
Mica	60,000	Paraffin	120,000
Transil Oil	90,000	Bakelite	150,000
Glass	100,000	Ebonite	600,000

Chapter XXVII

DIRECT CURRENTS

By A. P. HILL

As explained in the previous chapter, a current flow of electricity consists of a drift motion of electrons around a circuit from the negative to the positive pole of the voltage source.

Although, in moving in a vacuum, the electrons travel almost with the speed of light, i.e., approximately 186,300 miles a second, in the case of a direct current in a conductor the actual progression of the electrons in the circuit is very much slower, in some cases possibly only a fraction of a centimeter per second, and since, in order to obtain a current of one ampere, it is necessary for 6.3×10^{18} electrons to pass a given point in the circuit per second, it is obvious that a very dense flow of these particles must occur.

1. ELECTRICAL PRESSURE

In order to obtain a flow of water through a pipe it is necessary to have a difference in pressure between the two ends of the pipe. If the pressure at these two points was the same, no matter how high it might be, no flow of water would take place. Similarly, in an electrical circuit in order to produce a drift motion of the electrons along a wire, it is necessary that a difference in pressure or potential be applied. The

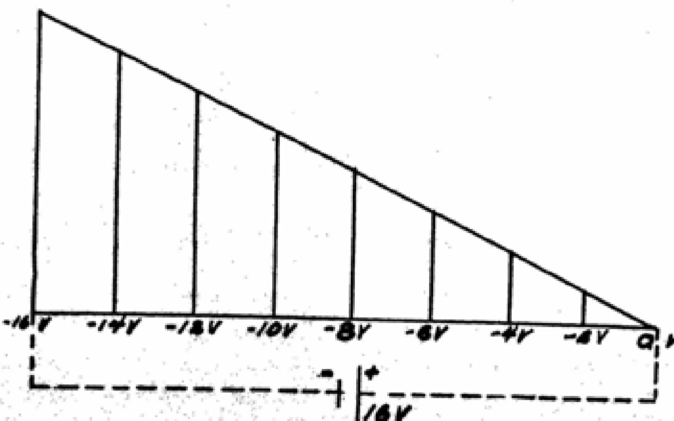


Figure 261 — Diagram illustrating drop in potential.

amount of flow will depend not on the absolute potential at any one point, but on the difference of the potential between the two points under consideration.

In Figure 261, if the wire is of uniform resistance, differences of potential along its length will be as indicated.

This conception of "drop in potential" is extremely important and should be clearly understood.

As previously mentioned, the unit of pressure or potential is the volt.

2. OHM'S LAW

The amount of current flow that takes place in a circuit is determined by the pressure applied and resistance of the circuit. The relationship which exists is stated in Ohm's Law as follows:

$$\text{Current} = \frac{\text{Pressure}}{\text{Resistance}}$$

from which

$$\text{Resistance} = \frac{\text{Pressure}}{\text{Current}}$$

and

$$\text{Pressure} = \text{Resistance} \times \text{Current}$$

The foregoing expressed mathematically, using the conventional abbreviations, is:

$$I = \frac{E}{R}$$

$$R = \frac{E}{I}$$

$$E = RI$$

where E = pressure in volts
R = resistance in ohms
I = current in amperes

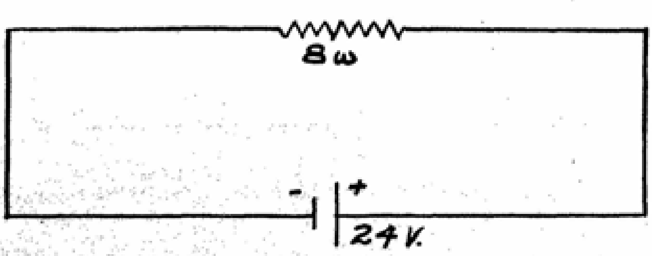


Figure 262 — The current may be calculated where resistance and voltage are given.

Applying Ohm's Law to the circuit of Figure 262.

$$I = \frac{E}{R} = \frac{24}{8} = 3 \text{ amperes}$$

If the current and voltage had been given, the resistance could have been calculated from:

$$R = \frac{E}{I} = \frac{24}{3} = 8 \text{ ohms}$$

or, if the values of resistance and current had been given, the voltage could have been calculated from:

$$E = RI = 8 \times 3 = 24 \text{ volts}$$

From this it will be seen that if any two of the three quantities involved are known, the third can be calculated by the application of Ohm's Law.

3. RESISTANCES IN SERIES AND IN PARALLEL

Resistances may be connected in series as shown in Figure 263, or in parallel as shown in Figure 264.

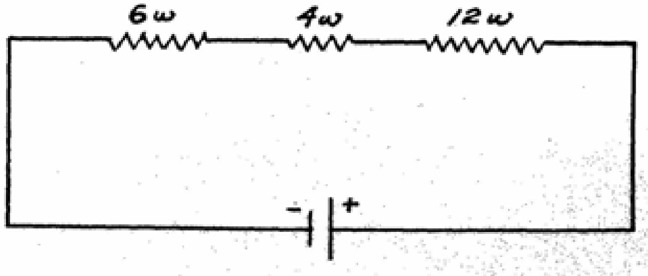


Figure 263 — The resistances of 6, 4, and 12 ohms are in series in the circuit.

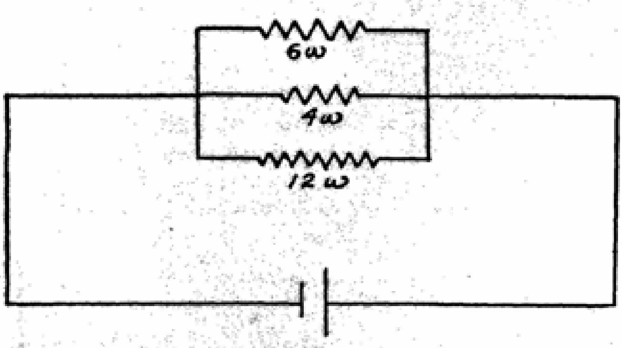


Figure 264 — The resistances of 6, 4, and 12 ohms are in parallel with one another.

The total resistance of two or more resistances in series is equal to their sum, i.e.:

$$R = r_1 + r_2 + r_3 + \text{........}$$

Thus in Figure 263 the total resistance $R = 6 + 4 + 12 = 22$ ohms.

The total resistance of resistances in parallel may be calculated from the formula:

$$\frac{1}{R} = \frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} + \text{........}$$

so that the total resistance of the circuit shown in Figure 264 may be calculated as follows:

$$\frac{1}{R} = \frac{1}{6} + \frac{1}{4} + \frac{1}{12} = \frac{2+3+1}{12} = \frac{6}{12}$$

$$\therefore \quad R = \frac{12}{6} = 2 \text{ ohms}$$

4. SERIES-PARALLEL ARRANGEMENT OF RESISTANCES

It will be realized that resistances in a circuit may be connected in a more complex manner than those shown above. The solution of the more complex circuits is, however, reducible to the elements described above.

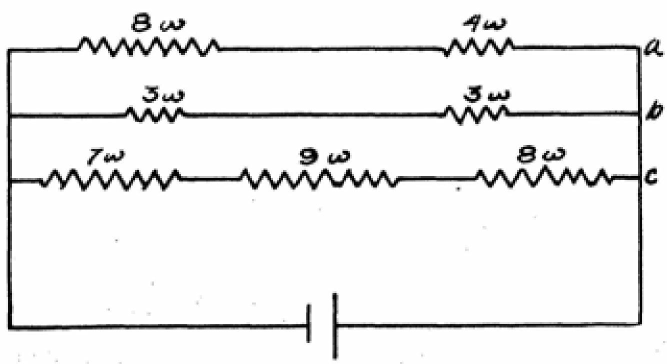


Figure 265 — Compound arrangement of three branches in parallel.

For example, in the case of the circuit shown in Figure 265, the three branches, *a*, *b*, and *c*, are in parallel with one another. The resistances in each branch, however, being in series with one another, should

first be added in order to obtain the total resistance of each branch; thus the total resistance of branch a is 12 ohms, branch b is 6 ohms, and branch c is 24 ohms. The total resistance of the complete circuit is therefore calculated as follows:

$$\frac{1}{R}=\frac{1}{12}+\frac{1}{6}+\frac{1}{24}=\frac{2+4+1}{24}=\frac{7}{24}$$

$$\therefore \quad R=\frac{24}{7}=3.428+\text{ohms}$$

Another type of circuit which will be discussed more in detail later is shown in Figure 266.

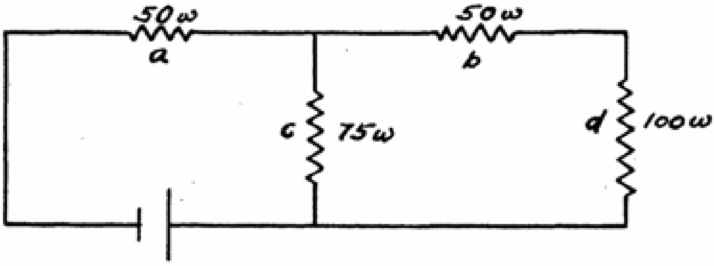


Figure 266 — "T" type network. Resistances b and d are together in parallel with c, and the combination in series with a.

In this instance, resistances b and d, which are in series with one another, are together in parallel with resistance c, and this whole combination is in series with resistance a.

The total resistance of this circuit is then calculated as follows:

$$b+d=150 \text{ ohms}$$

The total resistance of c in parallel with d and b is

$$\frac{1}{R}=\frac{1}{75}+\frac{1}{150}=\frac{2+1}{150}=\frac{3}{150}$$

$$\therefore \quad R \text{ (of } c, b, \text{ and } d) = 50 \text{ ohms}$$

and the resistance of the complete circuit is there for the sum of c, b, and d in series with a,

$$\text{or } R \text{ (total)} = 50+50=100 \text{ ohms}$$

Another method of statement for the resistance of two resistances in parallel is:

$$R=\frac{r_1 \times r_2}{r_1+r_2}$$

Using this method of expression a complete statement of the circuit shown in Figure 266 could be written as follows:

$$R = a + \frac{c\ (b + d)}{c + b + d}$$

$$= 50 + \frac{75\ (50 + 100)}{75 + 50 + 100}$$

$$= 50 + \frac{11{,}250}{225}$$

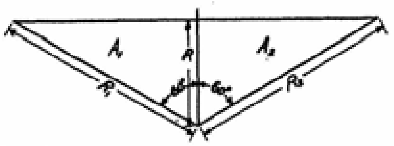
$$= 50 + 50$$

$$= 100 \text{ ohms}$$

There are several methods for the graphical solution of resistances in parallel, one of which is shown in Figure 267. The total resistance of any number of resistances in parallel may be found by using this chart in the manner indicated.

GRAPHICAL SOLUTION OF RESISTANCES IN PARALLEL *

Proof



To show that the line cuts the three radiating axes in the above figure in values of R, R_1 and R_2 that satisfy $\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$ where the R's are measured from the common point and the angle between the adjacent axes is 60°.

$A = A_1 + A_2$ where A, A_1 and A_2 are areas of triangles in the figure.

$\therefore \frac{R_1 R_2}{2} \sin 120° = \frac{R_1 R}{2} \sin 60° + \frac{R R_2}{2} \sin 60°.$

But $\sin 120° = \sin 60°.$

$\therefore R_1 R_2 = R_1 R + R R_2$

dividing by $R R_1 R_2$.

$\therefore \frac{1}{R} = \frac{1}{R_2} + \frac{1}{R_1}$

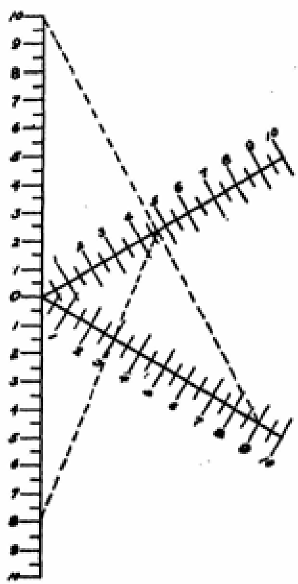


Figure 267

* NOTE: See page 526 for large scale figure.

5. POTENTIAL DROP

In Figure 268 is shown a water system in which the water pressure in the tank is 100 lbs. per sq. inch.

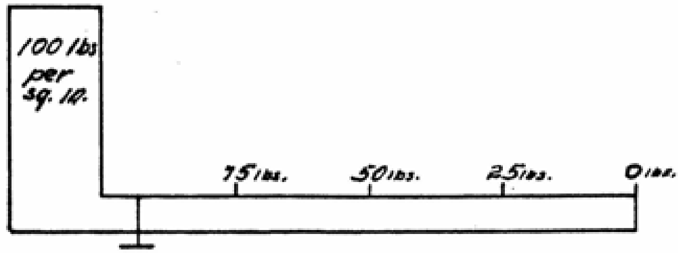


Figure 268 — Water analogy illustrating drop in potential.

If we assume the valve to be opened and the water flowing through the pipe there is a drop in pressure along the pipe due to the fact that it offers a resistance to the flow of water. If we assume the pipe to be of uniform diameter along its length, the pressure will have dropped to 50 lbs. per sq. inch at the half way point as indicated. Similarly at a point a quarter of the distance along the pipe the pressure will be 75 lbs. per sq. inch, and at three-quarters of the distance, 25 lbs. per sq. inch. A similar situation is true in the case of the electrical circuit. In Figure 269 the voltage applied to the circuit from the battery is assumed to be ten volts, i.e., there is a difference in potential between the two ends of the circuit of ten volts.

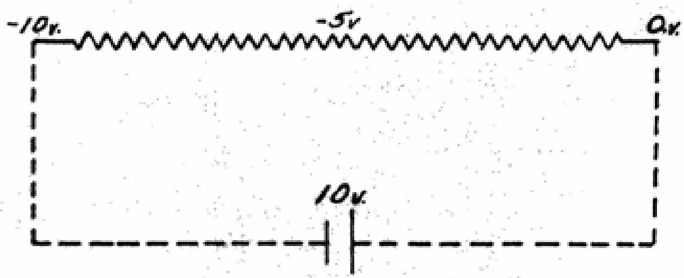


Figure 269 — There is a uniform drop in potential along this circuit as indicated.

Assuming the wire to be uniform throughout the circuit, the potential drop at a point half way along the circuit will be five volts. The potential drop can thus be calculated from Ohm's Law, $E = RI$.

It should also be realized that the battery or other generating device must of necessity have some internal resistance, so there is a potential

drop within the battery itself if a current is flowing. Thus, in Figure 270, if the battery on open circuit gives a pressure of 10 volts, and its internal resistance is 0.5 ohms and the external resistance of the circuit is 9.5 ohms, the total current may be calculated as follows:

$$I = \frac{E}{R} = \frac{10}{10}$$

$$= 1 \text{ ampere}$$

and the potential actually applied to the circuit is 9.5 volts, i.e., a potential drop has occurred within the battery itself.

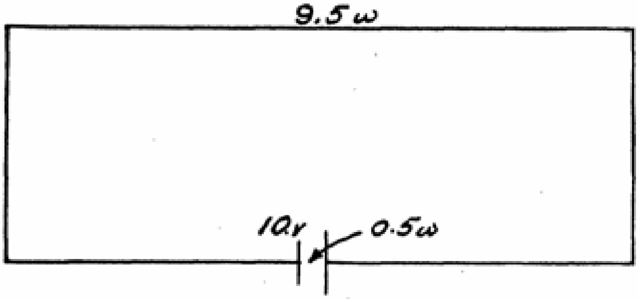


Figure 270 — The total resistance of a circuit is comprised of the internal resistance of the battery and external resistance of the circuit in series with one another.

If the internal resistance is represented by r and the external resistance by R, the potential drop across the battery equals $1 \times r$ and the potential drop across the external circuit equals $1 \times R$ (volts).

6. KIRCHOFF'S LAWS

Some of the laws governing circuits of this type were first formulated by Kirchoff as follows:

(1) At any point in a circuit there is as much current flowing away from the point as there is towards it.

(2) In any closed circuit the algebraic sum of the voltage applied and the potential drops is equal to zero.

By the combined application of Ohm's and Kirchoff's laws a solution may be obtained from most direct-current networks. To illustrate Kirchoff's first law reference may be made to Figure 271

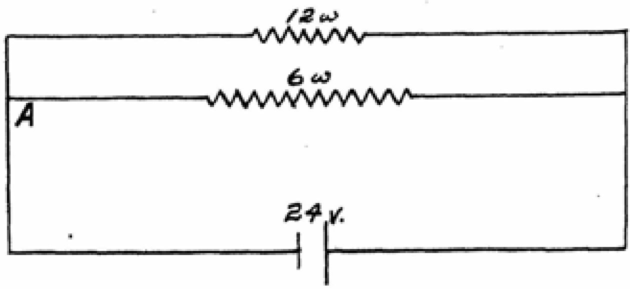


Figure 271 — Kirchoff's first law: The current flowing towards A equals the total current flowing away from A.

The total resistance of the circuit $R = \frac{6 \times 12}{6 + 12} = 4$ ohms

The total current in the circuit $I = \frac{24}{4} = 6$ amperes

This current divides at point A, part of it flowing through the upper and part through the lower resistance. Since the potential across these resistances is the same the current through the upper resistance $= \frac{24}{12} = 2$ amperes. The current through the lower resistance $= \frac{24}{6} = 4$ amperes.

It will thus be seen that this agrees with Kirchoff's first law in that the current flowing towards the point A is six amperes and that this current divides through the two resistances, the total flowing in the two branches being equal to six amperes.

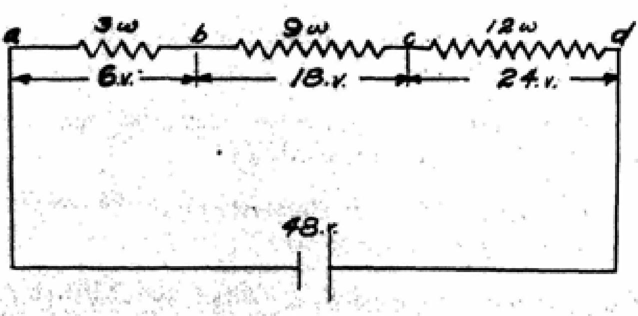


Figure 272 — Kirchoff's second law: The algebraic sum of the voltage applied and the potential drops equals zero.

Figure 272 will serve to illustrate Kirchoff's second law, namely: In any closed circuit the algebraic sum of the voltage applied and the potential drops is equal to zero. The drop in potential between a and b is 6 volts; between b and c, 18 volts, and between c and d, 24 volts. The total potential drop is therefore 48 volts. Since point a is 6 volts negative with respect to point b, and point b 18 volts negative with respect to c, and c again 24 volts negative with respect to d, the algebraic sum of the potential applied and the potential drops is equal to zero, or $E = 48 - 6 - 18 - 24 \doteq 0$.

Chapter XXVIII

ELECTRICAL POWER AND ENERGY

By A. P. HILL

An electrical current has been shown to consist of a drift motion of minute electrically charged particles (electrons) through a conductor. In order to create such a flow of electrons it is necessary that an electrical pressure or voltage be applied to the circuit. It is obvious that if a relatively small number of electrons are to be forced through a conductor they will require a lesser expenditure of power than if a greater number are to be so moved. It is therefore necessary that we have as a reference a unit of power.

1. UNIT OF POWER — THE WATT

The unit of electrical power is the watt and denotes the amount of power used when a pressure of one volt forces a current of one ampere through a circuit. The number of watts involved in forcing any amount of current through a circuit at a given pressure may be found by obtaining the product of the current and voltage:

i.e. $P = IE$

where $P =$ power in watts

$I =$ current in amperes

$E =$ pressure in volts

Since $I = \dfrac{E}{R}$

$$P = \frac{E^2}{R}$$

or since $E = IR$

$$P = I^2R$$

Thus in Figure 273, in which R represents the resistance of the heater of a Western Electric 294-A vacuum tube, which requires a current of 0.32 amperes at a pressure of 10 volts, the power $= 10 \times 0.32 = 3.2$ watts.

The resistance of this heater $R = \dfrac{E}{I} = \dfrac{10}{0.32} = 31.25$ ohms

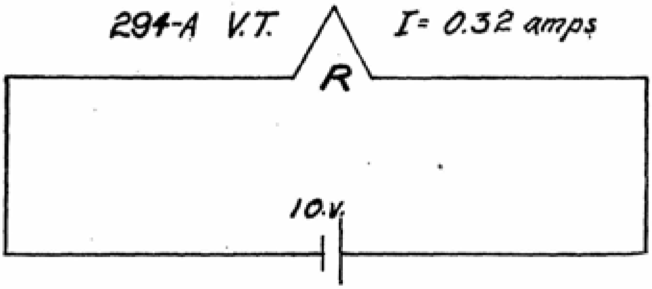


Figure 273 — Power is consumed in the filament of a vacuum tube.

As an alternative method, if this latter value were known, the power could be calculated from the formula:

$$P = \frac{E^2}{R} = \frac{100}{31.25} = 3.2 \text{ watts}$$

or, if the current and resistance were known, it could be calculated as follows:

$$P = I^2R = (0.32)^2 \times 31.25 = 0.1024 \times 31.25 = 3.2 \text{ watts}$$

2. KILOWATT AND HORSE POWER

Since the watt is too small a unit in which to conveniently express the output of modern electrical power machinery, a unit called the kilowatt equal to 1000 watts is frequently used.

$$1 \text{ kilowatt } (KW) = 1000 \text{ watts}$$

The horse power is another unit that is often used, its relation to the watt being 1 horse power $(HP) = 746$ watts.

Chapter XXIX

MAGNETISM

By A. P. HILL

Magnetism is a property of iron, nickel and cobalt. It is most pronounced, however, in iron and some iron alloys. This property has been known for many years. It was undoubtedly known to the early Greeks who discovered that certain magnetic ore when suspended always tended to turn in a direction approximating the north and south poles of the earth. The ore was discovered in Magnesia, Asia Minor, from which the name "magnetism" is derived. It is probable, however, that the magnetic properties were observed much earlier than this and that the principle was put to use for purposes of navigation in the very earliest civilizations.

1. MAGNETS

Magnets, as we know them, may be classed as either permanent or electro-magnets. A permanent magnet is one which retains its magnetism after the magnetizing force is removed, while an electro-magnet, is one which loses the greater part of its magnetism immediately upon the removal of such a force.

The earth itself acts as a large magnet and has a north and south magnetic pole. These poles do not coincide with the north and south geographical poles. The north magnetic pole is located in Boothia, Northern Canada, at approximately 71° north latitude and 91° west longitude, and it is to this magnetic pole that an ordinary compass needle points.

As is well known, *like* magnetic poles repel one another while *unlike* poles attract one another. Consequently, if we assume the magnetic pole nearest the north geographic pole to be the north pole, the south pole of a compass needle points towards that north magnetic pole. Such a pole was originally termed a north seeking pole. It is now more generally called simply a "North Pole." However, it should be realized in this and all other cases that unlike poles attract one another.

2. MAGNETIC LINES OF FORCE

According to the latest theories, magnetism is assumed to be the result of rotating electrical charges within the atom itself. Each atom consequently behaves like a magnet. When a large number of these atoms align themselves in a certain definite manner the substance as a whole becomes magnetized.*

If the individual atoms exhibit magnetic properties, it is then natural that the molecules should also appear to be small magnets, the adjacent north and south poles being attracted to one another as shown in Figure 274.

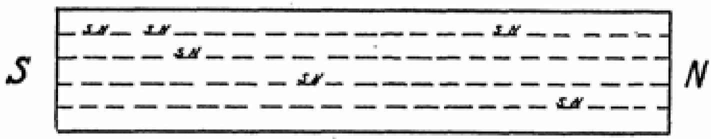


Figure 274 — Arrangement of molecules in a bar magnet.

In this figure it will be seen that the ends of all the molecules pointing to the right have a north polarity, while the left-hand ends have a south polarity. If, therefore, such a magnet were broken into two or more pieces each piece would have a definite polarity determined by that of the individual molecules as shown in Figure 275.

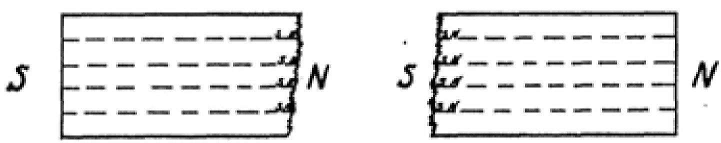


Figure 275 — When a bar magnet is broken, each piece becomes a magnet.

Figure 276 indicates the condition that exists external to the magnet itself. The magnetic effect from each chain of molecules continues through the surrounding space in the form of what are known as magnetic lines of force.

The magnetic lines of force are assumed to issue from the magnet at the north pole and reenter it at the south pole, continuing through the magnet itself from south to north. The space through which these lines of force pass is termed the "magnetic field."

It should be realized that the conception of lines of force is merely a convenient method of describing a field of force existing in space. The

* NOTE: For further information on this subject see Bell System Technical Journal, "*The Theory of Magnetism,*" by Karl K. Darrow, April, 1936, page 224.

strength of the field is represented by the number of lines of force existing per unit area of the field.

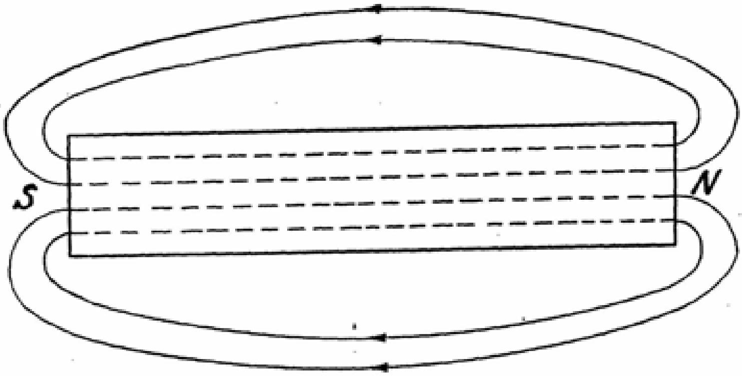


Figure 276 — A diagram of the field about a bar magnet.

3. RING MAGNETS

In certain cases it is possible to have a piece of magnetic material strongly magnetized without possessing any magnetic poles, as in the ring illustrated in Figure 277.

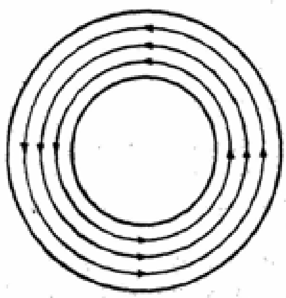


Figure 277 — A solid ring magnet.

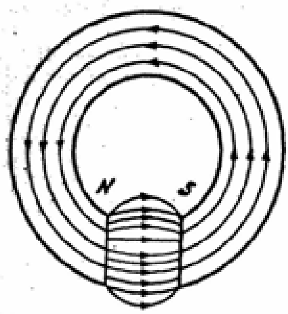


Figure 278 — A ring magnet with an air gap.

When such a ring is broken by the interposition of an air gap as shown in Figure 278, the magnetic polarities appear as shown.

4. GENERATOR FIELD POLES

Figure 279 shows the magnetic circuit of a two-pole generator. In this instrument electrical conductors are rotated between the poles of the magnet, thus cutting through the magnetic lines of force and so

generating an electrical potential in the conductors. It is on this principle that all electrical generators operate.

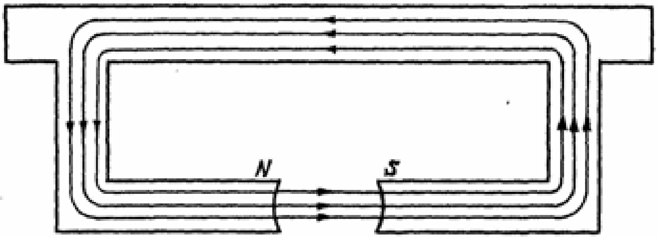


Figure 279 — The magnetic circuit of a two-pole generator.

Figure 280 shows the magnetic circuit of a four-pole generator. In this instrument the magnetic poles are arranged alternately, *like* poles being opposite one another. The lines of force will, however, pass from a north to a south pole as indicated.

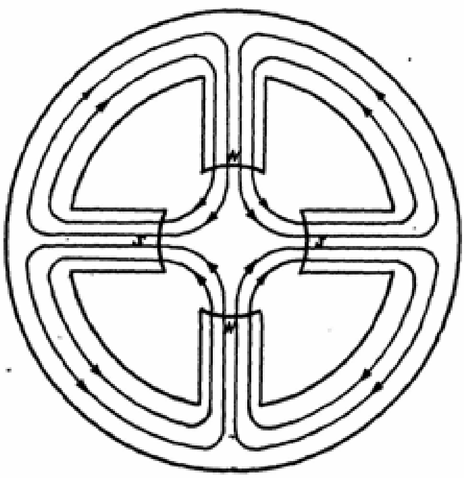


Figure 280 — The magnetic circuit of a four-pole generator.

5. PERMEABILITY

In the case of an electrical circuit it is known that some substances conduct electricity better than others. A list of a number of substances with their relative conductivities was given in Table III at the end of Chapter XXVI. In a somewhat similar manner some substances conduct magnetic lines of force better than others. In general, iron and iron alloys are the best conductors of such lines of force. The ability of a substance to conduct these lines of force, compared with air, is termed its permeability. Table V lists the permeabilities of a few of the more

commonly used magnetic materials. The reciprocal of permeability is called reluctance and corresponds to resistance in electrical conductors.

6. MAGNETIC SHIELDING

There are many instances where it is desirable to exclude magnetic lines of force from certain pieces of equipment, in which case we use what is termed a "magnetic shield." Such a shield is shown in Figure 281.

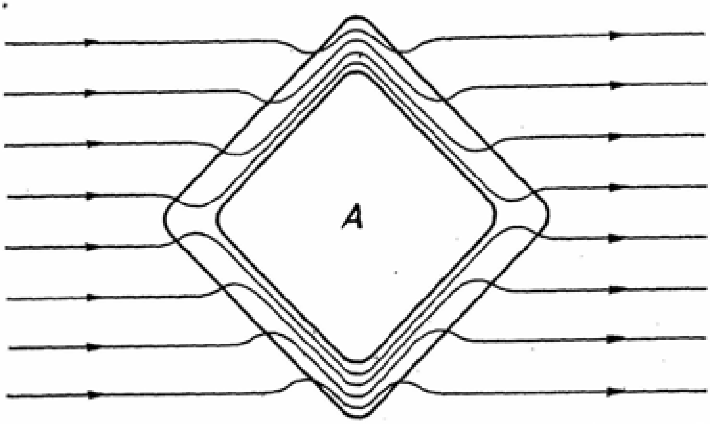


Figure 281 — The magnetic lines are nearly all shunted around the space A by the iron shield.

