PREFACE

This book is the outgrowth of a course in vacuum-tube design given for many years at Stanford University to senior and graduate students in electrical engineering and physics. It is concerned with the determination of vacuum-tube characteristics in terms of the electron action within the tube. The book attempts to bridge the gap between the physical laws that lie behind the electron behavior and the external characteristics of the tubes themselves.

It is hoped that the point of view taken will be acceptable to both physicists and engineers. The development of the physical laws involved is indicated, after which emphasis is placed upon their description and utilization. Although this book cannot pretend to give much design information, the attempt has been to include enough of the basic relations, physical data, and significant references to make it a useful reference source to vacuum experimenters and tube designers.

Vacuum tubes may seem a rather special subject to which to restrict the material in a book. Actually this is not so. In preparing the book so much material was collected that the contents had to be restricted to first-order effects. It is felt that although engineers and physicists working with vacuum tubes are primarily concerned with the utilization of already developed tubes, the successful application of these tubes is greatly enhanced by a knowledge of their limitations and an understanding of the origin of their characteristics. This is particularly true since there are many occasions when it is desired to use tubes under conditions different from those specified by the manufacturer. Under these conditions it is imperative to know how far one may depart from recommended operating conditions without exceeding some design limitation of the tube. This, in turn, requires a knowledge of how the tube operates.

Circuits and tube applications are so completely covered in the textbook and periodical literature that no effort has been made to include information on these subjects. Only in the case of ultra-high-frequency tubes where the tube cannot be completely separated from the circuit have circuit considerations been included.

The author is indebted to many people for assistance rendered in the preparation of this book. He is particularly indebted to Dr. F. E. Ter-

man, dean of the Stanford School of Engineering, who was a constant source of inspiration and encouragement, and who made many valuable suggestions and gave much direct assistance in checking the work. The author is also indebted to Prof. Paul Kirkpatrick, head of the Physics Department at Stanford, for suggestions on the material of Chaps. 3 to 6 and 9; to Prof. L. Marton for suggestions on the material of Chaps. 13 to 15 and 20; and to C. V. Litton for much information and suggestions relative to Chap. 21. He is indebted to Evelyn G. Sarson, who typed a large part of the manuscript in its final form. O. O. Pardee and Will Harman assisted in the correction of the entire work. Lastly, the author is more than a little indebted to his wife, who personally typed much of the manuscript and was a source of constant assistance.

KARL R. SPANGENBERG

PALO ALTO, CALIF.
January, 1948

CONTENTS

PAGE

PREFACE vii

CHAPTER 1—INTRODUCTION 1

1.1 Devices Using Electron Tubes 1
Radio Receivers—Radio Transmitters—Long-distance Wire Telephones—Television Systems—Measurement Devices—Industrial Control

1.2 Functions of Vacuum Tubes 2
Rectification—Amplification—Oscillation—Frequency Conversion—Modulation — Detection — Light-image Production — Photoelectric Action

CHAPTER 2—BASIC TUBE TYPES 5

2.1 Vacuum Diode 5
2.2 Vacuum Triode 6
2.3 Screen-grid Tube 7
2.4 Pentode 8
2.5 Beam-power Tube 9
2.6 Cathode-ray Tubes 12
2.7 Klystron 13
2.8 Magnetron 15
2.9 Phototubes 17

CHAPTER 3—ELECTRONS AND IONS 19

3.1 The Electron 19
3.2 The Proton 19
3.3 Other Fundamental Particles 20
3.4 Atoms and Molecules 20
3.5 Ions 22

CHAPTER 4—ELECTRONIC EMISSION 23

4.1 Theory of Thermionic Emission 24
Work Function—The Emission Equation—Types of Emitter

4.2 Emission of Pure Metals 35
Tungsten—Tantalum

4.3 Atomic-film Emitters 39
4.4 Oxide Emitters 42
Theory of Oxide Emission—Activation of Oxide Emitters—Specific Emission Characteristics—Transient Emission

4.5 Schottky Effect 46
4.6 Contact Difference of Potential 48

PAGE

4.7 Secondary Emission . . . 48
Variation of Secondary Emission with Primary-electron Potential—Velocity Distribution of Secondary Electrons—Variation of Secondary Emission of Composite Layers—Secondary Emission of Insulators

CHAPTER 5—DETERMINATION OF POTENTIAL FIELDS . . . 58
5.1 Units and Dimensions . . . 58
5.2 Fundamental Quantities and Definitions . . . 59
Forces between Charges
5.3 Solution of Potential Fields by Summation of Intensities . . . 61
5.4 Summation of Potentials . . . 62
5.5 Gauss's Law . . . 64
5.6 Poisson's and Laplace's Equations . . . 67
Interpretations of Laplace's Equation—Solutions of Laplace's Equation in Two Dimensions—Difference Form of Laplace's Equation
5.7 Elastic-membrane Models of Potential . . . 75
5.8 Current-flow Models of Potential . . . 76
5.9 Sketching of Flux and Potential Fields . . . 80
Properties Useful in Sketching Fields
5.10 Method of Conformal Transformations . . . 82
Complex Functions Satisfy Laplace's Equation—Definition of Analytic Functions—The Logarithmic Transformation—The Function $W = Z^{1/n}$

CHAPTER 6—LAWS OF ELECTRON MOTION . . . 97
6.1 Electron in a Uniform Electric Field . . . 97
6.2 Initial Velocity not Parallel to Field . . . 99
6.3 Electrostatic Deflection of Cathode-ray Beams . . . 101
6.4 Relativity Correction for Velocity . . . 103
6.5 Two-dimensional Electric Fields . . . 107
6.6 Electron in a Uniform Magnetic Field . . . 111
6.7 Behavior of Electrons in Nonuniform Magnetic Fields . . . 114
6.8 Combined Electric and Magnetic Fields . . . 116
6.9 Approximate Numerical and Graphical Methods for Determining Electron Paths . . . 121
Method of Joined Circular Segments—Use of Elastic-membrane Model of Potential to Determine Electron Paths—Application of the Principle of Least Action

CHAPTER 7—THE ELECTROSTATIC FIELD OF A TRIODE . . . 125
7.1 Method of Solution . . . 125
7.2 Electrostatic Field of a Plane-electrode Low-mu Triode . . . 125
Contour Representation of Potential Field—Profile Representation of Potential Field
7.3 Electrostatic Field of a Low-mu Cylindrical-electrode Triode . . . 135
Potential Contours of a Cylindrical Triode—Potential Profiles of a Cylindrical Triode
7.4 Analysis of the High-mu Triode . . . 142
Potential Contours and Profiles—Amplification Factor of a High-mu Plane-electrode Triode—Amplification Factor of a High-mu Cylindrical Triode

PAGE

7.5 The Equivalent Electrostatic Circuit of a Triode . . . 152
7.6 Equivalent-diode Spacing of a Triode . . . 153
Diode Equivalent to a Plane-electrode Triode—Diode Equivalent to a Cylindrical-electrode Triode
7.7 Application of Amplification-factor Formulas to Actual Triodes . . . 156
7.8 More Accurate Amplification-factor Formulas . . . 158
Formula for Small Grid-plate Spacings—Formulas for Small Screening Fraction—Formula for Small Cathode-grid Spacings
7.9 Amplification Factor of Unconventional Tubes . . . 165

CHAPTER 8—SPACE-CHARGE EFFECTS . . . 168
8.1 Effects of Current Flow . . . 168
8.2 Plane-electrode Space-charge Flow . . . 170
8.3 Cylindrical-electrode Space-charge Flow . . . 173
8.4 Space-charge Flow for Other Geometries . . . 181
Spherical Electrodes—The General Case
8.5 Current Law for Plane Triodes . . . 183
Current Law in Terms of Electrode Dimensions
8.6 Mutual Conductance of a Plane Triode . . . 188
8.7 Mutual Conductance of a Cylindrical Triode . . . 188
8.8 Effect of Filamentary Emitters . . . 189
8.9 Effect of Initial Electron Velocity . . . 191
8.10 Effect of Space Charge upon Transit Time in Diodes . . . 195
8.11 Summary . . . 198

CHAPTER 9—TRIODE CHARACTERISTICS . . . 201
9.1 Control Action of the Grid . . . 201
9.2 Current-voltage Characteristics of the Triode . . . 202
Plate-current–Grid-voltage Characteristics — Plate-current–Plate-voltage Characteristics—Contours of Constant Plate Current—The Plate-current Surface
9.3 Definition of Triode Constants . . . 205
Amplification Factor—Mutual Conductance—Plate Resistance—Relation between Tube Constants—Variation of Tube "Constants"
9.4 Effective Tube Constants of Combinations of Tubes . . . 212
9.5 Electron Paths . . . 213
9.6 Grid Current . . . 218
Grid-current–Grid-voltage Characteristics—Grid-current–Plate-voltage Characteristics—Constant-grid-current Contours—The Grid-current Surface—Effect of Secondary Electrons
9.7 Primary-grid-current Law . . . 224
Current-division Factor—Approximate Primary-grid-current Law—Current-division-factor Formula—Current-division Law in the Presence of Secondary Emission

CHAPTER 10—TETRODES . . . 238
10.1 Types of Tetrode . . . 238
10.2 Current-voltage Characteristics of the Screen-grid Tube . . . 238
Plate-current–Plate-voltage Characteristics of Screen-grid Tube—Screen-current–Plate-voltage Characteristics of the Screen-grid Tube—General Characteristics of Screen-grid Tubes

PAGE

10.3 Beam-power Tubes 245
10.4 The Electrostatic Field of a Beam-power Tube 245
10.5 Space-charge Effects in the Screen-grid–Anode Region of Beam-power Tubes 248
Type *A* Distribution—Type *B* Distribution—Type *C* Distribution—Type *D* Distributions
10.6 Dynamic Characteristics of Beam-power Tubes 259
Injected Current Varied, Potentials Constant—Plate Potential Varied, Screen Potential and Injected Current Constant

CHAPTER 11—PENTODES 266
11.1 Electrode Arrangement in a Pentode 266
11.2 Current-voltage Characteristics of the Pentode 267
Plate-current–Control-grid-voltage Characteristics—Plate-current–Plate-voltage Characteristics of a Pentode—Space-current–Plate-voltage Characteristics of the Pentode—Screen-grid-current–Plate-voltage Characteristics of a Pentode—Suppressor-grid Effects
11.3 Current Division in Pentodes 272
11.4 Amplification Factor of a Pentode 278
Electrostatic Field of a Pentode—Electrostatic Amplification Factor of a Pentode—True Amplification Factor of a Pentode
11.5 Transconductance of a Pentode 288
11.6 Plate Resistance of a Pentode 288
11.7 Design Considerations 289

CHAPTER 12—NOISE IN VACUUM TUBES 298
12.1 Noise as a Limiting Factor in the Ultimate Sensitivity of Electronic Devices 298
12.2 Noise in Resistors 299
12.3 Sources of Noise in Tubes 305
12.4 Shot Noise in Diodes with Temperature-limited Emission 306
12.5 Reduced Shot Effect in Diodes with Space-charge-limited Emission . . 308
12.6 Reduced Shot Effect in Triodes with Space-charge-limited Current . . 310
12.7 Noise Due to Gas in Tubes 312
12.8 Reduced Shot Effect in Multielectrode Tubes with Space-charge-limited Currents 313
12.9 Noise in Mixer Tubes 314
12.10 Noise Induced at Ultra-high Frequencies by Random Emission 316
12.11 Noise in Velocity-modulation Tubes 317
12.12 Noise in Phototubes 318
12.13 Noise in Secondary-emission Multipliers 319
12.14 Definition of Noise Figure 321
Noise Figure for Two Networks in Cascade
12.15 Measurement of Noise and Noise Figure 325
Typical Tube-noise Values

CHAPTER 13—ELECTROSTATIC ELECTRON OPTICS 328
13.1 Introduction 328
Snell's Law—The Principle of Least Action—Simple Lenses—Lens Formulas

PAGE

13.2 Electrostatic-lens Fields 337
General Form of Fields with Rotational Symmetry—The Equal-diameter Two-cylinder Lens—Equal-diameter Spaced Cylinders—Two-diameter Cylinder Lenses—Aperture Lenses
13.3 Electron Paths 349
13.4 General Lens Properties 350
Thin Lenses—Thick Lenses
13.5 Calculation of Lens Characteristics 360
Method of Linear Axial-potential Segments—Method of Equivalent Thin Lenses
13.6 Measurement of Lens Characteristics 365
Double-grid Method of Measuring Lens Characteristics
13.7 Optical Characteristics of Lenses 369
13.8 Calculation of Lens Characteristics 375
13.9 *P-Q* Curves of Lenses 377
Comparison of Lenses—The *Einzel* Lens
13.10 Aberrations 387
Chromatic Aberration—Coma—Astigmatism—Curvature of Field—Distortion of Field—Spherical Aberration—Other Lens Defects

CHAPTER 14—MAGNETIC LENSES 394
14.1 Focusing Action of Axial Magnetic Fields 394
14.2 Magnetic Fields with Rotational Symmetry 396
14.3 Electron Motion in a Magnetic Field Expressed in Cylindrical Coordinates 397
14.4 Differential Equations of Motion of the Paraxial Electron 398
14.5 Focusing Properties of Magnetic Lenses 399
General—Magnetic Lens of a Circular Turn of Wire—The Glazer Lens
14.6 Practical Magnetic Lenses 404
14.7 Magnetic-lens Defects 405
14.8 The General Equations of Motion in Combined Electric and Magnetic Fields 406

CHAPTER 15—CATHODE-RAY TUBES 412
15.1 The General Form of Cathode-ray Tubes 412
15.2 Electron-gun Design 414
Cutoff Relations in the Electron Gun—Electron Paths in the Electron Gun—Focusing System—Alternative Electrode Structures
15.3 Deflection Devices 425
Electrostatic Deflecting Plates—Magnetic Deflection—Relative Merits of Electrostatic and Magnetic Deflection—Visual versus Deflection Sensitivity—Postdeflection Acceleration
15.4 Fluorescent Materials 429
Definitions—General Make-up of Phosphors—Luminous Properties of Fluorescent Materials—Electrical Characteristics of Phosphors
15.5 Limitations of Spot Size 437
Effect of Thermal Velocity of Emission—Space-charge Limitation of Spot Size—Effect of Secondary Emission—Halation

PAGE

15.6 High-efficiency Cathodes . . . 449
Parallel Flow of a Rectangular Beam—Parallel-flow Cylindrical Beam—Convergent Radial Flow of a Conical Beam
15.7 Ultra-high-frequency Deflection Effects . . . 465
15.8 Photography of Cathode-ray Traces . . . 470
Beam Power—Screen Types—Writing Speed—Time, Stop, and Magnification—Film Sensitivity—Developers and Development

CHAPTER 16—ULTRA-HIGH-FREQUENCY EFFECTS IN CONVENTIONAL TUBES . . . 475
16.1 Introduction . . . 475
16.2 Causes of Decreased Output at Ultra-high Frequencies . . . 475
16.3 Onset of Tube-reactance Limitations . . . 477
16.4 The Nature of Currents Induced by Electron Motion at Ultra-high Frequencies . . . 482
The General Case—The Diode without Space Charge—The Diode with Space Charge—Currents Induced in the Electrodes of a Triode
16.5 Onset of Transit-time Effects in Triodes . . . 490
16.6 Transit-time Effects in the Space-charge-limited Diode . . . 495
16.7 Small-signal Transit-time Effects in the Space-charge-limited Triode . . . 501
16.8 Similitude and Scaling in Ultra-high-frequency Triodes . . . 504
16.9 High-frequency Limit of Triode Oscillation . . . 507
16.10 Large-signal Effects . . . 516
Transit-time Effects in Diodes—Transit-time Effects in Triodes—Transit-time Effects in Tetrodes—The Resnatron
16.11 Disk-seal Tubes . . . 524

CHAPTER 17—VELOCITY-MODULATED TUBES OR KLYSTRONS . . . 527
17.1 The Bunching Principle . . . 527
17.2 Cavity Resonators . . . 529
17.3 Mechanism of Energy Interchange between Electrons and Cavity Resonators . . . 537
17.4 First-order Bunching Theory . . . 541
17.5 The Klystron Amplifier . . . 556
Structure of the Klystron Amplifier—Output Power of the Klystron Amplifier—Efficiency of the Klystron Amplifier—Mutual Conductance of the Klystron Amplifier—Power Required to Bunch the Beam
17.6 The Cascade Amplifier . . . 564
17.7 Frequency-multiplier Klystrons . . . 566
17.8 Second-order Bunching Effects . . . 567
17.9 The Reflex-klystron Oscillator . . . 571
Behavior of Electrons in the Reflector Space—Distance-time Diagram of a Reflex-klystron Oscillator—Bunching Theory of the Reflex-klystron Oscillator—Self-admittance of the Beam—Mechanism by Which Oscillations Start—Variation of Beam Conductance with Amplitude of Oscillation—The Electronic-admittance Spiral—Reflex-klystron Oscillation with a Simple Resonant Circuit—Power Relations in the Reflex-klystron Oscillator—Voltage Stability of Reflex-klystron Oscillators
17.10 Broad-band Operation of Reflex-klystron Oscillators . . . 591
Equivalent Circuit of Concentric-line Resonator—Possible Modes of

PAGE

Oscillation—Method of Calculating Oscillation Mode Plot—Mode Interference
17.11 The Two-resonator Klystron Oscillator . . . 606
17.12 The Heil Tube . . . 616
17.13 Bunching Effects in Negative-grid Tubes . . . 616

CHAPTER 18—MAGNETRON OSCILLATORS . . . 621
18.1 Introduction . . . 621
18.2 Structural Form of Magnetrons . . . 622
18.3 Resonant Properties of Multicavity Magnetrons . . . 625
18.4 Electron Behavior in Crossed Static Magnetic and Static Electric Fields: Plane Case . . . 631
18.5 Electron Behavior in Crossed Magnetic and Alternating Electric Fields: Plane Case . . . 636
Alternating Transverse Electric-field Effect of a Traveling Electric Field
18.6 Electron Behavior in Crossed Magnetic and Radial Electric Fields . . . 642
18.7 The Effect of Space Charge . . . 648
18.8 Electron Behavior in Crossed Magnetic and Alternating Radial Electric Fields . . . 651
18.9 Basic Design Relations for Multicavity Magnetrons . . . 660
18.10 Dimensional Relations in Magnetrons . . . 665
18.11 Output Characteristics of Magnetrons . . . 667

CHAPTER 19—PHOTOELECTRIC TUBES . . . 675
19.1 The General Form of Photoelectric Tubes . . . 675
19.2 Fundamental Photoelectric Relations . . . 675
19.3 History of the Photoelectric Effect . . . 676
19.4 Specific Photoemission Characteristics . . . 677
19.5 Fundamental Theory of Photoemission . . . 681
19.6 Spectral Response Curves of Photoemissive Surfaces . . . 683
19.7 Vacuum-phototube Characteristics . . . 684
Current-voltage Characteristics—Spectral Characteristics
19.8 Gas-phototube Characteristics . . . 688
Factors in the Design of Gas Phototubes—Frequency Distortion in Gas Phototubes—Summary of Gas-phototube Characteristics
19.9 Utilization of Phototube Characteristics . . . 693
19.10 Photomultiplier Tubes . . . 694

CHAPTER 20—SPECIAL TUBES . . . 701
20.1 Introduction . . . 701
20.2 The Hexode . . . 702
20.3 The Heptode . . . 710
20.4 The Pentagrid Converter . . . 714
20.5 The Octode . . . 716
20.6 Space-charge Grid Tubes . . . 717
20.7 Negative-resistance Tubes . . . 718
Glow-discharge Tubes—The Dynatron—Direct-coupled Negative-resistance Devices—Negative Screen-grid Resistance of a Pentode—

PAGE

Push-pull Negative-resistance Circuit—Feedback Circuits—Special Negative-resistance Tubes
20.8 Negative-transconductance Tubes 723
20.9 Electron-ray Indicator Tubes 723
20.10 Directed-ray Electron Tubes 724
20.11 Deflection Tubes 727
20.12 Television Camera Tubes 728
The Image-dissector Tube—The Iconoscope—The Image Iconoscope—The Orthicon—The Image Orthicon—The Monoscope
20.13 The Electron Microscope 738
Structure of the Electron Microscope—Equivalent Wave Length of Electrons—Theoretical Resolving Power of the Electron Microscope—Operating Principle of the Electron Microscope—Limits of Resolving Power—Electrostatic Electron Microscopes

CHAPTER 21—HIGH-VACUUM PRACTICE 747
21.1 Introduction 747
21.2 Fundamental Gas Laws 749
Boyle's Law—Charles's Law or Gay-Lussac's Law—Avogadro's Law—General Gas-expansion Law—Distribution of Velocities in a Gas—Mean Free Path of a Gas Molecule—Mean Free Path of an Electron among Gas Molecules
21.3 Measurement of Vacuum 757
Manometers—The McLeod Gauge—The Spark-discharge Tube—The Pirani Gauge—The Thermocouple Gauge—Triode Ionization Gauge
21.4 Pumping Speed 775
Speed of an Aperture—Definition of Pump Speed—Speed of Tubing—Effect of Tubing Upon Pumping Speed
21.5 Production of Low Vacuum 780
21.6 Production of High Vacuum 781
The Mercury-diffusion Pump—Oil Pumps—Fractionating Pumps
21.7 Glass and Its Properties 791
Composition of Glass—Physical Properties of Glass—Working of Glass
21.8 Sealing of Glass to Other Materials 796
Sealing of Small Leads Into Glass—Copper-to-glass Seals—Kovar and Fernico—Glass-to-porcelain Seals—Glass-to-mica Seals
21.9 Metals Useful in Tube Construction 801
Nickel—Copper—Aluminum—Molybdenum—Tantalum—Tungsten—Relative Properties of the Metals—Spot Welding
21.10 Insulators 807
21.11 Degassing of Glass and Metals 808
21.12 Getters 809

APPENDIXES
I. Properties of the Elements 811
II. Differential Operators and Vector Notation 813
III. A Note on Mks Units 817
IV. Characteristics of Fluorescent Screens 821

PAGE

V. Surface Resistance and Depth of Penetration of Current Resulting from Skin Effects 822
VI. Principal Properties of the Bessel Functions 823
VII. Values of α^2 as a Function of r_c/r for use in Eq. (15.63) 825
VIII. Nomographic Chart Relating Object and Image Distance to the Focal Length of a Thin Lens 826
IX. Nomographic Chart Relating Tube Constants 827
X. Designation of Frequency Bands 828

PROBLEMS 829

NAME INDEX 849

SUBJECT INDEX 853

CHAPTER 1

INTRODUCTION

VACUUM tubes are found as basic or auxiliary elements in numerous technical devices now in use. They are indispensable in communication systems and industrial control. Their development has facilitated advances in the fields of power and transportation. Without the vacuum tube we should be back in the days of the gravity-cell telegraph and the ringer telephone.

In the United States the number of vacuum tubes in use is several times the number of human beings and household pets. The 50,000,000 radio sets manufactured in the United States in the year 1947¦ alone contained more vacuum tubes than the adult population of the country. Associated with the 25,000,000 telephones and 120,000,000 miles of telephone and telegraph wire in the United States are many more vacuum tubes. Various industrial devices include almost as many more. The United States uses nearly half the world's total of vacuum tubes.

One may conclude that there are many vacuum tubes in use. They must be of some importance. They are.

1.1. Devices Using Vacuum Tubes. This book is more concerned with the properties and functions of vacuum tubes than with the systems utilizing these properties. However, it is well to be reminded of the extent of vacuum-tube applications and the degree to which we are dependent upon them. The following devices are totally dependent upon vaccum tubes.

Radio Receivers. These are too well known to require much description. They range from portable receivers the size of a brick and capable of receiving local broadcast stations to large-size all-wave receivers capable of picking up a signal stronger than the noise level from any point on the globe. Even the smallest receivers use 4 or 5 vacuum tubes. The average home receiver has about 7 tubes. An all-wave receiver may have 20 or more tubes.

Radio Transmitters. Transmitters range from portable walkie-talkie sets to large power-broadcast and short-wave stations. In output power they vary from 0.1 watt to hundreds of kilowatts. In frequency they may range from 100 kc to 60,000 mc. The short-wave transmitters are capable of producing an audible signal at any point on the earth's surface.

Transmitters may use voice or code. They may incorporate static-elimination or secrecy features in their operation. A small transmitter may use only a few vacuum tubes. The largest transmitters may use 50 or more tubes.

Long-distance Wire Telephones. The connections between telephone stations on the same continent are effected by wire transmission lines rather than by radio. When the distance between telephone stations is large, it is necessary to amplify the speech energy about every 16 miles for cables and every 50 miles for open-wire lines. Each speech amplifier contains several vacuum tubes and amplifies the speech power from about 10 microwatts to about 1 milliwatt, a power amplification of 100. Thus a telephone call from San Francisco to New York passes through 30 or more speech amplifiers.

Television Systems. Television systems achieve the modern miracle of reproducing a visual scene at a point remote from the original. This is done entirely with vacuum tubes and electrical-circuit elements. No mechanical devices are needed. In its present stage of development the reproduced picture as viewed from 6 ft on an 8-in. cathode-ray-tube screen is as good as a motion picture seen from the first row of the balcony. Each television transmitter contains hundreds of vacuum tubes, including a special camera tube. Every television receiver contains 20 or more tubes, including a special viewing tube.

Measurement Devices. Electronic measurement devices are too numerous to mention. Quantities that can be measured, besides all the electrical quantities, are color, weight, light intensity, odor, time interval, and many others. In fact, it can be said that any quantity which can be measured at all can probably be measured by electronic means.

Industrial Control. The number of electronic industrial-control devices is legion. They include counting circuits, sorting systems, illumination-control systems, welding-control devices, and liquid- and gaseous-flow regulators. Typical devices are those which automatically regulate temperature or humidity. All these devices have their primary dependence upon the vacuum tubes in them.

In addition to the above devices, which are totally dependent upon vacuum tubes, there are many others that have acquired a strong dependence upon electronic devices. Thus all commercial flying makes constant use of radio communications to keep posted on the weather and on terminal traffic and to keep ground stations posted on plane positions as well as to guide the planes directly. The invasion of other fields by electronics has already been considerable and is bound to be greater in time to come.

1.2. Functions of Vacuum Tubes. Although the applications of vacuum tubes are almost infinite, the specific functions that vacuum tubes can perform by virtue of their own properties are relatively few. It is these few fundamental functions and their combinations that give rise to the numerous applications.

A list of the functions of vacuum tubes is bound to be an arbitrary one since the tube cannot function by itself without an associated circuit. However, some of the jobs that vacuum tubes can perform are so fundamental that they may be considered properties of the tube itself, independent of the associated circuits.

The principal functions that may be performed by vacuum tubes are listed below.

Rectification. Vacuum tubes are able to convert alternating currents to direct currents. This is known as "rectification." Rectification is an inherent property of vacuum tubes because current can flow in only one direction from a source of electrons.

If a sinusoidal wave of voltage is applied to a vacuum tube of the right type, current will flow in only one direction, giving rise to a succession of half-wave pulses all of the same polarity. It is possible to connect another like tube to insert half-wave pulses of the same polarity between the pulses of the first tube. The average of these pulses constitutes a direct current; the other frequency components are rejected by a filter circuit.

Rectification is important because electronic devices operate best on direct current, while power is usually generated and transmitted in alternating form. It is thus necessary to convert, or rectify, the a-c power to d-c power.

Amplification. The amplification of voltage or power is the outstanding function that vacuum tubes are able to perform. With the exception of the mechanical torque amplifier, no other device can do anything like it. Strictly speaking, the vacuum tube does not amplify power but rather controls the flow of a relatively large amount of power from one source with a small amount of power from another source. The British use the expression "electric valve" for certain types of electron tubes. This term is really better than ours, for it indicates the nature of the amplifying action.

Oscillation. The generation of high-frequency alternating currents, or oscillation, is another remarkable function that vacuum tubes can perform. Oscillation is obtained by causing part of the output of an amplifier to excite the amplifier and thus make the device self-excited and self-sustaining. Tubes can be built that will produce oscillations at frequencies as low as 1 cycle per sec, while other tubes can be built that will oscillate at frequencies as high as 60,000 mc per sec.

Frequency Conversion. Vacuum tubes are able to shift the frequency of a wave. This they are able to do by an electrical "beat" action.

Thus a wave of a given frequency can be mixed with a wave of another frequency in a vacuum tube, and among the products of the interaction is found the difference of the two frequencies. If one of the original waves had certain effects associated with it, these same effects are associated with the difference frequency. The beat action results from the nonlinear characteristics of the vacuum tube.

Modulation. The transmission of intelligence by radio waves or by certain types of wire telephony requires the use of frequencies higher than those audible. It is necessary to superimpose the audible frequencies upon the higher transmitted frequency. This superimposition is known as "modulation." Modulation is best performed by vacuum tubes.

Basically, modulation takes the form of varying some property of the r-f wave at the audible rate. The commonest form of modulation varies the amplitude of the r-f wave in accordance with the intelligence to be transmitted. This is known as "amplitude modulation." Frequency modulation is also common.

Detection. Detection is the inverse of modulation and is sometimes known as "demodulation." It is the process of extracting the intelligence from the modulated wave. In the case of amplitude modulation the detection may be effected by rectifying the r-f wave and then utilizing the average value of the rectified wave, since it follows the amplitude variations in magnitude. Detection of modulated radio signals is best performed by vacuum tubes over most of the range of radio frequencies.

Light-image Production. It is possible for vacuum tubes to convert part of their output energy into visible light. This is done in cathode-ray tubes in which a stream of electrons is caused to hit a fluorescent screen, causing light to be emitted. The cathode-ray tube can be used for viewing wave forms and for doing many other wonderful things, including the reproduction of visual scenes. The fundamental property involved here is the conversion of electrical energy into visual energy.

Photoelectric Action. Vacuum tubes can be made that will convert light energy into electrical energy. This is possible by virtue of the photoelectric effect, which is the emission of electrons from certain surfaces when illuminated with visual energy. The liberated electrons constitute an electric current whose measure is related to the frequency and intensity of the exciting light. Tubes making use of this principle are known as "photoelectric tubes." The photoelectric tube is one of the tubes most extensively used in industrial-control systems.

The above paragraphs have given a bird's-eye view of the functions of vacuum tubes. The reader is probably familiar with all the above functions, which are now commonly encountered in everyday life. The rest of the book is devoted to the description and explanation of the characteristics of the vacuum tubes themselves.

CHAPTER 2

BASIC TUBE TYPES

The electronic engineer has about a dozen *types* of vacuum tube he can call upon for his high-frequency and industrial-control circuits. This is a surprisingly small number of distinct tube types. The small number of types is balanced, however, by the large number of forms in which each type may appear, as determined by the required power capacity and frequency range.

The purpose of this chapter is to list the basic types and their fundamental characteristics as a prelude to a detailed study of their characteristics and the physical laws from which these are derived.

2.1. Vacuum Diode. The vacuum diode is a two-electrode vacuum tube. One electrode acts as an emitter of electrons and is called the "cathode." The other electrode acts as a collector of electrons and is called the "anode" or "plate." The emitter may be either directly or indirectly heated. In physical form the vacuum diode may vary from a small metal tube to a large glass rectifier tube.

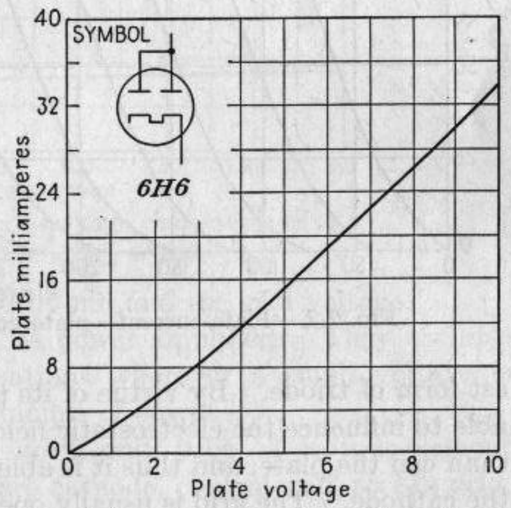


FIG. 2.1.—Plate-current—plate-voltage characteristics of a diode.

The current-voltage characteristics of a typical diode are shown in Fig. 2.1. The current follows a three-halves-power law of voltage over the normal range of operation. At high values of plate voltage or at low values of heater current the plate current tends to be limited by the cathode emission and to increase only very slowly with plate voltage.

The most useful property of the diode is that it passes current only in one direction. This property makes the diode useful as a detector and as a rectifier for d-c power supplies.

2.2. Vacuum Triode. A vacuum triode is a three-electrode tube containing an emitting electrode called the "cathode," a control electrode called the "grid," and a current-collecting electrode called the "anode" or "plate."

The emitting electrode may be an indirectly heated oxide cathode, an oxide-coated filament, or a filament of tungsten or thoriated tungsten.

The control electrode, usually in the form of a grid of fine wire, surrounds the emitter and is in turn surrounded by the plate in the common-

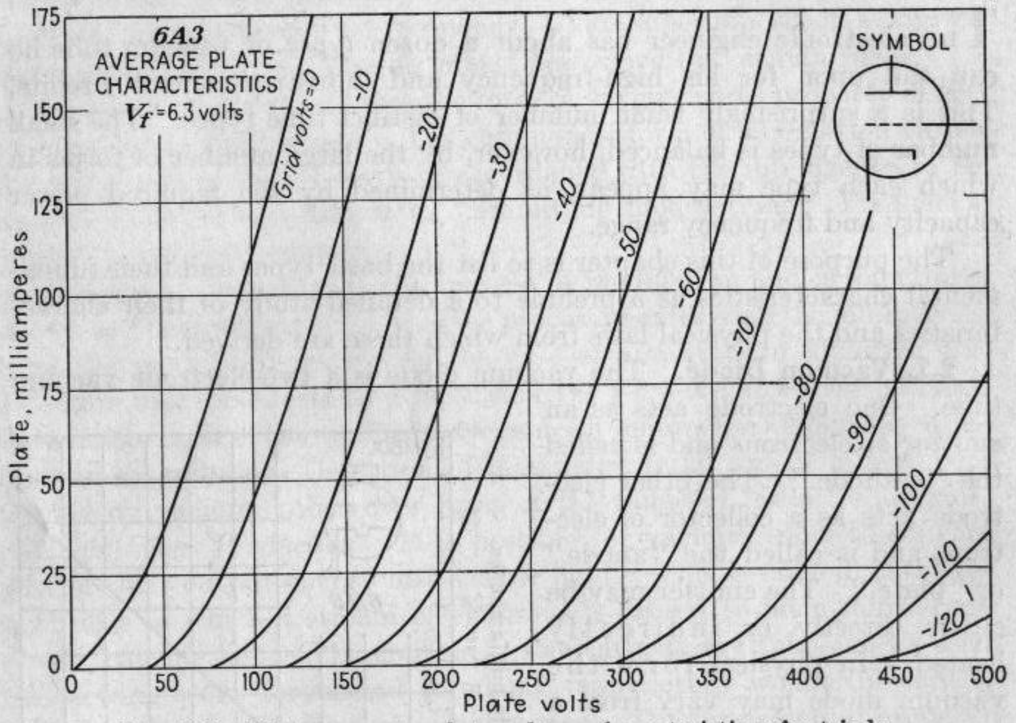


FIG. 2.2.—Plate-current—plate-voltage characteristics of a triode.

est form of triode. By virtue of its proximity to the cathode the grid is able to influence the electrostatic field at the cathode to a greater extent than can the plate, and thus it is able to control the flow of current from the cathode. The grid is usually operated on a slight negative potential so that the electrons will pass between the grid wires without hitting the wires themselves.

Some typical characteristics of a triode illustrating the variation of plate current with plate voltage for various fixed values of grid voltage are shown in Fig. 2.2. The plate current increases if either grid or plate voltage is increased. The increase in plate current for a given increase in grid voltage is always much larger than the increase in plate current for the same increase in plate voltage.

The relative effectiveness of the plate and grid potentials in controlling the plate current is known as the amplification factor of the tube (mu; symbol μ). The amplification factor is the *maximum* amplification that can be obtained by using the tube as an amplifier. With triodes the useful amplification is about two-thirds of the amplification factor.

Study of the family of curves of Fig. 2.2 shows that all the curves are alike in shape and further are somewhat similar to the characteristic of a diode. This is true in that the plate current of a triode is found to vary nearly as the three-halves power of an equivalent voltage which is

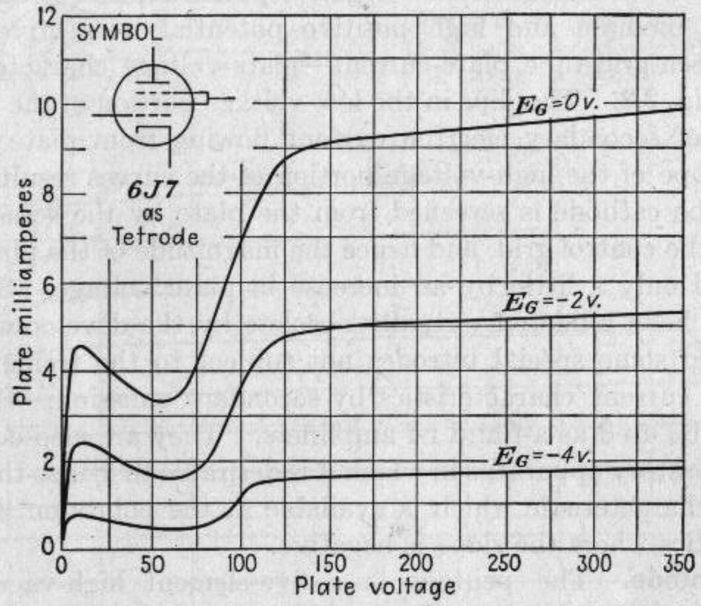


FIG. 2.3.—Plate-current—plate-voltage characteristics of a screen-grid tube. V_f = 6.3 volts, V_{sg} = 100 volts.

the sum of the plate voltage divided by mu and the grid voltage.

Triodes have their greatest use as power amplifiers. They are also used extensively in control applications wherever a small voltage is wanted to control an appreciable amount of current.

2.3. Screen-grid Tube. The screen-grid tube is a four-element vacuum tube. The four elements are cathode, control grid, screen grid, and plate. The electrode construction is similar to that of the triode except that an extra grid of mesh a little coarser than that of the control grid is inserted between the control grid and the plate.

The screen-grid tube is the historical predecessor of the pentode. Its invention was the result of an effort to overcome a limitation of the triode. Triodes do not work well as amplifiers of high frequencies, for the high interelectrode capacity between plate and grid causes the tube to regenerate and oscillate. In the screen-grid tube the capacity between

the control grid and plate is reduced by inserting the extra grid, known as the "screen grid," between these elements. The insertion of the screen grid and its operation at a constant potential succeeded in producing the low control-grid—plate capacity desired but caused distortions in the plate-current—plate-voltage characteristics, for the new electrode arrangement permitted secondary electron flow between plate and screen grid. This detrimental effect was overcome in the pentode by the addition of a coarse-mesh suppressor grid between screen grid and plate.

The screen-grid tube is usually operated with cathode near ground potential, control grid at a small negative potential, and screen grid and plate at a medium and high positive potential, respectively. Some typical screen-grid-tube plate-current—plate-voltage characteristics are shown in Fig. 2.3. The dips in the low-voltage portion of the curves are the result of secondary electron current flowing from plate to screen. The low slope of the high-voltage portion of the curves results from the fact that the cathode is screened from the plate by the screen grid as well as by the control grid, and hence the magnitude of the plate current is increased only a little by an increase in plate voltage. Screen-grid tubes have been rendered virtually obsolete by the development of the pentode and some special tetrodes not subject to the tremendous distortions of current characteristics by secondary emission. Screen-grid tubes may be used as a-f and r-f amplifiers. They are also occasionally used in laboratory apparatus in which it is desirable to utilize the negative resistance characteristic which is available at the points on the current characteristics where the slope is negative.

2.4. Pentode. The pentode is a five-element high-vacuum tube. The five electrodes, in the order in which they occur in the tube, are cathode, control grid, screen grid, suppressor grid, and plate. In normal operation the cathode is operated near ground potential, the control grid at a small negative potential, the screen grid at a relatively large positive potential, the suppressor grid at cathode potential, and the plate at the screen potential or a more positive potential.

Some typical plate-current—plate-voltage curves of a pentode are shown in Fig. 2.4. In these it is seen that the insertion of the suppressor grid at cathode potential between screen grid and plate has eliminated the distortions in the characteristic observed in the case of the screen-grid tube. This it does by causing a negative potential gradient at both the screen grid and plate, which suppresses the secondary electrons from these electrodes. The slope of the plate-current characteristic for high plate voltages is even less than in the screen-grid tube, for there is another screening grid between plate and cathode in the pentode. The result of this high screening action is to make the amplification factor of the

pentode extremely high, of the order of 1,000 or more, and to give the tube a high effective resistance in the plate circuit. The pentode is, in fact, very nearly a constant-current device. The variation of plate current with grid voltage, which is measured by a factor known as the "grid-plate transconductance" or, more commonly, the "mutual conductance" of the tube, is about the same as in the triode. Only about one-tenth of the high amplification factor of the pentode can be realized

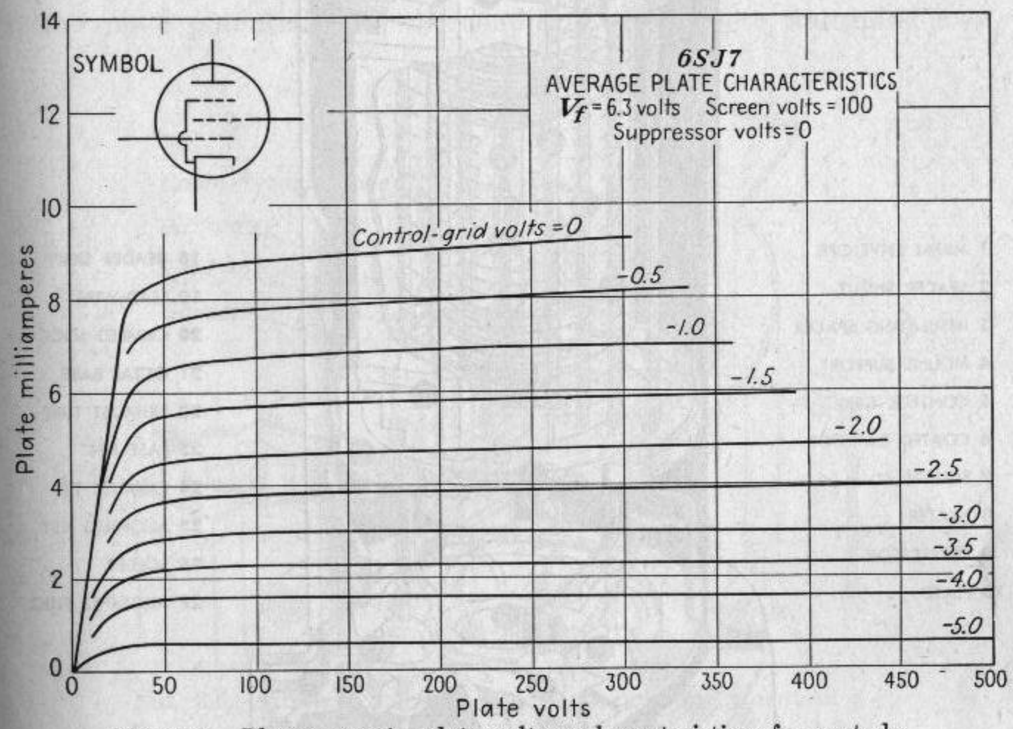


FIG. 2.4.—Plate-current—plate-voltage characteristics of a pentode.

in amplifier operation. However, the reduced plate—control-grid capacity makes the pentode a better tube in voltage-amplifier applications.

The pentode is a versatile tube. It can be connected to give diode, triode, and screen-grid as well as pentode action. It is available in constant- and variable-mu forms. It is probably the most extensively used tube in low-power applications. There are probably more pentodes in use today than any other type of electron tube. A cutaway drawing of a pentode showing the electrode structure is given in Fig. 2.5.

2.5. Beam-power Tube. The beam-power tube is a special type of tetrode. It is designed so that the electrons move from cathode to plate in dense sheets. This effect is achieved by making the control grid and screen grid of the same pitch and aligning the grid wires. The electrode structure of the tube is shown in Fig. 2.6.

The effect of the dense current sheets between the screen grid and plate is to depress the potential between these two electrodes within the

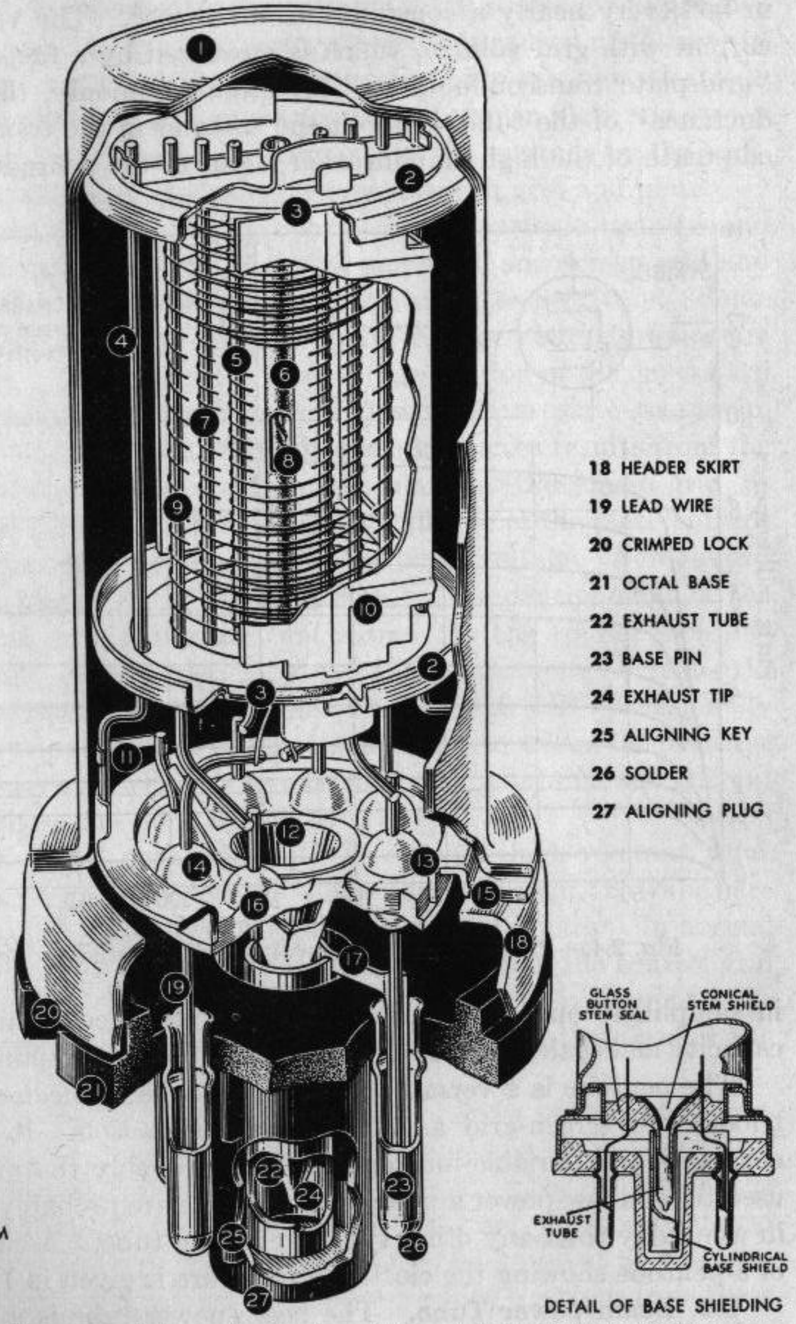


FIG. 2.5.—Cutaway picture of a single-ended metal-envelope pentode.

tube because of the high concentration of negative charge. The potential between screen grid and plate is depressed enough so that secondary electron flow from plate to grid is suppressed without the aid of a suppressor grid. Thus the tube represents another solution to the problem of overcoming the distortions in the current characteristics of the ordinary screen-grid tube.

The plate-current—plate-voltage characteristics of a beam-power tube are shown in Fig. 2.7. It is seen that these characteristics are free of the dip in the shoulder due to secondary electron flow. The distinctive features of the beam-power tube's characteristics as contrasted with the pentode characteristics are that the plate current rises much more rapidly at low plate potentials and the condition of complete transmission of

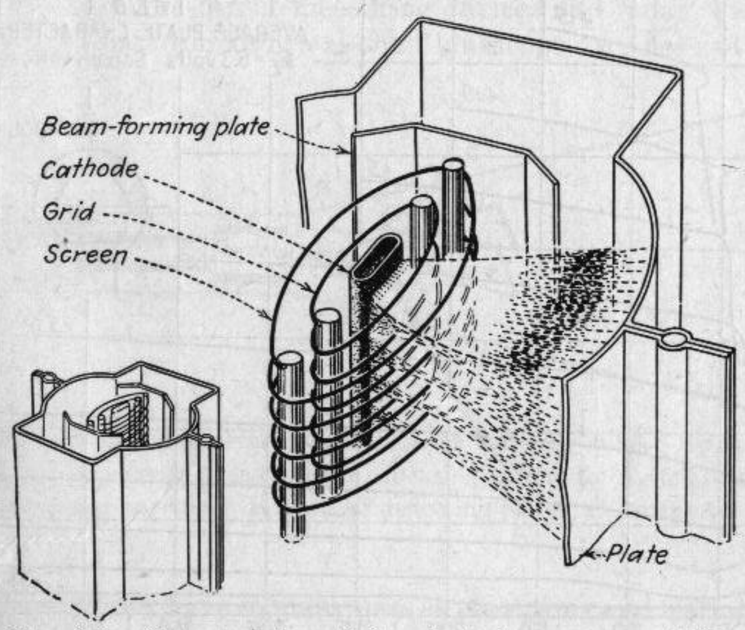


FIG. 2.6.—Cutaway view of the electrode arrangement in a beam-power tube. (*Courtesy of RCA.*)

current to the plate is reached at a lower plate potential. The plate current rises rapidly because the high space-charge density blocks the flow of electrons to the plate at low plate potentials, and this blocking action stops quite abruptly as the plate potential is increased. In the beam-power tube, complete transmission of current passed by the screen grid to the plate occurs when the plate potential has risen to about 20 per cent of the screen-grid potential, whereas in the pentode the transmission is not complete until the plate potential has risen to about 50 per cent of the screen-grid potential. This results from the behavior of the individual electrons, which, in the beam-power tube, are more uniform in direction and velocity than in the pentode, in which the electrons are strongly deflected by the suppressor grid.

The beam-power tube is made in small and medium-size metal tubes

and in a medium-size glass tube. The beam-power tube is used in many ways. It is extensively employed as an audio-power amplifier tube and also as a r-f amplifier and oscillator tube.

2.6. Cathode-ray Tubes. The cathode-ray tube is in a class by itself among the vacuum tubes. It makes use of the geometrical form rather than the intensity of its electron stream and converts the energy of its electron stream into a visual indication. In its commonest application

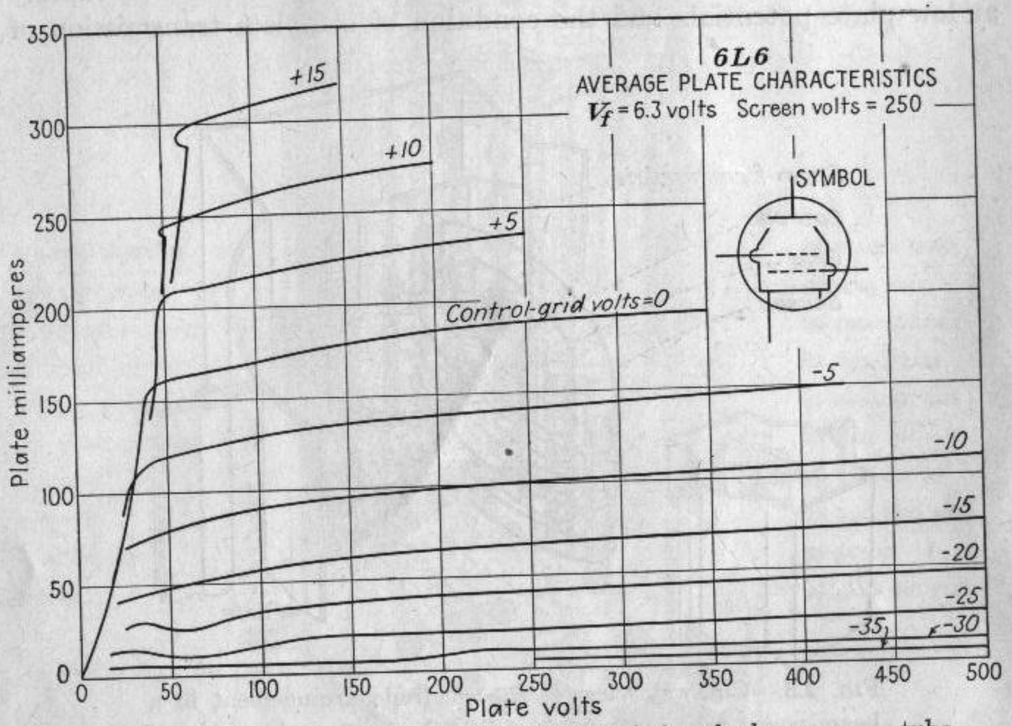


FIG. 2.7.—Plate-current—plate-voltage characteristics of a beam-power tube.

the cathode-ray tube uses its electron beam to show the shape of an applied voltage wave as a light trace upon a fluorescent screen. The cathode-ray tube is an electronic oscilloscope that produces on a screen a light spot that can be deflected in two dimensions.

The cathode-ray tube is generally housed in a large glass envelope shaped like an Erlenmeyer flask. In the neck of the glass envelope is located a set of electrodes known as the "electron gun." This gun serves to produce a circular beam of electrons that is fired at the large end of the envelope, which is covered with a fluorescent material. Also housed in the neck of the envelope are deflecting devices that serve to bend the beam in horizontal and vertical directions. The fluorescent screen on the inside of the large end of the envelope gives off light at the point at which the electron beam strikes.

In physical size the ordinary cathode-ray tubes range from 10 to 20 in. in length and have fluorescent screens from 3 to 5 in. in diameter. The tubes operate with a beam-accelerating potential between 800 and 10,000 volts. The electrode arrangement in a typical cathode-ray tube is shown in Fig. 2.8.

Cathode-ray tubes are principally used to observe electrical wave forms. They may also be used to compare frequencies, plot the *B-H* curves of iron, and plot the current-voltage characteristics of vacuum tubes. They are extensively employed as indicators of elapsed-time intervals in ionosphere height-measuring devices and radar sets. They are built in a special form known as the "kinescope" for use as television

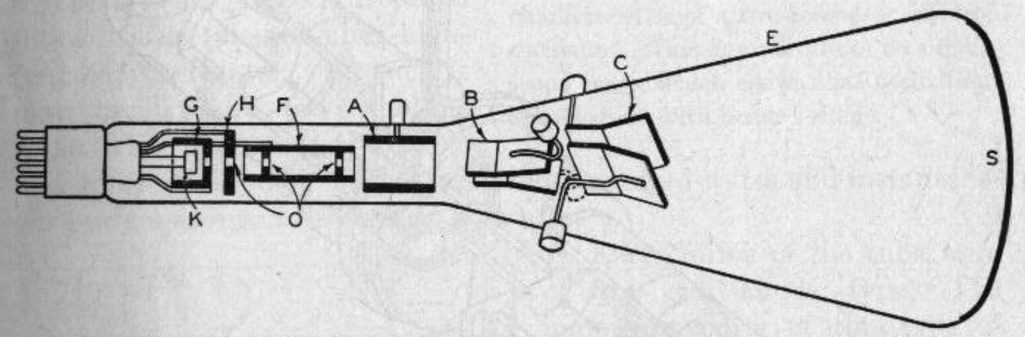


FIG. 2.8.—Typical electrode arrangement in a cathode-ray tube. K, cathode; G, control grid; H, accelerating electrode; F, focusing electrode; A, final accelerating electrode; O, limiting apertures; B, vertical deflecting plates; C, horizontal deflecting plates.

viewing tubes. They have so many uses as measuring and testing devices that no radio or electronic laboratory worthy of the name is without one.

2.7. Klystron. The klystron is a newcomer to the group of vacuum tubes in use today. It is a special ultra-high-frequency tube that is capable of generating, detecting, and amplifying radio waves ranging in frequency from 600 to 30,000 mc (50 to 1 cm).

The principle of operation of the klystron amplifier differs from that of other vacuum tubes. It makes use of a velocity-modulation principle that causes a stream of electrons, which initially has a uniform current density, to form in bunches. It is the periodic bunch impact that excites the output resonator, from which energy is extracted. This use of a beam passing through gaps in closed cavity resonators built into the tube made it possible for the klystron to overcome the transit-time limitations that the conventional negative-grid tubes encounter at high frequencies.

A cutaway drawing of an early type of klystron is shown in Fig. 2.9. The beam of electrons used in the tube is generated in a cathode at one

end of the tube. The electrons liberated from this cathode are accelerated toward the main body of the tube, where they pass through a tube and then through a set of grids in a cavity resonator. In passing through this first resonator some of the electrons are speeded up and some slowed down by an alternating axial electric field. This action, called "velocity modulation," causes the electrons to form in bunches by the time they pass through the grids of the second resonator, and it is the bunch

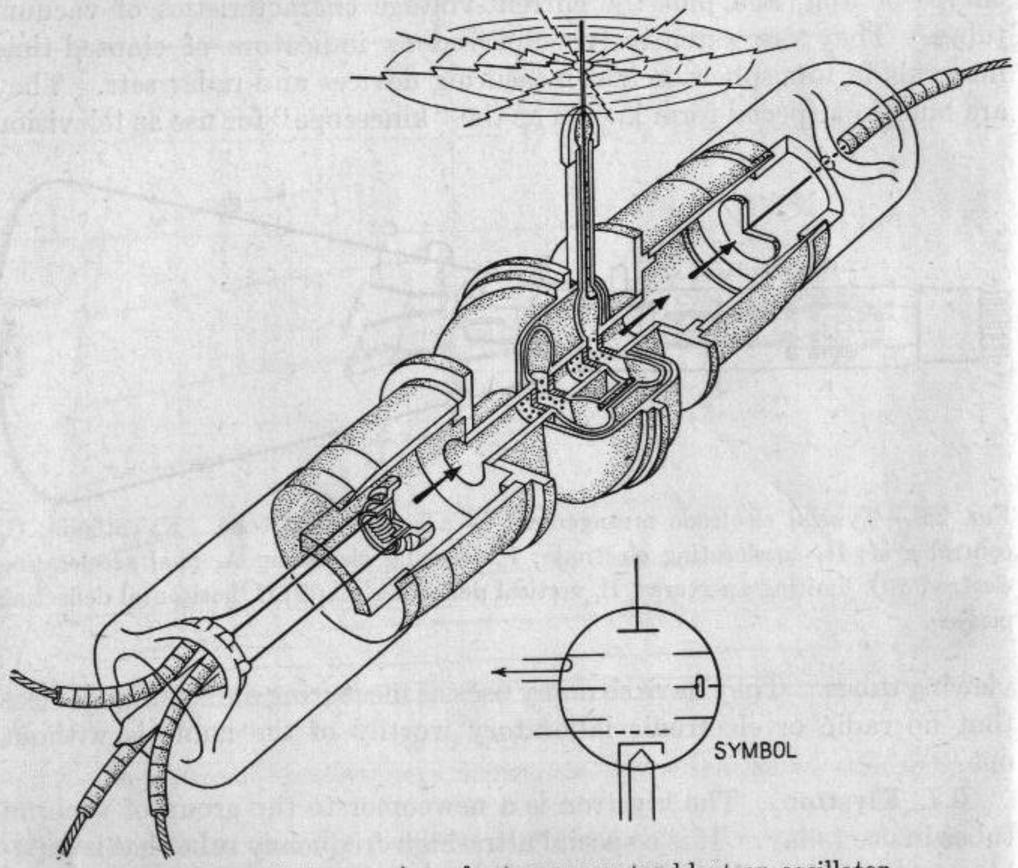


FIG. 2.9.—Cutaway view of a two-resonator klystron oscillator.

impact here that converts the kinetic energy of the electrons into high-frequency electromagnetic energy of the second, or catching, resonator.

A klystron tube may be used as an oscillator by feeding part of the output from the output resonator back to the input resonator. The tube will oscillate when the total phase shift around the circuit composed of the input resonator, the electron beam, the output resonator, and the coupling line back to the input resonator is some integral multiple of 360 deg. Because of this phase requirement it is found that the oscillating action is voltage selective; *i.e.*, the tube will oscillate at certain voltages but not at others since the phase-angle equivalent of the transit time of

the electrons along the beam is involved. In Fig. 2.10 is shown a curve of klystron output versus beam voltage. This shows how the tube oscillates at certain select bands of voltage. The maximum power output that can be obtained from a klystron is nearly inversely proportional to the frequency for which the tube is designed, being about 200 watts at 40 cm.

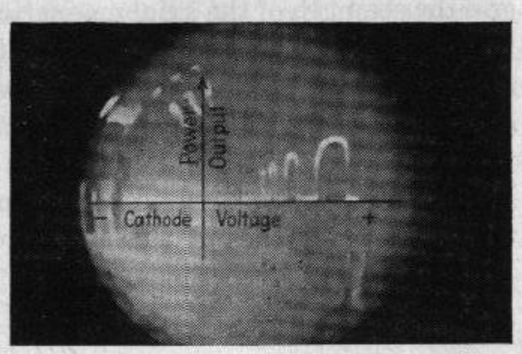


FIG. 2.10.—Power output—beam voltage characteristics of a two-resonator klystron oscillator. This is a picture of an oscilloscope trace, which shows that oscillations are selective with beam voltage.

2.8. Magnetron. The magnetron is a vacuum tube whose current may be influenced by a magnetic field. In certain special forms it is useful as an ultra-high-frequency oscillator. As such it may oscillate at wave lengths from 100 to 1 cm. It is capable of a continuous power output of several hundred watts and instantaneous powers of several thousand kilowatts.

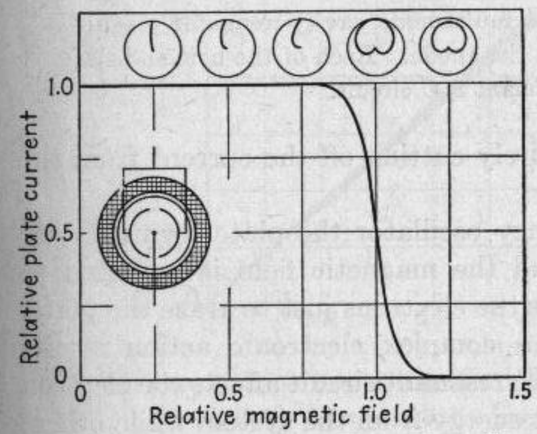


FIG. 2.11.—Cutoff characteristic of a split-anode magnetron. The curve shows that as the axial magnetic field is increased the plate current is at first constant and then suddenly drops rapidly to zero. This results from the electrons becoming progressively more curved in their paths until they finally are unable to reach the plate.

Early forms of the tube were of the split-anode type. The important parts of this type of magnetron are the cathode, frequently in the form of a straight wire filament, and the anode, in the form of a circular plate concentric with the cathode and split into an even number of similar segments. The segments of the plate are operated at the same positive d-c potential relative to the cathode, and a magnetic field is applied parallel to the tube axis. This combination of electric and magnetic fields causes the electrons emitted from the cathode to move in nearly circular paths in the region between cathode and anode.

The radii of the nearly circular electron paths in a magnetron depend upon the strength of the radial electric field and the axial magnetic field. The radii of the paths decrease as the electric field is

decreased or as the magnetic field is increased. Thus if a magnetron has its circular plate segments maintained at a constant d-c potential and the strength of the axial magnetic field is increased from zero strength to a large value, the electrons in the tube will at first move radially from the cathode to the plate and then move in paths which are more and more strongly curved until finally the magnetic-field strength is reached at which the electrons miss the plate entirely. This action is shown in Fig. 2.11, in which there is given a plate-current–magnetic-field characteristic and sketches of the associated electron paths. It is seen that

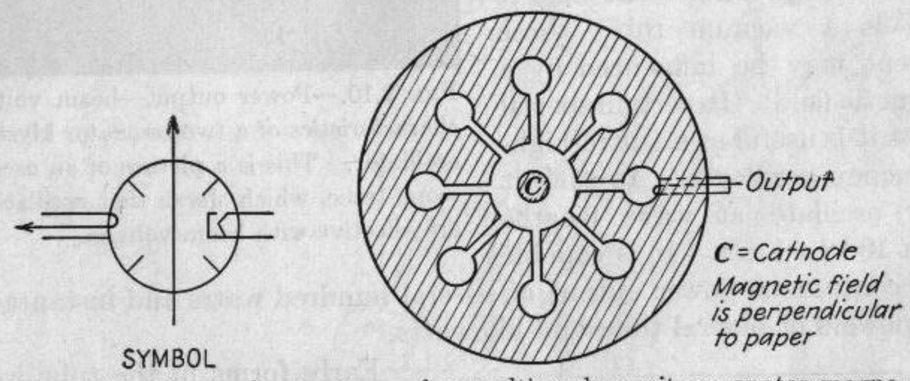


FIG. 2.12.—Electrode structure of a multianode cavity-resonator magnetron. The outer electrode serves as the anode. Each of the hole-and-slot combinations acts like a parallel resonant *L-C* circuit.

the magnetic field is capable of entirely cutting off the current from the plate.