It will be noticed that the magnetic lines of force are prevented from passing through the space "A," not by virtue of the resistance offered them by the magnetic shield, but because of the fact that this shield offers them a lower resistance path than does the air. The lines of force consequently pass through the magnetic shield (usually composed of iron) and so are deflected from the area it is desired to shield.

7. THE COMPASS

The compass needle consists of a bar magnet suspended so it can rotate freely in a horizontal direction, and as a result the south pole of the needle is attracted by the north magnetic pole of the earth.

As explained in Section I, Page 428, the south pole of the compass needle is called the "north seeking" or "north" pole. When a

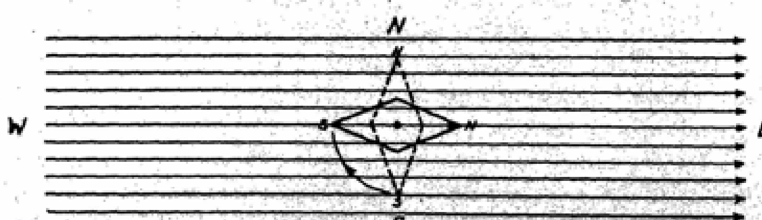


Figure 282 — A compass needle sets itself parallel to the lines of force near it.

compass needle is affected by another magnetic field in addition to that of the earth it tends to set itself parallel to these lines of force, which in Figure 282 are represented as moving from west to east, so that the lines of force within itself are parallel to, and in the same direction as, the lines of the external field. It is due to this fact that special caution must be taken on board ship to be sure that any magnetic fields produced in the ship itself due to its metallic construction do not affect the compass; if there are such fields, the effect they have on the compass must be accurately known so that the necessary correction may be applied to its reading.

TABLE V

Permeability of Iron and Steel

MATERIAL	MAXIMUM PERMEABILITY
Cast iron	240
Cast iron, annealed	600
Cast steel	3,550
Cast steel, annealed	14,800
Electrolytic iron	1,850
Electrolytic iron, annealed	14,400
Electrical iron in sheets, annealed	3,270
Steel, hardened	110

Chapter XXX

ELECTRO-MAGNETISM AND THE MAGNETIC CIRCUIT

By A. P. HILL

1. MAGNETIC FLUX ABOUT A CONDUCTOR

In the previous chapter it was stated that an atom exhibits magnetic properties due to the fact that in each atom there are minute electrical charges called electrons rotating around the nucleus. Consequently, it is logical to assume that since an electrical current consists of a drift motion of electrons along the length of a conductor, that if it were bent in the form of a circle such a wire would also exhibit magnetic properties, and this is found to be the case. In fact, a straight wire in which a current is flowing is found to have surrounding it magnetic lines of force as indicated in Figure 283. These magnetic lines of force circle about the wire in the direction shown in the figure; i.e., if the electrons in the circuit are moving towards the observer the magnetic lines of force circle around the conductor in a clock-wise direction.[1]

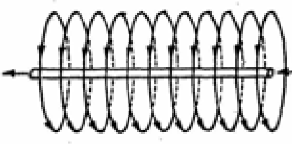


Figure 283 — Magnetic lines of force surround a conductor in which a current is flowing.

Another method of remembering the direction of flow of the magnetic lines of force is to assume that the wire in which the current is flowing is held in the left hand with the thumb pointing in the direction of electron flow. The lines of force then circle the wire in the direction in which the fingers point, as shown in Figure 284. Magnetic lines of force which are produced as the result of the flow of an electrical current are termed "electro-magnetic lines of force."

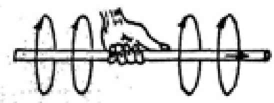


Figure 284 — If a wire is held in the left hand with the thumb pointing in the direction of motion of the electrons, the fingers will point in the direction of the lines of force.

2. MAGNETIC UNITS

In order that magnetic forces may be calculated, it is necessary to assume certain units. The total number of lines of force in a magnetic

field are referred to as the magnetic flux, and such a field is designated by the Greek letter ϕ. One line of force is termed a Maxwell. The magnetic field may be quite extensive and the density of the lines of force in the field may vary between different parts of the field. Consequently, it is necessary in many cases to know the density of the field (that is, the number of lines of force per square centimeter) at the point in which we are particularly interested. For such cases, we make use of a unit of magnetic field intensity, the name of which is the "gauss."

1 gauss = 1 line of force per square centimeter.

The letter H is used as the symbol representing the gauss.

If we assume the intensity of the field to be uniform throughout its area, the total number of lines of force must equal the total area of the field multiplied by its intensity:

$$\phi = AH$$

ϕ = total number of lines of force in a magnetic field

where A = area of the field in square centimeters

and H = number of lines of force per square centimeter (gausses).

3. MAGNETIC PRESSURE

In an electrical circuit a current flow is the result of applying an electrical potential (voltage) to a circuit which has a definite resistance. Without this pressure we would, of course, have no current flow. Similarly, in a magnetic circuit we must apply a magnetic pressure before we can obtain a flow of magnetic lines of force (or flux). The unit of magnetic pressure is the "Gilbert," and is designated by the letter F. It is known as the unit of magnetomotive force.

4. MAGNETIC RESISTANCE — RELUCTANCE

As mentioned in a previous paragraph, a current flow in an electrical circuit is the result of applying an electrical potential to a resistance. The relationship which exists between these three factors as stated in Ohm's Law is:

$$\text{Current} = \frac{\text{Voltage}}{\text{Resistance}}$$

Similarly in the magnetic circuit a magnetic flux is produced when the magnetomotive force is applied to a magnetic resistance, or, as it is termed, reluctance; i.e.:

$$\text{Magnetic flux} = \frac{\text{Magnetomotive force}}{\text{Reluctance}}$$

The unit of reluctance is the "Oersted" and is the reluctance of a centimeter cube of air. The reluctivity of practically all materials other than iron and the iron alloys is the same as that of air:

$$R = K \times \frac{l}{A}$$

where R = reluctance in oersteds

K = reluctivity of material under consideration

l = length in centimeters of the magnetic circuit

A = cross-sectional area of the magnetic circuit in square centimeters

The unit of magnetomotive force (the gilbert) is defined as that force which, when applied across one centimeter in air, will produce one line of force per square centimeter (one gauss).

It will thus be seen, that as in an electrical circuit, where there are three main factors to be considered, namely, pressure, resistance, and current, so in a magnetic circuit there are three corresponding factors, magnetomotive force, reluctance, and flux. The mathematical statement of the relationship between these factors is:

$$\text{Magnetic flux} = \frac{\text{Magnetic pressure}}{\text{Magnetic resistance}}$$

or $$\text{Maxwells} = \frac{\text{Gilberts}}{\text{Oersteds}}$$

or $$\phi = \frac{F}{R}$$

from which $$R = \frac{F}{\phi}$$

and $$F = \phi \times R$$

It should be realized that magnetic flux is not entirely analogous to electrical current, for as explained previously, an electrical current consists of a flow of electrons through a conductor, while a magnetic

flux exists in the form of a "state of strain" in the medium in which it is established.

5. UNIT MAGNETIC POLE

Since the strength of a magnetic pole will vary over a very wide range it is useful for purposes of calculation to assume a unit magnetic pole, and such a pole has been defined as one having a strength such that, if placed one centimeter away in air from a like pole it will repel it with a force of one dyne.

Another definition of the unit pole is that it is a pole of such strength that when placed in a field of one gauss it is acted upon by a force of one dyne:

$$\therefore \quad F = m \times H$$

where F = force in dynes
m = strength of magnet in unit poles
and H = field strength in gausses.

The mutual action between two magnets placed in proximity to one another may be stated in the following manner:

$$F = \frac{m \times m'}{d^2}$$

where F = force in dynes
m and m' = strength in unit poles of two magnets
and d = distance between them in centimeters.

6. MAGNETIC PROPERTIES OF IRON

As previously stated, some substances, notably iron and the iron alloys, conduct magnetic lines of force better than air. The conductivity of the latter is taken as unity. In other words, such substances have a lower reluctance than air. Consequently, if a magnetic field is established in air and a piece of iron introduced into the field, the lines of force per square centimeter (gausses) in the iron will be greater than in the air. This is illustrated in Figure 285.

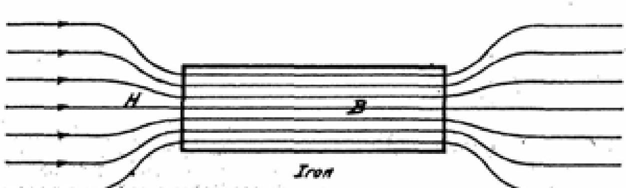


Figure 285 — The flux density in iron is greater than that in air under this condition.

In this figure the flux density of the magnetizing force in air is represented by the letter H. The resultant flux density in the iron is represented by the letter B. The latter is also measured in gausses. It is convenient to use different symbols for these two cases in order to differentiate between the densities in air and in iron.

If a piece of iron is placed in a magnetic field of low flux density, the density in the iron will be considerably greater than in the air. If, now, the density of the magnetizing flux be gradually increased, the flux density in the iron will also increase up to a certain point where it is impossible to obtain any greater density no matter what the magnetizing force may be. The iron under this condition is said to be saturated. This is due to the fact that practically all the molecules have been lined up into chains throughout the substance. No further increase in the total number of such chains of molecules is possible.

7. *B-H* CURVES

Figure 286 graphically indicates the condition described above, and it will be noted that saturation is reached at an H value of approximately ten in the case of wrought iron and annealed sheet steel; and somewhat higher in the case of soft steel castings. Such a curve may be drawn for all magnetic materials; their shape and saturation points will depend, however, upon a number of factors, such as:

(a) the type of magnetic material, whether cast iron, wrought iron, steel, or some alloy

(b) their degree of purity

(c) their previous magnetic history; i.e., whether or not they have previously been subjected to a high degree of magnetization.

The term "permeability" has already been described in Chapter XXIX, Section Five, and was stated qualitatively to mean the ability of a substance to conduct lines of force compared with air. It will now be seen that the permeability may be obtained from the *B-H* curve of the material.

$$\mu = \frac{B}{H}$$

where $\mu =$ permeability

$B =$ flux density in the material under consideration

and $H =$ flux density of the magnetizing field.

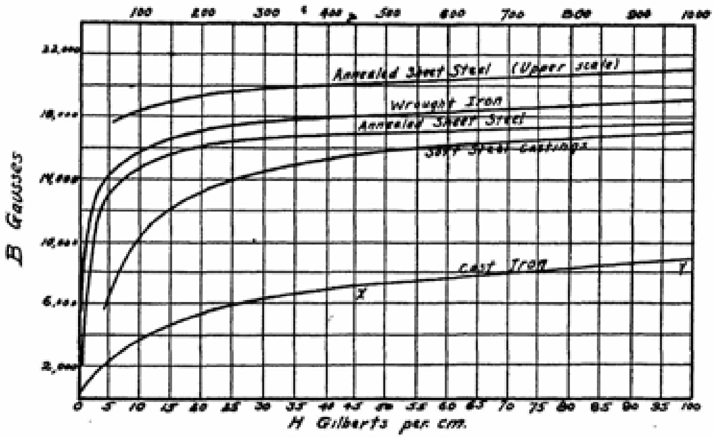


Figure 286 — Typical magnetization curves of iron and steel.

By reference to Figure 286 it will be noticed that the permeability of a substance is not a fixed value but varies with the magnetizing force applied. The permeability of most substances is calculated at the relatively straight line portion of a curve as between points X and Y in the case of cast iron.

The permeability of some of the magnetic alloys such as permalloy is very much greater at low flux densities than that of iron, principally below magnetizing forces of the order of one gauss.

Figure 287 shows a *B-H* curve for permalloy and Armco iron in this region. Due to the fact that permalloy of this type becomes saturated at low values of magnetizing force, for values very much greater than those shown in the figure, the permeability of the iron will be greater than that of the permalloy.

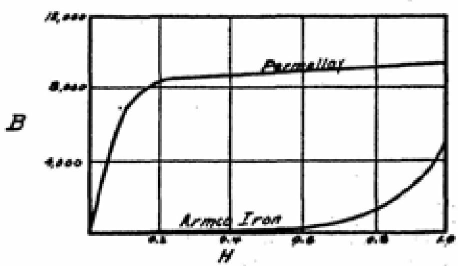


Figure 287 — *B-H* curve for permalloy and Armco iron.

8. HYSTERESIS LOOPS

If a piece of iron is subjected to a magnetizing force and this force is gradually increased until the point of saturation is reached and then

decreased to zero, the iron will retain some of its magnetism as shown in Figure 288.

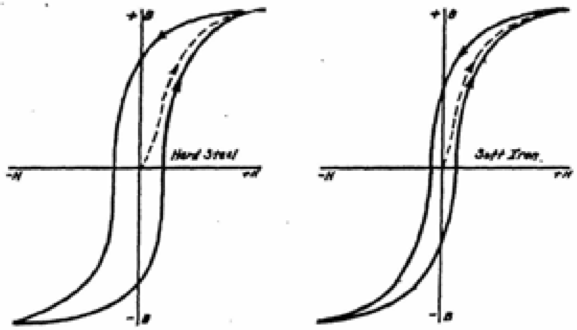


Figure 288 — Hysteresis loops.

If, now, the magnetizing force be reversed in direction, increased in that direction until saturation is reached, again reduced to zero, and increased in the original direction up to saturation, the *B-H* curve has the shape shown in Figure 288. This is called a hysteresis loop and indicates the amount of energy used up during the cycle of operations described above. It will be noticed that if the iron has once reached the saturation point it does not return to its original magnetic condition regardless of what magnetizing forces may be applied. Figure 289 shows hysteresis loops for Armco iron and permalloy.

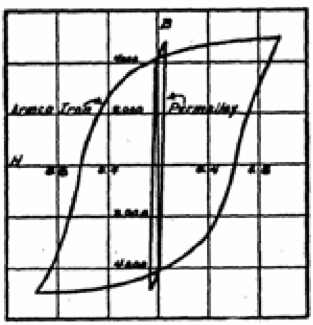


Figure 289 — Hysteresis loops for Armco iron and permalloy.

9. ELECTRO-MAGNETISM

If a straight vertical conductor in which a current is flowing pierces a horizontal sheet of cardboard, as shown in Figure 290, it may be shown that magnetic lines of force circle about the conductor on the same plane as the cardboard. This may be illustrated by sprinkling some iron filings on the cardboard, in which case, when the current starts to flow, the filings will arrange themselves in circles about the con-

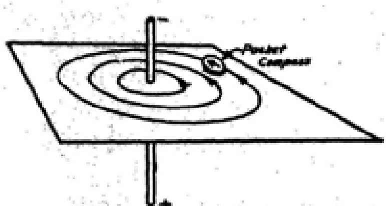


Figure 290 — By means of iron filings and a compass it may be shown that there is a circular magnetic field about a wire carrying a current.

ductor as the result of being magnetized by the flux about the wire.

As the result of such an experiment we learn that whenever an electrical current flows in a circuit, electro-magnetic lines of force circle about the conductor and these lines of force are always in a plane perpendicular to the conductor. If a small compass needle is placed in this magnetic field, as shown in the figure, the needle will set itself parallel to the lines of force, and since in the compass needle itself they move from south to north, the north pole of the compass will point in the direction in which the electro-magnetic lines of force surrounding the conductor are directed. As a result of such an experiment we may state the following rule:

> If a current flows in a conductor, and the electrons composing this current are moving towards the observer, the lines of force will surround the conductor in a clock-wise direction. If, on the other hand, the electrons are moving away from the observer, the lines of force will move in a counter-clockwise direction. This is indicated in Figures 291-A and 291-B. It should be realized that each of these lines of force exists as a closed circle about the conductor and the electro-magnetic effect produced is a function of the amount of current flowing in the conductor.

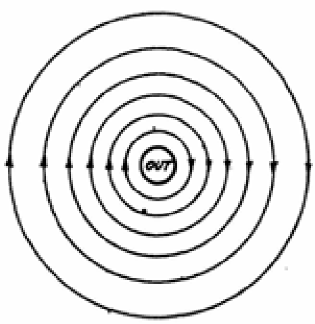


Figure 291-A — Lines of force produced by electrons moving towards the observer.

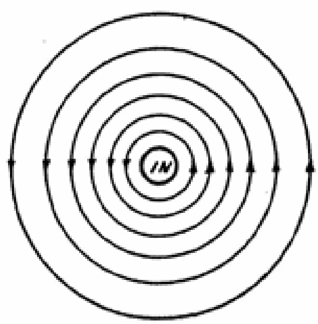


Figure 291-B — Lines of force produced by electrons moving away from the observer.

10. FIELD OF FORCE ABOUT A COIL

If, now, instead of considering a straight wire we assume this wire to be formed into a coil as shown in Figure 292, then at point A, since

the electrons are moving towards the observer, the lines of force will circle it in a clock-wise direction.

At the point B a similar condition exists with the result that between conductors A and B the lines of force due to turn of wire A, tend to move upwards while those due to turn of wire B, tend to move downward, the result is, therefore, that all lines of force tend to circle both conductors rather than each conductor individually. If, therefore, a coil is wound with the turns close together as shown in Figure 293, the tendency is for the lines of force to travel through the coil, leaving it at the left-hand end and reentering it at the right-hand end. The similarity between this and the condition existing in an ordinary bar magnet should be noted. Consequently, it is to be expected that such a coil will act as an ordinary bar magnet having a north pole to the left and a south pole to the right, and this is found to be the case. The intensity of the field inside the coil is dependent upon the number of turns of wire and the amount of current flowing in the wire. Such a coil having an air core is called a solenoid. If, however, the coil has an iron core it is commonly known as an electro-magnet.

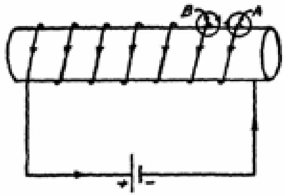


Figure 292 — The magnetic field about a coil of wire carrying a current.

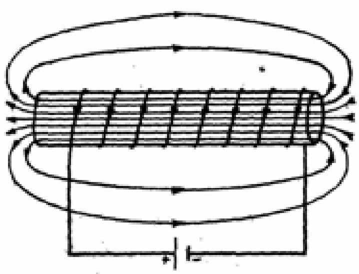


Figure 293 — The magnetic field about a closely wound coil carrying a current.

11. FLUX DENSITY WITHIN A SOLENOID

It has been found that if a solenoid whose length is at least ten times its diameter is wound so that there is one turn of wire to every centimeter of length and if one ampere is passed through the wire, the field intensity inside the solenoid is 1.26 or $\frac{4\pi}{10}$ gausses:

$$\text{i.e.}\quad H=\frac{4\pi NI}{10\,l}$$
$$=\frac{1.26\,NI}{l}$$

where H = field strength in gausses
N = total number of turns on the solenoid
I = current in amperes
and l = length of solenoid in centimeters

The total flux in a magnetic circuit equals the cross-sectional area of the circuit multiplied by the field density, or

$$\phi = AH$$

Substituting the value of H in the previous equation

$$\phi = \frac{4\pi NIA}{10\,l}$$

This equation gives the flux at the center of a single layer solenoid and does not allow for leakage effects which cut down the flux to approximately one-half at the ends of the coil.

---

[1] NOTE: As previously mentioned, all text books still make use of the old theory that an electrical current flows from the positive to the negative pole of a battery. Consequently, according to this theory the magnetic lines of force move in a clockwise direction when the current is flowing away from the observer. Since, however, as previously stated, we now know that a current consists of a movement of electrons from the negative to the positive pole of a battery, the statement used in text books, while not theoretically correct, is the accepted standard.

Chapter XXXI

ELECTRO-MAGNETIC INDUCTION

By A. P. HILL

1. ACCELERATION AND INERTIA

If a force be applied to some object at rest so that it is started in motion in a uniform straight line, one second after the initial application of such a force the object will have attained a certain velocity which can be measured in centimeters per second. At the end of the next second it will have attained a still greater velocity and this will continue as long as the force continues to be applied. If, for example, its velocity at the end of the first second is found to be two centimeters per second, at the end of the second second it will be traveling at the rate of four centimeters per second, and at the end of the third second it will be at the rate of six centimeters per second. Its velocity is therefore increasing at a rate of two centimeters per second each second. This increase in velocity is called its acceleration and is measured in centimeters per second per second. Thus in the case noted above, the acceleration of the object is two centimeters per second per second. From the above it may be seen that

$$\text{Force} = \text{Mass} \times \text{Acceleration}$$

or $F = MA$

where $F =$ force in dynes

$M =$ mass in grams

and $A =$ acceleration in centimeters per second per second

This may also be stated in the following form:

$$F = M\frac{V}{t}$$

where $V =$ velocity in centimeters

and $t =$ time in seconds

As has been previously discussed, when the current flows in a conductor, electro-magnetic lines of force circle about the conductor in the direction described above. If a current is started in one conductor and another conductor is placed near and parallel to it, the lines of force from the first conductor will cut the second, and due to such cutting, a voltage will be induced in it by what is known as electro-magnetic induction. Figure 294 shows a cross-sectional view of two such conductors.

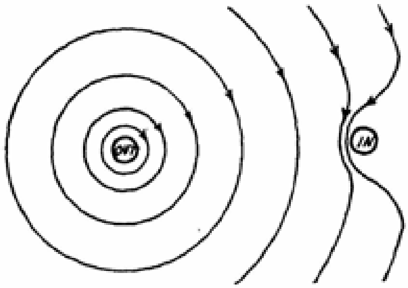


Figure 294 — Lines of force from one conductor cutting a second conductor.

The current flowing in the left hand conductor produces lines of force which circle about it in a clock-wise direction; the electrons are then flowing towards the observer. As the current builds up to its maximum value the lines of force expand farther and farther from it and eventually reach the second conductor, and it may be assumed that they will be bent by it in much the same manner that a ripple on the surface of a body of water would be bent by any obstruction in its path. It will be noticed then, that as these lines of force continue to expand they leave about the second conductor lines of force which circle it in a counter-clock-wise direction. The voltage induced, therefore, in this second conductor must be in the opposite direction to that applied to the first. There is, consequently, a universal law that whenever lines of force cut a conductor, a voltage is induced in it, the amount of voltage depending upon the rate of cutting of the lines of force; i.e., the number which cut the conductor per second. If we assume now that the current in the first conductor, as shown in Figure 295, reaches a maximum value and remains at such value for a period of time, the lines of force surrounding it will be neither expanding nor collapsing. Consequently they will not be cutting through the second conductor and no voltage will be induced in it.

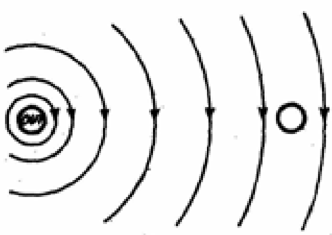


Figure 295 — Lines of force about a conductor in which a steady current is flowing.

If, now, the current in the first conductor is stopped, the lines of force will collapse and as they do so will cut through the second con-

ductor from right to left as indicated in Figure 296, leaving a circle of lines of force about the second conductor which now will be in a clock-wise direction as indicated in the figure. The voltage induced in it will then be in the same direction as that originally applied to the first conductor.

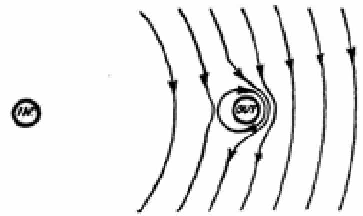


Figure 296 — Lines of force collapsing on a conductor in which the current has been cut off.

The following important law should therefore be remembered:

(a) If two conductors are placed parallel to one another and current is introduced into the first, lines of force will expand from it as the current rises in value, and induce in the second conductor a voltage in the opposite direction to that in the first.

(b) After this current has reached its maximum value, as long as it continues to flow steadily, no voltage will be induced in the second conductor.

(c) If the current in the first conductor is cut off, lines of force will collapse back upon it and induce a voltage in the second conductor in the same direction as that in the first.

In the case of a direct-current circuit, therefore, an induced voltage will be obtained only at the moment of making or breaking the first circuit.

Since, in the case described above, this action is obtained between two separate circuits it is called mutual induction.

2. VALUE OF INDUCED VOLTAGE

It is, of course, important that we be able to calculate the amount of voltage produced by electro-magnetic induction in a conductor. It is found that one volt is induced when 100,000,000 lines of force cut a conductor in a second, or $E = \frac{\phi}{10^8 t}$. Therefore, if 100,000,000 lines of force cut through two conductors in series per second, two volts will be induced. The induced voltage, therefore, may be said to depend upon the number of linkages of lines of force and turns of wire in a coil. Expressing this in the form of an equation

$$E = \frac{\phi N}{10^8 t}$$

3. SELF-INDUCTANCE

By referring to Figure 297 it will be seen that if a current is introduced into a coil the various turns of which are parallel to one another, when the current flows in conductor A the lines of force from it will expand and cut conductor B from right to left, thereby inducing in B a voltage in the opposite direction to that applied to the circuit as a whole. The same is, of course, true in the case of all turns of the coil. Consequently, it will be seen that when an attempt is made to introduce a current in such a circuit, a voltage is induced in each of the turns of the coil, tending to prevent the introduction of such a current. The moment after the current reaches its maximum value, this opposing induced voltage disappears. If, now, the circuit is opened and the current consequently starts to fall in value, the lines of force will cut back through the turns of the coil from left to right, producing a voltage in the same direction as that applied to the circuit, thus tending to prevent the current from falling to zero. This property of self-inductance in a circuit can therefore be defined as the property which tends to prevent the introduction, variation or extinction of a current flowing in the circuit. The induced voltage is sometimes referred to as the back e.m.f. of the coil.

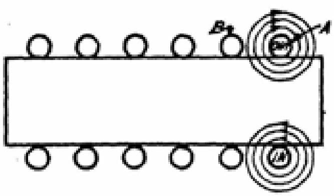


Figure 297 — Cross sectional view of coil in which current is flowing.

4. THE UNIT OF INDUCTANCE — HENRY

In order to be able to calculate the effect of electro-magnetic induction it is necessary that we have a unit of inductance, and this unit has been named the "henry."

The definition of this unit is that it is the inductance of a circuit in which a change of one ampere per second produces an induced e.m.f. of one volt. Since, as stated above, one volt is produced when 100,000,000 lines of force cut one turn of wire per second, the total voltage produced in any instance is equal to the total flux divided by 100,000,000 and also divided by the number of seconds required for such cuttings to take place, or

$$E = L\frac{I}{t}$$

where $\frac{I}{t}$ = the rate of change of current

E = induced voltage

and L = coil inductance in henries

The resemblance between this formula and the one given in Section 1 of this chapter, namely, $F = M\dfrac{V}{t}$ will be immediately noticed; and the inductance L is seen to correspond to the mass M in the mechanical case. It is a fact that inductance bears a close resemblance to inertia and is often referred to as the electrical inertia of a circuit.

Since, as stated above, $E = \dfrac{\phi}{10^8\, t}$ (255)

where a flux cuts only one turn of wire, if there are N turns of wire

$$E = \frac{\phi N}{10^8\, t} \tag{256}$$

but since $E = L\dfrac{I}{t}$ (257)

$$LI = \frac{\phi N}{10^8} \tag{258}$$

from which $L = \dfrac{\phi N}{10^8\, I}$ (259)

But in Chapter XXX, Section 7 on Permeability, it was stated that

$$\mu = \frac{B}{H}$$

or $B = \mu H$

It was also shown that $\phi = BA$

$\therefore\ \phi = \mu HA$

Using this value of ϕ in formula (5),

$$L = \frac{\mu HAN}{10^8\, I}$$

But in Chapter XXX, Section 11, it was shown that

$$H = \frac{1.26\, NI}{l}$$

$$\therefore\ L = \frac{1.26\, N^2 \mu A}{10^8\, l} \tag{260}$$

but, since $1.26 = \dfrac{4\pi}{10}$

and since the area of a circle $= \pi r^2$, equation (260) may be written

$$L = \frac{4\pi^2 N^2 r^2 \mu}{10^9 l}$$

in which $L =$ inductance in henries

$\mu =$ permeability of core

$l =$ length of core in centimeters

$N =$ number of turns of coil

$r =$ radius of core in centimeters

5. CALCULATION OF MUTUAL INDUCTANCE

As previously explained, when two separate circuits act mutually upon one another so that a current flow in one circuit induces a voltage in the second circuit, the effect is due to what is known as mutual induction. The unit of mutual inductance is the same as that of self-inductance, namely, the henry, and is defined as follows:

> "When a change of one ampere per second in one circuit induces one volt in a second circuit, the two circuits have a mutual inductance of one henry."

In order to calculate the amount of mutual inductance that exists between two circuits it is simplest to assume a condition where all the lines of force produced by one circuit cut all the turns of wire of the second circuit. Such a case may be assumed to exist in the circuit shown in Figure 298.

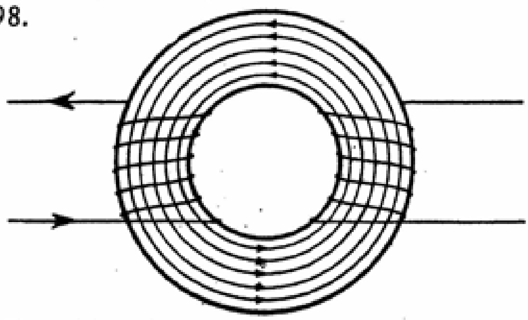


Figure 298 — Mutual inductance between two circuits having unity coefficient of coupling.

In Section 4 it was shown that $H = \dfrac{1.26\, NI}{l}$

Since however $\phi = \mu HA$

$$\phi = \frac{1.26\, NIA\,\mu}{l}$$

$$\text{the flux per ampere} = \frac{\phi}{I} = \frac{1.26\, NA\,\mu}{l}$$

and the flux cuttings of the second coil by a change of one ampere in the first coil equals

$$\frac{1.26\, N_1\, N_2\, \mu\, A}{l}$$

where N_1 = number of turns of wire in the first coil

and N_2 = number of turns of wire in the second coil

Since a mutual inductance of one henry exists when 10^8 flux cuttings are produced by a change of one ampere, the mutual inductance

$$M = \frac{1.26\, N_1\, N_2\, \mu\, A}{10^8\, l}$$

or since $1.26 = \frac{4\pi}{10}$

and $A = \pi r^2$

$$M = \frac{4\pi^2\, N_1\, N_2\, \mu\, r^2}{10^9\, l}$$

6. UNITY COEFFICIENT OF COUPLING

In the case considered above, it was assumed that all the flux set up in one circuit cut all the turns of wire in the second circuit. When this condition exists the two circuits are said to have unity coefficient of coupling. From the formula given above for mutual inductance, namely,

$$M = \frac{1.26\, N_1\, N_2\, \mu\, A}{10^8\, l}$$

it will be noticed that this is equivalent to $\sqrt{L_1 L_2}$

for $L_1 = \frac{1.26\, N_1^2\, \mu\, A}{10^8\, l}$

and $L_2 = \frac{1.26\, N_2^2\, \mu\, A}{10^8\, l}$

$$\therefore \quad M = \sqrt{L_1 L_2}$$

i.e. $\frac{M}{\sqrt{L_1 L_2}} = 1$

This factor $\frac{M}{\sqrt{L_1 L_2}}$ is called the coefficient of coupling and is equal to unity only when all the lines of force from one circuit cut all the turns of wire of the second. When this condition exists the circuits are said to have unity coefficient of coupling.