For operation as a high-frequency oscillator the plate segments are made part of resonant circuits, and the magnetic field is adjusted to approximately the value that causes the electrons just to graze the plate. If any small disturbance occurs, a complex electronic action results in which the damped oscillation of the resonant circuit affects the electron paths so that some electrons extract energy from the system while others give up part of their kinetic energy to the oscillating system. The tube can be adjusted so that energy is extracted from the majority of the electrons as they graze the plates, and thus powerful oscillations are maintained.

Modern super-high-frequency magnetrons are made in the form of a multianode cavity. The basic structure of such magnetrons is shown in Fig. 2.12. The cathode is in the form of a cylinder of appreciable diameter located in the center of the structure. The anodes are cut out of one piece of metal and have the form of a large circular hole in a block with radial slots leading out to smaller circular holes. Electrically, each slot and terminating hole are equivalent to a tuned resonant circuit,

the slot having a predominantly capacitive action and the terminating hole having a predominantly inductive action. One of the resonant conditions possible in this equivalent circuit is one in which alternate segments of the anode exhibit the same electrical polarity and thus give the same action as a split multisegment anode, with the advantage that the fields associated with this resonance are confined. Under proper conditions of voltage and magnetic-field strength parallel to the long axis of the cathode, energy will be transferred from the swarm of gyrating

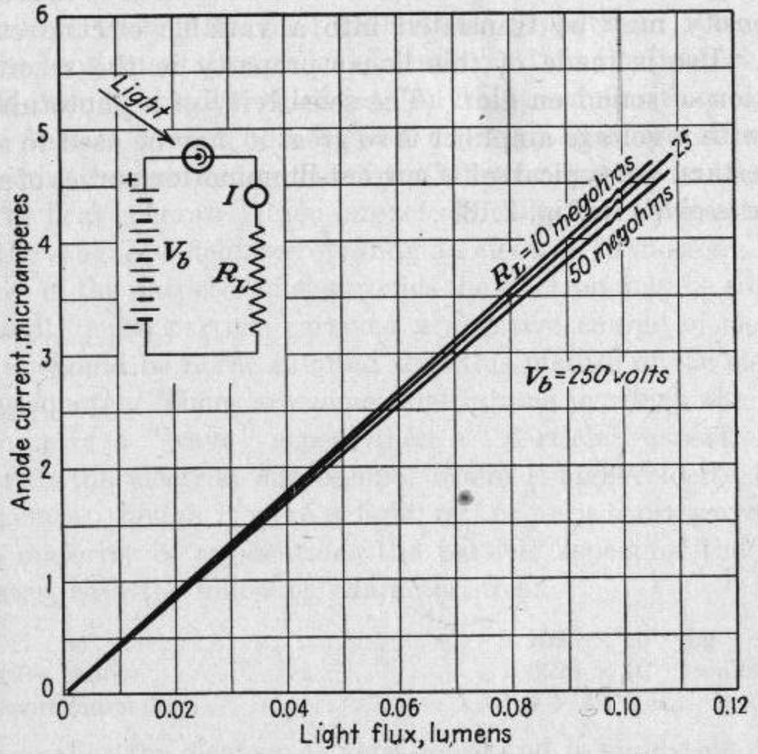


FIG. 2.13.—Current-illumination curves of a typical vacuum phototube. The current is linear with the illumination. Although actual currents are quite small, the voltage developed across the large series resistors used is ample for operating vacuum-tube devices.

electrons around the anode to the resonant circuit and powerful oscillations will be sustained.

2.9. Phototubes. The phototube is a vacuum tube that permits current to pass through it when light falls upon one of its electrodes. The tubes are generally small and contain an electrode in the form of a half cylinder coated with some photosensitive material such as caesium oxide. Various other light-sensitive materials enable the phototube to respond to light of different colors or even to invisible radiations.

Phototubes are extensively used for counting and sorting devices. They may be used to operate doors and drinking fountains, to turn on lights, and to provide safety devices for machine operators. They may be employed anywhere where the interruption or detection of a beam of light is to be correlated with some operation.

Phototubes are able to operate by virtue of an effect known as "photoemission." Certain materials exhibit the property of emitting electrons when exposed to light. The number of electrons emitted is directly proportional to the intensity of the illumination so that a variable light intensity may be translated into a variable electric current or potential. Use is made of this linear property in the recording and reproduction of sound on film. The sensitivity of a phototube in conjunction with a voltage amplifier is so great it may be used to study the light from stars. A typical set of current-illumination curves of a vacuum phototube is given in Fig. 2.13.

CHAPTER 3

ELECTRONS AND IONS

3.1. The Electron. It is the electron that makes vacuum tubes possible and endows them with their remarkable properties. The electron is one of the fundamental particles of matter. It is the lightest particle known. It cannot be subdivided into anything smaller than itself. It is so small that it cannot be observed directly; all observations of its properties must be made in terms of the effects associated with it, such effects as the heat generated upon impact of an electron with a stationary object or the magnetic field surrounding an electron in motion.

For most of the purposes of electronics the electron may be considered to be a small, dense particle carrying a negative charge of electricity. However, it should be borne in mind that this picture of the electron is far from adequate. There are some applications in which the electron displays more of a "wave" aspect than a "particle" aspect. This is the case with the electron microscope, where a high-velocity beam of electrons acts as though it were a light ray of very short wave length.

In the majority of applications the particle aspect of the electron predominates, with the following characteristics:

Mass	9.1066×10^{-31} kg
Negative charge	1.6020×10^{-19} coulomb
Apparent radius	1.9×10^{-13} cm

It is seen that the electron is very dense and is highly charged. It has an apparent density of 0.50×10^{11} g per cm^3, which is millions of times greater than that of our heaviest metals (the density of iron is 7.86 g per cm^3). Further, if the classic concepts of electrostatics be applied to the electron, it may be thought of as being charged to a potential of about 750 kv.

Electrons are a basic constituent of all matter, being the planetary unit of all atoms. No matter can exist without electrons, but electrons may exist by themselves. It is the free electrons that are responsible for most electrical phenomena. They are the units that carry the current in vacuum tubes. They constitute currents in conductors when in motion. Their motion in special conductors such as antennas gives rise to electromagnetic radiations. They constitute cathode rays and beta rays and are emitted from hot bodies.

3.2. The Proton. The proton is the companion piece to the electron. It is the fundamental particle carrying a positive charge of electricity. It, too, is a constituent of all matter, existing as it does in the nucleus of all atoms. The vital statistics of the proton are that its charge is the same in magnitude as that of the electron but with a positive sign, that its mass is 1,845 times that of the electron, and that its apparent diameter is a little less than 10^{-13} cm. The proton is not nearly so much in evidence as is the electron in vacuum tubes. It rarely exists as an isolated particle. Because of its great mass it has a smaller effect than does the electron in determining the characteristics of materials and in constituting a current flow.

3.3. Other Fundamental Particles. Until 1932 the electron and the proton were the only fundamental particles known. Then there were found a number of other fundamental particles whose rarity and short life had hitherto precluded their discovery.

Among these new particles is the *neutron*, which is basically a proton with no charge. There is also a *positron*, which is an electron with a positive charge. There is some evidence of a neutrino, which is a particle of small mass and with no charge. Strangest of all is the *mesotron*, often abbreviated as "meson," which is a particle with about one-tenth the mass of the proton and carrying either a positive or a negative charge. These particles, however, are of no concern to the electronic engineer since they seldom make their appearance in ordinary vacuum tubes.

Another "particle" that has been known for some time is the *photon*. The photon, though classed as a particle, exhibits a wave nature most of the time and is the one particle whose dual nature is most evident. It is a packet of electromagnetic energy whose apparent mass is directly proportional to the frequency of its wave aspect. It carries no charge.

3.4. Atoms and Molecules. Electrically neutral combinations of electrons and protons constitute atoms according to the atomic theory of Rutherford, Bohr, and subsequent workers. The word "atom" is derived from the Greek word meaning "indivisible." Atoms are indivisible in the sense that they are the smallest bits of matter which maintain the properties of the several elements of materials of which they are part. There are 92 types of atoms, corresponding to 92 materials known as "elements." Combinations of the different atoms form molecules, which are the smallest constituent parts of all other materials composing the physical world.

The basic structure of the atom is believed to be a kind of planetary system consisting of a nucleus, which is a group of neutrons and protons, and having a group of planetary electrons equal to the number of protons in the nucleus.

The simplest atom is that of hydrogen. It has 1 nuclear proton and 1 planetary electron. To get an idea of the size of such an atom, if the proton were 1 cm in radius, the electron would normally be spaced a distance of 5 km. The helium atom is the next simplest atom. Its nucleus consists of 2 protons and 2 neutrons. It has 2 planetary electrons. Other atoms are relatively more complicated. The oxygen atom has 8 nuclear protons and 8 nuclear neutrons, whose charge is balanced by that of 8 planetary electrons.

The weight of an atom is determined almost entirely by the sum of the number of protons and neutrons in its nucleus. The physical properties of the atom are determined by the number and arrangement of its planetary electrons. The number of the planetary electrons of an atom of an element is known as the "atomic number" of that element. The order of the elements when listed according to their atomic number is very nearly but not exactly the same as the order according to the atomic weights. If the elements are arranged in a periodic table according to their atomic weights and chemical affinity (valence), as was done by Mendelyeev, it is found that elements with similar characteristics are grouped in columns of equal valence (see Appendix I for a periodic table of the elements).

The planetary electrons of an atom were shown by Bohr to lie in restricted orbits. They were further found to lie in shells about the nucleus, each shell having a maximum capacity for electrons. The maximum capacity of the successive shells from the nucleus out is 2, 8, 18, 32, 18, 18, 2. Thus the atom of neon, whose atomic number is 10 and whose atomic weight is 20.183, has 10 planetary electrons arranged with 2 electrons in the first shell and 8 in the second. These 10 electrons balance the electrical charge of the nucleus, which consists of 10 protons and 12 neutrons.

The number of electrons in the outermost shell of an atom determines its valence and is the principal factor in determining the physical properties of the atom. Atoms with an outer shell filled to its capacity are relatively inactive, while atoms with only 1 electron in their outer shell are most active.

The atomic weights of the elements are taken as relative to that of oxygen, which is chosen to be 16. The fact that the atomic weights are not integers is due in most cases to the fact that there exist atoms of the same element with different numbers of neutrons in the nucleus. The atomic weight of a sample of an element is then determined by the relative number of these different atoms. Atoms with the same number of planetary electrons but with different numbers of nuclear neutrons are known as "isotopes" of the same element. Neon has isotopes with 20,

21, and 22 nuclear particles mixed in such a way that the atomic weight is 20.183. Hydrogen has isotopes with 0, 1, and 2 nuclear neutrons. The first isotope is the common one. The others are the relatively rare "heavy hydrogen" isotopes.

Combinations of the atoms of the elements form molecules. The molecule is the smallest particle of a compound which can exist without losing the characteristics of that compound. Molecules range in size from those of atomic size to a few large enough to be seen with an electron microscope. The molecules of some elements are not just single atoms but groups of two identical atoms.

3.5. Ions. An ion is a molecule or atom with a charge of electricity acquired by the loss or gain of one or more electrons. Electrons in the outer shell of an atom are rather loosely bound to the atom and so may be dislodged by impact of a particle or by exposure to X rays. Ionization of an atom of an element does not change it from one element to another. This is because the nucleus of the atom is unchanged and the form of the nucleus determines the arrangement of the electrons in neutral form.

Ions are important in vacuum tubes because they constitute a current when in motion and thus affect the characteristics of tubes, if they exist in sufficient number. Since even the most completely evacuated tubes contain billions of molecules per cubic centimeter, ions are always created by the impact of electrons and depending upon the type of tube may be a large factor in determining the tube characteristics.

Ions are of most importance in certain special tubes that contain considerable amounts of a definite gaseous element deliberately introduced in great quantities and are an important factor in the tube operation.

CHAPTER 4

ELECTRONIC EMISSION

Every vacuum tube depends for its action upon a stream of electrons that acts as a carrier of current. As necessary as the stream of electrons is the electrode that emits them. Whatever the nature of the tube and the arrangement of electrodes, an emitting electrode cannot be dispensed with. Even in cold-cathode tubes, one of the electrodes is treated with a low-work-function material to facilitate the production of some electrons that will initiate the action.

In general, the excellence of performance of a given tube depends upon the efficiency with which free electrons are produced. When the emission fails, the tube is useless. We infer correctly then, that the subject of electron emission is worthy of considerable study.

The types of electronic emission may be listed as follows:

1. Thermionic, or primary, emission.
2. Secondary emission.
3. Photoelectric emission.
4. Field emission.

The common feature of all types of emission is that energy is imparted to the free electrons in a solid in an amount sufficient to enable them to overcome the restraining forces at its surface and thus escape from the solid.

The types of emission differ only in the way in which the escape energy is imparted to the free electrons. *Thermionic emission* occurs when a material is heated to incandescence in a vacuum. In this case the escape energy is imparted by heating the material. *Secondary emission* occurs when a high-velocity electron or ion strikes a material in a vacuum and knocks out one or more electrons. In this case the energy that enables the free electrons to escape comes from the striking particle. *Photoelectric emission* occurs when energy in the form of light falls upon a surface. *Field emission* occurs at cold surfaces under the influence of extremely strong fields.

All types of emission are most effective in vacuum. If the emission did occur in air, the emitted electrons would not get very far through the relatively dense surrounding atmosphere. Most metals would burn

up in air at the temperatures to which they must be raised to emit satisfactorily. Only primary and secondary emission will be discussed in this chapter. Photoelectric emission will be discussed in a separate chapter. Field emission is not yet of much practical importance.

4.1. Theory of Thermionic Emission. Every metal has a crystalline structure of its atoms, *i.e.*, the atoms have an orderly arrangement in some sort of lattice pattern. The atoms in this lattice structure have certain of their outer electrons loosely bound. These loosely bound

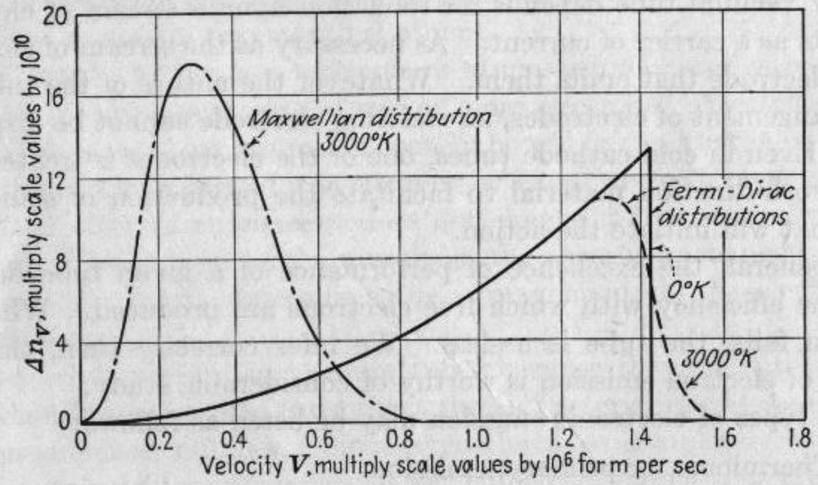


FIG. 4.1.—Maxwellian and Fermi-Dirac distribution of velocities in 1 cu mm of solid tungsten. The abscissa gives the relative number of electrons in a velocity increment of 10^{-3} meters per sec in units of 10^{10} electrons.

electrons may move from atom to atom in a relatively unrestricted fashion. Such electrons are known as the "free electrons" in the metal in that they are not bound to any one atom. The free electrons in a metal act much like the molecules in a gas. An increase in temperature increases their activity and average velocity. A potential gradient in the metal causes them to move progressively in one direction, giving rise to a conduction-current flow.

Because of the atomic restraints it is not expected that the velocity distribution in a metal is Maxwellian, as is almost exactly the case for gases. The true distribution was found by Fermi and Dirac from quantum-mechanical statistical considerations. For comparison there are shown in Fig. 4.1 the Maxwellian and Fermi-Dirac distribution of velocities. The distinctive feature of the Fermi-Dirac distribution of velocities is that at zero temperature only a small fraction of the electrons have zero velocity. As temperature increases, the velocity and corresponding energy distribution change so that more electrons have higher velocities.

The high-velocity electrons that escape from the metal constitute the emitted current.

The Maxwellian distribution of velocities referred to above and shown in Fig. 4.1 is one given by the equation

$$y = \frac{4x^2}{\sqrt{\pi}} \epsilon^{-x^2} \tag{4.1}$$

This is the general form of the probability y that a particle will have a velocity x times the most probable velocity. It applies perfectly for most gases but does not give the true picture for electrons in metals. For large velocities, however, the Maxwellian and Fermi-Dirac distributions differ only by a constant. Thus the electrons emitted from an incandescent surface do have a Maxwellian distribution, but the energies of the electrons are (at 3000°K) about 1,000 times those predicted from the simple Maxwellian theory. Upon converting Eq. (4.1) to a form involving energy instead of velocity and taking the derivative properly, the fraction $\frac{n}{n_0}$ of the emitted electrons that can move against a retarding field of V volts is given by

$$\frac{n}{n_0} = \epsilon^{-\frac{Ve}{kT}} \tag{4.2}$$

where $-e$ is charge of the electron, 1.602×10^{-19} coulomb
k is Boltzmann's constant, 1.380×10^{-23} watt-sec per °K
T is absolute temperature, 273+ °C

A nomographic chart of Eq. (4.2) is given in Fig. 4.2. From this it is seen that about 50 per cent of the electrons emitted from a cathode at 1500°K, typical oxide operating temperature, have velocities greater than 0.09 volt.

Work Function. The surface restraints that prevent the majority of the free electrons in a metal from leaving it are the electrostatic forces produced by the charges in the atoms. These come not only from the residual positive charges but also from a rearrangement of the negative charges. A free electron must have a certain minimum kinetic energy before it can tear itself free from these forces. The work per unit charge required to free an electron from the influence of the charges in the metal and thus to escape from it is known as the *work function* of the metal. The work function is usually expressed in volts.

The electrostatic forces within a metal are rather complex and not completely understood. Indications are that the forces are small within the metal, reach a maximum several atomic diameters outside the metal, and then decrease according to an inverse-square law at greater dis-

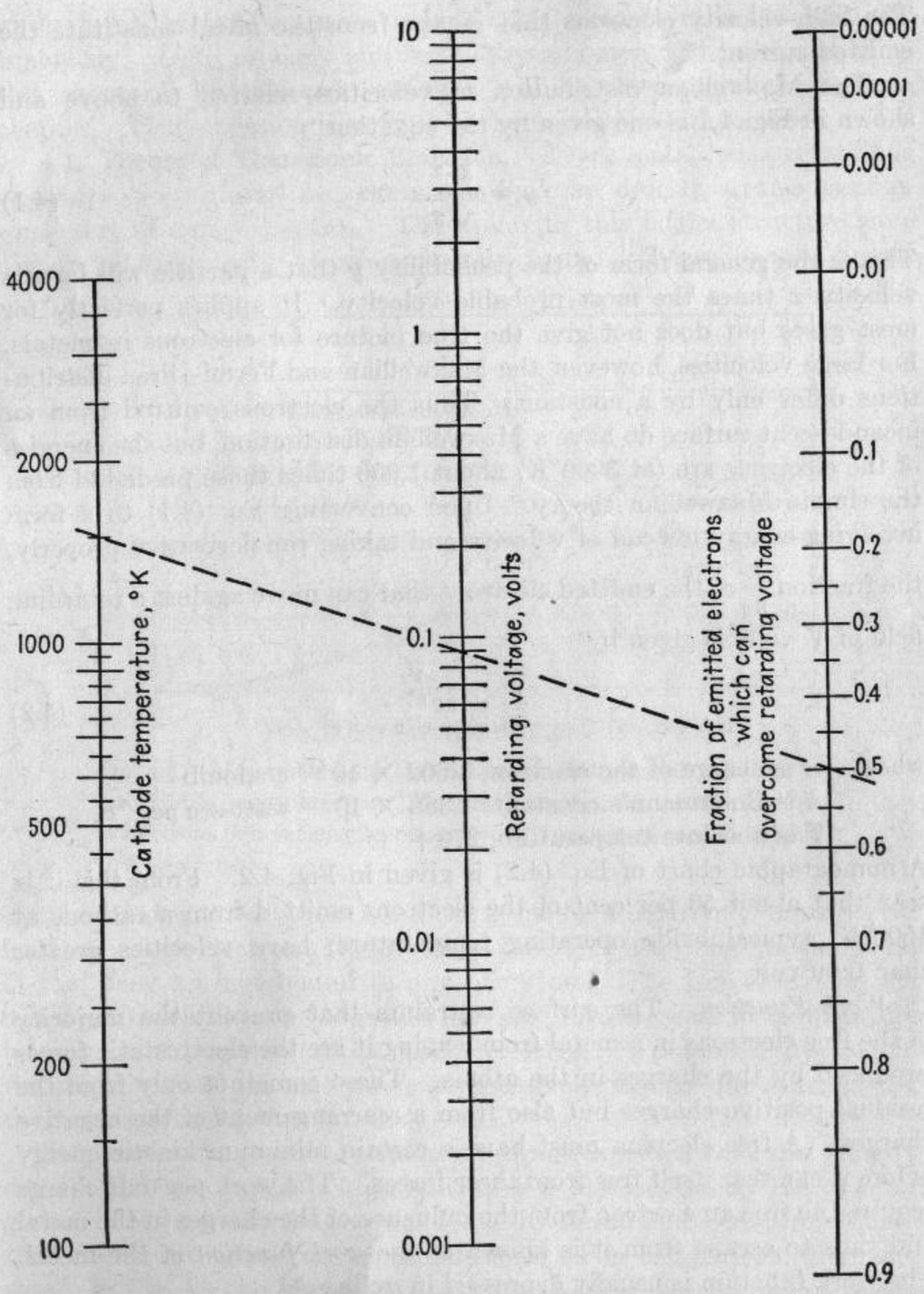


FIG. 4.2.—The velocity distribution of electrons resulting from thermionic emission as given by Eq. (4.2). The nomographic chart gives the fraction of the emitted electrons associated with a given cathode temperature that can overcome a retarding potential of a given number of volts.

tances, where an image action manifests itself. It would be expected that the work function would decrease as the distance between the atoms in the crystalline structure increased. This turns out to be the case, and experimentally, a curve of the work function of the alkali metals of the first column of the periodic table plotted against their atomic spacing is a smooth one, nearly inversely proportional to the square root of the atomic spacing, as may be seen in Fig. 4.3. Conclusions for other metals can hardly be drawn, for there are so few having the

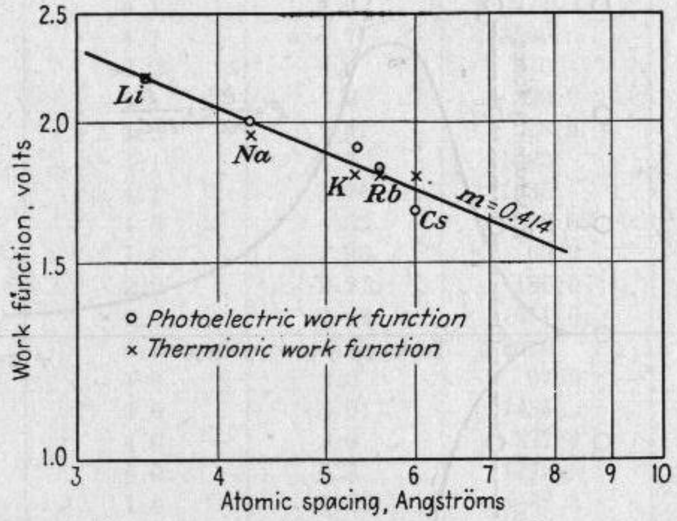


FIG. 4.3.—Work function of the alkali metals as a function of atomic spacing. The curve shows that for a given crystal structure, the further the atoms are apart the lower is the work function.

same valence and crystalline structure. Since the atomic spacing is a periodic function of the atomic number, the work function is also a periodic function of the atomic number.

No completely successful theoretical determination of the work function has apparently as yet been made. The general nature of the restraining forces is probably very much like that shown in Fig. 4.4. Within the metal the force has an average value of zero. Near the surface there are the attractive forces of atoms that have lost an electron by emission and forces due to rearrangement of residual charges. The forces are undoubtedly greatest near the surface, where the force-producing charges are closest and yet not symmetrically disposed with respect to the surface. Well outside the surface the force is probably one that varies with the inverse square of the distance from the metal, for in this region the charges in the metal arrange themselves so as to give the effect of an image charge of the electron escaping from the metal. The force

cannot be inverse-square law all the way out from the surface, for then an infinite energy would be needed for escape. It may be concluded that the work function depends in some complex way upon the atomic spacing, crystal structure, and valence of the metal.

The work function of materials is most accurately determined experimentally from observations of the photoelectric emission of the material,

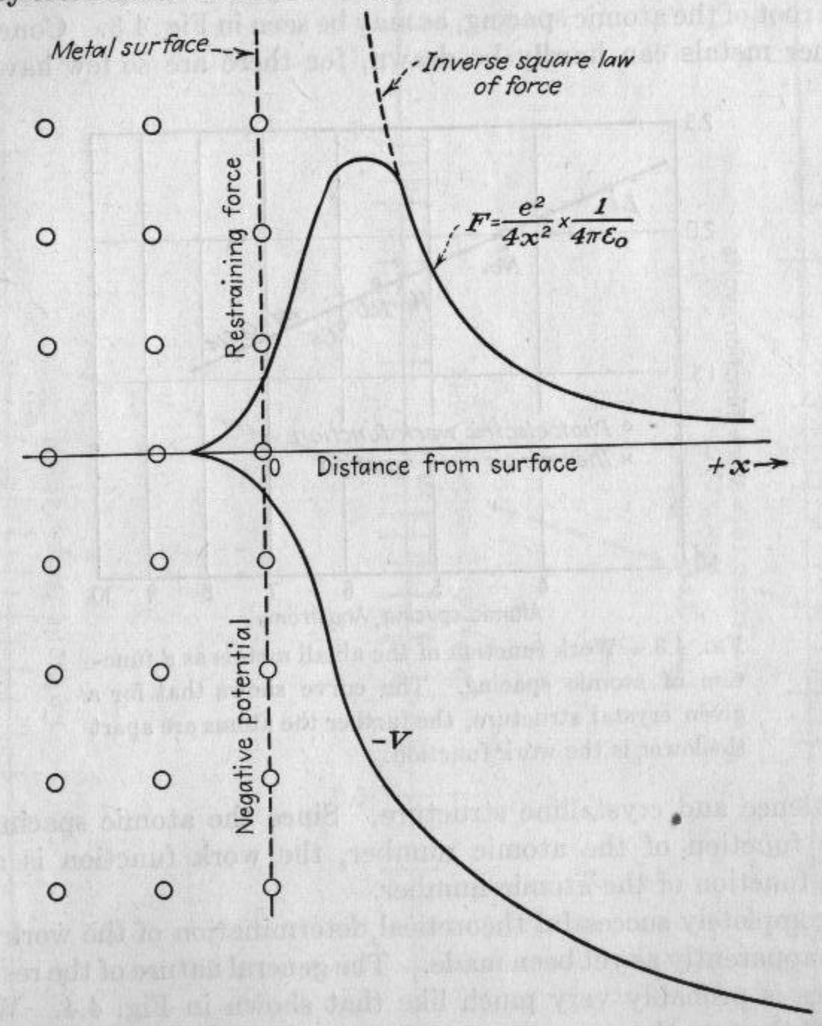


FIG. 4.4.—Restraining force on an electron near a metal surface. At large distances from the metal the force is that due to an image charge located in the metal.

but it may also be deduced from the thermionic-emission characteristics. A list of the work functions of the metal emitters most often used is given in Table I.[1,2]

[1] HUGHES, A. L., and L. A. DUBRIDGE, "Photoelectric Emission Phenomena," McGraw-Hill, New York, 1932.

[2] BECKER, J. A., Thermionic Emission and Adsorption, *Rev. Modern Phys.*, vol. 7, pp. 95–128, April, 1935.

TABLE I
EMISSION CONSTANTS OF THE METALS

	Probable ϕ_T,* volts	Average ϕ_P,† volts	Melting temp., °C	Lattice const, angstrom units
Ag	4.7	4.6	960.5	4.08
Al	3.0	3.0	659.7	4.04
Au	4.8	4.78	1063	4.07
Ba	2.52		850	5.015
Bi	4.1	4.2	271.3	4.75
C	4.7	4.77	>3500	2.455
Ca	3.2	3.0	810	5.56
Cd	4.1	4.0	320.9	2.97
Cs	1.8	1.67	28.5	6.05
Cu	4.1	4.3	1083	3.61
Fe	4.7	4.74	1535	2.90
Hg	4.5	4.53	−38.87	
K	1.8	1.90	62.3	5.33
Li	2.2	2.21	186.0	3.46
Mg	2.4	2.43	651.0	3.20
Mo	4.3	4.15	2620	3.14
Na	1.9	2.0	97.5	4.24
Ni	5.0	5.01	1455	2.66
Pb	4.0	3.9	327.4	4.94
Pt	6.0	6.3	1773.5	3.91
Rb	1.8	1.82	38.5	5.62
Sr	2.1	2.06	800	6.05
Ta	4.1	4.13	3269	3.28
Th	3.4	3.50	1845	5.07
W	4.52	4.61	3370	3.16
Zn	3.3	3.44	419.47	2.66
Zr	4.1	3.73	1900	3.22

* Work function as determined by thermionic measurements.
† Work function as determined by photoelectric measurements.

The Emission Equation. In view of the foregoing discussion it would be expected that the emission from a metal would depend upon its temperature and upon the work function. Richardson[1] and Dushman[2] have

[1] RICHARDSON, O. W., The Distribution of the Molecules of a Gas in a Field of Force, *Phil. Mag.*, vol. 28 (No. 5), pp. 633–647, 1914.

[2] DUSHMAN, S., Electron Emission from Metals as a Function of Temperature, *Phys. Rev.*, vol. 21 (No. 6), pp. 623–636, 1923. See also the summarizing source article, S. Dushman, Thermionic Emission, *Rev. Modern Phys.*, vol. 2, pp. 381–476, October, 1930, which gives a comprehensive survey of the subject as developed to that date. See also the book, A. L. Riemann, "Thermionic Emission," Wiley, New York, 1934.

shown this to be the case and have shown specifically that the thermionic emission from a metal is given by

$$J = AT^2\epsilon^{-\frac{b_0}{T}} \tag{4.3}$$

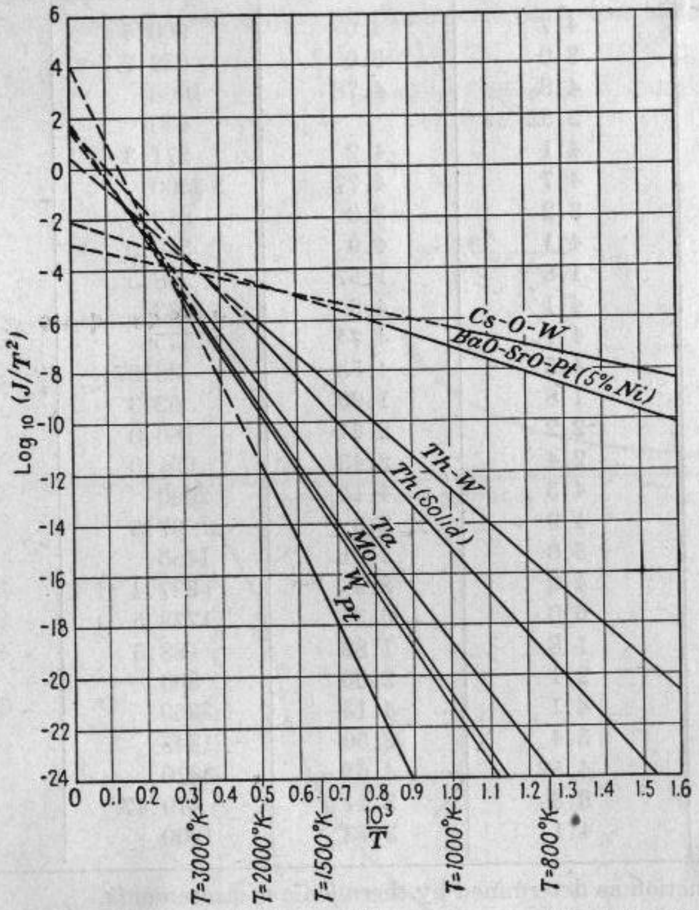


FIG. 4.5.—Characteristics of the common emitters shown as a curve of $\log J/T^2$ against $1/T$. This type of plot demonstrates the validity of the Richardson-Dushman equation (4.3). The y-axis intercepts give the emission constant A. The slope of the lines is proportional to the work function of the emitter.

where J is current density, amperes per cm^2

A is 120.4 amperes per cm^2 per deg^2, a universal theoretical constant

T is absolute temperature, °K (273 + °C)

b_0 is temperature equivalent of the work function, $11{,}600\phi_0$, °K

ϕ_0 is work function of the metal, volts

Equation (4.3) may be derived from either thermodynamic or quantum-mechanical considerations. The resulting equation is the same in either case.

From the form of the emission equation (4.3) it is seen that if the logarithm of $\frac{J}{T^2}$ be plotted against the reciprocal of T there will result a straight line whose slope is $-b_0$ and whose y-axis intercept is $\ln_\epsilon A$. The correctness of the emission equation has been verified by so plotting experimentally determined results. It is found in all cases that the results produce a straight line. A group of such curves for common emitters is given in Fig. 4.5. In this figure those lines with the lowest slope correspond to metals with the lowest work function. Theoretically, the intercept should be 2.08, corresponding to the $\log_{10} 120.4$. Actually, it is about 1.78, corresponding to a value of A of 60 instead of 120.4 for most of the pure metals. Values of A are found higher as well as lower than the theoretical values so that the theory is not discredited by this discrepancy. There is some evidence that the work function is not entirely independent of temperature as has been assumed in the derivation of the emission equation. The differences in the value of the work function as determined by thermionic and photoelectric methods may possibly be due to temperature. A decrease in the work function of 6 parts per 100,000 per degree would cause the observed discrepancy in the constant A.

The exponential term in the emission equation accounts for most of the variation of emission with temperature. The variation with the T^2 term is so small that the correctness of the exponent 2 can hardly be verified experimentally. In the case of tungsten at 2500°K a 1 per cent change in temperature changes the T^2 term by 2 per cent but changes the exponential term by 20 per cent. This causes the emission-temperature function to be one of the most rapidly varying functions found in nature. Doubling the temperature may increase the emission by a factor of 10^7. Halving the work function will have nearly the same effect as doubling.

The quantities of the curves of Fig. 4.5 are not in very convenient form for ordinary use, and therefore a better method of representing the emission characteristics of materials is sought. It is possible to plot emission current against temperature directly as in Fig. 4.6, but the variation of current with temperature is so rapid that such a curve is not very satisfactory. It would also be possible to plot emission against heating power by making use of the fact that at the high temperatures required for emission most of the power is lost through radiation according to the Stefan-Boltzmann law.

$$P = Ke_RT^4 \tag{4.4}$$

where P is radiated power, watts per cm^2

K is 5.73×10^{-12} watt per cm^2 per deg^4, a universal constant known as the "Stefan-Boltzmann constant"

e_R is radiation efficiency as fractional radiation of a black body or perfect radiator

Such a plot gives curves that are nearly but not quite straight lines because of the two temperature factors in the thermionic-emission equation. It

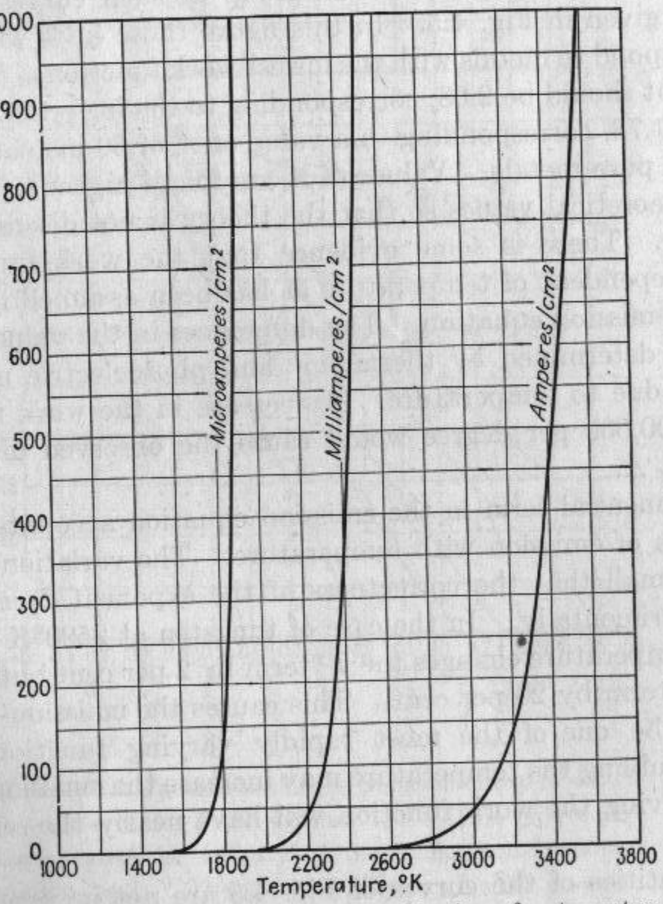


FIG. 4.6.—The emission-current density of a tungsten emitter as a function of temperature.

is possible, however, to warp the lines of the emission scale to take account of the nonuniform temperature variation and get a straight-line plot as shown in Fig. 4.7. The coordinate paper used in Fig. 4.7 is known as "power-emission paper." On it curves of emission against heating power are straight lines to the extent that the radiation efficiency of the emitter remains constant with temperature. Contours of emission efficiency in milliamperes per watt are also readily drawn. Since heat-

radiation efficiency varies rather slowly with temperature, as shown in Fig. 4.8, the emission-power curves can be extrapolated as straight lines with considerable assurance. Radiation efficiency is defined as the per

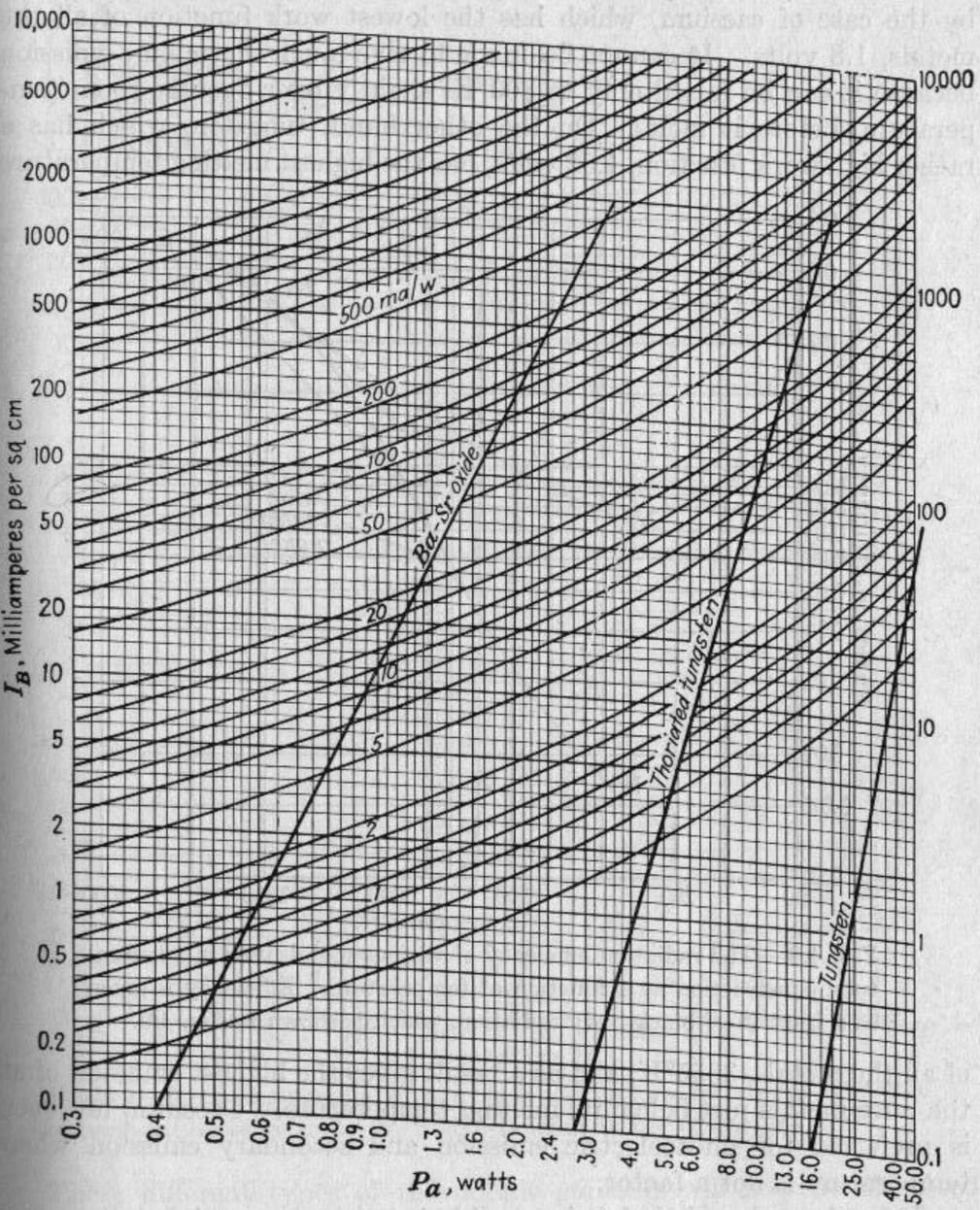


FIG. 4.7.—Emission-current density as a function of heating power for typical examples of various emitter types. The curves show the relative emission efficiencies of the different emitters.

cent of black-body, or perfect, radiation. Black-body radiation as given by Eq. (4.4) is shown in Fig. 4.9. Power-emission paper is manufactured and sold by the Keuffel and Esser Company.

Types of Emitter.—Because of the dependence of emission upon temperature and the work function it is not necessarily true that the metal with the lowest work function is the best emitter. This is shown by the case of caesium, which has the lowest work function of all the metals, 1.8 volts. It cannot be made to give much thermionic emission because it can be raised only to 300°K, slightly over average room temperature, before it melts. On the other hand, tungsten, which has a rather high work function, 4.52 volts, has the highest melting temperature

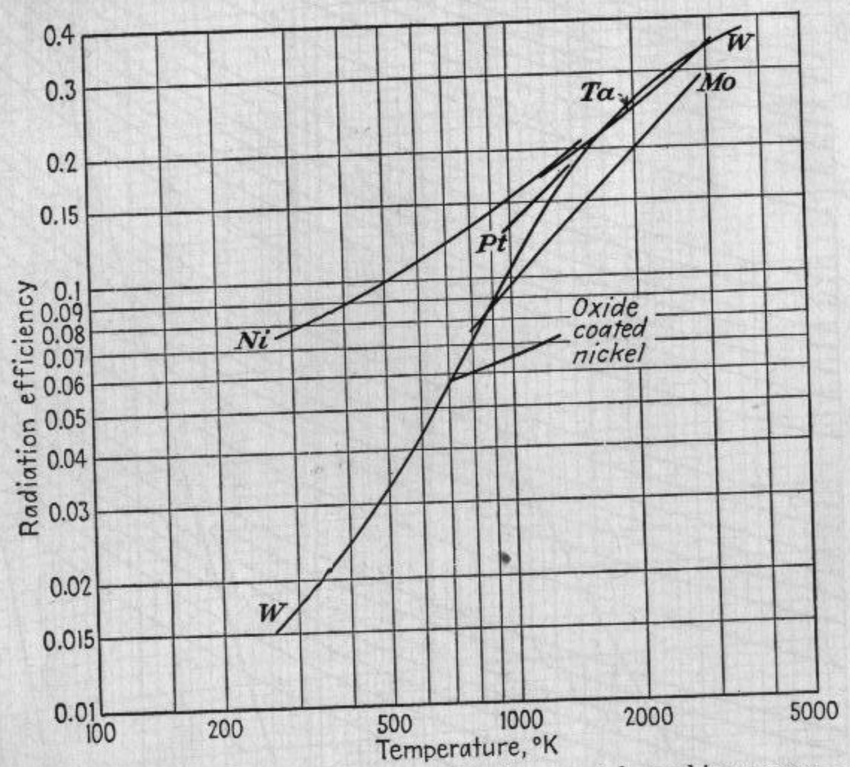


FIG. 4.8.—The radiation efficiency of various metals used in vacuum-tube construction as a function of temperature. Efficiency is given as a fraction of black-body radiation, which is shown in Fig. 4.9.

of all the metals, 3655°K, and as a result gives the highest emission of all the pure metals just below its melting temperature. Caesium, however, is preferred for photoelectric emission and secondary emission where temperature is not a factor.

It has been found that it is possible to raise some metals to temperatures higher than their melting temperatures in the pure state by using them in various chemical and physical combinations. Thus a monatomic layer of thorium on tungsten can be operated at or above the melting temperature of thorium itself. Also, it has been found that small bits of the pure metal can be made to diffuse out of an oxide in the case of

the rare-earth metals so that advantage can be taken of the low work function of these metals, which would otherwise melt at low temperatures.

From the above remarks it is seen that three classes of emitters exist. They are

1. Pure metals.
2. Atomic-film emitters.
3. Oxide emitters.

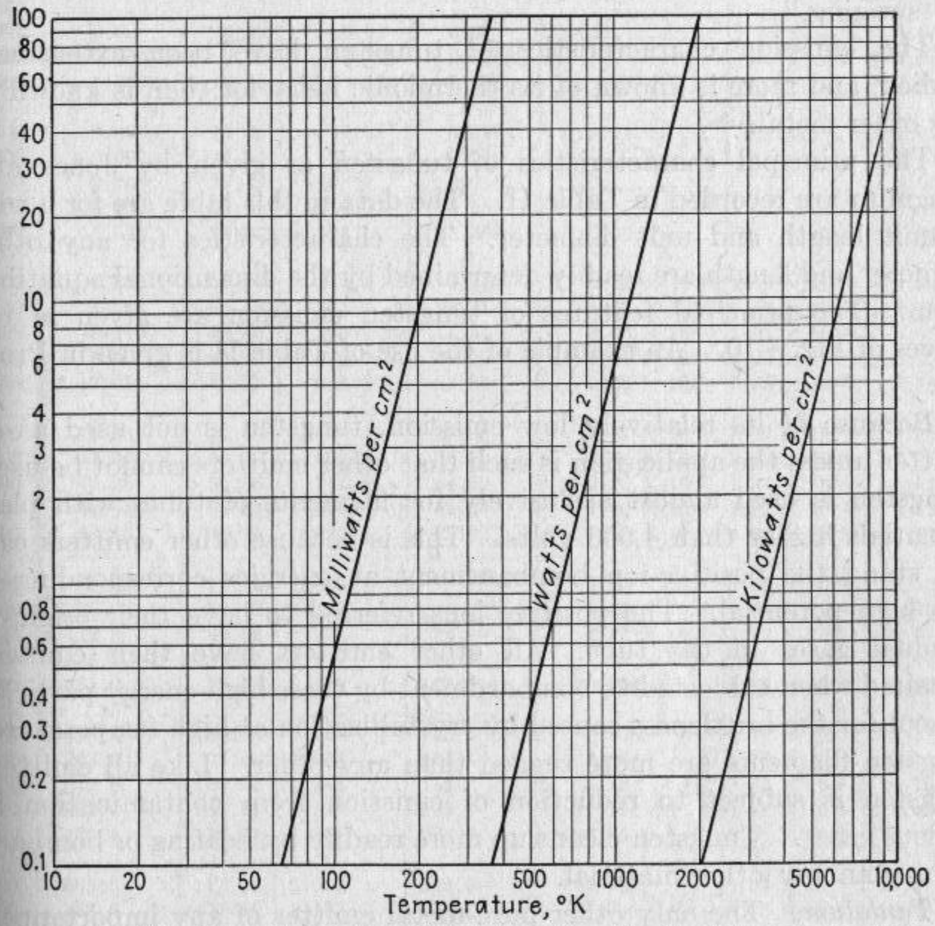


FIG. 4.9.—The thermal radiation of a black body or ideal radiator as a function of temperature as given by the Stefan-Boltzmann law of Eq. (4.4).

These different types of thermionic emitters will now be discussed separately.

4.2. Emission of Pure Metals. *Tungsten.* The pure metals follow the Richardson-Dushman emission equation as closely as can be determined experimentally. In general, the metals with suitable physical characteristics for emission have a relatively high work function and so even at best are not very good emitters. Of all the metals *tungsten*

is the most extensively used because it can be raised to a higher temperature without melting than any other metal. Although tungsten has a desirable high melting temperature, its other physical characteristics are less desirable. It is a hard metal to work because of its crystalline structure. It was not until 1908, when Coolidge discovered that tungsten becomes ductile when extensively worked, that it became practical to use the metal at all. Tungsten cannot be drawn into wire form as can most metals but must be hammered into shape, a process known as "swaging."

The emission characteristics of tungsten have been extensively studied, and more is known of its thermionic behavior than is known of any other metal.[1,2]

The principal characteristics of tungsten as given by Jones and Langmuir are recorded in Table II. The data in this table are for a wire of unit length and unit diameter. The characteristics for any other diameter and length are readily determined by the dimensional equations given. The principal features of tungsten emission are given in the curves of Fig. 4.10. An example of the use of Table II is given in Prob. 4.3.

Because of its relatively low emission, tungsten is not used as an emitter unless the application is such that other emitters cannot be used. Tungsten is used almost exclusively for filaments of tubes with plate potentials higher than 4,000 volts. This is because other emitters cannot stand the positive-ion bombardment at energies corresponding to this high potential. The positive ions referred to have their origin in residual gases in the tube. All other emitters have their emission impaired when subjected to bombardment by these high-energy particles. Except for the brittleness caused by crystallization at high temperatures, tungsten filaments are more rugged than any other. Like all emitters, tungsten is subject to reduction of emission from contamination by various gases. Tungsten cleans up more readily by heating or bombardment than any other material.

Tantalum. The only other pure-metal emitter of any importance is tantalum. Tantalum cannot be heated to as high a temperature as tungsten because its melting temperature is 3300°K. However, the work function of tantalum is relatively low, being 4.1 volts against 4.53 volts for tungsten, so that its emission is at least ten times that from

[1] JONES, H. A., and I. LANGMUIR, The Characteristics of Tungsten Filaments as Functions of Temperature, *Gen. Elec. Rev.*, vol. 30, Part I, pp. 310–319, June; Part II, pp. 354–361, July; Part III, pp. 408–412, August, 1927.

[2] FORSYTHE, W. E., and A. G. WORTHING, The Properties of Tungsten and the Characteristics of Tungsten Lamps, *Astrophys. Jour.*, vol. 61, pp. 146–185.

TABLE II
SPECIFIC CHARACTERISTICS OF IDEAL TUNGSTEN FILAMENTS*
(For a wire 1 cm in length and 1 cm in diameter)

T, °K	W', $\frac{W}{ld}$, watts per cm²	$R' \times 10^6$, $\frac{Rd^2}{l} \times 10^6$, ohm-cm	$I'_f\left(\frac{I_f}{d^{3/2}}\right)$, amps per cm$^{3/2}$	$V' \times 10^3$, $\frac{V\sqrt{d}}{l} \times 10^3$, volts per cm$^{1/2}$	I'_e, $\frac{I_e}{ld}$, amp per cm²	M', $\frac{M}{ld}$, g per cm² per sec, evaporation	$\frac{R'_T}{R'_{293°}}$, $\frac{R_T}{R_{293°}}$
273	.....	6.37	.....	.....	.....	.....	0.911
293	0.0	6.99	0.0	0.0	.....	.....	1.00
300	0.000100	7.20	3.727	0.02683	.....	.....	1.03
400	0.00624	10.26	24.67	0.2530	.....	.....	1.467
500	0.0305	13.45	47.62	0.6404	.....	.....	1.924
600	0.0954	16.85	75.25	1.268	.....	.....	2.41
700	0.240	20.49	108.2	2.218	.....	.....	4.93
800	0.530	24.19	148.0	3.581	.....	.....	3.46
900	1.041	27.94	193.1	5.393	.....	.....	4.00
1,000	1.891	31.74	244.1	7.749	3.36×10^{-15}	1.16×10^{-33}	4.54
1,100	3.223	35.58	301.0	10.71	4.77×10^{-13}	6.81×10^{-30}	5.08
1,200	5.210	39.46	363.4	14.34	3.06×10^{-11}	1.01×10^{-26}	5.65
1,300	8.060	43.40	430.9	18.70	1.01×10^{-9}	4.22×10^{-24}	6.22
1,400	12.01	47.37	503.5	23.85	2.08×10^{-8}	7.88×10^{-22}	6.78
1,500	17.33	51.40	580.6	29.85	2.87×10^{-7}	7.42×10^{-20}	7.36
1,600	24.32	55.46	662.2	36.73	2.91×10^{-6}	3.92×10^{-18}	7.93
1,700	33.28	59.58	747.3	44.52	2.22×10^{-5}	1.31×10^{-16}	8.52
1,800	44.54	63.74	836.0	53.28	1.40×10^{-4}	2.97×10^{-15}	9.12
1,900	58.45	67.94	927.4	63.02	7.15×10^{-4}	4.62×10^{-14}	9.72
2,000	75.37	72.19	1,022	73.75	3.15×10^{-3}	5.51×10^{-13}	10.33
2,100	95.69	76.49	1,119	85.57	1.23×10^{-2}	4.95×10^{-12}	10.93
2,200	119.8	80.83	1,217	98.40	4.17×10^{-2}	3.92×10^{-11}	11.57
2,300	148.2	85.22	1,319	112.4	1.28×10^{-1}	2.45×10^{-10}	12.19
2,400	181.2	89.65	1,422	127.5	0.364	1.37×10^{-9}	12.83
2,500	219.3	94.13	1,526	143.6	0.935	6.36×10^{-9}	13.47
2,600	263.0	98.66	1,632	161.1	2.25	2.76×10^{-8}	14.12
2,700	312.7	103.22	1,741	179.7	5.12	9.95×10^{-7}	14.76
2,800	368.9	107.85	1,849	199.5	11.11	3.51×10^{-7}	15.43
2,900	432.4	112.51	1,961	220.6	22.95	1.08×10^{-6}	16.10
3,000	503.5	117.21	2,072	243.0	44.40	3.04×10^{-6}	16.77
3,100	583.0	121.95	2,187	266.7	83.0	8.35×10^{-6}	17.46
3,200	671.5	126.76	2,301	291.7	150.2	2.09×10^{-5}	18.15
3,300	769.7	131.60	2,418	318.3	265.2	5.02×10^{-5}	18.83
3,400	878.3	136.49	2,537	346.2	446.0	1.12×10^{-4}	19.53
3,500	998.0	141.42	2,657	375.7	732.0	2.38×10^{-4}	20.24
3,600	1,130	146.40	2,777	406.7	1,173	4.86×10^{-4}	20.95
3,655	1,202	149.15	2,838	423.4	1,505	7.15×10^{-4}	21.34

* The values given are taken from H. A. Jones and I. Langmuir. The Characteristics of Tungsten Filaments, *Gen. Elec. Rev.*, vol. 30, pp. 312–313, 1927, Table I. The notation of Jones and Langmuir is retained in this table.

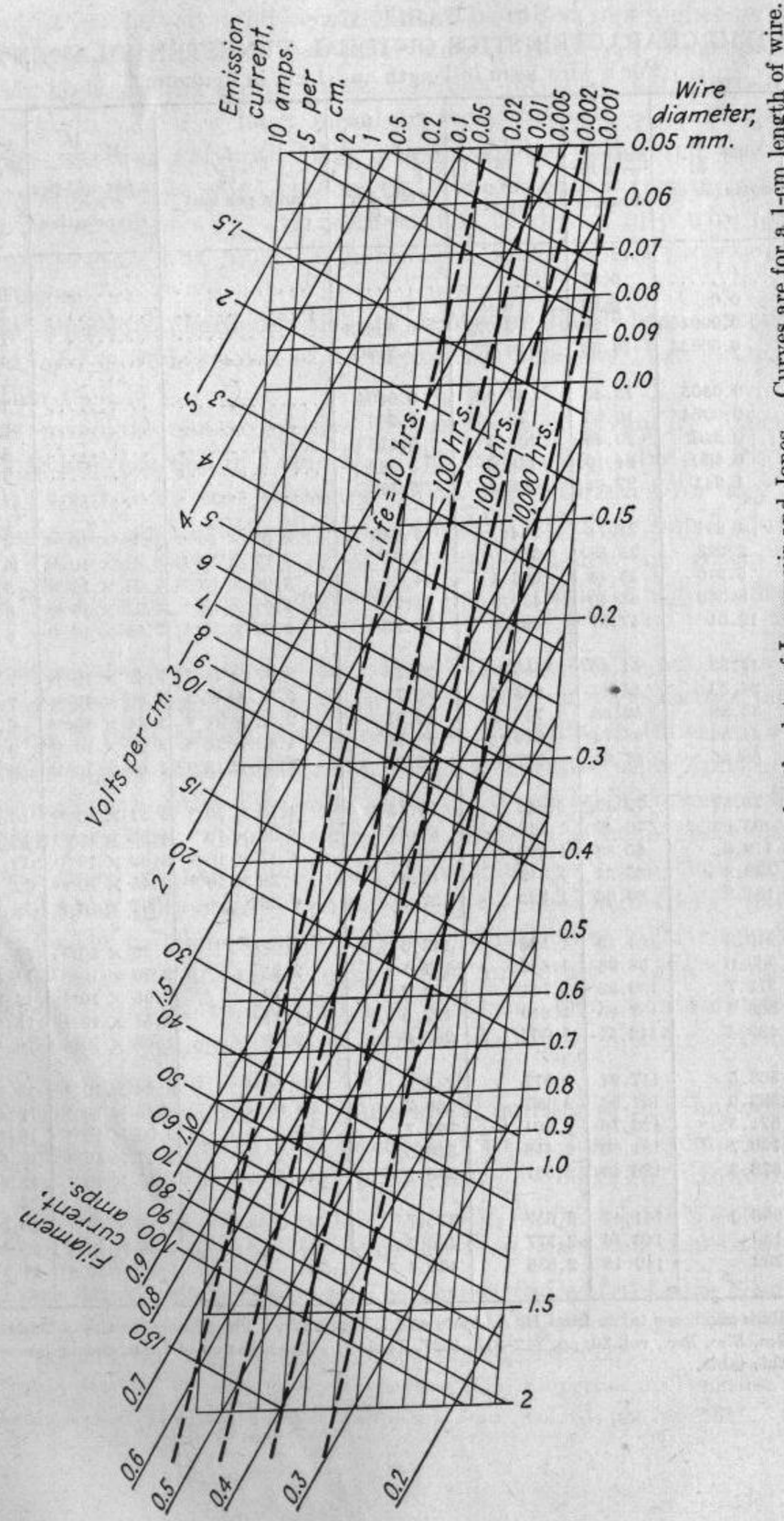


FIG. 4.10.—Characteristics of ideal tungsten filaments as given by the data of Langmuir and Jones. Curves are for a 1-cm length of wire. Life is defined as the time required for a 10 per cent reduction in mass through evaporation.

tungsten at any temperature less than 2500°K. Tantalum has the advantage over tungsten that it can be worked in sheet form to produce specially shaped cathodes, and the like. A disadvantage is that it is easily contaminated by residual gases, which form oxides that greatly reduce the emission.

4.3. Atomic-film Emitters. It is possible to get emission higher than that from pure metals from an atomic film of one metal on another. Of the various combinations that are possible, the most extensively used is that of *thorium on tungsten*. It was discovered by Langmuir and Rogers that the small amount of thorium put into tungsten to reduce the crystallization gave rise to very high emission under certain conditions. What apparently happens is that a certain amount of thorium in the metal diffuses to the surface, where it emits much as thorium would, with the advantage that the thorium can be heated above its own melting temperature and that the work function is reduced by the redistribution of charges in the tungsten and surface layer of thorium.

Thorium was originally added to tungsten to reduce crystallization. As now added to increase the emission, the amount is about 1½ per cent, and this amount is quite critical. If more than this amount is added, the tungsten wire is too hard to work. If less is added, there may not be enough to produce high emission. The thorium is added in the form of thoria (thorium oxide, ThO_2).

A rather intricate schedule of operations is required to produce and activate a film of thorium on tungsten. The process includes the following steps:

1. *Reduction of Thoria to Metallic Thorium.* This is achieved by heating the filament to 2800°K for 1 or 2 min. During this time, most of the thorium oxide is reduced to thorium, and such thorium as reaches the surface evaporates. If the emission is measured at this point, it will be found to be very nearly the emission of pure tungsten.

2. *Diffusion of Metallic Thorium to the Surface.* This takes place as the filament is held at a temperature of 2100°K for a period of 15 to 30 min. During this time the emission increases by a factor of about 1,000. The explanation of this behavior is that metallic thorium diffuses to the surface, where it builds up a monatomic layer of thorium. Studies with the electron microscope[1] show that the thorium arrives at the surface both through pores in the tungsten and at the grain boundaries, from which places it spreads over the surface. At this reduced temperature the evaporation is not very large. In the range of temperatures between 2100 and 2300°K the thorium diffuses to the surface faster than

[1] BRÜCHE, E., and H. MAHL, Ueber das Emissions bild von thorierten Volfram und thoriertem Molybdän, *Zeit. für Tech. Phys.*, vol. 16, pp. 623–627, December, 1935.

it evaporates, so that this is a suitable range for activation. In this range of temperature the percentage of the surface covered varies from 20 to 85 per cent, decreasing as temperature increases as shown in Fig. 4.11. The final layer of thorium that forms is believed to be monatomic.

3. *Operation.* After the above treatment the filament temperature is reduced to 1900°K, where it may be operated for long periods of time in a very stable fashion. At this temperature, both the diffusion and

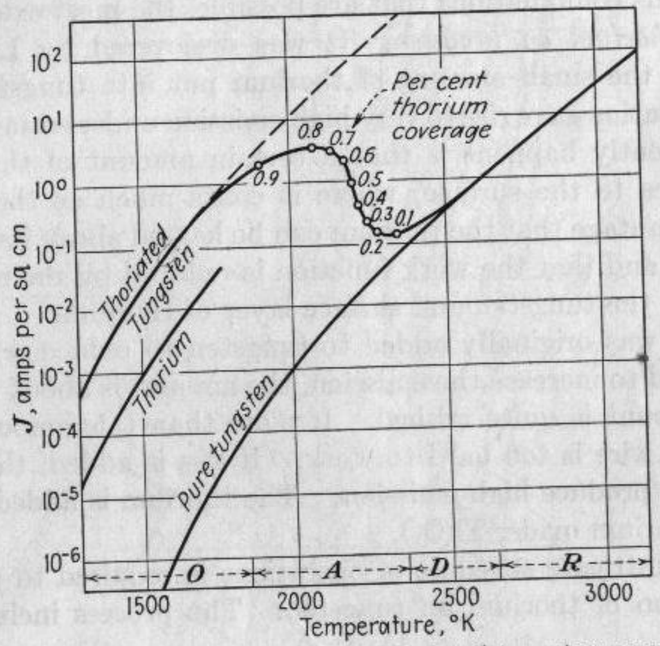


FIG. 4.11.—The emission of thoriated tungsten as a function of temperature. *O* indicates operating range of temperatures; *A*, activation range; *D*, diffusion range, and *R*, reduction range.

evaporation are low, but there is a sufficient preponderance of diffusion to maintain a good emitting surface. Any temperature below 1900°K is suitable for operation. At this temperature, the tungsten surface is about 85 per cent covered by thorium, and the life of the coating is several thousand hours. If the temperature is reduced, the effective work function is decreased, the life is increased, the percentage surface coverage is increased, but the emission is decreased.

It is interesting to note that the thoriated tungsten filaments are usually operated at 1900°K, which is nearly the melting temperature of thorium, something that could not be done with the pure metal because of its softness at this high temperature. Also, the work function of thoriated tungsten filaments is 2.6 volts for a 100 per cent covered surface, and this work function is lower than the work function either of

tungsten, 4.51 volts, or of thorium, 3.4 volts. The work function of thoriated tungsten is a linear function of the surface coverage given by $\phi = 4.51 - 1.9\theta$ volts, where θ is the fraction of the tungsten surface covered by thorium. The reason why the work function is reduced by having the metals in combination is that most of the electrons in the thorium layer are drawn toward the tungsten base. This produces a dipole layer on the surface, with its positive end outward. This means that in most of the surface region the electrostatic forces are outward, opposing the image forces and thus reducing the work function.

Thoriated tungsten surfaces are always carbonized to increase the life. It has been found that if some of the tungsten is converted to tungsten carbide (W_2C) the evaporation of thorium from its surface is greatly reduced.[1] The rate of evaporation of thorium from a tungsten carbide surface at 2200°K is only about one-sixth of that from an uncarbonized surface at this temperature. Carbonization may be achieved by heating the filament to a temperature of 1600°K in a vapor of some hydrocarbon such as naphthalene or acetylene. It may also be achieved by heating the filament to red heat in an atmosphere of hydrogen while in contact with a carbon surface. As the filament is converted to tungsten carbide, its electrical conductance decreases until when totally converted it is about 6 per cent of the original value. The electrical resistance is therefore an excellent index of the degree of conversion. In practice, it is found that the conversion cannot be carried beyond the point where the conductance is reduced to 80 per cent of its original value, for the tungsten carbide is so brittle that the filament would be dangerously weakened by further action.

The fact that the layer of thorium on tungsten is monatomic is evidenced by at least two aspects of the behavior of the composite emitter surface. (1) If the filament is deactivated by heating to a higher temperature after having been activated, the manner in which the emission reduces with time is independent of the length of time the film has been activated. This indicates that the activation beyond a certain point does not add any more emitting material to the surface, which can be the case only if the layer is monatomic and surplus atoms are lost by evaporation. (2) There is no discontinuity in the emission characteristics during the activation process.