Chapter XXXII

THE DECIBEL

By A. P. HILL

As in dealing with any other quantity, we need some kind of a unit by means of which we can express the efficiency of an electrical circuit as a medium for the transmission of power in communication work. For the same reason that we need some adopted standard as a unit of length, such as the foot or the meter to measure distance, we require some standard for the measurement of transmission loss or transmission gain in communication engineering. The practical unit that has been chosen is called the decibel.

In Figure 299 is shown a circuit consisting of three elements, each of which may represent either a gain or a loss of power.

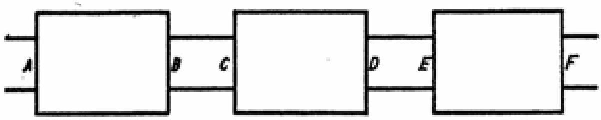


Figure 299 — Transmission losses or gains may be expressed in terms of the decibel.

If we assume that all three elements represent losses, and further, that the power at B is one-tenth that at A, the power at D one-tenth that at C, and the power at F one-tenth that at E, it is obvious that the power at F is one one-thousandth that at A. It is, of course, convenient to have a unit of transmission efficiency that can be treated by the ordinary processes of addition and subtraction in order to obtain the total losses or gains in a circuit. In order to do this it is necessary that the unit be an exponential one and it was for this reason that the decibel was chosen as the unit of transmission efficiency.

The international unit of transmission efficiency is the bel and is defined as the logarithm of the ratios of the powers at the two parts of the circuit under consideration; i.e., the bel $= \log \frac{P_2}{P_1}$,

where $P_2 =$ power at receiving end of circuit

and $P_1 =$ power at the transmitting end of the circuit

The decibel, the unit most often used in practical work, is equal to one-tenth of the bel. Consequently the number of decibels loss or gain in a circuit is expressed as follows:

$$\text{db} = 10 \log \frac{P_2}{P_1} \tag{261}$$

$$= 20 \log \frac{I_2}{I_1} \tag{262}$$

$$= 20 \log \frac{E_2}{E_1} \tag{263}$$

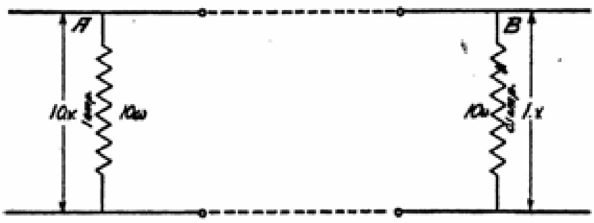


Figure 300 — A transmission loss in power may be expressed in decibels.

If a circuit such as is shown in Figure 300 be considered where an e.m.f. of ten volts forces a current of one ampere through a resistance of ten ohms, it is possible that at a point further along in the circuit a pressure of one volt might result, in which case, a current of one-tenth of an ampere would flow through a ten ohm resistance. Then the power at A is ten watts and the power at B one-tenth of a watt.

According to the above formula, therefore, the transmission loss in the circuit in decibels is calculated as follows:

$$\text{db} = 10 \log \frac{0.1}{10}$$

$$= 10 \log 0.01$$

$$= 20$$

or if we make use of current ratios,

$$\text{db} = 20 \log \frac{0.1}{1}$$

$$= 20 \log 0.1$$

$$= 20$$

Table VI gives the number of db corresponding to various gain or loss ratios. It should be realized that the basis of the decibel as a unit of transmission efficiency is founded on power ratios. Conse-

quently voltage or current ratios can only be used when measured in equal resistances. If unequal resistances are used a correction must be applied in the following manner.

If E_1 and E_2 represent the voltages measured at the transmitting and receiving ends of a circuit, respectively, and if they are not measured across equal resistances, in order to use their ratio as a measure of the transmission efficiency of the system, the voltage E_2 must be multiplied by the square root of $\frac{r_1}{r_2}$, where r_1 is the transmitting resistance and r_2 the receiving resistance, following which correction the new ratio may be used to determine the number of db loss or gain in the circuit. If I_1 and I_2 are the two currents concerned and are also measured in unequal resistances, I_2 must be multiplied by the square root of $\frac{r_1}{r_2}$ before their ratio can be used as a means of calculating the loss or gain of the circuit in db. The following is an example:

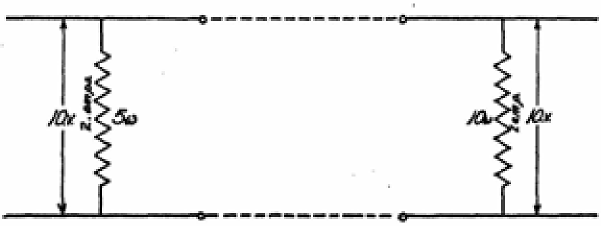


Figure 301 — Voltage and current ratios may be used when calculating db loss or gain, only if measured in equal resistances.

Figure 301 shows a circuit in which the resistances of the transmitting and receiving ends are different. The power at the transmitting end of the circuit is $10 \times 2 = 20$ watts. The power at the receiving end is $10 \times 1 = 10$ watts.

$$
\begin{aligned}
\text{The loss in db} &= 10 \log \frac{10}{20} \\
&= 10 \log 0.5 \\
&= 10 \times \bar{1}.699 \\
&\doteq 3 \text{ db}
\end{aligned}
$$

If, now, in order to calculate the number of db loss, we were to use the voltage ratio shown in the figure without the application of any correction factor, since the pressure at both points in the circuit is ten volts, it would indicate no loss or gain in the circuit, which is obviously

incorrect since the powers at these two points are not equal. If, on the other hand, we were to use the current ratio without correction, we obtain the following:

$$\begin{aligned} db &= 20 \log \frac{1}{2} \\ &= 20 \times \bar{1}.699 \\ &= 6 \text{ db} \end{aligned}$$

which again does not agree with the result obtained when we use the power ratios. If, however, we apply the proper corrections to the voltage and current as explained above, we obtain the following as the voltage ratio:

$$V.R. = \frac{10 \times \sqrt{\frac{5}{10}}}{10}$$

$$\begin{aligned} \therefore \quad db &= 20 \log \frac{10 \sqrt{\frac{5}{10}}}{10} \\ &= 20 \log \frac{10 \times 0.707}{10} \\ &= 20 \log 0.707 \\ &= 20 \times \bar{1}.85 \\ &= 3 \text{ db} \end{aligned}$$

Applying the correction to the current values involved, we obtain the following:

$$\begin{aligned} db &= 20 \log \frac{1 \sqrt{\frac{10}{5}}}{2} \\ &= 20 \log \frac{1 \times 1.414}{2} \\ &= 20 \log 0.707 \\ &= 20 \times \bar{1}.85 \\ &= 3 \text{ db} \end{aligned}$$

Thus, if we apply the proper corrections, voltage or current ratios may be used to obtain the number of db loss in any circuit.

In the above calculations if the correct voltage and current ratios had been obtained the number of db loss could have been read directly from Table VI.

TABLE VI

Relation Between Transmission Units, and Current, Voltage and Power Amplification and Attenuation Ratios

TRANSMISSION UNITS (DB) For Current or Voltage Ratio	For Ratio Power	Amplification Ratio (Gain)	Attenuation Ratio (Loss)	TRANSMISSION UNITS (DB) For Current or Voltage Ratio	For Power Ratio	Amplification Ratio (Gain)	Attenuation Ratio (Loss)
0.1	.05	1.012	0.989	4.1	2.05	1.603	0.624
0.2	.1	1.023	0.977	4.2	2.1	1.622	0.617
0.3	.15	1.035	0.966	4.3	2.15	1.641	0.610
0.4	.2	1.047	0.955	4.4	2.2	1.660	0.603
0.5	.25	1.059	0.944	4.5	2.25	1.679	0.596
0.6	.3	1.072	0.933	4.6	2.3	1.698	0.589
0.7	.35	1.084	0.923	4.7	2.35	1.718	0.582
0.8	.4	1.096	0.912	4.8	2.4	1.738	0.575
0.9	.45	1.109	0.902	4.9	2.45	1.758	0.569
1.0	.5	1.122	0.891	5.0	2.5	1.778	0.562
1.1	.55	1.135	0.881	5.1	2.55	1.799	0.556
1.2	.6	1.148	0.871	5.2	2.6	1.820	0.550
1.3	.65	1.162	0.861	5.3	2.65	1.841	0.543
1.4	.7	1.175	0.851	5.4	2.7	1.862	0.537
1.5	.75	1.188	0.841	5.5	2.75	1.884	0.531
1.6	.8	1.202	0.832	5.6	2.8	1.906	0.525
1.7	.85	1.216	0.822	5.7	2.85	1.928	0.519
1.8	.9	1.230	0.813	5.8	2.9	1.950	0.513
1.9	.95	1.245	0.804	5.9	2.95	1.972	0.507
2.0	1.0	1.259	0.794	6.0	3.0	1.995	0.501
2.1	1.05	1.274	0.785	6.1	3.05	2.018	0.495
2.2	1.1	1.288	0.776	6.2	3.1	2.04	0.490
2.3	1.15	1.303	0.767	6.3	3.15	2.06	0.484
2.4	1.2	1.318	0.759	6.4	3.2	2.09	0.479
2.5	1.25	1.334	0.750	6.5	3.25	2.11	0.473
2.6	1.3	1.349	0.741	6.6	3.3	2.14	0.468
2.7	1.35	1.365	0.733	6.7	3.35	2.16	0.462
2.8	1.4	1.388	0.724	6.8	3.4	2.19	0.457
2.9	1.45	1.396	0.716	6.9	3.45	2.21	0.452
3.0	1.5	1.413	0.708	7.0	3.5	2.24	0.447
3.1	1.55	1.429	0.700	7.1	3.55	2.26	0.442
3.2	1.6	1.445	0.692	7.2	3.6	2.29	0.437
3.3	1.65	1.462	0.684	7.3	3.65	2.32	0.432
3.4	1.7	1.479	0.676	7.4	3.7	2.34	0.427
3.5	1.75	1.496	0.668	7.5	3.75	2.37	0.422
3.6	1.8	1.514	0.661	7.6	3.8	2.40	0.417
3.7	1.85	1.531	0.653	7.7	3.85	2.42	0.412
3.8	1.9	1.549	0.645	7.8	3.9	2.45	0.407
3.9	1.95	1.567	0.638	7.9	3.95	2.48	0.403
4.0	2.0	1.585	0.631	8.0	4.0	2.51	0.398

TABLE VI (Continued)

TRANSMISSION UNITS (DB) For Current or Voltage Ratio	For Power Ratio	Amplification Ratio (Gain)	Attenuation Ratio (Loss)	TRANSMISSION UNITS (DB) For Current or Voltage Ratio	For Power Ratio	Amplification Ratio (Gain)	Attenuation Ratio (Loss)
8.1	4.05	2.54	0.394	12.1	6.05	4.03	0.248
8.2	4.1	2.57	0.389	12.2	6.1	4.07	0.245
8.3	4.15	2.60	0.385	12.3	6.15	4.12	0.242
8.4	4.2	2.63	0.380	12.4	6.2	4.17	0.240
8.5	4.25	2.66	0.376	12.5	6.25	4.22	0.237
8.6	4.3	2.69	0.372	12.6	6.3	4.27	0.234
8.7	4.35	2.72	0.367	12.7	6.35	4.32	0.232
8.8	4.4	2.75	0.363	12.8	6.4	4.37	0.229
8.9	4.45	2.79	0.359	12.9	6.45	4.42	0.226
9.0	4.5	2.82	0.355	13.0	6.5	4.47	0.224
9.1	4.55	2.85	0.351	13.1	6.55	4.52	0.221
9.2	4.6	2.88	0.347	13.2	6.6	4.57	0.219
9.3	4.65	2.91	0.343	13.3	6.65	4.62	0.216
9.4	4.7	2.95	0.339	13.4	6.7	4.68	0.214
9.5	4.75	2.98	0.335	13.5	6.75	4.73	0.211
9.6	4.8	3.02	0.331	13.6	6.8	4.79	0.209
9.7	4.85	3.05	0.327	13.7	6.85	4.84	0.206
9.8	4.9	3.09	0.324	13.8	6.9	4.90	0.204
9.9	4.95	3.13	0.320	13.9	6.95	4.95	0.202
10.0	5.0	3.16	0.316	14.0	7.0	5.01	0.200
10.1	5.05	3.20	0.313	14.1	7.05	5.07	0.197
10.2	5.1	3.24	0.309	14.2	7.1	5.13	0.195
10.3	5.15	3.27	0.305	14.3	7.15	5.19	0.193
10.4	5.2	3.31	0.302	14.4	7.2	5.25	0.191
10.5	5.25	3.35	0.298	14.5	7.25	5.31	0.188
10.6	5.3	3.39	0.295	14.6	7.3	5.37	0.186
10.7	5.35	3.43	0.291	14.7	7.35	5.43	0.184
10.8	5.4	3.47	0.288	14.8	7.4	5.50	0.182
10.9	5.45	3.51	0.285	14.9	7.45	5.56	0.180
11.0	5.5	3.55	0.282	15.0	7.5	5.62	0.178
11.1	5.55	3.59	0.279	15.1	7.55	5.69	0.176
11.2	5.6	3.63	0.275	15.2	7.6	5.75	0.174
11.3	5.65	3.67	0.272	15.3	7.65	5.82	0.172
11.4	5.7	3.72	0.269	15.4	7.7	5.89	0.170
11.5	5.75	3.76	0.266	15.5	7.75	5.96	0.168
11.6	5.8	3.80	0.263	15.6	7.8	6.03	0.166
11.7	5.85	3.85	0.260	15.7	7.85	6.10	0.164
11.8	5.9	3.89	0.257	15.8	7.9	6.17	0.162
11.9	5.95	3.94	0.254	15.9	7.95	6.24	0.160
12.0	6.0	3.98	0.251	16.0	8.0	6.31	0.158

TABLE VI (Continued)

TRANSMISSION UNITS (DB)				TRANSMISSION UNITS (DB)			
For Current or Voltage Ratio	For Power Ratio	Amplification Ratio (Gain)	Attenuation Ratio (Loss)	For Current or Voltage Ratio	For Power Ratio	Amplification Ratio (Gain)	Attenuation Ratio (Loss)
16.1	8.05	6.38	0.157	21	10.5	11.2	0.0891
16.2	8.1	6.45	0.155	22	11.0	12.6	0.0794
16.3	8.15	6.53	0.153	23	11.5	14.1	0.0708
16.4	8.2	6.61	0.151	24	12.0	15.8	0.0631
16.5	8.25	6.68	0.150	25	12.5	17.8	0.0562
16.6	8.3	6.76	0.148	26	13.0	20.0	0.0501
16.7	8.35	6.84	0.146	27	13.5	22.4	0.0447
16.8	8.4	6.92	0.144	28	14.0	25.1	0.0398
16.9	8.45	7.00	0.143	29	14.5	28.2	0.0355
17.0	8.5	7.08	0.141	30	15.0	31.6	0.0316
17.1	8.55	7.16	0.140	31	15.5	35.5	0.0282
17.2	8.6	7.24	0.138	32	16.0	39.8	0.0251
17.3	8.65	7.33	0.136	33	16.5	44.7	0.0224
17.4	8.7	7.41	0.135	34	17.0	50.1	0.0200
17.5	8.75	7.50	0.133	35	17.5	56.2	0.0178
17.6	8.8	7.59	0.132	36	18.0	63.1	0.0158
17.7	8.85	7.67	0.130	37	18.5	70.8	0.0141
17.8	8.9	7.76	0.129	38	19.0	79.4	0.0126
17.9	8.95	7.85	0.127	39	19.5	89.1	0.0112
18.0	9.0	7.94	0.126	40	20.0	100	0.0100
18.1	9.05	8.04	0.124	41	20.5	112	0.00891
18.2	9.1	8.13	0.123	42	21.0	126	0.00794
18.3	9.15	8.22	0.122	43	21.5	141	0.00708
18.4	9.2	8.32	0.120	44	22.0	158	0.00631
18.5	9.25	8.41	0.119	45	22.5	178	0.00562
18.6	0.3	8.51	0.118	46	23.0	200	0.00501
18.7	9.35	8.61	0.116	47	23.5	224	0.00447
18.8	9.4	8.71	0.115	48	24.0	251	0.00398
18.9	9.45	8.81	0.114	49	24.5	282	0.00355
19.0	9.5	8.91	0.112	50	25.0	316	0.00316
19.1	9.55	9.02	0.111	51	25.5	355	0.00282
19.2	9.6	9.12	0.110	52	26.0	398	0.00251
19.3	9.65	9.23	0.108	53	26.5	447	0.00224
19.4	9.7	9.33	0.107	54	27.0	501	0.00200
19.5	9.75	9.44	0.106	55	27.5	562	0.00178
19.6	9.8	9.55	0.105	56	28.0	631	0.00158
19.7	9.85	9.66	0.104	57	28.5	708	0.00141
19.8	9.9	9.77	0.102	58	29.0	794	0.00126
19.9	9.95	9.89	0.101	59	29.5	891	0.00112
20.0	10.0	10.00	0.100	60	30.0	1000	0.00100

TABLE VI (Continued)

TRANSMISSION UNITS (DB) For Current or Voltage Ratio	For Power Ratio	Amplification Ratio (Gain)	Attenuation Ratio (Loss)	TRANSMISSION UNITS (DB) For Current or Voltage Ratio	For Power Ratio	Amplification Ratio (Gain)	Attenuation Ratio (Loss)
61	30.5	1120	0.000891	81	40.5	11200	0.0000891
62	31.0	1260	0.000794	82	41.0	12600	0.0000794
63	31.5	1410	0.000708	83	41.5	14100	0.0000708
64	32.0	1580	0.000631	84	42.0	15800	0.0000631
65	32.5	1780	0.000562	85	42.5	17800	0.0000562
66	33.0	2000	0.000501	86	43.0	20000	0.0000501
67	33.5	2240	0.000447	87	43.5	22400	0.0000447
68	34.0	2510	0.000398	88	44.0	25100	0.0000398
69	34.5	2820	0.000355	89	44.5	28200	0.0000355
70	35.0	3160	0.000316	90	45.0	31600	0.0000316
71	35.5	3550	0.000282	91	45.5	35500	0.0000282
72	36.0	3980	0.000251	92	46.0	39800	0.0000251
73	36.5	4470	0.000224	93	46.5	44700	0.0000224
74	37.0	5010	0.000200	94	47.0	50100	0.0000200
75	37.5	5620	0.000178	95	47.5	56200	0.0000178
76	38.0	6310	0.000158	96	48.0	63100	0.0000158
77	38.5	7080	0.000141	97	48.5	70800	0.0000141
78	39.0	7940	0.000126	98	49.0	79400	0.0000126
79	39.5	8910	0.000112	99	49.5	89100	0.0000112
80	40.0	10000	0.000100	100	50.0	100000	0.0000100

Chapter XXXIII

RESISTANCE ATTENUATION NETWORKS

By A. P. HILL

It is at times necessary to insert in a circuit an arrangement of resistances in order to introduce a definite number of db loss. These resistances may be connected in various ways and are called resistance attenuation networks. There are four types of these networks which are commonly used and they are known as:

(A) *T* type networks
(B) *H* type networks
(C) π type networks
(D) *U* type networks.

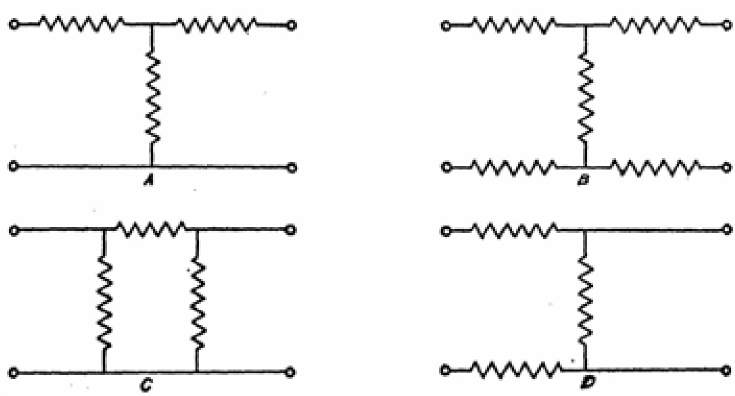


Figure 302 — Resistance attenuation networks — (A) *T* type, (B) *H* type, (C) π type, (D) *U* type.

Figure 302, A, B, C, and D, indicates the arrangement of the resistances used in these four types of networks. The *T* and *H* types are electrically similar as will be shown later.*

* NOTE: Only the *T*, *H*, and *U* type networks will be described in detail. For information regarding the π type network and still more detailed information regarding other types, reference may be made to "*Transmission Circuits for Telephone Communication*," by K. S. Johnson, and "*Communication Engineering*," by W. L. Everitt.

1. *T* TYPE NETWORKS

Networks of the *T* type consist of series and shunt elements. Two different solutions are given below:

Case I, when working between equal resistances;

Case II, when working between unequal resistances.

Figure 303 shows a *T* type network connected between two resistances Z_1 and Z_2.

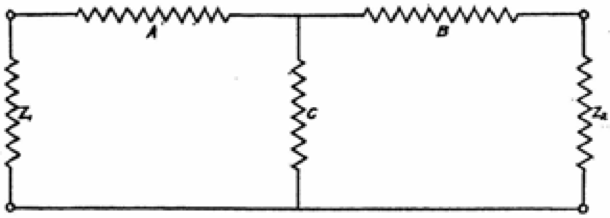


Figure 303 — *T* type network.

Following is a general solution for the two cases mentioned above:

Case I

$$\text{Where } Z_1 = Z_2$$
$$\text{In this case } A = B$$

Since in a circuit having resistances in parallel the current divides inversely as the values of the resistances,

$$\text{if } R = \text{the current ratio } \frac{I_2}{I_1}$$

$$\text{then } R = \frac{C\ (B + Z_2)}{B + C + Z_2} \div (B + Z_2)$$

$$= \frac{C}{B + C + Z_2} \tag{264}$$

The total resistance looking into the network from Z_1 must equal Z_1.

By reference to the figure

$$Z_1 = A + \frac{C\ (B + Z_2)}{B + C + Z_2} \tag{265}$$

Substituting (264) into (265)

$$Z_1 = A + R\ (B + Z_2) \tag{266}$$
$$= A + RB + RZ_2$$

$$\text{but } A = B$$
$$\text{and } Z_1 = Z_2$$
$$\therefore Z_1 = A\ (1 + R) + RZ_1$$

$$\therefore \quad B = A = \frac{Z\,(1 - R)}{1 + R} \tag{267}$$

For convenience, using Z in place of Z_1 or Z_2

From (264) $\quad C = R\,(B + C + Z_2)$

Substituting the value of B from (267)

$$C = R\left[\frac{Z\,(1 - R)}{1 + R} + C + Z_2\right]$$

$$= \frac{RZ\,(1 - R)}{1 + R} + RC + RZ_2 \tag{268}$$

$$\therefore \quad C\,(1 - R) = \frac{RZ\,(1 - R)}{1 + R} + RZ_2$$

$$\therefore \quad C = \left[\frac{RZ\,(1 - R)}{1 + R} + RZ_2\right]\left(\frac{1}{1 - R}\right)$$

but since $\quad Z = Z_2$

$$C = \frac{2\,RZ}{1 - R^2} \tag{269}$$

Case II $\qquad$ Where $Z_1 \gtrless Z_2$

$$\frac{I_2}{I_1} = R = \frac{C}{C + B + Z_2}\sqrt{\frac{Z_2}{Z_1}} \tag{270}$$

$$\frac{I_1}{I_2} = R = \frac{C}{C + A + Z_1}\sqrt{\frac{Z_1}{Z_2}} \tag{271}$$

$$Z_1 = A + \frac{C\,(B + Z_2)}{C + B + Z_2} \tag{272}$$

$$Z_2 = B + \frac{C\,(A + Z_1)}{C + A + Z_1} \tag{273}$$

Substituting (270) into (272)

$$Z_1 = A + R\sqrt{\frac{Z_1}{Z_2}}\,(B + Z_2)$$

$$\therefore \quad A = Z_1 - R\sqrt{\frac{Z_1}{Z_2}}\,(B + Z_2) \tag{274}$$

Substituting (271) into (273)

$$Z_2 = B + R\sqrt{\frac{Z_2}{Z_1}}\,(A + Z_1)$$

$$\therefore \quad B = Z_2 - R\sqrt{\frac{Z_2}{Z_1}}\,(A + Z_1) \tag{275}$$

Substituting (275) into (274)

$$A = Z_1 - R\left(\sqrt{\frac{Z_1}{Z_2}}\right)\left[Z_2 - AR\sqrt{\frac{Z_2}{Z_1}} - R\sqrt{Z_1 Z_2} + Z_2\right]$$

$$= Z_1 - R\sqrt{Z_1 Z_2} + AR^2 + R^2 Z_1 - R\sqrt{Z_1 Z_2}$$

$$A - AR^2 = Z_1 - 2R\sqrt{Z_1 Z_2} + R^2 Z_1$$

$$A = \frac{Z_1(1+R^2) - 2R\sqrt{Z_1 Z_2}}{1-R^2}$$

$$= \frac{2Z_1 - 2R\sqrt{Z_1 Z_2}}{1-R^2} - Z_1 \qquad (276)$$

From (275), $B = Z_2 - R\sqrt{\frac{Z_2}{Z_1}}(A + Z_1)$

Substituting from (274)

$$B = Z_2 - R\sqrt{\frac{Z_2}{Z_1}}\left[Z_1 - BR\sqrt{\frac{Z_1}{Z_2}} - R\sqrt{Z_1 Z_2} + Z_1\right]$$

$$= Z_2 - R\sqrt{Z_1 Z_2} + BR^2 + R^2 Z_2 - R\sqrt{Z_1 Z_2}$$

$$B - BR^2 = Z_2 - 2R\sqrt{Z_1 Z_2} + R^2 Z_2$$

$$\therefore \quad B = \frac{Z_2(1+R^2) - 2R\sqrt{Z_1 Z_2}}{1-R^2}$$

$$= \frac{2Z_2 - 2R\sqrt{Z_1 Z_2}}{1-R^2} - Z_2 \qquad (277)$$

From (270) $C = R\sqrt{\frac{Z_1}{Z_2}}(C + B + Z_2)$

$$C\sqrt{\frac{Z_2}{Z_1}} = CR + BR + Z_2 R$$

Substituting B from (277)

$$C\sqrt{\frac{Z_2}{Z_1}} - CR = R\left[\frac{Z_2 + Z_2 R^2 - 2R\sqrt{Z_1 Z_2}}{1-R^2}\right] + RZ_2$$

$$C\left[\sqrt{\frac{Z_2}{Z_1}} - R\right] = \frac{RZ_2 + R^3 Z_2 - 2R^2\sqrt{Z_1 Z_2} + RZ_2 - R^3 Z_2}{1-R^2}$$

$$= \frac{2RZ_2 - 2R^2\sqrt{Z_1 Z_2}}{1-R^2}$$

$$\therefore \quad C = \frac{2R\sqrt{Z_1 Z_2}}{1-R^2} \qquad (278)$$

For minimum attenuation to match terminal impedances

$$A = 0$$

and $R = \sqrt{\frac{Z_2}{Z_1}} - \sqrt{\frac{Z_2 - Z_1}{Z_1}}$

In both the above cases an *H* type network, such as is shown in Figure 304, may be designed by merely placing half of the values of the *A* and *B* resistances in each series leg. This type of network is used where it is necessary to have both sides of the circuit balanced in respect to one another, the *T* type being used primarily where it is desired to ground one side of the circuit.

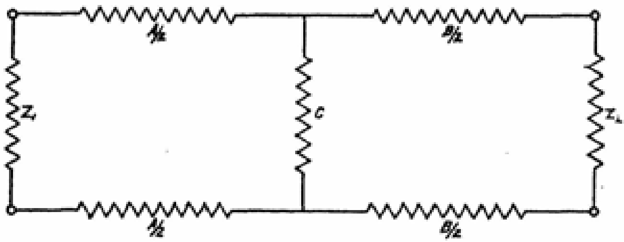


Figure 304 — *H* type network.

2. *U* TYPE NETWORKS

A *U* type network is used primarily where it is necessary to match two circuits which have unequal impedances, and to do so by the interposition of a minimum loss. As will be realized from the description of *T* and *H* type networks, the *U* type network is a limiting case of the *H* type network working between unequal impedances. For this latter type of structure, where a minimum attenuation is required, the series elements on one side of the network become zero. Under this condition the network becomes one of the *U* type.

Figure 305 shows a network of this type. For simplicity of calculation, however, it is better to consider all the series resistances as lumped in one leg of the circuit. The side of the network in which the series element occurs must of course be connected to the higher of the two resistances.

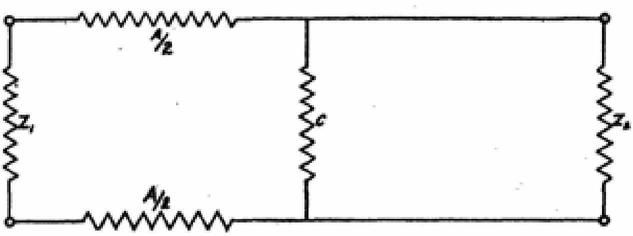


Figure 305 — *U* type network.

A solution of the *U* type network follows:

Where $Z_1 > Z_2$

$$\frac{I_2}{I_1} = R = \frac{C}{C + Z_2}\sqrt{\frac{Z_2}{Z_1}} \qquad (279)$$

$$Z_1 = A + \frac{CZ_2}{C + Z_2} \tag{280}$$

$$Z_2 = \frac{C(A + Z_1)}{C + A + Z_1} \tag{281}$$

From (280) $A = Z_1 - \dfrac{CZ_2}{C + Z_2}$

Substituting from (279)

$$A = Z_1 - \frac{RZ_2}{\sqrt{\dfrac{Z_2}{Z_1}}} \tag{282}$$

From (279) $C = \dfrac{CR + RZ_2}{\sqrt{\dfrac{Z_2}{Z_1}}}$

$$C\sqrt{\frac{Z_2}{Z_1}} - CR = RZ_2$$

$$\therefore \quad C = \frac{RZ_2}{\sqrt{\dfrac{Z_2}{Z_1}} - R} \tag{283}$$

For balanced networks the value of A is equally divided between both series arms. A graphical solution for a minimum loss network of this type is shown in Figure 306.

Artificial Lines for Use Between Unequal Impedances

Minimum Loss "U" Lines

$R = Z_2/Z_1$
$x = XZ_1$
$y = YZ_1$

Formulae:

$X = \frac{\sqrt{1-R}}{2}$

$Y = \frac{R}{\sqrt{1-R}}$

$\text{db loss} = 20 \log_{10} \frac{\sqrt{R}}{1+\sqrt{1-R}}$

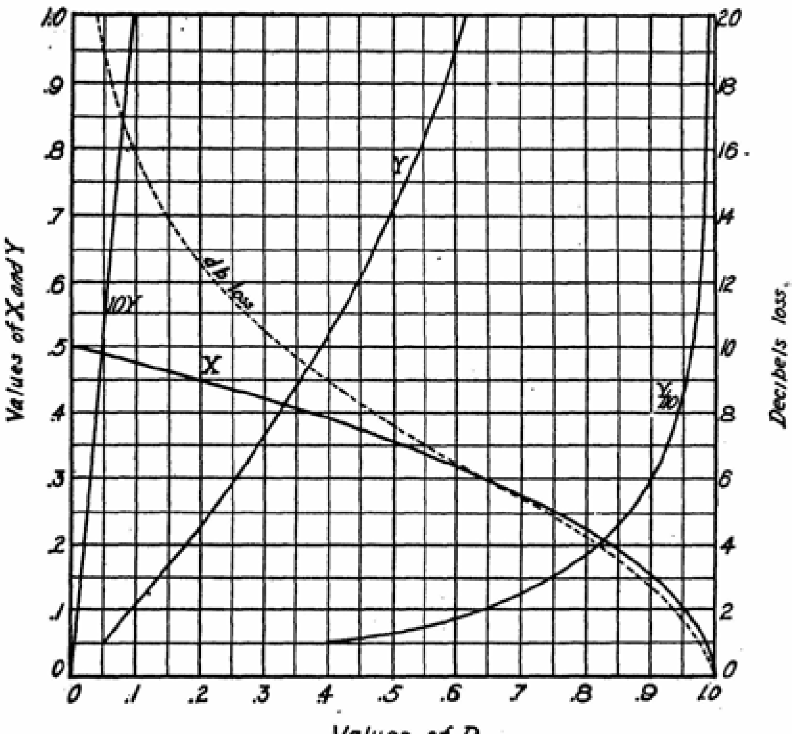


Figure 306.

Chapter XXXIV

GENERATORS AND MOTORS

By A. P. HILL

1. INDUCED ELECTROMOTIVE FORCE

Chapter XXXI described some of the factors relating to electromagnetic induction and it was there explained that whenever electromagnetic lines of force cut a conductor, a voltage was induced in it — its value depending upon the number of linkages per second of the lines of force and turns of wire. It is immaterial whether the conductor remains stationary and the lines of force move, or whether the reverse action takes place. In either case the conductor will be cut by the lines of force and a voltage will result.

The generator depends upon this principle of electromagnetic induction for its operation, the usual procedure being to move the conductors through a stationary magnetic field. Figure 307 shows the elements of such a system. In this figure a north and south pole are placed opposite one another with a magnetic field existing between them as shown.

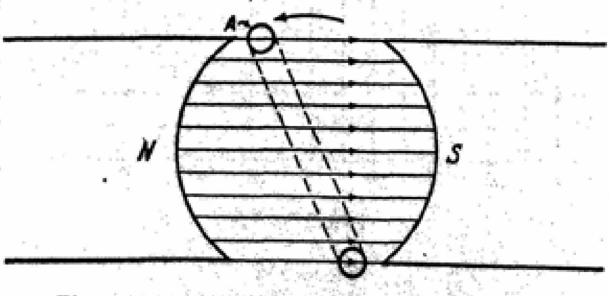


Figure 307 — Elements of a direct-current generator.

"A" represents a cross-sectional view of a conductor which is assumed to be rotated through the magnetic field in the direction shown by the arrow. Consequently, as it rotates, it will cut through the lines of force of the field and a voltage will be induced in it.

A convenient method of remembering in which direction this

voltage will be produced is to use the so-called left-hand rule as illustrated in Figure 308.

In this figure the forefinger represents the direction of the lines of force of the field, i.e., pointing from north to south. The thumb represents the direction of the motion of the conductor and the second finger will then indicate the direction of electron drift in the conductor; i.e., it will point towards the negative end of the conductor.

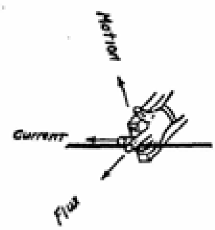


Figure 308 — The thumb indicates the direction of motion of the conductor, the forefinger the direction of flux, and the second finger the direction of the induced e.m.f.

If the conductor mentioned above is an open circuit, a potential will be produced, but there will be no current flow. If, however, its two ends are connected together a current flow (electron drift) will result in the direction indicated by the second finger in the left-hand rule described above.

To summarize the above, we may state that whenever a conductor is made to cut lines of force there will be an e.m.f. induced in it, and the direction of the e.m.f., the direction of the flux and the direction of the motion of the conductor, have a perpendicular relation to one another as shown by the left-hand rule. The amount of induced e.m.f. depends upon the rate of cutting of the lines of force by the conductor, one volt being induced when one conductor cuts one hundred million lines of force per second.

2. E.M.F. INDUCED IN A REVOLVING LOOP

Instead of having a single conductor cutting lines of force a loop of wire may revolve in a magnetic field as shown in Figure 309, so that the conductors *ab* and *cd* move in a direction shown by the arrows. With this direction of rotation, according to the left-hand rule described above, the electron drift in the conductor opposite the north pole will be from *b* to *a* and in the other conductor from *d* to *c*. The drift motion of electrons will therefore take place in the loop in the direction of the arrows.

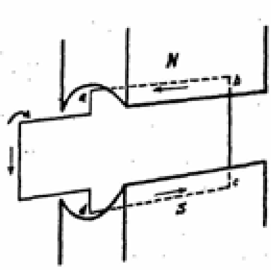


Figure 309 — A loop of wire rotating in a magnetic field.

With such an arrangement we would, of course, have no means of utilizing the voltage and consequent current generated in the loop. It

is, therefore, necessary to have some means of connecting such a loop to an external circuit, which is done by connecting the points *a* and *d* of the loop to two rings mounted on the shaft of the machine on which two brushes make sliding contact and so make possible the connection of the external circuit as shown in Figure 310.

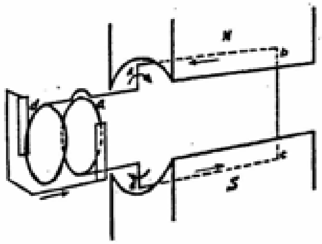


Figure 310 — Method of connecting armature to external circuit.

If the loop now rotates in the direction indicated by the arrow, the electron drift will take place as shown, giving the ring *d* a negative potential and the ring *a* a positive potential.

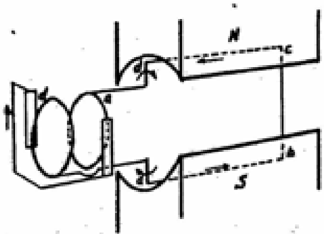


Figure 311—Armature circuit after rotation of 180° from position shown in Figure 310.

If we now consider that this loop of wire has rotated through 180° so that conductor *cd* is moving opposite the north pole and conductor *ab* moving opposite the south pole as shown in Figure 311, the direction of current in the external circuit is seen to have changed so that for each complete revolution of the loop between the magnetic poles there are two reversals of current in the external circuit.

Referring now to Figure 312, it will be seen that since the voltage is induced in the two wires of the loop (which loop is called the armature winding of the machine, the magnetic poles being called field poles, and the two rings to which the armature wires are connected being called slip or connector rings) the generated voltage at any instant will depend upon the angular position of the conductors.

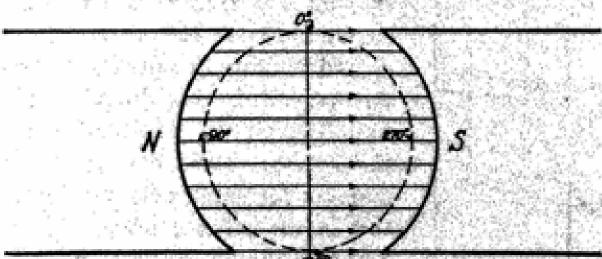


Figure 312 — The rate of cutting of the lines of force varies with the angular position of the armature windings.

When a conductor is at the position indicated as zero degrees it is traveling for an instant parallel to the lines of force and is therefore not cutting through them. Consequently no voltage is induced.

When the conductor is at the position marked 90°, i. e., opposite the center of the pole face, the potential induced in it will be a maximum since it is moving at right angles to the lines of force and is consequently cutting them at a maximum speed. When the conductor gets down to the 180° position, it is again traveling parallel to the lines of force, with the result that there is no voltage induced in it. This state of affairs now reverses and at the 270° position the voltage is a maximum in a direction opposite to that at the 90° position. When it has completely rotated through 360° the voltage has again dropped to zero, following which this cycle of events repeats itself.

The direction and value of the voltages induced, as has just been explained, is shown graphically in Figure 313, and the resultant voltage wave is called a sine wave of alternating current. The reason it is given this name is because the value of the voltage at any instant is equal to the maximum voltage multiplied by the sine of the angle through which the conductor has rotated. A generator producing this type of current is called an alternating-current generator.

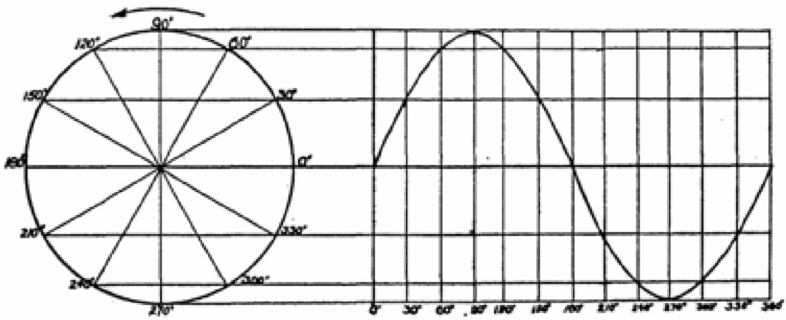


Figure 313 — A sine wave of alternating current.

3. DIRECT-CURRENT GENERATOR

The alternating current produced by the machine described in the last section can be changed to a direct current by a rather simple expedient. If, instead of connecting the two ends of the armature conductor to slip rings, we connect them to the two halves of a split ring as shown in Figure 314, the current in the external circuit will flow continuously in one direction, for as may be seen in Figure 314, the top brush will always be connected to the conductor that is moving opposite the north pole, and the bottom brush to the conductor opposite the south pole.

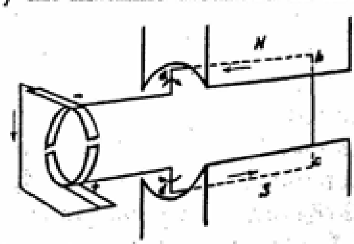


Figure 314 — Simple direct-current generator.

As a result the current in the external circuit would be as shown in Figure 315. It illustrates the principle involved in the operation of the direct-current generator. There are several different types of armature windings, information concerning which should be obtained by reference to any standard textbook on the subject.

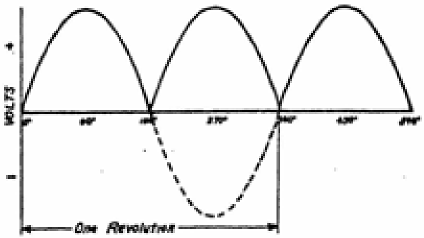


Figure 315 — Pulsating direct current produced by two-pole generator.

Instead of the type of armature winding shown in Figure 314, which, although it results in a form of direct current, nevertheless is of a type that does not have a particularly steady value as is indicated in Figure 315, it is possible by using a larger number of armature windings spaced evenly around the armature core, as shown in Figure 316, to produce a direct current of the type shown in Figure 317. Here it will be noticed that the pulsations are much smaller in amount and by adding still more windings to the armature it is possible to produce a direct current which will be practically steady, having only a very small ripple component.

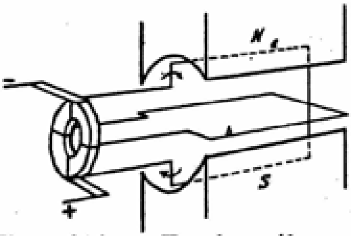


Figure 316 — Two-loop (four-segment commutator) generator.

4. METHOD OF CONNECTING FIELD WINDINGS

Since the field windings of a generator must be supplied with current in order to produce the required magnetic flux between the field poles, it is obvious that they may be supplied with this current either from a separate source of supply, such as a battery or another generator, or may be supplied from the armature of the machine itself. If the first of these two methods is used, the machine is called a separately-excited generator, and if the second method is used it is a self-excited generator.

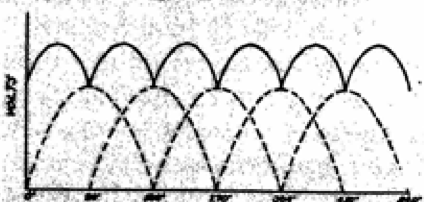


Figure 317 — E.m.f. resulting from generator shown in Figure 316.

It is obvious that if the machine is of the self-excited type, the field windings may be connected either in series or parallel with the

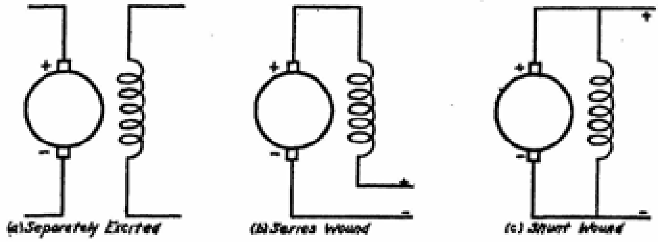


Figure 318 — Direct-current generator (a) separately-excited, (b) series-wound, and (c) shunt-wound.

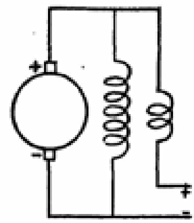


Figure 319 — Compound-wound direct-current generator.

armature windings. Figure 318 shows in schematic form a separately-excited machine and also self-excited of both the series- and shunt-wound types.

Still another type of machine is one that makes use of both a series and a shunt winding. This is called a compound-wound machine and is shown schematically in Figure 319. The operation of these various types of generators is described below.

5. THE SEPARATELY-EXCITED GENERATOR

The separately-excited generator is not in very general use. Since the voltage produced by any generator depends upon the number of lines of force which the armature windings cut per second, this voltage may be controlled by varying the speed of rotation of the armature, or by keeping the speed of rotation constant and varying the density of the magnetic flux between the field poles. The customary practice is to vary the density of the magnetic flux between the field poles, which is accomplished by varying the amount of current flowing in the field windings by the insertion of a variable resistance, called a field rheostat, in series with them. A schematic circuit of such an arrangement is shown in Figure 320.

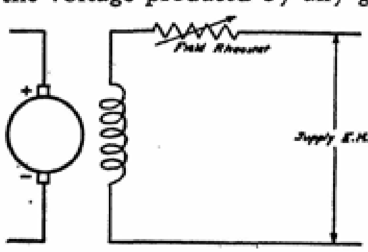


Figure 320 — Schematic circuit of separately-excited generator with field rheostat connected.

6. SELF-EXCITED GENERATORS

(a) Series-wound Generator

A series-wound generator is shown in schematic form in Figure 318 (b). In a machine of this type the total current that flows in the external circuit of the machine also flows through the field windings. Consequently, these windings consist of a few turns of relatively large sized

wire. If now the resistance of the load circuit is decreased, resulting in an increase of current, this increased current will also flow through the field windings, creating a denser magnetic field and a resultant rise in generated voltage; thus the voltage generated by a series-wound dynamo varies with the load.

(b) Shunt-wound Generator

A schematic circuit of a shunt-wound generator is also shown in Figure 318 (c). In this type of machine if the current flowing in the circuit to which power is being delivered is increased, a greater drop in voltage results at the brushes. Since the field windings are connected across these brushes the voltage reduction will result in the flow of a smaller field current. This will cause the magnetic field between the poles to be less dense and the voltage produced by the machine will drop. It will therefore be seen that the shunt-wound generator acts in exactly the opposite manner to the series-wound machine insofar as generated voltage is concerned.

(c) Compound-wound Generator

A compound-wound generator, such as is shown in Figure 319, is equipped with both a series and shunt winding. Its operation, therefore, is affected by both these windings, the series winding tending to cause a rise in voltage with an increased current, the shunt winding tending to cause a drop. The result is that if the number of turns of both these windings is of the correct value, the machine will maintain a constant voltage with a load which may vary considerably.

Compound-wound generators can be designed by varying the number of turns of the two field windings so that they may have either a rising or falling voltage characteristic with increasing load.

7. MOTORS

Any generator will run as a motor if supplied with the necessary driving voltage, there being no fundamental differences in design although certain modifications may have to be made to obtain the greatest efficiency.

In order to understand the principle upon which a motor operates, the following points already discussed should be remembered:

When a current flows through a wire, lines of force surround the wire in a clockwise direction if the electron motion in the wire is towards

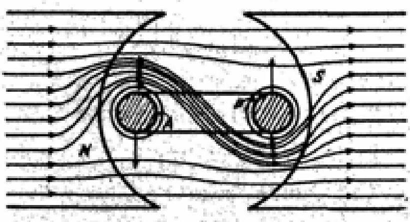


Figure 321 — The lines of force surrounding the armature winding distort those of the field, producing a mechanical pressure on the armature.

the observer. The density of the lines of force surrounding the wire depends upon the amount of current flowing through it. In Figure 321 the lines of force of the field are moving from the north to the south pole.

If the electrons in conductor A are moving towards the observer, the lines of force surrounding it will be as shown in a clockwise direction. These will distort the lines of force of the field, bending them in the manner shown, and thus resulting in a downward pull on conductor A.

The electrons in conductor B will, of course, be moving away from the observer. Consequently the lines of force surrounding it will be moving in a counter-clockwise direction and the distortion of the field lines of force will be as shown, and will result in an upward pull on B. This will produce a counter-clockwise rotation of the armature. If the direction of either the lines of force of the field or the armature current is reversed, the direction of rotation of the motor will be changed, but if, however, both of these are reversed, the direction of rotation will remain the same.

8. MOTOR FIELD WINDINGS

While in a motor it is necessary to supply an e.m.f. to both the field and armature windings, only one source of supply is necessary, and, as in the case of the generator, the machine may be either series, shunt, or compound wound. Figure 322 shows in schematic form the connections for these three types of machines. The method of speed control for each will next be considered.

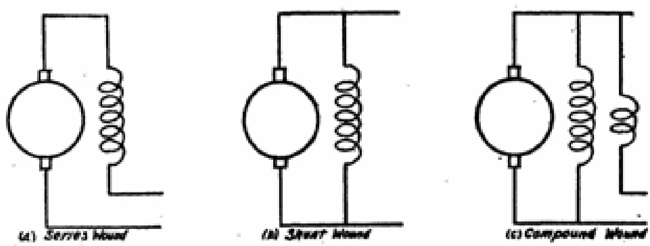


Figure 322 — Schematic circuit of motor, (a) series-wound, (b) shunt-wound, (c) compound-wound.

9. MOTOR SPEED CONTROL

When current is applied to the armature and field of a motor the magnetic flux is produced as described above and the machine will start to rotate in a direction determined by the direction of current in the two circuits. Immediately this happens, however, it will be seen

that the armature conductors start to cut through the lines of force of the field and as a result a voltage will be generated in these conductors which, as will be realized from application of the left-hand generator rule, will be in the opposite direction to that which is applied and its value will increase as the speed of rotation of the machine increases. Theoretically this would continue until the back e.m.f. became equal to the applied e.m.f., at which time no current would flow through the armature, the lines of force surrounding the armature conductors would disappear and the machine would slow down. Immediately after the speed was reduced the back e.m.f. would drop and consequently the lines of force surrounding the armature would again appear, resulting in a speeding up of the machine to the point where these two e.m.f.'s again became equal. In actual practice, however, this does not occur, since even without any mechanical load being applied to the motor shaft, a certain amount of power is required to overcome the frictional resistance of the bearings, etc. Consequently, the back e.m.f. will always be slightly less than the applied e.m.f.; the difference between the two resulting in the generation of an amount of power just sufficient to do the work required to keep the machine in rotation. This back e.m.f. of a motor is one of the most important factors to be considered in its operation.

From the preceding paragraph it will be seen that a motor will always run at such speed that the difference between the back e.m.f. and applied e.m.f. is sufficient to perform the work required of the motor.

The speed with which the armature must rotate in order to generate this back e.m.f. will, of course, depend upon the density of field flux. If this density is small, the armature will have to rotate faster in order to cut the required number of lines of force per second than would be necessary if a denser field flux were produced. Consequently, a simple method of controlling the speed of rotation of such a machine is to control the current flowing in the field windings by means of a field rheostat and thus control the density of the field flux.

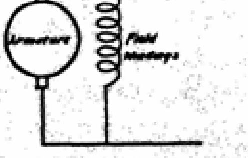


Figure 323—Shunt-wound motor with field rheostat.

Such a rheostat is shown in Figure 323. If its resistance is increased, it will reduce the current flow in the field coils, reducing the field flux density and resulting in an increase of speed of rotation of the machine. Decreasing this resistance would, of course, cause the motor to slow down.

10. SERIES-WOUND MOTOR

In the case of a series-wound motor, since all the current taken by the armature also flows through the field windings, these windings will of necessity consist of a few turns of large sized wire.

If we now consider the operation of such a machine, it will be seen that if it is rotating under light mechanical load (in which case the back e.m.f. will be almost equal to the applied), and the mechanical load is increased, the machine will tend to slow down. This will result in a decreased back e.m.f. and a consequent increase in the actual amount of current flowing in the armature windings. Since this current also flows through the field windings, an increase in flux density of the field will result. This will cause the motor to slow down still more. The net result then is that as the mechanical load is increased on a series-wound motor, it results in a considerable slowing down of its speed of rotation.

11. SHUNT-WOUND MOTOR

In the case of a shunt-wound motor as shown in Figure 323, the action under load is different. If, as before, we assume a condition under which the motor is rotating under light load, the generated back e.m.f. will be almost equal to the applied. If now the mechanical load is increased, the motor will tend to slow down, and this will result in a reduction in potential difference at the brushes. Since the field windings are connected across these brushes, it will tend to reduce the amount of current flowing in these windings, reduce the density of field flux, and consequently the armature will have to rotate faster in order to produce the required back e.m.f. Within certain limits, therefore, such a machine tends to speed up under increased load.

12. COMPOUND-WOUND MOTOR

In a compound-wound motor, since there are both series and shunt windings, any tendency to decreased speed due to the series winding will be offset by the tendency to increase due to the shunt winding, and if the ratio of the two windings is correctly chosen the machine may be designed to operate at practically constant speed under varying loads, or it may be designed to give either a slightly increased or slightly decreased speed as required.

13. CAUTION IN THE USE OF SERIES AND SHUNT MOTORS

From the above discussion it will be realized that if the field circuit of a shunt-wound motor were opened during running, it would rapidly

increase in speed in trying to produce the required back e.m.f. and since the lines of force of the field never entirely disappear due to residual magnetism, the machine would tend to run away and damage would probably result.

A series-wound motor should never be started without any load being applied to the armature. Neither should the load be removed while the machine is running. In either instance it would tend to run away with resulting damage.

14. SOME POINTS OF COMPARISON BETWEEN A MOTOR AND A GENERATOR

As explained above, when a motor starts to rotate it tends to act as a generator, due to the fact that the armature windings cut through the lines of force of the field and generate a voltage. Similarly a generator will tend to act as a motor due to the fact that the moment a current starts to flow in the armature windings lines of force are produced which surround these windings and distort the lines of force of the field with the result that the machine tends to rotate in the opposite direction to that in which it is being turned as a generator. Such a reversal of motion never actually takes place. However, there is a back mechanical action in the motor corresponding to the back e.m.f. produced in a generator. This back mechanical action makes it harder to rotate the motor when it is delivering mechanical power than is the case when it is being rotated under no load conditions. Thus the larger the amount of mechanical power that the machine is delivering the greater must be the amount of electrical power supplied to it.

Chapter XXXV

ALTERNATING CURRENTS

By A. P. HILL

Up to the present time we have considered currents in which the electron flow is always in one direction: the analogy that has been used is that of the steady flow of water in one direction through a pipe. It is obvious, however, that since a current flow of electricity consists of a drift motion of electrons through a wire, that under certain conditions the direction of drift might change, being first in one direction and then in the opposite. A current of this type is called an alternating current.

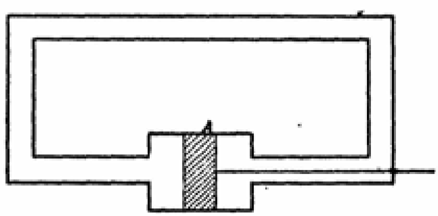


Figure 324 — Water analogy for an alternating current.

In Figure 324 is shown a mechanical analogy in which *A* represents a water pump that is capable of forcing the water either to the right or left. When the piston of the pump moves to the right a flow of water results in that direction which will start from rest, build up to a maximum velocity and die down to rest when the piston has reached its point of maximum displacement. When the reverse piston motion takes place, the flow of water is reversed, again starting from rest, rising to a maximum in this reversed direction and finally coming to rest, after which this cycle of events repeats itself so that there is produced a flow of water reversing in direction and also changing in actual velocity. Such a flow of water could be called an alternating flow. In Figure 325 is shown the corresponding electrical circuit. The generator is capable of producing voltages that reverse in direction at periodic intervals. As a result the flow of electrons through the circuit will also reverse and an alternating current will flow.

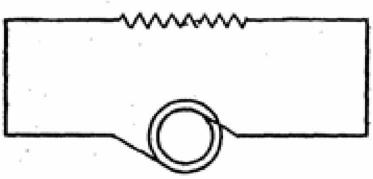


Figure 325 — Alternating-current generator connected to a resistance.

1. FREQUENCY

A current of the type described above can reverse itself quite slowly or the reversals may occur many times per second. The number of complete reversals that take place every second is called the frequency of the current. Current used for ordinary electric lighting purposes has a frequency of 60 cycles per second. Since each complete reversal consists of a flow in one direction followed by a similar flow in the reverse direction, each complete reversal of current is called a cycle, each cycle consisting of two alternations in opposite directions. If therefore the frequency of the current is 60 cycles per second there will be actually 120 alternations comprising the 60 cycles.

2. GENERATION OF ALTERNATING CURRENT

Figure 326 shows in simplified form the elements of an alternating-current generator. The conductor *A* is made to rotate in a counter-clockwise direction opposite the face of the north and south poles of the field magnets. Let us assume that the rotation starts from the position shown and that it requires one second to complete a revolution. In the position shown the conductor will be traveling parallel to the lines of force of the field, and therefore will not be cutting through them. Consequently no voltage will be induced in it. When it has rotated through an angle of 30° it will be cutting these lines of force at an angle and consequently a voltage will be induced, the value of which will depend upon the number of lines of force cut per second. (It will be remembered that if a conductor cuts one hundred million lines of force per second one volt will be induced in it.) As the angle through which the armature rotates increases, it will cut through the lines of force faster and faster until it reaches the 90° position, at which time, since it is moving through these lines of force at right angles, it will be cutting them at a maximum speed and consequently will have the maximum voltage induced in it. After the 90° position is passed it will be cutting the lines of force at a sharper angle; the voltage will gradually drop until it reaches the 180° position when it will again be traveling parallel to the lines of force and the voltage will consequently have dropped to zero.

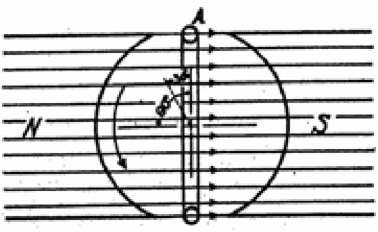


Figure 326 — Simple alternating-current generator.

By applying the left-hand rule described in Chapter XXXIV, Section 1, it will be seen that the electron flow in the armature conductor A will be away from the observer. However, when it leaves the 180° position and starts to move across the face of the south pole the induced voltage will be reversed in direction and the electron motion will then be towards the observer, its value starting from zero at 180°, rising to a maximum in this direction at 270° and falling to zero when it has completed one revolution.

It is now possible to plot the value and direction of voltage generated by a machine of this type in the manner shown in Figure 327.

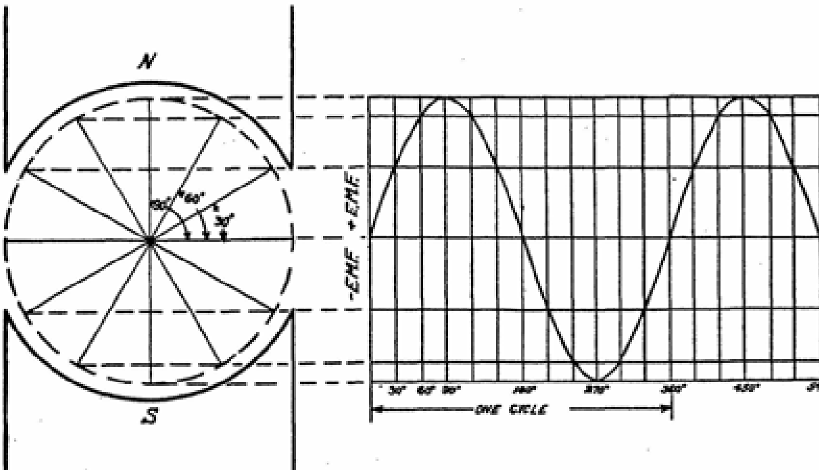


Figure 327 — A sine wave of alternating current.

Here the same direction of rotation is assumed as in Figure 326, the curve above the zero axis marked "positive" representing a voltage generated in one direction and the portion below the zero axis marked "negative" representing a voltage in the opposite direction. The height of the curve above the zero axis represents the amount of voltage generated at any instant. As is shown by this curve the maximum voltage in the positive direction is reached at the 90° position, the corresponding maximum in the reverse direction occurring at 270°. A curve of the type shown is called a sine curve due to the fact that the voltage value at any instant is equal to the maximum value multiplied by the sine of the angle through which the armature has rotated at the instant under consideration.

3. THE VALUE OF ALTERNATING CURRENTS

From a consideration of the previous section it will be realized

that there are several ways in which the value of an alternating current can be specified. They are as follows:

(a) The maximum value,

(b) The instantaneous value,

(c) The effective value,

(d) The average value.