Monatomic films other than thorium on tungsten may be used. It is found that they are not as stable as a thorium layer because of more rapid diffusion and evaporation, and hence they are not much used. Curves showing the emission characteristics of various combinations are

[1] KOLLER, L. R., "The Physics of Electron Tubes," 1st ed., McGraw-Hill, New York, 1934.

shown in Fig. 4.12. In this figure the notation *O-W* means that the emitting metal is on an oxidized tungsten surface.

4.4. Oxide Emitters. In 1904, Wehnelt discovered that copious electron emission could be obtained from alkaline-earth oxide coatings. The entire development of small vacuum tubes is based upon this discovery, for oxide coatings are used almost exclusively as a source of emission in them. The alkaline-earth metals that are readily available

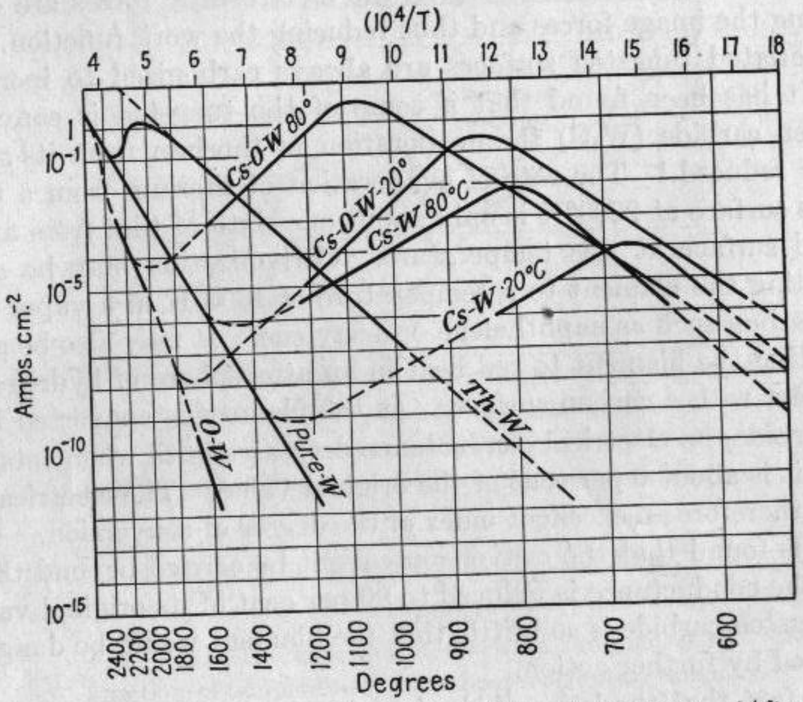


FIG. 4.12.—The emission of monatomic films on tungsten. (*After Dushman.*)

are barium, strontium, and calcium, and it is their oxides that have been found to give such high emission. Modern oxide coatings are usually a half-and-half mixture of the oxides of barium and strontium. Such a coating will give high emission at low temperatures with a high thermal efficiency; thus at 1000°K an emission of 100 ma per cm² at an efficiency of 20 ma per watt input is readily obtained. This is about the same emission as is given by a tungsten filament at 2300°K, but the emission efficiency here is only 1 ma per watt. The oxide coatings may be applied either to an indirectly heated cathode surface or directly to a filament. They are particularly well adapted to making specially shaped unipotential cathodes.

Theory of Oxide Emission. Oxide emission has been the subject of extensive study for the last 30 years though it has not been until recently

that an explanation of the action has been available in fairly complete form.[1–5]

The accumulated evidence indicates that the emission takes place from particles of free metal on the surface of the oxide coating. The free metal is made available by the following mechanisms:

1. Chemical reaction of the oxide with the core metal.
2. Electrolytic reduction by the gradient of potential through the oxide coating.
3. Reduction of the oxide by positive-ion bombardment.

It was discovered early that the core metal played a part in the electron emission. It was even believed that electrons were liberated at the core. This was disproved by showing that there was no emission from the core metal when the coating was removed by mechanical shock. Further, the emission was shown to be independent of the size and shape of the core. Also, the photoelectric work function of the oxide surface was the same as the thermionic work function. However, the most conclusive evidence that the emission is from the surface is that the same emission characteristics are obtained from an oxide coating if metal is vaporized onto the surface as is obtained by the normal process of activation. Different core metals do, however, exhibit different effects upon the emission. In the order of their reaction titanium, tantulum, nickel, and molybdenum will react with the alkaline-earth oxides to produce core-metal oxide and free alkaline earth. The action is evidenced by the fact that oxides can be activated by heating alone. The titanium reaction is probably responsible for the excellent performance obtained with cores of "Konel" metal, which is an alloy of nickel, iron, cobalt, and titanium. The metal most used for core metals is nickel, which is preferred because of its excellent physical properties and low cost.

Free alkaline-earth metal is also made available by the electrolytic action associated with the passage of current through the coating. The earth oxides dissociate under the usual condition of polarity. The metal ion goes to the core, and the oxygen ion is liberated. This action can be detected by the liberation of oxygen.

Dissociation of the oxides is also caused by positive-ion bombardment.

[1] BLEWETT, J. P., Properties of Oxide Coated Cathodes, *Jour. Appl. Phys.*, vol. 10, Part I, October, 1939, pp. 668–679; Part II, pp. 831–848, December, 1939.

[2] DUSHMAN, S., Thermionic Emission, *Rev. Modern Phys.*, vol. 2, pp. 381–476, October, 1930.

[3] RIEMANN, *op. cit.*

[4] BECKER, *op. cit.*

[5] BLEWETT, J. P., Oxide Coated Cathode Literature, 1940–1945, *Jour. Appl. Phys.*, vol. 17, pp. 643–647, August, 1946.

Even in the best vacuums there are enough ions present to give an appreciable action. The fact that activation is greatly facilitated by application of a positive potential to the tube in processing is considered sufficient evidence of the existence of this action.

As with the atomic-film emitters the resultant work function is lower than that of the pure metals alone, and these are already very low. Reported values of the work function of oxides have shown a tremendous variation until recently, when improved vacuum techniques and a better understanding of the mechanism have given rise to some fairly consistent values. The work functions of the oxides are now believed to lie within 25 per cent of the following values:

BaO	1.1 volts
SrO	1.4 volts
CaO	1.9 volts
BaO + SrO	1.0 volts

Emission from the combination of barium and strontium oxides is seen to be better than from either one alone. The reduction in work function over that of the pure metals is again probably due to an electrical double layer formed by a monatomic coating of the pure metal on the oxide. Values of the emission constant A also show a great range of variation as reported by various observers. It has been found that both the emission constant and the work function change with the degree of activation of the oxide coatings. Both decrease with activation, and experimentally it is found that the work function is a linear function of the logarithm of the emission constant. Properly speaking, it is not correct to ascribe an emission constant to oxide coatings, for the emission law in this case is slightly different from the Richardson-Dushman law.[1] An equivalent emission constant is of the order of 0.01 amperes per cm^2 per deg.

Electron-microscope studies of oxide emission show that there is no relation between surface irregularities and emission.[2] Variations in work function are observed with orientation of crystal faces. The emission surface does not change much with degree of activation though the emission may change greatly. Emission is improved by reducing oxide particle size, as may be done by using colloidal particles.

In operation, an oxide cathode has to establish an equilibrium between rate of production of free emitting metal and evaporation of the same. This means the establishment of an equilibrium between electrolysis, diffusion, and evaporation. This latter will be disturbed if the temperature of the oxide or the amount of current is changed. Under normal

[1] BLEWETT, Properties of Oxide Coated Cathodes, Part I, *op. cit.*

[2] HEINZE, W., and S. WAGENER, Vorgänge bei Aktivierung von Oxydkathoden, *Zeit. fur Tech. Phys.*, vol. 17 (No. 12), pp. 645–653, 1936.

conditions the equilibrium adjusts itself to the current drawn so quickly that no change is evident. If, however, the cathode temperature is low or if the emission is partly contaminated or partly exhausted, there will be evident an adjustment of emission over a period of seconds or even minutes as the current drawn is suddenly changed.

The adjustment is of the following nature: If the voltage on a tube is increased, the current immediately increases and then drops slowly, coming to rest at a value between the previous and initial value. If the voltage is decreased, the current will immediately decrease and then slowly rise to a value between the previous and initial value.

Activation of Oxide Emitters. Since the alkaline-earth oxides are not stable in air, the coating must be applied to the cathode or filament in the form of a carbonate or hydroxide. The carbonates are most extensively used, being held to the surface with an organic binder. Coatings of a thickness of 0.010 to 0.020 in. work well. When a coating has been applied and the tube evacuated, the coating is activated by first heating it to a temperature of about 1500°K for a few minutes. This reduces the carbonates to oxides, and during this time copious CO_2 is evolved. Considerable thermal reduction also occurs, with attendant evaporation of liberated metal. The oxide coating is then operated at a temperature of about 1000°K with a potential of about 100 volts applied to an adjacent electrode through a protective resistor. Electrolysis and positive-ion bombardment then occur, and the emission will build up slowly to a final value, when the filament will be ready for use.

Various other methods of applying coatings may be used. Heating in air is recommended to eliminate the organic binder. For a water paste the coating should be baked in an inactive gas to get good adherence. Hydroxides, which are very good for coating tungsten, may be dipped and then baked in air to get a so-called "combined coating."

Specific Emission Characteristics. The lines of Fig. 4.5 show the behavior of oxide coatings in comparison with other emitters. The low work function is evident from the small negative slope of the curve. Emission as a function of power is shown in Fig. 4.7 in contrast with other emitters. The higher emission efficiencies are evident. The emission obtainable from oxide coatings has increased with the years. This may continue, though an increase over present values by more than a factor of 10 is not probable. Some comparative emission efficiencies are

	Ma per Cm^2 per Watt
Pure tungsten filament	2–10
Thoriated tungsten filaments	5–100
Oxide-coated indirectly heated cathodes	10–200
Oxide-coated filaments	200–1,000

Under normal conditions the life of an oxide coating should be several thousand hours. Cessation of emission is due to exhaustion of free metal in the oxide. In mixed coatings there is a preferential evaporation of the barium, which finally leaves the relatively less efficient strontium to give a greatly reduced emission.

Oxide coatings are more easily damaged or poisoned than any other type of coating. They are particularly susceptible to poisoning by oxygen. Emission may be reduced by several powers of 10 by the presence of oxygen at a pressure of 10^{-4} mm of mercury, while a pressure of 10^{-3} mm will inhibit emission completely. Oxide coatings are seldom used on tubes where they will be subjected to bombardment of more than 1,000 volts. Bombardment by particles of higher energy will disintegrate an oxide coating completely.

Transient Emission. The monatomic layer of barium of the oxide coating has tremendous instantaneous-emission potentialities. Such a layer may yield instantaneous emission as great as 100 amperes per cm^2. When short-time high voltages are applied, such large emission may be realized.[1,2] The high voltage exhausts the available emission in a time of the order of milliseconds. When this happens, the supply of free barium must be resupplied through processes of reduction and diffusion. Since this takes an appreciable time, a current-voltage plot of a diode operated under these conditions at 60 cycles exhibits pronounced exhaustion effects, giving rise to a loop in the retrace characteristic. When a very sharp pulse of voltage is applied to an emitting surface, the emitted current consists of a capacitive displacement component as well as the conduction component. As a result, the current pulse will generally have an initial peak with a subsequent rapid decay.

4.5. Schottky Effect. A departure from the Richardson-Dushman emission equation occurs when the emitting surface is subjected to a strong positive potential gradient. Effectively the field reduces the work function. As a result, the current from an emitter increases with the potential applied even though the temperature is kept constant and the emission is not affected by the space charge of the electrons.

The action may be understood by referring to Fig. 4.13, in which the effect of a constant gradient of potential upon the normal potential barrier at the surface of the emitter is shown. The combination of the constant gradient and the normal potential barrier is seen to give a new potential barrier, which has a maximum at a certain distance d_c from

[1] SCHADE, O. H., Analysis of Rectifier Operation, *Proc. I.R.E.*, vol. 31 (No. 7), pp. 341–361, 1943.

[2] COOMBES, E. A., Pulsed Properties of Oxide Cathodes, *Jour. Appl. Phys.*, vol. 17, pp. 647–654, August, 1946.

the surface. This distance is known as the "critical escape distance" because once an electron gets beyond this distance the electrostatic forces are outward rather than restraining and thus an electron keeps on moving. Upon equating the image field with the gradient, the maximum of the restraining potential is found to occur at a distance

$$d_c = \frac{1}{2}\sqrt{\frac{e}{4\pi\varepsilon_0 E}} \tag{4.5}$$

where e is the charge on the electron, E is the potential gradient, and ε_0 is the dielectric constant of free space of value 8.85×10^{-12} for rationalized mks units. The crest of the potential barrier has been reduced by the work the electron would have to do to overcome the image force from the surface from d_c to infinity. This amount of work is $\frac{e}{2}\sqrt{\frac{eE}{4\pi\varepsilon_0}}$ volts. The work function is further reduced the same amount owing to the fact that the potential at the distance d_c is reduced by the amount $d_c E$. The total reduction in the effective work function is thus $e\sqrt{\frac{eE}{4\pi\varepsilon_0}}$ volts. When this correction is made for the work function in the Richardson-Dushman equation, it is found that the ratio of the emitted current in the presence of the strong electric field to the normal emission current is given by

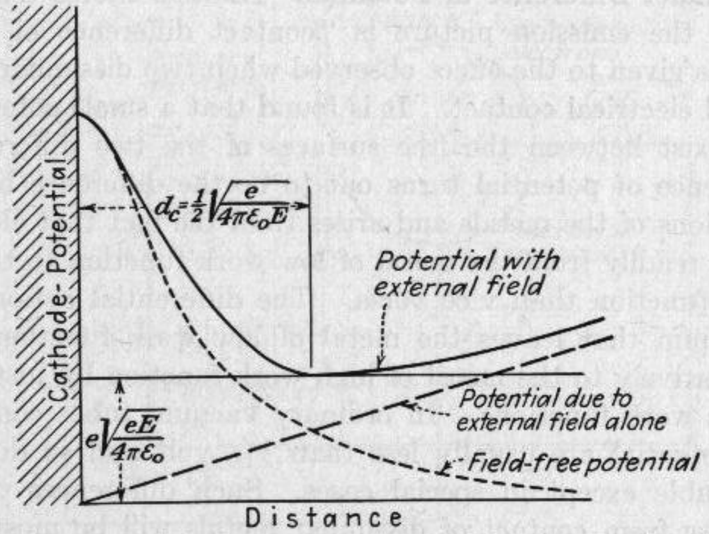


FIG. 4 13.—Diagram of the potential barrier associated with the Schottky effect.

$$\frac{J_E}{J} = \epsilon^{4.403E^{1/2}/T} \tag{4.6}$$

where J_E is the emission-current density in the presence of the strong electric field, J is the normal emission-current density, ϵ is the Napierian base 2.718, E is now the negative gradient of potential in volts per centimeter, and T is the temperature in degrees Kelvin. This equation may be verified experimentally by plotting the logarithm of J_E against the square root of E. The experimental results are found to give a good straight line for all but low values of gradient at which the current drops more rapidly than this simple theory predicts. The slope of the line is, of course, $\frac{4.403 \log_{10} \epsilon}{T}$.

4.6. Contact Difference of Potential. Another factor that occasionally enters the emission picture is "contact difference of potential." This term is given to the effect observed when two dissimilar metals are put in good electrical contact. It is found that a small potential difference will exist between the free surfaces of the two different metals. This difference of potential turns out to be the difference between the work functions of the metals and arises from the fact that electrons can move more readily from the metal of low work function to the metal of high work function than vice versa. The differential action results in an equilibrium that leaves the metal of low work function positively charged relatively to the metal of high work function by just the difference of the work functions. In ordinary vacuum tubes contact differences of potential are usually less than $\frac{1}{10}$ volt and so do not cause serious trouble except in special cases. Such differences of potential as may arise from contact of dissimilar metals will be most serious in such places as the cathode–control-grid circuit.

4.7. Secondary Emission. Another form of emission that plays an important role in vacuum tubes is secondary emission. This occurs when a surface is struck by electrons or ions of appreciable velocity. Secondary emission caused by the bombardment of electrons is the more important case and occurs whenever the striking electrons have energies corresponding to a few volts or more. When this happens, the striking electrons may knock one or more electrons out of the material, giving rise to a reverse component of current. The electrons knocked out of a material, known as "secondary" electrons, may number more than the "striking," or "primary," electrons. There is no violation of the conservation of energy law when this happens, for the velocity of the secondary electrons is for the most part very low. Secondary emission is commonly encountered in multiple-electrode tubes, where it has the effect of altering somewhat the normal primary-electron current characteristics. It occurs in cathode-ray tubes where the beam electrons hit the fluorescent screen, and is necessary there to complete the circuit for the current flow. It

is deliberately used in a number of types of electron-multiplier tubes, where it makes possible a high amplification of current by a purely electronic action.

Secondary-emission characteristics of materials are measured by means of the apparatus shown schematically in Fig. 4.14. In the arrangement shown a beam of electrons is directed at a target inside of a sphere at a higher potential, which attracts the secondary electrons liberated at the target. The ratio of secondary- to primary-electron current can be read for any primary-electron potential.[1] For a long time there were great discrepancies in the reported secondary-emission characteristics of

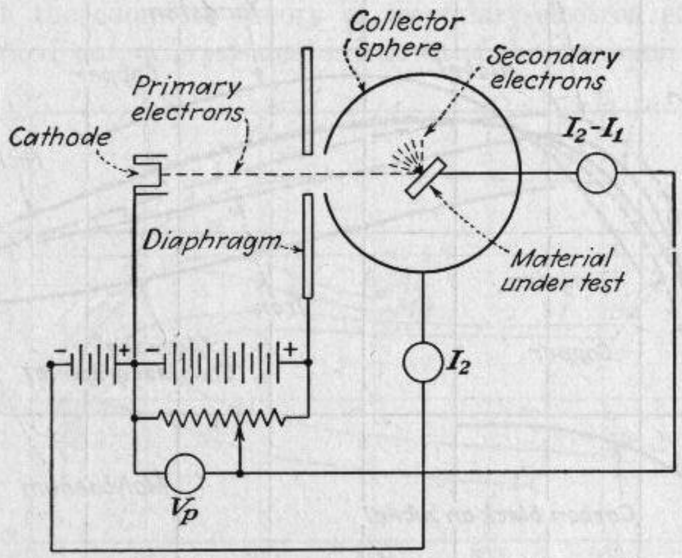


FIG. 4.14.—Apparatus for the measurement of secondary-emission characteristics.

the various metals. It was evident that small traces of impurities or surface contaminations made a great difference in the secondary-emission characteristics. Techniques have now been refined to the point where the values reported by various investigators are fairly consistent. The average secondary-emission characteristics of the materials commonly used in vacuum tubes when only the ordinary precautions against contamination are taken are shown in Fig. 4.15.[2]

Variation of Secondary Emission with Primary-electron Potential. In Fig. 4.16 are shown the secondary characteristics of the common metals presented in curve form, giving the ratio of secondary- to primary-elec-

[1] See KOLLATH, R., Sekundarelektronemission fester Körper, *Physik. Zeit.*, vol. 38, pp. 202–224, Mar. 15, 1937, for an excellent discussion of methods of measurement and results obtained up to that date.

[2] HARRIES, J. H. OWEN, Secondary Electron Radiation, *Electronics*, vol. 17, pp. 100–108, 180, September, 1944.

tron current as a function of the primary-electron potential as reported by Bruining and DeBoer.[1] These results probably are more reliable than any previously reported, for the investigators used a special apparatus in which the metal to be tested was evaporated onto the target in a vacuum just before the measurement was made. The results presented show lower ratios of secondary- to primary-electron current than those

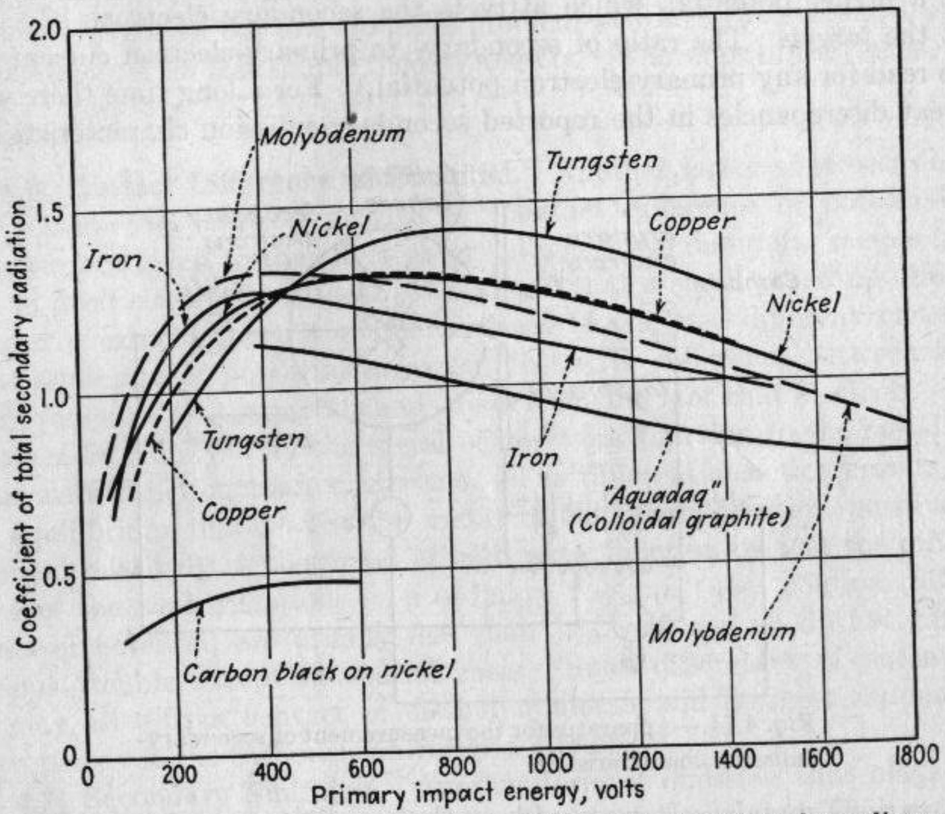


FIG. 4.15.—Secondary-emission characteristics of the metals under ordinary conditions. The curve shows the ratio of the number of secondary to primary electrons for various primary-electron impact velocities expressed in volts. (*After Harries.*)

previously reported. This is probably due to the fact that with previous handling the metals became partly oxidized and oxidized surfaces are

[1] BRUINING, H., and J. H. DEBOER, Secondary Emission, Part I, Secondary Emission of Metals, *Physica*, vol. 5, pp. 17–30, January, 1938; Part II, Absorption of Secondary Electrons, *Physica*, vol. 5, pp. 901–912, December, 1938; Part III, Secondary Electron Emission Caused by Bombardment with Slow Primary Electrons, *Physica*, vol. 5, pp. 913–917, December, 1938; Part IV, Compounds with a High Capacity for Secondary Electron Emission, *Physica*, vol. 6, pp. 823–833, August, 1939; Part V, Mechanism of Secondary Electron Emission, *Physica*, vol. 6, pp. 834–839, August, 1939; Part VI, Influence of Externally Adsorbed Ions and Atoms, on the Secondary Electron Emission of Metals, *Physica*, vol. 6, pp. 941–950, October, 1939.

known to have higher secondary emission than those which are not. The curves of Fig. 4.15 show that all the metals have a low secondary emission at low primary-electron potentials. Most of the metals have a maximum secondary emission between 200 and 400 volts of primary potential, which then decreases slowly, becoming constant at a value between 50 and 95 per cent of the maximum value. Most of the uncontaminated metals have a maximum ratio of secondary- to primary-electron currents less than 1 though it should be remembered that metals as encountered in tubes are seldom uncontaminated and will have maximum ratios of the order of 1 to 5.

Although the complete theory of secondary-electron emission is as yet not worked out, a great deal is known of the mechanism.[1,2] When

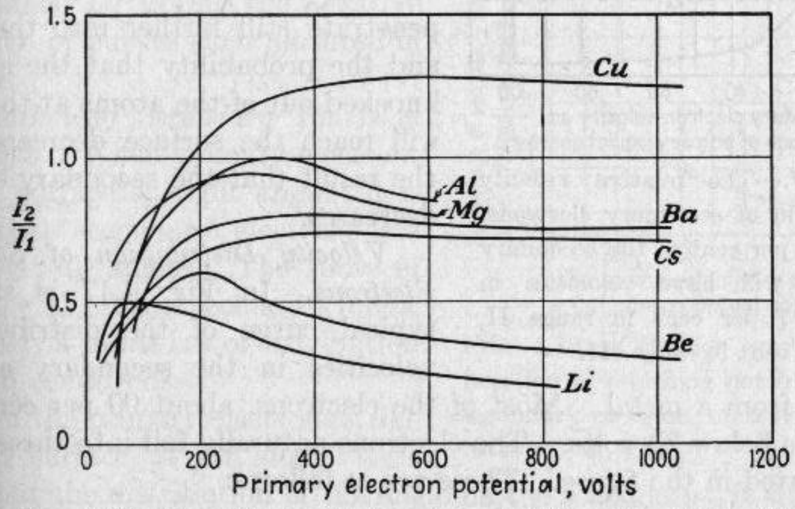


FIG. 4.16.—Secondary-emission characteristics of metals with inappreciable surface contamination. (*After Bruining and DeBoer.*)

primary electrons strike a surface at right angles, they may knock electrons out of the atoms near the surface and those with velocity components directed toward the surface may be able to overcome the surface-potential restraints and escape from the metal. Each primary electron may shake up several atoms, thus giving rise to several electrons emitted per primary electron. It should be noted that the source of secondary electrons lies almost entirely in the electrons of the surface atoms and not in the free electrons of the metal. If a normally directed primary electron strikes a free electron, it cannot give it a component of velocity directed toward the surface. Electrons knocked out of atoms, however, may have such

[1] *Ibid.*, Part V.

[2] WOOLDRIDGE, D. E., Theory of Secondary Emission, *Phys. Rev.*, vol. 56, pp. 562–578, Sept. 15, 1939.

a component. As the potential of the primary electron is increased, it will at first knock out more and more secondary electrons. However, as the potential is further increased, the surface atoms are exposed to the primary-electron forces for a shorter time, *i.e.*, the so-called "collision diameter" decreases, and the primary electron will first knock electrons out of atoms when it has slowed down upon penetration into the metal. Thus at the maximum of emission it is believed that the majority of the secondary electrons are liberated a depth of several atoms into the metal.[1] Beyond this potential the primaries penetrate still farther into the metal, and the probability that the electrons knocked out of the atoms at this depth will reach the surface decreases, with the result that the secondary emission decreases.

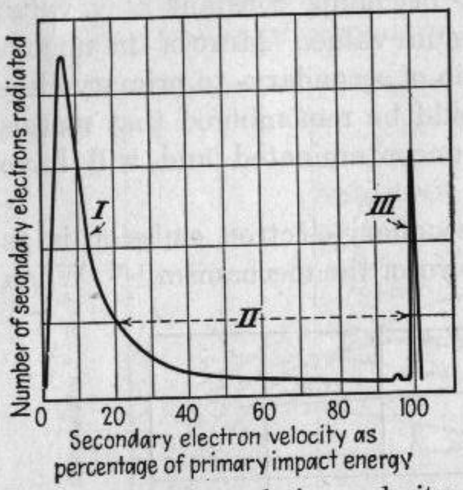


FIG. 4.17.—The relative velocity distribution of secondary electrons. About 90 per cent of the secondary electrons will have velocities in range I, 7 per cent in range II, and 3 per cent in range III.

Velocity Distribution of Secondary Electrons. In Fig. 4.17 is shown a typical curve of the distribution of velocities in the secondary electrons emitted from a metal. Most of the electrons, about 90 per cent, have velocities below 20 volts. The electrons naturally fall into three groups as indicated in the figure. These are as follows:

Group I—0 *to* 20 *volts.* This group comprises about 90 per cent of all the secondaries for primary potentials of 50 volts or more. There is a pronounced maximum in this group at about 10 volts. These are the electrons which are shaken out of the atoms as a result of the passage of the primary electrons and do not have much energy.

Group II—20 *volts to* 98 *per cent of the primary-electron potential.* These comprise about 7 per cent of the total secondary current. They represent high-energy electrons knocked out of atoms and elastic reflections of the primary electrons at a considerable depth in the metal.

Group III—98 *to* 100 *per cent of primary-electron potential.* This group comprises only about 3 per cent of the secondary current and has a maximum at about 99 per cent of the primary-electron potential. This group arises from elastic reflections of primary electrons from atoms near the surface of the metal, not really secondary electrons at all.

[1] BRUINING, H., Depth at Which Secondary Electrons Are Liberated, *Physica*, vol. 3, pp. 1046–1052, September, 1936.

Another representation of the secondary-electron velocity distribution is obtained if potential between sphere and target of the apparatus of Fig. 4.14 is made negative instead of positive and the current of the sphere is measured against the retarding potential. The resultant curve is shown in Fig. 4.18. This curve is an average for measurements on various metals with primary-electron potentials in the range of 275 to 1,000 volts. Curves like those in Fig. 4.17 are obtained by taking the negative derivative of curves such as those in Fig. 4.18.

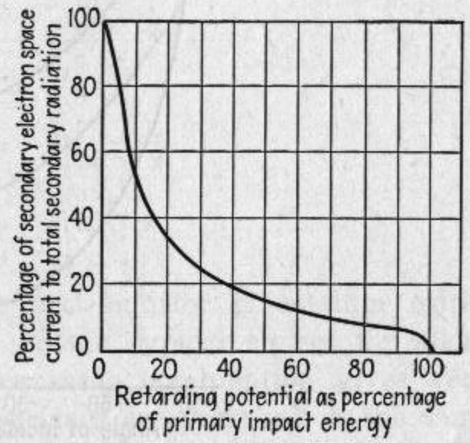


FIG. 4.18.—Collector current as a function of retarding potential of the secondary-emission measuring apparatus of Fig. 4.14.

Variation of Secondary Emission with Angle. When primary electrons strike a surface at right angles, it is found that secondary electrons are emitted at all angles. The spray of secondary electrons seems to follow very nearly a cosine law of distribution under all conditions.

When the primary electrons strike a metal surface at an angle, it is found that the distribution of the angle on the secondaries is still nearly a cosine-law variation. More important than this is the fact that the secondary- to primary-emission ratio increases as the primary electrons strike more nearly parallel to the surface. Some typical curves showing the variation of the secondary- to primary-emission ratio are given in Fig. 4.19. The increase in secondary emission with angle is largely due to the fact that at angles other than normal the primary electron may knock free electrons out of the metal as well as electrons out of atoms. The variation of emission is given quite closely by[1]

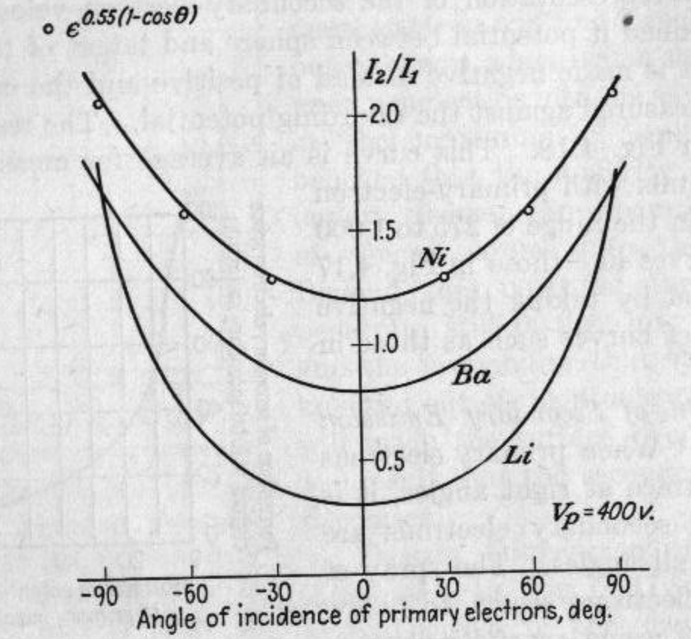
$$R_\theta = R_0 \epsilon^{p(1-\cos\theta)} \tag{4.7}$$

where θ is angle between normal and direction of primary electrons
R_θ is ratio of secondary to primary electrons at angle θ
R_0 is ratio of secondary to primary electrons at angle zero
ϵ is Napierian base 2.718

[1] BRUINING and DEBOER, *op. cit.*, Part II.

p is a coefficient that increases with primary potential and is proportional to the primary-electron penetration

Secondary Emission of Composite Layers. Certain combination surfaces have been found to have pronouncedly higher secondary emission



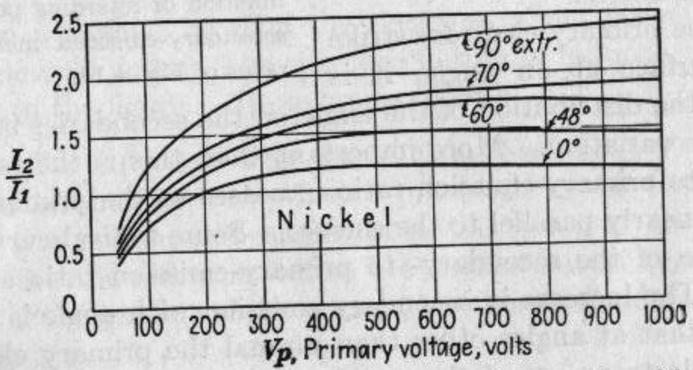


FIG. 4.19.—Variation of secondary-emission ratio with angle of primary impact. Note that the secondary-emission ratio increases as the angle of incidence becomes more nearly grazing. (*After Bruining and DeBoer.*)

than the pure metals. Such surfaces are the alkalihalides on a base of the alkali metal and alkali oxides on various metal bases. All these combinations show the same general secondary-emission characteristics as do the pure metals except that the current ratios instead of being in the vicinity of unity may be as high as 8 to 11. The velocity distribution for composite surfaces is much narrower than for the pure metals, *i.e.*, a given percentage of the total electrons are included in a lower range of velocities, 85 per cent in the first 3 volts. Below are given data on some of the alkali halides.[1]

TABLE III
MAXIMUM SECONDARY-EMISSION RATIOS OF ALKALI HALIDES

Compound	Maximum Ratio
LiF	5.6
NaF	5.7
CaF_2	3.15
NaCl	6.8
KCl	7.5
RbCl	5.8
CsCl	6.5
NaBr	6.25
NaI	5.5
KI	5.6

Of the alkali oxides, by far the best emitter is caesium oxide, partly reduced, on a base of silver. Some typical curves for alkali oxides are shown in Fig. 4.20. This same combination gives very high photoemission. Photoemissive surfaces are prepared in the same way.

In connection with composite surfaces it should be noted that a combination with a low work function does not necessarily have a high secondary- to primary-electron ratio, and vice versa. Thus tungsten with a work function of 4.52 volts has a maximum ratio of 1.5. Contamination with oxygen increases the work function to 9.25 volts but increases rather than decreases the maximum ratio.[2] This probably means that electrons are more readily knocked out of the surface atoms and so give increased secondary emission even though they require more energy to escape from the surface. For a given combination of elements, however, the secondary emission usually increases with decreasing work function. Thus, if caesium on caesium oxide on silver is contaminated with oxygen, the work function increases and the secondary emission decreases. Also, in the case of molybdenum partly coated with barium the work function passes through a maximum with a given percentage of the surface covered, as is evidenced by the photoelectric emission. The secondary emission passes through a maximum with the same

[1] BRUINING and DEBOER, *op. cit.*, Part V.

[2] ZWORYKIN, V. K., and G. A. MORTON, "Television," p. 32, Wiley, New York, 1940.

percentage of surface coverage though the maximum is not nearly so pronounced.[1]

Secondary Emission of Insulators. Insulators as well as conductors may emit secondary electrons. Measurements on insulators are more difficult to make because the potential of the insulator cannot be measured directly. The characteristics can, however, be deduced from the potential that the insulator assumes relative to a spherical collector

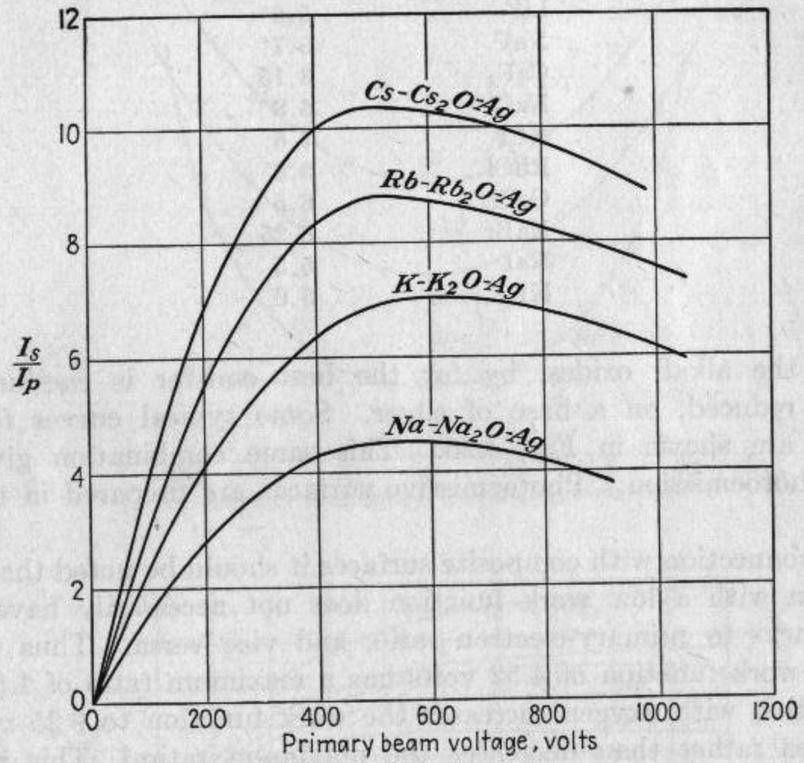


FIG. 4.20.—Secondary emission of the alkali oxides. (*Reprinted by permission from "Television" by V. K. Zworykin and G. A. Morton, Wiley, New York,* 1940.)

electrode when bombarded with electrons of different potentials. The general features of the secondary emission of insulators may be summed as follows:[2] Insulators exhibit curves of ratio of secondary- to primary-electron current versus primary-electron potential that are similar to those of the metals. Ratios usually exceed 1 over a considerable range of potentials, a maximum occurring between 300 and 800 volts. As with the metals, the ratio rises rapidly to a maximum and then drops slowly. As with the metals, most of the secondary electrons are emitted perpendicularly to the surface, following very nearly a cosine law of distribution

[1] BRUINING and DEBOER, *op. cit.*, Part VI.
[2] KOLLATH, *op. cit.*

regardless of the direction of the primary electrons. Upon bombardment at an angle the secondary- to primary-electron ratio increases as the primary electrons strike more nearly parallel to the surface up to a critical angle, beyond which the ratio drops to a small value less than unity and remains nearly constant. The critical angle depends upon the material and is a function of temperature, the angle with the normal increasing with temperature. The explanation of the sudden drop in emission with increased angle with the normal of primary-electron bombardment seems to be that a layer of negative charge forms on the surface which traps, by a space-charge action, the primary electrons and the secondary electrons they would have freed.

In normal action an insulator will have its potential influenced by its secondary-emission characteristics. The action will depend upon the primary-electron energy relative to the secondary-emission characteristics. Action can be divided into three cases as follows:

1. Primary-electron potential below that at which secondary- to primary-current ratio is unity. Here the number of secondary electrons emitted is less than the number of primaries, and so the insulator acquires a negative potential that is large enough to repel most of the primaries. This constitutes a blocking action. The insulator is finally in stable equilibrium at zero potential.
2. Secondary- to primary-current ratio greater than unity. Under this condition the insulator gives off more electrons than it acquires and so becomes more positive than its surroundings. When this happens, the insulator reattracts the slow secondaries and so remains a few volts more positive than the potential through which the primary electrons have been accelerated.
3. Primary-electron potential greater than that at which secondary- to primary-current ratio has dropped to unity. In such cases the insulator will gain more electrons than it loses and so will become more negative in potential until the primary electrons are retarded to the point where the ratio of secondary to primary current is unity. At this potential, the primary- and secondary-electron currents are equal, and the insulator is in stable equilibrium.

CHAPTER 5

DETERMINATION OF POTENTIAL FIELDS

THE fundamental theoretical technique necessary for the study of the internal behavior of a vacuum tube is that of determining the distribution of the electric potential within the tube. From the determination of the electric potential within a tube can be deduced the amplification factor of the tube, the focusing properties of the electrodes, and the current-voltage characteristics. In short, the determination of the distribution of the electric potential within a tube is the point of departure for the study of almost all its characteristics.

The methods of determining the potential fields of vacuum tubes are rather special. The most extensive information is obtained from conformal transformations and from solutions of the Laplace differential equation. The particular transformations and functional forms most frequently encountered in tubes are ordinarily given only a fraction of the total space allotted to the entire subject of electrostatics in books devoted to this subject. For this reason a brief review will be given of all the standard methods of determining potential fields, including some numerical and graphical methods, so that the elegance of the special methods mentioned will be appreciated.

5.1. Units and Dimensions. In this book there will be used the system of rationalized mks units. For this system the units of length, mass, and time are the meter, kilogram, and second, respectively and the electrical units are the usual practical ones—the volt, the ampere, the coulomb, etc. The term "rationalized" indicates that the factor 4π has been incorporated into the arbitrary constants in such a way that the greatest over-all simplicity of all relations is obtained. This is done in such a way that the factor 4π does not appear in relations involving plane geometry and rectangular coordinates but does appear in relations involving spherical geometry. A further feature of the rationalized mks system of units is that the equivalent dielectric constant of free space and the equivalent permeability of free space are not unity but have some specific values. These are the only two values that need to be known in this system to work practical problems, whereas in some of the other systems a whole table of conversion factors has to be invoked every time a practical problem is solved.

5.2. Fundamental Quantities and Definitions: Forces between Charges. All electrostatic relations are based upon the application of the observed effects of charges upon one another. Qualitatively, the observations are that there are two kinds of charges, that like repel and unlike attract, that the force between charges decreases as the distance between them increases. Quantitatively, all this is expressed by *Coulomb's law*,

$$F = \frac{q_1 q_2}{4\pi\varepsilon r^2} \tag{5.1}$$

where F is the radially directed force in newtons (1 newton equals 10^5 dynes) between charges q_1 and q_2 in coulombs, r is the distance between charges in meters, and ε is the so-called "dielectric constant" of the medium. The dielectric constant is equal to the product of the relative dielectric constant and the dielectric constant of free space,

$$\varepsilon = \varepsilon_r \varepsilon_0 \tag{5.2}$$

where ε_r is the relative dielectric constant as would be determined by the ratio of capacity of a condenser using the medium and free space as dielectric and ε_0 is the equivalent dielectric constant of free space whose value turns out to be 8.85×10^{-12} farad per meter in rationalized mks units.

The region in the vicinity of electric charges is referred to as the *electric field*. The *electric intensity* E at any point in such a field is the *force per unit charge* on a small test charge placed at the point. The intensity, which will also be shown to be the negative gradient of the electric potential, is a *vector* quantity in that it has both magnitude and direction.

Intensity at a distance r from a charge q is, by Coulomb's law,

$$|E| = \frac{q}{4\pi\varepsilon r^2} \tag{5.3}$$

Where more than one charge is concerned,

$$E_x = \frac{\sum_n q_n \cos(x,r_n)}{4\pi\varepsilon {r_n}^2} \tag{5.4}$$

$$E_y = \frac{\sum_n q_n \cos(y,r_n)}{4\pi\varepsilon {r_n}^2} \tag{5.5}$$

The summation must be taken by a summation of components where (x,r_n) is the angle between a line parallel to the x axis and the vector from the charge q_n to the point at which the intensity is being determined.

A line of force, or a line of flux, is a line drawn so that it has everywhere the direction of the electric intensity. Lines of flux originate on positive charges and terminate on negative charges. In the rationalized mks system of units *one line of flux emanates from every unit positive charge.* The density of the flux lines is known as the *displacement* or flux density. Displacement and intensity are related by the expression

$$\boldsymbol{D} = \varepsilon \boldsymbol{E}^* \tag{5.6}$$

where $\boldsymbol{D}$ is the displacement, or number of flux lines per square meter, and ε is the dielectric constant of the medium. Equation (5.6) is, for homogeneous isotropic dielectrics, strictly analogous to the expression $\boldsymbol{B} = \mu \boldsymbol{H}$, which applies for magnetic fields.

The potential at any point in an electric field is defined as the *work per unit charge* required to bring a small positive test charge from infinity to the point in question (symbol V). Potential is a scalar quantity, *i.e.*, completely specified when its magnitude alone is given. Applying this definition to obtain the potential at a distance r from a charge q,

$$V = \int_{\infty}^{r} F\, dr = -\int_{\infty}^{r} \frac{q}{4\pi\varepsilon r^2}\, dr = \frac{q}{4\pi\varepsilon r} \tag{5.7}$$

The minus sign appears because the work is being done against the force. The potential obtained above is in volts if q is in coulombs and r is in meters. The work is independent of the path. The potential at a point due to a number of charges is equal to the sum of the potentials due to the separate charges,

$$V = \sum_n \frac{q_n}{4\pi\varepsilon r_n} \tag{5.8}$$

For a continuous distribution of charge over a surface,

$$V = \frac{1}{4\pi\varepsilon} \int \frac{\sigma}{r}\, da \tag{5.9a}$$

where σ is the surface density of charge, da is the element of area, and the integration is taken over the area of the surface. For a continuous distribution of charge throughout a volume,

$$V = \frac{1}{4\pi\varepsilon} \int \frac{\rho}{r}\, dv \tag{5.9b}$$

* Bold-faced capitals will be employed to designate vector quantities when used in the vector sense. Components of vectors are themselves vectors but may usually be treated as scalar quantities when dealt with separately.

where ρ is the volume density of charge, dv is an element of volume, and the integration is taken over the volume.

The difference of potential between two points in an electric field is defined as the work per unit charge required to bring a small positive test charge from one point to the other. This difference is independent of the path by which it is evaluated.

From the definition of potential it is seen that the intensity is the negative gradient of potential, the negative sign indicating that the force is exerted in a direction opposite to that of increasing potential. The gradient of the potential is a vector having the magnitude and direction of the maximum variation of potential. Thus

$$|\boldsymbol{E}| = -\frac{\partial V}{\partial s} \tag{5.10}$$

The force per unit charge in any general direction is given by

$$E \cos \alpha = -\frac{\partial V}{\partial l} \tag{5.11}$$

where α is the angle between the direction considered and the gradient of potential. Components of intensity are conveniently related to potential by

$$E_x = -\frac{\partial V}{\partial x} \tag{5.12a}$$

$$E_y = -\frac{\partial V}{\partial y} \tag{5.12b}$$

The form that components of intensity have in terms of derivatives of potential depends upon the coordinates in which the potential and distances are expressed. In all cases the component expressions corresponding to Eq. (5.10) have the form of the limiting value of the ratio of an increment of potential to an increment of length in the direction of the variable considered. Expressions for the intensity as a negative gradient of potential are given in Appendix II for the coordinate systems most commonly used.

5.3. Solution of Potential Fields by Summation of Intensities. The electric field around any distribution of charges may be found by summing the forces due to the charges by means of Eq. (5.4). Forces are best summed one component at a time. The procedure can usually be simplified by choosing the axes to take advantage of any symmetries. When an expression for each of the components of intensity has been found, the resultant intensity has a magnitude that is the square root of the sum of the squares of the components. The direction cosines of

the resultant vector are given by the ratio of the respective components to the magnitude of the resultant.

Example: Find the electric intensity on the axis of a right-circular cylinder of radius a and length h at a distance x_0 from the end of the cylinder if the cylinder has a charge uniformly distributed throughout its volume of density ρ. In the configuration of Fig. 5.1 let x be the distance from the point P to the point on the axis corresponding to an element of volume in the cylinder. The elementary volume is given by

$$dv = r\,dr\,d\theta\,dx$$

and the corresponding element of charge is given by

$$dq = \rho\,dv$$

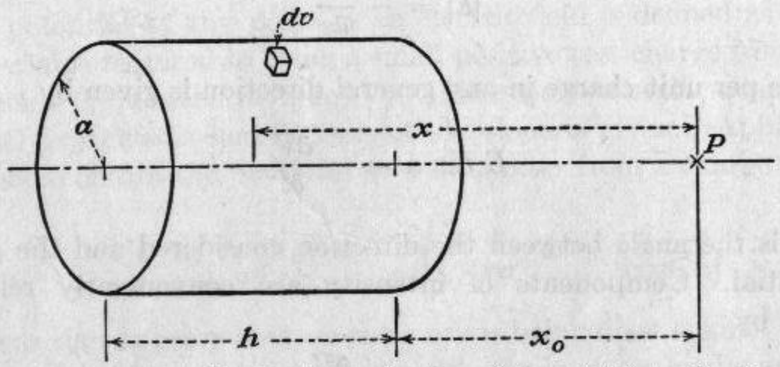


FIG. 5.1.—Notation for the evaluation of the axial intensity due to a cylindrical distribution of charge.

By symmetry there will be only an x component of intensity at the point P on the axis to which the element of charge will contribute

$$dE_x = \frac{\rho r\,dr\,d\theta\,dx}{r^2 + x^2}\,\frac{x}{(r^2 + x^2)^{1/2}}\,\frac{1}{4\pi\varepsilon}$$

which will be recognized as being of the form

$$\frac{dq\cos\alpha}{R^2}$$

This differential expression must be integrated with respect to its three variables, θ from 0 to 2π, r from 0 to a, and x from x_0 to $x_0 + h$. When this triple integration has been performed, the resulting expression for the intensity on the axis is

$$E_x = 2\pi\rho\left[h - \sqrt{(x_0 + h)^2 + a^2} + \sqrt{x_0^2 + a^2}\right]\frac{1}{4\pi\varepsilon}$$

5.4. Summation of Potentials. The potential at any point in a field may similarly be obtained by application of Eq. (5.7). This procedure is in general easier to apply than the direct evaluation of the intensity,

for the summation for potentials is algebraic, whereas that for intensities must be vectorial. The expression for the components of intensity is, of course, derivable from the expression for potential.

Example: Find the potential at a distance c from the center of a spherical shell with inner and outer radii r_1 and r_2 and with a charge uniformly distributed throughout its volume of charge density ρ.

The element of volume in spherical coordinates is

$$dv = r^2 \sin\theta\,dr\,d\theta\,d\phi$$

where the symbols have the significance indicated in Fig. 5.2 and ϕ is the azimuthal angle. Then the potential at the point P due to the element of charge associated with the above element of volume is

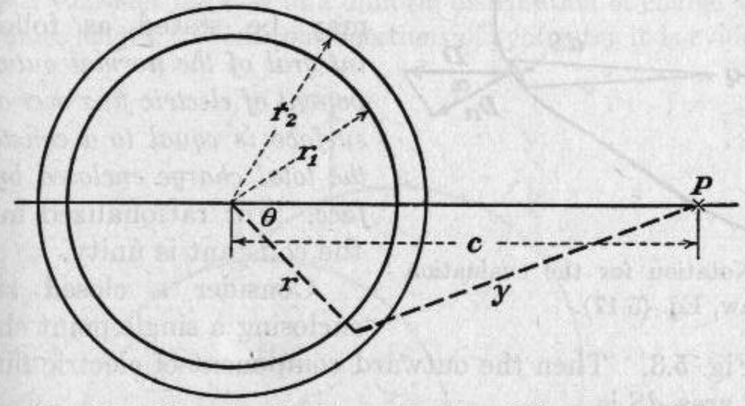


FIG. 5.2.—Notation for the evaluation of the potential due to a charge uniformly distributed throughout a spherical shell.

$$dV = \frac{\rho r^2 \sin\theta\,dr\,d\theta\,d\phi}{y}\,\frac{1}{4\pi\varepsilon}$$

It is convenient to use the distance y instead of the angle θ as a variable. The two quantities are related by the law of cosines

$$y = (c^2 + r^2 - 2cr\cos\theta)^{1/2}$$

so that, for constant r,

$$dy = \frac{cr\sin\theta\,d\theta}{y}$$

Making this substitution into the expression for the element of volume,

$$dV = \frac{\rho r}{c}\,dy\,d\phi\,dr\,\frac{1}{4\pi\varepsilon}$$

so that

$$V = \frac{\rho}{c}\int_{r_1}^{r_2}\int_0^{2\pi}\int_{c-r}^{c+r} r\,dy\,d\phi\,dr\,\frac{1}{4\pi\varepsilon}$$

The result of this integration gives

$$V = \frac{4\pi\rho}{3}\frac{r_2^3 - r_1^3}{c}\frac{1}{4\pi\varepsilon}$$

From this it is seen that the potential at the point P is the same as though the entire charge of the shell were concentrated at its center.

5.5. Gauss's Law. Gauss's law is one of the most useful relations in electrostatics. It enables one to determine quickly the field and potential around any symmetrical distributions of charge. The law may be stated as follows: *The integral of the normal outward component of electric flux over any closed surface is equal to a constant times the total charge enclosed by the surface.* For rationalized mks units, the constant is unity.

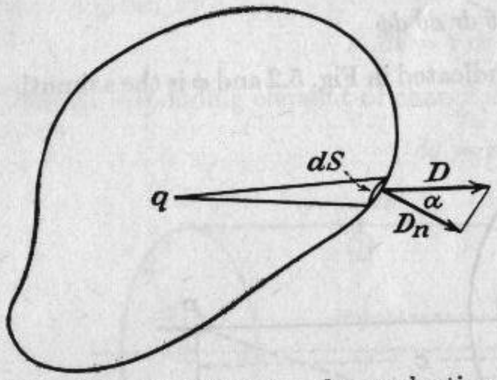


FIG. 5.3.—Notation for the evaluation of Gauss's law, Eq. (5.17).

Consider a closed surface S enclosing a single point charge q as shown in Fig. 5.3. Then the outward component of electric flux for the element of area dS is

$$D_n\,dS = D\cos\alpha\,dS \tag{5.13a}$$

$$D_n\,dS = \frac{q}{4\pi r^2}\cos\alpha\,dS \tag{5.13b}$$

It will be recognized that $\frac{dS}{r^2}\cos\alpha$ is the element of solid angle about the point charge intercepted by the area dS, since solid angle is measured by area intercepted on a unit sphere just as linear angle may be measured by arc length on a unit circle. Thus

$$d\Omega = \frac{dS\cos\alpha}{r^2} \tag{5.14}$$

where $d\Omega$ is an element of solid angle. Then

$$D_n\,dS = \frac{q\,d\Omega}{4\pi} \tag{5.15}$$

If this is integrated over the entire surface surrounding the point charge,

$$\int D_n\,dS = q \tag{5.16}$$

since there are 4π units of solid angle around a point.

Since the law of superposition holds for the potentials due to charges, the integral of the outward normal component of flux is equal to the total charge enclosed when the closed surface contains more than a single charge.

For a volume distribution of charge the law can be written

$$\int D\cos\alpha\,dS = \int\rho\,dv \tag{5.17}$$

where ρ is the volume charge density, v indicates volume, and the other symbols have the previous significance.

Example: Consider the case of a uniform distribution of charge on a circular wire of infinite length. From considerations of symmetry it is evident that the

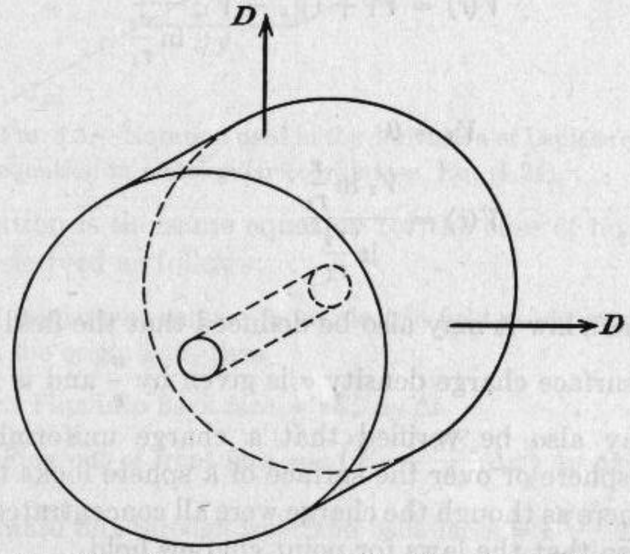


FIG. 5.4.—The flux associated with a linear distribution of charge.

electric field will everywhere be radial and will be constant along the length of the wire. The equipotential surfaces will be cylinders concentric about the wire, and the flux lines will be straight radial lines.

Let the charge be uniformly distributed along the wire with a density of λ units per unit length. Draw a cylinder of radius r about the wire of radius a. Then the electric flux $D = \varepsilon E$ is everywhere outwardly directed as shown in Fig. 5.4. The integral of normal component of flux per unit length of this wire is equal to the product of the displacement and the area of the cylinder per unit length. This product must be equal to the linear charge density, so that

$$2\pi\varepsilon E = \lambda$$

or

$$E = \frac{\lambda}{2\pi\varepsilon r}$$

This gives the intensity at any distance r from a wire with a linear charge density.

The potential at any distance r from the wire is found by integrating the negative of the field with respect to r, giving

$$V = -\frac{\lambda}{2\pi\epsilon}\ln r + C$$

The constant is necessary to adjust the potential to a prescribed value at some particular distance since the potential about a cylindrical wire, unlike that about a point charge, does not vanish at an infinite value of the radius.

In the case of two concentric cylinders of radii r_2 and r_1 having potentials V_2 and V_1, respectively, the potential between them is

$$V(r) = V_1 + (V_2 - V_1)\frac{\ln\frac{r}{r_1}}{\ln\frac{r_2}{r_1}}$$

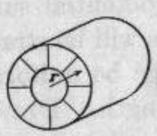
If

$$V_1 = 0$$

$$V(r) = \frac{V_2 \ln\frac{r}{r_1}}{\ln\frac{r_2}{r_1}}$$

From Gauss's law it may also be deduced that the field adjacent to a plane with a surface charge density σ is given by $\frac{\sigma}{\epsilon}$ and is normal to the plane. It may also be verified that a charge uniformly distributed throughout a sphere or over the surface of a sphere looks to an observer outside the sphere as though the charge were all concentrated at the center of the sphere so that the laws for point charges hold.

The above results are summarized in the following table:

Geometry	Plane 	Cylindrical 	Spherical 
Total positive charge	$\sigma \times$ area	$\lambda \times$ length	q
Charge density	σ per unit area	λ per unit axial length	$\frac{q}{\text{area}}$
Intensity E	$\frac{\sigma}{\epsilon}$	$\frac{\lambda}{2\pi\epsilon r}$	$\frac{q}{4\pi\epsilon r^2}$
Potential V	$-\frac{\sigma x}{\epsilon} + C$	$-\frac{\lambda}{2\pi\epsilon}\ln r + C$	$\frac{q}{4\pi\epsilon r}$

5.6. Poisson's and Laplace's Equations. Poisson's and Laplace's equations are differential expressions of Gauss's law applied to an element of volume. Poisson's equation applies to regions containing charge.

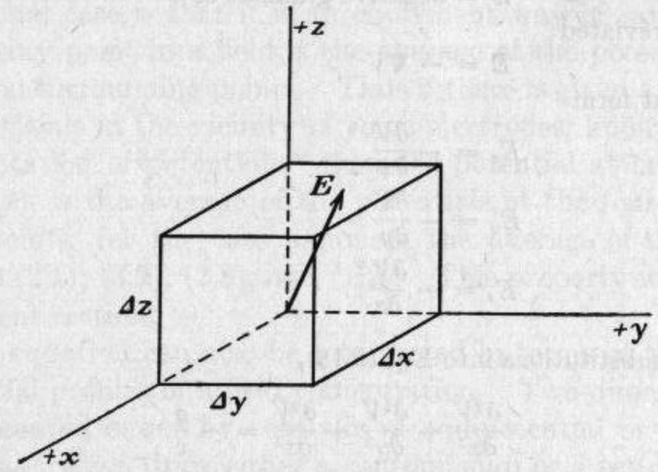


FIG. 5.5.—Notation used in the derivation of Laplace's equation in rectangular coordinates, Eq. (5.24).

Laplace's equation is the same equation for the case of no charge. The equations are derived as follows:

Consider an element of volume in an electric field as shown in Fig. 5.5. If the intensity at the origin is E, then

$$\text{Flux into back face} = \epsilon E_x\,\Delta y\,\Delta z$$

$$\text{Flux out of front face} = \epsilon\left(E_x + \frac{\partial E_x}{\partial x}\Delta x\right)\Delta y\,\Delta z$$

$$\text{Net outward flux through front and back faces} = \epsilon\frac{\partial E_x}{\partial x}\Delta x\,\Delta y\,\Delta z$$

Similarly

$$\text{Net outward flux through left and right faces} = \epsilon\frac{\partial E_y}{\partial y}\Delta x\,\Delta y\,\Delta z$$

and

$$\text{Net outward flux through bottom and top faces} = \epsilon\frac{\partial E_z}{\partial z}\Delta x\,\Delta y\,\Delta z$$

Upon combining these, the outward flux through all faces is

$$\epsilon\left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}\right)\Delta x\,\Delta y\,\Delta z = \rho\,\Delta x\,\Delta y\,\Delta z \qquad (5.18)$$

by Gauss's law where ρ is the volume charge density. The above equality is abbreviated

$$\text{Divergence } \boldsymbol{E} = \nabla\cdot\boldsymbol{E} = \frac{\rho}{\epsilon} \qquad (5.19)$$

in which the element of volume has been cancelled and the term "divergence" has

been applied to the limiting value of the net outward flux per element of volume as the element of volume approaches zero.

But

$$\boldsymbol{E} = \text{negative gradient of } V \tag{5.20a}$$

frequently abbreviated

$$\boldsymbol{E} = -\nabla V \tag{5.20b}$$

or in component form

$$E_x = -\frac{\partial V}{\partial x}$$
$$E_y = -\frac{\partial V}{\partial y} \tag{5.21}$$
$$E_z = -\frac{\partial V^*}{\partial z}$$

Making these substitutions into Eq. (5.18),

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = -\frac{\rho}{\varepsilon} \tag{5.22}$$

which is Poisson's equation. This is abbreviated

$$\nabla^2 V = -\frac{\rho}{\varepsilon} \tag{5.23}$$

In a region free of charge, $\rho = 0$ so that

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0 \tag{5.24}$$

which is Laplace's equation. This is abbreviated

$$\nabla^2 V = 0 \tag{5.25a}$$

If the derivation is made in terms of general coordinates u_1, u_2, and u_3 with scale factors h_1, h_2, and h_3, respectively, so that an element of arc length is related to the coordinates and scale factors by

$$ds^2 = h_1^2\, du_1^2 + h_2^2\, du_2^2 + h_3^2\, du_3^2$$

then Laplace's equation assumes the general form

$$\nabla^2 V = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial u_1}\left(\frac{h_2 h_3}{h_1}\frac{\partial V}{\partial u_1}\right) + \frac{\partial}{\partial u_2}\left(\frac{h_1 h_3}{h_2}\frac{\partial V}{\partial u_2}\right) + \frac{\partial}{\partial u_3}\left(\frac{h_1 h_2}{h_3}\frac{\partial V}{\partial u_3}\right)\right] \tag{5.25b}$$

Interpretations of Laplace's Equation. As has been mentioned before and as is evident from the development of the equation, Laplace's equation is a differential expression of Gauss's law for an element of volume.

* See Appendix II.

In the language of differential equations it says that the net electric flux emerging from an element of volume in a region free of charge is zero.

Another interpretation that can be given to Laplace's equation in the two-dimensional case is that it is an equivalent way of saying that the potential at any point in a field is the average of the potentials at four equally spaced surrounding points. Thus if there is given a set of curves of equal potentials in the vicinity of some electrodes, known as a "contour representation of potentials," then the potential at any point, say the point (2,2), is the average of the potentials at the four surrounding coordinate points, for the case assumed, the average of the potentials at the points (2,1), (3,2), (2,3), and (1,2). This property will be proved in a subsequent section.

Laplace's equation can also be interpreted in terms of the curvature of the potential profiles of a field configuration. Two-dimensional fields can be represented either by contours of equipotential or by potential profiles just as we can draw either a contour map or a set of profiles for a topographic representation of terrain. In the profile representation we draw potential as an ordinate against distance along some line as abscissa. It will be remembered from elementary calculus that the curvature of any curve is given by

$$\frac{1}{R} = K = \frac{\dfrac{d^2y}{dx^2}}{\left[1 + \left(\dfrac{dy}{dx}\right)^2\right]^{3/2}} \tag{5.26}$$

from which it is seen that the *sign* of the curvature is determined by the sign of the second derivative in the numerator since the denominator is always positive. If we now examine Laplace's equation in two dimensions,

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial^2 y} = 0 \tag{5.27}$$

we see that the two terms may be interpreted as giving the sign of the curvature of the profiles in the x and y directions. By Eq. (5.27) the curvatures must be of opposite nature since the sum of the terms is zero; and hence if the profile in the x cut at some point in the field is concave upward, then the profile in the y cut at the same point must be concave downward.

Examination of a simple case will illustrate the property described above. In Fig. 5.6*a* is shown the contour representation for the case of a concentric line with a circular inner conductor and a rectangular outer conductor. The solid lines represent the equipotential lines or contours. In Fig. 5.6*b* is shown the potential profile along the line *ab*,

a finite potential along the axis of symmetry, as in electrostatic electron lenses where there is no conductor along the axis.