(a) The Maximum Value

The maximum value of an alternating current is reached twice in each cycle for the type of generator shown, namely at the 90° and 270° positions. This is the peak value which the voltage attains for any given generator.

(b) The Instantaneous Value

The instantaneous value of a current is, as its name indicates, the voltage existing in the machine at the instant under consideration. As explained above, it is equal to the maximum value multiplied by the sine of the angle through which the armature conductor has rotated up to the instant under consideration. This angle is called the phase angle and is designed by the Greek letter θ. The instantaneous value of voltage is usually written E_{inst}. Consequently, if the maximum value is written E_{max}

$$E_{inst} = E_{max} \sin \theta$$

(c) The Effective Value

Probably the most useful method of describing the value of an alternating current is its effective value.

When a direct current flows through a circuit it will do a definite amount of work. In the case of an ordinary resistance it will generate a certain amount of heat in the wire. The effective value of an alternating current is the value of current that will generate the same amount of heat in a circuit as is produced by a similar amount of direct current. Its value is equal to the maximum value multiplied by 0.707.

The letter E followed by no subscript always represents the effective value of the voltage. Consequently $E = 0.707\ E_{max}$.

(d) The Average Value

The average value of an alternating current is as its name indicates, the average of all the instantaneous values of the current during a half cycle. Its value is equal to 0.636 of the maximum, or

$$E_{av} = 0.636\ E_{max}$$

It is obvious that the average of the instantaneous values during a half cycle only must be considered since the average of all the instantaneous values over a whole cycle duration must equal zero.

4. OHM'S LAW FOR ALTERNATING CURRENTS

Corresponding to the direct-current circuit, Ohm's Law for alternating-current circuits may be expressed as follows:

$$I_{\text{max}} = \frac{E_{\text{max}}}{R}$$

$$I_{\text{inst}} = \frac{E_{\text{inst}}}{R}$$

$$I_{\text{av}} = \frac{E_{\text{av}}}{R}$$

$$I\,(\text{eff}) = \frac{E\,(\text{eff})}{R}$$

5. INDUCTIVE REACTANCE

By reference to Chapter XXXI, Section 3, on inductance, it will be seen that when a current is introduced into a coil of wire as the value of the current rises in the circuit a voltage is induced in a direction opposite to that of the applied voltage. This effect is due to what is known as the self-inductance of the circuit. Similarly, when the current in a circuit decreases in value, a voltage is induced in the same direction as that applied, tending to keep the current flowing in the circuit. As long as the current is at a steady value however this effect is not produced.

In the case of an alternating current however, its value is never steady but is always either increasing or decreasing, first in one direction and then in the other. Consequently this effect is always present and the self-inductance of a coil of wire becomes a matter that must be considered at all times when an alternating voltage is applied to it. This effect tends to limit the amount of current that flows and must therefore be taken into consideration in all cases involving alternating-current calculations. The effect that is produced in the coil is termed reactance and is measured in ohms. Since the value of the back e.m.f. produced in such a coil when an alternating voltage is applied to it depends, as previously explained, on the number of lines of force cut per second by each turn of the coil, it will be affected not only by the number of turns of wire of the coil and the density of the flux which cuts it but also by the number of times the current changes in direction per second;

i.e., the number of cycles per second of the alternating current. The symbol used to describe the reactance of such a coil is X_L where X indicates reactance and the sub L shows that it is due to the inductance of the coil. Its value numerically is as follows:

$$X_L = 2\,\pi\, fL$$

where X_L = reactance in ohms
f = frequency in cycles per second
L = inductance in henries

It is now necessary to consider the manner in which the current in such a coil will flow when an alternating voltage is applied to it.

Figure 328 shows a sine wave of alternating current flowing through such a coil. At the instant of zero current, marked A, the rate of change is a maximum, as can be seen by considering the change between points A and B as compared to the extremely small change occurring between points C and D. The lines of force will therefore be expanding faster at point A than at any other non-similar point on the curve. This means that the back e.m.f. induced will be a maximum at this time. At point C on the curve no change of current is momentarily taking place, consequently the induced voltage will be zero. At point E the induced voltage will again be a maximum. We can consequently draw the induced e.m.f. curve as shown.

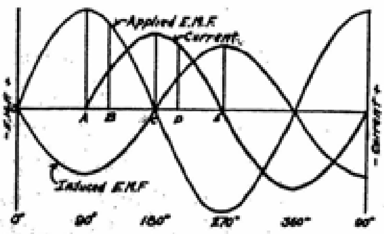


Figure 328 — Voltage and current curves produced when alternating e.m.f. is applied to an inductance coil.

Since the induced e.m.f. is opposite in phase to the applied e.m.f. it will be seen that the current lags behind the applied e.m.f. by an angle of 90°. Such a current is consequently called a lagging current and where reactance only is considered the angle of lag is 90°.

6. CAPACITIVE REACTANCE

A condenser acts somewhat similarly to a coil of wire when an alternating current is applied to it, in that it also impedes the current flow, the effect being due to what is known as capacitive reactance. This may be understood by considering the following analogy, which corresponds rather closely to the action of a condenser. If air is forced into a small rubber balloon, as the sides of the balloon

are distended they exert a pressure tending to prevent the flow of air into it. The greater the amount of air that flows into the balloon the greater will be the back pressure exerted. If the pressure forcing the air into the balloon be gradually raised to maximum and then reduced, the moment it starts to reduce, the back pressure will force the air out of the balloon. A similar action takes place in the case of an electrical condenser. At the moment the pressure is applied there is no back e.m.f. and the current consequently flows into the condenser with a minimum amount of impedance. As the current rises in value the back pressure builds up and reaches a maximum at the moment that the applied voltage has dropped to zero. From an inspection of Figure 329 it will be seen that the current leads the voltage by 90°. This is called a leading current and reaches an angle of 90° only when no other factor than reactance is considered.

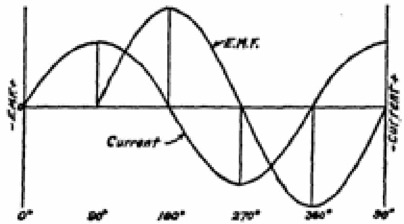


Figure 329 — A capacitive reactance causes the current to lead the voltage by 90°.

As in the case of inductive reactance its value is measured in ohms. Its numerical value is as follows:

$$X_C = \frac{1}{2\,\pi f C}$$

where X_C = capacitive reactance

f = frequency in cycles per second

C = capacity in farads

We thus have two forms of reactance occurring in alternating-current circuits, both of which impede the flow of current and must be considered in all cases where inductance or capacity are present. The manner in which reactance values combine with pure resistance values will be described in the following section.

7. INDUCTANCE AND RESISTANCE IN SERIES

When inductance alone is present in an alternating-current circuit the current will lag behind the voltage by 90° as explained in Section 5. If resistance alone is present the current will be in phase with the voltage. If, on the other hand, inductance and resistance are both present the current will lag behind the voltage by some angle less than 90°, the exact angle being determined by the relative amount of inductive reactance and resistance present.

When reactance and resistance values are combined in a circuit, the total opposition that they offer to the flow of current is termed impedance and is designated by the letter Z. The total impedance may be calculated as follows:

$$Z=\sqrt{R^2+X_L^2}$$

$$\text{and since} \quad X_L = 2\pi fL$$

$$\therefore \quad Z=\sqrt{R^2+(2\pi fL)^2}$$

The reason that resistance and reactance values are not combined arithmetically is, of course, due to the fact that they act 90° out of phase with one another and must consequently be combined vectorially in a manner that will be described more in detail later.

8. CAPACITY AND RESISTANCE IN SERIES

If capacity alone is present in a circuit it causes the current to lead the voltage by 90°. When resistance and capacity are both present they must be combined in a manner similar to that described in the last section in the case of resistance and inductance. The total effect is also termed impedance and may be expressed mathematically as follows:

$$Z=\sqrt{R^2+X_C^2}$$

$$\text{but since} \quad X_C=\frac{1}{2\pi fC}$$

$$Z=\sqrt{R^2+\left(\frac{1}{2\pi fC}\right)^2}$$

9. INDUCTANCE, CAPACITY, AND RESISTANCE IN SERIES

When inductance and capacity are both present in a circuit it is apparent that since inductance causes the current to lag behind the voltage by 90° and capacity causes the current to lead the voltage by 90°, these two act 180° out-of-phase with one another and so tend to cancel one another. The individual reactances must therefore be subtracted from one another in order to determine the total reactance. The resulting impedance may be stated as follows:

$$Z=\sqrt{R^2+(X_L-X_C)^2}$$

$$\text{or} \quad Z=\sqrt{R^2+\left(2\pi fL-\frac{1}{2\pi fC}\right)^2}$$

10. GENERAL LAW FOR ALTERNATING-CURRENT CIRCUITS

It will be obvious from the foregoing that since the combination of resistance, inductance and capacity results in a factor termed imped-

ance that the current in a circuit of this type must be equal to the effective voltage divided by the total impedance of the circuit:

$$\text{or} \quad I = \frac{E}{Z}$$

$$\therefore \quad E = IZ$$

$$\text{and} \quad Z = \frac{E}{I}$$

The solution of alternating-current circuits using the above formulae will of course give the amount of voltage, current or impedance in the circuit, but will give no information as to the phase angle of the current; i.e., whether it leads, lags, or is in-phase with the voltage. The solution to this type of problem will be given in subsequent sections.

Chapter XXXVI

VECTOR NOTATION

By A. P. HILL

If a force be applied to an object in motion or at rest, the amount of force so applied may be expressed in any convenient system of units. The value given would define the amount of force applied but would give no indication of the direction in which it was exerted. It is possible, however, to make use of an expression that will describe both the amount of force and its direction of application. Such a quantity is called a vector.

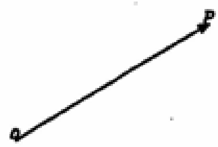


Figure 330 — OP is a vector indicating amount and direction of force.

In Figure 330 the length of the line OP may represent the amount of force and the angle at which this line is drawn may represent the direction in which the force is applied. OP is then a vector. Any vector such as OP may be considered to be made up of two components at right angles to one another, with OP as the hypotenuse of the right angle triangle thus formed. Figure 331 shows such a triangle in which the sides R and X are the two components acting at right angles to one another and Z the resultant of these two forces.

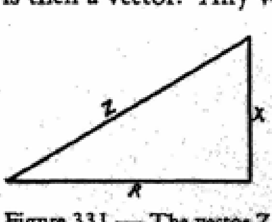


Figure 331 — The vector Z is made up of two components, R and X at right angles to one another.

The magnitude of OP =

$$Z = \sqrt{R^2 + X^2}$$

It now becomes necessary to have some means of describing at what angle Z is applied. To do this we make use of the letter j, which represents a rotation of 90° counter-clockwise of any quantity to which it is attached. Thus, the expression $\overline{Z} = R + jX$ indicates that $\overline{Z}$ is a vector whose component parts are R and X and since j is associated with the X quantity, it indicates that X has been rotated 90° in a counter-clockwise direction from R. When used in this manner the quantity

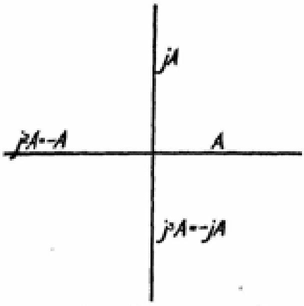


Figure 332—The operator j indicates a rotation in a counter-clockwise direction of 90°.

is called an "operator" and its meaning has already been described. Referring now to Figure 332, if A is a vector applied in the direction shown, jA represents this vector when rotated 90° in a counter-clockwise direction.

$j(jA)$ or j^2A represents a counter-clockwise rotation of 180° from the original position. If the plus sign represents a force applied in one direction and the minus sign represents a force applied in the opposite direction, then $j^2A = -A$. j^3A will then represent a counter-clockwise rotation of 270° and is equal to $-jA$. The positions of these various vectors are indicated in Figure 332.

From the above:

$$\begin{aligned} j\ (jA) &= j^2A \\ &= -A \\ &= (-1)\ A \\ \therefore\quad j^2 &= -1 \\ \text{and}\quad j &= \sqrt{-1} \end{aligned}$$

From the above:

$$\sqrt{R^2 + X^2} = R + jX$$

where $\sqrt{R^2 + X^2}$ gives only the magnitude of a vector, but $R + jX$ gives both the magnitude and angular position.

Referring to Figure 333 and by application of the fundamental laws of trigonometry

$$\begin{aligned} R &= Z \cos\theta \\ \text{and}\quad X &= Z \sin\theta \\ \therefore\quad \overline{Z} &= R + jX \\ &= Z\cos\theta + jZ\sin\theta \\ &= Z\ (\cos\theta + j\sin\theta) \end{aligned}$$

thus the magnitude of the vector equals Z and the angle equals $(\cos\theta + j\sin\theta)$.

This is usually written $Z \underline{/\theta}$ or $Z \overline{/\theta}$ where $\underline{/}$ represents a positive angle and $\overline{/}$ represents a negative angle. In this connection an

inductance is said to result in a positive angle and a capacitance in a negative angle.

In making computations with j and non j terms, it should be remembered that j terms can be combined algebraically only with other j terms. j terms can be combined with non j terms only geometrically.

From the above it will be seen that vectors may be expressed in either the $R + jX$ or $Z \underline{/\theta}$ form.

It may also be shown that the $\underline{/\theta}$ portion of the vector is an exponential factor and should be treated as such in all calculations.

Vectors may be added, subtracted, multiplied, divided, etc., and these various methods of treatment are explained below.

1. ADDITION OF VECTORS

In order to add vectors to one another they should first be expressed in the $R + jX$ form and treated as follows:

$$(R_1 + jX_1) + (R_2 + jX_2) = (R_1 + R_2) + j\,(X_1 + X_2)$$

2. SUBTRACTION OF VECTORS

In order to subtract vectors from one another they should first be expressed in the $R + jX$ form and treated as follows:

$$(R_1 + jX_1) - (R_2 + jX_2) = (R_1 - R_2) + j\,(X_1 - X_2)$$

3. MULTIPLICATION OF VECTORS

In order to multiply two vectors together they should first be expressed in the $Z \underline{/\theta}$ form and the magnitudes (moduli) should be multiplied and the angles added.

Thus $Z_1 \underline{/\theta_1} \times Z_2 \underline{/\theta_2} = Z_1 Z_2 \underline{/\theta_1 + \theta_2}$

4. DIVISION OF VECTORS

In order to divide one vector by another they should be expressed in the $Z \underline{/\theta}$ form, the magnitudes divided and the angles subtracted, as

$$\frac{Z_1 \underline{/\theta_1}}{Z_2 \underline{/\theta_2}} = \frac{Z_1}{Z_2} \underline{/\theta_1 - \theta_2}$$

5. CONVERSION

From the above it is apparent that it will be frequently necessary to convert vectors from the $R + jX$ to the $Z \underline{/\theta}$ form and vice-versa. This may be performed as follows:

Referring to Figure 333:

$$\theta = \tan^{-1}\frac{X}{R}$$

$$Z = \frac{R}{\cos\theta} \text{ or } \frac{X}{\sin\theta}$$

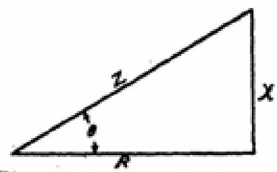


Figure 333 — Application of vectors to impedance, resistance and reactance values.

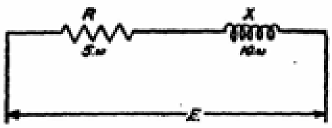


Figure 334 — A resistance of five ohms in series with an inductance of ten ohms.

Example: From Figure 334

$$\tan\theta = \frac{10}{5} = 2$$

$$\therefore \quad \theta = 63° \; 25'$$

$$Z = \frac{R}{\cos\theta} = \frac{5}{0.447} = 11.2$$

$$\therefore \quad 5 + j10 = 11.2 \underline{/63° \; 25'}$$

Vectors in the $Z \underline{/\theta}$ form may be converted into the $R + jX$ form as follows:

$$R = Z\cos\theta$$

$$\text{and} \quad X = Z\sin\theta$$

Referring again to Figure 334

$$R = 11.2 \cos 63° \; 25' = 11.2 \times 0.447 = 5$$

$$X = 11.2 \sin 63° \; 25' = 11.2 \times 0.894 = 10$$

An example illustrating the addition of vectors is as follows:

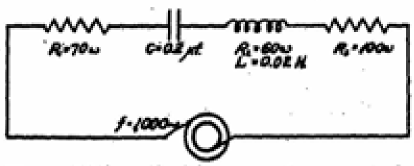


Figure 335 — Resistances and reactances in series with one another.

To find the impedance of the circuit shown in Figure 335

$$\overline{Z} = (R_1 + R_2 + R_3) + j\,(X_1 + X_2)$$
$$= (70 + 60 + 100) + j\,(X_1 + X_2)$$

$$\text{but} \quad X_1 = \frac{1{,}000{,}000}{2\pi fC}$$
$$= -796 \text{ ohms}$$
$$X_2 = 2\pi fL$$
$$= 125.6 \text{ ohms}$$
$$\therefore \; Z = (70 + 60 + 100) + j\,(-796 + 125.6)$$
$$= 230 + j\,(-670.4)$$

In order to express this vector in the $Z \angle \theta$ form, proceed as follows:

$$\tan\theta = \frac{-670.4}{230}$$
$$= -2.9$$
$$\therefore \quad \theta = -71^\circ$$
$$Z = \frac{R}{\cos\theta}$$
$$= \frac{230}{0.3256}$$
$$= 709 \text{ ohms.}$$
$$\therefore \quad \overline{Z} = 709 \,\overline{\angle 71^\circ}$$

Following is another example of this type of calculation. To find the total impedance of the circuit shown in Figure 336.

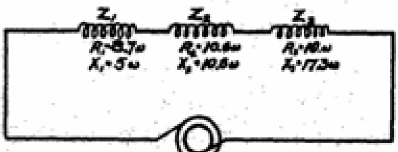


Figure 336 — Three resistances and reactances in series with one another.

$$\overline{Z} = (R_1 + R_2 + R_3) + j\,(X_1 + X_2 + X_3)$$
$$= (8.7 + 10.6 + 10) + j\,(5 + 10.6 + 17.3)$$
$$= 29.3 + j\,(32.9)$$

Expressing this vector in the $Z \angle \theta$ form:

$$\theta = \tan^{-1}\frac{X}{R}$$
$$= \tan^{-1}\frac{32.9}{29.3}$$
$$= \tan^{-1} 1.12$$
$$= 48^\circ$$

$$Z=\frac{R}{\cos\theta}$$

$$=\frac{29.3}{0.6691}=44 \text{ ohms}$$

$$\therefore \quad \bar{Z}=44\,\underline{/\,48^\circ}$$

In the foregoing example, note that $\bar{Z}_1=10\,\underline{/\,30^\circ}$

$$\bar{Z}_2=15\,\underline{/\,45^\circ}$$

$$\bar{Z}_3=20\,\underline{/\,60^\circ}$$

By reference to Figure 337, the foregoing mathematical proof is shown geometrically.

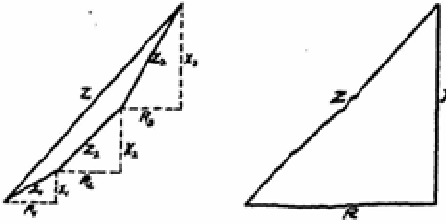


Figure 337—Geometrical proof of Figure 336.

Example showing multiplication and division of vectors:

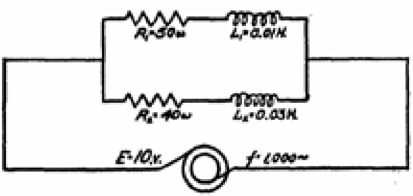


Figure 338 — Multiplication and division of vectors.

Figure 338 shows an example where it is necessary to consider both multiplication and division of vectors. In this case it is necessary to solve for two parallel circuits. In the ordinary direct-current case the total resistance of two resistances in parallel, $R=\dfrac{R_1\times R_2}{R_1+R_2}$ where R_1 and R_2 are the two resistances which are connected in parallel and R represents the total resistance of the combination.

A similar solution is made in the case of an alternating-current circuit of the type shown in Figure 338 where, if $Z\,\underline{/\,\theta}$ represents the total impedance of the combination, $Z_1\,\underline{/\,\theta_1}$ the impedance of the upper portion, and $Z_2\,\underline{/\,\theta_2}$ that of the lower, then

$$Z\,\underline{/\,\theta}=\frac{Z_1\,\underline{/\,\theta_1}\times Z_2\,\underline{/\,\theta_2}}{Z_1\,\underline{/\,\theta_1}+Z_2\,\underline{/\,\theta_2}}$$

Since quantities in the numerator must be multiplied and those in the denominator added, it is necessary to express the numerator in the $Z\,\underline{/\,\theta}$ form and the denominator in the $R+jX$ form. This may be done as follows:

$$\theta_1=\tan^{-1}\frac{X_1}{R_1}$$

$$= \tan^{-1} \frac{62.8}{50}$$

$$= \tan^{-1} 1.255$$

$$= 51^\circ\ 27'$$

$$Z_1 = \frac{R_1}{\cos \theta_1}$$

$$= \frac{50}{0.6232}$$

$$= 80.3 \text{ ohms}$$

$$\therefore \quad Z_1 \underline{/\theta_1} = 50 + j62.8 = 80.3 \underline{/51^\circ\ 27'}$$

Similarly $$\theta_2 = \tan^{-1} \frac{188.4}{40}$$

$$= \tan^{-1} 4.71$$

$$= 78^\circ\ 1'$$

$$Z_2 = \frac{40}{0.2076}$$

$$= 193 \text{ ohms}$$

$$\therefore \quad Z_2 \underline{/\theta_2} = 40 + j188.4 = 193 \underline{/78^\circ\ 1'}$$

$$Z \underline{/\theta} = \frac{Z_1 Z_2 \underline{/\theta_1 + \theta_2}}{(R_1 + R_2) + j\,(X_1 + X_2)}$$

$$= \frac{80.3 \times 193 \underline{/51^\circ\ 27' + 78^\circ\ 1'}}{(50 + 40) + j\,(62.8 + 188.4)}$$

$$= \frac{15{,}500 \underline{/129^\circ\ 28'}}{90 + j251}$$

In order to divide the denominator into the numerator it must now be put in the $Z \underline{/\theta}$ form as follows:

$$90 + j251 = \frac{R}{\cos \theta} \underline{/\tan^{-1} \frac{R}{X}}$$

$$= \frac{90}{\cos\left(\tan^{-1} \frac{251}{90}\right)} \underline{/\tan^{-1} \frac{251}{90}}$$

$$= \frac{90}{0.3374} \underline{/70^\circ\ 17'}$$

$$= 267 \underline{/70^\circ\ 17'}$$

$$Z\,\underline{/\theta} = \frac{15{,}500\,\underline{/129^\circ\,28'}}{267\,\underline{/70^\circ\,17'}}$$

$$= \frac{15{,}500}{267}\,\underline{/129^\circ\,28' - 70^\circ\,17'}$$

$$= 58\,\underline{/59^\circ\,11'}$$

From the foregoing it will be seen that this method of solution is not only quicker than that shown in the previous chapter, but in addition gives the phase angle as well as the numerical value of the impedance.

Chapter XXXVII

RESONANT CIRCUITS

By A. P. HILL

In ChapterXXXV on the subject of the impedance of circuits having both inductance and capacity in series, it was pointed out that the total reactance was equal to the difference between the individual reactances, since they act 180° out-of-phase with one another. If in any circuit these two reactances happen to be equal to one another, the net reactance will be zero and the total impedance of the circuit will be equal to the resistance value. When this condition occurs a state of resonance is said to exist and the current flowing in the circuit will be the maximum that it is possible to obtain.

If a condition of resonance exists, then

$$X_L = X_C$$

$$\text{or} \quad 2\pi fL = \frac{1}{2\pi fC}$$

and solving this equation for f it will be seen that

$$f = \frac{1}{2\pi\sqrt{LC}}$$

where $L =$ inductance in henries

$C =$ capacity in farads

If L is expressed in microhenries and C in microfarads, this equation becomes:

$$f = \frac{3 \times 10^8}{1884.96\sqrt{LC}}$$

Figure 343, Page 498, shows graphically the impedance of such a circuit when the conditions are as shown. The impedance and phase angle are shown for two conditions where $R = 0$ and where $R = 500$ ohms.

Figure 344, Page 498, shows, for a similar circuit, the curve representing the current in the circuit when the resistance is varied.

Figure 345, Page 499, shows the effect of varying the inductance on the current in the circuit. It will be noted here that the point at which resonance occurs varies with change of inductance.

Figure 346, Page 499, shows the effect of varying the capacity in the circuit. As in the last case the resonant frequency changes with any variation in capacity. From the foregoing curves it will be realized that any circuit of this type will have some frequency at which it becomes resonant. This is obvious from the fact that if we take any circuit containing inductance and capacity in series and start with some low frequency, gradually raising this frequency to a high value, the inductive reactance will gradually increase, while the capacitive reactance decreases. Consequently there must be some frequency at which these two become equal. This is called the resonant frequency.

It may be well at this point to review one or two of the conditions occurring in alternating current circuits.

Figure 339 — Resistance only.

$$I = \frac{E}{Z} \qquad Z = R \qquad E = IZ$$

Figure 340 — Resistance and inductance in series.

$$I = \frac{E}{Z} \qquad E = IZ \qquad E = E_r + E_L$$

$$Z = \sqrt{R^2 + X_L^2} \qquad \theta = \tan^{-1} \frac{X_L}{R}$$

Figure 341 — Resistance and capacity in series.

$$I = \frac{E}{Z} \qquad E = IZ \qquad E = E_r + E_C$$

$$Z = \sqrt{R^2 + X_C^2} \qquad \theta = \tan^{-1} \frac{X_C}{R}$$

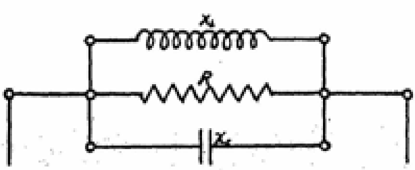


Figure 342 — Resistance, inductance, and capacity in parallel

If we now consider a simple form of parallel circuit as shown in Figure 342:

$$E = IX_L = IR = IX_C$$
$$\bar{I} = I_L + I_r + I_C$$
$$= \sqrt{(I_L - I_C)^2 + I_r^2}$$

If we now wish to find the impedance of a circuit of the type shown in Figure 342 it may be solved as follows:

$$\text{Let } E = 12 \text{ volts}$$
$$R = 2 \text{ ohms}, X_C = 3 \text{ ohms}, X_L = 4 \text{ ohms}$$
$$\therefore \quad I_r = 6 \text{ amps.}, I_C = 4 \text{ amps.}, I_L = 3 \text{ amps.}$$
$$I = \sqrt{I_r^2 + (I_L - I_C)^2}$$
$$= \sqrt{6^2 + 1^2} = \sqrt{37} = 6.0828 \text{ amps.}$$
$$Z = \frac{E}{I} = \frac{12}{6.0828} = 1.97 \text{ ohms}$$
$$\theta = \tan^{-1}\left(\frac{I_L - I_C}{I_r}\right) = \tan^{-1}\left(\frac{-1}{6}\right) = \tan^{-1}(-0.167) = /\overline{9^\circ 27'}$$

Another method of solution for this circuit is as follows:

$$\bar{Z}_C = 0 - j3 = 3\,/\overline{90^\circ} \qquad \bar{Z}_L = 0 + j4 = 4\,\angle 90^\circ$$

$$Z_{CL} = \text{impedance of } C \text{ and } L \text{ in parallel}$$

$$\bar{Z}_{CL} = \frac{12\,\angle 0^\circ}{0 + j1} = \frac{12}{1}\,\frac{\angle 0^\circ}{\angle 90^\circ} = 12\,/\overline{90^\circ}$$
$$= 0 - j12$$

The impedance of Z_{CL} and R in parallel

$$\bar{Z} = \frac{24\,/\overline{90^\circ}}{2 - j12} = \frac{24\,/\overline{90^\circ}}{12.2\,/\overline{80^\circ 33'}}$$
$$= 1.97\,/\overline{9^\circ 27'} \text{ as above.}$$

Still another method follows:

$$\bar{Z} = \frac{1}{\frac{1}{2} + j\left(\frac{1}{3} - \frac{1}{4}\right)} = \frac{1}{\frac{1}{2} + j\frac{1}{12}}$$
$$= \frac{1}{0.5 + j\,0.0834} = \frac{1}{0.507\,\angle 9^\circ 27'}$$
$$= 1.97\,/\overline{9^\circ 27'} \text{ as above}$$

Figure 347, Page 500, shows the distribution of voltages which exist in a resonant parallel circuit.

Figure 348, Page 500, shows the distribution of currents in an anti-resonant circuit, and Figure 349, Page 501, the total current in the external circuit is an arrangement of this type.

Figure 350, Page 501, shows the impedance values occurring in an anti-resonant circuit.

Only the simplest forms of parallel resonant circuits have been mentioned above. It will be realized that there are many possible combinations of such circuits and it will consequently be impossible to treat of them all. For further information on this subject reference should be made to standard textbooks such as *"Radio Engineering,"* by Terman, etc.

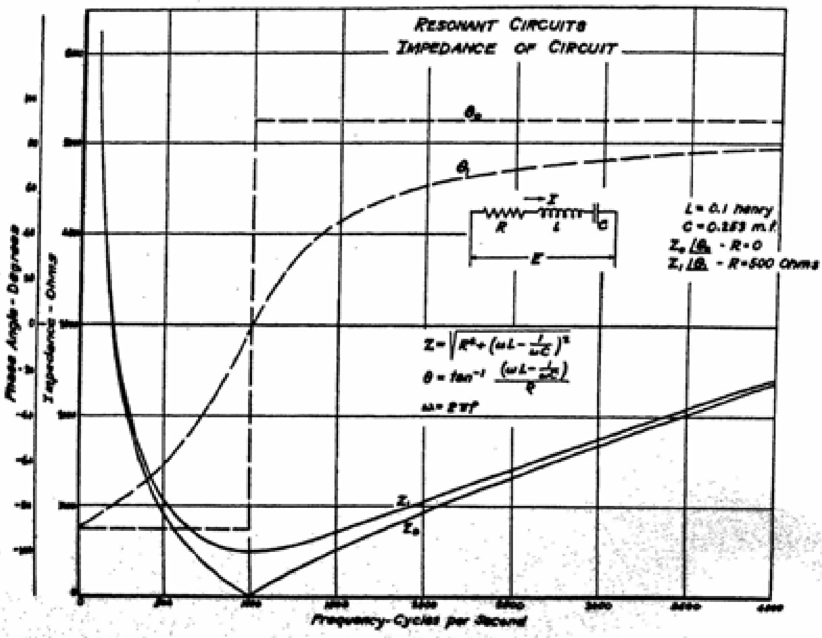


Figure 343.

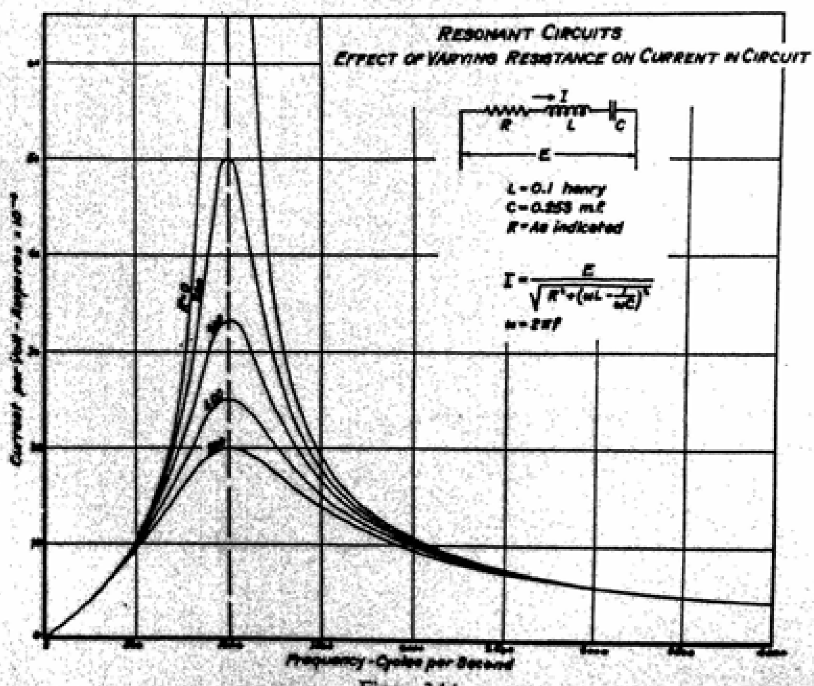


Figure 344.

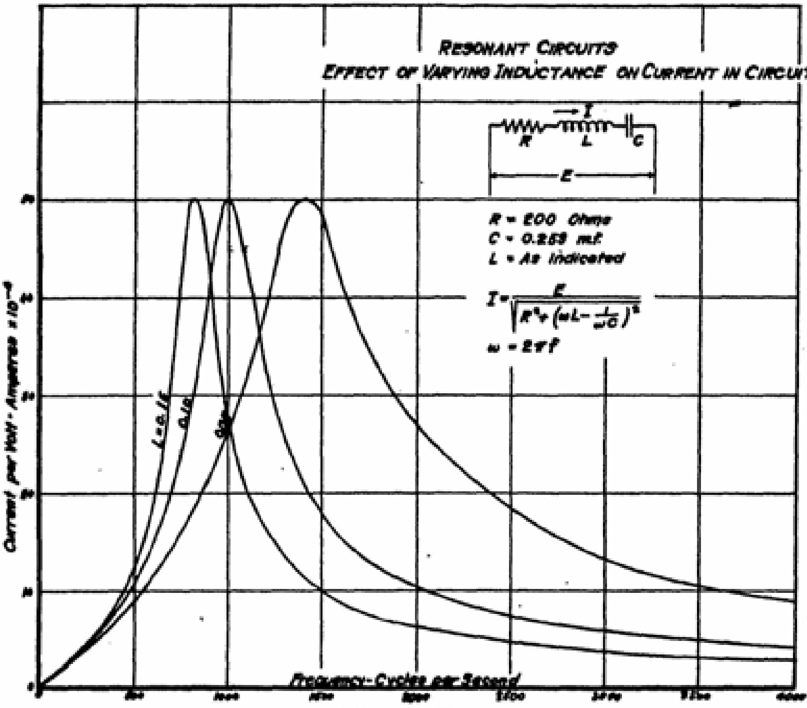


Figure 345.

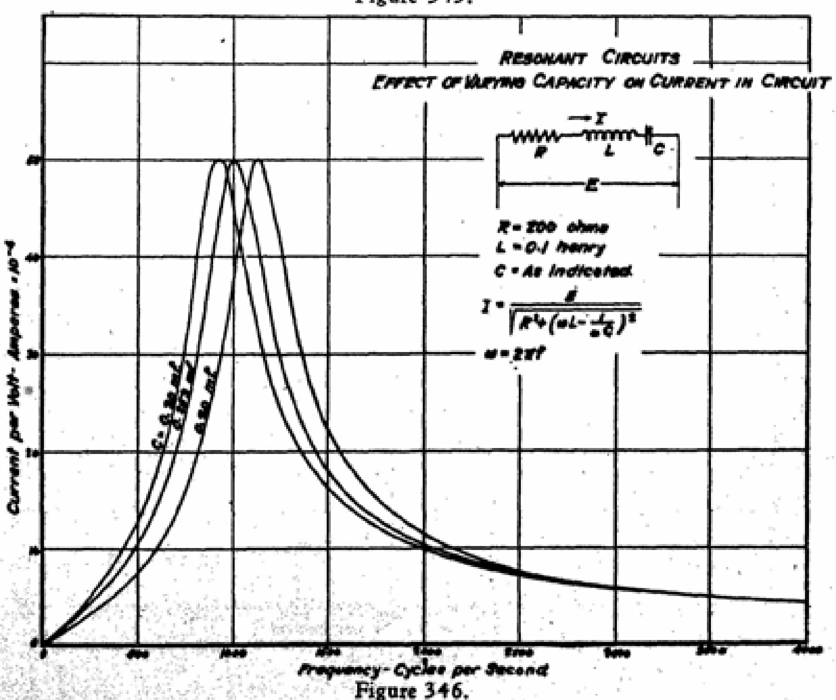


Figure 346.

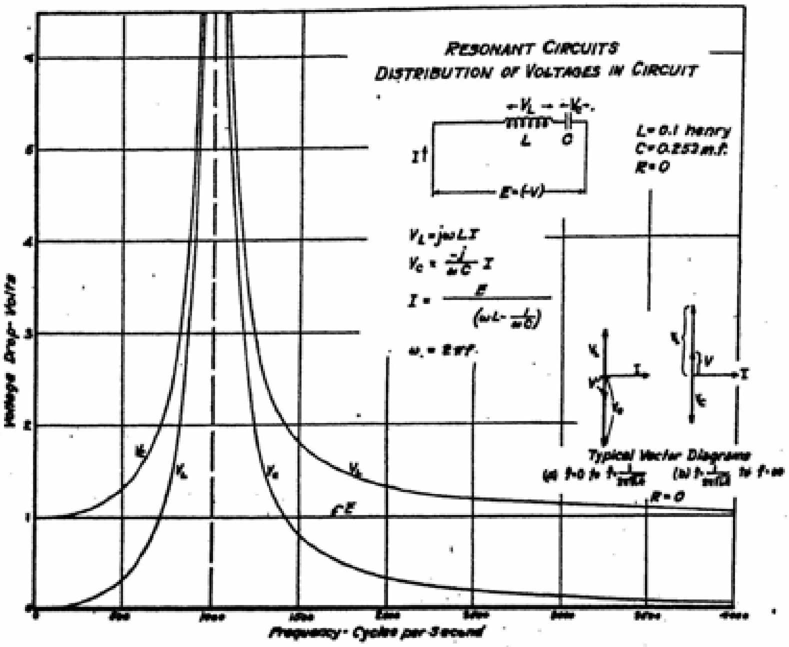


Figure 347.

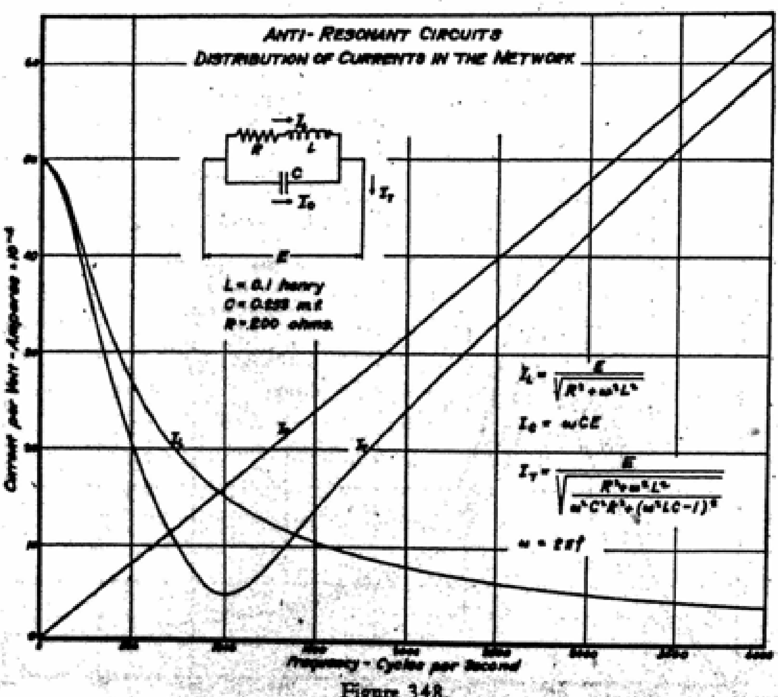


Figure 348.

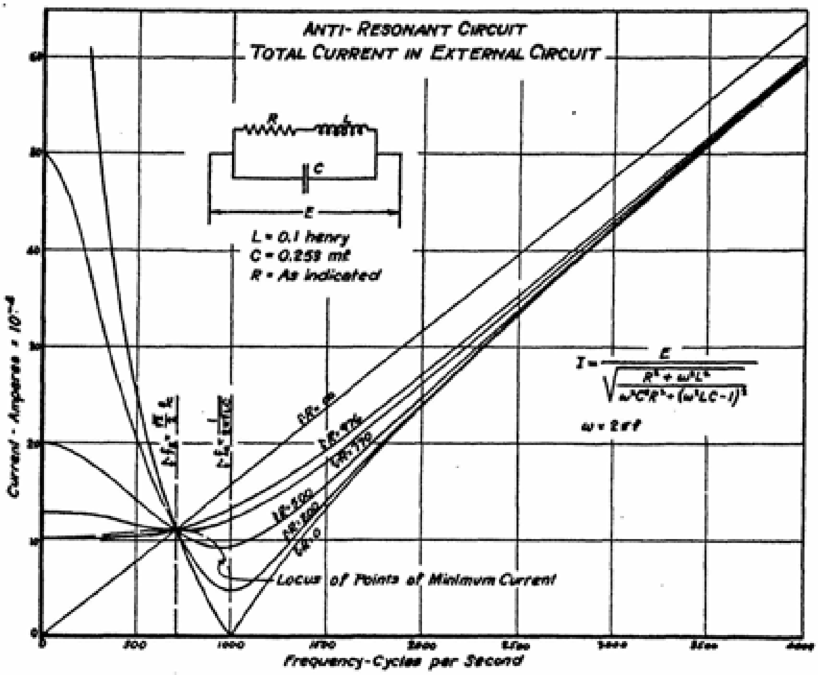


Figure 349.

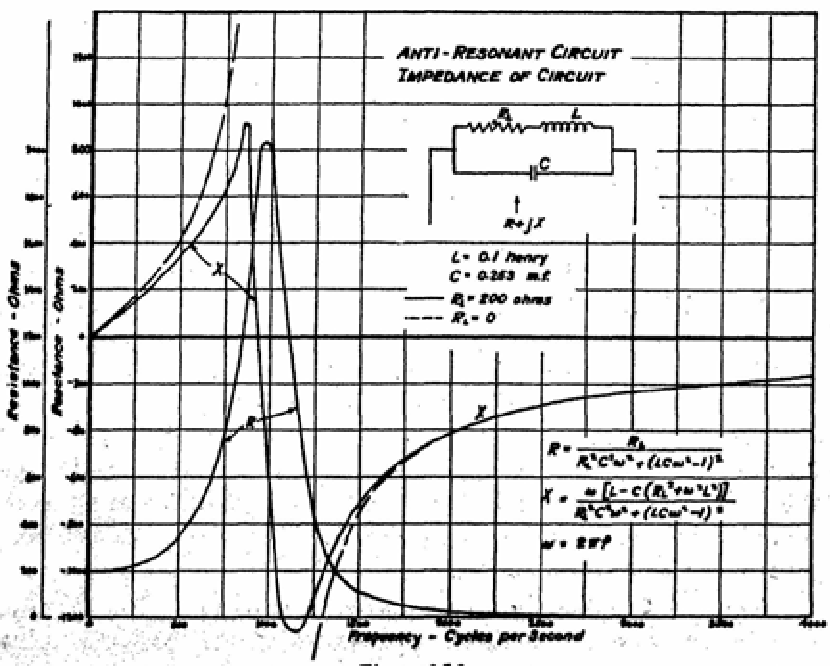


Figure 350.

Chapter XXXVIII

VACUUM TUBES

By A. P. HILL

It was explained in Chapter XXVI that a current flow of electricity consists of a drift motion of electrons through a conductor in a direction from the negative to the positive pole of the source of supply.

It is not necessary that this electron motion take place through a physical conductor. It can equally well occur through a vacuum and still constitute a current flow of electricity. Electrons can not of course be made to pass through air except at extremely high potentials, in which case an electric spark occurs. An extreme condition of this type is a flash of lightning which is the result of the passage of electrons through the air due to an extremely high potential being produced, usually between a cloud and some other object, the accompanying thunder being due to the air disturbance created by the passage of these electrons through it.

In order to understand the factors producing a flow of electrons through a vacuum, it may be helpful to consider an analogy which possesses some points of resemblance to the electrical case. If a body of water be gradually raised to the boiling point, as the temperature of the water increases the molecules of water move within the substance at higher and higher velocities until eventually some of them attain such a velocity that on rising to the surface of the water they have sufficient energy to break away from the main body, in which case they appear above the surface in the form of steam. A somewhat similar action takes place if an electrical conductor is heated in a vacuum by the flow of an electric current through it. As the temperature of the metal is raised the velocities of the molecules and electrons increase, and by the time that it has reached a red heat the electrons have attained a velocity of approximately 300 miles per second. When a velocity of about 600 miles per second is reached many of them will escape through the surface of the metal and be thrown off into the surrounding space. The electrons that escape from the metal are the same no matter from what metal they are evaporated.

Figure 351 shows such a conductor enclosed in a glass envelope from which most of the air has been excluded, the conductor being

heated by the passage of an electrical current through it supplied by the battery marked A and called the "A" battery. Under this condition, after the electrons leave the conductor, which is called the filament, they appear in the space surrounding it and some of them will actually bombard the glass envelope and eventually return to the filament.

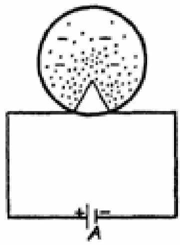


Figure 351 — When a metal is heated in a vacuum, electrons are evaporated from it and appear in the surrounding space as negatively charged particles.

From Chapter XXVI it will be clear that since the electrons are negatively charged particles, the moment that any of them leave the filament some of its negative charge has been removed and it therefore becomes positively charged. Since there is then a positively charged body surrounded by negatively charged particles (electrons) these latter will be attracted by the positive charge, thus all the electrons emitted will eventually be recaptured by the filament.

1. TWO-ELEMENT VACUUM TUBE

The first vacuum tube to be developed was of the two-element type consisting of a filament and a plate as shown in Figure 352.

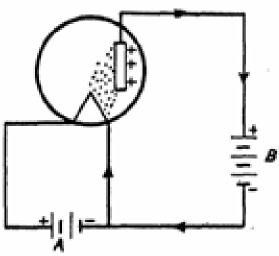


Figure 352 — Electrons emitted by the filament are attracted by the positive charge on the plate and continue through the connecting wires back to the filament.

In this type a plate is inserted in the glass envelope and is given a positive charge by the battery marked "B" in the figure, the positive pole of which is connected to the plate and the negative to the filament. When the current from the "A" battery starts to flow through the filament it raises its temperature and electrons are emitted in the manner described in the previous section. Since now the plate is given a positive charge by the "B" battery, many of the electrons emitted will be attracted by this positive charge and will consequently stream across the space between these two elements, continuing through the connecting wires and battery, eventually returning to the filament in the direction shown by the arrows. This constitutes a flow of direct current, not only in the wires comprising the circuit, but also in the space between the filament and plate. There is thus produced a direct-current flow in a vacuum which is just as truly an electrical current as when this takes place in a metal conductor.

Vacuum tubes of this type are commonly used as rectifiers.

2. AMOUNT OF CURRENT FLOW IN A VACUUM TUBE

From fundamental considerations discussed earlier it is apparent that we can ascertain the amount of current flowing in a vacuum tube under given conditions. In Chapter XXV the diameter, mass, and electric charge of an electron were given. For convenience these values are repeated:

Diameter - - - - - -	3.2×10^{-13} centimeters
Mass - - - - - - - -	8.9×10^{-28} grams
Electric charge - - - -	1.59×10^{-19} coulombs

One coulomb therefore consists of 6.29×10^{18} electrons, and since one ampere is equal to a flow of one coulomb per second, it is produced by a flow of 6.29×10^{18} electrons per second. If we therefore are able to determine the number of electrons being emitted by a filament under given temperature conditions, it is possible to calculate the maximum amount of current that can be produced in the plate circuit.

3. VELOCITY OF ELECTRONS IN A VACUUM TUBE

The velocity with which electrons move in a vacuum tube in the space between the filament and plate will depend partly on the potential difference between these two elements.

The unit of work in an electrical circuit is the joule and is the amount of work done by a current of one ampere flowing through a resistance of one ohm.

$$W = Q \times E$$

where $W =$ joules
$Q =$ coulombs
$E =$ volts

The work done on an electron equals QV where V is the potential difference of the points between which it is moved, and Q its electrical charge.

A mechanical body in motion possesses a certain amount of energy equal to $\frac{1}{2}\, MV^2$ where $M =$ Mass in grams
$V =$ Velocity in cms. per second.

The resultant unit of energy is the Erg.

Since one joule $= 10^7$ ergs

$$\tfrac{1}{2}\, MV^2 = QV \times 10^7$$

$$\text{from which} \quad V = \sqrt{\frac{2QV \times 10^7}{M}}$$

If, in this equation, we substitute for Q and M the corresponding values

for an electron and assume this electron to be affected by a difference of potential of 10 volts, the equation becomes:

$$V = \sqrt{\frac{2 \times 1.6 \times 10^{-19} \times 10 \times 10^{7}}{8.8 \times 10^{-28}}}$$

From which $V = 1900$ kilometers per second, or about 1180 miles per second. From this it will be seen that electrons under these conditions attain extremely high velocities. In the case of a water-cooled vacuum tube with a plate potential of 10,000 volts, electrons may attain a velocity as high as 37,000 miles per second, and it is because of this high velocity bombardment of the plate that under such operating conditions it frequently becomes red hot.

4. AMOUNT OF PLATE CURRENT

Referring again to Figure 352, there are two factors which determine the amount of plate current that will flow. These are (a) filament temperature, (b) plate potential.

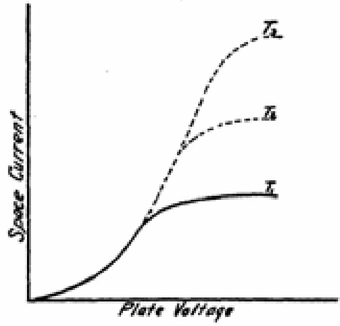


Figure 353 — The plate current may be increased by raising the filament temperature.

Figure 353 shows the way in which the plate current will vary when the plate potential (E_p) is changed and three different values of filament temperature are considered. It will be noticed that for the temperature marked T_1 as the plate potential is raised from a low value, the plate current (I_p) gradually increases up to a certain point beyond which the curve flattens off, indicating that all of the electrons that have been emitted by the filament at that temperature have reached the plate and consequently no matter how much the plate potential is raised beyond this point no further increase in plate current is possible. When the filament temperature is increased to the values marked T_2 and T_3 more electrons are emitted and a higher maximum plate current is consequently obtainable.

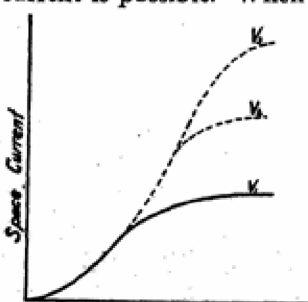


Figure 354—The plate current may be increased by maintaining a constant filament temperature and raising the plate potential.

Referring now to Figure 354, it will be seen that for a given value of plate potential (V_1) the plate current rises as the filament temperature is increased up to the point where the maximum number of electrons which that plate voltage can attract are reaching it. If now a higher plate voltage is applied, such as shown by V_2 or V_3, higher values of plate current

will result, but in each case a maximum amount will be reached as indicated by the flattening off of the curves.

5. SPACE CHARGE

When the filament is emitting electrons which are being attracted to the plate, due to the fact that at any point in the space between the filament and plate a certain number of electrons are always present, this point will have a negative charge, the amount of which depends upon the density of electron flow. This charge is called the space charge of the tube and acts as a limiting factor on the plate current. The denser the flow of electrons the higher will be the space charge. As a result, if we consider any point in the space between the filament and plate, its negative charge will tend to repel the electrons between it and the filament, thus cutting down the total number of electrons reaching the plate, and as a result limiting the plate current. The space charge is therefore an important factor in the operation of a vacuum tube.

6. THREE-ELEMENT VACUUM TUBE

If a third element is inserted in a vacuum tube between the filament and plate, and if a negative potential be applied to this element, it may be used as a control element in limiting the plate current to any desired value.

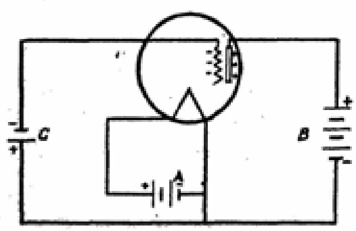


Figure 355 — Circuit of three-element vacuum tube.

Figure 355 shows such a tube in which the third element or grid has a battery, commonly called a "C" battery, connected to it which gives it the desired negative potential. If this potential is high very few of the electrons which leave the filament will reach the plate. If, however, it is low in value most of them will arrive at the plate and so constitute a relatively large plate current. It will thus be seen that the potential applied to the grid determines the amount of plate current. If this potential be sinusoidal, so also will be the plate current.

7. VOLTAGE AMPLIFICATION FACTOR

Referring to Figure 356 we have assumed a case in which the plate is held at a 100 volt positive potential with respect to the negative end of the filament.

If we assume a uniform drop in potential across the plate-filament space, at a point half way between them the potential would be 50 volts above the filament. If we pick the point in this space that is 10 volts above the filament potential and insert at this point a grid to which a negative potential of 10 volts is applied, it is obvious that these two potentials would exactly counteract one another and would therefore result in no attractive force being applied to the electrons leaving the filament.

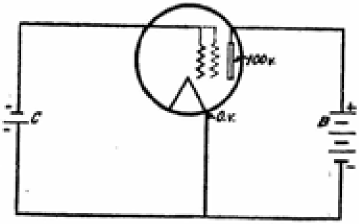


Figure 356 — The position of the grid affects the plate current.

If the grid now be moved closer to the plate its effect on the plate potential will become less and less so that its power to control the plate current will depend not only on the potential applied to it but also to its position with respect to the other elements of the tube. If we assume any given position for the grid, apply a negative potential to it, measure the plate current and then reduce the negative potential on the grid by one volt, we would obtain a larger plate current. If we now return the grid to its original potential and increase the plate voltage until the original value of plate current is obtained, we might find that it was necessary to increase this voltage by 10 volts.

Consequently, a change of one volt on the grid is as effective in this case as a change of 10 volts on the plate. In other words, we have obtained a voltage amplification factor of 10 in the particular tube under consideration. This voltage amplification factor is another of the important characteristics of a vacuum tube and enters into practically all calculations on amplifier characteristics. Amplification factors, varying from 3 up to 30 or more, are obtained from three-element vacuum tubes. The Greek letter μ is the symbol used to represent this constant and may be shown to be equal to the rate of change of plate voltage with respect to the grid voltage, or equals $\frac{dE_p}{dE_g}$ where d represents a small change in the quantity with which it is associated.

8. PLATE-FILAMENT RESISTANCE

In common with other electrical circuits the plate-filament resistance of a vacuum tube may be determined by dividing the differences of potential between these two elements by the resultant current flow; i.e.,

$$R_p = \frac{E_p}{I_p}$$

where R_p = plate to filament resistance

E_p = plate voltage

I_p = plate current

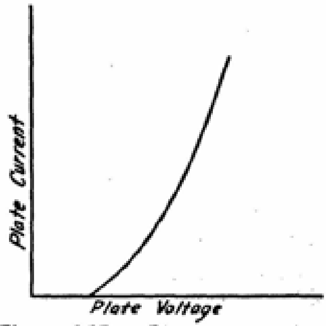


Figure 357 — Plate current-plate voltage characteristic of three-element vacuum tube.

Figure 357 shows a typical plate voltage-plate current curve for a triode and it will be noticed that the value obtained by dividing the plate voltage by the plate current will vary, depending upon what point on the curve is considered. Since, under operating conditions, we utilize a considerable portion of the curve, it is obvious that the direct-current plate to filament resistance would not necessarily give a true indication of the situation under operating conditions. We, therefore, make use of what is known as the alternating-current plate to filament resistance.

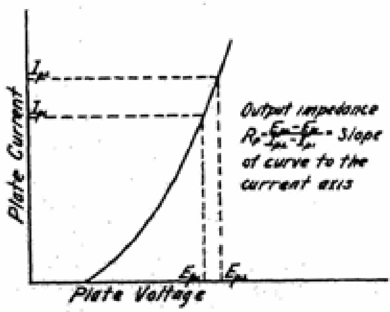


Figure 358 — The plate-filament resistance may be determined from the plate voltage-plate current curve.

By reference to Figure 358 it may be seen that it is determined as follows:

$$R_p = \frac{E_{p2} - E_{p1}}{I_{p2} - I_{p1}}$$

This is known as the alternating-current plate-filament resistance and is also sometimes called the plate-filament impedance.

9. MUTUAL CONDUCTANCE

The conductance of an ordinary direct-current circuit is considered as the reciprocal of the resistance. If we designate the conductance by the letter G, then

$$G = \frac{I}{E}$$

In the three-element vacuum tube, however, we are principally interested in the variation in plate current produced by a change of grid voltage rather than by a change of plate voltage. The mutual conductance therefore must equal the plate current divided by the grid voltage rather than the plate current divided by the plate voltage, and if we designate the mutual conductance by the symbol G_m, then

$$G_m = \frac{I_p}{E_g}$$

From Section 7 it was shown that

$$\mu = \frac{E_p}{E_g}$$

and in Section 8 it was shown that

$$R_p = \frac{E_p}{I_p}$$

Consequently $\quad G_m = \frac{\mu}{R_p}$

$$= \frac{E_p}{E_g} \times \frac{I_p}{E_p}$$

$$= \frac{I_p}{E_g}$$

Thus, the mutual conductance of a three-element tube equals the amplification constant divided by the plate to filament resistance.

Chapter XXXIX

TRIODE AMPLIFIERS

By A. P. HILL

1. THREE-ELEMENT TUBE AS AN AMPLIFIER

The amplifying property of a vacuum tube results from the fact that when a potential is applied to the grid no grid current results as long as the grid is maintained at a negative potential. There is, however, an increased voltage released in the plate circuit which results in a change in plate current. In this way it is possible for a voltage representing the expenditure of practically no power to be applied to the grid of a vacuum tube and to release an appreciable amount of power in the plate circuit.

Amplifiers may be classified in several ways. They may be designed to increase either the voltage or power in a circuit and are consequently called voltage and power amplifiers, respectively.

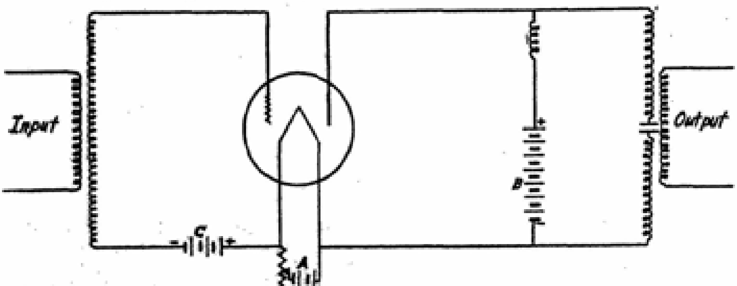


Figure 359 — Simple circuit of triode amplifier.

Figure 359 shows the fundamental circuit of a triode amplifier in which "A" is the battery that supplies power to the filament in order to raise the temperature to the point where satisfactory electron emission takes place. "B" is the battery that applies a positive potential to the plate in order to attract to it electrons that are emitted by the filament. "C" is the battery applied to the grid which maintains it at some desired negative potential, thus preventing the flow of any grid current. If now an alternating potential is applied to the grid its voltage will vary about the mean value determined by the "C" battery and as a result the plate current will decrease and increase as the grid becomes alternately more or less negative with respect to the filament. Such a condition is shown graphically in Figure 360.

In order to obtain amplification with a minimum amount of distortion, one of the essentials is that the grid never be permitted to become

positive, for such a positive grid would capture some of the electrons which are being emitted by the filament and these would pass around the grid circuit back to the filament, thus constituting a grid current. Neither should the grid potential be allowed to swing so far negative that the lower bend of the grid potential-plate current curve is reached. Consequently the "C" battery value or grid bias is selected so as to fall about the mid-point between the zero grid value and the point at which the plate-current curve reaches its lower bend, which, in the case of Figure 360, is approximately minus ten volts.

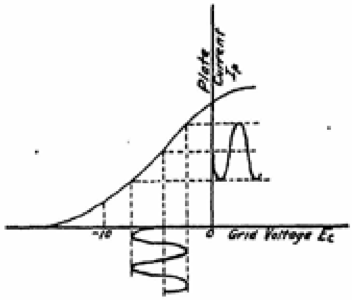


Figure 360 — A change in grid potential results in a corresponding change in plate current.

If consequently, we pick a value of minus five volts as the point at which to apply the alternating-current potential, we may apply an alternating voltage to the grid which may have a peak value of five volts, in which case, the grid will never be driven positive nor will it pass beyond the minus ten volt value previously mentioned. Figure 360 shows the effect produced on the plate current by applying an alternating-current potential to the grid.

2. MAXIMUM VOLTAGE AMPLIFICATION OBTAINABLE

The maximum voltage amplification obtainable by the use of a single triode is limited by its amplification constant, since the amplification constant is the total amount of voltage released in the plate circuit divided by that applied to the grid. If it were possible to use all the voltage released in the plate circuit we would only obtain a voltage amplification equal to this amplification constant. The voltage developed across the tube itself is, however, wasted since the proportion of the total voltage that is applied to the grid of the second tube is that developed across the coupling device.