In order to apply the above solutions to definite problems it is necessary to evaluate the constants in such a way that the potential function fits the prescribed boundary conditions. If the constants can be selected so that the function fits all the boundaries (electrodes), then it will define the potential at all points in the field. The potential solutions frequently appear as a series summation of terms of the form indicated above.

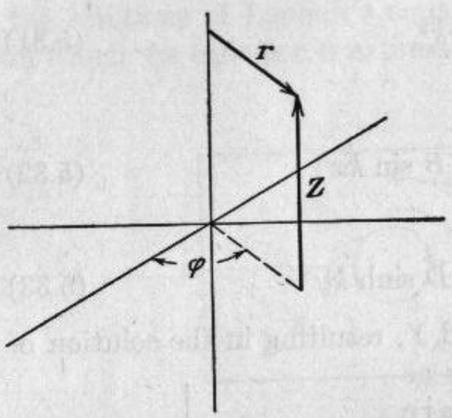


FIG. 5.8.—Cylindrical coordinate notation.

Difference Form of Laplace's Equation. We may write Laplace's equation in the form of a difference equation of which the differential equation is the limiting form. To do this we shall assume that the potential is known at a number of points whose spacing is finite though small. We shall assume that the points are at the intersections of a rectangular lattice as shown in Fig. 5.9 and that the spacing between the points is h,[1-3] The conclusions that we shall draw from the difference equation set up on this basis will apply also to the differential equation and its solution.

Consider the first derivative of potential at the point (0) in the xy plane. The difference operators corresponding to the partial derivatives are given by

$$\frac{\partial V}{\partial x} = \frac{1}{h}(V_1 - V_0) \tag{5.38}$$

$$\frac{\partial V}{\partial x} = \frac{1}{h}(V_0 - V_3) \tag{5.39}$$

$$\frac{\partial V}{\partial y} = \frac{1}{h}(V_2 - V_0) \tag{5.40}$$

$$\frac{\partial V}{\partial y} = \frac{1}{h}(V_0 - V_4) \tag{5.41}$$

[1] MORSE, P. M., and HERMAN FESHBACH, "Methods of Theoretical Physics," Massachusetts Institute of Technology, 1946, pp. 139–147.

[2] SHORTLEY, G. H., and R. WELLER, The Numerical Solution of LaPlace's Equation, *Jour. Appl. Phys.*, vol. 9, pp. 334–348, May, 1938. Probably the best single reference on this subject.

[3] FROCHT, M. M., and M. M. LEVIN, A Rational Approach to the Numerical Solution of LaPlace's Equation, *Jour. Appl. Phys.*, vol. 12, pp. 596–604, August, 1941.

The difference operators corresponding to the second derivatives are given by

$$\frac{\partial^2 V}{\partial x^2} = \frac{1}{h^2}[(V_1 - V_0) - (V_0 - V_3)] \tag{5.42}$$

$$\frac{\partial^2 V}{\partial y^2} = \frac{1}{h^2}[(V_2 - V_0) - (V_0 - V_4)] \tag{5.43}$$

Upon substituting these values in Laplace's equation there results

$$V_0 = \tfrac{1}{4}(V_1 + V_2 + V_3 + V_4) \tag{5.44}$$

which states that the potential at the center of a square is the average of the potentials at the corners of the square.

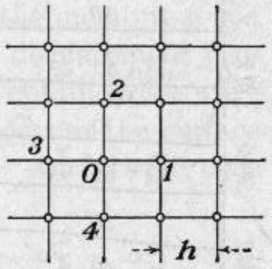


FIG. 5.9.—Arrangement of net points for the difference form of Laplace's equation in two-dimensional rectangular coordinates, Eq. (5.44).

It is possible to obtain numerical values of potential for various electrode configurations by means of Eq. (5.44). The procedure is to break up the field whose potential is desired into a suitable lattice, assume reasonable values of potential at each point in the lattice, and then apply Eq. (5.44) successively to each of the points, always using any new values of potential obtained. Successive applications of this procedure will correct any errors in the original assumptions, and the values of potential at any point will converge quite rapidly to the correct value. It is well to start with a coarse network and then make it finer.

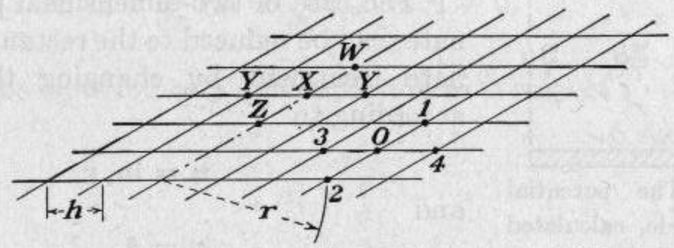


FIG. 5.10.—Arrangement of net points for the difference form of Laplace's equation in two-dimensional cylindrical coordinates, Eqs. (5.45) and (5.46).

The expression given in Eq. (5.44) was derived for two-dimensional rectangular coordinates. For two-dimensional problems of axial symmetry expressed in terms of the cylindrical coordinates of Fig. 5.8 that hold for electrostatic electron lenses, and the like; the corresponding

and it will be seen that the profile is curved *away from the axis* at the point *e*. In Fig. 5.6*c* is shown the potential profile along the line *cd*, and it will be seen that the profile is curved *toward the axis* at the point *e*.

Solutions of Laplace's Equation in Two Dimensions. The form which the solutions of Laplace's equation take depends upon the coordinates in which the equation is expressed. For *rectangular coordinates* Laplace's

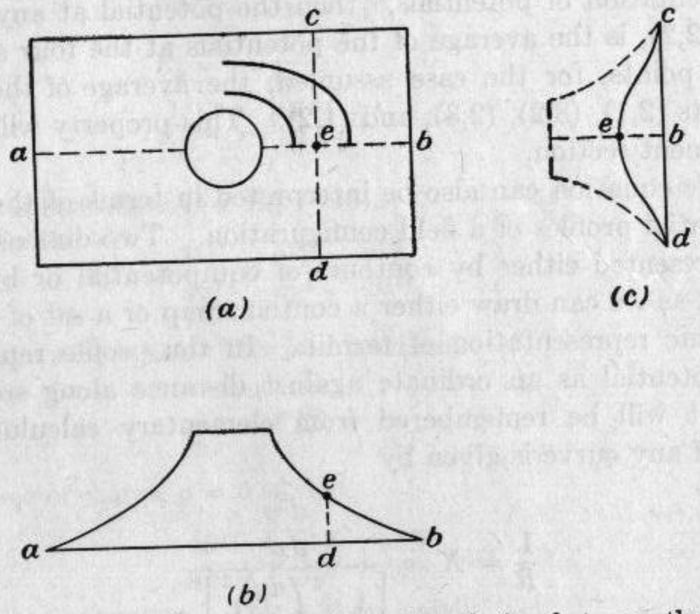


FIG. 5.6.—Example showing the relation between the curvatures of the profiles of a potential field.

equation has the form of Eq. (5.27). The solutions of this equation have the form

$$V = (A \cos kx + B \sin kx)(C \cosh ky + D \sinh ky) \quad (5.28a)$$

or

$$V = (A \cos ky + B \sin ky)(C \cosh kx + D \sinh kx) \quad (5.28b)$$

The above results are arrived at by assuming that V has a solution of the form XY where X is a function of x alone and Y is a function of y alone. If the product XY is substituted for V in Eq. (5.27), there results upon differentiation and rearrangement

$$\frac{1}{X}\frac{d^2X}{dx^2} = -\frac{1}{Y}\frac{d^2Y}{dy^2} \quad (5.29)$$

It is seen that the left-hand member is a function of x alone and that the right-hand member is a function of y alone. These can be equal only

if each equals the same constant. If this constant is taken as $-k^2$, then we may write two component equations in the place of Eq. (5.29),

$$\frac{1}{X}\frac{d^2X}{dx^2} = -k^2 \quad (5.30)$$

and

$$\frac{1}{Y}\frac{d^2Y}{dy^2} = +k^2 \quad (5.31)$$

The solution of Eq. (5.30) is

$$X = A \cos kx + B \sin kx \quad (5.32)$$

and the solution of Eq. (5.31) is

$$Y = C \cosh ky + D \sinh ky \quad (5.33)$$

Thus V is given by the product of X and Y, resulting in the solution of Eq. (5.28*a*) where multiple values of k as determined by imposed conditions are allowed. If the separation constant is chosen as $+k^2$ instead of $-k^2$, then the solution of Eq. (5.28*b*) results.

For the *polar coordinates* of Fig. 5.7 Laplace's equation has the form

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 V}{\partial \theta^2} = 0 \quad (5.34)$$

when the problem is one of axial symmetry. This has a solution in the form

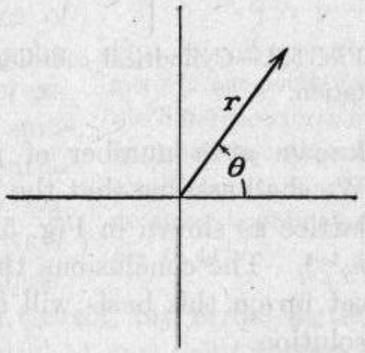


FIG. 5.7.—Polar coordinate notation.

$$V = (a \cos n\theta + b \sin n\theta)(cr^n + dr^{-n}) \quad (5.35)$$

as may be shown by the method demonstrated above using n^2 as the separation constant. When n equals zero the second factor in Eq. (5.35) is $c + d \ln r$.

For the *cylindrical coordinates* of Fig. 5.8 Laplace's equation, for cases of axial symmetry, has the form

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right) + \frac{\partial^2 V}{\partial z^2} = 0 \quad (5.36)$$

This has a solution of the form

$$V = [aJ_0(kr) + bN_0(kr)](c \sinh kz + d \cosh kz) \quad (5.37)$$

where J_0 and N_0 are the zero-order Bessel and Neumann functions. Since the Neumann function of zero is infinite, this term is not often encountered in electronics problems. Most potential configurations have

expression for the lattice of Fig. 5.10 is

$$V_0 = \frac{1}{4}(V_1 + V_2 + V_3 + V_4) + \frac{h}{8r}(V_4 - V_3) \qquad (5.45)$$

in which the points 1 and 2 are on a line parallel to the axis and point 3 is closer to the axis than point 4. The expression [Eq. (5.45)] works for all parts of the field except points on the axis for which

$$V_x = \tfrac{1}{6}(V_w + V_z + 4V_y) \qquad (5.46)$$

It will be noticed that Eq. (5.45) reduces to Eq. (5.44) for large values of r.

The above manipulations for the cylindrical case can be simplified by a change of variable. If as a new variable there be taken

$$y = r^{1/2}V \qquad (5.47)$$

then the Laplace equation reduces to

$$\frac{\partial^2 y}{\partial r^2} + \frac{\partial^2 y}{\partial z^2} + \frac{y}{4r^2} = 0 \qquad (5.48)$$

The corresponding net-point equation is

$$y_0 = \frac{y_1 + y_2 + y_3 + y_4}{\left(4 - \frac{h^2}{4r^2}\right)} \qquad (5.49)$$

which is much simpler to apply than Eq. (5.45).

The case of two-dimensional polar coordinates can be reduced to the rectangular coordinate treatment by changing the variables according to

$$u = \ln_e r \qquad (5.50)$$

and

$$v = \theta \qquad (5.51)$$

For a lattice of equal increments of u and v, Eq. (5.44) applies directly. The reasons for this will become evident when the transformation $W = \ln_e Z$ has been studied.

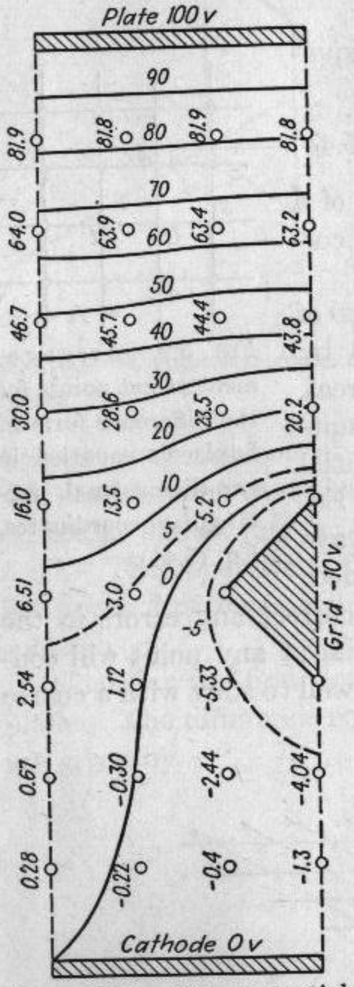


FIG. 5.11.—The potential field of a triode, calculated from the difference form of Laplace's equation, Eq. (5.44).

In Fig. 5.11 is shown the potential field inside a half section of a plane-electrode triode as calculated from repeated application of the difference form of Laplace's equation, Eq. (5.44).

5.7. Elastic-membrane Models of Potential. It is possible to represent two-dimensional potential problems having a *z-axis* symmetry by the elevation of a deformed elastic membrane. If an elastic membrane is uniformly stretched in all directions and leveled when suspended in a plane frame and is then deformed from its original plane by displacing the membrane distances proportional to electrode voltages with blocks shaped like the electrodes to be studied, then the displacement of the membrane at any point from the original plane is proprotional to the potential at that point in the field. In other words, the membrane is a topographic model of the potential field with vertical displacement proportional to potential. The deformed surface that is obtained is a very good representation of the potential field. This is because the surface will deform itself so that its area will be a minimum. Analytically this is expressed by making

$$S = \int\int \sqrt{1 + \left(\frac{dz}{dx}\right)^2 + \left(\frac{dz}{dy}\right)^2}\, dx\, dy = \min \qquad (5.52)$$

where z is the elevation and x and y are the coordinates in the horizontal plane. This is a problem in the calculus of variations that is converted into a problem in differential equations by applying to the equation for S that is in the form

$$S = \int\int F\left(x,y,z,\frac{dz}{dx},\frac{dz}{dy}\right) dx\, dy \qquad (5.53)$$

The Euler differential equation

$$\frac{\partial}{\partial x}F_{z^x} + \frac{\partial}{\partial y}F_{z^y} + F_z = 0 \qquad (5.54)$$

where the subscripts indicate differentiation with respect to the following factors:

$$z_x = \frac{\partial z}{\partial x}$$

$$z_y = \frac{\partial z}{\partial y}$$

$$F_{z_x} = \frac{\partial F}{\partial z_x}$$

$$F_{z_y} = \frac{\partial F}{\partial z_y}$$

$$F_z = \frac{\partial F}{\partial z}$$

Application of the Euler equation yields as the differential equation of the deformed surface[1-3]

$$\left[1+\left(\frac{\partial z}{\partial y}\right)^2\right]\frac{\partial^2 z}{\partial y^2}+\left[1+\left(\frac{\partial z}{\partial x}\right)^2\right]\frac{\partial^2 z}{\partial x^2}+2\,\frac{\partial^2 z}{\partial x\,\partial y}\,\frac{\partial z}{\partial x}\,\frac{\partial z}{\partial y}=0 \quad (5.55)$$

which reduces to Laplace's equation,

$$\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial y^2}=0 \quad (5.56)$$

for $\frac{\partial z}{\partial x} \ll 1$ and $\frac{\partial z}{\partial y} \ll 1$. If the angles of all lines on the surface are kept below 6 deg with the horizontal, the departure in deformation from that representing the true potential at any point will be less than 1 per cent. Some practical considerations are of importance. A No. 30 rubber surgical dam makes a good membrane. It should be stretched enough so that it will be tight and not sag and yet not be too close to the rubber's elastic limit. A linear stretch of about ⅓ works well. It helps in obtaining a uniform stretch to mark coordinate lines on the sheet before stretching and then stretch so that these are straight and of the proper spacing.

The applications of the elastic-membrane model of potential are somewhat limited, for it is accurate only for small deformations, it can be used to represent only two-dimensional problems with a z-axis (stacking) symmetry—it cannot exactly represent problems with a rotational symmetry about an axis—and it cannot be modified to include space-charge effects. In spite of these limitations, models of this sort have been used extensively by various laboratories in their studies of potential fields and electron paths; in the latter regard it yields much information in a short time. The use of the membrane in determining electron paths will be mentioned in a later section. Figure 5.12 shows the elastic membrane model used in the Electrical Engineering Department of Stanford University.

5.8. Current-flow Models of Potential. The laws which govern the flow of current in a uniformly conducting medium are the same as those which govern the "flow" of electrostatic-flux lines in a vacuum. This

[1] KLEYNEN, P. H. J. A., Motion of an Electron in Two Dimensional Electrostatic Field, *Philips Tech. Rev.*, vol. 2 (No. 11), pp. 338–345, 1937. Original article on this subject.

[2] STRUTT, M. J. O., "Moderne Mehrgitter-Elektronenroehren," pp. 3–6, Springer, Berlin, 1938.

[3] ZWORYKIN, V. K., and J. A. RAJCHMAN, The Electrostatic Electron Multiplier, *Proc. I.R.E.*, vol. 27, pp. 558–565, September, 1939.

makes it possible to set up current-flow models of electrode systems and to measure the potential at any point.

The equations for the components of current density in a continuous and uniform medium such as some electrolyte are given by Ohm's law in terms of the gradient of potential as

$$J_x = -g\,\frac{\partial V}{\partial x} \quad (5.57a)$$

$$J_y = -g\,\frac{\partial V}{\partial y} \quad (5.57b)$$

for the two-dimensional case, where J is current density and g is the specific conductivity of the medium.

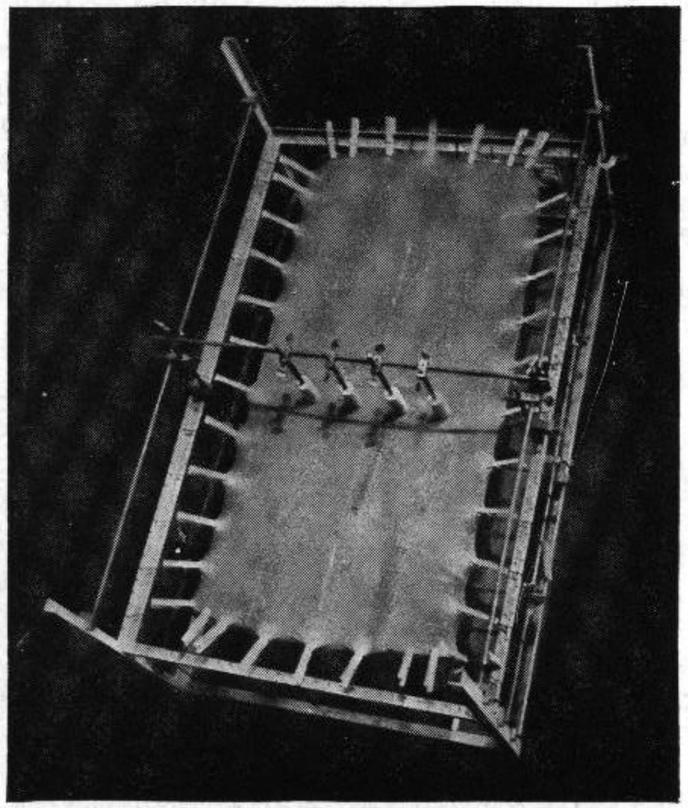


FIG. 5.12.—Elastic-membrane model of potential.

Since the flow of current in a medium of constant conductivity corresponds to an irrotational flow of an incompressible fluid, as much current will flow into any element of volume as flows out of it. This condition is sometimes expressed by saying that the divergence of the current is zero, which may be expressed mathematically as

$$\frac{\partial J_x}{\partial x}+\frac{\partial J_y}{\partial y}=0 \quad (5.58)$$

Substitution of the components of Eq. (5.57) into Eq. (5.58) yields Laplace's equation in the form

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0 \tag{5.59}$$

Upon comparing the above equations with those developed for the electrostatic field it is seen that an exact correspondence can be established. The relations may be tabulated in one-to-one correspondence as follows:

Currents		Electrostatic Fields
	Quantities	
J Current density		D Displacement flux
g Specific conductivity		ε_r Dielectric constant
V Potential		V Potential
	Relations	
$J_x = -g\frac{\partial V}{\partial x}, \quad J_y = -g\frac{\partial V}{\partial y}$		$D_x = -\varepsilon\frac{\partial V}{\partial x}, \quad D_y = -\varepsilon\frac{\partial V}{\partial y}$
$\frac{\partial J_x}{\partial x} + \frac{\partial J_y}{\partial y} = 0$		$\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} = 0$
$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0$		$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0$

From the above tabulation it is seen that the correspondence between the current flow and electrostatic field is quite complete. It is thus

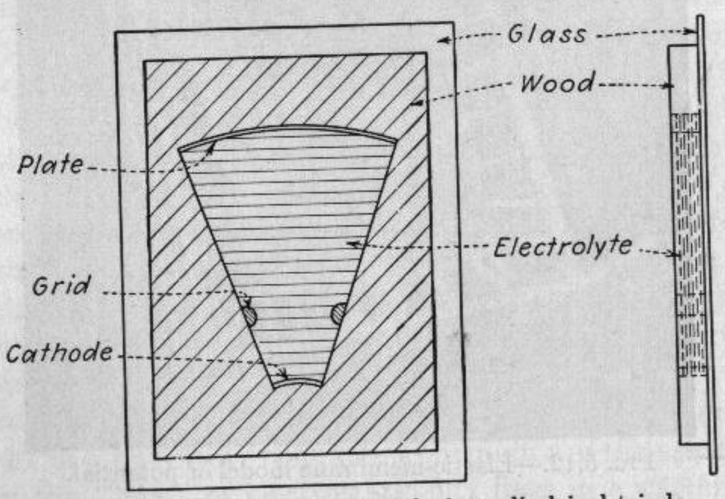


FIG. 5.13.—Current-flow model of a cylindrical triode.

necessary only to set up a current-flow model with electrodes geometrically similar to those of the electrostatic problem whose solution is desired and to measure the potential contours. The model is easily set up for two-dimensional problems by means of a flat tank. A weak solution of copper sulphate may be used as an electrolyte. This has a fairly good conductivity and has no polarizing action with copper electrodes. The

equipotential contours can be traced with a probe connected to two resistances forming two arms of a bridge. The other arms of the bridge are in the electrolyte. By setting the external resistances any equipotential contour can be traced by observing the points at which a null indication is received. Tanks may be made of wood cemented to a glass plate. A large sheet of coordinate paper may be put under the glass to identify the location of points and thus facilitate their transfer to another sheet for plotting. It is also possible to use a pantograph for plotting directly. In using a flat tank it is absolutely necessary that the liquid be of the same depth at all points. Placing the tank upon a board with leveling screws makes it easy to level. Figure 5.13 shows a tank of the type described. This particular tank represents a section of a cylindrical electrode triode.

The arrangement of resistors used with triode current-flow models is shown in Fig. 5.14. The resistors R_p and R_g are used to set the relative positive plate and negative grid potentials. The resistors R_1 and R_2 in the bridge arms are used to determine the potential of the contour to be traced. If the resistors R_2 and R_1 are set in the ratio of 2 to 8, the probe will trace out the contour having 80 per cent of the plate-cathode potential, since the percentage voltage of the contour is given by $\frac{R_2}{R_1 + R_2}$. Headphones are conveniently used as balance detectors. It is also possible to use a cathode-ray oscilloscope. If the probe and R_1, R_2 junction are connected to the vertical plates and a voltage in phase with the electrode potentials is connected to the horizontal plates, there will result a straight-line Lissajous figure whose slope will be zero when the probe is in the proper position. The advantage of this arrangement is that the slope of the line will be negative or positive according to whether the probe is to one side or the other of the proper position. A low frequency of the order of 50 to 100 cycles should be used. If the frequency is too low, it is difficult to detect a null. If it is too high, the distributed capacities affect the balance.

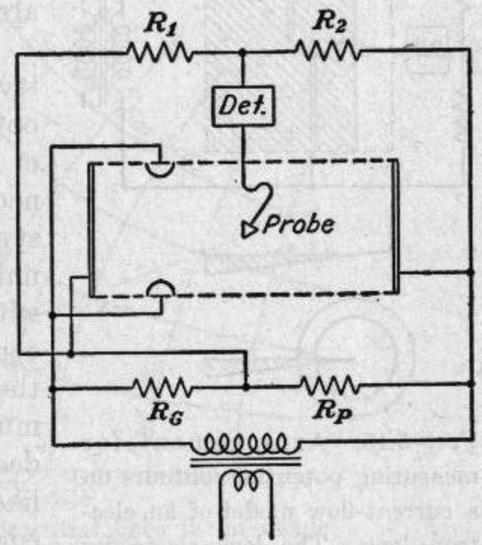


FIG. 5.14.—Circuit arrangement for measuring potential contours on a current-flow model of a plane-electrode triode section.

It should be observed in the model of Fig. 5.14 that the proper

conditions of symmetry are obtained along the nonconducting boundaries which are indicated by dotted lines. Here the current flow must be parallel to the boundary, which ensures that the potential contours are at right angles to the nonconducting boundaries, since the equipotentials are perpendicular to the flow lines. It will also be true that the equipotentials will be perpendicular to all lines of symmetry running in the direction of the flow. Flow lines will be perpendicular to conducting surfaces. The flow will also be parallel to the top and bottom of the liquid layer since the air above and the glass below are nonconducting.

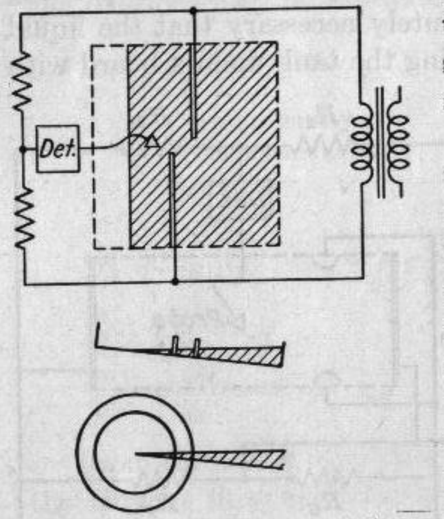


FIG. 5.15.—Arrangement for measuring potential contours on a current-flow model of an electron lens. The lens electrodes are cylinders of revolution that require a tilted tank to represent a wedge-shaped portion of the potential field. The edge of the wedge-shaped portion of electrolyte corresponds to the axis of the electrodes.

For problems involving axial symmetry such as are encountered in electron optics, a slightly different arrangement of electrodes must be used. Here it is necessary to reproduce conditions of axial symmetry and it is not correct to use a uniform depth of electrolyte as in Fig. 5.14 without special electrodes. To obtain correct results, either the electrodes or the volumetric shape of the electrolyte must be changed. It is possible to use a deep flat tank if the electrodes are shaped like portions of half cylinders with their edges at the surface of the electrolyte. For such an arrangement the probe should be kept at the surface of the electrolyte. A more convenient arrangement is to use a wedge-shaped electrolyte. Use of such a section corresponds to a pie-shaped section of small angle cut out of the field of revolution as shown in Fig. 5.15. The wedge-shaped volume of electrolyte is obtained by simply tilting a flat tank. Properly speaking the electrodes should be portions of cylinders, but if the angle of the electrolyte wedge is small enough, say less than 5 deg, they may be portions of planes without introducing any appreciable error.

5.9. Sketching of Flux and Potential Fields. The properties of electrostatic fields are such that, with a little practice, it is possible to sketch fields with considerable accuracy without recourse to mathematical methods. It is known, for instance, that flux and potential lines are everywhere at right angles to each other. Flux lines emerge at right

angles from conducting surfaces. Potential lines near conductors tend to have the same shape as the conductor. These and other useful properties may be summarized as follows:

PROPERTIES USEFUL IN SKETCHING FIELDS

1. Flux and potential lines form orthogonal families of curves.
2. Flux lines are perpendicular to conductors at conductor surface.
3. Potential contours close to conductors tend to have the same form as the conductors.

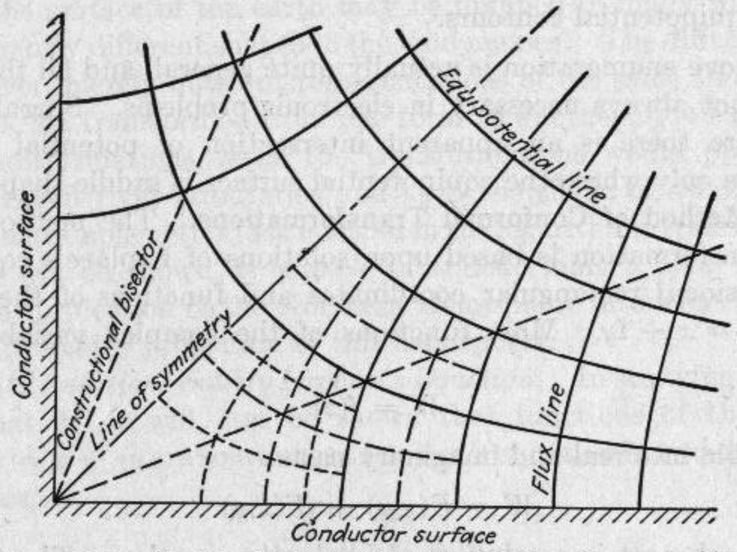


FIG. 5.16.—Sketch of flux and potential lines in an inside right-angled corner. This sketch was made by the method of Sec. 5.9 without mechanical or numerical aids.

4. Potential lines are parallel or perpendicular to lines of symmetry; constructional bisectors may exist.
5. Flux-potential patterns should be drawn with curvilinear squares, *i.e.*, a four-sided figure, with right angles at the corners and with equal average lengths of opposite sides, which maintains these properties upon infinite subdivision.

a. Same potential difference exists across each square.
b. Same flux passes through each square.
c. Each square has the same attraction for the conductor face.
d. Each square has the same energy storage.
e. Each flux line represents the same increment of capacity between electrodes.

Most of the above properties are self-evident. In Fig. 5.16 is shown a plot of the flux and potential inside of a right-angled corner. This plot was sketched, not calculated. It will be observed that all the curvi-

linear squares upon infinite subdivision will still be curvilinear squares. The principal line of symmetry is shown by the center line. The constructional bisectors are shown by dotted lines.

A flux plot to be of value should include

1. Shape of fields at large distances as well as small distances from the charges (conductors).
2. Location of all conductors and charges.
3. Geometrical symmetries of any kind.
4. All singular points, *i.e.*, "saddle" points, giving rise to a crossing of equipotential contours.

The above enumeration is actually quite general, and all these inclusions are not always necessary in electronic problems. Singular points occur where there is an apparent intersection of potential contours. This occurs only where the equipotential surface is saddle-shaped.[1,2]

5.10. Method of Conformal Transformations. The method of conformal transformation is based upon solutions of Laplace's equation in two-dimensional rectangular coordinates and functions of the complex variable $z = x + iy$. Most functions of the complex variable of the form

$$W = f(z) \tag{5.60}$$

are separable into real and imaginary parts

$$W = E(x,y) + iF(x,y) \tag{5.61}$$

in which each part is a solution of Laplace's equation. The two parts of the complex function, $E(x,y)$ and $F(x,y)$, further represent orthogonal families of curves. They may hence be taken as representing flux and equipotential lines. The functions having the above properties are known as *analytic* functions (to be defined more explicitly).

Every analytic function of the complex variable may thus be considered to represent the flux and potential field of some set of electrodes. Fields may further be transformed by means of analytic functions from one form to another. Thus, given a function that gives the field corresponding to one set of electrodes, the application of another function will transform this field into one corresponding to another set of electrodes. In the course of this transformation all the properties of flux and potential fields are preserved.

Analytic functions when used for making transformations have the property of preserving the angles between lines and of making corresponding incremental areas similar in shape. It is for these reasons that the transformations are called "conformal."

[1] Moore, A. D., Mapping of Magnetic and Electric Fields, *Elec. Jour.*, vol. 23, pp. 355–362, July, 1926.

[2] Stevenson, A. R., and R. H. Park, Graphical Determination of Magnetic Fields, *Trans. A.I.E.E.*, vol. 46, pp. 112–135, February, 1927.

Application of the method of conformal transformations usually takes the form of finding a transformation which converts the electrodes (equipotentials) of some simple field to the structure of which the field is desired. The field of the simple electrodes can usually be determined, and then the transformation converts the entire field to that of the more complex arrangement.

Conformal transformations are familiar to everyone in the form of maps. The surface of the earth may be mapped in many ways, which give apparently different shapes to the land masses. The different shapes are, however, merely different representations of the same thing. Most maps could be transformed from one form to another by means of conformal transformations, since the transformations would preserve the angles between river tributaries and keep the shape of small areas the same. An example of this idea is found in the logarithmic transformation, which, as will be shown, is capable of transforming a polar azimuthal equidistant projection of the Northern Hemisphere into what is approximately a Mercator projection of this hemisphere.

Complex Functions Satisfy Laplace's Equation. In studying conformal transformations it will first be shown that functions of the complex variable $z = x + iy$ are solutions of Laplace's equation in two-dimensional coordinates,

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0 \tag{5.62}$$

where

$$f(x,y) = f(x + iy) = f(z) \tag{5.63}$$

This follows since

$$\frac{\partial f}{\partial x} = \frac{df}{dz}\frac{\partial z}{\partial x} = \frac{df}{dz} \tag{5.64}$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{d^2 f}{dz^2} \tag{5.65}$$

Similarly

$$\frac{\partial f}{\partial y} = \frac{df}{dz}\frac{\partial z}{\partial y} = i\frac{df}{dz} \tag{5.66}$$

$$\frac{\partial^2 f}{\partial y^2} = -\frac{d^2 f}{dz^2} \tag{5.67}$$

It is evident that these partial derivatives are such as to satisfy Laplace's equation in the form of Eq. (5.62). The converse of this property is also true, *viz.*, that solutions of Laplace's equation in two-dimensional rectangular coordinates are expressible as functions of the complex variable.

Example: Let

$$f(z) = Az^2 + Bz$$
$$= Ax^2 + 2Aixy - Ay^2 + Bx + iBy$$

Then

$$\frac{\partial f}{\partial x} = 2Ax + 2Aiy + 0 + B + 0$$
$$\frac{\partial^2 f}{\partial x^2} = 2A$$

And

$$\frac{\partial f}{\partial y} = 0 + 2Aix - 2Ay + 0 + iB$$
$$\frac{\partial^2 f}{\partial y^2} = -2A$$

Laplace's equation is seen to be satisfied.

It is also true that the real and imaginary parts of the function are solutions of Laplace's equation.

$$f(z) = E(x,y) + iF(x,y)$$

where

$$E(x,y) = Ax^2 - Ay^2 + Bx$$

and

$$F(x,y) = 2Axy + By$$

It is evident that

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} = 2A - 2A = 0$$

and that

$$\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} = 0 - 0 = 0$$

Definition of Analytic Functions. The properties of functions of the complex variable will now be considered. It was mentioned above that a large group of functions had the desired properties, and such functions were referred to as "analytic functions." It will be remembered that, in the study of functions of the real variable, attention is usually restricted to functions which are continuous and functions of which the derivative at any point is independent of the direction in which we approach the point as we take the limit of the ratio of the increment of the function to the increment of the variable. Similarly in studying functions of the complex variable we shall restrict attention to functions having a derivative that is independent of the direction of approach to the point in question. This is necessary because only functions having this property also have the desired properties of potential functions. Mathematicians use the term *analytic* to describe such functions.

Consider

$$W = f(z) = f(x + iy) \tag{5.68}$$

Such a function is said to be *analytic* if it has a derivative that is independent of the direction of the increment of the variable Δz as it approaches zero.

For a single real variable

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} \tag{5.69}$$

For the complex variable $z = x + iy$

$$f'(z) = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} = \lim_{\Delta z \to 0} \frac{\Delta f}{\Delta z} \tag{5.70}$$

Let

$$W = u + iv = f(x + iy)$$

If the function is analytic,

$$\lim_{\substack{\Delta z \to 0 \\ \Delta z = \Delta x + i\,\Delta y}} \frac{\Delta W}{\Delta z} = \lim_{\substack{\Delta x \to 0 \\ \Delta z = \Delta x + i0}} \frac{\Delta W}{\Delta x} = \lim_{\substack{\Delta y \to 0 \\ \Delta z = 0 + i\,\Delta y}} \frac{\Delta W}{i\,\Delta y} \tag{5.71}$$

In derivative form

$$\frac{dW}{dz} = \frac{\partial W}{\partial x} = -i\frac{\partial W}{\partial y} \tag{5.72}$$

But

$$W = u + iv$$

Therefore

$$\frac{\partial W}{\partial x} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} \tag{5.73}$$

and

$$-i\frac{\partial W}{\partial y} = -i\left(\frac{\partial u}{\partial y} + i\frac{\partial v}{\partial y}\right) \tag{5.74}$$

$$-i\frac{\partial W}{\partial y} = \frac{\partial v}{\partial y} - i\frac{\partial u}{\partial y} \tag{5.75}$$

Hence

$$\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i\frac{\partial u}{\partial y} \tag{5.76}$$

Equating real and imaginary parts,

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \tag{5.77}$$

$$\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y} \tag{5.78}$$

These equations are known as the "Cauchy-Riemann conditions" and serve to identify analytic functions.

Dividing Eq. (5.77) By Eq. (5.78),

$$\frac{\frac{\partial u}{\partial x}}{\frac{\partial u}{\partial y}} = -\frac{\frac{\partial v}{\partial y}}{\frac{\partial v}{\partial x}} \tag{5.79}$$

which is the orthogonality condition for two functions since the derivative of a function of x and y is

$$\frac{dy}{dx} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}} \tag{5.80}$$

and curves are perpendicular if the derivative of one curve is the negative reciprocal of the derivative of the other. If we take derivatives of the Cauchy-Riemann equations with respect to x and y, respectively, then

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 v}{\partial x\,\partial y} \tag{5.81}$$

$$\frac{\partial^2 v}{\partial x\,\partial y} = -\frac{\partial^2 u}{\partial y^2} \tag{5.82}$$

Subtracting these gives

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{5.83}$$

or Laplace's equation holds for the real part of the function. Similarly, Laplace's equation holds for the imaginary part.

To summarize, an analytic function is one whose derivative is independent of the direction of the increment of the variable as the increment approaches zero. For such a function the Cauchy-Riemann conditions hold. Analytic functions have real and imaginary parts which are orthogonal to each other and each one of which is a solution of Laplace's equation.

It will be recognized that functions may be analytic except at certain points just as functions of a real variable may be continuous except at certain points. Such points are frequently those at which the function has a pole, *i.e.*, assumes an infinite value. It is possible to use such functions if the regions in which the function is not analytic are excluded from consideration.

A serious limitation of the method of conformal transformations is that it is not always possible to find the transformation which will convert one set of electrodes to another. In general, there is no definite method by which the transformation which fits a set of electrodes can be found. An exception to this remark is the Schwartz-Christoffel transformation, which transforms the real axis in the W plane into any polygon in the Z plane, but this transformation does not find much use in the field of electronics. However, it is also possible to use methods of successive approximations and series expansions. *Fortunately, the transformations necessary for the most important vacuum-tube problems are known.*

The Logarithmic Transformation. The transformation that solves the problem of determining fields in the plane-electrode triode is known. It is the logarithmic transformation

$$W = \ln_\epsilon Z \tag{5.84}$$

where $\epsilon = 2.718$ is the Napierian base.
This is analytic for all finite values of x and y other than zero.

The nature of the logarithmic transformation can best be understood by studying its component relations. It is most convenient to use polar coordinates in the Z plane and rectangular coordinates in the W plane. Thus let

$$Z = re^{i\theta} \tag{5.85}$$

and

$$W = u + iv \tag{5.86}$$

In these coordinates

$$u + iv = \ln_\epsilon r + i\theta \tag{5.87}$$

so that the component equations relating the real and imaginary parts are

$$u = \ln_\epsilon r \tag{5.88}$$

$$v = \theta \tag{5.89}$$

or, solved for r and θ,

$$r = \epsilon^u \tag{5.90}$$

$$\theta = v \tag{5.91}$$

This function is readily proved to be analytic for finite values of the argument by application of the Cauchy-Riemann equations.

Examination of the v component of W shows that it is multiple-valued, in fact infinitely so. This occurs because any angle in the Z plane can be written as an angle less than 2π plus any integral multiple of 2π. The angle θ can be written as $\theta + 2\pi n$, where n is any positive or negative integer. Thus, corresponding to any point in the Z plane there are an infinite number of points in the W plane evenly spaced by a distance 2π along a vertical line.

From Eq. (5.88) it is seen that any circle about the origin in the Z plane, $r = k$, transforms into a line parallel to the v axis in the W plane, $u = \ln k$.* Circles with radii less than 1 give lines in the left half of the

* Hereafter, the notation ln r will be used to designate the natural logarithm of r.

W plane since the logarithms of numbers less than 1 are negative, and circles with radii greater than 1 give lines in the right half of the W plane since the logarithms of numbers greater than 1 are positive. Any radial line through the origin, $\theta = k$, transforms into a set of lines in the W plane parallel to the u axis and spaced a distance 2π, $v + 2\pi n = k$. These relations are shown in Fig. 5.17. From this it is seen that a single point in the Z plane such as $r = 1.5$, $\theta = \frac{\pi}{4}$ transforms into a series of points $u = \ln 1.5$, $v = \frac{\pi}{4} + 2n\pi$ in the W plane. Thus a single point in the Z plane that may be taken as representing a line charge transforms

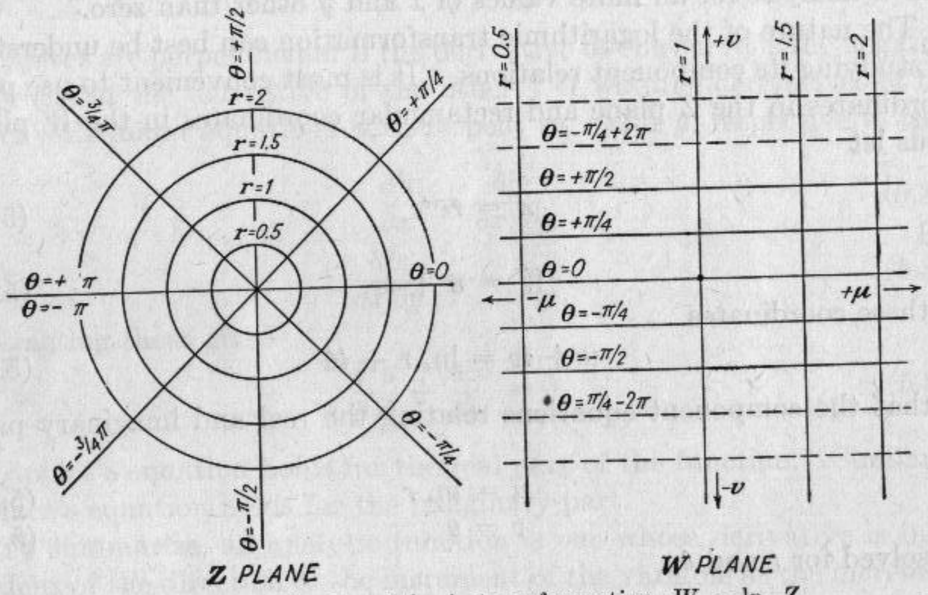


FIG. 5.17.—The logarithmic transformation, $W = \ln_\epsilon Z$.

into a row of line charges evenly spaced in the W plane. This gives the arrangement corresponding to a grid of evenly spaced parallel wires and is the basis for the plane-electrode representation.

The nature of the logarithmic transformation is better understood if the transformation be effected in a series of steps. Imagine the Z-coordinate plane to be a stretched elastic membrane. If the polar-coordinate diagram of the Z plane shown in Fig. 5.18*a* be split along the negative x axis and the upper and lower edges be rotated clockwise and counterclockwise, respectively, the pie-shaped section of Fig. 5.18*b* will result. If now the point on the pie is stretched to the left and the outer edge is compressed, the configuration shown in Fig. 5.18*c* results. Finally the left and right edges are made the same length and are stretched to negative and positive infinity, respectively, to give a strip of the W plane as shown in Fig. 5.18*d*.

The nature of the logarithmic transformation is also well illustrated by relationship between certain types of maps. Thus, if the Northern Hemisphere on a polar azimuthal equidistant projection be taken as the Z plane, then the Northern Hemisphere on a Mercator projection corresponds very closely to the $W = \ln_\epsilon Z$ plane. It will be recognized that each of these two common maps is but a different representation of a part of the earth's surface. In Fig. 5.19 is shown a polar azimuthal projection of the Northern Hemisphere. In Fig. 5.20 is shown a Mercator projection of the Northern Hemisphere.

The polar azimuthal equidistant projection is made by unfolding the earth's surface and stretching it out until it is a plane tangent to the

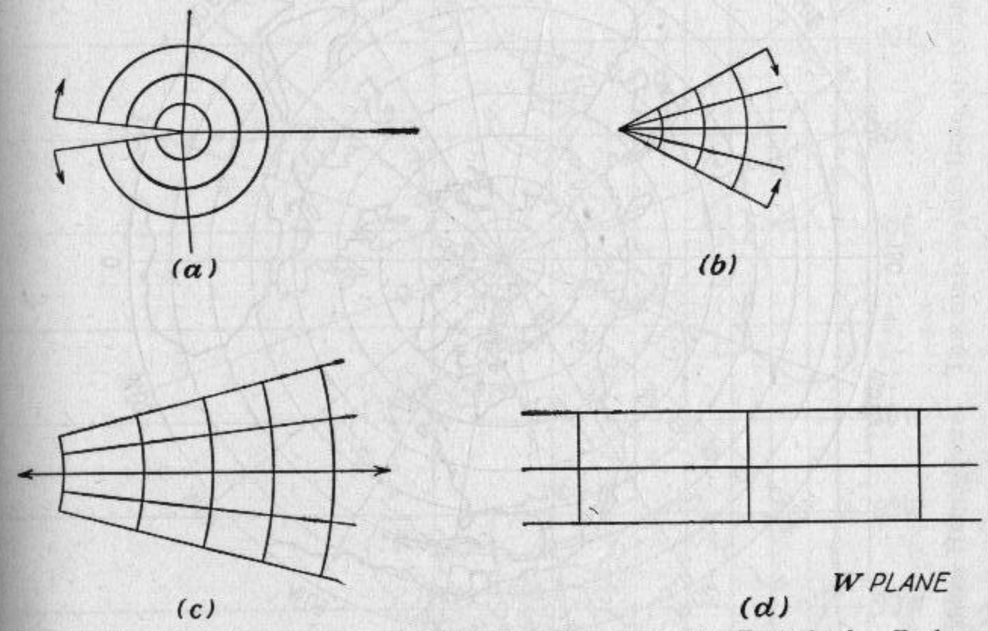


FIG. 5.18.—Steps in a progressive transformation from the Z to the $\ln_\epsilon Z$ plane.

pole with distances from the pole made equal to the great-circle distances on the actual sphere. This is indicated in Fig. 5.21. The longitude lines become straight lines through the pole, and the latitude circles remain circles.

Mercator's projection is approximated by surrounding the earth with a circular cylinder tangent to the earth at the equator as in Fig. 5.22 and extending to infinity in both directions. Points on the earth's surface are then projected onto this cylinder by drawing a line from the earth's center through the point in question and extending it until it hits the cylinder. The cylinder is then cut and unfolded to give a plane surface. The latitude circles on the sphere become a series of parallel straight lines on the Mercator projection. The longitude circles become another

set of equidistant parallel straight lines perpendicular to the latitude lines.

It is easily seen that the latitude circles, $r = k$, in the polar azimuthal equidistant projection become straight lines parallel to the equator, $u = \ln k$, in the Mercator projection. The longitude lines through the pole in the polar azimuthal equidistant projection, $\theta = K$, become a set of evenly spaced lines perpendicular to the equator in the Mercator

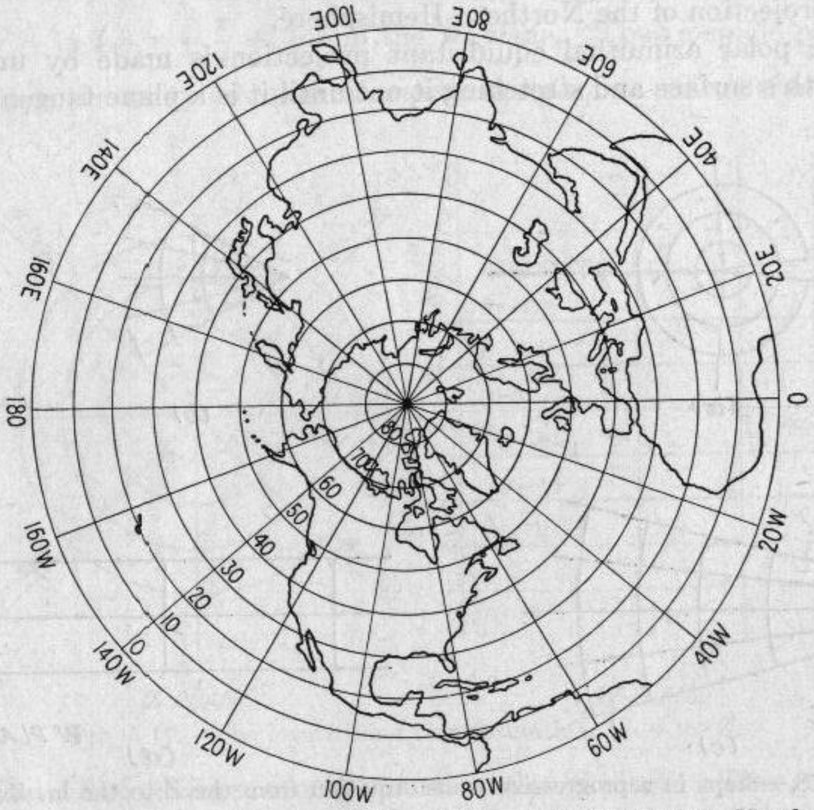
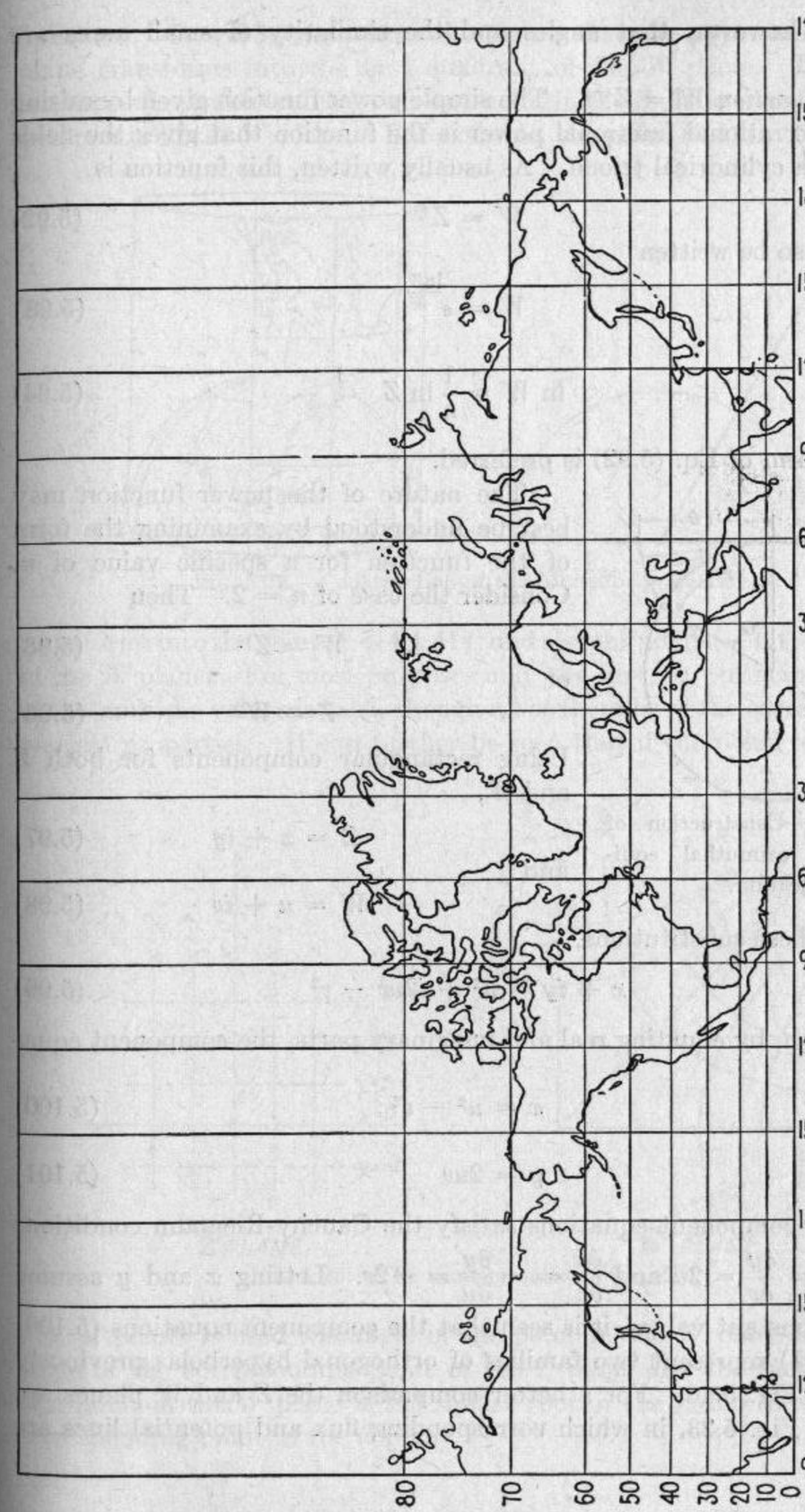


FIG. 5.19.—The polar azimuthal equidistant projection of the Northern Hemisphere. This may be considered as a polar-coordinate representation of the Z plane.

projection, $v = K$. The pole, which is the center of things in the polar azimuthal equidistant projection, recedes to infinity in the Mercator projection. Distortions in the different representations are evident. The polar azimuthal equidistant projection gives its most accurate representation near the pole but stretches out the equator disproportionately, causing Africa to be too wide. The Mercator projection is most accurate in the band around the equator but causes areas near the poles to be disproportionately large. Greenland on a Mercator projection looks larger than South America but is actually only one-tenth as large.



This map results from applying the transformation $W = \ln_e Z$ to the

FIG. 5.20.—The Mercator projection of the Northern Hemisphere. This is the polar azimuthal equidistant projection of Fig. 5.19.

Observe, however, that angles and the similarity of small areas are preserved.

The Function $W = Z^{1/n}$. The simple power function given by raising Z to some rational fractional power is the function that gives the fields inside of a cylindrical triode. As usually written, this function is

$$W = Z^{1/n} \tag{5.92}$$

It may also be written

$$W = \epsilon^{\frac{\ln Z}{n}} \tag{5.93}$$

or

$$\ln W = \frac{1}{n} \ln Z \tag{5.94}$$

but the form of Eq. (5.92) is preferred.

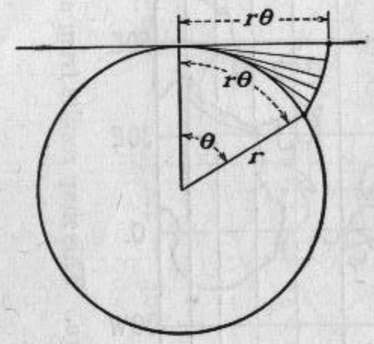


FIG. 5.21.—Construction of the polar azimuthal equidistant projection.

The nature of the power function may best be understood by examining the form of the function for a specific value of n. Consider the case of $n = 2$. Then

$$W = Z^{1/2} \tag{5.95}$$

or

$$Z = W^2 \tag{5.96}$$

Using rectangular components for both Z and W,

$$Z = x + iy \tag{5.97}$$

and

$$W = u + iv \tag{5.98}$$

Making these substitutions,

$$x + iy = u^2 + i2uv - v^2 \tag{5.99}$$

from which, by equating real and imaginary parts, the component equations are

$$x = u^2 - v^2 \tag{5.100}$$

and

$$y = 2uv \tag{5.101}$$

These component equations satisfy the Cauchy-Riemann conditions since $\frac{\partial x}{\partial u} = \frac{\partial y}{\partial v} = 2u$ and $\frac{\partial x}{\partial v} = -\frac{\partial y}{\partial u} = -2v$. Letting x and y assume various constant values, it is seen that the component equations (5.100) and (5.101) represent two families of orthogonal hyperbolas previously shown in Fig. 5.16. For a better comparison the Z and W planes are shown in Fig. 5.23, in which corresponding flux and potential lines are

indicated. It will be seen from this figure that the upper half of the Z plane transforms into the first quadrant of the W plane. The transformation is double-valued, *i.e.*, one point in the Z plane transforms into two points in the W plane. For example, the point (0,4) in the Z plane

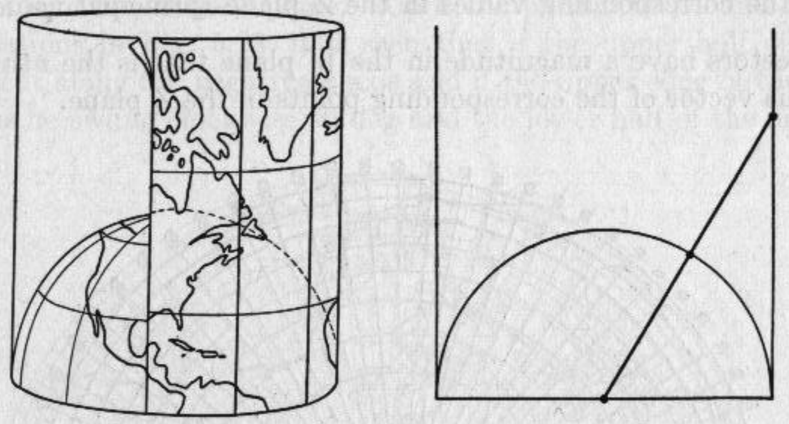


FIG. 5.22.—Construction of the Mercator projection.

transforms into the point (1.414,1.414) and also the point (−1.414, −1.414) in the W plane. For most purposes only the first, or "principal," value of the multiple values is used, though all of them have the correct mathematical properties. It can further be seen that if the polar representation of points is used the angle of the point in the W plane is half the angle of the corresponding point in the Z plane and the radius vector of a point in the W plane is the square root of the radius vector of the corresponding point in the Z plane.

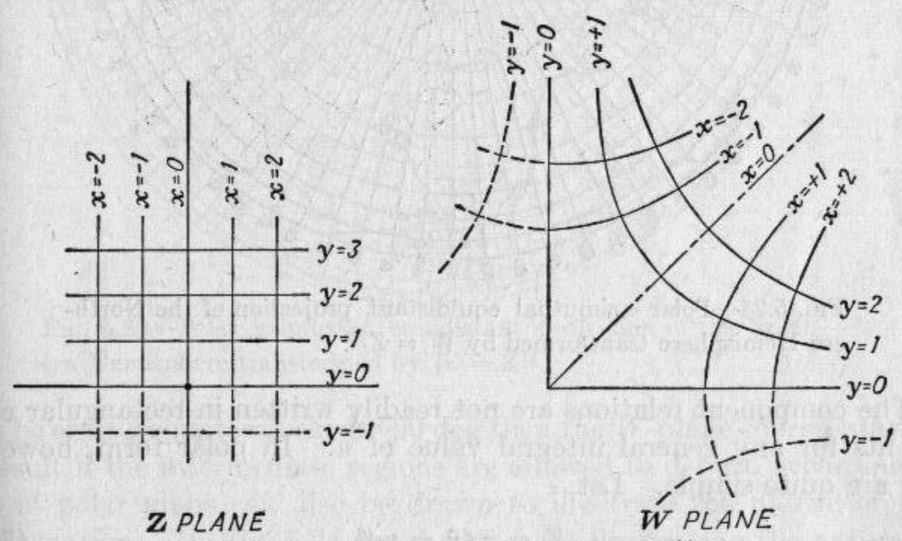


FIG. 5.23.—The transformation $W = Z^{1/2}$.

In the general case of the function $W = Z^{1/n}$, the function is n-valued if n is an integer. As a result, the upper half of the Z plane transforms into a segment of the W plane having an angle $\frac{\pi}{n}$. Angles in the W plane are $\frac{1}{n}$th the corresponding values in the Z plane (principal values), and radius vectors have a magnitude in the W plane that is the nth root of the radius vector of the corresponding points in the Z plane.

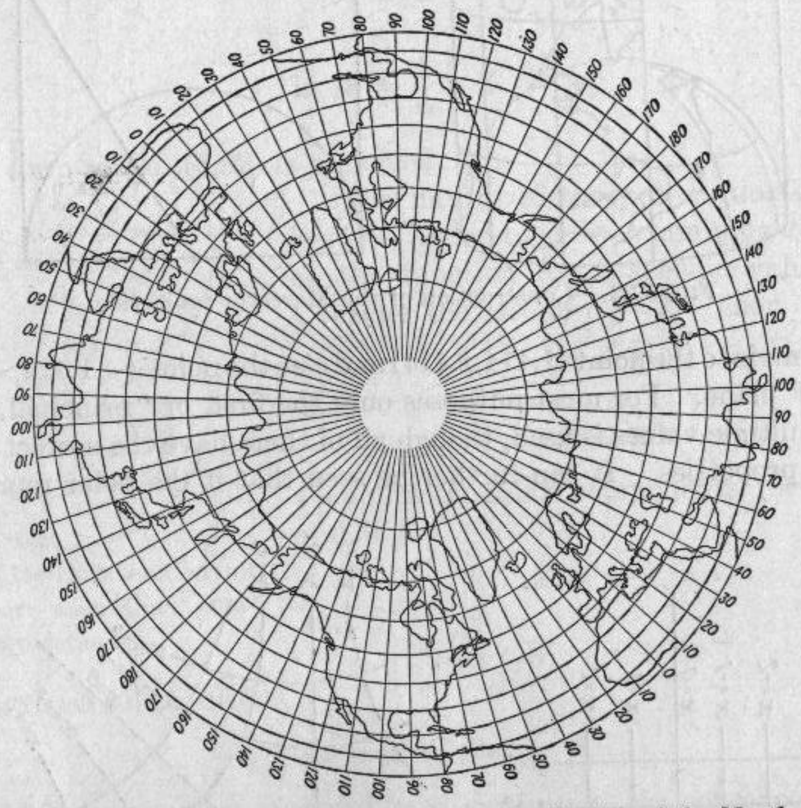


FIG. 5.24—Polar azimuthal equidistant projection of the Northern Hemisphere transformed by $W = Z^{1/2}$.

The component relations are not readily written in rectangular components for any general integral value of n. In polar form, however, they are quite simple. Let

$$Z = r\angle\theta = re^{i\theta} \tag{5.85}$$

as before; and let

$$W = R\angle\phi = Re^{i\phi} \tag{5.102}$$

Then the component equations in polar form are

$$R = r^{1/n} \tag{5.103}$$

and

$$\phi = \frac{\theta}{n} \tag{5.104}$$

The nature of the transformation $W = Z^{1/n}$ may be indicated as a kind of deformation of the Z plane. Upon comparing the W- and Z-plane representations in Fig. 5.23, it is seen that if the upper half of the Z plane be cut along the negative x axis and if the upper edge of the negative x axis be swung clockwise 90 deg and the lower half of the negative

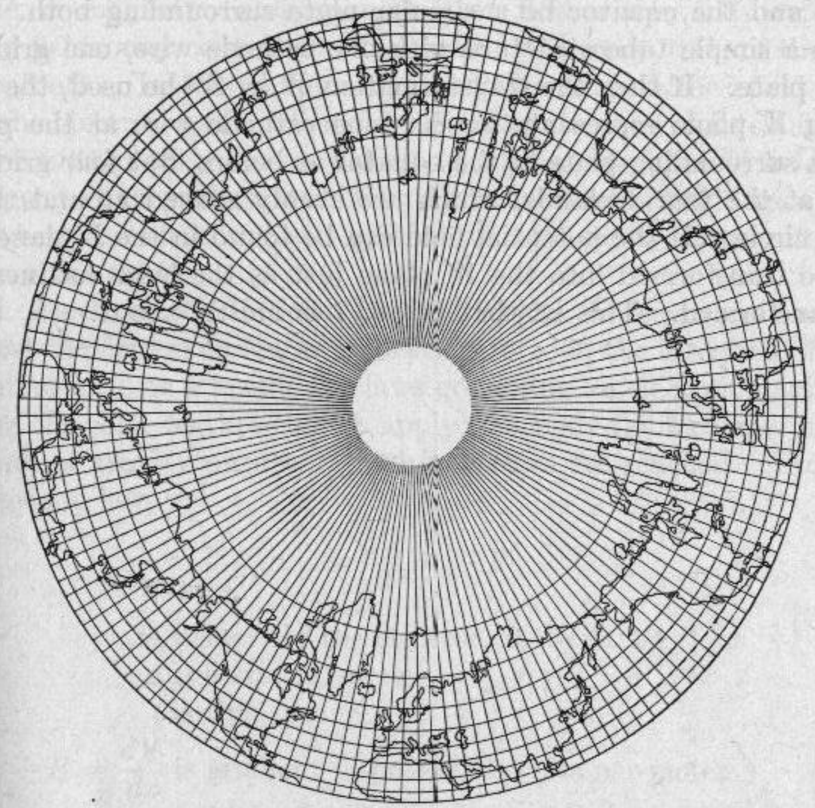


FIG. 5.25—Polar azimuthal equidistant projection of the Northern Hemisphere transformed by $W = Z^{1/4}$.

x axis be swung counterclockwise 90 deg then the W-plane representation will result if the intermediate regions are allowed to deform accordingly. A set of polar maps can also be drawn to illustrate the nature of the transformation. In Fig. 5.24 are shown maps illustrating the nature of the transformation $W = Z^{1/2}$. It is seen that the representation is double-valued and that the scale of distances from the pole in the W-plane representation is quadratic rather than linear; the land areas are pushed out from the pole toward the equator though the map as a whole differs surprisingly little from the usual representation.

In Fig. 5.25 are shown maps illustrating the nature of the transformation $W = Z^{1/4}$. This transformation is quadruple-valued, *i.e.*, every point in the Northern Hemisphere is repeated four times in the W-plane representation. The scale of distances from the pole is quartic in the W plane, with the result that the land masses are compressed strongly near the equator.

An inkling of how this transformation is used is obtained if we consider that in the polar azimuthal equidistant or Z-plane representation a cathode wire be located at the North Pole, a grid wire be located at Iceland, and the equator be a circular plate surrounding both. Then we have a simple tube structure with one cathode wire, one grid wire, and one plate. If then the transformation $W = Z^{1/4}$ be used, the corresponding W-plane representation has one cathode wire at the pole as before, a surrounding plate at the equator as before, but *four* grid wires located at the four Icelands, which are evenly spaced around the 66° latitude circle. If the potential field can be found in the Z plane, then it can be transformed into the W plane just as the land outlines have been transformed. This is what Chap. 7 is mostly about.

CHAPTER 6

LAWS OF ELECTRON MOTION

All electronic devices depend for their action upon the effect of applied electric or magnetic fields upon electron flow within the device. The applied fields may control the direction or the magnitude of the current flow or both. In this chapter there will be studied the effect of fields upon the electron paths when the electrons are present in small enough number so that their presence does not change the applied fields. In a subsequent chapter there will be studied the effect of fields upon electron flow when the electrons are present in sufficiently large numbers to influence the fields.

6.1. Electron in a Uniform Electric Field. An electron in a uniform electrostatic field experiences a constant force in the direction of increasing potential. As a result, the laws governing an electron starting from rest are the same as those which apply to a body falling freely under the influence of gravity until very high velocities are reached. From Newton's second law,

$$m \frac{d^2x}{dt^2} = -Ee \tag{6.1}$$

where m is mass of the electron, 9.107×10^{-31} kg
x is distance, meters
t is time, sec
$-E = \dfrac{dV}{dx}$ is gradient of potential, volts per meter
e is *magnitude* of the charge of the electron, 1.602×10^{-19} coulomb

A first integration of Eq. (6.1) gives

$$v = \frac{dx}{dt} = -\frac{e}{m} Et \quad \text{meters per sec} \tag{6.2}$$

the constant being zero because the velocity is taken as zero when time is zero. A second integration gives

$$x = -\frac{1}{2} \frac{e}{m} Et^2 \quad \text{meters} \tag{6.3}$$

in which the constant is again zero for an electron starting from rest at a point of zero potential.

If time is eliminated between Eqs. (6.2) and (6.3), there results

$$\tfrac{1}{2}mv^2 = eV \tag{6.4a}$$

where $V = -Ex$ is the potential through which the electron has fallen. Equation (6.4a) states that the kinetic energy acquired by an electron starting from rest is equal to the potential energy which it has lost. Solving for v,

$$v = \sqrt{\frac{-2Exe}{m}} = \sqrt{\frac{2eV}{m}} \qquad \text{meters per sec} \tag{6.4b}$$

The relation between Eqs. (6.2), (6.3), and (6.4a) and the corresponding equations for a falling body is evident. It is seen that the quantity $-\frac{eE}{m}$ is the equivalent of the gravitational constant.

If the values for charge and mass be substituted and all quantities be expressed in practical units, then

$$v = 1.758 \times 10^{11} Et \qquad \text{meters per sec} \tag{6.5}$$

$$x = 0.879 \times 10^{11} Et^2 \qquad \text{meters} \tag{6.6}$$

$$v = 5.93 \times 10^5 \sqrt{V} \qquad \text{meters per sec} \tag{6.7a}$$

where v is velocity, meters per sec
E is gradient, volts per meter
V is potential, volts
x is distance through which the electron has been accelerated

The above expressions are not accurate for potentials exceeding 30,000 volts.

The ratio of the charge to the mass of the electron is so high that a small voltage will impart a tremendous velocity to the electron. It takes only three-tenths of a microvolt to give an electron a velocity of 700 mph which is approximately the velocity of sound. Although the speeds of electrons are very high, their energy is low because of their minute mass.

Electron speeds are frequently expressed in terms of the corresponding voltage. Energies are also designated in terms of electron volts, 1 electron volt being equal to 1.602×10^{-19} watt-sec. An electron that has "fallen" through 1 volt of potential is said to have acquired an energy of 1 electron volt.

If an electron enters a region of uniform field at a point x_0 with an initial velocity v_0 parallel to the field, then

$$v = -\frac{e}{m} Et + v_0 \qquad \text{meters per sec} \tag{6.8}$$

$$x = -\frac{e}{2m} Et^2 + v_0 t + x_0 \qquad \text{meters} \tag{6.9}$$

where x_0 is the initial distance in meters.

Eliminating t between Eqs. (6.8) and (6.9),

$$\tfrac{1}{2}m(v^2 - v_0^2) = e(V - V_0) \tag{6.10a}$$

or

$$v = 5.93 \times 10^5 \sqrt{V - V_0} \qquad \text{meters per sec} \tag{6.10b}$$

Equations (6.4), (6.7), and (6.10b), which give velocity in terms of potential, are not restricted to uniform fields or to one-dimensional fields. This is due to the fact that these equations express the conservation of energy and hence are independent of the electron path and the nature of the potential field.

6.2. Initial Velocity Not Parallel to Field. When an electron enters a region of a uniform field with an initial velocity that is at an angle with the gradient of potential, the electron follows a parabolic trajectory. This is because it experiences a constant force in the direction of the gradient and no force at right angles to this. The case is analogous to the mechanical case of a projectile fired from a gun in the absence of friction. The projectile is subjected to a constant downward force but has no force affecting the component of velocity parallel to the earth's surface.

The differential equations for the components of electron motion when the electron meets a retarding component of field are

$$\frac{d^2y}{dt^2} = -\frac{e}{m} E \tag{6.11}$$

and

$$\frac{d^2x}{dt^2} = 0 \tag{6.12}$$

The initial conditions that determine the solution of these equations are as follows:
When $t = 0$,

$$\frac{dy}{dt} = v_0 \cos\theta \qquad y = 0$$

$$\frac{dx}{dt} = v_0 \sin\theta \qquad x = 0$$

where θ is the angle that the initial velocity makes with the gradient of potential and v_0 is the initial velocity.

A first integration gives

$$\frac{dy}{dt} = -\frac{e}{m}Et + v_0 \cos\theta \tag{6.13}$$

and

$$\frac{dx}{dt} = v_0 \sin\theta \tag{6.14}$$

A second integration gives

$$y = -\frac{e}{2m}Et^2 + v_0 \cos\theta\, t \tag{6.15}$$

and

$$x = v_0 \sin\theta\, t \tag{6.16}$$

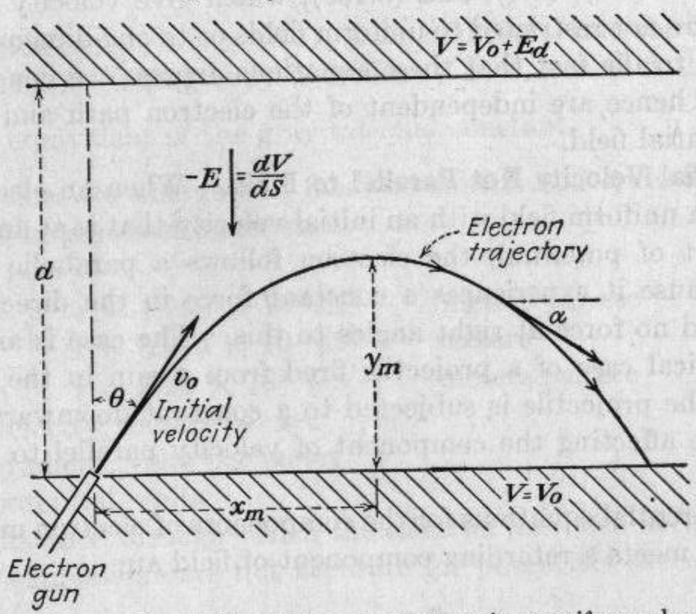


FIG. 6.1.—Parabolic electron trajectory in a uniform electric field. This case results from the injection of an electron with an initial velocity into a region where the electric field has a uniform retarding action.

Elimination of the time factor between Eqs. (6.15) and (6.16) gives the equation of the parabolic trajectory

$$y = \frac{-Ex^2}{4V_0 \sin^2\theta} + \frac{x}{\tan\theta} \tag{6.17}$$

where V_0 is the potential corresponding to the initial velocity. This is observed to be the equation of a parabola in x and y and to be independent of the system of electrical units used. The notation used in all the above equations corresponds to that shown in Fig. 6.1.

The slope of the trajectory at any point is given by

$$\tan\alpha = \frac{-Ex}{2V_0 \sin^2\theta} + \frac{1}{\tan\theta} \tag{6.18}$$

where α is the angle that the tangent to the parabola makes with the horizontal axis.

The maximum height to which the electron rises is

$$y_m = \frac{V_0 \cos^2\theta}{E} \tag{6.19}$$

and the horizontal displacement corresponding to this is

$$x_m = \frac{2V_0 \sin\theta \cos\theta}{E} \tag{6.20}$$

6.3. Electrostatic Deflection of Cathode-ray Beams. An application of the situation analyzed in the last section is found in the deflecting plates

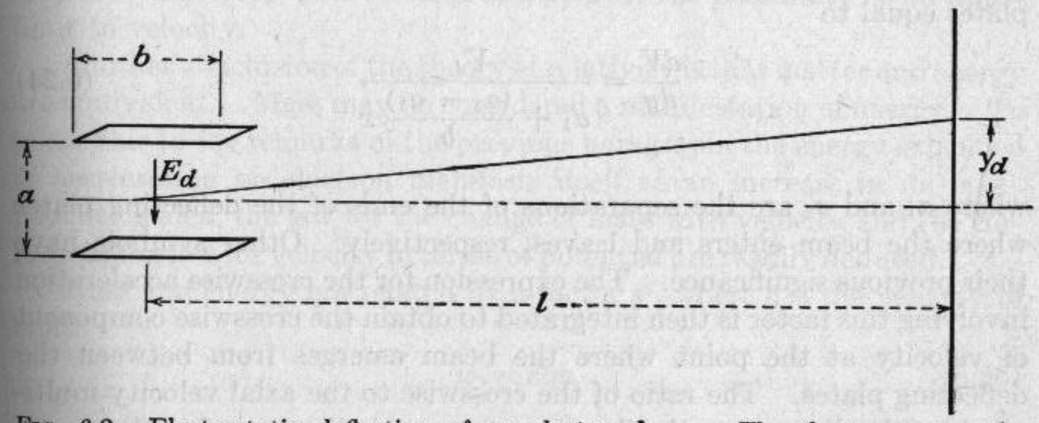


FIG. 6.2.—Electrostatic deflection of an electron beam. The electron enters the region of deflecting field at right angles to the field. The trajectory is parabolic between the plates.

of the ordinary cathode-ray tube. Here a stream of electrons enters a region of a uniform field, traverses a parabolic path while under the influence of this field, and leaves the region between the plates at a different angle from that at which it enters. It then travels in a straight line until it hits the fluorescent screen.

In this case, as may be seen by reference to Fig. 6.2, the electron enters the deflecting field at right angles, making the angle θ equal to 90 deg. The potential gradient is V_d/a. For this condition the slope of the trajectory upon emerging from the plates after a distance of travel b in the horizontal direction is, by Eq. (6.18)

$$\tan\alpha = \frac{V_d b}{2aV_0} + \frac{1}{\infty} \tag{6.21}$$

PROVIDENCE
COLLEGE

where V_d is the potential between plates and V_0 is the potential corresponding to the initial velocity. But

$$\tan \alpha \cong \frac{y_d}{l} \tag{6.22}$$

so that

$$y_d = \frac{l\,b V_d}{2a V_0} \tag{6.23}$$

This expression is only an approximate one, for it neglects the fringing effect of the flux lines around the end of the deflecting plates.

In most cathode-ray tubes the deflecting plates are not parallel but slope apart so that the electron in passing between them is subjected to a constantly decreasing gradient. When this is the case, the expressions obtained previously cannot be used and the problem must be solved anew. This is readily done by setting the gradient between the plates equal to

$$\frac{dV}{dy} = \frac{V_d}{a_1 + \dfrac{(a_2 - a_1)}{b}\,x} \tag{6.24}$$

where a_1 and a_2 are the separations of the ends of the deflecting plates where the beam enters and leaves, respectively. Other symbols have their previous significance. The expression for the crosswise acceleration involving this factor is then integrated to obtain the crosswise component of velocity at the point where the beam emerges from between the deflecting plates. The ratio of the crosswise to the axial velocity multiplied by the distance to the fluorescent screen is then equal to the screen displacement. This has the form

$$y = \frac{blV_d}{2V_0 a_1} \frac{ln\left(\dfrac{a_2}{a_1}\right)}{\left(\dfrac{a_2}{a_1} - 1\right)} \tag{6.25}$$

which reduces to Eq. (6.23) when $a_1 = a_2$. From this it is seen that the effect of spreading the deflecting plates at one end is to decrease the deflection. If the separation of the plates is increased 50 per cent at the far end, the deflection is decreased to 81.2 per cent of its value for the parallel plates having the near-end spacing. The deflection for divergent plates is, however, slightly greater than for parallel plates having their spacing equal to the average spacing of the divergent plates. Equation (6.25) is still in error because it takes no account of the flux fringing at

the ends of the deflecting plates.[1,2] The effect of the fringing is to increase the effective length of the plates.

6.4. Relativity Correction for Velocity. The general expression developed in Sec. 6.1 giving electron velocity as proportional to the square root of potential is valid only for velocities low compared with the velocity of light. This is due to the fact that according to the theory of relativity the mass of a particle changes with its velocity, and in the derivation of the expressions of Sec. 6.1 the mass was assumed constant.

One of the postulates of the theory of relativity is that nothing can move with a speed greater than the velocity of light. As a consequence of this upper limit on velocity, it is seen that a body subjected to a constant force must have its mass increase as it is accelerated, or otherwise its velocity would increase indefinitely and finally violate the postulate by exceeding the velocity of light. If, however, the mass of the particle increases as its velocity increases, a constant force produces an acceleration that decreases with velocity and permits the possibility of an upper limit to velocity.

Another conclusion of the theory of relativity is that matter and energy are equivalent. Mass may be considered a manifestation of energy. To relate this to the remarks of the previous paragraph, the energy expended in accelerating an electron manifests itself as an increase in its mass. From this idea, the law for the change of mass with velocity and the corresponding law for velocity in terms of potential are readily deduced.

Mass and energy are related by a factor c^2, where c is the velocity of light.

$$w = c^2 m \tag{6.26}$$

where w is energy in watt-seconds, c is the velocity of light, 3×10^8 meters per sec, and m is mass in kilograms.

Consider the increase in mass that an electron experiences as it is accelerated. Then

$$c^2\,dm = dw = F\,ds \tag{6.27}$$

where dm is the increase in mass, dw is the energy expended in accelerating the particle, F is the applied force, and s is the distance factor.

According to Newton's second law,

$$F = \frac{d}{dt}(mv) \tag{6.28}$$

[1] See also BENHAM, W., Inclined Deflecting Plates, *Wireless Eng.*, vol. 13 (No. 148) pp. 10–13, 1936.

[2] HINTERBERGER, O., Correction for End Effects in Oscilloscope Deflecting Plates, *Zeit. für Phys.*, vol. 105, pp. 501–512, July, August, 1937.

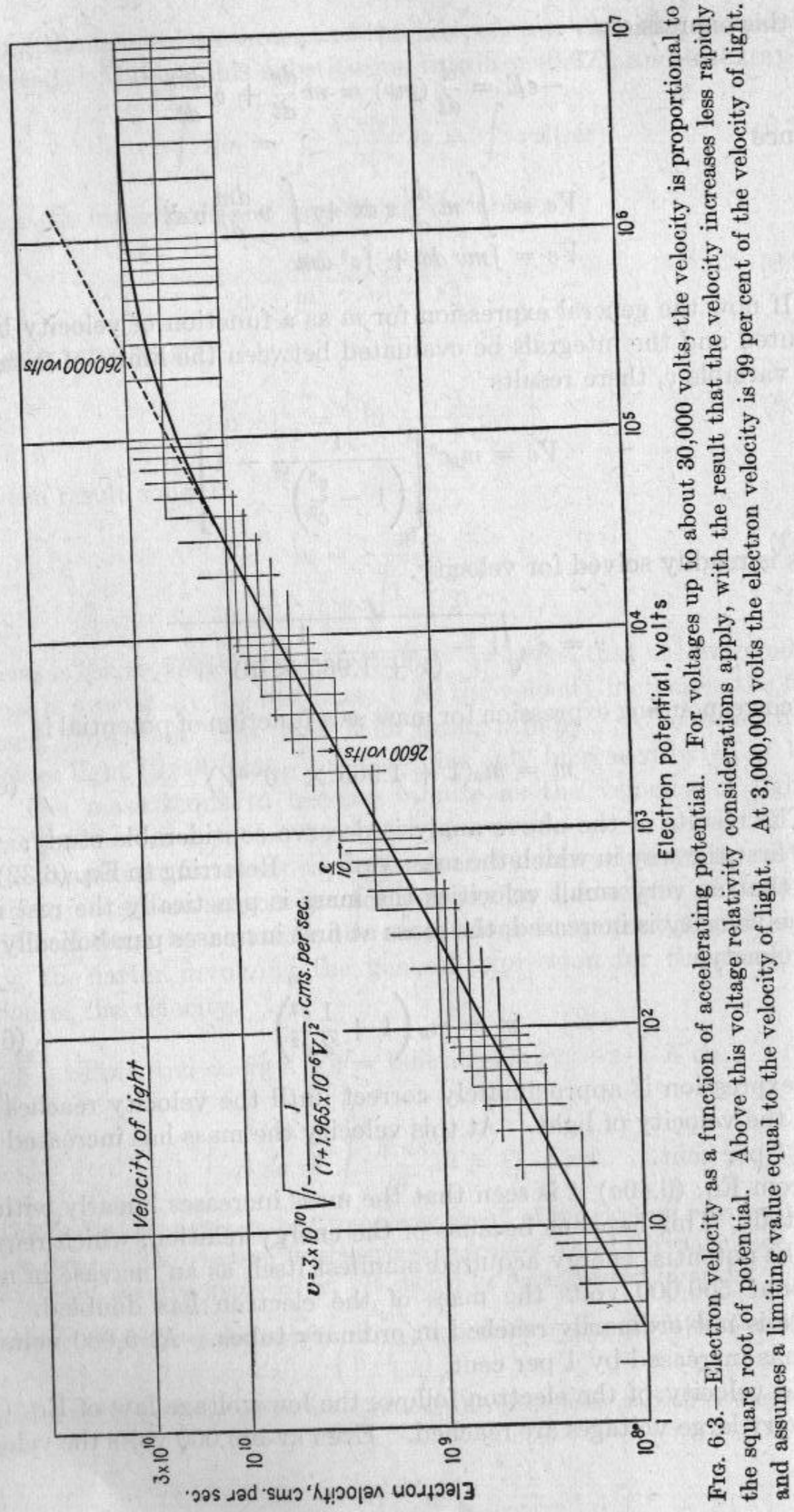


FIG. 6.3.—Electron velocity as a function of accelerating potential. For voltages up to about 30,000 volts, the velocity is proportional to the square root of potential. Above this voltage relativity considerations apply, with the result that the velocity increases less rapidly and assumes a limiting value equal to the velocity of light. At 3,000,000 volts the electron velocity is 99 per cent of the velocity of light.

has dropped only 7½ per cent from the low-voltage value given by Eq. (6.7*a*), which may be written

$$\frac{v}{c} = \frac{\sqrt{V}}{506} \tag{6.7b}$$

By 1,000,000 volts, however, the velocity has reached 93 per cent of the velocity of light. Above 1,000,000 volts the velocity becomes closer and closer to the velocity of light but experiences no rapid change with voltage.

The above relations are shown in Fig. 6.3. On this log-log plot it is seen that velocity follows the half-power law of potential well up to about 100,000 volts. Between 100,000 and 1,000,000 volts the change from the half-power law occurs, and above 1,000,000 volts the velocity is practically constant. Several convenient reference points may be taken from this curve. An electron reaches one-tenth the velocity of light at about 2,600 volts. If there were no change of mass with velocity, the electron would reach the velocity of light at about 260,000 volts.

6.5. Two-dimensional Electric Fields. Electrons are frequently exposed to fields that are not uniform but that are two-dimensional or more. It is generally quite difficult to determine exactly what the electron path is by analytical methods.

The fundamental differential equations involved are quite simple, but they are usually difficult if not impossible to solve. In rectangular coordinates the differential equations are

$$\frac{d^2x}{dt^2} = -\frac{e}{m}E_x(x,y) \tag{6.41}$$

and

$$\frac{d^2y}{dt^2} = -\frac{e}{m}E_y(x,y) \tag{6.42}$$

where

$$E_x = -\frac{\partial V\,(x,y)}{\partial x} \tag{6.43}$$

and

$$E_y = -\frac{\partial V\,(x,y)}{\partial y} \tag{6.44}$$

When these equations can be solved, they give the components of electron displacement parametrically in terms of t.

When the potentials are given in two-dimensional circular-cylinder coordinates with an axial symmetry, as is the case in most electron-optical problems, the equations have the same form as those above. It is necessary only to substitute r for x and z for y to get the corresponding equations for this case.

this being the general statement of the law when a variation in mass is encountered. Making this substitution into Eq. (6.27), and integrating

$$c^2 \int dm = \int \frac{d(mv)}{dt}\, ds = \int v\, d(mv) \tag{6.29}$$

Equating the integrands and separating variables,

$$\frac{dm}{m} = \frac{v\, dv}{c^2 - v^2} \tag{6.30}$$

which integrates to give

$$[\ln m]_{m_0}^{m} = \left[\ln \frac{1}{(c^2 - v^2)^{1/2}}\right]_0^v \tag{6.31}$$

giving the result sought,

$$m = \frac{m_0}{\sqrt{1 - \dfrac{v^2}{c^2}}} \tag{6.32}$$

where m_0 is the rest mass of the electron.[1] It is seen that at low velocities the mass is practically the rest mass. As the velocity increases, the mass increases, slowly at first and then quite rapidly. At one-tenth the velocity of light (2,600 volts) the mass has only increased by ½ of 1 per cent. The mass tends to become infinite as the velocity of light is approached.

The expression for mass as determined by the velocity can now be applied to obtain an expression for velocity as a function of potential. This is best done by equating the expressions for potential and kinetic energy, the latter involving the general expression for the mass as a function of the velocity.

$$\text{Potential energy, } Ve = \text{kinetic energy, } -e \int E\, ds \tag{6.33}$$

But

$$\int E\, ds = \int \frac{E\, ds}{dt}\, dt = \int Ev\, dt \tag{6.34}$$

[1] This is what is known as the "transverse mass" of the electron because it is the effective mass of the electron to transverse deflection where the magnitude of the velocity is not changed appreciably. It should be distinguished from the "longitudinal mass," which has the value $\dfrac{m_0}{\left(1 - \dfrac{v^2}{c^2}\right)^{3/2}}$, which is the effective mass that an electron presents to longitudinal acceleration where the mass as well as the velocity changes.

In this expression

$$-eE = \frac{d}{dt}(mv) = m\frac{dv}{dt} + v\frac{dm}{dt} \tag{6.35}$$

Hence

$$Ve = \int m\frac{dv}{dt}\, v\, dt + \int v\frac{dm}{dt}\, v\, dt \tag{6.36}$$

$$Ve = \int mv\, dv + \int v^2\, dm \tag{6.37}$$

If now the general expression for m as a function of velocity be substituted and the integrals be evaluated between the limits of 0 and v on the variable v, there results

$$Ve = m_0c^2\left[\frac{1}{\left(1 - \dfrac{v^2}{c^2}\right)^{1/2}} - 1\right] \tag{6.38}$$

This is readily solved for velocity.

$$v = c\sqrt{1 - \frac{1}{(1 + 1.965 \times 10^{-6}V)^2}} \tag{6.39}$$

The corresponding expression for mass as a function of potential is

$$m = m_0(1 + 1.965 \times 10^{-6}V) \tag{6.40a}$$

The results of the above analysis deserve considerable study. Consider first the way in which the mass varies. Referring to Eq. (6.32), it is seen that at very small velocities the mass is practically the rest mass. As the velocity is increased, the mass at first increases parabolically with the velocity,

$$m = m_0\left(1 + \frac{1}{2}\frac{v^2}{c^2}\right) \tag{6.40b}$$

This expression is approximately correct until the velocity reaches one-tenth the velocity of light. At this velocity the mass has increased only ½ of 1 per cent.

From Eq. (6.40a) it is seen that the mass increases linearly with the potential. This happens because of the energy relation, which requires that the potential energy acquired manifest itself as an increase in mass. At about 500,000 volts the mass of the electron has doubled. This voltage is not ordinarily reached in ordinary tubes. At 5,000 volts the mass has increased by 1 per cent.

The velocity of the electron follows the low-voltage law of Eq. (6.7) until very large voltages are reached. Even at 100,000 volts the velocity

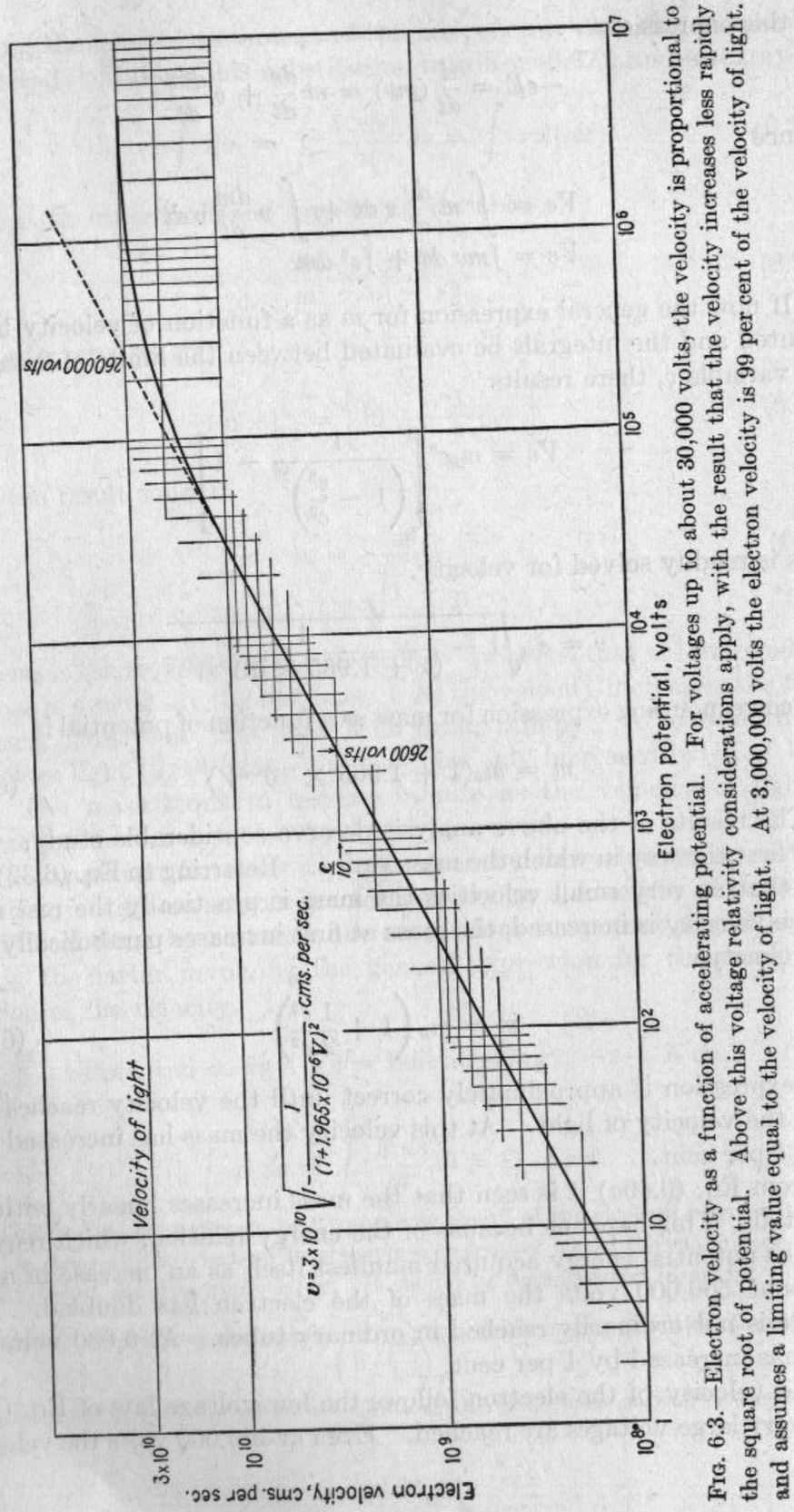


FIG. 6.3.—Electron velocity as a function of accelerating potential. For voltages up to about 30,000 volts, the velocity is proportional to the square root of potential. Above this voltage relativity considerations apply, with the result that the velocity increases less rapidly and assumes a limiting value equal to the velocity of light. At 3,000,000 volts the electron velocity is 99 per cent of the velocity of light.

has dropped only 7½ per cent from the low-voltage value given by Eq. (6.7*a*), which may be written

$$\frac{v}{c} = \frac{\sqrt{V}}{506} \tag{6.7b}$$

By 1,000,000 volts, however, the velocity has reached 93 per cent of the velocity of light. Above 1,000,000 volts the velocity becomes closer and closer to the velocity of light but experiences no rapid change with voltage.

The above relations are shown in Fig. 6.3. On this log-log plot it is seen that velocity follows the half-power law of potential well up to about 100,000 volts. Between 100,000 and 1,000,000 volts the change from the half-power law occurs, and above 1,000,000 volts the velocity is practically constant. Several convenient reference points may be taken from this curve. An electron reaches one-tenth the velocity of light at about 2,600 volts. If there were no change of mass with velocity, the electron would reach the velocity of light at about 260,000 volts.

6.5. Two-dimensional Electric Fields. Electrons are frequently exposed to fields that are not uniform but that are two-dimensional or more. It is generally quite difficult to determine exactly what the electron path is by analytical methods.

The fundamental differential equations involved are quite simple, but they are usually difficult if not impossible to solve. In rectangular coordinates the differential equations are

$$\frac{d^2x}{dt^2} = -\frac{e}{m}E_x(x,y) \tag{6.41}$$

and

$$\frac{d^2y}{dt^2} = -\frac{e}{m}E_y(x,y) \tag{6.42}$$

where

$$E_x = -\frac{\partial V\,(x,y)}{\partial x} \tag{6.43}$$

and

$$E_y = -\frac{\partial V\,(x,y)}{\partial y} \tag{6.44}$$

When these equations can be solved, they give the components of electron displacement parametrically in terms of t.

When the potentials are given in two-dimensional circular-cylinder coordinates with an axial symmetry, as is the case in most electron-optical problems, the equations have the same form as those above. It is necessary only to substitute r for x and z for y to get the corresponding equations for this case.

For two-dimensional polar coordinates, such as are used for the cylindrical triode, the equations are quite different and still more difficult to solve. In terms of a radial variable r and an angular variable θ the differential equations of motion are

$$\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 = \frac{e}{m}\frac{\partial V}{\partial r} \tag{6.45}$$

$$r\frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\frac{d\theta}{dt} = \frac{1}{r}\frac{e}{m}\frac{\partial V}{\partial \theta} \tag{6.46}$$

These equations are most readily obtained by applying the Lagrangian operator to the energy equation, which in these coordinates has the form

$$\frac{m}{2}\left[\left(\frac{dr}{dt}\right)^2 + r^2\left(\frac{d\theta}{dt}\right)^2\right] = eV \tag{6.47}$$

The difficulty in solving these two-dimensional problems arises from the fact that the variables in the component equations are rarely separable.

Example: One of the few two-dimensional problems that can be solved exactly is that of an electron released from a point on the side of an interior right-angled conducting corner at zero potential. The potential configuration is shown in Fig. 5.16. The equation for the potential is $V = kxy$ so that the components of electric intensity are $E_x = -ky$ and $E_y = -kx$. The differential equations of motion are then

$$\frac{d^2x}{dt^2} = \frac{e}{m}ky \tag{6.48}$$

and

$$\frac{d^2y}{dt^2} = \frac{e}{m}kx \tag{6.49}$$

It is convenient to make the substitution $\frac{e}{m}k = \omega^2$. If each equation is differentiated twice and the relations from the original equations substituted, there result

$$\frac{d^4x}{dt^4} = \omega^2 x \tag{6.50}$$

and

$$\frac{d^4y}{dt^4} = \omega^2 y \tag{6.51}$$

in which a separation of the variables has been achieved. When these equations are solved subject to the initial conditions that when

$$t = 0$$
$$x = 0$$
$$y = a$$

$$\frac{dx}{dt} = 0$$
$$\frac{dy}{dt} = 0$$

there results

$$x = \frac{a}{2}(\cosh \omega t - \cos \omega t) \tag{6.52}$$

and

$$y = \frac{a}{2}(\cosh \omega t + \cos \omega t) \tag{6.53}$$

The above solutions may be obtained either by standard methods or by the operational calculus. The nature of the solution is more apparent if the component displacements are referred to the line $y = x$, that is, if the system be rotated 45 deg clockwise. When this is done,

$$x_1 = \frac{a}{\sqrt{2}}\cosh \omega t \tag{6.54}$$

$$y_1 = \frac{a}{\sqrt{2}}\cos \omega t \tag{6.55}$$

This same result may be obtained more quickly if the original potential field be rotated 45 deg clockwise before formulating the differential equations. When this is done, the field has the form shown in Fig. 6.4 and the potential is

$$V = \frac{k}{2}(x_1^2 - y_1^2) \quad \text{and} \quad E_x = -kx_1, \quad E_y = +ky_1$$

The differential equations are then

$$\frac{d^2x_1}{dt^2} = \omega x_1 \tag{6.56}$$

$$\frac{d^2y_1}{dt^2} = -\omega y_1 \tag{6.57}$$

Here the variables are already separated in equations of lower order; and when these are solved subject to the conditions that when

$$t = 0$$
$$x_1 = \frac{a}{\sqrt{2}}$$
$$y_1 = \frac{a}{\sqrt{2}}$$
$$\frac{dx_1}{dt} = 0$$
$$\frac{dy_1}{dt} = 0$$

the same solution as was obtained above results.

Examination of the solution shows that the electron oscillates about the line of symmetry while moving outward at a constantly increasing rate. It is also seen that the *shape* of the trajectory is independent of the strength of the field and also of the charge and mass of the electron. This is a general characteristic of such problems. The transit time, however, does depend upon all three of these factors. This means that a heavier particle starting under the same conditions will trace out the same path but be slower in doing so.

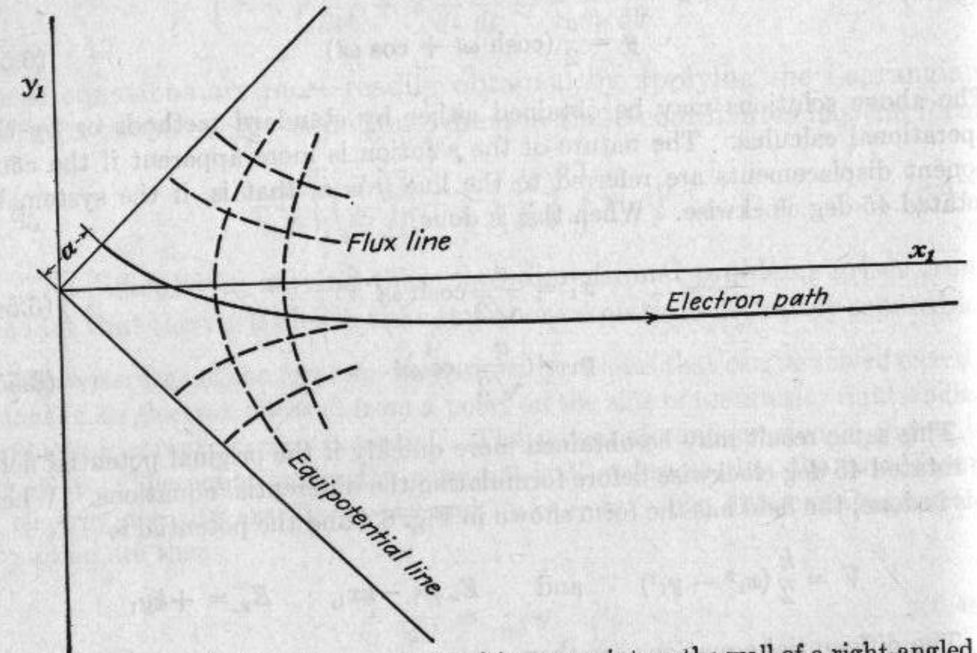


FIG. 6.4.—Path of an electron released from a point on the wall of a right-angled corner. Note that the electron does not follow a flux line but, because of its finite mass, overshoots the line of symmetry and subsequently oscillates about it.

If the general differential equations (6.41) and (6.42) are combined with the energy equation

$$\frac{m}{2}\left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2\right] = eV \tag{6.58}$$

and the factor t be eliminated between them, there results a differential equation in the coordinates x and y alone,

$$2V(x,y)\,\frac{d^2y}{dx^2} = \left[E_x\frac{dy}{dx} - E_y\right]\left[1 + \left(\frac{dy}{dx}\right)^2\right] \tag{6.59}$$

This equation is no easier to solve than those previously given, but several important properties of electron trajectories can be deduced from it.

1. The mass and charge of the electron do not appear in the equation. This means that the path taken will be independent of these factors.
2. The equation is not changed if either voltage or distance is changed by a constant factor. This means that the path will be the same for all magnitudes of voltage as long as the form of the field is not changed.
3. If the tube structure is enlarged by any factor, then the trajectory will be enlarged by the same factor.

6.6. Electron in a Uniform Magnetic Field. An electron in motion constitutes a minute electric current of magnitude $-ev$, where e is the magnitude of the charge on the electron and v is its velocity. As such, an electron in a magnetic field experiences a sidewise force just as does a wire carrying current. The magnitude of this force in newtons is $Bev \sin \theta$, where B is the magnetic-flux density in webers per square meter (1 weber per meter^2 equals 10^4 gausses) and θ is the angle between the vectors representing the field and the velocity, the latter being in units of meters per second. When the electron enters the field at right angles to it, the force is simply Bev directed at right angles to the velocity. The relative directions of field, velocity, and force are shown in Fig. 6.5. The force changes the direction but not the magnitude of the velocity and in this case is continuously exerted at right angles to the instantaneous velocity because the direction of the force changes with the direction of the velocity. This fulfills the conditions necessary for a circular motion of the electron in a plane normal to the magnetic field.

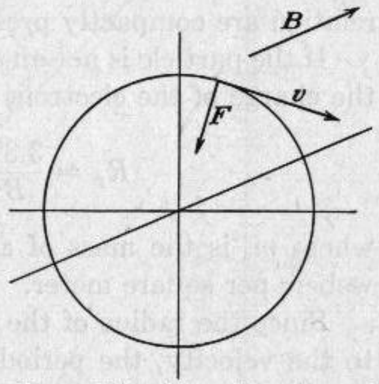


FIG. 6.5.—The direction of the force on an electron relative to the velocity and magnetic field that produce it. The force is the vector product of the magnetic-flux density and the velocity. If B is turned into v, then F advances like a right-hand screw.

The force developed by the magnetic field may be considered as a centripetal force that must equal the centrifugal force developed by the circular motion of the electron. Equating these forces,

$$Bev = \frac{mv^2}{R} \quad \text{newtons} \tag{6.60}$$

where R is the radius of the circular electron path. From this the radius of the circular path is

$$R = \frac{m}{e}\frac{v}{B} \quad \text{meters} \tag{6.61}$$

It is more convenient for application of this formula to express the physical quantities in numerical form and to use the potential corresponding to the velocity. With these changes the expression for the radius of the circular path becomes

$$R = 3.37 \times 10^{-6} \frac{\sqrt{V}}{B} \quad \text{meters} \tag{6.62}$$

where V is the electron velocity in equivalent volts and B is the magnetic-flux density in webers per square meter (10^4 gausses). This relation shows that, the stronger the field and the smaller the velocity of the electron, the smaller the circle in which it moves. The results of this relation are compactly presented in the nomogram of Fig. 6.6.

If the particle is not an electron but an ion of mass m_p and with n times the charge of the electron, the radius is given by

$$R_p = \frac{3.37}{B} \times 10^{-6} \sqrt{\frac{m_p}{m_e} \frac{V}{n}} \quad \text{meters} \tag{6.63}$$

where m_e is the mass of an electron and B is magnetic-flux density in webers per square meter.

Since the radius of the circle followed by the particle is proportional to the velocity, the period corresponding to one loop is independent of the initial velocity and depends only upon the magnetic field. The period is given by the circumference of the circle divided by the velocity,

$$T = \frac{2\pi m}{e} \frac{1}{B} \quad \text{sec} \tag{6.64}$$

In practical units this is

$$T = \frac{35.5}{B} \quad \text{micromicroseconds} \tag{6.65}$$

The value of the period can be obtained from the nomogram of Fig. 6.6 by observing that the period in microseconds is the same as the radius in centimeters when the potential is 11.22 mv.

For particles with a mass m_p and having n times the charge of the electron the period is

$$T = \frac{35.5}{nB} \frac{m_p}{m_e} \quad \text{micromicroseconds} \tag{6.66}$$

The fact that the period is independent of the velocity is significant and useful. If a number of electrons of different velocities be injected into a uniform magnetic field, they will trace out circles of different size but they will all return to the starting point at the same time. Use is made of this property in magnetic focusing of electron beams.

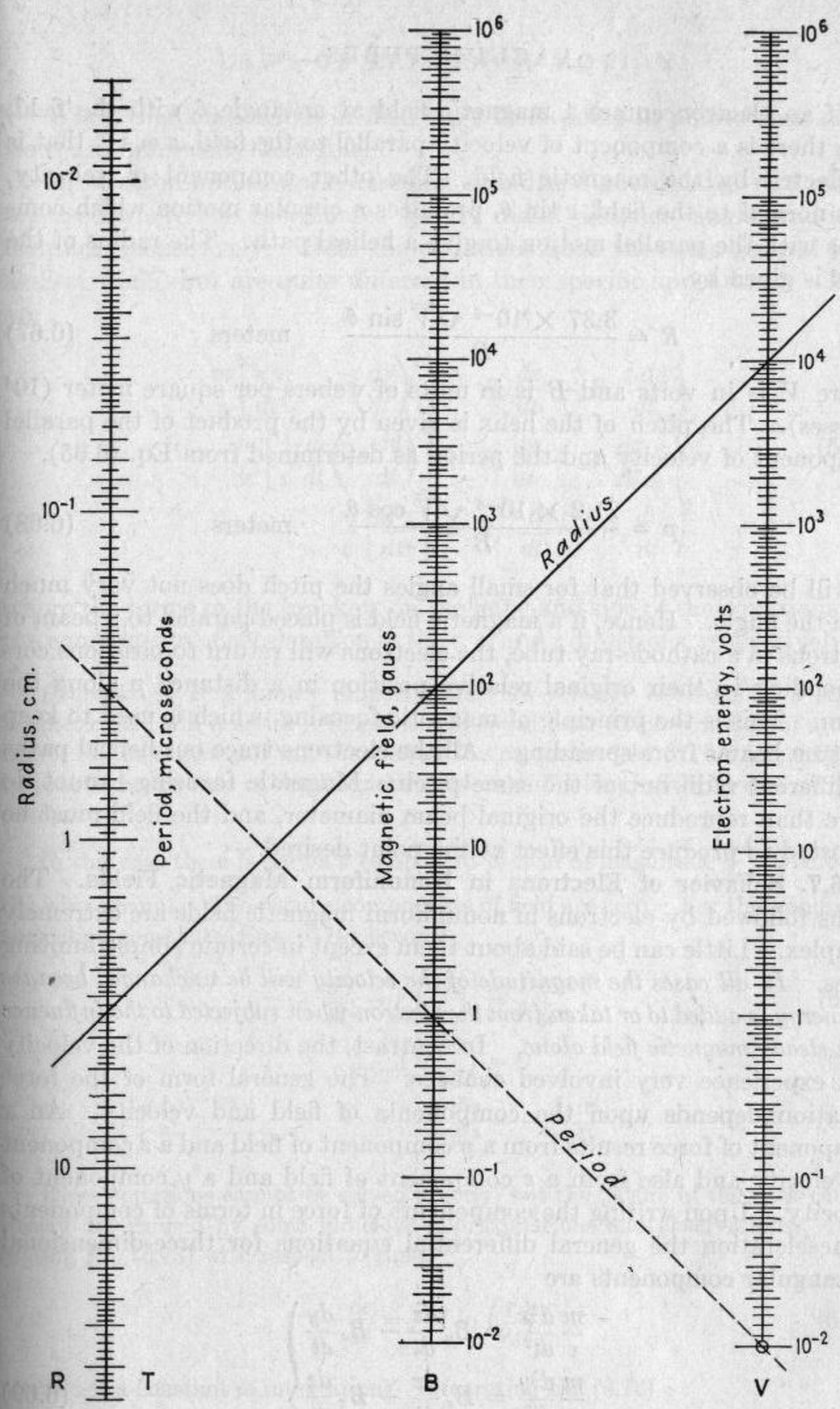


FIG. 6.6.—Nomographic chart giving the radius of the circular path of an electron in a uniform magnetic field as a function of the magnetic field strength and the electron energy in volts, Eq. (6.62). The chart also gives the period of a single rotation as a function of magnetic field strength, Eq. (6.65).

If an electron enters a magnetic field at an angle θ with the field, then there is a component of velocity parallel to the field, $v \cos \theta$, that is unaffected by the magnetic field. The other component of velocity, that normal to the field, $v \sin \theta$, produces a circular motion which combines with the parallel motion to give a helical path. The radius of the helix is given by

$$R = \frac{3.37 \times 10^{-6} \sqrt{V} \sin \theta}{B} \quad \text{meters} \tag{6.67}$$

where V is in volts and B is in units of webers per square meter (10^4 gausses). The pitch of the helix is given by the product of the parallel component of velocity and the period as determined from Eq. (6.65).

$$p = \frac{21.2 \times 10^{-6} \sqrt{V} \cos \theta}{B} \quad \text{meters} \tag{6.68}$$

It will be observed that for small angles the pitch does not vary much with the angle. Hence, if a magnetic field is placed parallel to a beam of electrons in a cathode-ray tube, the electrons will return to positions corresponding to their original relative position in a distance p along the beam. This is the principle of magnetic focusing, which is used to keep electron beams from spreading. All the electrons trace out helical paths of different radii but of the same pitch. Magnetic focusing cannot do more than reproduce the original beam diameter, and the field must be adjusted to produce this effect at the point desired.

6.7. Behavior of Electrons in Nonuniform Magnetic Fields. The paths followed by electrons in nonuniform magnetic fields are extremely complex. Little can be said about them except in certain simple limiting cases. *In all cases the magnitude of the velocity will be unchanged because no energy is added to or taken from the electron when subjected to the influence of a steady magnetic field alone.* In contrast, the direction of the velocity can experience very involved changes. The general form of the force equation depends upon the components of field and velocity. An x component of force results from a y component of field and a z component of velocity and also from a z component of field and a y component of velocity. Upon writing the components of force in terms of components of acceleration the general differential equations for three-dimensional rectangular components are

$$\left.\begin{aligned} \frac{m}{e}\frac{d^2x}{dt^2} &= B_y \frac{dz}{dt} - B_z \frac{dy}{dt} \\ \frac{m}{e}\frac{d^2y}{dt^2} &= B_z \frac{dx}{dt} - B_x \frac{dz}{dt} \\ \frac{m}{e}\frac{d^2z}{dt^2} &= B_x \frac{dy}{dt} - B_y \frac{dx}{dt} \end{aligned}\right\} \tag{6.69}$$

When the components of field vary from point to point, these equations are practically insoluble.

In electron-optics work, circular cylindrical coordinates r, θ, and z are used where the coordinates specify radial distance, angle, and axial distance, respectively. Here the equations have the same general form as Eqs. (6.69) but are quite different in their specific appearance. They are

$$\left.\begin{aligned} \frac{m}{e}\left[\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2\right] &= B_\theta \frac{dz}{dt} - B_z \frac{r\,d\theta}{dt} \\ \frac{m}{e}\left[\frac{1}{r}\frac{d}{dt}\left(r^2\frac{d\theta}{dt}\right)\right] &= B_z \frac{dr}{dt} - B_r \frac{dz}{dt} \\ \frac{m}{e}\left[\frac{d^2z}{dt^2}\right] &= B_r \frac{r\,d\theta}{dt} - B_\theta \frac{dr}{dt} \end{aligned}\right\} \tag{6.70}$$

where the terms in the brackets on the left-hand side of the equations are the components of acceleration in the r, θ, and z directions, respectively.

Example: It is a known property that low-velocity electrons in a strong magnetic field will describe a tightly coiled spiral path which wraps itself around one of the flux lines and will thus follow the magnetic field. This property will be proved in the case of the magnetic field around a long, straight wire carrying current.

In this case there is only a θ component of field of magnitude $\frac{\mu I}{2\pi r}$, where I is the wire current. The r and z components of field are zero. For this condition, neglecting constants, Eqs. (6.70) become

$$\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 = \frac{I}{r}\frac{dz}{dt} \tag{6.71}$$

$$\frac{1}{r}\frac{d}{dt}\left(r^2\frac{d\theta}{dt}\right) = 0 \tag{6.72}$$

$$\frac{d^2z}{dt^2} = -\frac{I}{r}\frac{dr}{dt} \tag{6.73}$$

These equations cannot be solved exactly, but the nature of the path can be closely determined by some judicious approximations and observations. Integrating Eq. (6.73) with respect to time,

$$\frac{dz}{dt} = -I \ln \left(\frac{r}{r_0}\right) \tag{6.74}$$

where r_0 is a constant of integration. Integrating Eq. (6.72)

$$r^2\frac{d\theta}{dt} = C_\theta \tag{6.75}$$

where C_θ is a constant of integration. Substituting Eqs. (6.74) and (6.75) into Eq. (6.71),

$$\frac{d^2r}{dt^2} = \frac{I}{r}\left[-I \ln \left(\frac{r}{r_0}\right)\right] + \frac{C_\theta^2}{r^3} \tag{6.76}$$

This expression would be difficult to integrate exactly, but the form of the variation in r can be determined. It is seen that there is a value of r for which the acceleration is zero and that for values of r slightly less than this the acceleration is positive, while for values of r slightly greater than this the acceleration is negative. This means that, if the initial r component of velocity is small, the electron will oscillate about the value of r for which the acceleration is zero. Hence the expression for the r component of position is of the form

$$r = r_1 - a \sin kt \tag{6.77}$$

With this variation of r the z component of velocity is also seen to be periodic of small variation from Eq. (6.74), and hence z oscillates about its original value. Similarly the θ component of velocity is periodic and of small variation from Eq. (6.75) so that the value of $\frac{d\theta}{dt}$ is a constant with a superimposed periodic variation. The net result of these component displacements is that the electron will spiral around a flux line in some fashion, keeping a constant average value of r and z, and progress in the θ direction with a constant average velocity as shown in Fig. 6.7. Use of this property is made in television pickup tubes of the Orthicon type.[1]

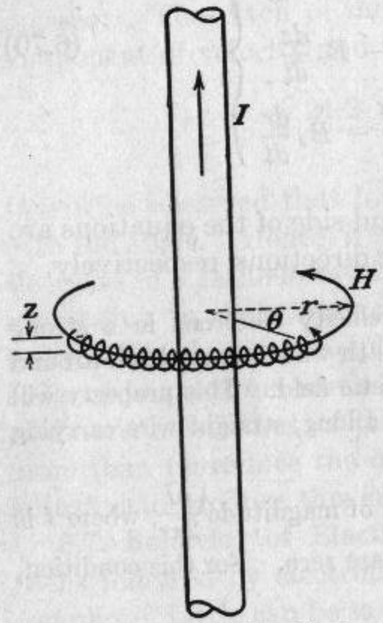


FIG. 6.7.—The motion of a low-velocity electron about a magnetic flux line. In the absence of strong electric fields, low-velocity electrons will spiral about magnetic flux lines.

6.8. Combined Electric and Magnetic Fields. When an electron is subjected to the combined action of both electric and magnetic fields, the paths tend to become quite complex. Some simple cases can be studied, however.

When an electron starts from rest under the influence of parallel electric and magnetic fields, the electron moves in the direction of the electric field and is unaffected by the magnetic field. The path in this case is a straight line, and the electron behaves as though the magnetic field did not exist.

If an electron with a given velocity is injected into a region containing electric and magnetic fields at right angles to each other and each at right angles to the initial velocity, then there is a certain ratio of electric- to magnetic-field strength for which the electron is not deflected in its path. This occurs when the force due to the electric field is equal and opposite to that produced by the magnetic field. For this condition

$$\frac{E}{B} = 5.93 \times 10^5 \sqrt{V_0} = v \quad \text{meters per sec} \tag{6.78}$$

where v is the original velocity and V_0 the potential that produced it. As long as the above relation holds, the electron moves in a straight line. If any of the quantities involved is changed, the electron will be deflected from the straight-line path.

If an electron starts from rest in the presence of uniform electric and magnetic fields that are mutually perpendicular, it first experiences a force in the direction of the electric field and is unaffected by the magnetic field because of the low velocity. As it acquires velocity, it is deflected sidewise by the magnetic field. This action turns it around and brings it to rest at a point corresponding to its original position but displaced to one side. If the electron is then free to move, the action is repeated and the resulting path is a cycloid. The cycloidal nature of the path can be seen by considering that, if the magnetic field were moving in a direction mutually perpendicular to the electric field and to itself at a velocity given by Eq. (6.78), then to an observer moving with the magnetic field the effects of the two fields would cancel as far as forces parallel to the electric field were concerned. To this same observer the electron would behave as though it were injected into a magnetic field alone with a velocity given by Eq. (6.78) in a direction opposite to that of the observer's motion, and the resulting path would be a circle to this observer. To someone standing still relative to the fields the motion would be a circular motion combined with a translational motion, which in this case because of the equality of the velocity components gives rise to a cycloidal path.

For the relative position of the fields shown in Fig. 6.8, where B is in the negative z direction, the differential equations of motion are

$$\frac{d^2y}{dt^2} = \frac{e}{m}\frac{dV}{dy} - \frac{Be}{m}\frac{dx}{dt} \tag{6.79}$$

and

$$\frac{d^2x}{dt^2} = \frac{Be}{m}\frac{dy}{dt} \tag{6.80}$$

These equations are more simply written in the form

$$\ddot{y} = (a - \omega\dot{x}) \tag{6.81}$$

$$\ddot{x} = \omega\dot{y} \tag{6.82}$$

[1] ROSE, A., and H. IAMS, Television Pickup Tubes Using Low-velocity Electron Beam Scanning, *Proc. I.R.E.*, vol. 27, pp. 547–555, September, 1939.

where the dots indicate derivatives with respect to time and $a = \frac{e}{m}\frac{dV}{dy}$ and $\omega = \frac{Be}{m}$. The initial conditions are that, when t is zero, y, x, $\dot{y}$, and $\dot{x}$ are also zero.

Integrating Eq. (6.82) with respect to time,

$$\dot{x} = \omega y \tag{6.83}$$

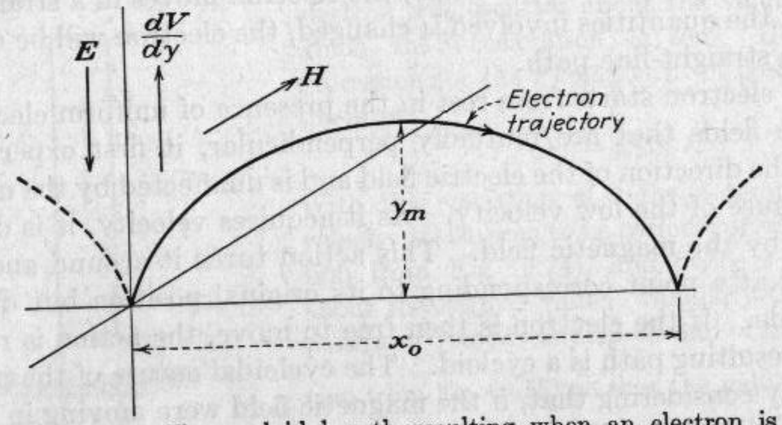


FIG. 6.8.—The cycloidal path resulting when an electron is liberated at zero velocity in crossed uniform electric and magnetic fields. The electron progresses in the positive x direction when the gradient of the electric field is in the positive y direction and the magnetic field is in the negative z direction.

since, when t equals zero, $\dot{x}$ and y are also zero. When this value of $\dot{x}$ is substituted in Eq. (6.81), there results

$$\ddot{y} = a - \omega^2 y \tag{6.84}$$

This can be solved either by standard methods or by the operational calculus to give

$$y = \frac{a}{\omega^2}(1 - \cos \omega t) \tag{6.85}$$

and the corresponding expression for x is from an integration of Eq. (6.83),

$$x = \frac{a}{\omega^2}(\omega t - \sin \omega t) \tag{6.86}$$

The last two equations above give the motion of the electron parametrically in terms of t. The motion is seen to consist of a uniform translation in the x direction with a superimposed circular motion.

The maximum displacement in the y direction is

$$y_m = \frac{2a}{\omega^2} = 1.138 \times 10^{-11}\frac{\frac{dV}{dy}}{B^2} \tag{6.87}$$

The displacement in the x direction corresponding to one cycle of the motion is found by substituting the value of time that restores the value of y to zero. This occurs for ωt equal to 2π so that

$$x_0 = \pi y_m \tag{6.88}$$

which is also to be expected from the ratio of the circumference to the radius of the generating circle that produces the cycloidal motion.

When an electron is injected into a region with uniform electric and magnetic fields at right angles to each other but with a finite initial velocity normal to the magnetic field, it will follow a trochoidal path in a plane normal to the magnetic field. Geometrically the trochoidal path is generated by a point on the rim of a wheel that is rolling along a straight line on a smaller diameter hub. The cycloid is the special case of the trochoid for which the diameters of the rolling and tracing circles are the same.

The differential equations for the case of an initial velocity are the same as for the cycloidal case [Eqs. (6.81) and (6.82)]; but in this case the initial conditions are different, and the form of the solution is hence different. When t is zero, y and x are zero, but $\dot{y} = v_{0y}$, $\dot{x} = v_{0x}$. Hence the first integration of Eq. (6.83) for the configuration of Fig. 6.8 gives

$$\dot{x} = \omega y - v_{0x} \tag{6.89}$$

When this substitution is made in Eq. (6.81) and this expression integrated twice to obtain the value for y, there results

$$y = \frac{(a - \omega v_{0x})(1 - \cos \omega t)}{\omega^2} + \frac{v_{0y}}{\omega}\sin \omega t \tag{6.90}$$

Substituting this in Eq. (6.89) and integrating to get the corresponding expression for x,

$$x = \frac{at}{\omega} + (1 - \cos \omega t)\frac{v_{0y}}{\omega} - \frac{a - \omega v_{0x}}{\omega^2}\sin \omega t \tag{6.91}$$

The two equations (6.90) and (6.91) determine completely the nature of the trochoidal path. The corresponding expressions for the components of velocity are

$$\dot{y} = v_{0y}\cos \omega t - \frac{a - \omega v_{0x}}{\omega}\sin \omega t \tag{6.92}$$

$$\dot{x} = \frac{a}{\omega} + v_{0y}\sin \omega t - \frac{a - \omega v_{0x}}{\omega}\cos \omega t \tag{6.93}$$

From the velocity-component equations it is seen that there is a constant x component of velocity of magnitude $\frac{a}{\omega}$. This corresponds to the translational velocity of the circles that generate the trochoidal motion. To achieve this translational velocity the radius of the rolling circle must be $\frac{a}{\omega^2}$ since the angular velocity of the velocity vectors is ω. The initial conditions also require that the instantaneous velocity of rotation of a point on the tracing circle be equal to the vector difference of the initial

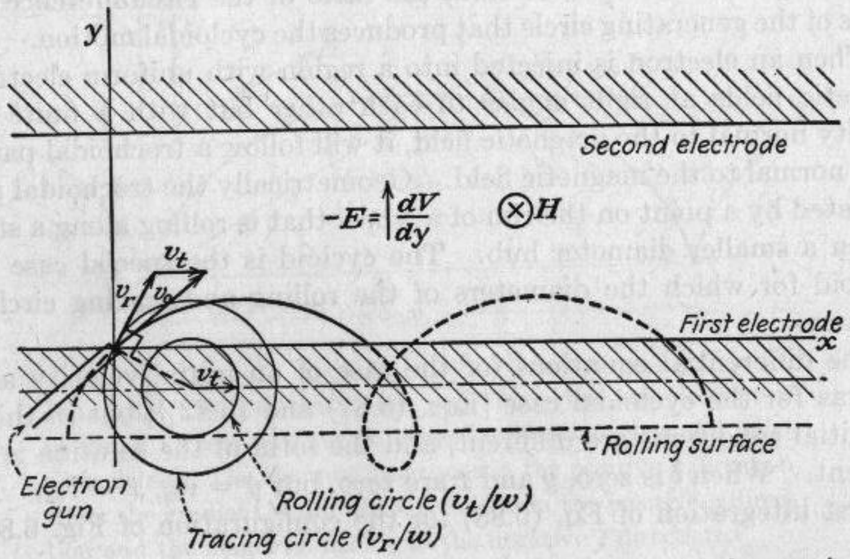


FIG. 6.9.—Trochoidal electron path resulting when an electron is injected with a finite velocity into a region of uniform crossed electric and magnetic fields. The electron will progress in the positive x direction when the gradient of the electric potential is in the positive y direction and the magnetic field is in the negative z direction.

velocity and the translational velocity. This relation is shown by the vector diagram of Fig. 6.9.

If the scale of the velocity-vector diagram be taken the same as that of the diagram showing the generating circles and the resulting path, the electron path can be constructed graphically in quite a simple manner. It will be observed that the terms in Eqs. (6.90) and (6.91) giving the instantaneous displacements are the same as those in Eqs. (6.92) and (6.93) for the component velocities except for the ω factors. The center of the generating circles in the initial position is given by rotating the rotational vector v_r, 90 deg in a clockwise direction. The radius of the tracing circle is then given by the length of the rotational vector v_r, and the radius of the rolling circle is given by the length of the vector

v_t, the translational velocity. With the rolling and tracing circles and their initial position given, the path is readily constructed geometrically for any case. This construction is also illustrated in Fig. 6.9.

6.9. Approximate Numerical and Graphical Methods for Determining Electron Paths. The number of cases in which the motion of an electron under the influence of applied fields can be determined exactly is actually quite small and restricted to very simple cases. Hence the need exists for methods that will give an approximate answer when the fields are more complex, as they usually are.

Method of Joined Circular Segments. When an electron is moving through a potential field, the instantaneous radius of curvature of its path is determined by its velocity and by the sidewise force that is exerted on it by the field.[1] The sidewise force exerted on the electron depends upon the component of the gradient of potential normal to the instantaneous direction. This component of the gradient will be designated by $\nabla_n V$. The actual sidewise force is $e\nabla_n V$. This force must equal the centrifugal force of the electron in its instantaneous circular motion, and this is given by $\frac{mv^2}{R}$, where R is the instantaneous radius of curvature. Equating these two forces and substituting $2eV$ for mv^2,

$$\frac{2eV}{R} = e\nabla_n V \tag{6.94}$$

from which

$$R = \frac{2V}{\nabla_n V} \tag{6.95}$$

This is the instantaneous radius of curvature of the electron path at any point in the field, as shown in Fig. 6.10, on the assumption that the electron started from rest at a point of zero potential. It will be observed that the radius of curvature is independent of the mass and charge of the electron and also of the scale of potential, checking the conclusions drawn from the differential equations of the electron path.

By calculating the radius of curvature at a point in the field by Eq. (6.95), drawing a small segment of arc, and then applying this process repeatedly a good approximation to the actual curve traced by the electron is obtained. The potential at any point in the field is easily obtained, and the normal component of gradient is the projection of the vector giving the magnitude and direction of the greatest variation of potential upon a line normal to the electron's path. The method is subject to cumulative error unless the average potential and average normal

[1] SALINGER, H., Tracing Electron Paths in Electric Fields, *Electronics*, vol. 10, pp. 50–54, October, 1937.

gradient over each segment of arc is used. If the segment of arc drawn at each step is kept a constant fraction of the radius of curvature, say one-twentieth, the error will not be great. It is also possible to derive simple expressions for the position of the next step in terms of the displacement and change of angle when the radius of curvature is so large that the arc segment is not easily drawn, as is frequently the case.

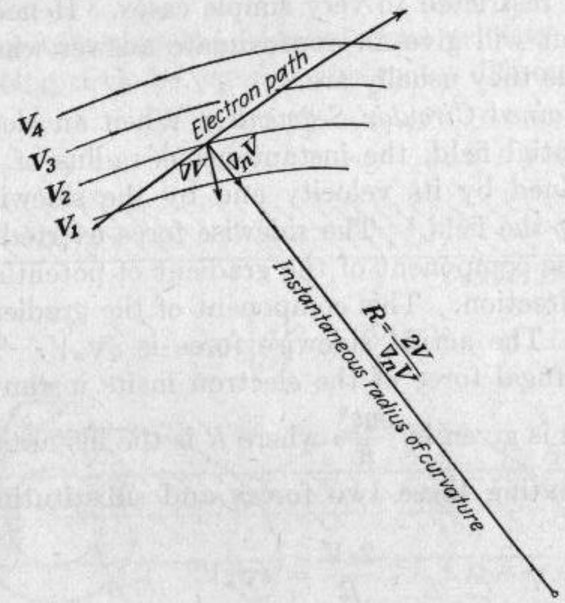


FIG. 6.10.—The instantaneous radius of curvature of an electron path in a region of varying potential. The instantaneous radius of curvature is equal to twice the potential at the point in question divided by the component of the gradient of potential perpendicular to the path.