If we now consider a case where a microphone is used to convert acoustical into electrical energy, as for example in a loud-speaker system, it will be necessary, in order to satisfactorily operate the loud-speaker, to increase the amount of power existing in the microphone circuit possibly several thousand times. This means that the vacuum tube immediately preceding the loud-speaker must be capable of delivering in its plate circuit an amount of power equal to that which the loud-speaker requires. In order to release this amount of power from a vacuum tube it means that a considerable voltage charge must be applied

to its grid and since the voltage output of a microphone is quite small, there must be one or more stages of amplification preceding the final power stage which consecutively build up the voltage to the value required for application to the grid of this tube. Such an amplifying system will normally consist of several stages of voltage amplification and one final stage of power amplification, and as previously stated, the design of voltage and power amplifiers is quite different. These will now be discussed in some detail.

3. VOLTAGE AMPLIFIERS

In order to connect two or more tubes together so that the voltage released in the plate circuit of one may be applied to the grid circuit of the next, they may be coupled together in three different ways. These are known as: (a), resistance coupling; (b), impedance coupling; and (c), transformer coupling.

(a) Resistance Coupled Amplifiers

In a voltage amplifier the object is to obtain from the plate circuit of one tube as high a potential as possible for application to the grid of the next tube. In this way, by utilizing several stages of amplification, the voltage may be built up gradually until at the grid of the final power stage there is delivered a potential large enough to release the required amount of power in its plate circuit. A resistance coupled amplifier is one in which a resistor is inserted in the plate circuit of one tube and across this resistor the grid and filament of the next tube are connected. Thus, all of the voltage developed across this resistor is applied to the next stage. Any voltages developed in other parts of the plate circuit will, however, be lost. Figure 361 shows a simplified circuit of an amplifier of this type in which R_C is the coupling resistor and C_G and R_{GL} are the grid condenser and grid leak respectively, the purpose of the former being to prevent the plate potential of the first tube being applied to the grid

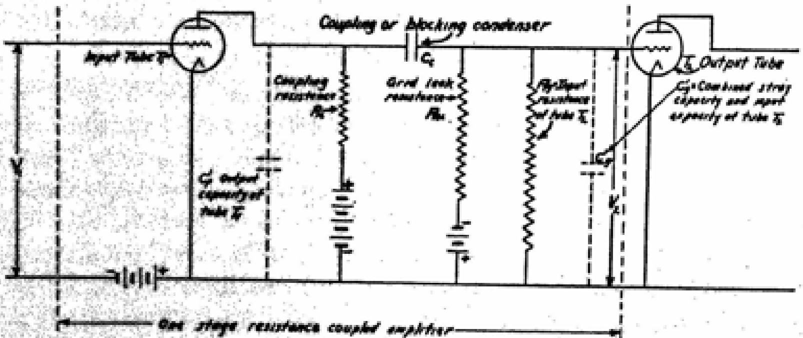


Figure 361 — Simplified circuit of resistance coupled amplifier.

of the second tube, and of the latter to prevent the "C" battery from short circuiting the grid and filament of the second tube.

If we now consider the plate circuit of the first tube we see that it consists of two resistances in series, namely, the resistance of the plate to filament (R_{p1}) of the first tube and the coupling resistor R_G. Any voltage developed in this circuit would, therefore, be divided between these two elements, the larger voltage being, of course, developed across the larger resistance. As stated above, the voltage developed across R_G is the only portion of the total released voltage that is applied to the grid and filament of the second tube. It is, therefore, advisable to make this resistance as large as practicable in relation to the plate-filament resistance of the tube in order that the voltage developed across it may be as large a proportion of the total voltage possible.

We should, of course, be able to calculate the amount of voltage amplification that can be obtained for a single stage where the constants of the circuit are known. Referring again to Figure 361, the voltage applied across the grid and filament of the first tube is designated by V_1 and that across the grid and filament of the second tube by V_2. The voltage amplification obtained is, therefore, the ratio of these two voltages or $A = \frac{V_2}{V_1}$, where A represents the amplification per stage. The signal current in the plate circuit of the first tube equals the total signal voltage released in this circuit divided by the total resistance of the circuit. This signal voltage will be the voltage applied in the grid circuit multiplied by the amplification constant of the tube, or $\mu \times V_1$,

$$\therefore \quad I_p = \frac{\mu \times V_1}{R_p + R_G}$$

If we designate the portion of the signal voltage that is developed across R_G by E_{rG}, this must equal the value of the resistance multiplied by the current flowing through it, or

$$E_{rG} = I_p \times R_G$$

Substituting in this equation the value given for I_p above:

$$E_{rG} = \frac{\mu \times V_1 \times R_G}{R_p + R_G}$$

But, as stated above:

$$A = \frac{V_2}{V_1}$$

So, if we now neglect the small voltage drop that occurs across R_{GL} and C_G, V_2 will be seen to be equal to E_r,

$$\therefore \quad A = \frac{V_2}{V_1} = \frac{\mu R_G}{R_p + R_G}$$

As stated above, Figure 361 is a simplified circuit of such an amplifier. Figure 362-A shows its exact equivalent circuit.

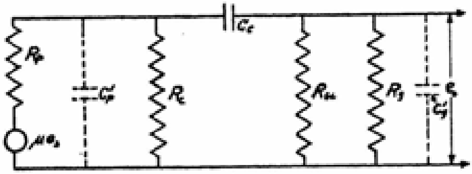


Figure 362-A — Exact equivalent circuit of amplifier.

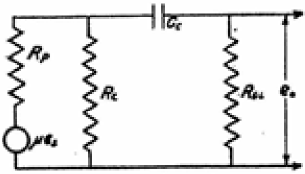


Figure 362-C — Simplified circuit accurate for low frequencies.

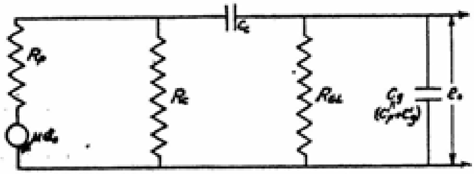


Figure 362-B — Practical equivalent circuit of amplifier.

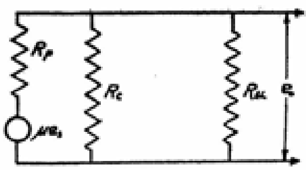


Figure 362-D—Simplified circuit accurate for intermediate frequencies.

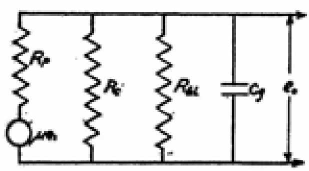


Figure 362-E — Simplified circuit accurate for high frequencies.

The capacity C'_g and the resistance R_g represent the input capacity plus any stray wiring capacities, and the input resistance respectively of the second tube. C'_p represents the plate-filament capacity of the first tube together with any stray wiring capacities that are in parallel with its plate circuit. For ordinary analyses the capacities C'_p and C'_g may be combined into a single capacity as shown in Figure 362-B, which represents the practical equivalent circuit of the amplifier. Figure 362-C represents a simplified circuit that is accurate for low frequencies. In this drawing the capacity C_g has been omitted since its shunting capacity at low frequencies is negligible.

Figure 362-D shows a simplified circuit that is accurate for intermediate frequencies in which it will be noticed that the capacity C_c shown in the previous figure has been omitted since the series reactance which it introduces is negligible at the frequencies being considered. Figure 362-E shows a simplified circuit that is accurate for high frequencies. In this figure the capacity C_g must be included since its shunting effect may become quite appreciable at these frequencies.

Figure 365, shows the voltage amplification per stage obtainable under the conditions indicated. Here amplification values were calculated for various values of coupling resistor and a curve drawn through the points so obtained. It will be seen that the amplification increases rapidly from low values of coupling resistor up to a value of approximately 60,000 ohms, and beyond this point the curve slopes off and the addition of a large amount of resistance in the coupling resistor results in but little increase in amplification. It will also be noted that an amplification value equal to 90 per cent of the μ of the tube is obtainable when a coupling resistor is used that is equal to ten times its plate-filament resistance. Reference to this fact will again be made in the following section.

(b) Impedance Coupled Amplifiers

A simplified circuit of an impedance coupled amplifier is shown in Figure 363.

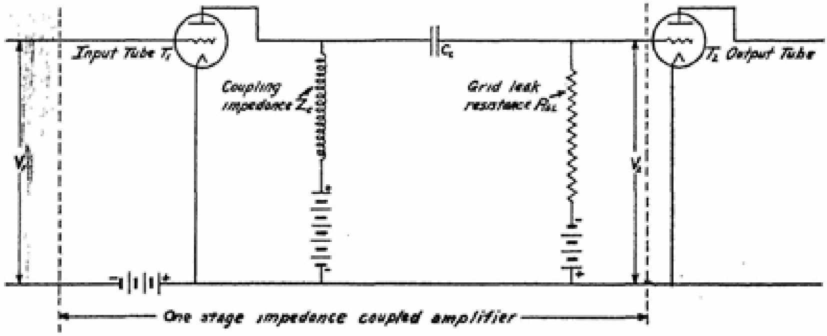


Figure 363 — Simplified circuit of impedance coupled amplifier.

Here the coupling resistor R_C used in the resistance coupled amplifier has been replaced by an inductance coil marked Z_C. Apart from this change the circuit is similar to the resistance coupled type. We may calculate the voltage amplification per stage obtained with this type of amplifier in a manner similar to that previously discussed, substituting for the value of the coupling resistance an impedance factor; thus, referring to Figure 363, the amplification per stage $A = \frac{V_2}{V_1}$

The voltage across the coupling impedance $E_Z = I_p \times Z$.

In calculating the value of I_p care should be taken to maintain the correct phase relationships between the resistance and reactance components of the circuit. Keeping these in mind it will be seen that

$$I_p = \frac{\mu \times V_1}{\sqrt{(R_p + R)^2 + X^2}}$$

$$\therefore \quad E_z = \frac{\mu V_1 Z}{\sqrt{(R_p + R)^2 + X^2}}$$

But since $A = \dfrac{V_2}{V_1}$

$$A = \frac{\mu \times Z}{\sqrt{(R_p + R)^2 + X^2}}$$

Figure 366, shows the voltage amplification obtainable per stage where two 262-A tubes are coupled by means of an impedance, the constants of the circuit being as shown.

By comparing this curve to that of the resistance coupled amplifier using the same tubes, it will be noted that the curve rises more rapidly for lower values of impedance and that an amplification per stage equal to 90 per cent of the amplification constant of the tube is obtained when coupling impedance is equal to twice the plate to filament resistance.

Since the reactance value of any inductance coil will fall off at the lower frequencies and since the amplification per stage obtainable for an amplifier of this type depends upon the impedance in the plate circuit, it will be seen that unless care is taken in the design of such an amplifier the amplification per stage will be less at low frequencies than at high. Figure 367 shows some frequency amplification curves calculated for various values of inductance using 262-A vacuum tubes under the conditions indicated. Here it will be seen that in order to obtain a flat frequency response down to 50 cycles it is necessary in this instance to use an inductance as high as 400 henries. At lower values of inductance the amplification falls off until when using only 10 henries inductance it starts to drop at a frequency of 2000 cycles.

4. COMPARISON OF RESISTANCE AND IMPEDANCE COUPLED AMPLIFIERS

In the case of a resistance coupled amplifier, since the value of the coupling resistor must be several times as great as the plate filament resistance, there will be a considerable drop in direct-current potential across it. Consequently, in order to maintain any specified voltage across the plate and filament of the tube it will be necessary to maintain the battery voltage several times higher.

In the case of the impedance coupled amplifier it is possible to obtain a high value of impedance with a relatively small amount of direct-current resistance in the coil, thus avoiding the necessity for maintaining high "B" battery voltages.

In either of these two types of amplifiers the maximum value of amplification obtainable will be somewhat less than the amplification constant of the tube that is used. Usually a value of 70 per cent or 80 per cent of this amount can be conveniently obtained. The principal advantage of the impedance coupled amplifier over the resistance coupled is that lower "B" battery voltages may be used.

By careful design a satisfactory frequency characteristic may be obtained with either type.

5. TRANSFORMER COUPLED AMPLIFIERS

A transformer coupled amplifier, as the name implies, makes use of a transformer to couple together the plate and grid circuits of two successive tubes in place of the resistance or inductance coils previously described. Such transformers may have varying turn ratios depending upon the types of tubes involved, their function, however, being to raise the voltage applied to the grid of a tube to as high a value as possible in order to obtain as great a voltage amplification per stage as possible.

In the design and measurement of amplifiers certain fundamental considerations must be remembered, otherwise entirely erroneous results may be obtained.

The gain of an amplifier, frequently referred to as insertion gain, represents the amount of amplification obtained with the amplifier inserted in the circuit compared to a condition where the amplifier is removed from the circuit and the equipment that is connected to the input and output of the amplifier is tied directly together. Assuming a condition such as is shown in Figure 364 where we have two resistances, R_t and R_r, which are equal in value and between which the amplifier is inserted, the gain of the amplifier would be the amplification obtained with the circuit as shown in the figure as compared to the condition where R_t and R_r are connected directly together.

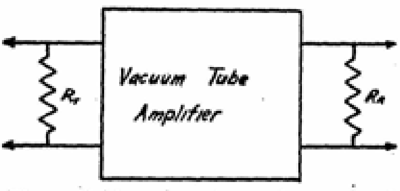


Figure 364 — Insertion gain produced by an amplifier.

In order to ascertain the amount of voltage amplification obtained it is only necessary to determine the ratio of the voltages across R_t and R_r with the amplifier inserted as compared to that where the amplifier is not inserted. If then R_t and R_r are equal in value we have by this measurement the voltage amplification of the amplifier. If, however, they are unequal, it will be necessary to apply the correction factor as described in Chapter XXXII on the Decibel.

Figure 369, Page 523, shows a simplified diagram of a single stage transformer coupled amplifier. In this diagram the resistance R_t represents the resistance of the circuit to which the amplifier is connected and E_t represents the voltage of the source. The effect produced would be the same as if a generator of internal resistance R_t and voltage E_t were supplying voltage to the amplifier. R_r represents the resistance of the circuit to which the output of the amplifier is connected. R_1 is used where it is desired to match the impedance of the input of the amplifier to the circuit to which it is connected. R_2 might be a potentiometer by means of which the gain of the amplifier could be controlled. If the amplifier were not connected in the circuit the resistance R_r would be in the position shown for R_1 and the gain of the amplifier is, therefore, the voltage obtained across R_r with the amplifier inserted, compared to the voltage across it when connected in the position just described.

In order to calculate the gain of the amplifier it is convenient to assume that the voltage generated by E_t is two volts, in which case, if R_t and R_r were directly connected together and were equal to one another, one volt would be generated across each. If now the amplifier were inserted and with two volts generated at E_t, there were ten volts produced across R_r the voltage amplification would be ten. Since instead of obtaining one volt across R_r as was the case when there was no amplifier inserted, we now obtain ten volts. It is consequently convenient in calculating gain to always assume two volts as the value of the source of supply, in which case the voltage actually developed across R_r with the amplifier inserted is the actual amount of voltage amplification obtained.

If we now consider Case A of Figure 369, Page 523, where R_1 and R_2 are equal to infinity and are, therefore, non-existent, and where R_t equals R_r we may calculate the gain of the amplifier as follows: two volts are generated at E_t, and since R_1 is infinite, the whole of this voltage will appear across the primary of transformer T_1. Since T_1 represents the turn ratio of this transformer, the voltage across the secondary will be T_1 times two and the whole of this voltage will appear across the grid and filament of the tube. The total voltage released in the plate circuit will be this value multiplied by the μ of the tube, or $\mu \times T_1 \times 2$.

We have, however, assumed that $R_p = \dfrac{R_r}{T_2^2}$; this means that the impedance looking into the primary of transformer T_2 is equal to the plate to filament resistance of the tube. Consequently one-half of the total voltage released in the plate circuit will be developed across the

primary of the transformer. The voltage across the secondary of this transformer will consequently be T_2 times as great; i.e., the total amplification obtained equals $T_2 \times \frac{1}{2} \times \mu \times T_1 \times 2 = \mu \times T_1 \times T_2$, as shown in Figure 369.

The other cases shown in this figure are simple extensions of Case A and no difficulty should be experienced in following the reasoning involved.

Figure 370, Page 524, and Figure 371, Page 525, give similar data for two and three stage amplifiers of this type and by applying a similar method of solution to that used for the single stage amplifier it should be possible to check the results shown for each of the cases considered.

6. POWER OUTPUT OF AMPLIFIERS

The stage of amplification that delivers the power output of an amplifier is called the power stage and the tube employed is called the power tube. There are two different types of requirements that this stage may be called upon to meet. The first of these has as its object the production of the maximum possible amount of undistorted power that can be obtained where a given tube is used and when a given signal voltage acts upon its grid. The second is the production of the maximum undistorted power output that can be obtained irrespective of the amount of signal voltage applied to that grid. The first of these two requirements is met by designing the circuit so that the tube works into an impedance equal to its own plate-filament resistance. The second is met by designing the circuit so that the tube works into an impedance that is equal to twice its own plate-filament resistance. The amount of power delivered to the output circuit under these two conditions may be calculated as follows:

CASE I — Where $R_p = R_r$:

The effective grid voltage $= E_G \times 0.707$.

The voltage acting in the output circuit $= E_G \times 0.707 \times \mu$.

The voltage available at the output transformer

$$= \frac{E_G \times 0.707 \times \mu}{2}$$

Since power in watts $= \frac{E^2}{R}$,

the power output to the transformer

$$= \frac{\left(\frac{E_G \times 0.707 \times \mu}{2}\right)^2}{R_r}$$

$$= \frac{(E_G\,\mu)^2}{8\,R_r} \text{ watts}$$

CASE II — Where $R_p \gtrless R_r$:

The effective grid voltage $= E_G \times 0.707$.

The voltage acting in the output circuit

$$= E_G \times 0.707 \times \mu \frac{R_r}{R_p + R_r}.$$

Power output to the transformer

$$= \frac{\left(E_G \times 0.707 \times \mu \frac{R_r}{R_p + R_r}\right)^2}{R_r}$$

$$= \frac{E_G^2 \, \mu^2 \, R_r}{2 \, (R_p + R_r)^2} \text{ watts}$$

7. REFERENCE POWER LEVEL

In communication work it is convenient to be able to express the power existing at any point in a circuit as so many db above or below a standard amount of power. The standard chosen is called the reference power level and is arbitrarily fixed as 0.006 watts. Since the number of db equals $10 \log \frac{P_2}{P_1}$

$$\therefore \quad \text{Power level} = 10 \log \frac{\text{power (watts)}}{0.006} \text{ db.}$$

If we take as an example the case where an amplifier delivers 0.06 watts to its output circuit, its power output level

$$= 10 \log \frac{0.06}{0.006} \text{ db}$$

$$= + 10 \text{ db}$$

This is normally written:

$$\text{Power level} = 10 \text{ db} \, / \, 0.006 \text{ watts}$$

This means that the amplifier under consideration delivers a power level which is 10 db above the arbitrary reference power level of 0.006 watts.

Figure 368, Page 522, gives in graphical form a means of deriving the number of db above or below this reference power level for any value of power between 0.00006 and 60 watts.

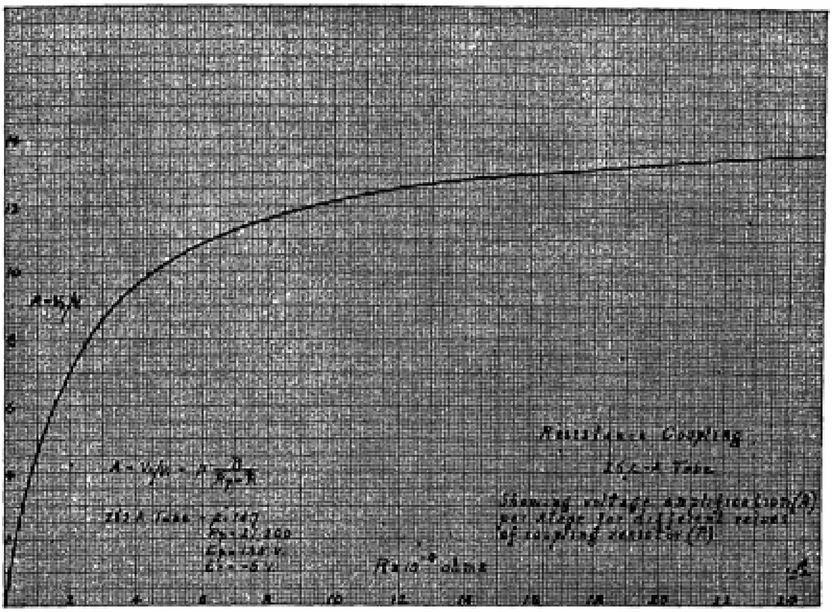


Figure 365.

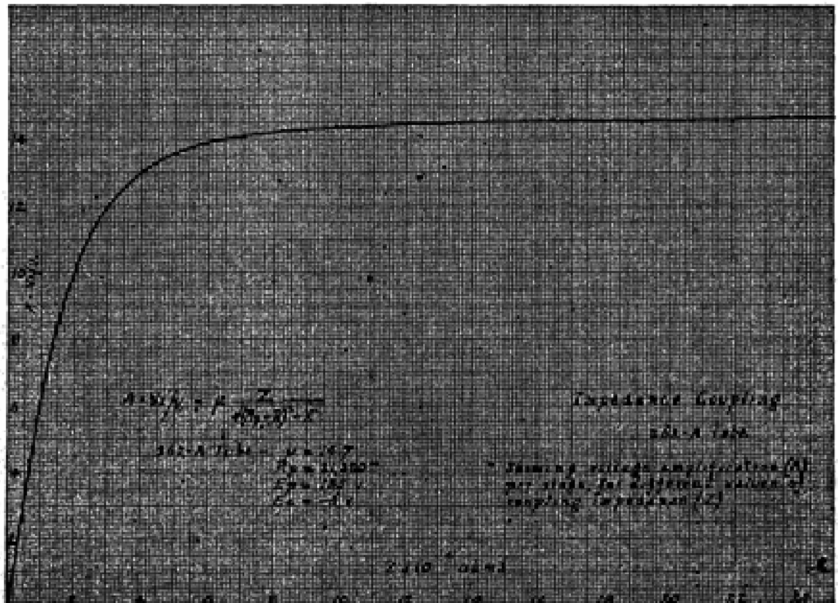


Figure 366.

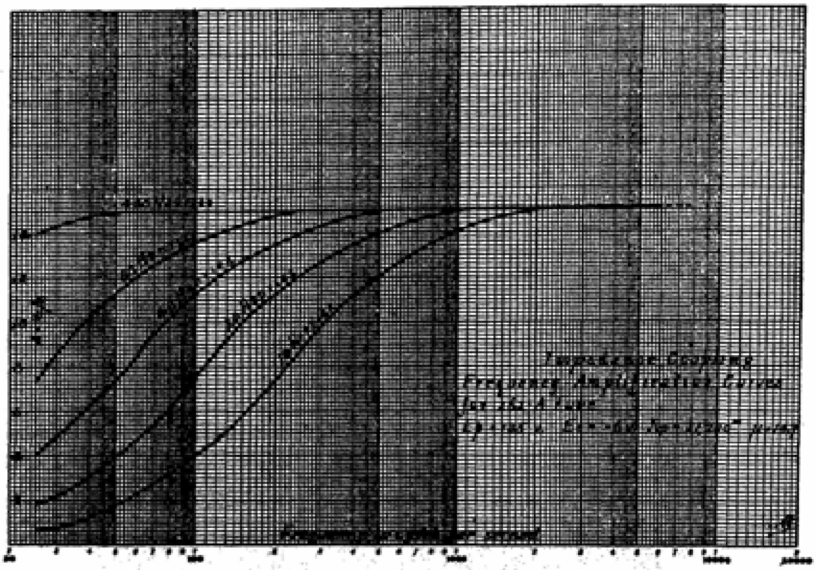


Figure 367

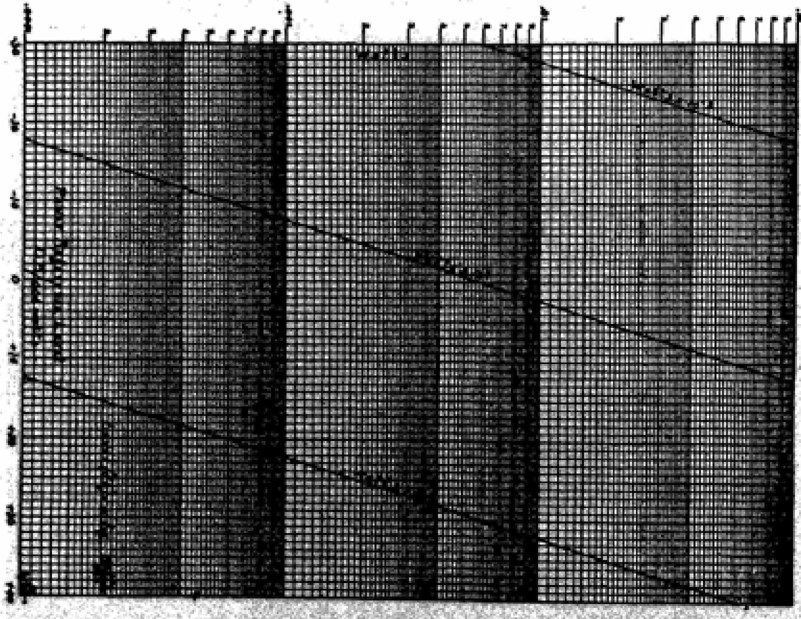
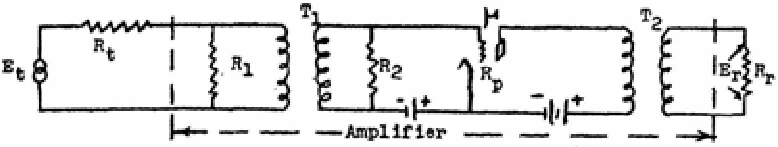


Figure 368.

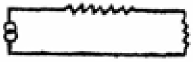
Transformer Coupled Amplifiers — Single Stage



R_t = Impedance of source. R_r = Receiving Impedance.

T_1 & T_2 = Turn ratios (S/P) of transformers.

With no amplifier inserted the following circuit condition would exist: -



and maximum power would be obtained in R_r when $R_t = R_r$ i.e. $E_r = 1/2\ E_t$.

This is equivalent to assuming $2E_t$ as the initial voltage, in which case E_r = the voltage obtained by the use of an amplifier, over that existing without the use of an amplifier.

In the following cases A = Voltage amplification.

(A) $R_p = R_r/T_2^2$ $\quad R_1 = \infty$ $\quad R_2 = \infty$ $\quad R_t = R_r$

$$A = T_1 T_2 \mu$$

(B) $R_p = R_r/T_2^2$ $\quad R_1 = \infty$ $\quad R_2 = \infty$

$$A = \mu T_1 T_2 \sqrt{R_t/R_r}$$

(C) $R_p = R_r/T_2^2$ $\quad R_2 = \infty$ $\quad R_1 = R_r$

$$A = 1/2 \mu T_1 T_2 \sqrt{R_t/R_r}$$

(D) $R_p = R_r/T_2^2$ $\quad R_2 = \infty$

$$A = \mu T_1 T_2 \sqrt{R_t/R_r} \left(\frac{R_1}{R_1 + R_t} \right)$$

(E) $R_p = R_r/T_2^2$ $\quad R_1 = \infty$

$$A = \mu T_1 T_2 \sqrt{R_t/R_r} \left(\frac{R_2}{R_2 + R_t T_1^2} \right)$$

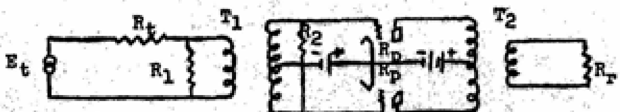
(F) $R_p = R_r/T_2^2$

$$A = \mu T_1 T_2 \sqrt{R_t/R_r} \left(\frac{R_1 R_2}{R_1 R_2 + R_t R_2 + R_1 R_t T_1^2} \right)$$

(G)

$$A = 2\mu T_1 T_2 \sqrt{R_t/R_r} \left(\frac{R_1 R_2}{R_1 R_2 + R_t R_2 + R_1 R_t T_1^2} \right) \left(\frac{R_r}{R_p T_2^2 + R_r} \right)$$

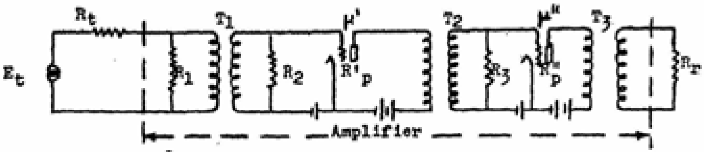
(H)



In push-pull circuits apply same formulae as above except replace R_p by $2\ R_p$

Figure 369.

Transformer Coupled Amplifiers — Two Stage



(A) $R''_p = R_r/T_3^2$. $R_1 = \infty$ $R_2 = \infty$ $R_3 = \infty$ $R_t = R_r$

$$\underline{A = \mu'\mu'' T_1 T_2 T_3}$$

(B) $R''_p = R_r/T_3^2$. $R_1 = R_2 = R_3 = \infty$

$$\underline{A = \mu'\mu'' T_1 T_2 T_3 \sqrt{R_t/R_r}}$$

(C) $R''_p = R_r/T_3^2$. $R_1 = R_t$ $R_2 = R_3 = \infty$

$$\underline{A = 1/2\,\mu'\mu'' T_1 T_2 T_3 \sqrt{R_t/R_r}}$$

(D) $R''_p = R_r/T_3^2$. $R_3 = \infty$

$$\underline{A = \mu'\mu'' T_1 T_2 T_3 \sqrt{R_t/R_r}\left(\frac{R_1R_2}{R_1R_2 + R_tR_2 + R_1R_tT_1^2}\right)}$$

(E) $R''_p = R_r/T_3^2$. $R_2 = \infty$

$$\underline{A = \mu'\mu'' T_1 T_2 T_3 \sqrt{R_t/R_r}\left(\frac{R_1}{R_1 + R_t}\right)\left(\frac{R_3}{R_3 + R'_pT_2^2}\right)}$$

(F) $R''_p = R_r/T_3^2$. $R_1 = \infty$

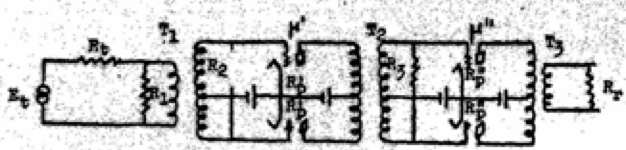
$$\underline{A = \mu'\mu'' T_1 T_2 T_3 \sqrt{R_t/R_r}\left(\frac{R_2}{R_2 + R_tT_1^2}\right)\left(\frac{R_3}{R_3 + R'_pT_2^2}\right)}$$

(G) $R''_p = R_r/T_3^2$.

$$\underline{A = \mu'\mu'' T_1 T_2 T_3 \sqrt{R_t/R_r}\left(\frac{R_1R_2}{R_1R_2 + R_tR_2 + R_1R_tT_1^2}\right)\left(\frac{R_3}{R'_pT_2^2 + R_3}\right)}$$

(H) $$A = 2\mu'\mu'' T_1 T_2 T_3 \sqrt{R_t/R_r}\left(\frac{R_1R_2}{R_1R_2 + R_tT_2 + R_1R_tT_1^2}\right)\left(\frac{R_3}{R_3 + R'_pT_2^2}\right)\left(\frac{R_r}{R_r + R''_pT_3^2}\right)$$

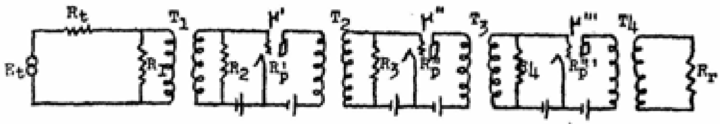
(I)



In push-pull circuits apply same formulae as above except replace R'_p by $2R'_p$ and R''_p by $2R''_p$.