Several ingenious gadgets have been devised that make the application of the principle outlined above purely a mechanical one.[1,2] These make use of a double probe in a current-flow model that has been set up to give the electric field involved. The double probe picks up a voltage proportional to the component of gradient in the direction of its alignment, and the average potential of the probes gives the potential at the point. The probe is connected to a small cart attached to a pantograph. The cart is steered in such a way that the instantaneous curvature of path which it is tracing is determined by the relation Eq. (6.95). Adjustments on the steering are made as continuously as possible from the information picked up by the double probe as it traces out the path. These devices are capable of considerable accuracy.

[1] GABOR, D., Mechanical Tracer for Electron Trajectories, *Nature*, vol. 139, p. 373, February, 1937.

[2] LANGMUIR, D., Automatic Plotting of Electron Trajectories, *Nature*, vol. 139, pp. 1066–1067, June 19, 1937.

Use of Elastic-membrane Model of Potential to Determine Electron Paths. In a previous chapter it was pointed out that the elevation of the surface of a stretched elastic membrane approximated closely the solution of Laplace's equation when the deformations were small. Actually such a model of potential fields is of more use in determining electron paths than in solving potential problems. This is because it is found that the laws governing the motion of a small sphere rolling on the membrane are strictly analogous to the laws governing the motion of an electron in an electric field, and hence the path of such a sphere is a good approximation of the path of an electron in the corresponding electric field.

Except for frictional effects the kinetic energy picked up by a small sphere, say a $\frac{3}{16}$-in. ball bearing, is equal to the potential energy it has lost owing to its change in elevation. This is exactly what happens to the electron. In the case of the mechanical model, however, the kinetic energy is divided between translational and rotational components. As long as the sphere rolls with a given circle of contact, the proportionality between these two components of the kinetic energy is constant and the path of the sphere will be similar to that of the electron. Although it is difficult to prove mathematically, it can readily be shown by experiments with a large sphere on a hard, curved surface that the sphere will turn relatively sharp corners and finish with the same rolling circle of contact as it had initially. The sphere can actually change its direction by about 300 deg without losing its original circle of rolling contact. If the angle is more than 300 deg, the turn introduces a spinning action that spoils the energy relations indicated above. Actually, it is the radius of curvature of the path rather than the angle that matters. Roughly, the limiting radius of the path is five times the radius of the sphere.

Application of the Principle of Least Action. In many electron-trajectory problems use can be made of the principle of least action. This principle states that in a potential field of the type encountered in vacuum-tube problems a particle will move between two points by such a path that the action, defined as the integral of momentum with distance, will have a minimum value. This means that, if the paths are known and conditions are such that only one electron goes through each point, contours of constant action calculated from the defining integral will be everywhere perpendicular to the electron paths. Hence, if the electron paths are known, contours of constant action can be found that give the path of all electrons. If the electron paths are not known, as is generally

the case, it is still possible to calculate the contours of constant action by methods involving successive approximation.

This is done by assuming a path of the electrons that is known to be close to the true path and then calculating the action along the assumed path. The process is easy, for the square root of potential can be substituted for velocity so that action is given by

$$A = \int mv\, ds = \sqrt{2em} \int \sqrt{V}\, ds \tag{6.96}$$

When the action along the assumed paths has been calculated, a first approximation to the action function has been obtained and corresponding contours of constant action can be drawn. If now curves be drawn normal to these contours of constant action, these will give a more accurate representation of the electron paths than those originally assumed. The second approximation to the action function can then be calculated along the improved paths and the process repeated to give any desired degree of accuracy. This is seen to be a perturbation process between action and potential.

In actual application in cases where the electron deflections are slight the first step of the process gives results that are sufficiently accurate.[1,2] The errors involved compensate because of the fact that, when the assumed path is shorter than the actual path, the potentials involved are smaller. In the determination of electron paths in tubes the assumption of straight-line paths initially is usually sufficiently good for cases in which the electron deflections are slight. The method is not accurate in the vicinity of any line of symmetry.

[1] LANGE, H., Current Division in Triodes and Its Significance in the Determination of Contact Potential, *Zeit. für Hochfrequenz.*, vol. 31, pp. 105–109, 133–140, 191–196, 1928.

[2] SPANGENBERG, KARL, Current Division in Plane-electrode Triodes, *Proc. I.R.E.*, vol. 28, pp. 226–236, May, 1940.

CHAPTER 7

ELECTROSTATIC FIELD OF A TRIODE

7.1. Method of Solution. The electrostatic fields within tubes are most readily obtained by means of the conformal transformations given in a previous chapter. These transformations give potential configurations that represent closely the fields encountered in tubes, whose electrode configuration is somewhat idealized. The cathode is assumed to be a plane or cylindrical surface, which it rarely is in practice. The electrodes are assumed to be infinite in length and breadth so that tube constants per unit area evaluated on this assumption do not include end effects.

It should be pointed out that, since the solutions obtained are not mathematically exact, various degrees of approximation are possible. In general, the more accurate the solution, the more complex and cumbersome the expressions obtained. Where extreme accuracy is desired, the method of conformal transformations is used as a starting point for series representations. Imaging or series procedures may also be used, but these have not proved of great value as a general method.

In spite of the above-mentioned departures from exactness the formulas obtained by the application of the method of conformal transformations meet the accuracy requirements of modern engineering.

7.2. Electrostatic Field of a Plane-electrode Low-mu Triode. The field of a low-mu triode may be determined by a method outlined by Maxwell.[1] Vacuum tubes had not yet been invented in Maxwell's time, but his analysis of the electrostatic field about a shielding screen of parallel wires is readily applied to the problem of the triode field.

The field analysis is based upon the Z-plane configuration shown in Fig. 7.1*a*. This consists of two line charges located within a large cylinder. One line charge is located at the origin and has a linear-charge density of $+q_c$. The other is located at the point (1,0) and has a linear-charge density of $+q_g$. The field at great distances from these lines is nearly circular and may be fitted to that of a circular electrode whose radius is large compared with the distance between the line charges. It may be seen that the Z-plane representation represents a simple tube

[1] MAXWELL, J. CLERK, "Electricity and Magnetism," 3d ed., Vol. I, Sec. 203, Cambridge, London, 1904.

with a cathode wire at the origin and a single grid wire at the point (1,0) surrounded by a circular plate. This simple tube has electrical characteristics the same as those of the plane-electrode and cylindrical-electrode structures that may be derived from it.

To obtain the field within the plane-electrode tube it is necessary to obtain an expression for the field in the Z plane of Fig. 7.1*a* and then transform this expression by the logarithmic transformation to fit that of the electrode configuration of Fig. 7.1*b*, which closely represents the structure of a practical tube. The potential at any point in the Z

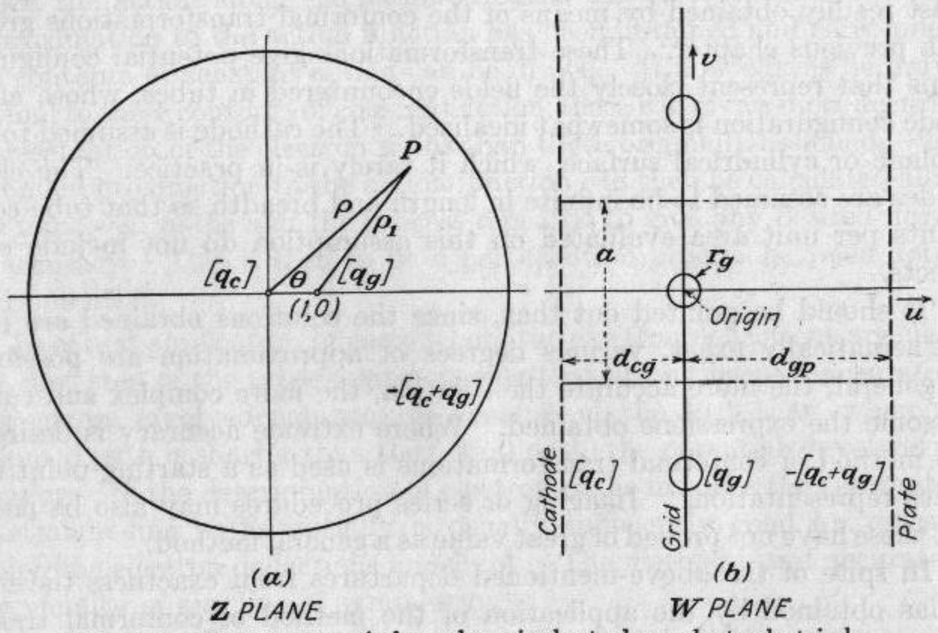


FIG. 7.1.—Elementary triode and equivalent plane-electrode triode.

plane is given by the sum of the potentials resulting from each of the line charges. Polar coordinates will be used in the Z-plane relations.

The potential at any point (ρ,θ) is given by

$$V_z = -\frac{q_g}{2\pi\epsilon_0}\ln\rho_1 - \frac{q_c}{2\pi\epsilon_0}\ln\rho + C \tag{7.1}$$

where ρ_1 is the distance from the point in question to the grid-wire charge at the point (1,0) and C is a constant that adjusts the level of potential, and *ln* will be used hereafter to denote the natural logarithm. Making use of the law of cosines,

$$V_z = -\frac{q_g}{4\pi\epsilon_0}\ln(\rho^2 + 1 - 2\rho\cos\theta) - \frac{q_c}{4\pi\epsilon_0}\ln\rho^2 + C \tag{7.2}$$

The logarithmic transformation with a suitable coefficient will be used. The coefficient is selected so that in the plane-electrode structure of Fig. 7.1*b* the grid wires are spaced a distance a.

$$W = \frac{a}{2\pi}\ln Z \tag{7.3}$$

The component parts of this equation are

$$u = \frac{a}{2\pi}\ln\rho \quad \text{or} \quad \rho = \epsilon^{\frac{2\pi u}{a}} \tag{7.4}$$

$$v = \frac{a\theta}{2\pi} \quad \text{or} \quad \theta = \frac{2\pi v}{a} \tag{7.5}$$

in which u and v are the real and imaginary parts of W, respectively, and ρ and θ are the polar coordinates in the Z plane. Making the above substitutions in Eq. (7.2),

$$V_w = -\frac{q_g}{4\pi\epsilon_0}\ln\left(\epsilon^{\frac{4\pi u}{a}} + 1 - 2\epsilon^{\frac{2\pi u}{a}}\cos\frac{2\pi v}{a}\right) - \frac{q_c}{4\pi\epsilon_0}\ln\epsilon^{\frac{4\pi u}{a}} + C \tag{7.6}$$

The above expression gives very closely the potential inside of a plane triode. Examination of its form will show that the equipotential lines in the vicinity of the origin and the points $(0,\pm na)$ are circles, one set of which may be fitted to the grid wires. For large positive and negative values of u the equipotentials are almost planes that may be fitted to the plate and cathode planes, respectively. The general potential expression of Eq. (7.6) gives potential in terms of the charges q_g, q_c, and the constant C. For application it is also necessary to evaluate these constants in terms of the electrode potentials.

To evaluate the constants of Eq. (7.6) let the plate plane be located at $u = +d_{gp}$ where $d_{gp} \geqq a$. When this relation between d_{gp} and a holds, the second and third terms of the argument of the first logarithm will be less than 1 per cent of the first term and may thus be neglected. Making the substitution $u = +d_{gp}$ into Eq. (7.6),

$$V_p = -\frac{d_{gp}q_g}{a\epsilon_0} - \frac{d_{gp}q_c}{a\epsilon_0} + C \tag{7.7}$$

Let the cathode plane be located at $-d_{cg}$, where $d_{cg} \geqq a$. In this case the first and third terms of the first logarithm argument will be small compared with 1 so that the first term is substantially zero. Making the substitution $v = -d_{cg}$ into Eq. (7.6),

$$V_c = 0 + \frac{d_{cg}q_c}{a\epsilon_0} + C \tag{7.8}$$

Let the grid wires be located at the points $(0,\pm na)$ and be of radius r_g and potential V_g. If $r_g \leqq \frac{a}{20}$, the potentials at points $(0,r_g)$ and $(r_g,0)$

differ by only a few per cent. Into Eq. (7.6) substitute the values $u = 0$, $v = r_g$.

$$V_g = -\frac{q_g}{2\pi\epsilon_0} \ln\left(2 \sin\frac{\pi r_g}{a}\right) + C \tag{7.9}$$

If the cathode potential is set at a reference level of potential of zero, then

$$C = -\frac{d_{cg}q_c}{a\epsilon_0} \tag{7.10}$$

from which Eqs. (7.7) and (7.9) become

$$V_p = -\frac{d_{gp}q_g}{a\epsilon_0} - \frac{(d_{cg} + d_{gp})q_c}{a\epsilon_0} \tag{7.11}$$

and

$$V_g = -\frac{\ln\left(2 \sin\frac{\pi r_g}{a}\right) q_g}{2\pi\epsilon_0} - \frac{d_{cg}q_c}{a\epsilon_0} \tag{7.12}$$

It is already possible to obtain the amplification factor of the tube from Eqs. (7.11) and (7.12). The amplification factor of a tube is the ratio of the plate voltage to the negative of grid voltage for a condition of cutoff.[1] In terms of the electric field within the tube, cutoff exists when the gradient of potential at the cathode is zero, which in turn occurs when the cathode charge is zero. If q_c is made zero in the above two equations and the ratio taken as indicated,

$$\mu = \frac{-2\pi d_{gp}}{a \ln\left(2 \sin\frac{\pi r_g}{a}\right)} \tag{7.13}$$

If Eqs. (7.11) and (7.12) are solved for q_c and q_g in terms of V_p and V_g and the expressions simplified by use of Eq. (7.13),

$$q_c = -\frac{\epsilon_0 a(V_p + \mu V_g)}{(d_{gp} + d_{cg} + \mu d_{cg})} \tag{7.14a}$$

and

$$q_g = \frac{\epsilon_0 a\mu[(d_{gp} + d_{cg})V_g - d_{cg}V_p]}{d_{gp}(d_{gp} + d_{cg} + \mu d_{cg})} \tag{7.14b}$$

The expression for μ of Eq. (7.13) given above is the simplest expression that adequately gives the amplification factor of a plane-electrode triode. Examination of this expression shows that the amplification factor increases as the grid-plate distance increases. This is in accord

[1] For a more general definition of the amplification factor see the chapter on Triode Characteristics.

with physical reasoning since the more remote the plate is the more influence the grid has. Amplification factor is also increased if the grid-wire spacing is decreased since this makes the grid more effective in controlling the off-cathode gradient of potential. Amplification is also increased if the grid-wire radius is increased, as would be expected. It will be observed that according to Eq. (7.13) the amplification factor is independent of the cathode-grid distance. This is approximately true as long as the approximations made in deriving the expression are not exceeded, *i.e.*, as long as the cathode-grid distance is not less than the grid-wire spacing. This may be understood by considering that the cathode charge is zero at cutoff. Thus for a cutoff condition all the flux lines originating on the grid terminate on the plate, and though some of them start toward the cathode they turn and end on the plate so that as long as the cathode is not too close to the grid the field pattern is not disturbed and the amplification factor is independent of cathode-grid distance. This interpretation will be discussed further in connection with equipotential contours and potential-profile plots.

Contour Representation of Potential Field. The form of the potential field resulting from the equations developed above may be best studied by examining the plots of the equipotential lines. A group of these equipotential contour plots of a typical plane triode are shown in Fig. 7.2 for various potentials. The contours of Fig. 7.2*a* show the field configuration for the case of the grid biased beyond cutoff. It will be observed that the gradient of potential at the cathode is negative. In the line of the grids the potential is increasingly negative in moving from cathode to grid. Along this same line the potential is increasingly positive in moving from grid to plate. Along the line from cathode to plate midway between grid wires the potential is at first negative and then positive. The dotted lines shown represent the boundary between the various types of equipotential lines. In the area within the dotted lines including the grid wires the equipotential lines are closed curves about the grid wires. In the other areas the equipotential lines run continuously from one section of the triode to adjoining sections, always on one side of the grid plane.

The other equipotential plots may be interpreted in a similar fashion. It will be observed that all the plots have some features in common. The equipotential lines close to the grid wires are nearly circular in all cases. The equipotential lines close to the plate and cathode are nearly straight lines. The equipotential lines may be divided into two groups, those which completely enclose the grid and those which run along from section to section. It will be observed that in some cases the equipotential lines of the second type listed above cross the grid plane between

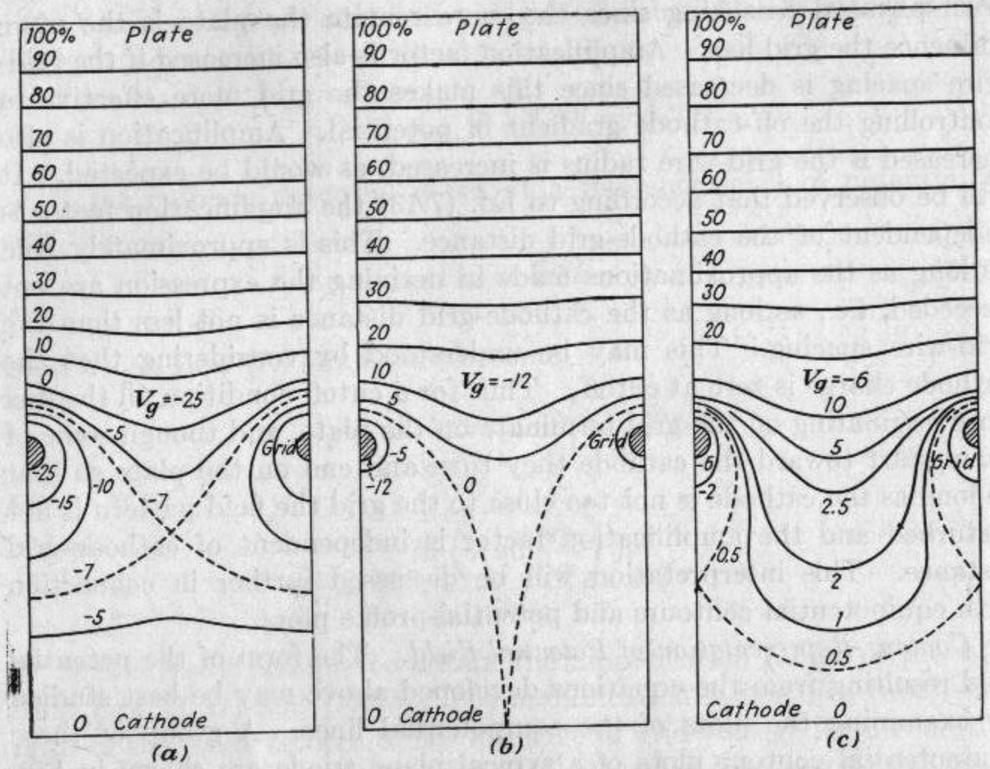


FIG. 7.2.—Equipotential contours in the plane-electrode triode: (*a*) grid beyond cutoff potential; (*b*) grid at cutoff potential; (*c*) grid negative at half cutoff value; (*d*) grid at zero potential; (*e*) grid positive.

grid wires but always pass the grid wires on the same side in moving from section to section. This behavior is observed in the case of the 10 per cent contour for $V_g = 0$. In all the cases shown the equipotentials are bowed toward the cathode through the grid plane. Only if the grid is more positive than the potential that gives a uniform positive potential gradient from plate to grid will the contours that cross the grid plane be bowed toward the plate. The equipotential plots shown in Fig. 7.2 were obtained with an electrolytic-trough model of potential. The equipotentials calculated from Eq. (7.6) would be almost the same in shape. For comparison a contour plot calculated by Eq. (7.6) is shown in Fig. 7.3. This plot represents an extreme condition of potential and dimensions. The grid-wire radius is $\frac{1}{20}$ of the grid-wire spacing. It will be observed that the grid-wire contour is not quite circular. It is of proper width in the plane of the grid wires but is longer in the direction at right angles to this. Because of this distortion of shape, which increases as the ratio of grid-wire diameter to grid-wire spacing increases, the formula for the amplification factor of Eq. (7.13) becomes inaccurate when the above ratio, known as the *screening fraction*, becomes greater than $\frac{1}{10}$. In the following section a more accurate formula is given, which is good up to screening fractions of $\frac{1}{6}$.

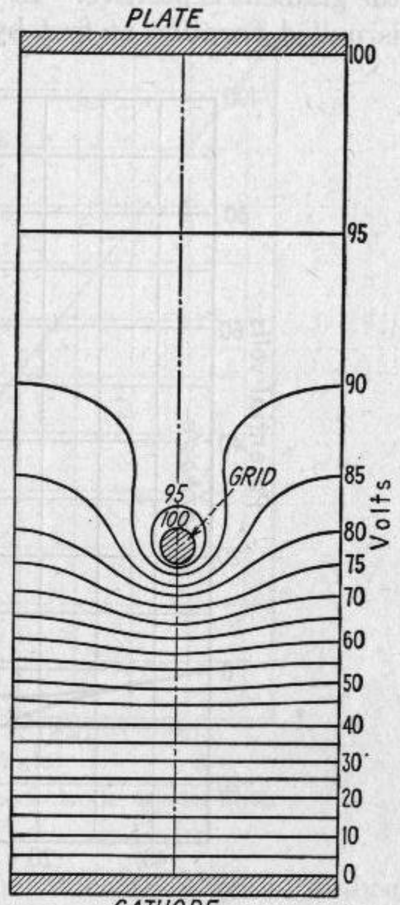


FIG. 7.3.—Equipotential contours in a plane-electrode triode with equal positive grid and plate potentials.

Profile Representation of Potential Field. The potential fields of a low-mu triode may also be studied by reference to profile representations of potential. These curves show how the potential varies along certain lines within the tube. The most common profile representations are shown along lines running from cathode to plate. In particular, two profiles are particularly informative. These are the profiles through the grid wire in a line running from cathode to plate at right angles to each of the latter, and in a line midway between grid wires. Such profiles are shown in Fig. 7.4.

In Fig. 7.4*a* are shown the profiles for a condition of a tube biased beyond cutoff. Here it is seen that the gradient of potential at the

cathode is negative, thus making it impossible for electrons to leave the cathode. This is true because most electrons that do succeed in getting away because of some initial velocity are driven back by the negative gradient of potential. In the line of the grids the potential goes strongly negative until it reaches the negative grid potential. Beyond the grid the gradient is positive. In the line between the grid wires the potential is pulled negative at first by the presence of the negative grid, and it

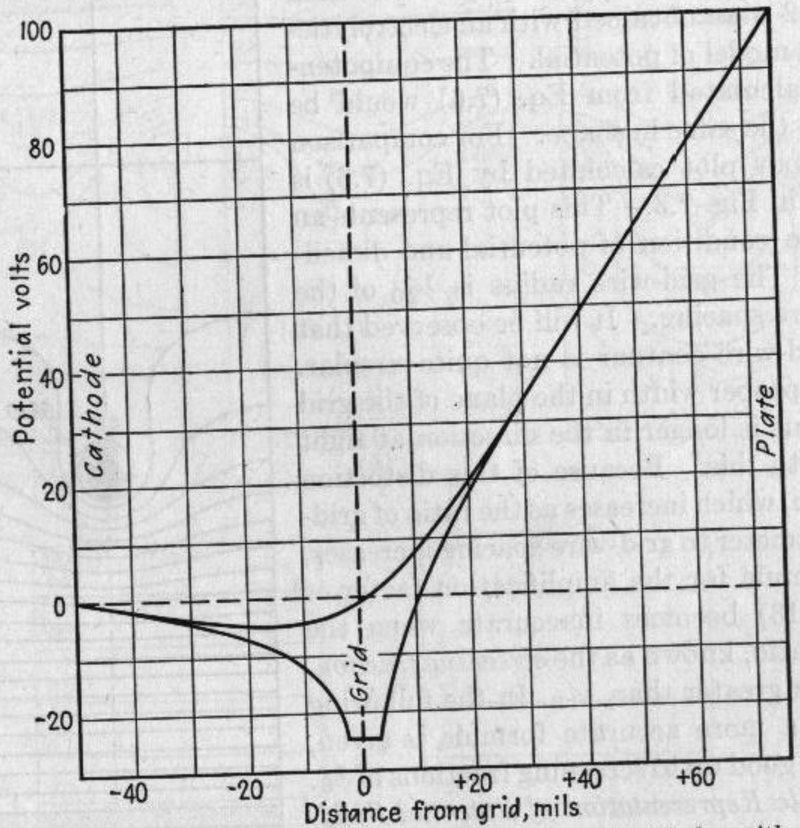


FIG. 7.4*a*.—Potential profiles of a plane-electrode triode, with grid at twice the cutoff value of potential.

then becomes positive. It will be observed that the potential profiles are straight lines near the cathode and also near the plate. Further, it is only in the vicinity of the grid that there is a great variation in the value of potential in moving parallel to the cathode and plate planes.

In Fig. 7.4*b* are shown the potential profiles for the case of the grid biased to approximately cutoff. Here it is seen that the gradient at the cathode is zero. In the line of the grid wires the potential first goes negative to the value of grid potential and then positive. In the line between the grid wires the potential becomes increasingly positive in moving from cathode to plate. In this representation the amplification

factor of the tube is given by the ratio of positive plate to negative grid potential. It is evident from these profiles why the amplification factor is independent of the cathode-grid distance provided that this is not too small. Up to a distance of about half the cathode-grid spacing the potential on both profiles is substantially zero for the particular dimensions shown. Hence in this particular case a cathode at zero potential

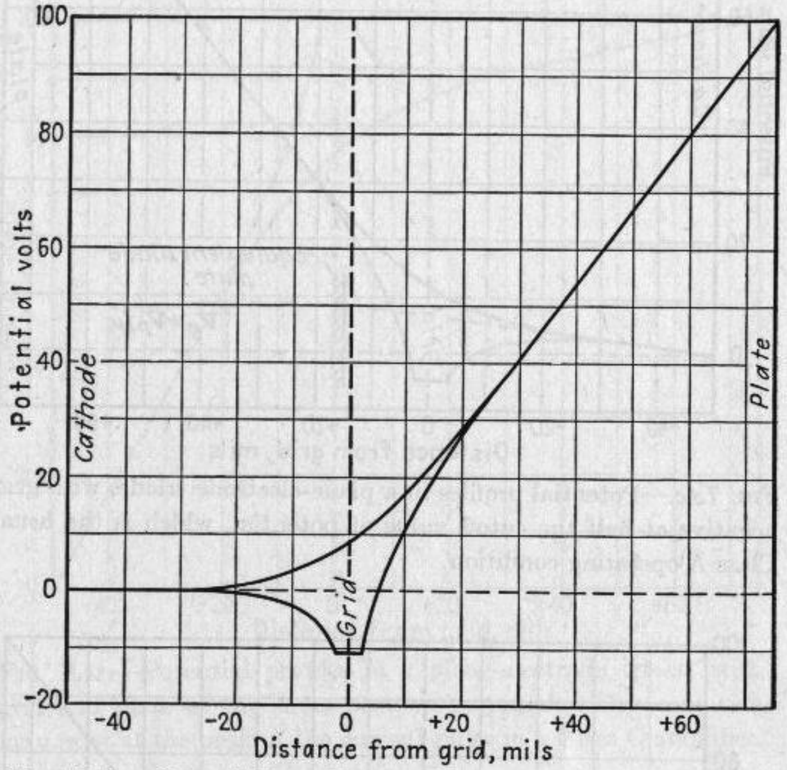


FIG. 7.4*b*.—Potential profiles of a plane-electrode triode, with grid at the cutoff value of potential.

could be put at any distance greater than half the cathode-grid distance shown without changing the shape or position of the potential profiles to the right of the profiles. The curves of Fig. 7.4*b* show the potential conditions that will just allow current to flow.

In Fig. 7.4*c* are shown profiles for a negative grid potential greater than that which gives the cutoff condition. Here the gradient of potential at the cathode is positive even though the grid is negative. The curves shown represent the potentials that would exist in the absence of current, say in a cold tube. Although this condition of potential would permit current to flow, the actual flow would depress the profiles in the vicinity of the cathode, as will be described in a later chapter. In Fig. 7.4*d* the grid is at zero potential, and it is now possible for electrons to reach the grid, which has not previously been possible for nega-

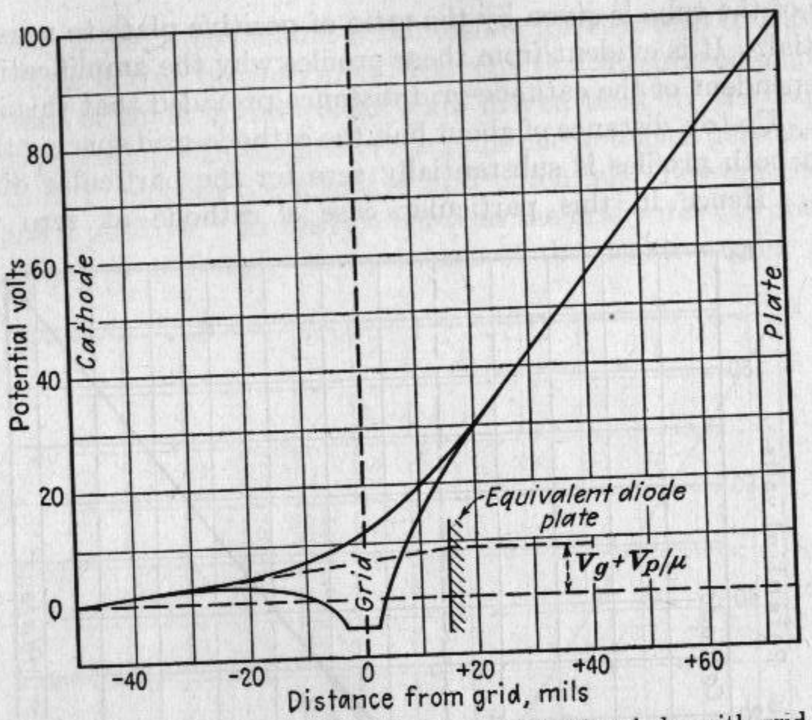


FIG. 7.4*c*.—Potential profiles in a plane-electrode triode, with grid negative at half the cutoff value of potential, which is the usual Class A operating condition.

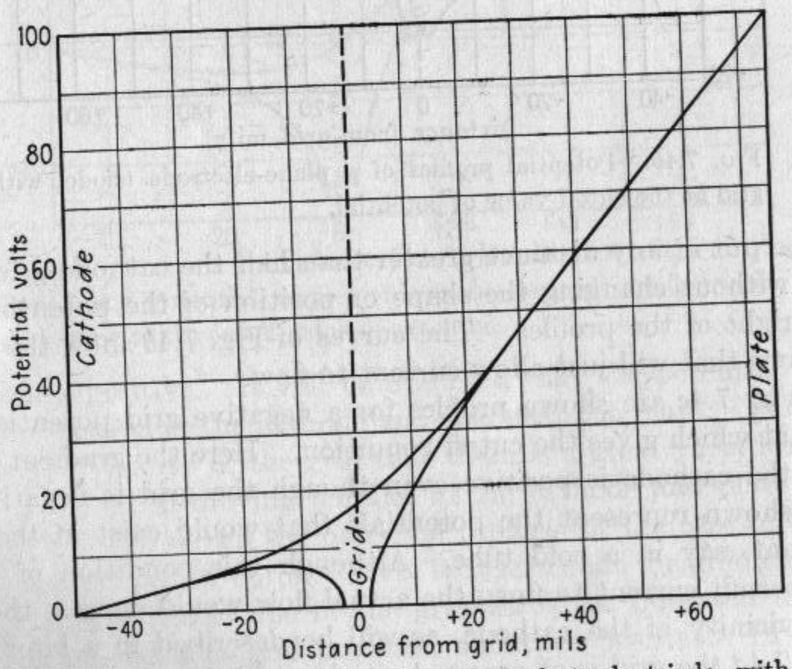


FIG. 7.4*d*.—Potential profiles in a plane-electrode triode, with grid at zero potential.

tive grid potentials. In Fig. 7.4*e* is shown an extreme condition of positive grid potential. Here the grid is as positive as the plate. This condition may be reached at the peak of the cycle in Class C power amplifiers.

In all cases the profiles are straight lines in the vicinity of the cathode and plate. For a condition of grid potential more negative than that

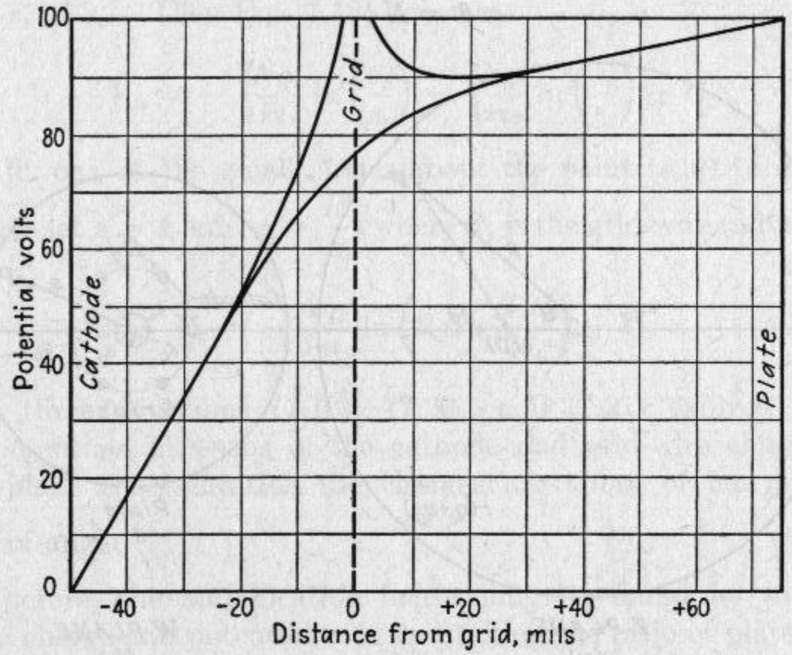


FIG. 7.4*e*.—Potential profiles in a plane-electrode triode, with grid and plate at the same positive potential. This condition may exist at the peak of the current pulse in a Class C amplifier.

of cutoff the slope of the straight-line portion at the cathode is negative. Above cutoff it is positive. The general form of the profiles corresponds to that which one would expect from a deformed elastic membrane. In each case the grid pushes a hole in what would otherwise be a straight-line profile from cathode to plate. Curvature requirements are met here. It will be observed that when one profile is concave upward the other is concave downward.

7.3. Electrostatic Field of a Low-mu Cylindrical-electrode Triode. The same fundamental tube configuration as was used for the plane-electrode triode in Fig. 7.1*a* can be used to develop the cylindrical-electrode triode. In this case, however, the transformation equation takes the form

$$W = s_g Z^{1/N} \tag{7.15}$$

to give the electrode arrangement of a cylindrical triode with N grid wires as shown in Fig. 7.5*b*. Let the coordinates in the Z plane be ρ

and θ and in the W plane s and ϕ; let s_g be the radius of the grid-wire circle. Then the components of the transformation equation in polar coordinates are

$$\rho = \left(\frac{s}{s_g}\right)^N \tag{7.16}$$

and

$$\theta = N\phi \tag{7.17}$$

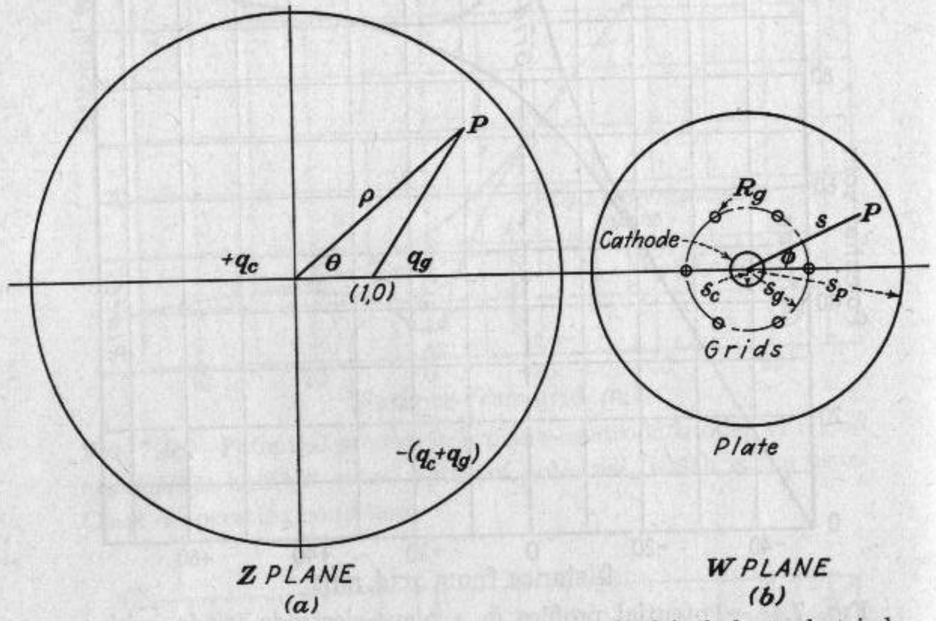


FIG. 7.5.—Elementary triode and equivalent cylindrical-electrode triode.

As before, the equation for the potential at any point in the Z plane is

$$V_z = -\frac{q_g}{4\pi\varepsilon_0}\ln(\rho^2 + 1 - 2\rho\cos\theta) - \frac{q_c}{4\pi\varepsilon_0}\ln\rho^2 + C \tag{7.2}$$

Substitution of the component transformation equations gives

$$V_w = -\frac{q_g}{4\pi\varepsilon_0}\ln\left[\left(\frac{s}{s_g}\right)^{2N} + 1 - 2\left(\frac{s}{s_g}\right)^N\cos N\phi\right] - \frac{2q_c}{4\pi\varepsilon_0}\ln\left(\frac{s}{s_g}\right)^N + C \tag{7.18}$$

This gives the equation of a potential field in which the contours are circles close to the origin and at great distances from the origin. The contours are also circles about the points $\left(s = s_g, \phi = \frac{2\pi k}{N}\right)$, where k assumes integral values from zero to N.

The three sets of circles can be fitted to the cathode, plate, and grid

wires, respectively. To fit the cathode to one of the circles close to and about the origin let $s = s_c \ll s_g$. Then

$$V_c = 0 - \frac{2q_cN}{4\pi\varepsilon_0}\ln\left(\frac{s_c}{s_g}\right) + C \tag{7.19}$$

To fit the large circles centered at the origin to the plate electrode let $s = s_p \gg s_g$. Then Eq. (7.18) becomes

$$V_p = -\frac{2Nq_g}{4\pi\varepsilon_0}\ln\left(\frac{s_p}{s_g}\right) - \frac{2Nq_c}{4\pi\varepsilon_0}\ln\left(\frac{s_p}{s_g}\right) + C \tag{7.20}$$

To fit one of the small circles about the point $(s_g,0)$ to one of the grid wires let $s = s_g$ and $\phi = \frac{R_g}{s_g}$, where R_g is the grid-wire radius. Then

$$V_g = -\frac{2q_g}{4\pi\varepsilon_0}\ln\left(2\sin\frac{NR_g}{2s_g}\right) + C \tag{7.21}$$

The three equations (7.19), (7.20), and (7.21) express the electrode potentials in terms of the cathode and grid-wire charges. For the W-plane representation the charges are those of one pie-shaped section of angle $\frac{2\pi}{N}$.

As before, the amplification factor may be found by setting the cathode charge and potential zero and taking the ratio of plate to negative grid potential. From this operation

$$\mu = -\frac{N\ln\left(\frac{s_p}{s_g}\right)}{\ln\left(2\sin\frac{NR_g}{2s_g}\right)} \tag{7.22}$$

The way in which the amplification factor of a cylindrical triode varies with the various electrode dimensions can be seen by inspection of Eq. (7.22). As the number of grid wires is increased, the amplification factor increases since N appears as a linear factor in the numerator and as a logarithmic factor in the denominator. This is to be expected from physical reasoning since an increase in the number of grid wires increases the effectiveness of the grid in controlling the off-cathode gradient and hence in controlling the current. The amplification factor increases as the plate radius is increased, also to be expected since this makes the plate less effective in controlling the current. The amplification factor increases as the radius of the grid-wire cylinder decreases since the factor in the numerator is more effective than that in the denominator. The

amplification factor also increases as the grid-wire radius is increased. Because of the way in which the factors combine, high amplification factors may be obtained more readily with cylindrical-electrode structures than with plane-electrode structures.

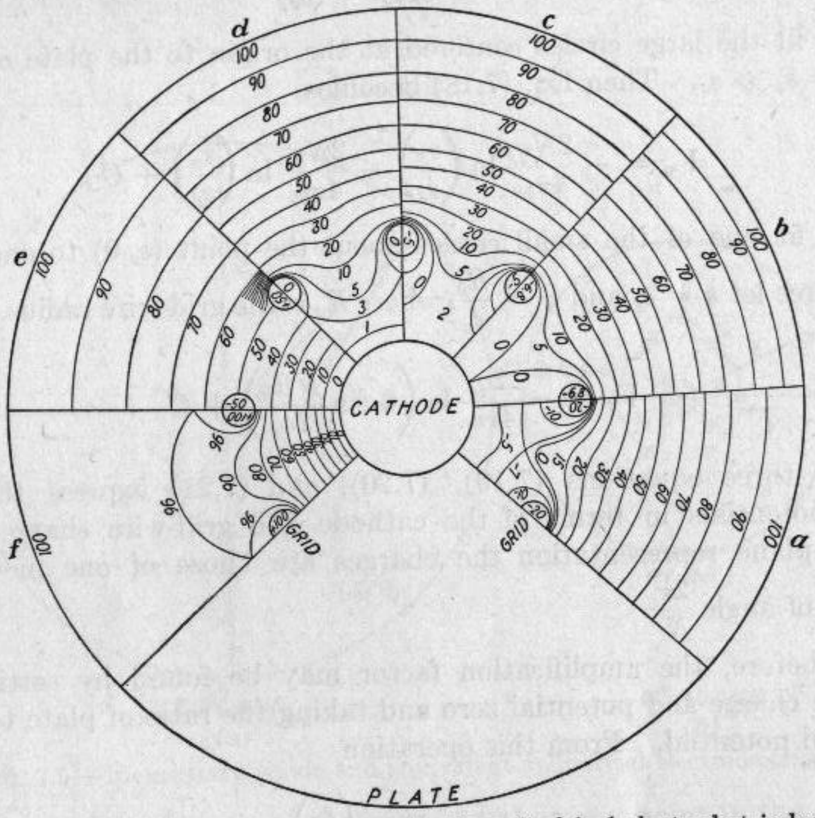


FIG. 7.6.—Equipotential contours in the cylindrical-electrode triode: (*a*) grid beyond cutoff potential; (*b*) grid at cutoff potential; (*c*) grid negative but above cutoff potential; (*d*) grid at zero potential; (*e*) grid at "natural" potential; (*f*) grid at positive plate potential.

As in the case of the parallel-electrode tube it is desirable to express the charges in terms of the electrode potentials. This is done by setting the cathode potential equal to zero and solving for q_c and q_g.

$$q_c = \frac{2\pi\varepsilon_0 \left[V_g \ln\left(\frac{s_p}{s_g}\right) - \frac{V_p}{N} \ln\left(2 \sin \frac{NR_g}{2s_g}\right)\right]}{N\left[\ln\left(\frac{s_p}{s_g}\right)\ln\left(\frac{s_c}{s_g}\right) + \frac{1}{N}\ln\left(\frac{s_p}{s_c}\right)\ln\left(2 \sin \frac{NR_g}{2s_g}\right)\right]} \tag{7.23}$$

$$q_g = -\frac{2\pi\varepsilon_0 \left[V_p \ln\left(\frac{s_c}{s_g}\right) + V_g \ln\left(\frac{s_p}{s_c}\right)\right]}{N\left[\ln\left(\frac{s_p}{s_g}\right)\ln\left(\frac{s_c}{s_g}\right) + \frac{1}{N}\ln\left(\frac{s_p}{s_c}\right)\ln\left(2 \sin \frac{NR_g}{2s_g}\right)\right]} \tag{7.24}$$

Potential Contours of a Cylindrical Triode. Contour representations of potential are shown in Fig. 7.6 for various relative electrode potentials. The contours of the cylindrical triode exhibit the same general characteristics as those of the plane-electrode triode. In each case the contours near any electrode have the same shape as the electrode. This means that the contours about the cathode and just inside the plate are

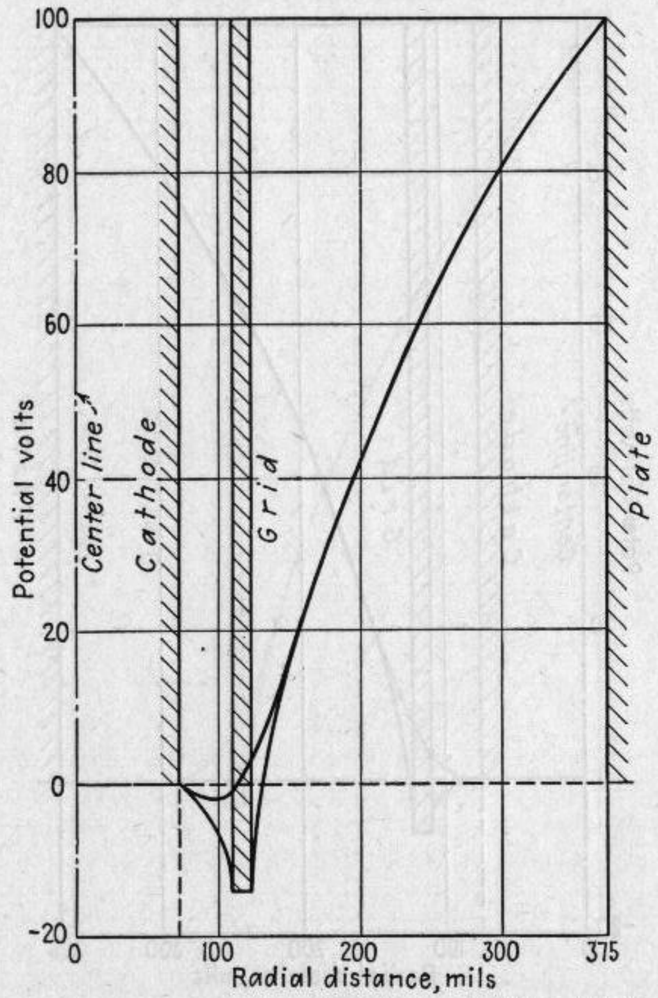


FIG. 7.7*a*.—Potential profiles in a cylindrical triode, with grid at twice the cutoff value of potential.

circles concentric about the center of the tube. There are also circles about each of the grid wires. The circles inside the plate have a nonlinear spacing in the case of the cylindrical triode. This is better understood by reference to the potential profiles.

Potential Profiles of a Cylindrical Triode. Reference to the potential profiles of Fig. 7.7 reveals several striking differences between the plane-electrode and cylindrical-electrode cases. Although the profiles have the

same general trend, they are characterized by different curvature characteristics. In the plane-electrode case the profiles through and between the grid wires coincided near the cathode and plate and were nearly straight lines there. In the case of the cylindrical triode they do again coincide but are curved instead of straight. The coincident profiles near the plate tend to be logarithmic in shape, as would be the case in a

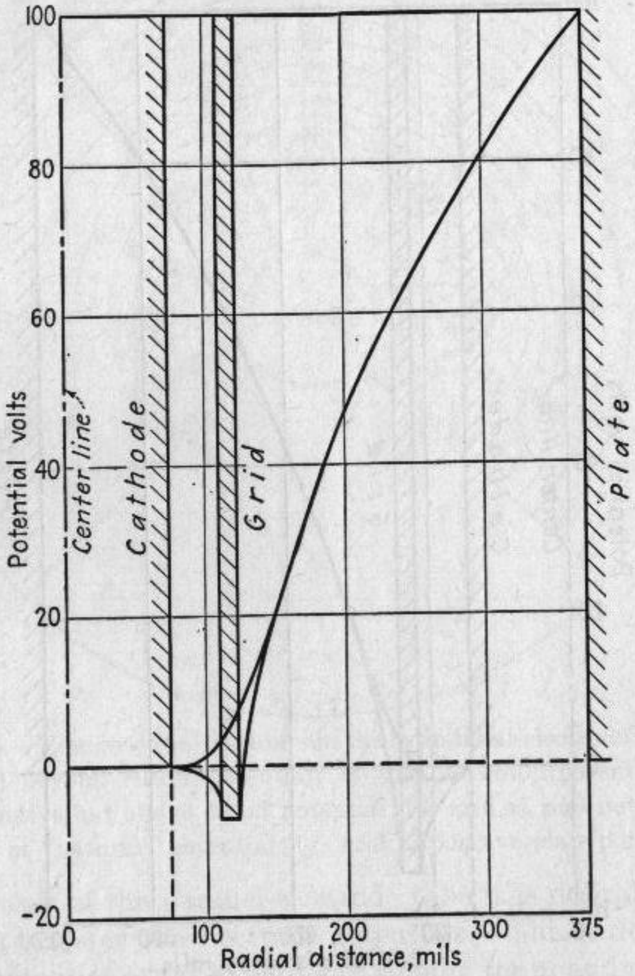


FIG. 7.7*b*.—Potential profiles in a cylindrical triode, with grid at the cutoff value of potential.

cylindrical diode. The same is true for the profiles near the cathode, though in the particular case of the relatively high-mu tube shown, the region in which the profiles coincide near the cathode is small because of the short cathode-grid distance.

Below cutoff in Fig. 7.7*a* the cathode gradient of potential is negative. At cutoff as in Fig. 7.7*b* it is zero, and it can again be seen that this condition of zero cathode gradient is independent of the grid-cathode

distance provided that the distance is not too small. In the particular figure shown any larger cathode-grid distance would not change the amplification factor of the tube. The remaining figures show the profiles for a grid negative, but above cutoff in Fig. 7.7*c*, for a grid at zero potential in Fig. 7.7*d*, and for an extreme condition of positive grid potential in Fig. 7.7*e*.

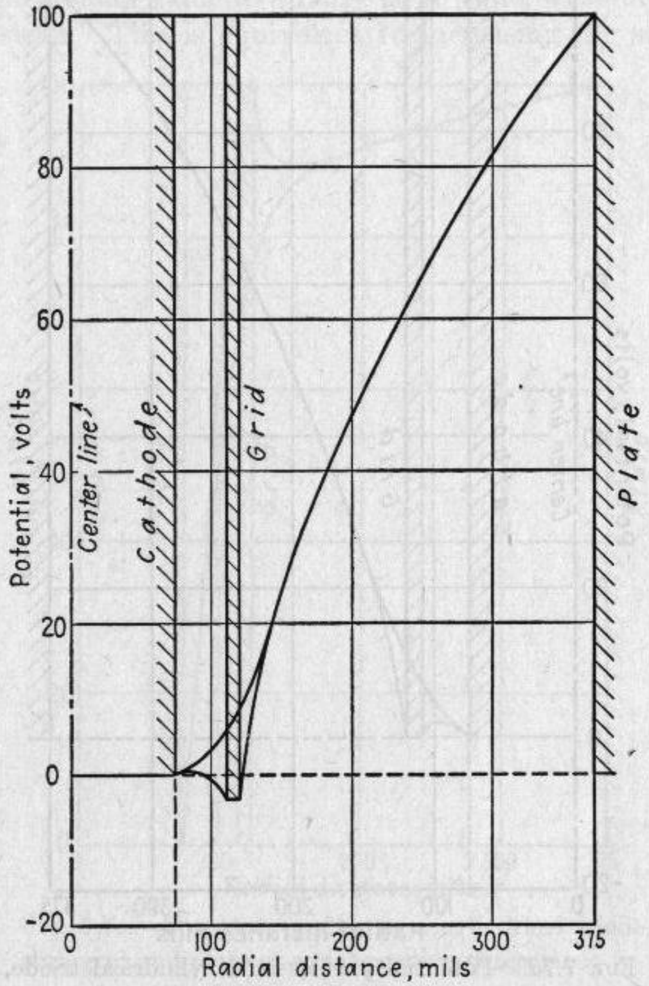


FIG. 7.7*c*.—Potential profiles in a cylindrical triode, with grid at half the cutoff value of potential.

The curvature conditions that were noted in the case of the plane-electrode triode are no longer valid in the case of the cylindrical triode. It is no longer true that if one profile is concave upward the other is concave downward. This follows from the fact that the Laplace equation for polar coordinates can no longer be interpreted so simply in terms of curvatures. If the coordinates were changed so that the profiles were plotted against the logarithm of r, then the curvature conditions

that held for the plane-electrode case would be fulfilled. This follows because with this change of variables the Laplace equation becomes identical with that for Cartesian coordinates. In Fig. 7.8 are shown the potential profiles of Fig. 7.7*a* plotted with a logarithmic scale of radius. It is seen that the profiles become straight lines in the vicinity of the

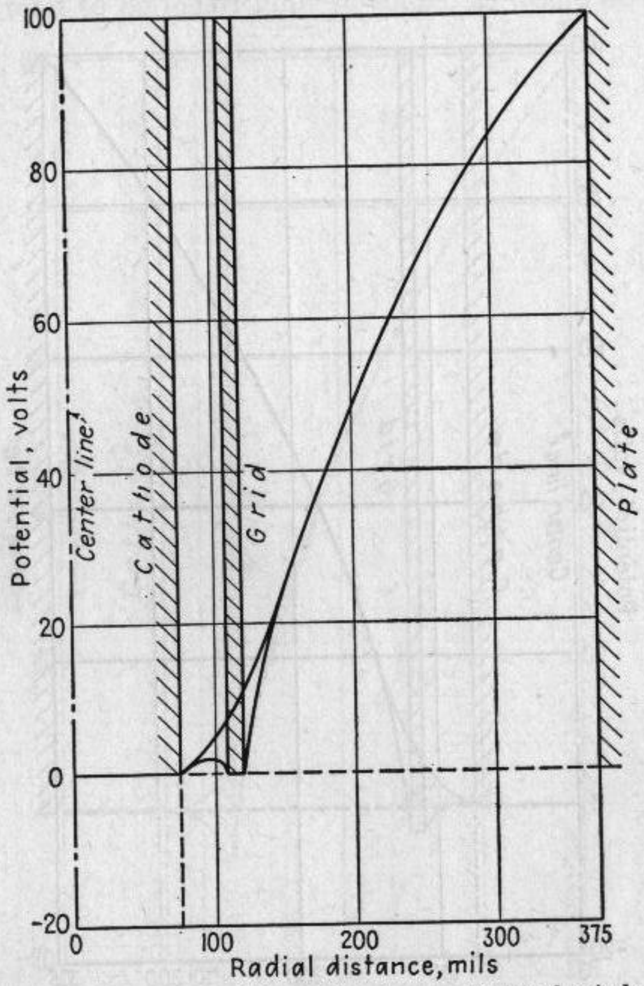


FIG. 7.7*d*.—Potential profiles in a cylindrical triode, with grid at zero potential.

cathode and plate and thus resemble the plane-triode profiles in this type of plot.

7.4. Analysis of the High-mu Triode. *Potential Contours and Profiles.* The method of Maxwell discussed in the previous sections has some limitations that make the results inaccurate when the attempt is made to apply them to a high-mu triode. In the previous analysis it was assumed that the equipotentials about the grid line charge in the fundamental tube of the Z-plane representation were circles concentric about

the line. This is very nearly true provided that the circles are not too large. But if we examine the shape the grid wire in the Z plane must have as determined by circles in Fig. 7.1*b* transformed back to the Z plane, it is found that departures from small circles centered about the point (1,0) are soon encountered as the grid-wire radius is increased. Consider the grid wires of Fig. 7.1*b*, and let the grid-wire radius be increased from a small value to a fairly large value without changing the other dimensions. This is equivalent to increasing the screening frac-

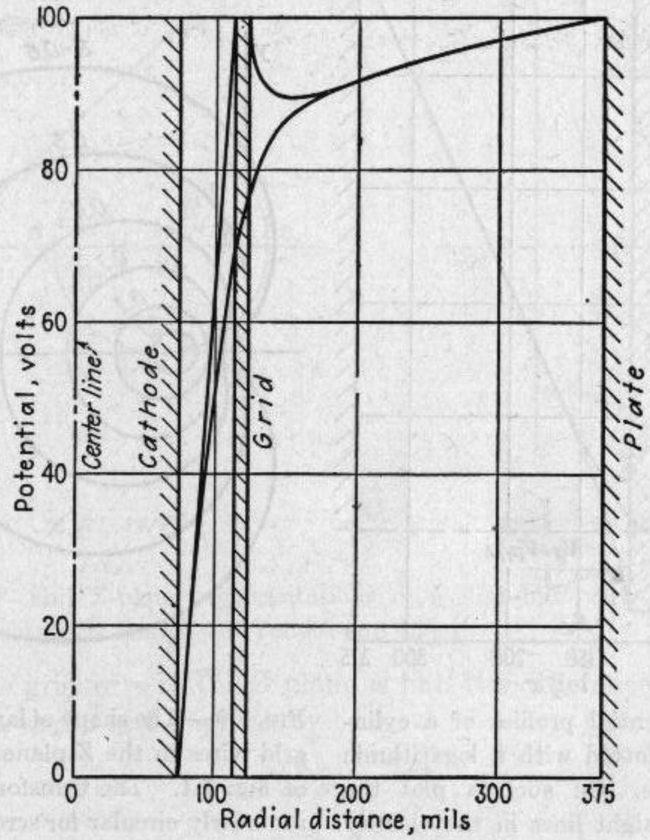


FIG. 7.7*e*.—Potential profiles in a cylindrical triode, with grid and plate at equal positive potentials.

tion and increases the mu of the tube. As the grid-wire radii are increased, the corresponding contours in the Z plane, which are at first small circles with centers at the point (1,0), become larger curves that are nearly circular in shape but that are shifted in position so that their centers are not at the point (1,0) but to the right of this point. This applies as the screening fraction is increased from 0.1 to 0.2. The progressive changes in shape encountered are shown in Fig. 7.9. In the analysis indicated in this section another line charge is introduced to take account of the shift in position of the circular grid-potential contour. As the screening frac-

tion is increased beyond 0.3, the transformed contour of the grid wire loses its circular shape, becoming dented on one side, and the improved analysis is no longer valid. The accuracy of the formulas developed can be extended, however, so that they may be used for tubes having screening fractions as low as $\frac{1}{6}$ instead of merely $\frac{1}{10}$. The resulting expressions are considered the most accurate simple expressions available.[1,2]

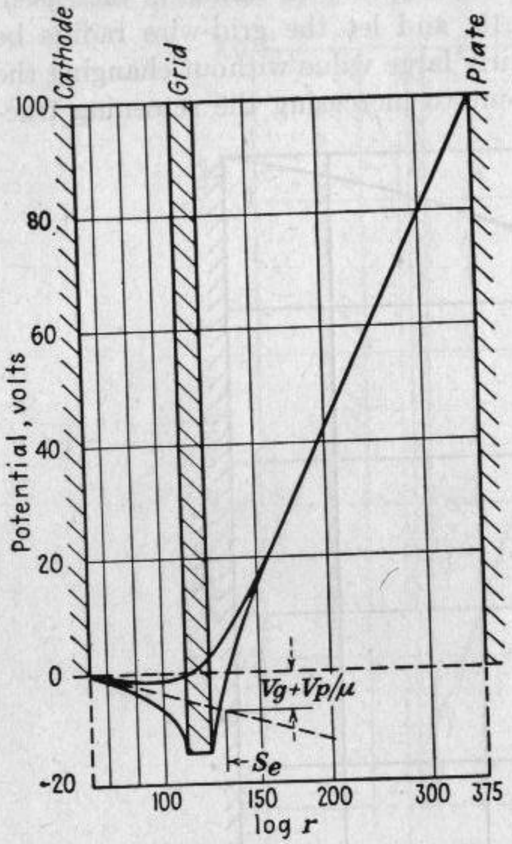


FIG. 7.8.—Potential profiles of a cylindrical triode plotted with a logarithmic scale of radius. On such a plot the profiles are straight lines in the vicinity of the cathode and plate. Also shown is the graphical construction for determining the equivalent diode radius.

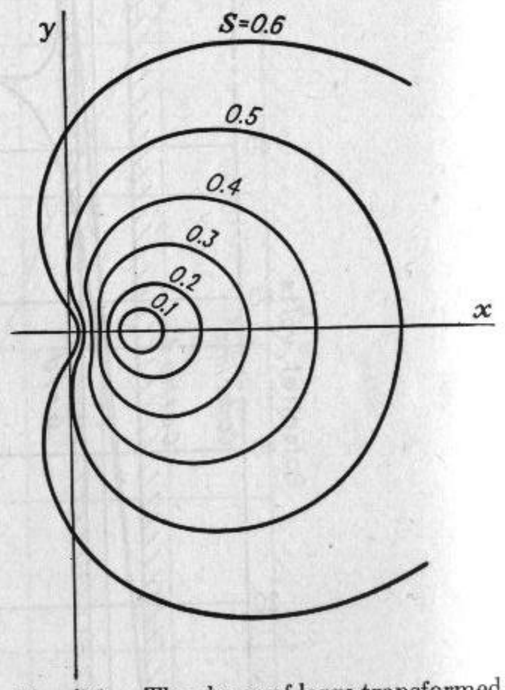


FIG. 7.9.—The shape of large transformed grid wires in the Z-plane representation of Fig. 7.1. The transformed grid wires are nearly circular for screening fractions as large as 0.2. Beyond this value, the grid contour is noncircular and cannot be represented by two line charges.

As before, use the plane-electrode transformation equations,

$$W = \frac{a}{2\pi} \ln Z = \frac{a}{2\pi} \ln \rho + \frac{i\theta a}{2\pi} \tag{7.3}$$

[1] VODGES, F. B., and F. R. ELDER, Formulas for the Amplification Constant for Three-element Tubes, *Phys. Rev.*, vol. 24, pp. 683–689, December, 1924.

[2] Dow, W. G., "Engineering Electronics," pp. 24–53, Wiley, New York, 1937.

which have the component transformation equations given by Eqs. (7.4) and (7.5), ρ being Z-plane radius. Then the points a and b in Fig. 7.10b which have the coordinates $(r_g,0)$ and $(-r_g,0)$ respectively, transform into the points α and β, which in Fig. 7.10a have the coordinates $(e^{\frac{2\pi r_g}{a}},0)$ and $(e^{-\frac{2\pi r_g}{a}},0)$. If the screening fraction is less than 0.16, the transformed grid wire is nearly a circle through these two points. The

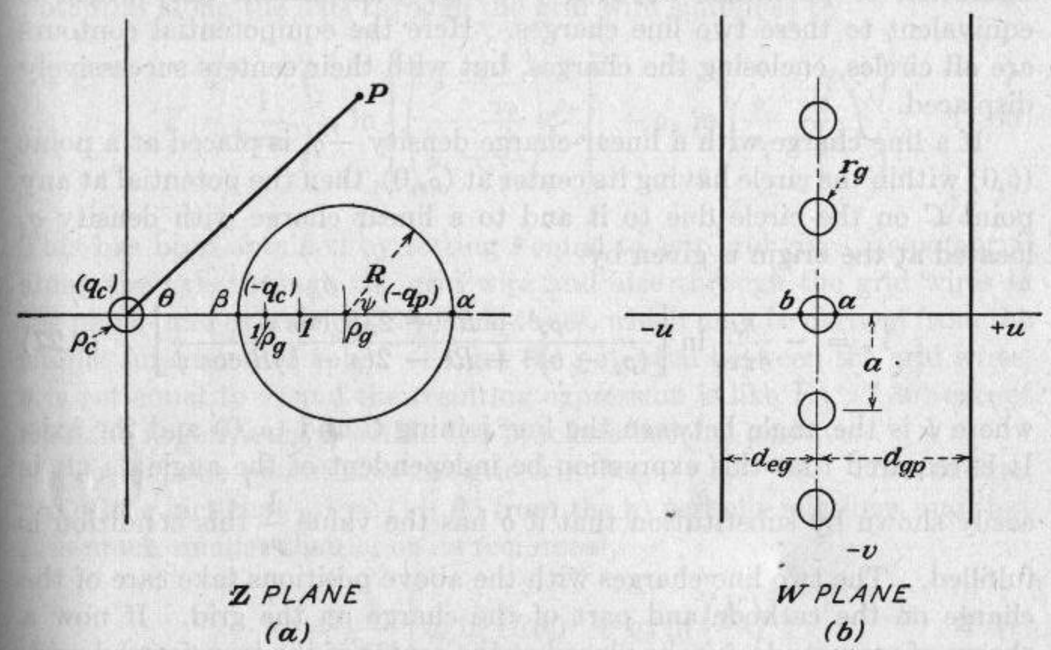


FIG. 7.10.—W- and Z-plane representations of a high-mu plane-electrode triode. This is the basis of the analysis of Vodges and Elder.

radius of the grid wire in the Z plane is half the difference of the ρ components of α and β and is given by

$$R = \sinh \frac{2\pi r_g}{a} \tag{7.25}$$

The location of the center of the grid-wire circle is given by the average of the values of α and β and is given by

$$\rho_g = \cosh \frac{2\pi r_g}{a} \tag{7.26}$$

It is now necessary to locate line charges so that this circle will be an equipotential contour for all combinations of cathode, grid, and plate potential.

In the analysis given for the low-mu triodes a single line charge was placed within the circle at its center. When the analysis is extended so that the circle is no longer small, this is not adequate. A line charge at

the center of the circle will still give an equipotential contour that will fit the circle if this charge is the only one present, which is not the case. However, it is also possible to find another position within the circle such that a line charge placed there together with a line charge at the origin with an equal charge of opposite sign will give an equipotential contour on the transformed grid-wire circle.[1] This follows from the well-known configuration of potentials about a two-wire transmission line, which is equivalent to these two line charges. Here the equipotential contours are all circles, enclosing the charges, but with their centers successively displaced.

If a line charge with a linear-charge density $-q_c$ is placed at a point $(b,0)$ within the circle having its center at $(\rho_g,0)$, then the potential at any point C on the circle due to it and to a linear charge with density q_c located at the origin is given by

$$V_c = -\frac{q_c}{4\pi\varepsilon_0}\ln\left[\frac{\rho_g^2 + R^2 + 2\rho R\cos\psi}{(\rho_g - b)^2 + R^2 - 2(\rho - b)R\cos\psi}\right] \tag{7.27}$$

where ψ is the angle between the line joining C and $(\rho_g,0)$ and the axis. It is required that this expression be independent of the angle ψ. It is easily shown by substitution that if b has the value $\frac{1}{\rho_g}$ this condition is fulfilled. The two line charges with the above positions take care of the charge on the cathode and part of the charge on the grid. If now a charge of magnitude $-q_p$ be placed at the center of the transformed grid-wire circle, it will be the source of flux lines that will extend in all directions, becoming radial at great distances and terminating on a large plate circle. Since any value of either q_c or q_p gives rise to an equipotential circle of radius R with center at ρ_g, this circle can be made the grid-wire circle for any combination of charges and hence of potentials. With this location of line charges it is easy to write the potential at any point in the tube.

At any point (ρ,θ) within the tube the potential is given by

$$V = -\frac{1}{4\pi\varepsilon_0}\left[q_c\ln\rho^2 - q_c\ln\left(\rho^2 + \frac{1}{\rho_g^2} - \frac{2\rho}{\rho_g}\cos\theta\right) - q_p\ln(\rho^2 + \rho_g^2 - 2\rho\rho_g\cos\theta) + C\right] \tag{7.28}$$

in which the constant C is introduced to adjust the level of the potential. In most tubes the cathode can be fitted to a small circle about the origin.

[1] This attack on the problem was first successfully applied by W. G. Dow.

Let $\rho = \rho_c$, $\theta = 0$; then

$$V_c = -\frac{1}{4\pi\varepsilon_0}\left[q_c\ln\rho_c^2 - q_c\ln\left(\rho_c - \frac{1}{\rho_g}\right)^2 - q_p\ln(\rho_c - \rho_g)^2 + C\right] \tag{7.29}$$

If the cathode potential be taken as zero and the value of C then obtained from Eq. (7.29) be substituted into Eq. (7.28), the expression for the potential along the axis through the grid wire simplifies to

$$V = \frac{1}{4\pi\varepsilon_0}\left\{q_c\ln\left[\left(\frac{\rho - \frac{1}{\rho_g}}{\rho_c - \frac{1}{\rho_g}}\right)\frac{\rho_c}{\rho}\right]^2 + q_p\ln\left(\frac{\rho - \rho_g}{\rho_c - \rho_g}\right)^2\right\} \tag{7.30}$$

This has been obtained by setting θ equal to zero and gives the potential along the axis through the grid wire and also through the grid wires in the plane- and cylindrical-electrode tubes, which may be derived from the simple fundamental tube. To get the potential between the grid wires, θ is set equal to π, and the resulting expression is like Eq. (7.30) except that the negative signs within the brackets become positive.

To find the potential of the grid wire let $\rho = \rho_g - R$. Then making use of the fact that $\rho_g^2 = 1 + R^2$ from the hyperbolic relations, and that ρ_c is much smaller than ρ_g or its reciprocal,

$$V_g = \frac{1}{4\pi\varepsilon_0}\left[q_c\ln(R\rho_c)^2 + q_p\ln\left(\frac{R}{\rho_g}\right)^2\right] \tag{7.31}$$

Similarly, to find the plate potential let $\rho = \rho_p$, and make use of the fact that ρ_c is much less than 1, which in turn is much less than ρ_p. Then

$$V_p = \frac{1}{4\pi\varepsilon_0}\left[q_c\ln(\rho_c\rho_g)^2 + q_p\ln\left(\frac{\rho_p}{\rho_g}\right)^2\right] \tag{7.32}$$

Equations (7.31) and (7.32) give the grid and plate potential in terms of cathode and grid charges. In order to calculate potential profiles it is desirable to know the charges in terms of the potentials. The above equations are readily rearranged to give this form.

$$q_c = 2\pi\varepsilon_0\left[\frac{V_g\ln\left(\frac{\rho_p}{\rho_g}\right) - V_p\ln\left(\frac{R}{\rho_g}\right)}{\ln(R\rho_c)\ln\left(\frac{\rho_p}{\rho_g}\right) - \ln\left(\frac{R}{\rho_g}\right)\ln(\rho_c\rho_g)}\right] \tag{7.33}$$

$$q_p = 2\pi\varepsilon_0\left[\frac{V_p\ln(R\rho_c) - V_g\ln(\rho_c\rho_g)}{\ln(R\rho_c)\ln\left(\frac{\rho_p}{\rho_g}\right) - \ln\left(\frac{R}{\rho_g}\right)\ln(\rho_c\rho_g)}\right] \tag{7.34}$$

Amplification Factor of a High-mu Plane-electrode Triode. Again the amplification factor is given by the ratio of plate to negative grid potential for zero cathode charge. It will be noticed that the cathode charge is a linear function of plate and grid potentials. In multielectrode tubes the cathode charge is a linear function of all the electrode potentials. In the case of a pentode, for instance, the cathode charge can be written

$$q_c = \frac{V_g + \frac{V_1}{\mu_1} + \frac{V_2}{\mu_2} + \frac{V_3}{\mu_3}}{D} \tag{7.35}$$

in which the various μ's indicate the relative effectiveness of the control grid and the electrode in question in controlling the off-cathode potential gradient. In the case of the triode considered here, the amplification factor is

$$\mu = -\frac{\ln\left(\frac{\rho_p}{\rho_g}\right)}{\ln\left(\frac{R}{\rho_g}\right)} \tag{7.36}$$

Substitution of values from Eqs. (7.3), (7.25), and (7.26) gives the amplification factor of a plane-electrode triode as

$$\mu = \frac{\left(\frac{2\pi d_{gp}}{a}\right) - \ln\cosh\left(\frac{2\pi r_g}{a}\right)}{\ln\coth\left(\frac{2\pi r_g}{a}\right)} \tag{7.37}$$

Making use of the definition of the screening fraction as the percentage of the area in the grid-wire plane occupied by the grid wires, numerically equal to $\frac{2r_g}{a}$, and denoting the screening fraction by S, the expression for the amplification factor can also be written

$$\mu = \frac{\left(\frac{2\pi d_{gp}}{a}\right) - \ln\cosh \pi S}{\ln\coth \pi S} \tag{7.38}$$

From this it is seen that the amplification factor depends upon only two factors, the screening fraction and the ratio of grid-plate distance to grid-wire spacing. The way in which the amplification factor varies with these two factors is shown in Fig. 7.11. It is seen that the amplification factor increases with the screening fraction and increases as the ratio of grid-plate to grid-wire spacing increases. The upper solid curve represents the limit of accuracy of the formula given by Eq. (7.38).

If the screening fraction is small, then Eqs. (7.37) and (7.38) reduce to the same expression that results from Eq. (7.13) so that these two different expressions give substantially the same numerical result when the screening fraction is less than $\frac{1}{10}$.

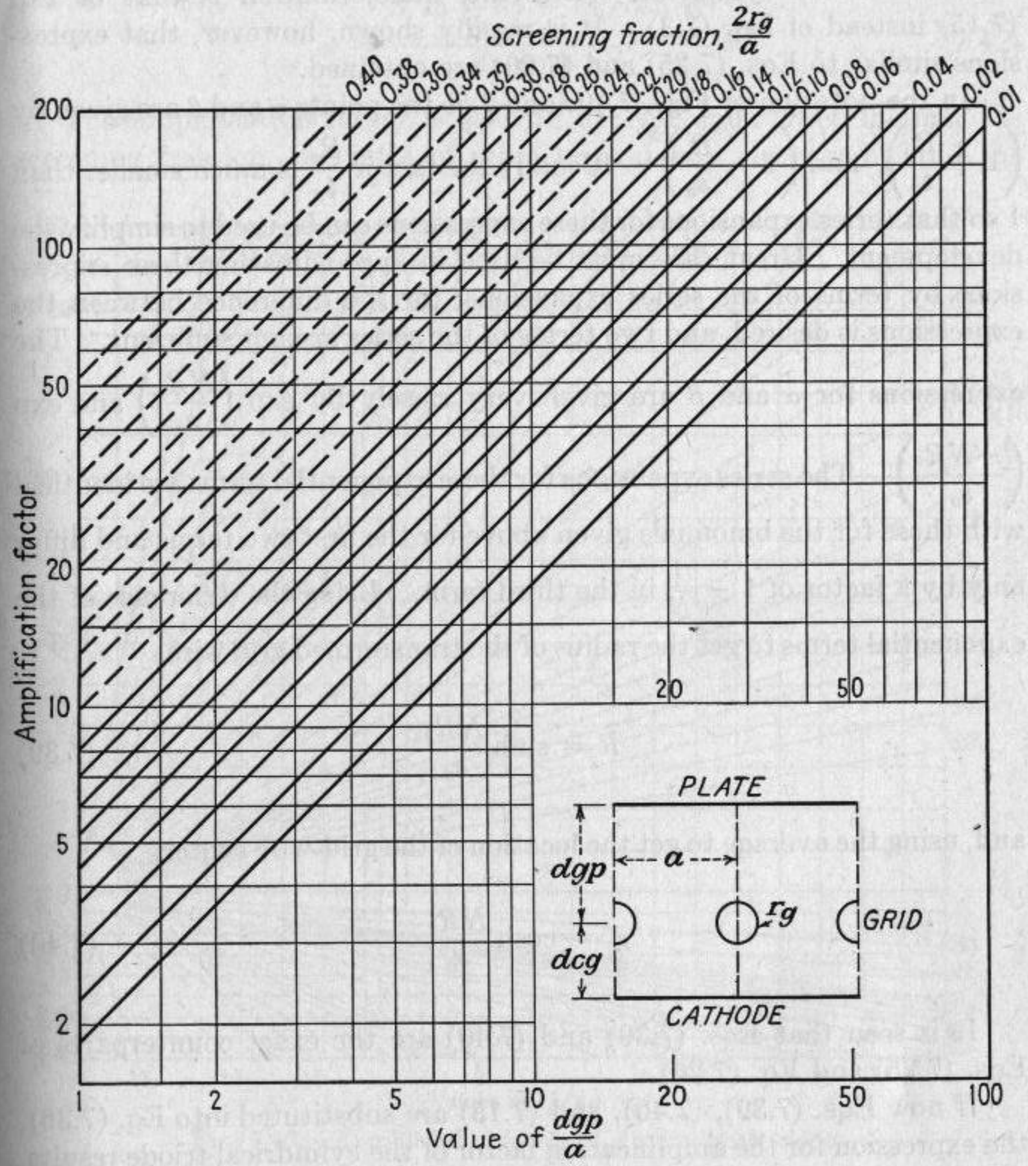


FIG. 7.11.—Chart giving the amplification factor of a plane-electrode triode. The solid curves were obtained from Eq. (7.38), which is accurate up to screening fractions of 0.16. Dotted curves were obtained from Eq. (7.71), which is accurate up to screening fractions of 0.4.

Amplification Factor of a High-mu Cylindrical Triode. When the transformation relations of Eqs. (7.16) and (7.17) are applied to Eq. (7.36),

the expression for the amplification factor of a cylindrical triode is obtained. This development requires some intermediate justification because for the cylindrical triode the transformation is that of Eq. (7.15) instead of Eq. (7.3). It is readily shown, however, that expressions similar to Eqs. (7.25) and (7.26) are obtained.

In the notation of Figs. 7.10*a* and 7.5*b* the points α and β are given by $\left(1 + \frac{R_g}{s_g}\right)^N$ and $\left(1 - \frac{R_g}{s_g}\right)^N$. Usually the factor $\frac{R_g}{s_g}$ is much smaller than 1 so that series expansions for these expressions can be used to simplify the development. Great care must be used in approximating these expressions by terms of the series expansions; for the difference between the expressions is desired, and two terms of the series are not sufficient. The expressions for α and β are given very closely by $\exp\left(\frac{NR_g}{s_g}\right)$ and $\exp\left(\frac{-NR_g}{s_g}\right)$. The series expansions for these exponential terms are identical with those for the binomials given above for the first two terms and differ only by a factor of $1 - \frac{1}{N}$ in the third term. Using the difference of the exponential terms to get the radius of the transformed grid wire,

$$R = \sinh \frac{NR_g}{s_g} \tag{7.39}$$

and, using the average to get the location of the grid-wire center,

$$\rho_g = \cosh \frac{NR_g}{s_g} \tag{7.40}$$

It is seen that Eqs. (7.39) and (7.40) are the exact counterparts of Eqs. (7.25) and Eq. (7.26).

If now Eqs. (7.39), (7.40), and (7.15) are substituted into Eq. (7.36), the expression for the amplification factor of the cylindrical triode results.

$$\mu = \frac{N \ln\left(\frac{s_p}{s_g}\right) - \ln \cosh\left(\frac{NR_g}{s_g}\right)}{\ln \coth\left(\frac{NR_g}{s_g}\right)} \tag{7.41}$$

Since the screening fraction for the cylindrical triode is $S = \frac{NR_g}{\pi s_g}$, the

above expression may be written

$$\mu = \frac{N \ln\left(\frac{s_p}{s_g}\right) - \ln \cosh \pi S}{\ln \coth \pi S} \tag{7.42}$$

The amplification factor is seen to depend upon three factors, the screening fraction, the ratio of plate to grid-wire circle radii, and the

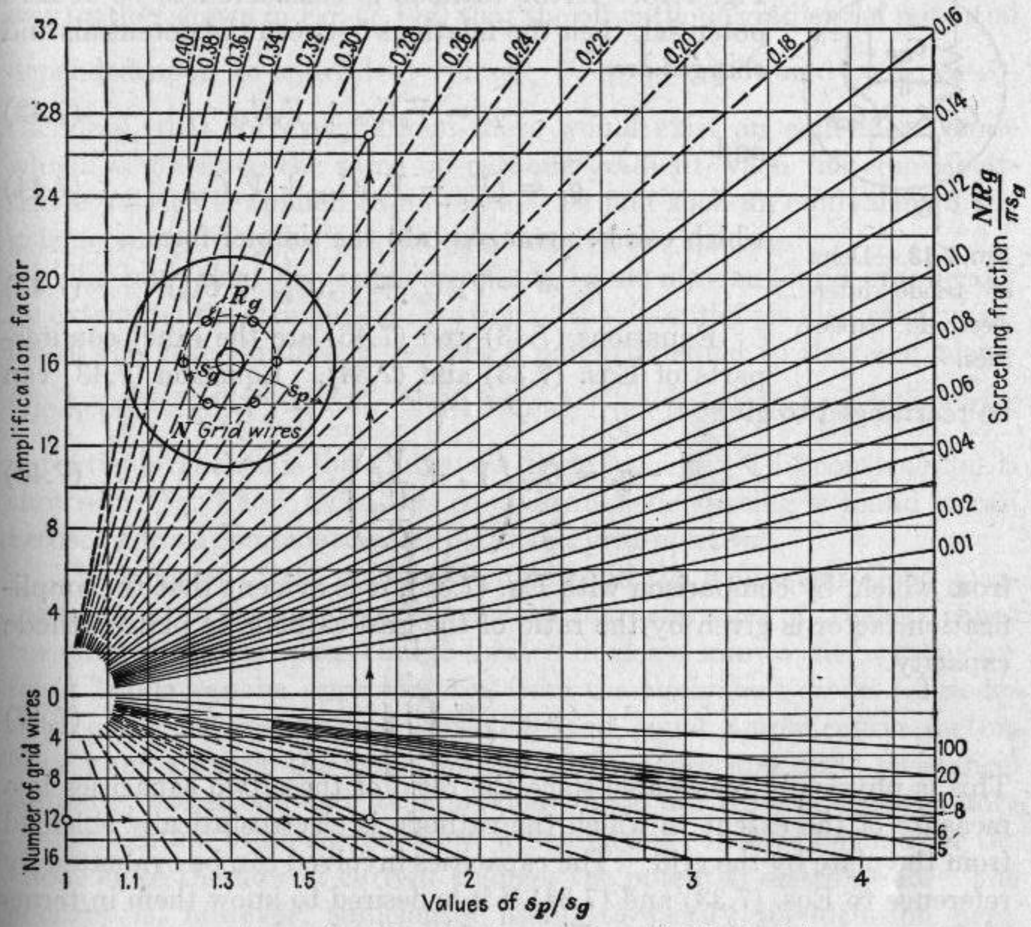


FIG. 7.12.—Amplification factor of a cylindrical triode.

number of grid wires. This is much more difficult to plot but may be done by using a number of axes each corresponding to a different number of grid wires. Such a plot of the amplification factor of the cylindrical triode is shown in Fig. 7.12. It is seen that the amplification factor increases with the screening fraction, with the number of grid wires, and with the ratio $\frac{s_p}{s_g}$. The upper solid curve represents the limit of an accuracy of about 2 per cent. The formulas of Vodges and Elder given

in Eqs. (7.38) and (7.42) are considered the best practical formulas available for the amplification factor of triodes from the standpoint of simplicity and accuracy.

7.5. The Equivalent Electrostatic Circuit of a Triode. Examination of Eqs. (7.33) and (7.34) shows that these expressions have the form of equations for a delta of capacities. A delta of capacities such as could be used to represent a triode is shown in Fig. 7.13. If the cathode is considered to be at zero potential, then the relations between the potentials and charges are

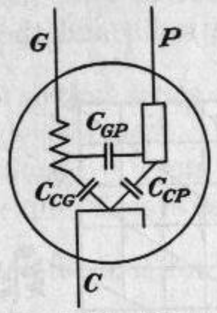


FIG. 7.13.—Delta of triode interelectrode capacities.

$$-q_c = V_g C_{cg} + V_p C_{cp} \tag{7.43}$$

and

$$q_p = (V_p - V_g) C_{gp} + V_p C_{cp} \tag{7.44}$$

which can be arranged into the simpler form

$$q_p = -V_g C_{gp} + V_p (C_{gp} + C_{cp}) \tag{7.45}$$

Equations (7.43) and (7.45) are the exact counterparts of Eqs. (7.33) and (7.34). Equation (7.43) can be rearranged to give

$$-q_c = C_{cg}\left(V_g + \frac{V_p}{\frac{C_{cg}}{C_{cp}}}\right) \tag{7.46}$$

from which, by comparison with Eq. (7.35), it is evident that the amplification factor is given by the ratio of the grid-cathode to plate-cathode capacity.

$$\mu = \frac{C_{cg}}{C_{cp}} \tag{7.47}$$

This is physically reasonable since the ratio of these two capacities is a measure of the extent to which the cathode is electrostatically shielded from the plate by the grid. The capacities involved can be evaluated by reference to Eqs. (7.33) and (7.34) if it is desired to know them in terms of the geometry of the tube. The capacities in the above expressions are in farads per unit length (meter) per grid-wire section of the elementary tube in the Z plane of Fig. 7.10. When the dimensions are transformed to other planes, the capacities of the corresponding tubes result.

Reference to any tube manual will show that the numerical ratio of the grid-cathode to grid-plate capacities listed there differs considerably from the amplification factor of the tube. This is because the capacities listed in the manual include the capacities of the leads and supports as well as those of the parts of the tube in which the electrons are effective. In most triodes the capacities between the leads and supports may be as

large as those of the active portions of the tube or larger, the apparent discrepancy being thus accounted for.

7.6. Equivalent-diode Spacing of a Triode. In the case of the plane-electrode triode, examination of the potential profiles showed that the profiles were straight lines in the vicinity of the cathode for all conditions of potential. This means that the potential distribution in a triode as seen from the cathode is the same as that in a diode. The cathode has no data by which it can tell whether it is part of a diode or triode. It was further shown in Eq. (7.14*a*) that the off-cathode gradient of potential depended upon an equivalent voltage, $V_g + \frac{V_p}{\mu}$. It would be expected, therefore, that for every triode there would exist an equivalent diode which would have the same off-cathode gradient when the equivalent-triode voltage is applied to its plate. To find such an equivalent diode it is necessary only to find the equivalent-diode spacing.

The equivalent-diode spacing of any triode may be found graphically by extending the straight-line portion of the potential profile in the vicinity of the cathode until it reaches a potential equal to the equivalent-triode potential, $V_g + \frac{V_p}{\mu}$. The distance from the cathode at which this potential is reached is the equivalent-diode spacing. This construction is shown in Fig. 7.4*c*. Once this equivalent-diode spacing is found it can be used for all combinations of plate and grid potential.

The concept of the equivalent diode and the equivalent-diode spacing is useful in the study of the current characteristics of triodes. Since the current law for space-charge-limited diodes is known, the current law for a triode can be approximated from the equivalent diode. Triodes with equal equivalent-diode spacings and equal amplification factors have *approximately* the same mutual conductance and plate resistance. Actually, the triode and its equivalent diode are truly equivalent only for a condition of no current flow, which means cutoff or beyond for the triode since the flow of current changes the potential distribution. The concept is, however, sufficiently useful to justify its inclusion here. The subject of current flow and mutual conductance will be discussed in the chapter on Space-charge Effects.

Diode Equivalent to a Plane-electrode Triode. An analytical expression for the equivalent-diode spacing can be found from the expression for cathode charge in terms of the equivalent-triode potential and the geometry of the tube. Since the cathode charge was taken as $+q_c$ per grid-wire section, the off-cathode gradient of potential is given by

$$\left(\frac{dV}{dx}\right)_c = \frac{-q_c}{a\varepsilon_0} \tag{7.48}$$

or

$$\left(\frac{dV}{dx}\right)_c = -\left[\frac{V_g + \left(\frac{V_p}{\mu}\right)}{D}\right]\frac{1}{a\varepsilon_0} \tag{7.49}$$

in which all the symbols have the previous significance and D is the denominator of Eq. (7.33) or (7.35) when the equivalent-voltage factor is extracted.

For a plane-electrode diode the electrode spacing is the ratio of the potential difference to the gradient. For the triode the equivalent-diode spacing is the ratio of the equivalent-triode potential to the off-cathode gradient. Thus, from Eq. (7.33)

$$d_e = \frac{a}{2\pi}\left[\frac{\ln(\rho_c\rho_g)\ln\left(\frac{R}{\rho_g}\right)}{\ln(\rho_p/\rho_g)} - \ln(\rho_c R)\right] \tag{7.50}$$

where d_e is the equivalent-diode spacing in terms of the Z-plane dimensions. It is desirable to express d_e in terms of the cathode-grid distance, the grid-plate distance, and the amplification factor. If the expression for the amplification factor [Eq. (7.36)] is used to eliminate R, then

$$d_e = \frac{-a}{2\pi}\left[\left(\frac{1}{\mu}+1\right)\ln(\rho_c\rho_g) - \frac{1}{\mu}\ln\left(\frac{\rho_p}{\rho_g}\right)\right] \tag{7.51}$$

If now the plane-electrode transformation is applied to insert the triode dimensions,

$$d_e = d_{cg} + \frac{d_{gp}+d_{cg}}{\mu} - \frac{a}{2\pi}\left(1+\frac{2}{\mu}\right)\ln\cosh\pi S \tag{7.52}$$

For tubes in which the screening fraction is less than $\frac{1}{10}$, the last term of Eq. (7.52) is negligible so that the expression reduces to

$$d_e \cong d_{cg}\left[1 + \frac{d_{cg}+d_{gp}}{\mu d_{cg}}\right] \tag{7.53}$$

Examination of Eq. (7.53) shows that the plate of the equivalent diode always lies beyond the actual grid of the triode. The distance beyond is relatively less if the amplification factor and cathode-grid distance are large and relatively more if the grid-plate distance is large. The mutual conductance of a tube is an inverse function of the equivalent-diode spacing so that the influence of the various tube dimensions is readily determined from Eq. (7.53).

Diode Equivalent to a Cylindrical-electrode Triode. The procedure that was used to find the equivalent-diode spacing of a plane-electrode triode

is quite general and may be applied to cylindrical triodes as well, though the form of the resulting expressions is quite different, as may be expected from the fact that the potential profile for the plane-electrode diode is a straight line, while for the cylindrical diode it is a logarithmic curve. For a cylindrical diode the relation between cathode and plate radius, potential difference, and electrode charge per unit length of outer electrode is

$$\ln\left(\frac{r_p}{r_c}\right) = \frac{2\pi\varepsilon_0 V}{q} \tag{7.54}$$

For the cylindrical triode the equivalent-diode radius is given by

$$\ln\left(\frac{s_e}{s_c}\right) = -\frac{2\pi\varepsilon_0}{q_c N}\left(V_g + \frac{V_p}{\mu}\right) \tag{7.55}$$

where s_e is the equivalent-diode radius and q_c is the cathode charge per unit length for each of N grid-wire sections. Substitution of the value of q_c from Eq. (7.33) and using Eq. (7.36) gives

$$\ln\left(\frac{s_e}{s_c}\right) = -\left[\frac{\ln(\rho_c\rho_g)}{\mu} + \ln(\rho_c R)\right]\frac{1}{N} \tag{7.56}$$

Eliminating R by means of Eq. (7.39) and making the substitutions of Eqs. (7.16), (7.40), and Eq. (7.42),

$$\ln\left(\frac{s_e}{s_c}\right) = \left[\frac{1}{\mu}\left(\ln\frac{s_p}{s_g} - \frac{1}{N}\ln\cosh\pi S\right) - \left(1+\frac{1}{\mu}\right)\left(\ln\frac{s_c}{s_g} + \frac{1}{N}\ln\cosh\pi S\right)\right] \tag{7.57}$$

For tubes with screening fractions less than $\frac{1}{10}$ the terms involving the hyperbolic cosine can be neglected so that the expression takes the form

$$\ln\left(\frac{s_e}{s_c}\right) = \ln\left(\frac{s_g}{s_c}\right)\left[1 + \frac{1}{\mu} + \frac{1}{\mu}\frac{\ln\left(\frac{s_p}{s_g}\right)}{\ln\left(\frac{s_g}{s_c}\right)}\right] \tag{7.58}$$

The same remarks that applied to the diode spacing for plane electrodes apply here. The equivalent-diode plate lies outside of the grid-wire circle. The equivalent-diode radius for the cylindrical triode is readily obtained by graphical construction. If the profiles of Fig. 7.7 are plotted against a logarithmic abscissa as in Fig. 7.8, then the profiles are straight lines in the vicinity of the cathode and plate. The equivalent-diode radius can be found by projecting the straight-line portion of the potential

profile at the cathode until it reaches a potential equal to the equivalent-grid potential.

7.7. Application of Amplification-factor Formulas to Actual Triodes. The amplification-factor formulas previously developed have been derived for simple idealized structures not always encountered in actual tubes. It is, however, possible in most cases to interpret these formulas so that they will apply to tubes whose structure departs somewhat from that for which the formulas were derived.

The formula for the amplification factor of the cylindrical triode [Eq. (7.42)] was given for a grid in which the wires had a squirrel-cage structure of evenly spaced wires parallel to the axis of the tube. In this expression the quantity N is the number of grid wires and is also the active length of grid wire per unit axial length of the tube. The expression may therefore be generalized by letting

$$N = L_g \tag{7.59}$$

where L_g is the active length of grid wire per unit axial length of the tube.[1]

In case the grid structure differs from that postulated in the derivation of the amplification-factor formula, the screening fraction S may always be interpreted as the ratio of the actual area of the grid structure to the total area of the surface containing the grid.

If the cylindrical grid consists of a square mesh of fine wires, then

$$L_g = 4\pi \frac{s_g}{d} \tag{7.60}$$

where d is the spacing of the square mesh and s_g is the radius of the grid-wire circle. If the diameter of the wires in the square mesh is appreciable, then

$$L_g = \frac{4\pi s_g}{d}\left(1 - \frac{r_g}{d}\right) \tag{7.61}$$

where r_g is the radius of the grid wires as shown in Fig. 7.14*a*. For a cylindrical grid of parallel rings having supports parallel to the axis,

$$L_g = \frac{2\pi s_g}{s} + N_s - \frac{2N_s r_s}{s} \tag{7.62}$$

where N_s is the number of supports and r_s is the radius of the support wires and s is the spacing of the grid wires as shown in Fig. 7.14*b*.

If the grid is a helix of diameter $2s_g$ and of pitch d as in Fig. 7.14*c*,

[1] KUSUNOSE, Y., Design of Triodes, *Proc. I.R.E.*, vol. 17, pp. 1706–1749, October, 1929.

$$L_g = \sqrt{1 + \left(\frac{2\pi s_g}{d}\right)^2} \tag{7.63}$$

The screening fractions for the cases listed above are readily evaluated. In general if all the wires have the same radius,

$$\pi S = L_g \frac{r_g}{s_g} \tag{7.64}$$

For parallel grid rings with supports,

$$\pi S = \frac{2r_g(L_g - N_s) + 2N_s r_s}{2s_g} \tag{7.65}$$

in which the symbols have the same significance as in Eq. (7.62). Equation (7.64) also holds for a helical grid. Equation (7.65) reduces to Eq. (7.64) for $r_g = r_s$.

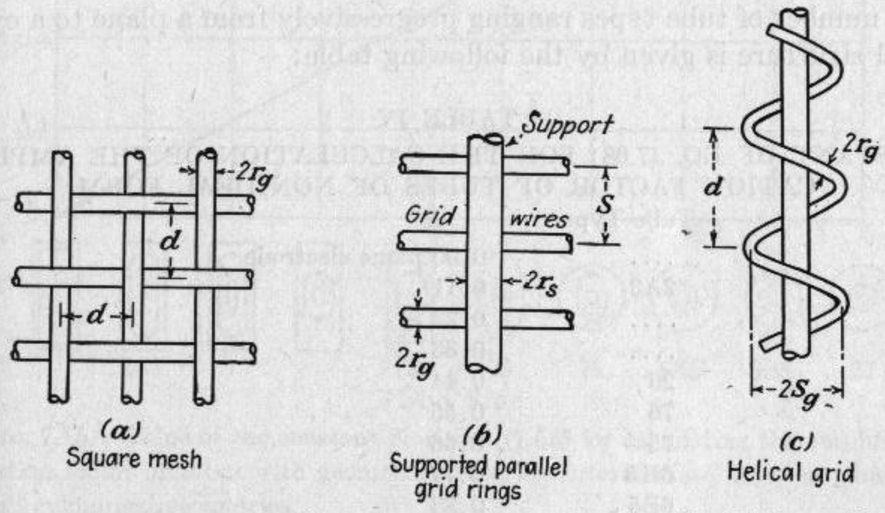


FIG. 7.14.—Practical grid structures: (*a*) square mesh; (*b*) grid rings with supports parallel to the axis; (*c*) helical grid.

The generalized amplification-factor formula for the cylindrical-electrode triode is

$$\mu_c = \frac{L_g \ln \frac{s_p}{s_g} - \ln \cosh \pi S}{\ln \coth \pi S} \tag{7.66}$$

The generalized amplification-factor formula for the plane-electrode triode is

$$\mu_p = \frac{2\pi d_{gp} L_g' - \ln \cosh \pi S}{\ln \coth \pi S} \tag{7.67}$$

The relation of this to Eq. (7.38) is evident. L_g' is the length of grid wire per unit area of the grid plane. For parallel grid wires, $L_g' = \frac{1}{a}$.

In many tubes the structure is neither plane nor cylindrical but something intermediate. In such cases it has been found empirically that a combination of the plane and cylindrical amplification-factor formulas gives very nearly the amplification factor of the tube.[1] The combination formula is

$$\mu = \mu_p - K(\mu_p - \mu_c) \tag{7.68}$$

where μ_p is the amplification factor as calculated by the plane-electrode formula, μ_c is the amplification factor as calculated by the cylindrical-electrode formula, K is a constant depending upon the tube structure, and μ is the amplification factor of the actual tube. The constant K for a number of tube types ranging progressively from a plane to a cylindrical structure is given by the following table:

TABLE IV

CONSTANT OF EQ. (7.68) FOR THE CALCULATION OF THE AMPLIFICATION FACTOR OF TUBES OF NONIDEAL FORM

Tube Type	K
....	0.00 plane electrode
2A3	0.11
....	0.22
....	0.33
26	0.44
76	0.55
75	0.66
6K5	0.77
6B5	0.88
100TL	1.00 cylindrical electrode

The results of Table IV are shown graphically in Fig. 7.15. The empirical constant K given here includes the effect of the grid supports.

7.8. More Accurate Amplification-factor Formulas. The amplification-factor formula of Vodges and Elder given in Eq. (7.38) is accurate only for certain ranges of electrode dimensions. Specifically, the formula breaks down if

(a) $d_{gp} < a$
(b) $S > 0.16$
(c) $d_{cg} < a$

In general, the formula breaks down if any of the electrodes are too close together. Since many modern tubes are built with very close-

[1] JERVIS, E. R., Amplification Factor Chart, *Electronics*, vol. 12, p. 45, June, 1939.

spaced electrodes, it is desirable to extend the range of the above relations. Various formulas have been worked out that extend the range of any of the three limitations listed above, but as yet no formula has appeared that is valid over the complete range of all variables.

Formula for Small Grid-plate Spacings. When the grid-plate spacing is small, an improved amplification-factor formula may be worked out by

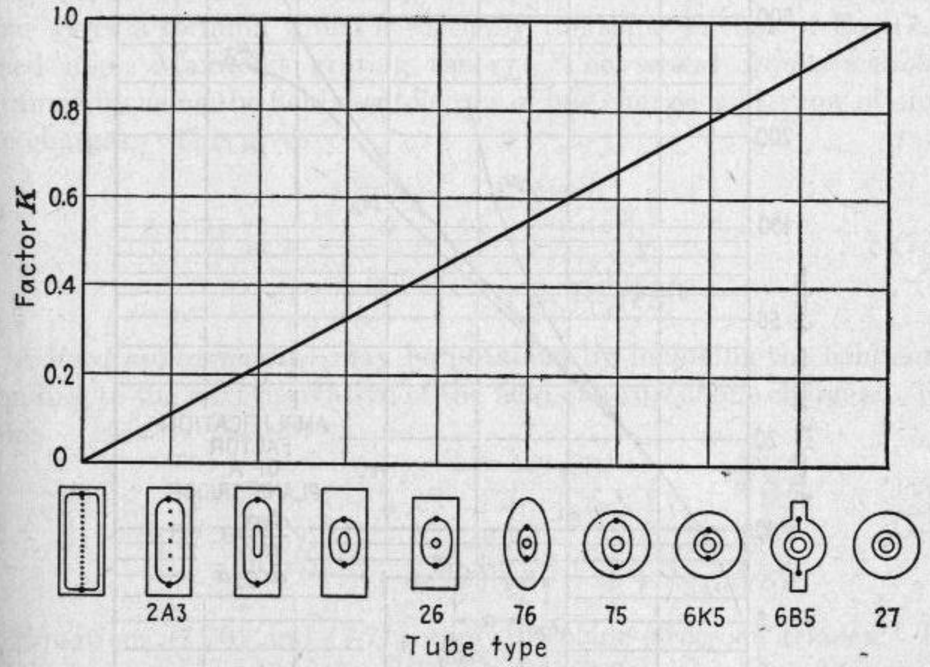


FIG. 7.15.—Value of the constant K of Eq. (7.68) for calculating the amplification factor of tubes with geometries that are intermediate between plane and cylindrical geometries.

placing an image set of grids outside of the plate position and then fitting the plate to the equipotential curve midway between.[1] The resulting expression is

$$\mu = \frac{\ln \cosh \pi S - 2\pi \frac{d_{gp}}{a}}{\ln \tanh \pi S - \ln \left[1 - \cosh^2 \pi S \exp\left(\frac{-4\pi d_{gp}}{a}\right)\right]} \tag{7.69}$$

This expression is valid for grid-plate spacings as low as one-quarter of the spacing between grid wires but still assumes that the cathode-grid spacing and the screening fraction are large.

[1] SALZBERG, BERNARD, Formulas for the Amplification Factor of Triodes, *Proc. I.R.E.*, vol. 30, pp. 134–138, March, 1942.

Formulas for Large Screening Fraction. Perhaps the most serious limitation to the amplification formula of Vodges and Elder is that it begins to be in error for a screening fraction of $\frac{1}{6}$ and is 10 per cent low for a screening fraction of $\frac{1}{3}$. Many modern tubes are built with very

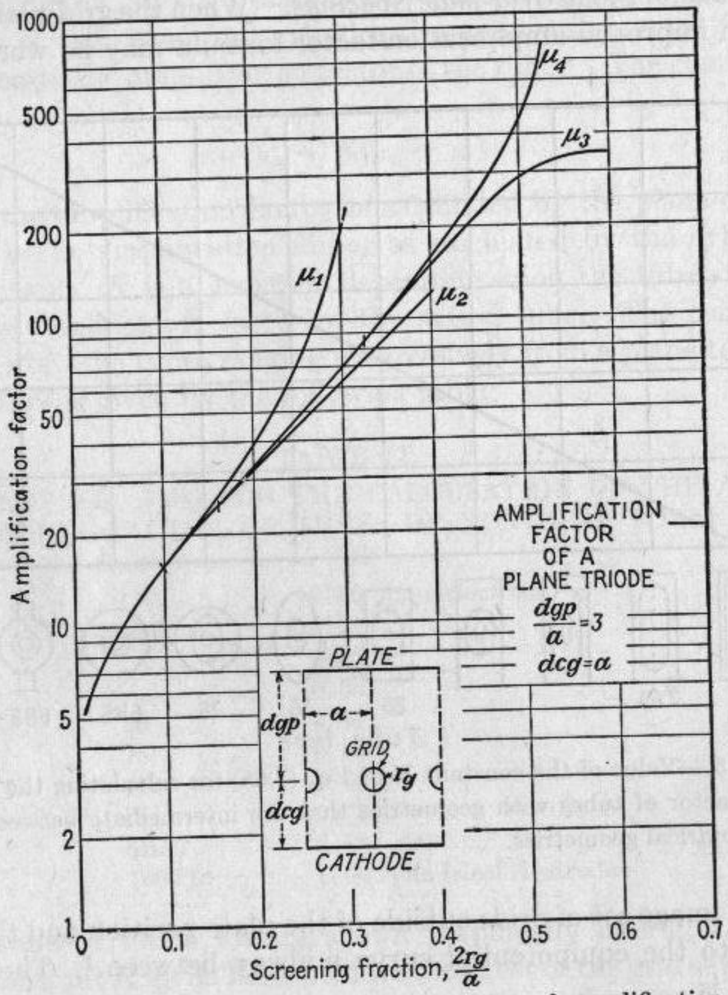


FIG. 7.16.—A comparison of the values of amplification factor of a plane-electrode triode as a function of screening fraction as indicated by various formulas. The extension of the region of validity by the successive refinements in formulas is evident. (See discussion on page 161.)

large screening fractions, and it therefore is desirable to have a formula valid in this range.

Such a formula has been evaluated[1,2] by an analysis based upon the

[1] OLLENDORF, FRANZ, Berechnung des Durchgriffes durch enge Steggitter, *Elektrotech. u. Maschinenbau*, vol. 52, pp. 585–591, Dec. 16, 1934.

[2] See also HERNE, H., Valve Amplification Factors, *Wireless Engr.*, vol. 21, pp. 59–64, February, 1944.

expression $W = \ln \sin Z$, which gives the potential due to a row of evenly spaced line charges positioned along a straight line. Derivatives of this expression give the potential fields due to a row of dipoles, quadripoles, and so on. By combining in series form such a succession of fields and fitting the resultant field to the circular grid wires, some highly accurate formulas for amplification factor are obtained.

The first approximation using the expression for the row of line charges alone gives a formula which is virtually the same as that of Eq. (7.13) based upon Maxwell's grating theory. The *second approximation* is obtained by using the field due to a row of line charges and a row of dipole line charges. This gives

$$\mu = \frac{\frac{2\pi d_{gp}}{a} - \frac{\frac{1}{2}(\pi S)^2}{1 + \frac{1}{12}(\pi S)^2}}{-\ln(\pi S) + \frac{\frac{1}{4}(\pi S)^2}{1 + \frac{1}{12}(\pi S)^2}} \tag{7.70}$$

A *third approximation* may be obtained by including the field corresponding to the next derivative of the field of a row of line charges. This yields

$$\mu = \frac{\frac{2\pi d_{gp}}{a} - \frac{\frac{1}{2}(\pi S)^2}{1 + \frac{1}{12}(\pi S)^2}}{-\ln(\pi S) + \frac{\frac{1}{4}(\pi S)^2}{1 + \frac{1}{12}(\pi S)^2} - \frac{\frac{1}{288}(\pi S)^4}{1 + \frac{1}{240}(\pi S)^4}} \tag{7.71}$$

Equations (7.70) and (7.71) apply to plane-electrode triodes. The corresponding expressions for cylindrical electrodes may be obtained by substituting $s_g \ln\left(\frac{s_p}{s_g}\right)$ for d_{gp} provided that the grid-wire radius is small compared with the radius of the grid-wire circle.

A comparison of the values of amplification factor given by the various formulas presented is shown in Fig. 7.16. In this figure the notation is as follows:

μ_1, Eq. (7.13), Maxwell.
μ_2, Eq. (7.38), Vodges and Elder.
μ_3, Eq. (7.70), Ollendorf second approximation.
μ_4, Eq. (7.71) Ollendorf third approximation.

From Fig. 7.16 it is seen that the range of validity of the various formulas within a 2 per cent error is

μ_1, Eq. (7.13), S less than 0.1.
μ_2, Eq. (7.38), S less than 0.16.
μ_3, Eq. (7.70), S less than 0.325.
μ_4, Eq. (7.71), S less than 0.4.

In the chart of Fig. 7.11 giving the amplification factor of a plane triode, the solid lines were obtained from the formula of Vodges and Elder, while the dotted lines were obtained from the Ollendorf approximations.

Formula for Small Cathode-grid Spacings. All the amplification-factor formulas previously given are restricted to electrode configurations in which the cathode-grid spacing is equal to or larger than the grid-wire spacing. Many modern tubes have grid-cathode spacings that are less than the grid-wire spacing. When this is true, it is really no longer proper to speak of the amplification factor, for the gradient of potential is not constant along the cathode but varies with position relative to the grid wires. Such a tube exhibits no true cutoff condition since as the grid is made more and more negative the cathode gradient opposite the grid wires will become negative while the cathode gradient between the wires is yet positive. This gives rise to a condition, sometimes referred to as *Insel Bildung*, in which little island strips of the cathode are emitting while other parts are not. Such a tube acts as a variable-mu tube in that every part of the cathode surface has a different amplification factor.

It is possible to find the effective amplification factor of a tube with small cathode-grid spacing if this effective amplification factor is understood to be dependent upon position on the cathode. When the cathode gradient is not uniform, as is the case with small cathode-grid spacings, the field's configuration can still be found if use is made of the theory of images.[1] The true field that exists is the same that would result from a line of grid wires of one charge and an image line of wires located as though mirrored in the cathode but having charge of opposite sign. The field can therefore be obtained by studying the potential within a fictitious tube consisting of two parallel plates of opposite charge and potential, between which there are two sets of identical grids of opposite charge and potential, symmetrically disposed with respect to the center plane. Under the conditions stated the mid-plane will be a surface of zero potential and can be identified as the cathode.

The potential due to a grid of equally charged parallel wires, equally spaced, a distance a along the y axis of the Z plane is

$$V(x,y) = -\frac{q}{4\pi\epsilon_0} \ln\left[2\left(\cosh\frac{2\pi x}{a} - \cos\frac{2\pi y}{a}\right)\right] + C \qquad (7.72)$$

where q is the charge per unit length of each grid wire. This expression may be obtained by considering the potential due to a line charge of density q located at the point (1,0) and one of density $-\frac{q}{2}$ located at the

[1] FREMLIN, J. H., Calculation of Triode Constants, *Phil. Mag.*, vol. 27, pp. 709–741, June, 1939. Also published in *Elec. Commun.*, July, 1939.

point (0,0) in the Z plane and then transforming this potential to the W plane by means of the transformation $W = \frac{a}{2\pi} \ln Z$. Equation (7.72) can also be shown to be the real part of the expression $\ln\left(\sin\frac{i\pi Z}{a}\right)$. A plot of Eq. (7.72) is given in Fig. 7.17.

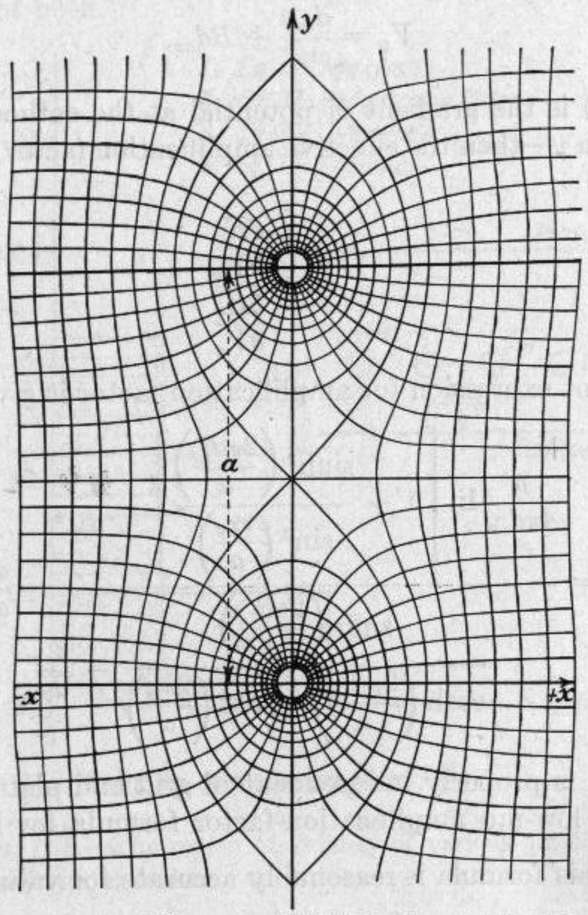


FIG. 7.17.—The potential field in the vicinity of a row of line charges as given by Eq. (7.72).

Upon using Eq. (7.72) to obtain the field due to grids of opposite charge located a distance d_{cg} from the cathode and adding a linear component of field to account for the effect of the plate the expression for the potential within the tube is obtained. It is

$$V = +\frac{q}{4\pi\epsilon_0} \ln\left[\frac{\cosh\frac{2\pi}{a}(x + d_{cg}) - \cos\frac{2\pi}{a}y}{\cosh\frac{2\pi}{a}(x - d_{cg}) - \cos\frac{2\pi}{a}y}\right] + Bx \qquad (7.73)$$

in which q and B are related to the electrode potentials by

$$V_g = +\frac{q}{4\pi\epsilon_0}\ln\left[1+\frac{\sinh^2\left(\frac{2\pi d_{cg}}{a}\right)}{\sin^2\left(\frac{\pi r_g}{a}\right)}\right]+Bd_{cg} \tag{7.74}$$

and

$$V_p = \frac{d_{cg}q}{a\epsilon_0}+Bd_{cp}, \tag{7.75}$$

If $-E_c(0,y)$ is the gradient of potential at the cathode—a quantity that varies with y—then the effective amplification factor may be defined as

$$\frac{1}{\mu} = +\frac{\frac{\partial E_c}{\partial V_p}}{\frac{\partial E_c}{\partial V_g}} \tag{7.76}$$

The resultant expression for amplification factor is given by

$$\frac{1}{\mu} = \frac{\frac{a}{4\pi d_{cp}}\ln\left[1+\frac{\sinh^2\left(\frac{2\pi d_{cg}}{a}\right)}{\sin^2\left(\frac{\pi r_g}{a}\right)}\right]-\frac{{d_{cg}}^2}{{d_{cp}}^2}}{\frac{\sinh\left(\frac{2\pi d_{cg}}{a}\right)}{\cosh\left(\frac{2\pi d_{cg}}{a}\right)-\cos\left(\frac{2\pi y}{a}\right)}-\frac{d_{cg}}{d_{cp}}}-\frac{d_{cg}}{d_{cp}} \tag{7.77}$$

This expression is properly independent of grid and plate potential and reduces to the low-mu amplification-factor formula for large values of $\frac{d_{cg}}{a}$ and $\frac{d_{cp}}{d_{cg}}$. This formula is reasonably accurate for values of $\frac{d_{cg}}{a}$ as low as 0.4. For small values of the ratio $\frac{d_{cg}}{a}$ the reciprocal of the amplification factor, sometimes referred to as the penetration factor or *Durchgriff* since it is a measure of the shielding effect of the grid, exhibits what is nearly a sinusoidal variation with distance parallel to the grid. For the case in which the grid-cathode distance is 0.4 of the grid-wire spacing, the relative amplification factor may vary between 0.7 and 1.4 times the average value, the average value being very nearly equal to that given by Eq. (7.13).

Expressions for the amplification factor of a plane triode have now been given that cover nearly the entire range of practical electrode

dimensions. All the expressions given are limited, however, to some definite range of electrode dimensions. The expressions given in this subsection, for instance, are limited to screening fractions of 0.1 or less. The region of validity of the various amplification-factor formulas is shown in Fig. 7.18. Here it is seen that formulas are good for either large screening fractions or large ratio of grid-wire spacing to cathode-grid distance but not both.[1]

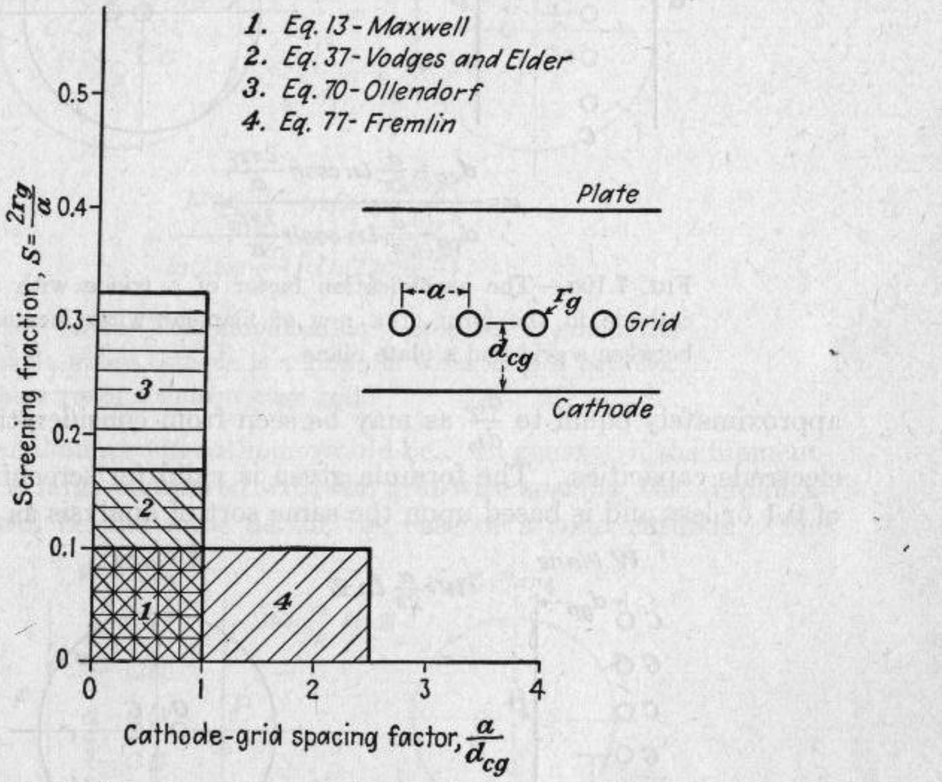


FIG. 7.18.—The region of validity of various amplification-factor formulas.

7.9. Amplification Factor of Unconventional Tubes. The methods that have been studied in this chapter may be applied to numerous structures other than the idealized plane and cylindrical structures so far treated. Where simple geometries are involved, it is usually possible to find a correspondingly simple arrangement of line charges that can be transformed into the desired structure. In Fig. 7.19 are shown some sample unconventional tube structures along with their elementary forms and the corresponding amplification-factor formula.

[1] See FREMLIN, J. H., R. N. HALL, and P. A. SHAFFORD, Triode Amplification Factors, *Elec. Commun.*, vol. 23, no. 4, pp. 426–435, 1946, for a semiempirical formula good for both small screening fraction and small cathode-grid distance.

In Fig. 7.19*a* is shown the famous gammatron "gridless wonder." This consists of a row of filament wires between grid and plate planes. The tube has an inherently low amplification factor whose value is

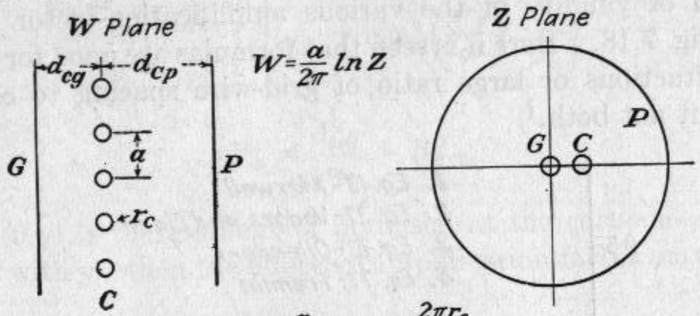


FIG. 7.19*a*.—The amplification factor of a triode with a cathode in the form of a row of filament wires located between a grid and a plate plane.

approximately equal to $\frac{d_{cp}}{d_{cg}}$ as may be seen from consideration of interelectrode capacities. The formula given is valid for screening fractions of 0.1 or less and is based upon the same sort of analysis as was used to

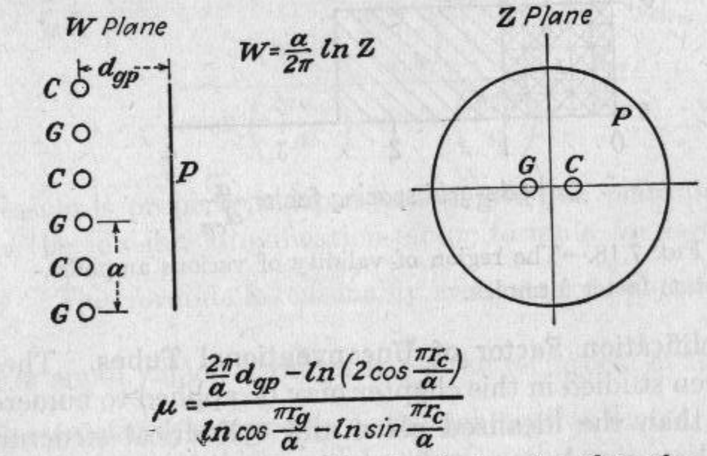


FIG. 7.19*b*.—The amplification factor of a plane-electrode triode with a cathode consisting of a row of filament wires spaced midway between the grid wires and having a single plate.

treat the low-mu triode. It is seen that the amplification factor has a second-order dependence upon the cathode radius.

In Figs. 7.19*b*, *c*, and *d* are shown tubes in which filament wires are placed between the grid wires. The formulas are not greatly different

from those for tubes which have a grid between cathode and plate except that they exhibit a second-order dependence upon cathode radius.

The question arises as to what the effect of having a filamentary

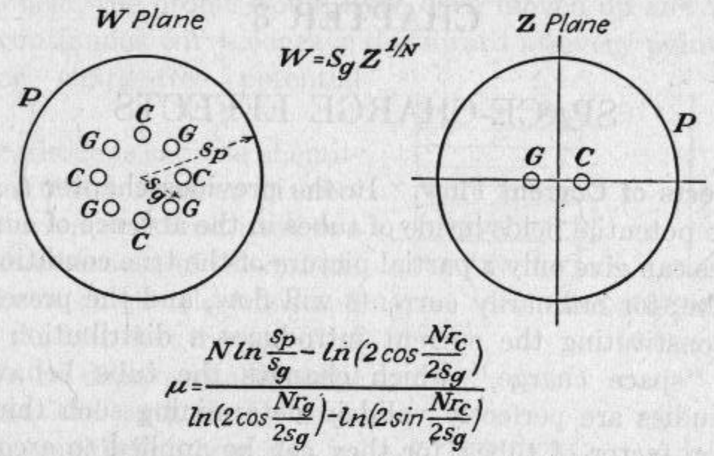


FIG. 7.19*c*.—The amplification factor of a cylindrical triode whose cathode is a group of wires located between the wires of a squirrel-cage grid.

emitter rather than a solid cathode would be. In general, if the filament-grid spacing is large compared with the grid-wire spacing, the amplification factor will be the same as for the case of a solid cathode. The

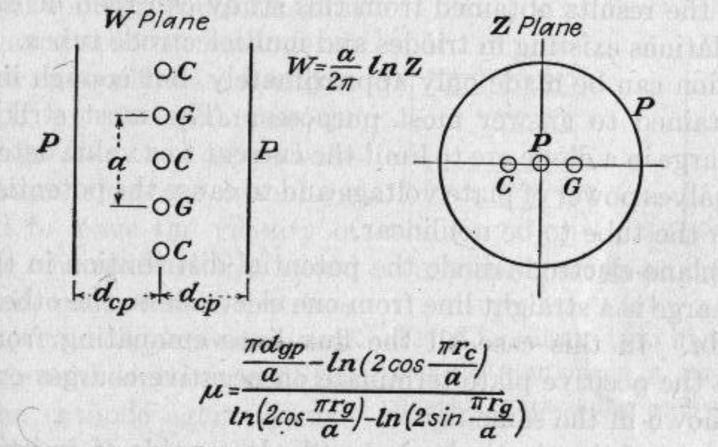


FIG. 7.19*d*.—The amplification factor of a plane-electrode triode whose grid and cathode wires are alternate and equally spaced in a plane between two equidistant plate planes.

equipotential lines around the filament will be circles very near to it but become nearly straight lines, the same as for the solid cathode, between the filament and grid. The current-voltage characteristics may, however, be considerably different.

CHAPTER 8

SPACE-CHARGE EFFECTS

8.1. Effects of Current Flow. In the previous chapter a study was made of the potential fields inside of tubes in the absence of current flow. Such studies can give only a partial picture of the true condition within a vacuum tube; for ordinarily currents will flow, and the presence of the electrons constituting the current introduces a distribution of charge known as "space charge," which changes the tube behavior. The previous studies are perfectly valid in determining such things as the amplification factor of tubes, for they can be applied to a condition of cutoff at which no current flows and yet at which the relative influence of the various electrodes is the same as for current flow. For studies of such subjects as the variation of current with potential and the determination of mutual conductance it is necessary to take into account the effect of currents and the corresponding space charge.

The effect of space charge is most readily studied in the case of the diode, and the results obtained from this study can then be extended to give the relations existing in triodes and multielectrode tubes. Actually, this extension can be made only approximately, but enough information can be obtained to answer most purposes. The most striking effects of space charge in a diode are to limit the current to a value determined by the three-halves power of plate voltage and to cause the potential distribution within the tube to be nonlinear.

In the plane-electrode diode the potential distribution in the absence of space charge is a straight line from one electrode to the other as shown in Fig. 8.1*a*. In this case all the flux lines emanating from positive charges on the positive plate terminate on negative charges on the other plate, as shown in the same figure.

If now one of the electrodes is a cathode capable of emitting a small number of electrons and the other electrode is positive with respect to this, there will be some electrons in the field between the two electrodes moving toward the plate and some of the flux lines will terminate on electrons in the field as shown in Fig. 8.1*b*. The drawing has been conventionalized and the unit of flux density chosen so that there is only one flux line to each electron. Because of the negative charge in the field, the potential at any point will be less than in the previous case. Accordingly

the potential profile will be pulled down and will be a curve concave upward at every point below the straight line, with the same initial and final values. If instead of negative electrons there had been positive ions in the field, the profile would have been moved up and would have formed a continuous curve concave downward at every point above the straight-line charge-free potential profile.

If the cathode is capable of emitting an unlimited number of electrons, the current will limit itself to a definite value because of the mutual repulsion between the electrons and because of the fact that the potential contour can be depressed only until its slope at the cathode is zero. There is an equilibrium here; for, as shown in Fig. 8.1*c*, if in some manner the slope could be made less than zero at the cathode, the electrons starting out would be forced to return to the cathode, no current would flow, the space charge would be reduced, and the potential contour would lift until its slope was zero at the cathode, at which point an equilibrium would be reached. If in some manner the slope were greater than zero at the cathode, more electrons would be encouraged to leave the vicinity of the cathode, the space charge would be increased, the potential-distribution curve would be depressed, and this action would continue until the slope at the cathode again became zero. The zero slope at the cathode indicates that the charge on the cathode is zero, which means that the flux lines emanating from charges on the positive plate all terminate on electrons in the field and that none of them get through to the cathode. In the case of the current limited by space charge, the potential distribution is a four-thirds-power law, as will be shown presently. In this case also, the current varies as the three-halves power of the potential on the plate.

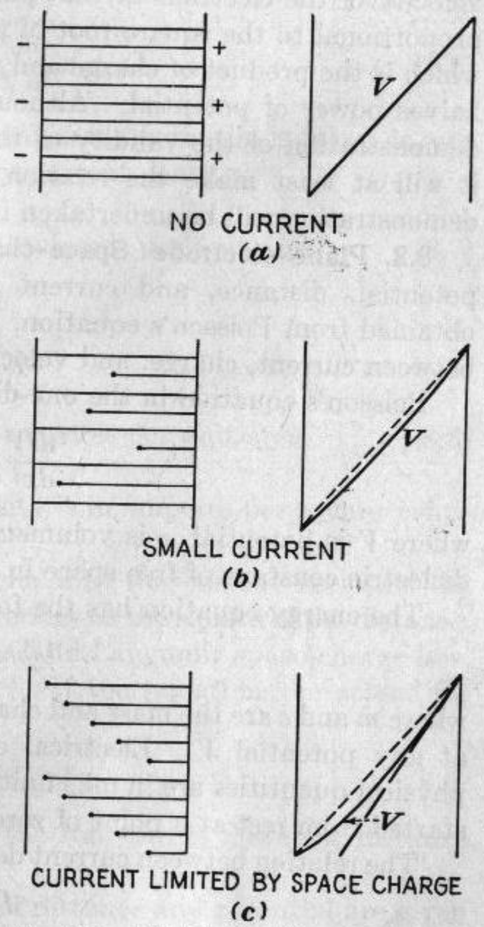


FIG. 8.1.—Electric flux lines and potential distribution in a plane-electrode diode for various degrees of space charge.

The reason for the three-halves-power variation of current with poten-

tial can be shown from simple theoretical considerations. In Fig. 8.1*c* it is seen that the space-charge-limited diode is roughly equivalent to a condenser. As is the case in a condenser, charge on the plate and hence also in the interelectrode space is proportional to the potential. The current density is the product of the space-charge density and the velocity of the electrons at that point. The velocity of the electrons is proportional to the square root of potential so that the current density, which is the product of charge and velocity, is proportional to the three-halves power of potential. Although the above is not a very rigorous demonstration of the validity of the three-halves-power law of current, it will at least make the relation seem reasonable. A more rigorous demonstration will be undertaken in the next section.

8.2. Plane-electrode Space-charge Flow. The relations between potential, distance, and current in the plane-electrode case can be obtained from Poisson's equation, the energy equation, and the relation between current, charge, and velocity.

Poisson's equation in the one-dimensional case reduces to

$$\frac{d^2V}{dx^2} = -\frac{\rho}{\varepsilon_0} \tag{8.1}$$

where V is potential, ρ is volumetric space-charge density, and ε_0 is the dielectric constant of free space in rationalized mks units.

The energy equation has the form

$$\tfrac{1}{2}mv^2 = Ve \tag{8.2}$$

where m and e are the mass and charge of the electron and v is the velocity at any potential V. Electrical quantities are in practical units, and physical quantities are in mks units. This assumes that the electron has started from rest at a point of zero potential.

The relation between current density, charge, and velocity is

$$J = \rho v \tag{8.3}$$

The three equations above suffice for a determination of all the relations involved in a parallel-electrode space-charge flow. If ρ is expressed in terms of J and V from Eqs. (8.2) and (8.3) and the resulting expression substituted into Eq. (8.1),

$$\frac{d^2V}{dx^2} = \frac{J}{\varepsilon_0}\sqrt{\frac{m}{2e}}\,V^{-\frac{1}{2}} \tag{8.4}$$

where J is now the magnitude of the current density, actually electronic flow in the positive x direction is negative. A first integration is achieved by multiplying both sides of Eq. (8.4) by $2\dfrac{dV}{dx}$ and integrating,

$$\left(\frac{dV}{dx}\right)^2 = \frac{4J}{\varepsilon_0}\sqrt{\frac{m}{2e}}\,V^{\frac{1}{2}} + C_1 \tag{8.5}$$

The constant of integration is zero because the gradient is zero when the potential is zero.

A second integration gives

$$\frac{4V^{\frac{3}{4}}}{3} = \sqrt{\frac{4J}{\varepsilon_0}\sqrt{\frac{m}{2e}}}\,x + C_2 \tag{8.6}$$

in which the constant is again zero because the potential is taken as zero when the distance is zero. Solving for current density,

$$J = \frac{4\varepsilon_0}{9\sqrt{\dfrac{m}{2e}}}\frac{V^{\frac{3}{2}}}{x^2} \quad \text{amperes per unit area} \tag{8.7}$$

Numerically this is equal to

$$J = \frac{2.335 \times 10^{-6}V^{\frac{3}{2}}}{x^2} \quad \text{amperes per unit area} \tag{8.8}$$

If x is in centimeters, the current density is in amperes per square centimeter.