Figure 370.

Transformer Coupled Amplifiers — Three Stage



(A) $R'''_p = R_r/T_4^2$. $R_1 = R_2 = R_3 = R_4 = \infty$ $R_t = R_r$

$$A = \mu'\mu''\mu'''T_1T_2T_3T_4$$

(B) $R'''_p = R_r/T_4^2$. $R_1 = R_2 = R_3 = R_4 = \infty$

$$A = \sqrt{R_t/R_r}\,\mu'\mu''\mu'''T_1T_2T_3T_4$$

(C) $R'''_p = R_r/T_4^2$. $R_1 = R_t$ $R_2 = R_3 = R_4 = \infty$

$$A = 1/2\sqrt{R_r/R_r}\,\mu'\mu''\mu'''T_1T_2T_3T_4$$

(D) $R'''_p = R_r/T_4^2$. $R_3 = R_4 = \infty$

$$A = \mu'\mu''\mu'''T_1T_2T_3T_4\left(\frac{R_1R_2}{R_1R_2 + R_tR_2 + R_1R_tT_1^2}\right)\sqrt{R_t/R_r}$$

(E) $R'''_p = R_r/T_4^2$. $R_2 = \infty$

$$A = \mu'\mu''\mu'''T_1T_2T_3T_4\left(\frac{R_1}{R_1 + R_t}\right)\left(\frac{R_3}{R_3 + R'_pT_2^2}\right)\left(\frac{R_4}{R_4 + R''_pT_3^2}\right)\sqrt{R_t/R_r}$$

(F) $R'''_p = R_r/T_4^2$. $R_1 = \infty$

$$A = \mu'\mu''\mu'''T_1T_2T_3T_4\left(\frac{R_2}{R_2+R_t+T_1^2}\right)\left(\frac{R_3}{R_3+R'_pT_2^2}\right)\left(\frac{R_4}{R_4+R''_pT_3^2}\right)\sqrt{R_tR_r}$$

(G) $R'''_p = R_r/T_4^2$.

$$A = \mu'\mu''\mu'''T_1T_2T_3T_4\left(\frac{R_1R_2}{R_1R_2 + R_tR_2 + R_1R_tT_1^2}\right)\left(\frac{R_3}{R_3+R_pT_2^2}\right)\left(\frac{R_4}{R_4+R''_pT_3^2}\right)\sqrt{R_tR_r}$$

(H) $$A = 2\mu'\mu''\mu'''T_1T_2T_3T_4\left(\frac{R_1R_2}{R_1R_2 + R_tR_2 + R_1R_tT_1^2}\right)\left(\frac{R_3}{R_3 + R'_pT_2^2}\right)\left(\frac{R_4}{R_4 + R''_pT_3^2}\right)\left(\frac{R_r}{R_4+R'''_pT_4^2}\right)\sqrt{R_t/R_r}$$

Figure 371.

Graphical Solution of Resistances in Parallel

(See Page 421)

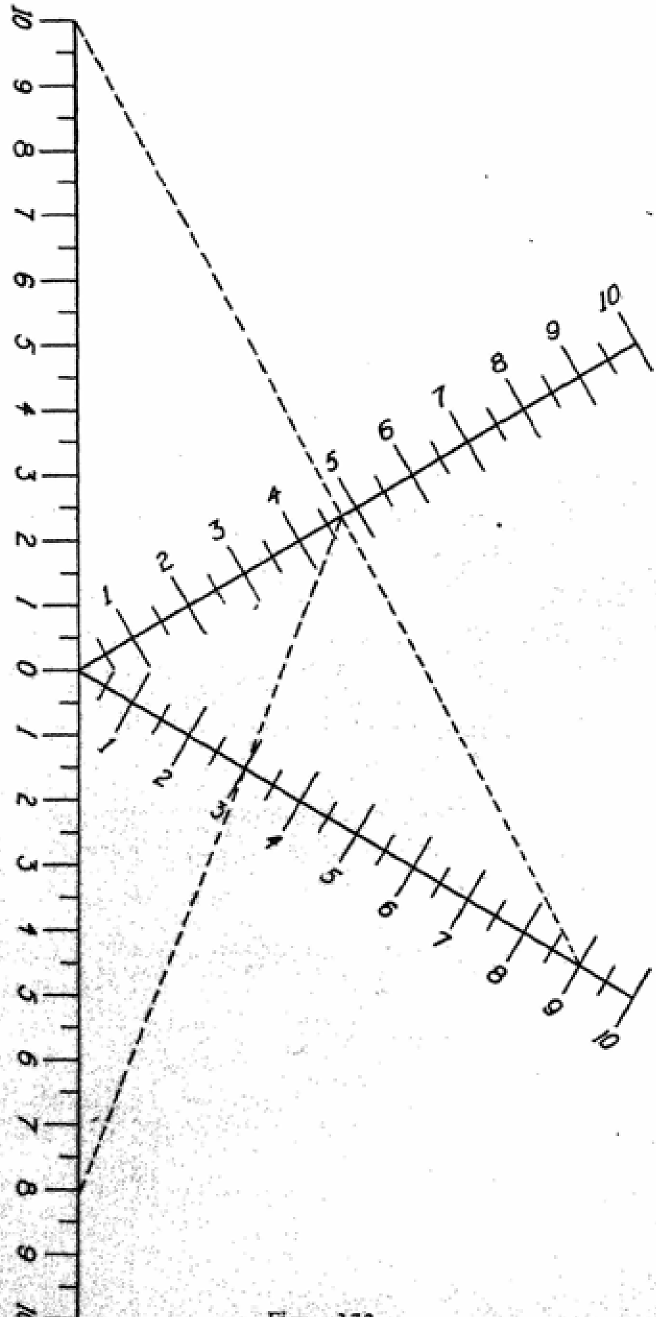
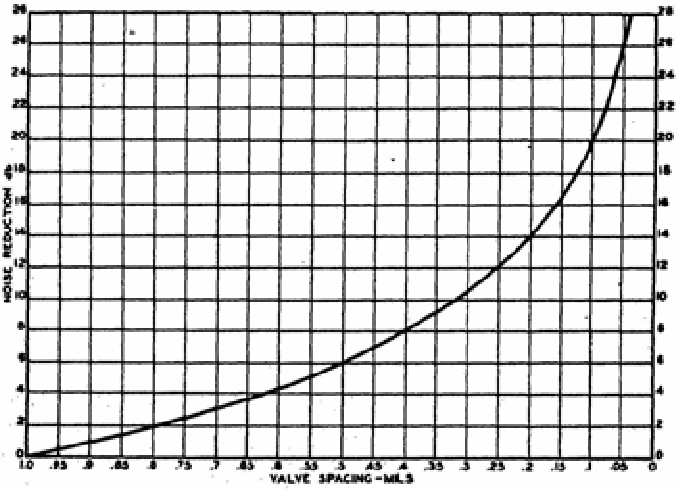


Figure 372.

APPENDIX

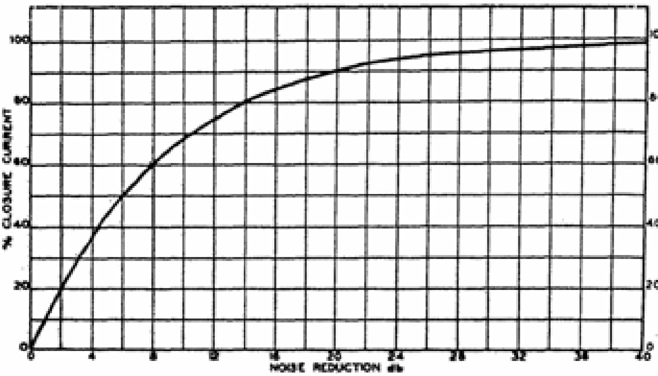
Charts XXX, XXXI, and XXXII are Noise Reduction Charts giving the valve spacing of a 1 mil light valve to obtain different amounts of noise reduction in terms of percentage closure current, and valve level in the terms of percentage modulation.

CHART XXX



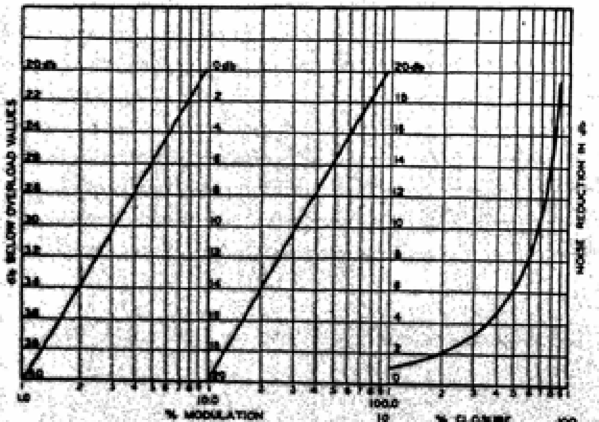
Valve Spacing vs. Noise Reduction for 1 mil Light Valve.

CHART XXXI



Noise Reduction vs. Valve Closure Current.

CHART XXXII

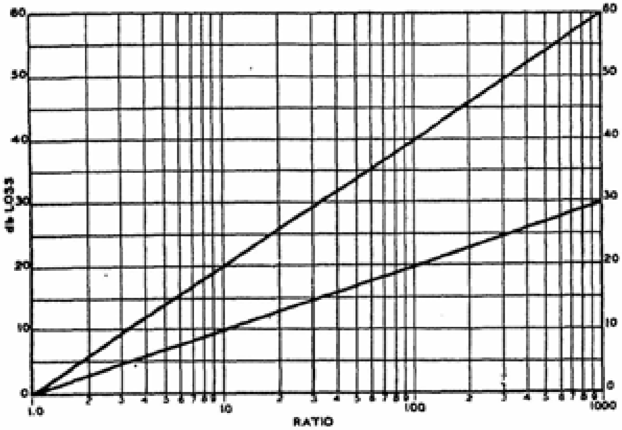


Valve Level vs. Percentage Modulation.

Noise Reduction vs. Light Valve Closure Current.

CHART XXXIII

Conversion Chart—Voltage, Current and Power Ratios to Transmission Units (db)



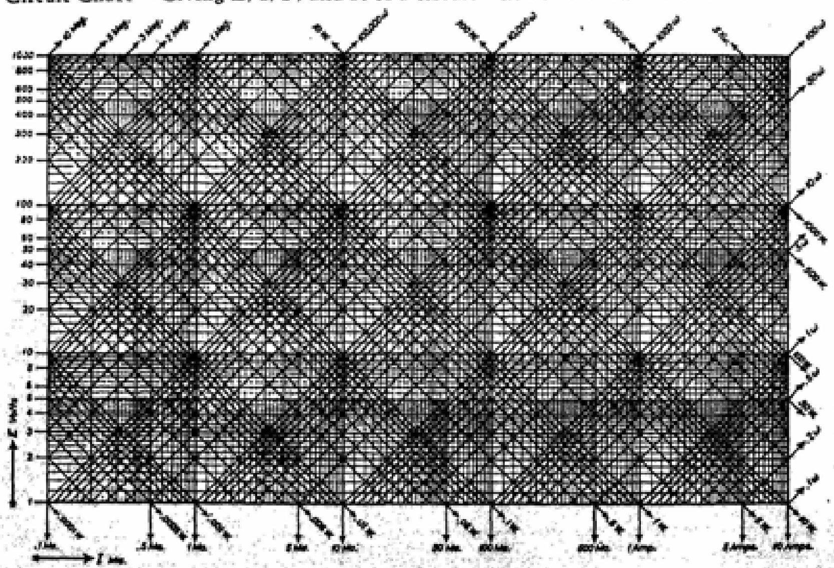
Top Curve—Voltage and Current ratios.

$$\text{Loss in db} = 20 \log \frac{E_1}{E_2} = 20 \log \frac{I_1}{I_2}$$

$$\text{Bottom Curve—Power ratios loss in db} = 10 \log \frac{P_1}{P_2}$$

CHART XXXIV

Circuit Chart—Giving E, I, P, and R of a circuit when two of these values are known.



USE OF THE CHART

Suppose the emf in a circuit is 100 volts and the current 100 ma; what is the resistance of the circuit, and what power is being dissipated? Starting with the 100-ma point on the abscissa, follow the vertical line to its intersection with the horizontal 100-volt line. This is also an intersection of two sloping lines. Following the line sloping to the left, the power, 10 watts, is read on the top of the chart. The line to the right gives the resistance, 1000 ohms, at the upper margin.

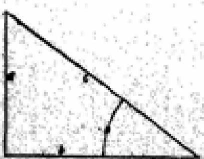
If a 5000-ohm resistance has a rating of 20 watts, what is the maximum permissible current, and at what voltage? Follow the 5000-ohm line to its intersection with the 20-watt line. Follow the vertical line from this point downward to the abscissa and obtain, by interpolation, 63 ma. Follow the horizontal line to the left of the intersection to the margin and read 316 volts.

Table VII
NATURAL SINES, COSINES, TANGENTS AND COTANGENTS

Degrees	Sine	Cosine	Tangent	Cotangent	Degrees
0	0	1	0	∞	90
0.5	0.0087	1	0.0087	114.6	89.5
1.	.0175	0.9998	.0175	57.29	89.
1.5	.0262	.9997	.0262	38.19	88.5
2.	.0349	.9994	.0349	28.64	88.
2.5	.0436	.9990	.0437	22.90	87.5
3.	0.0523	0.9986	0.0524	19.08	87.
3.5	.0610	.9981	.0612	16.35	86.5
4.	.0698	.9976	.0699	14.30	86.
4.5	.0785	.9969	.0787	12.71	85.5
5.	.0872	.9962	.0875	11.43	85.
5.5	0.0958	0.9954	0.0963	10.39	84.5
6.	.1045	.9945	.1051	9.5144	84.
6.5	.1132	.9936	.1139	8.7769	83.5
7.	.1219	.9925	.1228	8.1443	83.
7.5	.1305	.9914	.1317	7.5958	82.5
8.	0.1392	0.9903	0.1405	7.1154	82.
8.5	.1478	.9890	.1495	6.6912	81.5
9.	.1564	.9877	.1584	6.3138	81.
9.5	.1650	.9863	.1673	5.9758	80.5
10.	.1736	.9848	.1763	5.6713	80.
10.5	0.1822	0.9833	0.1853	5.3955	79.5
11.	.1908	.9816	.1944	5.1446	79.
11.5	.1994	.9799	.2035	4.9152	78.5
12.	.2079	.9781	.2126	4.7046	78.
12.5	.2164	.9763	.2217	4.5107	77.5
13.	0.2250	0.9744	0.2309	4.3315	77.
13.5	.2334	.9724	.2401	4.1653	76.5
14.	.2419	.9703	.2493	4.0108	76.
14.5	.2504	.9681	.2586	3.8667	75.5
15.	.2588	.9659	.2679	3.7321	75.
15.5	0.2672	0.9636	0.2773	3.6059	74.5
16.	.2756	.9613	.2867	3.4874	74.
16.5	.2840	.9588	.2962	3.3759	73.5
17.	.2924	.9563	.3057	3.2709	73.
17.5	.3007	.9537	.3153	3.1716	72.5
18.	0.3090	0.9511	0.3249	3.0777	72.
18.5	.3173	.9483	.3346	2.9887	71.5
19.	.3256	.9455	.3443	2.9042	71.
19.5	.3338	.9426	.3541	2.8239	70.5
20.	.3420	.9397	.3640	2.7475	70.
20.5	0.3502	0.9367	0.3739	2.6746	69.5
21.	.3584	.9336	.3839	2.6051	69.
21.5	.3665	.9304	.3939	2.5386	68.5
22.	.3746	.9272	.4040	2.4751	68.
22.5	.3827	.9239	.4142	2.4142	67.5
Degrees	Cosine	Sine	Cotangent	Tangent	Degrees

NOTE:—For angles from 0 to 45 degrees, use column headings at top of table. For angles from 45 to 90 degrees, use column headings at bottom of table. For angles greater than 90 degrees, subtract the given angle from 180, or an even multiple of 180 (360, 540, etc.), to obtain a value between 0 and 90 and use the function of that value; e. g., the sine of 94 degrees is equal to the sine of 180 — 94 = 86 degrees which, from the table, is found to be 0.9976.

Degrees	Sine	Cosine	Tangent	Cotangent	Degrees
23.	0.3907	0.9205	0.4245	2.3559	67.
23.5	.3987	.9171	.4348	2.2998	66.5
24.	.4067	.9135	.4452	2.2460	66.
24.5	.4147	.9100	.4557	2.1943	65.5
25.	.4226	.9063	.4663	2.1445	65.
25.5	0.4305	0.9026	0.4770	2.0965	64.5
26.	.4384	.8988	.4877	2.0503	64.
26.5	.4462	.8949	.4986	2.0057	63.5
27.	.4540	.8910	.5095	1.9626	63.
27.5	.4617	.8870	.5206	1.9210	62.5
28.	0.4695	0.8829	0.5317	1.8807	62.
28.5	.4772	.8788	.5430	1.8418	61.5
29.	.4848	.8746	.5543	1.8040	61.
29.5	.4924	.8704	.5658	1.7675	60.5
30.	.5000	.8660	.5774	1.7321	60.
30.5	0.5075	0.8616	0.5890	1.6977	59.5
31.	.5150	.8572	.6009	1.6643	59.
31.5	.5225	.8526	.6128	1.6319	58.5
32.	.5299	.8480	.6249	1.6003	58.
32.5	.5373	.8434	.6371	1.5697	57.5
33.	0.5446	0.8387	0.6494	1.5399	57.
33.5	.5519	.8339	.6619	1.5108	56.5
34.	.5592	.8290	.6745	1.4826	56.
34.5	.5664	.8241	.6873	1.4550	55.5
35.	.5736	.8192	.7002	1.4281	55.
35.5	0.5807	0.8141	0.7133	1.4019	54.5
36.	.5878	.8090	.7265	1.3764	54.
36.5	.5948	.8039	.7400	1.3514	53.5
37.	.6018	.7986	.7536	1.3270	53.
37.5	.6088	.7934	.7673	1.3032	52.5
38.	0.6157	0.7880	0.7813	1.2799	52.
38.5	.6225	.7826	.7954	1.2572	51.5
39.	.6293	.7771	.8098	1.2349	51.
39.5	.6361	.7716	.8243	1.2131	50.5
40.	.6428	.7660	.8391	1.1918	50.
40.5	0.6494	0.7604	0.8541	1.1708	49.5
41.	.6561	.7547	.8693	1.1504	49.
41.5	.6626	.7490	.8847	1.1303	48.5
42.	.6691	.7431	.9004	1.1106	48.
42.5	.6756	.7373	.9163	1.0913	47.5
43.	0.6820	0.7314	0.9325	1.0724	47.
43.5	.6884	.7254	.9490	1.0538	46.5
44.	.6947	.7193	.9657	1.0355	46.
44.5	.7009	.7133	.9827	1.1076	45.5
45.	.7071	.7071	1.0000	1.0000	45.
Degrees	Cosine	Sine	Cotangent	Tangent	Degrees



$\sin\theta = \frac{a}{c}$ $\tan\theta = \frac{a}{b}$ $\sec\theta = \frac{1}{\cos\theta}$

$\cos\theta = \frac{b}{c}$ $\cot\theta = \frac{b}{a}$ $\operatorname{cosec}\theta = \frac{1}{\sin\theta}$

Four-Place Table of Common Logarithms

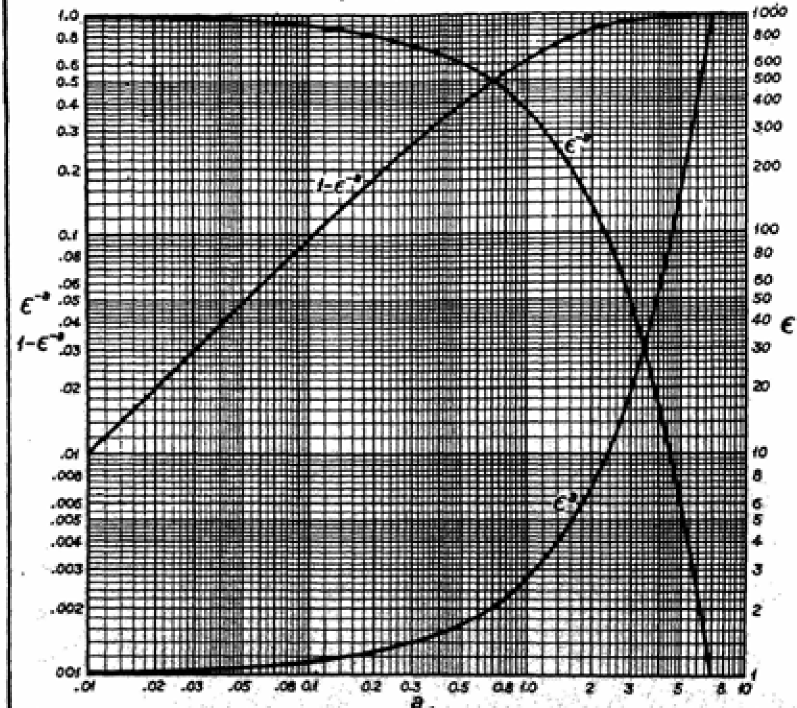
N	0	1	2	3	4	5	6	7	8	9
10	0000	0043	0086	0128	0170	0212	0253	0294	0334	0374
11	0414	0453	0492	0531	0569	0607	0645	0682	0719	0755
12	0792	0828	0864	0899	0934	0969	1004	1038	1072	1106
13	1139	1173	1206	1239	1271	1303	1335	1367	1399	1430
14	1461	1492	1523	1553	1584	1614	1644	1673	1703	1732
15	1761	1790	1818	1847	1875	1903	1931	1959	1987	2014
16	2041	2068	2095	2122	2148	2175	2201	2227	2253	2279
17	2304	2330	2355	2380	2405	2430	2455	2480	2504	2529
18	2553	2577	2601	2625	2648	2672	2695	2718	2742	2765
19	2788	2810	2833	2856	2878	2900	2923	2945	2967	2989
20	3010	3032	3054	3075	3096	3118	3139	3160	3181	3201
21	3222	3243	3263	3284	3304	3324	3345	3365	3385	3404
22	3424	3444	3464	3483	3502	3522	3541	3560	3579	3598
23	3617	3636	3655	3674	3692	3711	3729	3747	3766	3784
24	3802	3820	3838	3856	3874	3892	3909	3927	3945	3962
25	3979	3997	4014	4031	4048	4065	4082	4099	4116	4133
26	4150	4166	4183	4200	4216	4232	4249	4265	4281	4298
27	4314	4330	4346	4362	4378	4393	4409	4425	4440	4456
28	4472	4487	4502	4518	4533	4548	4564	4579	4594	4609
29	4624	4639	4654	4669	4683	4698	4713	4728	4742	4757
30	4771	4786	4800	4814	4829	4843	4857	4871	4886	4900
31	4914	4928	4942	4955	4969	4983	4997	5011	5024	5038
32	5051	5065	5079	5092	5105	5119	5132	5145	5159	5172
33	5185	5198	5211	5224	5237	5252	5263	5276	5289	5302
34	5315	5228	5340	5353	5366	5378	5391	5403	5416	5428
35	5441	5453	5465	5478	5490	5502	5514	5527	5539	5551
36	5563	5575	5587	5599	5611	5623	5635	5647	5658	5670
37	5682	5694	5705	5717	5729	5740	5752	5763	5775	5786
38	5798	5809	5821	5832	5843	5855	5866	5877	5888	5899
39	5911	5922	5933	5944	5955	5966	5977	5988	5999	6010
40	6021	6031	6042	6053	6064	6075	6085	6096	6107	6117
41	6128	6138	6149	6160	6170	6180	6191	6201	6212	6222
42	6232	6243	6253	6263	6274	6284	6294	6304	6314	6325
43	6335	6345	6355	6365	6375	6385	6395	6405	6415	6425
44	6435	6444	6454	6464	6474	6484	6493	6503	6513	6522
45	6532	6542	6551	6561	6571	6580	6590	6599	6609	6618
46	6628	6637	6646	6656	6665	6675	6684	6693	6702	6712
47	6721	6730	6739	6749	6758	6767	6776	6785	6794	6803
48	6812	6821	6830	6839	6848	6857	6866	6875	6884	6893
49	6902	6911	6920	6928	6937	6946	6955	6964	6972	6981
50	6990	6998	7007	7016	7024	7033	7042	7050	7059	7067
51	7076	7084	7093	7101	7110	7118	7126	7135	7143	7152
52	7160	7168	7177	7185	7193	7202	7210	7218	7226	7235
53	7243	7251	7259	7267	7275	7284	7292	7300	7308	7316
54	7324	7332	7340	7348	7356	7364	7372	7380	7388	7396
N	0	1	2	3	4	5	6	7	8	9

Four-Place Table of Common Logarithms

N	0	1	2	3	4	5	6	7	8	9
55	7404	7412	7419	7427	7435	7443	7451	7459	7466	7474
56	7482	7490	7497	7505	7513	7520	7528	7536	7543	7551
57	7559	7566	7574	7582	7589	7597	7604	7612	7619	7627
58	7634	7642	7649	7657	7664	7672	7679	7686	7694	7701
59	7709	7716	7723	7731	7738	7745	7752	7760	7767	7774
60	7782	7789	7796	7803	7810	7818	7825	7832	7839	7846
61	7853	7860	7868	7875	7882	7889	7896	7903	7910	7017
62	7924	7931	7938	7945	7952	7959	7966	7973	7980	7987
63	7993	8000	8007	8014	8021	8028	8035	8041	8048	8055
64	8062	8069	8075	8082	8089	8096	8102	8109	8116	8122
65	8129	8136	8142	8149	8156	8162	8169	8176	8182	8189
66	8195	8202	8209	8215	8222	8228	8235	8241	8248	8254
67	8261	8267	8274	8280	8287	8293	8299	8306	8312	8319
68	8325	8331	8338	8344	8351	8357	8363	8370	8376	8382
69	8388	8395	8401	8407	8414	8420	8426	8432	8439	8445
70	8451	8457	8463	8470	8476	8482	8488	8494	8500	8506
71	8513	8519	8525	8531	8537	8543	8549	8555	8561	8567
72	8573	8579	8585	8591	8597	8603	8609	8615	8621	8627
73	8633	8639	8645	8651	8657	8663	8669	8675	8681	8686
74	8692	8698	8704	8710	8716	8722	8727	8733	8739	8745
75	8751	8756	8762	8768	8774	8779	8785	8791	8797	8802
76	8808	8814	8820	8825	8831	8837	8842	8848	8854	8859
77	8865	8871	8876	8882	8887	8893	8899	8904	8910	8915
78	8921	8927	8932	8938	8943	8949	8954	8960	8965	8971
79	8976	8982	8987	8993	8998	9004	9009	9015	9020	9025
80	9031	9036	9042	9047	9053	9058	9063	9069	9074	9079
81	9085	9090	9096	9101	9106	9112	9117	9122	9128	9133
82	9138	9143	9149	9154	9159	9165	9170	9175	9180	9186
83	9191	9196	9201	9206	9212	9217	9222	9227	9232	9238
84	9243	9248	9253	9258	9263	9269	9274	9279	9284	9289
85	9294	9299	9304	9309	9315	9320	9325	9330	9335	9340
86	9345	9350	9355	9360	9365	9370	9375	9380	9385	9390
87	9395	9400	9405	9410	9415	9420	9425	9430	9435	9440
88	9445	9450	9455	9460	9465	9469	9474	9479	9484	9489
89	9494	9499	9504	9509	9513	9518	9523	9528	9533	9538
90	9542	9547	9552	9557	9562	9566	9571	9576	9581	9586
91	9590	9595	9600	9605	9609	9614	9619	9624	9628	9633
92	9638	9643	9647	9652	9657	9661	9666	9671	9675	9680
93	9685	9689	9694	9699	9703	9708	9713	9717	9722	9727
94	9731	9736	9741	9745	9750	9754	9759	9763	9768	9773
95	9777	9782	9786	9791	9795	9800	9805	9809	9814	9818
96	9823	9827	9832	9836	9841	9845	9850	9854	9859	9863
97	9868	9872	9877	9881	9886	9890	9894	9899	9903	9908
98	9912	9917	9921	9926	9930	9934	9939	9943	9948	9952
99	9956	9961	9965	9969	9974	9978	9983	9987	9991	9996
N	0	1	2	3	4	5	6	7	8	9

CHART XXXV

Curves of Exponential Functions



Some Uses of Exponential Function Curves

Charging condenser: $i_c = \frac{E}{R}\epsilon^{-\frac{t}{CR}}$

$E_C = E\left(1-\epsilon^{-\frac{t}{CR}}\right)$

Discharging condenser: $i_c = -\frac{E}{R}\epsilon^{-\frac{t}{CR}}$

$E_c = E\epsilon^{-\frac{t}{CR}}$

Growth of current in an inductance: $i_L = \frac{E}{R}\left(1-\epsilon^{-\frac{Rt}{L}}\right)$ $\qquad E_L = E\epsilon^{-\frac{Rt}{L}}$

Hyperbolic Functions

$\sinh a = \frac{\epsilon^{a}-\epsilon^{-a}}{2}$ $\qquad \cosh a = \frac{\epsilon^{a}+\epsilon^{-a}}{2}$ $\qquad \tanh a = \frac{\epsilon^{a}-\epsilon^{-a}}{\epsilon^{a}+\epsilon^{-a}}$

Where a is in radians

INDEX

G

H

R

T

W

Photography >< Motion Picture
Sound engineering

Cat

N. 4.1.75.