From the above equations it is seen that the current varies as the three-halves power of potential and inversely as the square of the distance. The last two equations constitute the Child-Langmuir space-charge law. It has been verified experimentally.[1–3] If the equations be solved for potential, there results

$$V = 5{,}680J^{\frac{2}{3}}x^{\frac{4}{3}} \tag{8.9}$$

showing that the potential varies as the four-thirds power of distance between cathode and plate.

Values of current density in terms of distance and potential are given in the curves of Fig. 8.2.

It is of interest to note how various other factors vary with distance. The gradient of potential is given by the derivative of potential with distance. Hence

$$E = k_1x^{\frac{1}{3}} \tag{8.10}$$

[1] CHILD, D. C., Discharge from Hot CaO, *Phys. Rev.*, vol. 32, pp. 492–511, May, 1911.

[2] LANGMUIR, I., The Effect of Space Charge and Residual Gases on Thermionic Currents in High Vacuum, *Phys. Rev.*, Ser. 2, vol. 2, pp. 450–486, December, 1913.

[3] LANGMUIR, I., and K. B. BLODGETT, Currents Limited by Space Charge between Coaxial Cylinders, *Phys. Rev.*, Ser. 2, vol. 22, pp. 347–357, October, 1923.

Velocity varies as the square root of potential, and therefore

$$v = k_2 x^{\frac{2}{3}} \tag{8.11}$$

Space-charge density varies inversely as velocity, from Eq. (8.3), since the current density is constant; thus

$$\rho = k_3 x^{-\frac{2}{3}} \tag{8.12}$$

Curves showing the variation of these various factors are shown in Figs. 8.3*a* and *b*. Here it is seen that the potential and gradient are zero

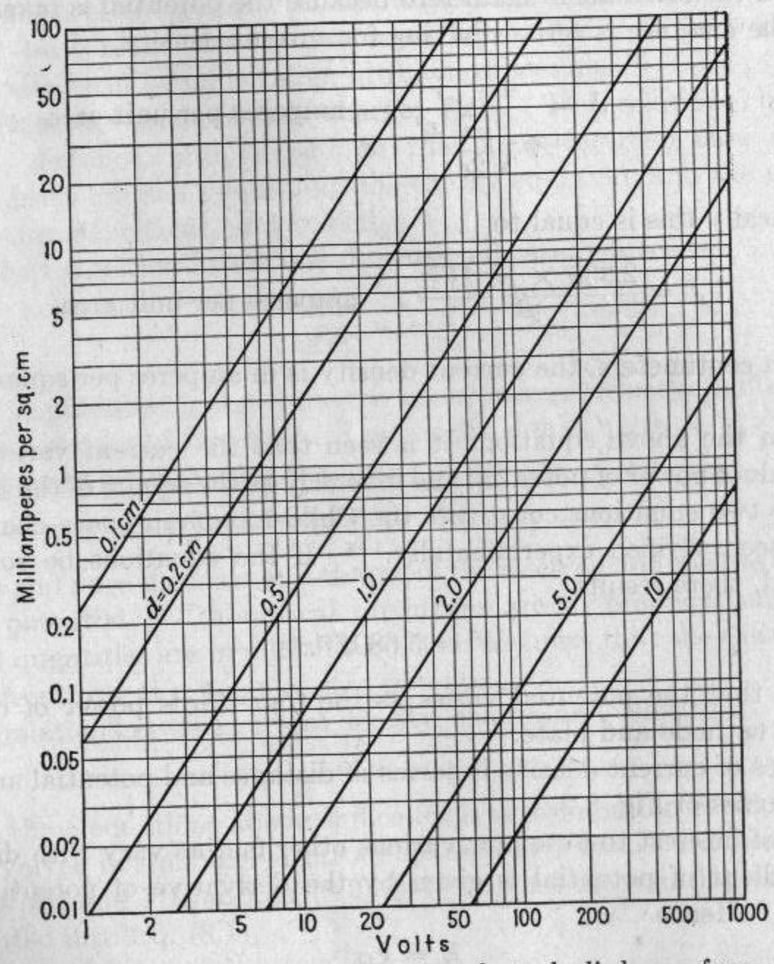


FIG. 8.2.—Current density in a plane-electrode diode as a function of voltage and electrode spacing, Eq. (8.8).

at the cathode. Since the velocity was assumed zero at this point, the space-charge density is theoretically infinite here. Actually, the electrons start with a small finite velocity rather than with zero velocity so that the gradient is initially negative for a small distance, passes through a

minimum at a small negative potential, and then increases. In spite of this difference, the Child-Langmuir law is quite accurate except for very low values of potential corresponding to the average velocity of emission, which is of the order of a few volts.

8.3. Cylindrical-electrode Space-charge Flow. When the electrode structure consists of two concentric cylinders the inner of which is capable of emitting electrons, the space-charge-limited current still varies with the

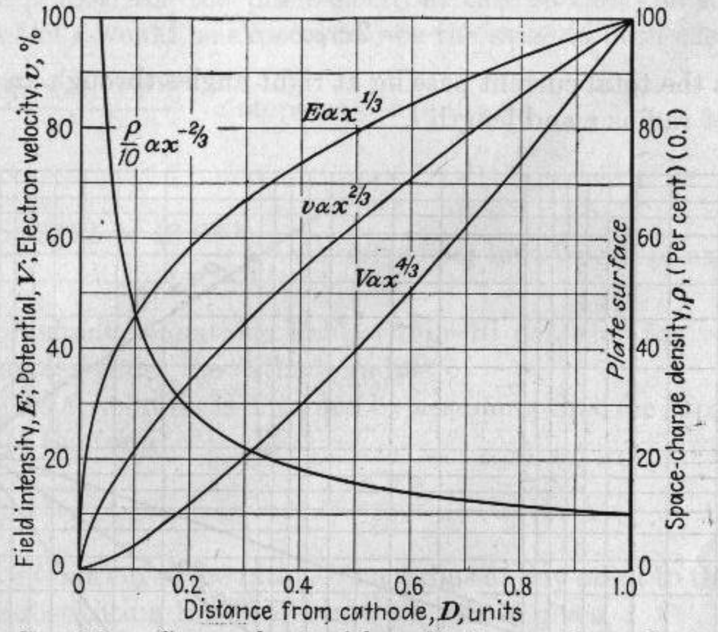


FIG. 8.3*a*.—Curves of potential, gradient, space-charge density, and velocity as a function of interelectrode distance in a space-charge-saturated plane diode, linear scales.

three-halves power of the voltage but the effect of the electrode dimensions is a little more complex.

The behavior of the cylindrical diode can be studied in just the same way as the plane-electrode diode, but in this case it is necessary to use cylindrical coordinates. Actually, this does not complicate the problem too much, for the conditions of symmetry are such that, at any fixed radial distance, conditions are the same regardless of angle. The problem is thus still a one-dimensional one.

For this case Poisson's equation reduces to

$$\frac{1}{r}\frac{d}{dr}\left(\frac{r\,dV}{dr}\right) = -\frac{\rho}{\varepsilon_0} \tag{8.13}$$

The same energy equation as for the plane-electrode case holds,

$$\tfrac{1}{2}mv^2 = eV \tag{8.2}$$

Current density at any radial distance r is given by

$$J(r) = \rho v \tag{8.14}$$

Since the current density varies with radial distance, it is more convenient to express the space-charge density in terms of total current per unit length of axis.

$$\frac{I}{l} = 2\pi r J(r) \tag{8.15}$$

$$= 2\pi r \rho v \tag{8.16}$$

where I is the total current passing at right angles through an imaginary cylinder of radius r and length l.

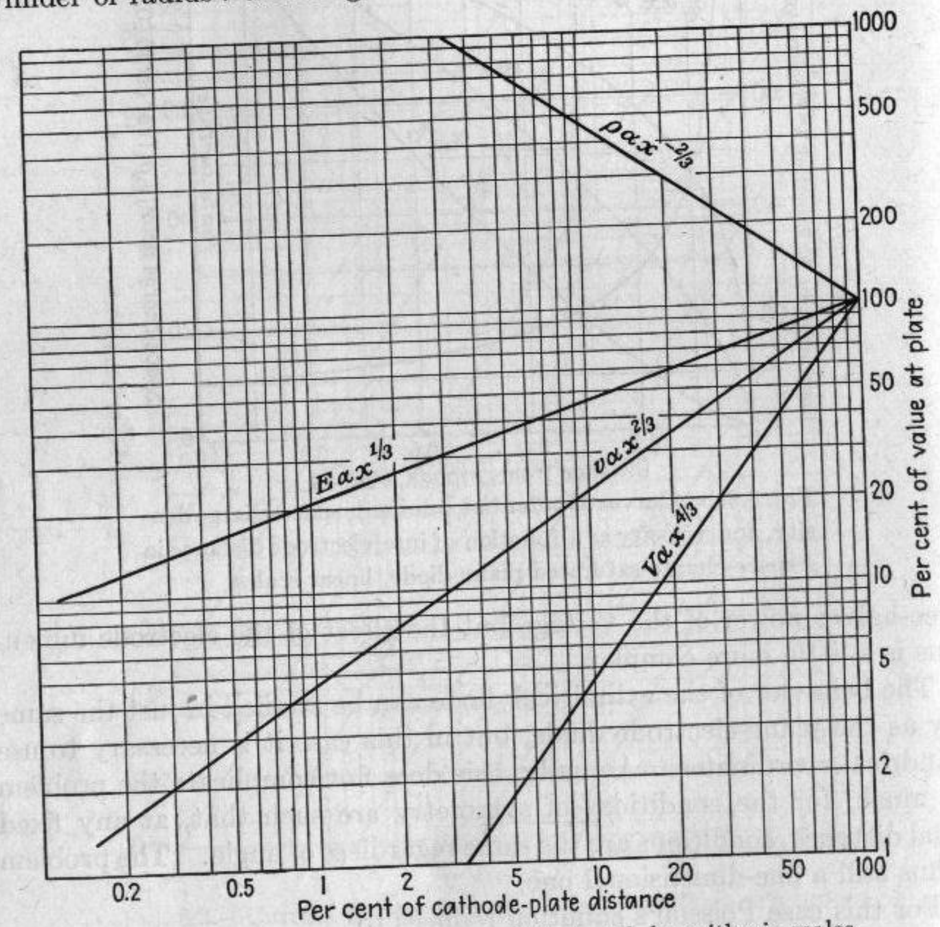


FIG. 8.3*b*.—Same as Fig. 8.3*a*, plotted with logarithmic scales.

Expressing the right side of Eq. (8.13) in terms of I and V from Eqs. (8.2) and (8.16),

$$\frac{d}{dr}\left(r\frac{dV}{dr}\right) = \frac{I}{2\pi\varepsilon_0 l}\sqrt{\frac{m}{2eV}} \tag{8.17}$$

which may also be written as

$$\frac{d^2V}{dr^2} + \frac{1}{r}\frac{dV}{dr} = \frac{J}{\varepsilon_0}\sqrt{\frac{m}{2eV}} \tag{8.18}$$

An approximate solution of Eq. (8.18) can be obtained by observing that for large values of r the second term of the left member of the equation is negligible. If this term is dropped, the resulting equation is the same as Eq. (8.4) for the plane-electrode case so that the solution for large values of r would be expected to be the same in both cases. Thus

$$J = \frac{2.335 \times 10^{-6} V^{3/2}}{r^2} \quad \text{amperes per unit area} \tag{8.19}$$

or

$$\frac{I}{l} = \frac{2\pi \times 2.335 \times 10^{-6} V^{3/2}}{r} \quad \text{amperes per unit length of axis} \tag{8.20}$$

These approximate equations hold within 10 per cent for values of r greater than ten times the cathode radius.

A more exact solution is obtained by assuming that the expression for current is of the form

$$\frac{I}{l} = \frac{KV^{3/2}}{r\beta^2} \tag{8.21}$$

where β^2 is a function of the ratio of the radius at any point to the cathode radius.[1] Substituting Eq. (8.21) into Eq. (8.17) gives

$$3\beta r^2 \frac{d^2\beta}{dr^2} + r^2\left(\frac{d\beta}{dr}\right)^2 + 7\beta r\frac{d\beta}{dr} + \beta^2 - 1 = 0 \tag{8.22}$$

This can be simplified slightly by letting $u = \ln\left(\frac{r}{r_c}\right)$, a logical substitution because the space-charge-free potential is expressible in this form. With this change of variable,

$$3\beta\frac{d^2\beta}{du^2} + \left(\frac{d\beta}{du}\right)^2 + 4\beta\frac{d\beta}{du} + \beta^2 - 1 = 0 \tag{8.23}$$

This may be solved by series, the solution being

$$\beta = u - \frac{2u^2}{5} + \frac{11u^3}{120} - \frac{47u^4}{3{,}300} + \cdots \tag{8.24}$$

This expression is valid for either an internal or an external cathode.

[1] LANGMUIR, *op. cit.*

Values of β^2 have been calculated by Langmuir by means of this series and other equivalent expressions which are more convenient for large values of u. In Table V are tabulated the values of β^2 and in Fig. 8.4 is shown a curve of β^2 as a function of $\frac{r}{r_c}$. It is seen that for ratios of $\frac{r}{r_c}$ greater than 7 the value of β^2 differs from 1 by less than 10 per cent, thus

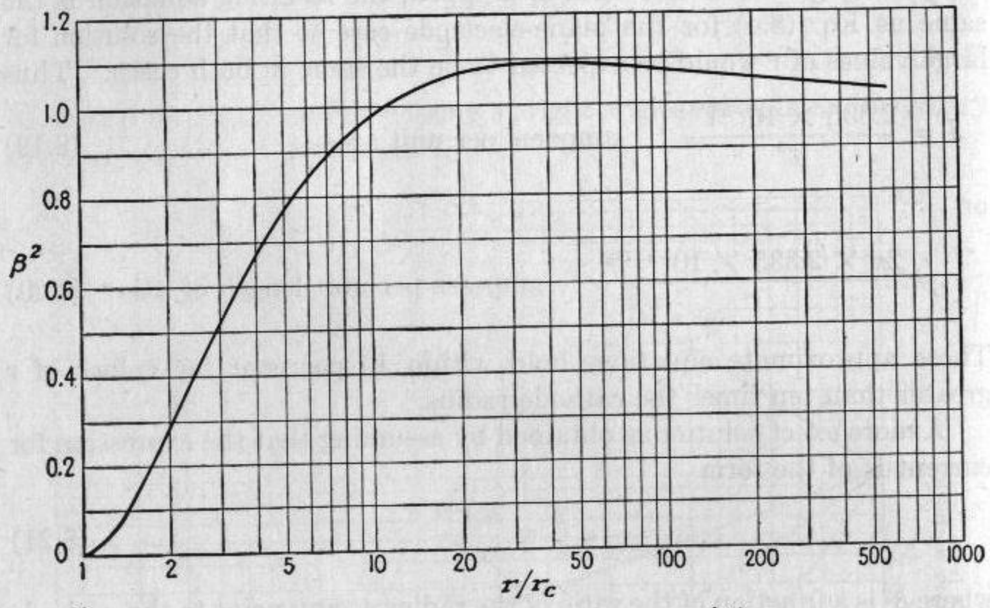


FIG. 8.4.—Curve of β^2 as a function of $\left(\frac{r}{r_c}\right)$.

substantiating the approximate form of the equation for current, previously given. In the chart of Fig. 8.5 is given the current per unit length of axis for various voltages and various values of plate radius for the case in which β^2 has the value of 1. For other values of β^2 the current obtained from this chart must be divided by the value of β^2 as obtained from Fig. 8.4.

In practical units the expression for current is

$$\frac{I}{l} = \frac{14.66 \times 10^{-6} V^{3/2}}{r\beta^2} \quad \text{amperes per unit length of axis} \qquad (8.25)$$

Since the current per unit length of axis is a constant, for a constant plate voltage, the variation of potential with radial distance r is given by

$$V = k_1(r\beta^2)^{2/3} \qquad (8.26)$$

in which it must be recognized that β^2 is not a constant but a function of the radial distance r. Velocity of the electron is given by the square

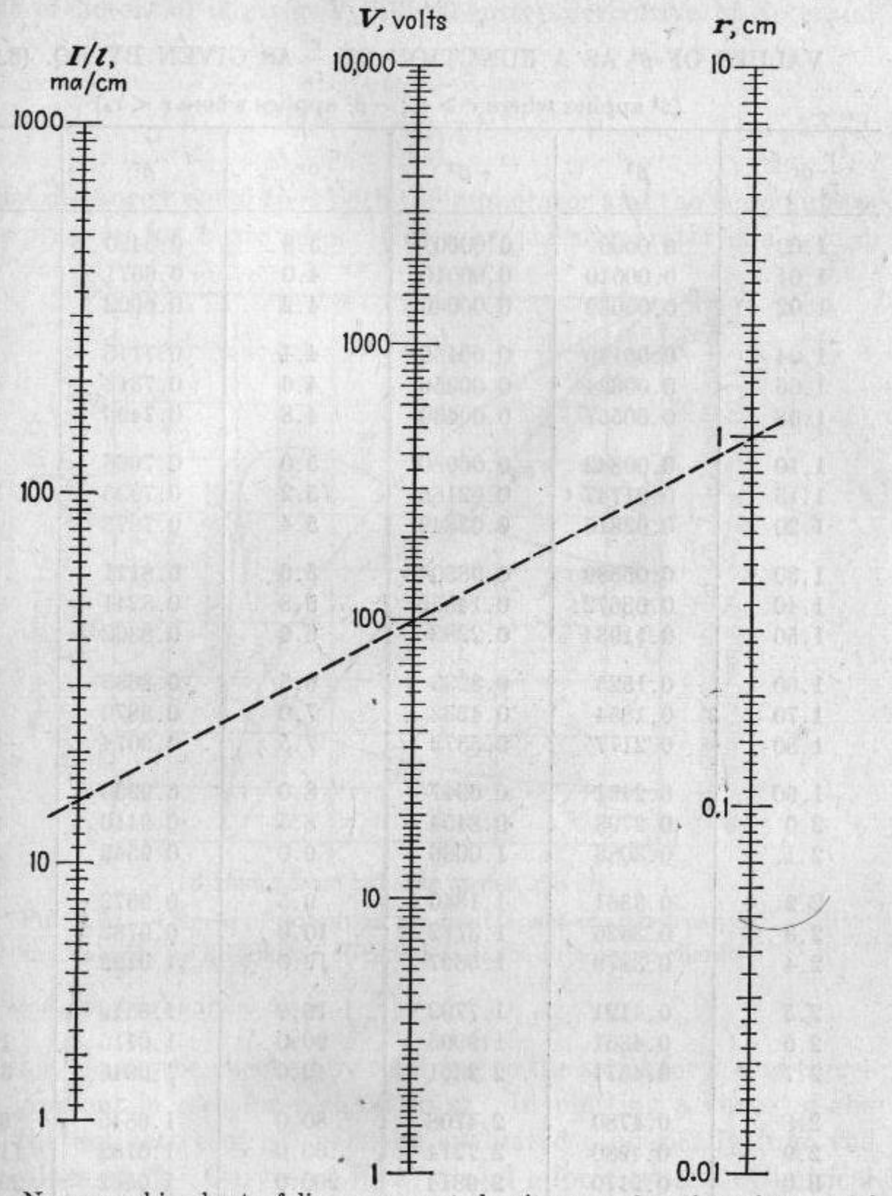


FIG. 8.5.—Nomographic chart of linear current density as a function of potential and plate radius in a cylindrical diode ($\beta^2 = 1$).

root of potential, and thus

$$v = k_2(r\beta^2)^{1/3} \qquad (8.27)$$

From Eq. (8.16), space-charge density is the reciprocal of the product of velocity and radial distance r,

$$\rho = \frac{k_3}{r(r\beta^2)^{1/3}} \qquad (8.28)$$

TABLE V

VALUES OF β^2 AS A FUNCTION OF $\frac{r}{r_c}$ AS GIVEN BY EQ. (8.24)

(β^2 applies where $r > r_c$; $-\beta^2$ applies where $r < r_c$)

$\frac{r}{r_c}$ or $\frac{r_c}{r}$	β^2	$-\beta^2$	$\frac{r}{r_c}$ or $\frac{r_c}{r}$	β^2	$-\beta^2$
1.00	0.0000	0.0000	3.8	0.6420	5.3795
1.01	0.00010	0.00010	4.0	0.6671	6.0601
1.02	0.00039	0.00040	4.2	0.6902	6.7705
1.04	0.00149	0.00159	4.4	0.7115	7.5096
1.06	0.00324	0.00356	4.6	0.7313	8.2763
1.08	0.00557	0.00630	4.8	0.7496	9.0696
1.10	0.00842	0.00980	5.0	0.7666	9.887
1.15	0.01747	0.02186	5.2	0.7825	10.733
1.20	0.02815	0.03849	5.4	0.7973	11.601
1.30	0.05589	0.08504	5.6	0.8111	12.493
1.40	0.08672	0.14856	5.8	0.8241	13.407
1.50	0.11934	0.2282	6.0	0.8362	14.343
1.60	0.1525	0.3233	6.5	0.8635	16.777
1.70	0.1854	0.4332	7.0	0.8870	19.337
1.80	0.2177	0.5572	7.5	0.9074	22.015
1.90	0.2491	0.6947	8.0	0.9253	24.805
2.0	0.2793	0.8454	8.5	0.9410	27.701
2.1	0.3083	1.0086	9.0	0 9548	30.698
2.2	0.3361	1.1840	9.5	0.9672	33.791
2.3	0.3626	1.3712	10.0	0.9782	36.976
2.4	0.3879	1.5697	12.0	1.0122	50.559
2.5	0.4121	1.7792	16.0	1.0513	81.203
2.6	0.4351	1.9995	20.0	1.0715	115.64
2.7	0.4571	2.2301	40.0	1.0946	327.01
2.8	0.4780	2.4708	80.0	1.0845	867.11
2.9	0.4980	2.7214	100.0	1.0782	1174.9
3.0	0.5170	2.9814	200.0	1.0562	2946.1
3.2	0.5526	3.5293	500.0	1.0307	9502.2
3.4	0.5851	4.1126	∞	1.000	∞
3.6	0.6148	4.7298			

For small u

$$\beta^2 = u^2(1 - 0.8u + 0.344u^2 + \cdots)$$

$$\frac{d\beta^2}{du} = 2u - 2.4u^2 + 1.374u^3 - 0.509u^4 + \cdots$$

where $u = \ln \frac{r}{r_c}$.

Gradient of potential is given by the negative derivative of potential,

$$E = \frac{-k_4 \frac{d(r\beta^2)}{dr}}{(r\beta^2)^{1/3}} \tag{8.29}$$

For radial distance r equal to r_c both the numerator and the denominator of the expression for E are zero. However, the numerator is a zero of

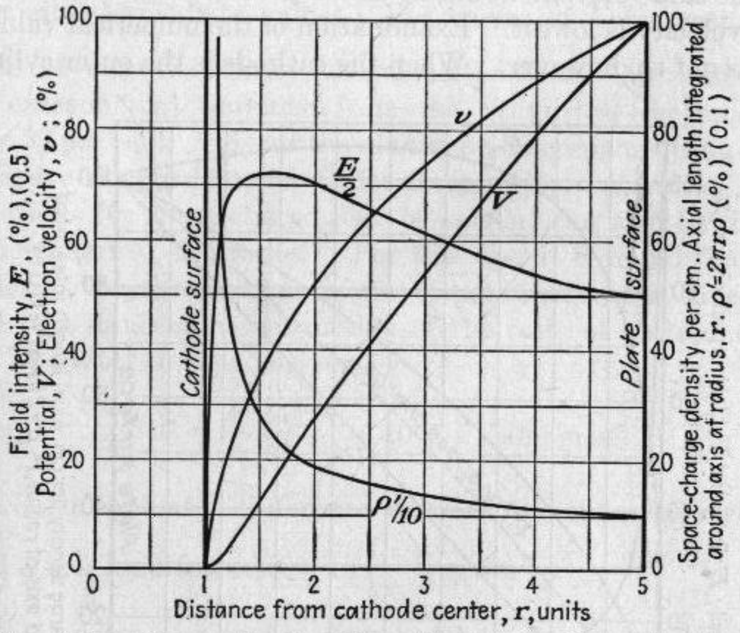


FIG. 8.6*a*.—Curves of potential, gradient, space-charge density, and velocity as a function of radial distance in a space-charge-saturated cylindrical diode, $\left(\frac{r_p}{r_c}\right) = 5$.

higher order, as may be checked by referring to the series for β, and therefore the gradient is zero for r equal to r_c. In plotting a curve of the gradient the derivative of $r\beta^2$ must be evaluated numerically from the table of values of β^2. Curves of V, E, v, and ρ for a typical cylindrical triode are shown in Fig. 8.6 for two ratios of plate to cathode radii. Curves for other radii have forms similar to those shown. Although not evident from the appearance of the curves, the potential profile leaves the cathode with zero slope, having a considerable change of slope in a short distance. This can also be seen from the series expansion of the expression for potential near the cathode, the first term of which is

$$V = k_1 r_c^{2/3} \left(\frac{r - r_c}{r_c}\right)^{4/3} \tag{8.30}$$

From this it can also be seen that the gradient has the value of zero at the cathode.

In all the foregoing it has been assumed that the cathode cylinder is smaller than the plate cylinder. The formulas are also valid if the cathode is the outer cylinder, though tubes are rarely built this way. It might be thought that if the cathode were the outer cylinder a greater current would flow for a given voltage than if it were the inner cylinder because the same current would be distributed through a greater volume where its velocity is lowest. Examination of the numerical values shows that this is not so, however. When the cathode is the outer cylinder, the

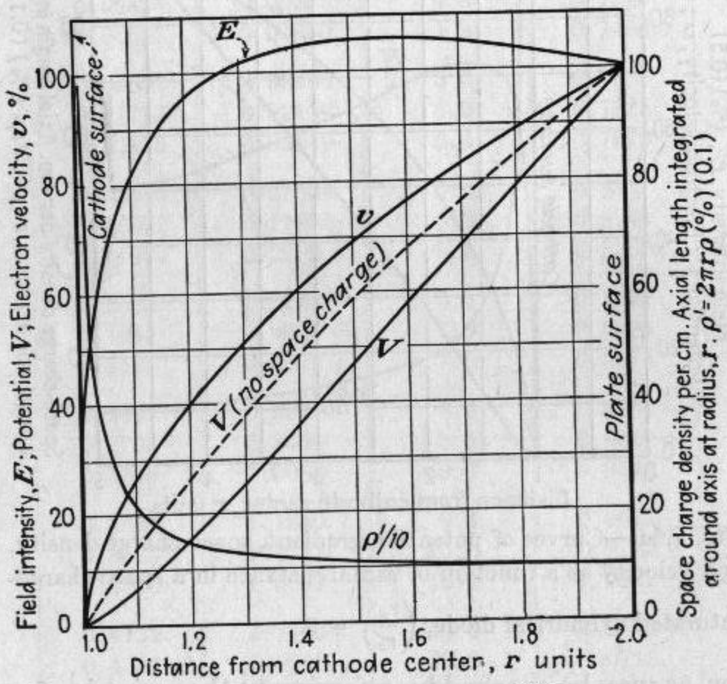


FIG. 8.6*b*.—Same as Fig. 8.6*a*, $\left(\frac{r_p}{r_c}\right) = 2$.

functon β^2 increases very rapidly as the ratio of plate to cathode radius decreases. As a result, the current is actually less when the cathode is the outer cylinder; for although the factor r in Eq. (8.21) is decreased, the factor β^2 has increased more than enough to offset this.

It is of interest to compare Eqs. (8.8) and (8.25) for the plane and cylinder case. If Eq. (8.25) is divided by $2\pi r$, it then resembles Eq. (8.8) except that x is replaced by $r\beta$. When r is very large, β approaches unity and the expressions become identical. When r is only slightly larger than r_c, β is approximately equal to $\ln\left(\frac{r}{r_c}\right)$ and hence to $\frac{r - r_c}{r_c}$, with the result that the factor $r\beta$ nearly equals $r - r_c$ and the expressions for current density again approach identity with x replaced by $r - r_c$. This identity is expected for values of $\frac{r}{r_c}$ near unity; for here the electrode spacing is small compared with the radii, and a plane structure is approximated. The fact that Eqs. (8.8) and (8.25) are identical for limiting values of the ratio $\frac{r}{r_c}$ does not mean that they are nearly equal for all intermediate values. For values of $\frac{r}{r_c}$ between 4 and 20 the value of current density as calculated from the plane-electrode formula exceeds that obtained from the cylindrical-electrode formula by nearly 20 per cent. This is the maximum discrepancy that can occur.

8.4. Space-charge Flow for Other Geometries. *Spherical Electrodes.* The equations for space-charge flow of current can also be derived for concentric spherical electrodes.[1] For this case it is found that the current varies as the three-halves power of potential and is inversely proportional to a dimensionless function of the ratio of plate and cathode radius. The total current is given by

$$I = \frac{29.34V^{3/2}}{\alpha^2} \times 10^{-6} \quad \text{amperes} \tag{8.31}$$

where V is the potential difference between the spherical electrodes in volts and α is a function of $u = \ln\left(\frac{r}{r_c}\right)$, given by

$$\alpha = u - 0.3u^2 + 0.075u^3 - 0.00143u^4 + 0.00216u^5 - \cdots \tag{8.32}$$

This expression for α is valid whether the cathode is external or internal.

Values of α for the spherical case and of β for the cylindrical case are equal within 2 per cent for values of $\frac{r}{r_c}$ between 0.65 and 1.35. For values of $\frac{r}{r_c}$ less than 1, α differs not more than 10 per cent from the larger value of β.[2] For values of $\frac{r}{r_c}$ very nearly equal to 1 the current density approaches that for plane electrodes.

The General Case. The observation that the current varies with the three-halves power of potential for plane, cylindrical, and spherical electrodes leads one to believe that this is the case for electrodes of any shape. Actually this is so, but the conclusion must be examined carefully, for the three cases enumerated are special cases in which the elec-

[1] LANGMUIR, I., and K. T. COMPTON, Electrical Discharges in Gases, Part II, *Rev. Modern Phys.*, vol. 13, pp. 191–257, April, 1931.

[2] See Appendix VII for values of α^2.

trons move in straight lines perpendicular to the equipotential surfaces. For other geometries this is not necessarily so. For electrodes of other shapes the electrons will in general cut across the lines of electric force and move in curved paths.

The validity of the three-halves-power law of potential can however be shown by a simple dimensional analysis of the basic equations from which the current laws were developed. These are Poisson's equations,

$$\nabla^2 V = -\frac{\rho}{\varepsilon_0} \tag{8.33}$$

the energy equation

$$\tfrac{1}{2}mv^2 = eV \tag{8.34}$$

and the current-density expression

$$J = \rho v \tag{8.35}$$

In previous cases the quantities in the last equation above have been treated as scalar quantities, but in the general case current density and velocity must be treated as vector quantities because they do not necessarily have the direction of the gradient of potential.

Consider now what happens if the potential is increased by a factor k. If the electrons move in a curved path, their centrifugal force $\frac{mv^2}{R}$, where R is the instantaneous radius of curvature, must equal the force due to the component of the gradient of potential normal to the path, $e\nabla_n V$. From Eq. (8.34) the centrifugal force will have increased by a factor k, and likewise the gradient will have increased by the same factor, so that the shape of the electron paths will be unchanged. This is the same conclusion that was reached in the case of the space-charge-free fields. Once this is established, the final conclusion follows immediately. From Eq. (8.33) the space-charge density is increased by a factor k, and from Eq. (8.34) the velocity is increased by a factor $k^{1/2}$. Hence, by Eq. (8.35), the current density is increased by a factor $k^{3/2}$, and the general validity of the three-halves-power law of potential for electrons starting from rest is established.

It can also be shown by the same type of reasoning that, if an electrode structure is enlarged by a factor n maintaining geometrical similarity and voltages are kept unchanged, then if the current is space-charge-limited, the current is also unchanged. From Eq. (8.33) space-charge density is decreased by a factor n^2, and hence current density is decreased by a factor of n^2 from Eq. (8.35), velocity at corresponding points in the original and enlarged structure being the same. However, since the total area over which the current density is summed is increased by a factor of

n^2, it is seen that the total current is unchanged if the voltages are unchanged.

8.5. Current Law for Plane Triodes. It is found experimentally in triodes that the total current released from the emitter is very nearly proportional to the three-halves power of the equivalent voltage, $V_g + \frac{V_p}{\mu}$. This is most readily shown by plotting curves of constant space current,

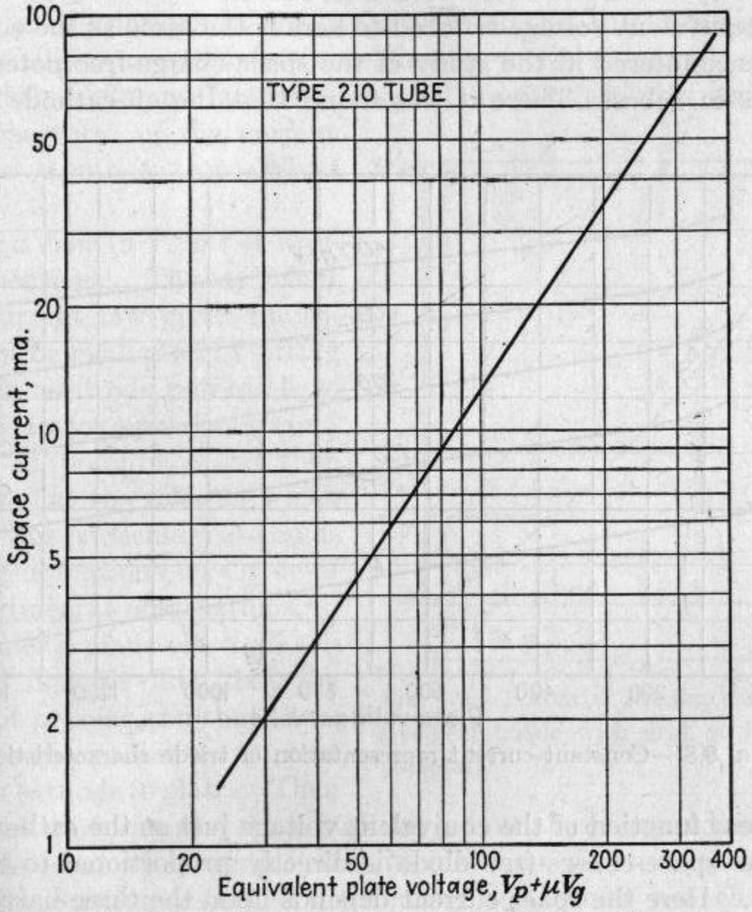


FIG. 8.7.—Current as a function of equivalent voltage in a 210 triode.

i.e., the sum of plate and grid current, against equivalent voltage on logarithmic paper. All points and curves so plotted tend to fall on the same straight line, which has a slope of nearly $\frac{3}{2}$. Slight departures from the slope of $\frac{3}{2}$ are sometimes encountered because of initial electron velocity, the Schottky effect, and because of potential drop along the emitter. The relation holds whether the grid is negative or positive even though the space current in the first case is all plate current and in the

second case is the sum of grid and plate current. This is in accord with the reasoning presented in the previous section.

The correlation between the theory and experiment is sufficiently good so that in general it is a very good approximation to write

$$J = J_p + J_g = k\left(V_g + \frac{V_p}{\mu}\right)^{3/2} \tag{8.36}$$

in which k is a constant to be determined later.

The equivalent voltage referred to here is the same as the equivalent voltage encountered in the study of the space-charge-free potential distributions in tubes. There it was found that the off-cathode gradient

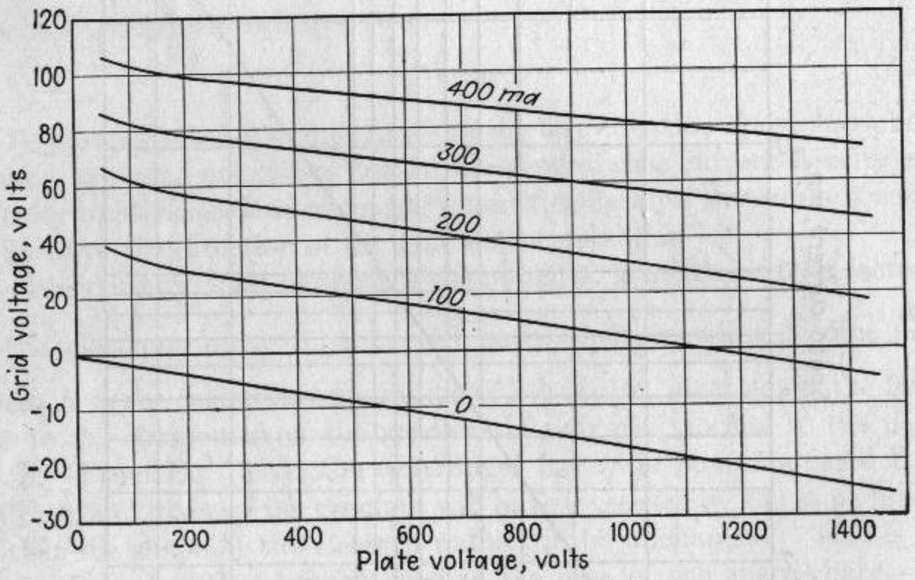


FIG. 8.8.—Constant-current representation of triode characteristics.

was a linear function of the equivalent voltage just as the cathode gradient of the space-charge-free diode is directly proportional to the plate potential. Here the space current depends upon the three-halves power of the equivalent potential just as the saturated diode current is proportional to the three-halves power of the plate potential. This is strictly true only when the space charge in the grid-plate region is negligible. A curve showing how the space current varies with equivalent voltage in a typical triode is given in Fig. 8.7.

The fact that the relative effectiveness of the plate and grid in controlling current flow is the same for a great range of current is shown by plotting contours of constant space current against axes of grid and plate voltage. Such a set of curves is shown in Fig. 8.8. It is seen that the

curves are substantially straight lines whose slope by definition is $-\frac{1}{\mu}$. The constancy of the slope of these straight-line curves attests to the constancy of the amplification factor. It will be observed that for very small currents the slope decreases, indicating a lower amplification factor. This is explained by the fact that the amplification factor of all parts of the electrode structure is not the same, with the result that the parts with the highest μ cut off first, leaving current to pass through the parts of lower μ. Variations in the magnitude of the amplification factor, which actually is an equivalent amplification factor of a number of areas with slightly different factors connected in parallel, also account for slight departures of the current law from a strict three-halves power law.

Current Law in Terms of Electrode Dimensions. The coefficient of the current law given in Eq. (8.36) can be evaluated by fitting the triode electrode potentials to the diode law for a particular combination of potentials and then assuming that the relation which holds for this particular case holds for all combinations by virtue of the experimental observations.[1]

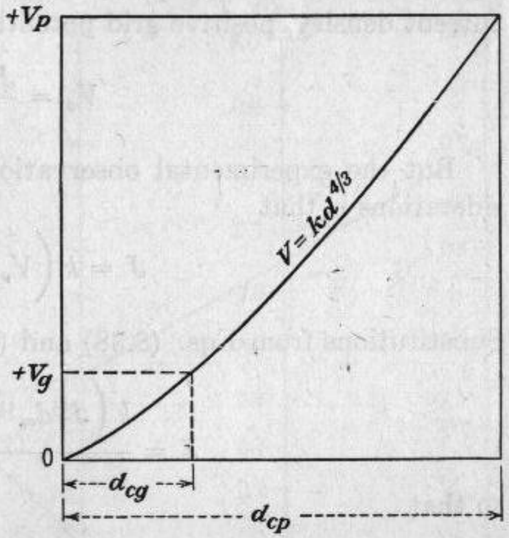


FIG. 8.9.—Potential distribution in a positive-grid triode with grid at its natural potential.

Consider a plane-electrode triode, and imagine first that the grid is not present and that there is a space-charge-limited current flow from cathode to plate. Then if the grid were inserted at a positive potential corresponding to that which existed at its location before its insertion, its presence would not disturb the existing potential distribution and would not change the magnitude of the plate current. Since the equivalent voltage can now be determined for a given current, the constant of Eq. (8.36) can be evaluated. When this constant is known for one combination of potentials, our experimental observations show that it is the same for all combinations of potentials and thus the current law for triodes is determined in terms of the electrode dimensions.

The distribution of potentials referred to above is shown in Fig. 8.9. Here it is seen that the potential distribution from cathode to plate is a

[1] FREMLIN, J. H., Calculation of Triode Constants, *Elec. Commun.*, July, 1939.

four-thirds-power-law curve as in the case of the diode. On this basis, the relation between the current density, the diode plate potential V_p, and the cathode-plate distance is

$$J = \frac{KV_p^{3/2}}{d_{cp}^2} \quad \text{amperes per unit area} \tag{8.37}$$

where K is 2.335×10^{-6}. Let this be written in the form

$$V_p = \frac{J^{2/3} d_{cp}^{4/3}}{K^{2/3}} \tag{8.38}$$

The grid is inserted with the potential that would exist in the diode at the location of the grid plane as shown in the figure. The relation for current density, positive grid potential, and cathode-grid distance is

$$V_g = \frac{J^{2/3} d_{cg}^{4/3}}{K^{2/3}} \tag{8.39}$$

But the experimental observation in keeping with theoretical considerations is that

$$J = k\left(V_g + \frac{V_p}{\mu}\right)^{3/2} \tag{8.40}$$

Substitutions from Eqs. (8.38) and (8.39) give

$$J = \frac{k\left(J^{2/3} d_{cg}^{4/3} + \dfrac{J^{2/3} d_{cp}^{4/3}}{\mu}\right)^{3/2}}{K} \tag{8.41}$$

so that

$$k = \frac{K}{\left(d_{cg}^{4/3} + \dfrac{d_{cp}^{4/3}}{\mu}\right)^{3/2}} \tag{8.42}$$

or, numerically,

$$k = \frac{2.335 \times 10^{-6}}{\left(d_{cg}^{4/3} + \dfrac{d_{cp}^{4/3}}{\mu}\right)^{3/2}} \tag{8.43}$$

As a result, the expression for current is

$$J = \frac{2.335 \times 10^{-6}\left(V_g + \dfrac{V_p}{\mu}\right)^{3/2}}{d_{cg}^2\left[1 + \dfrac{1}{\mu}\left(\dfrac{d_{cp}}{d_{cg}}\right)^{4/3}\right]^{3/2}} \quad \text{amperes per unit area} \tag{8.44}$$

This is the expression that has been sought and that has been the object of the above development. It is seen to be of the same form as the expression for the diode current density of Eq. (8.8). There is good experimental verification of Eq. (8.44).

It may also be seen that the space-charge-saturated equivalent-diode spacing of the triode is

$$d_{se} = d_{cg}\left[1 + \frac{1}{\mu}\left(\frac{d_{cp}}{d_{cg}}\right)^{4/3}\right]^{3/4} \tag{8.45}$$

A nomograph giving the space-charge-saturated equivalent-diode spacing of a triode in terms of the cathode-grid distance, the cathode-plate distance, and the amplification factor is given in Fig. 8.10. It should be noted that the space-charge-saturated equivalent-diode spacing of a triode given here is somewhat different from the space-charge-free equivalent-diode spacing given in Eq. (7.53) of the chapter on Electrostatic Field of a Triode. Each is slightly greater than the cathode-grid distance. For a μ of 10 and a ratio of cathode-plate to cathode-grid spacing of 5 the value of diode spacing from Eq. (8.45) is only 6 per cent greater than that from Eq. (7.53). For large values of μ the difference is even less. Various attempts have been made to devise expressions for the triode current and mutual conductance in terms of the space-charge-

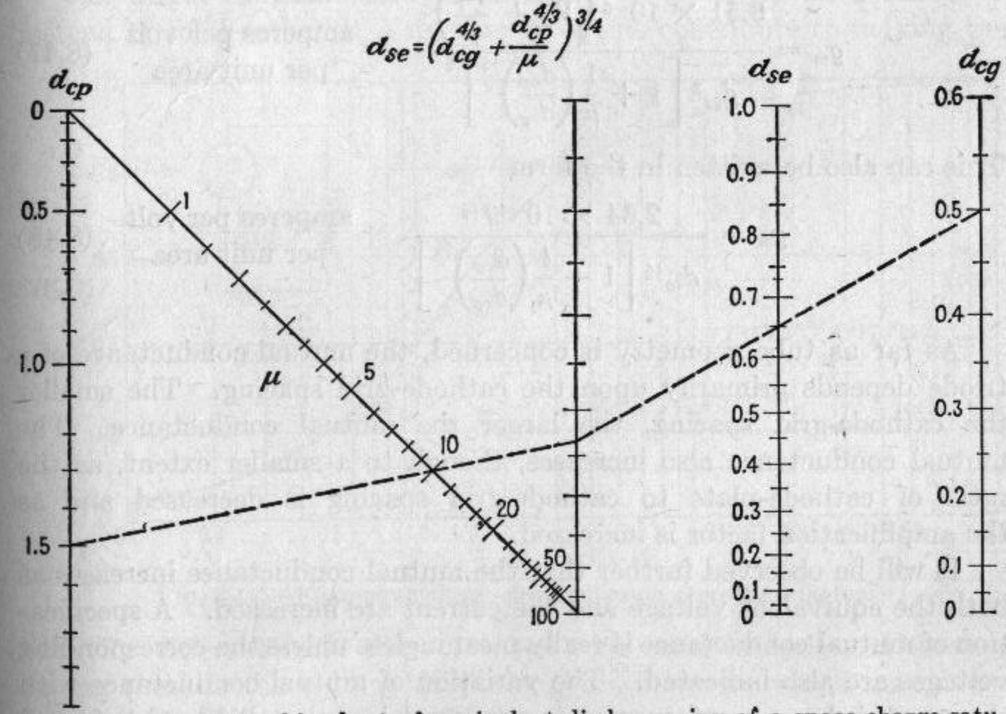


FIG. 8.10.—Nomographic chart of equivalent-diode spacing of a space-charge-saturated plane-electrode triode.

free equivalent-diode spacing, but all these are subject to an inescapable error.[1]

8.6. Mutual Conductance of a Plane Triode. The mutual conductance is easily obtained from the expression for current density of Eq. (8.44). By definition,

$$g_m = \frac{\partial I}{\partial V_g} = \left(\frac{dI_p}{dV_g}\right)_{V_p=k} \tag{8.46}$$

Performing this operation on Eq. (8.44),

$$g_m = \frac{3.51 \times 10^{-6}\left(V_g + \frac{V_p}{\mu}\right)^{1/2}}{d_{cg}^2\left[1 + \frac{1}{\mu}\left(\frac{d_{cp}}{d_{cg}}\right)^{1/3}\right]^{3/2}} \quad \text{amperes per volt per unit area} \tag{8.47}$$

This can also be written in the form

$$g_m = \frac{2.64 \times 10^{-4} J^{1/3}}{d_{cg}^{4/3}\left[1 + \frac{1}{\mu}\left(\frac{d_{cp}}{d_{cg}}\right)^{1/3}\right]} \quad \text{amperes per volt per unit area} \tag{8.48}$$

As far as tube geometry is concerned, the mutual conductance of a triode depends primarily upon the cathode-grid spacing. The smaller the cathode-grid spacing, the larger the mutual conductance. The mutual conductance also increases, though to a smaller extent, as the ratio of cathode-plate to cathode-grid spacing is decreased and as the amplification factor is increased.

It will be observed further that the mutual conductance increases as both the equivalent voltage and the current are increased. A specification of mutual conductance is really meaningless unless the corresponding voltages are also indicated. The variation of mutual conductance with the one-third power of current is a general law that holds well for all types of tubes, including pentodes as well as triodes.

8.7. Mutual Conductance of a Cylindrical Triode. The current law and mutual conductance of cylindrical triodes are readily evaluated by an analysis similar to that used for the plane-electrode triode. The current is given by

$$\frac{I}{l} = \frac{14.66 \times 10^{-6}\left(V_g + \frac{V_p}{\mu}\right)^{3/2}}{r_g\beta_{cg}^2\left[1 + \frac{1}{\mu}\left(\frac{r_p\beta_{cp}^2}{r_g\beta_{cg}^2}\right)^{2/3}\right]^{3/2}} \quad \text{amperes per unit length} \tag{8.49}$$

[1] WALKER, G. B., Theory of the Equivalent Diode, *Wireless Engr.*, vol. 24, pp. 5–7, January, 1947.

where r_g and r_p are grid and plate radii and the combination subscripts indicate that β^2 is to be determined by the ratio of the radii of the electrodes indicated by the subscripts.

The mutual conductance is

$$g_m = \frac{22.0 \times 10^{-6}\left(V_g + \frac{V_p}{\mu}\right)^{1/2}}{r_g\beta_{cg}^2\left[1 + \frac{1}{\mu}\left(\frac{r_p\beta_{cp}^2}{r_g\beta_{cg}^2}\right)^{2/3}\right]^{3/2}} \quad \text{amperes per volt per unit length} \tag{8.50}$$

8.8. Effect of Filamentary Emitters. When tubes have filaments instead of solid cathodes, a number of effects contribute to making the

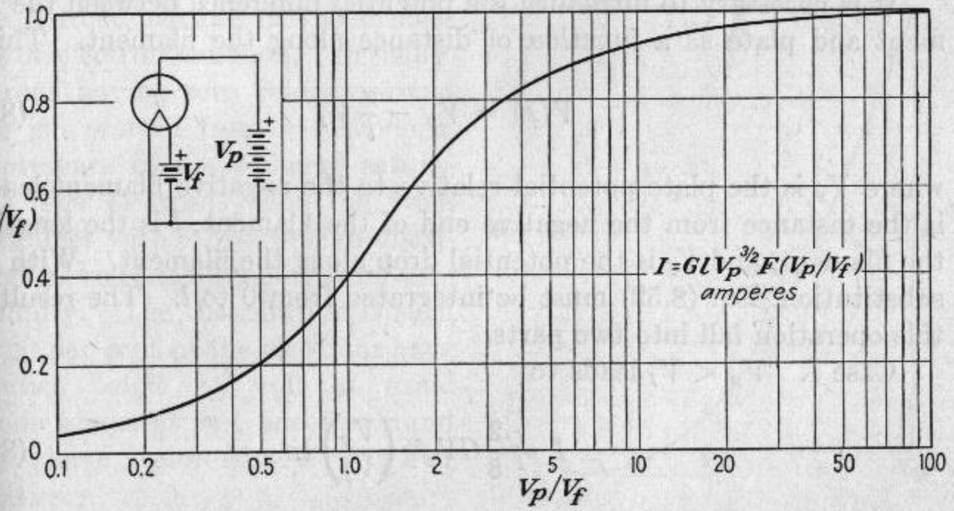


FIG. 8.11.—Filament voltage-drop emission correction factor.

behavior different from that of tubes with solid cathodes. Foremost among these is the voltage drop along the filament, which may cause considerable divergence from the simple three-halves power of current with voltage. If the plate current is returned to the negative filament lead, the current is at all times less than that for a unipotential cathode but becomes nearly equal to that value for large values of the ratio of plate to filament voltage. The ratio of the current without to that with a unipotential filament is shown in Fig. 8.11. This is essentially a correction factor for the fact that the filament potential is not uniform. The ratio of currents is 0.4 for a plate- to filament-voltage ratio of 1, dropping linearly to 0 with this ratio for plate voltages less than the filament voltage. When the plate voltage exceeds the filament voltage by more than a factor of 15, the current ratio is within 5 per cent of unity.

The above results are arrived at by integrating the emission effect along the filament, taking into account the different potential differences

to the plate at each point. Every element of length of the filament contributes a current

$$dI = \frac{14.66 \times 10^{-6}[V(x)]^{3/2}\,dx}{r_p\beta^2} \tag{8.51}$$

which may be written more simply as

$$dI = G[V(x)]^{3/2}\,dx \tag{8.52}$$

in which G is the so-called "perveance" of the tube. The perveance is simply the coefficient of the voltage factor in the Child-Langmuir law. It contains the geometrical factors of the tube and has dimensions of current per unit length per volt$^{3/2}$.

It is necessary to introduce the potential difference between the filament and plate as a function of distance along the filament. This is

$$V(x) = V_p - \frac{x}{l}V_f \tag{8.53}$$

where V_p is the plate potential relative to the negative filament lead, x is the distance from the negative end of the filament, l is the length of the filament, and V_f is the potential drop along the filament. With this substitution, Eq. (8.52) must be integrated from 0 to l. The results of this operation fall into two parts.

Case I. $V_p < V_f$ leads to

$$I = \frac{2}{5}GV_p^{3/2}\left(\frac{V_p}{V_f}\right)l. \tag{8.54}$$

Case II. $V_p > V_f$ leads to

$$I = \frac{2}{5}\frac{Gl}{V_f}[V_p^{5/2} - (V_p - V_f)^{5/2}] \tag{8.55}$$

For purposes of computation Eq. (8.55) is best put in the form of the series

$$I = GlV_p^{3/2}\left[1 - \frac{3V_f}{4V_p} + \frac{3}{24}\left(\frac{V_f}{V_p}\right)^2 + \frac{3}{192}\left(\frac{V_f}{V_p}\right)^3 + \cdots\right] \tag{8.56}$$

All the above equations for current can be put into the form

$$I = GlV_p^{3/2}F\left(\frac{V_p}{V_f}\right) \tag{8.57}$$

in which $F\left(\frac{V_p}{V_f}\right)$ is the current ratio plotted in Fig. 8.11.

Some difficulty is occasionally encountered in calculating currents

in tubes with filaments and plane electrodes. Neither the plane- nor the cylindrical-electrode formulas will fit directly here. Experimentally it is found that there is an equivalent filament area which may be applied. This equivalent area is obtained by projecting the filament onto the filament plane and surrounding it by a band twice as wide as the distance from the filament to the nearest electrode. This equivalent area may be used either to obtain current in diodes or to obtain current or mutual conductance in triodes. The same concept may be applied to helical filaments in cylindrical-electrode tubes.

8.9. Effect of Initial Electron Velocity. In all the foregoing analyses it has been assumed that the electrons start with zero velocity from a point of zero potential. This is not quite correct because of the mechanism of electron emission. Actually, electrons having zero velocity would never get started from a cathode in the presence of space-charge saturation. The electrons come off from the emitter with a Maxwellian distribution of velocities ranging from zero to infinity. The distribution is such that 90 per cent of the electrons have velocities below ½ volt at usual cathode temperatures, and fewer and fewer have successively higher velocities.

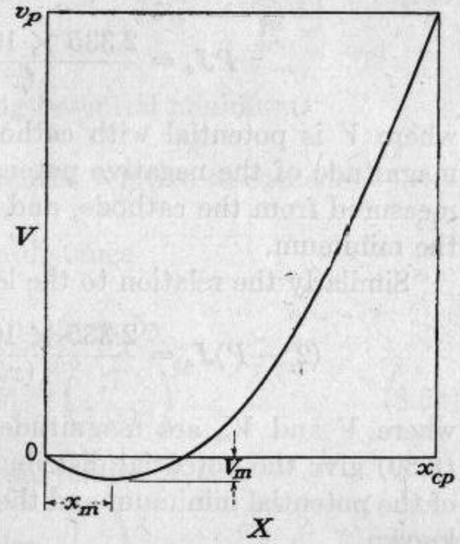


FIG. 8.12.—Potential distribution in a plane-electrode diode for the case of uniform initial velocity of emission.

A good idea of what the actual potential distribution is when the initial velocities are considered may be obtained by assuming that all electrons leave the cathode with the same normal velocity. Because of the initial velocity, the gradient at the cathode may and does become negative, and the potential curve moves down until it has a minimum somewhere close to the cathode at which the potential is negative and corresponds to the velocity of emission. For this condition the electrons are slowed down until they all come to rest at the potential minimum. From this point, which acts like an ideal cathode and which is called a "virtual" cathode, the electrons may start in either direction, either being returned to the cathode or going to the plate.

For a condition of space-charge saturation the potential distribution on either side of the virtual cathode will follow the four-thirds-power law. The location of the potential minimum will be determined by the fraction

of the current that goes on to the plate. A typical potential distribution is shown in Fig. 8.12.

Let it be assumed that the fraction of the emitted current which continues on to the plate is given by P. Then if the emitted current per unit area is J_e, the current to the right of the potential minimum is PJ_e and that to the left is $(2 - P)J_e$. The current to the left of the minimum is made up of the emitted current J_e going in one direction and the returning current $(1 - P)J_e$ going in the other. As far as the space-charge effects are concerned, the directions of these current components are immaterial because the charge densities add regardless of sign.

The relation between current, potential, and distance to the right of the minimum is given by

$$PJ_e = \frac{2.335 \times 10^{-6}(V + V_m)^{3/2}}{(x - x_m)^2} \quad \text{amperes per unit area} \tag{8.58}$$

where V is potential with cathode potential taken as zero, V_m is the magnitude of the negative potential at the minimum, x is the distance measured from the cathode, and x_m is the distance from the cathode to the minimum.

Similarly the relation to the left of the minimum is

$$(2 - P)J_e = \frac{2.335 \times 10^{-6}(V_m + V)^{3/2}}{(x_m - x)^2} \quad \text{amperes per unit area} \tag{8.59}$$

where V and V_m are magnitudes of potential. Equations (8.58) and (8.59) give the potential distribution if the magnitude and the location of the potential minimum and the fraction of the transmitted current are known.

To determine the factors in terms of which Eqs. (8.58) and (8.59) are expressed let these two equations be evaluated at the plate and cathode, respectively. Then

$$PJ_e = \frac{2.335 \times 10^{-6}(V_p + V_m)^{3/2}}{(x_{cp} - x_m)^2} \quad \text{amperes per unit area} \tag{8.60}$$

and

$$(2 - P)J_e = \frac{2.335 \times 10^{-6}(V_m)^{3/2}}{x_m^2} \quad \text{amperes per unit area} \tag{8.61}$$

Taking the ratio of these two expressions and solving for x_m,

$$x_m = \frac{x_{cp}}{1 + \left(\frac{V_p}{V_m} + 1\right)^{3/4}\left(\frac{2}{P} - 1\right)^{1/2}} \tag{8.62}$$

from which the location of the potential minimum can be determined

for any assumed fraction of transmitted current. The magnitude of the potential at the minimum is known from the initial velocity.

The results of the above analysis are not exact because in the actual case there is a distribution of velocities, with the result that the position and magnitude of the potential minimum and also the fraction of the transmitted current are uniquely determined from the potential and cathode temperature. An exact analysis considering the velocity distribution is given by Langmuir.[1] The exact expressions are quite involved, but some approximate expressions which are accurate to within about 2 per cent take the following form:

$$PJ_e = \frac{2.335 \times 10^{-6}(V_p - V_m)^{3/2}}{(x_{cp} - x_m)^2}\left[1 - \frac{0.0247T^{1/2}}{(V_p - V_m)^{1/2}}\right] \tag{8.63}$$

where J_e is emitted current per unit area
PJ_e is current per unit area passing potential minimum
V_p is plate potential
V_m is magnitude of minimum potential relative to cathode
x_{cp} is cathode-plate distance
x_m is cathode–potential-minimum distance
T is cathode temperature, °K

The location of the potential minimum is given by

$$x_m = 0.0156(1{,}000J)^{-1/2}\left(\frac{T}{1{,}000}\right)^{3/4} \tag{8.64}$$

The corresponding magnitude of potential minimum is

$$V_m = \left(\frac{T}{5{,}040}\right)\log_{10}\left(\frac{1}{P}\right) \tag{8.65}$$

where P is the fraction of the emitted current transmitted to the plate.

The exact relations for space-charge-saturated flow with initial electron velocity are given by Langmuir in the form of the universal curve I of Fig. 8.13. This curve gives potential as a function of distance with the origin arbitrarily taken at the potential minimum. For comparison there are also shown the potential-distribution curves of Eqs. (8.63) and (8.8) as curves II and III, respectively. It is seen that the actual potential distribution is considerably different from that of the Child-Langmuir law beyond the potential minimum and is totally different from a four-thirds-power law to the left of the potential minimum except in its immediate vicinity.

A study of the approximate relations given above and of the universal potential-distribution curve reveals the following effects of initial velocity

[1] LANGMUIR and COMPTON, *op. cit.*, Part II, p. 241.

upon space-charge flow: The larger the emitted current and the cathode temperature, the greater the magnitude of the potential minimum. The lower the cathode temperature and the larger the plate potential, the closer the potential minimum approaches the cathode and the lower it becomes. The plate current considering initial velocity is larger than that obtained from the Child-Langmuir law because the potential difference between the virtual cathode at the potential minimum and the plate

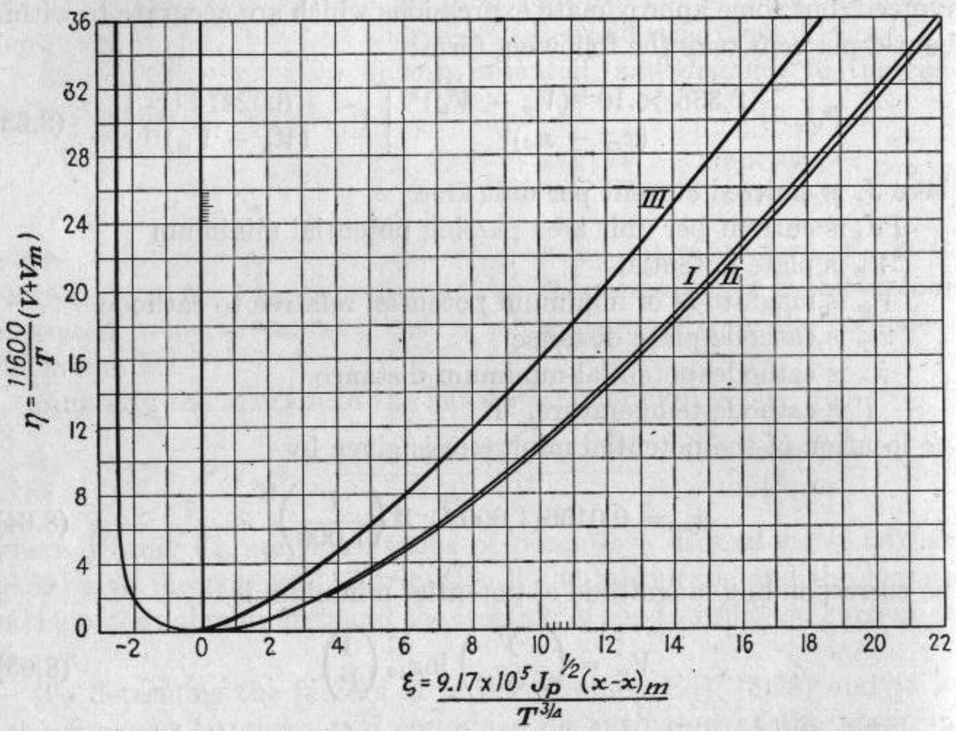


FIG. 8.13.—Universal diode potential-distribution curve including effect of Maxwellian distribution of emission velocity (Langmuir). (See discussion on page 193.)

is greater than the actual plate potential, because the distance from the virtual cathode to the plate is less than the actual cathode-plate distance, and because the electrons leave the virtual cathode with an average velocity that is greater than zero.

The distance of the virtual cathode from the actual cathode may be appreciable. For a cathode temperature of 1000°K and a transmitted current density of 1 ma per cm², the distance from the cathode to the virtual cathode is approximately 0.006 in. In modern close-spaced electrode tubes, this distance is by no means inappreciable.

When the fraction of the emitted current transmitted beyond the virtual cathode is not known, it is necessary to solve Eqs. (8.63) to (8.65)

by trial to determine the transmitted current for a given electrode spacing, plate potential, and cathode temperature.

8.10. Effect of Space Charge upon Transit Time in Diodes. In general, the transit time in an electric field is given by the integral of the reciprocal of velocity with respect to distance.

$$T = \int_{x_1}^{x_2} \frac{dx}{v} \tag{8.66}$$

For the plane-electrode diode the transit time with and without space charge is easily determined. Without space charge the potential profile is a straight line so that

$$V(x) = \frac{x}{d_{cp}} V_p \tag{8.67}$$

where $V(x)$ is the potential at any distance x from the cathode, V_p is the plate potential, and d_{cp} is the cathode-plate distance. The velocity at any point, *assuming zero initial velocity*, is then given by

$$v(x) = \left(\frac{x}{d_{cp}}\right)^{1/2} v_p \tag{8.68}$$

so that the transit time is

$$T = \frac{d_{cp}^{1/2}}{v_p} \int_0^{d_{cp}} \frac{dx}{x^{1/2}} \tag{8.69}$$

with the result that

$$T = \frac{2d_{cp}}{v_p} \tag{8.70}$$

When space charge is present in the plane-electrode diode, then the potential follows a four-thirds-power law so that

$$V(x) = \left(\frac{x}{d_{cp}}\right)^{4/3} V_p \tag{8.71}$$

The velocity at any point is then given by

$$v(x) = \left(\frac{x}{d_{cp}}\right)^{2/3} v_p \tag{8.72}$$

so that the transit time is

$$T = \frac{d_{cp}^{2/3}}{v_p} \int_0^{d_{cp}} x^{-2/3}\, dx \tag{8.73}$$

with the result that

$$T = \frac{3d_{cp}}{v_p} \tag{8.74}$$

This is seen to be of the same form as for the space-charge-free case, the only difference being that the time is 50 per cent greater.

For the cylindrical-diode case in the absence of space charge the potential profile is a logarithmic function.

$$V = \frac{\ln\left(\frac{r}{r_c}\right)}{\ln\left(\frac{r_p}{r_c}\right)} V_p \tag{8.75}$$

where r is the radial distance to any point and r_c and r_p are cathode and plate radius, respectively. The velocity at any point is

$$v = \sqrt{\frac{\ln\left(\frac{r}{r_c}\right)}{\ln\left(\frac{r_p}{r_c}\right)}}\, v_p \tag{8.76}$$

The transit time for this case is

$$T = \frac{\sqrt{\ln\left(\frac{r_p}{r_c}\right)}}{v_p} \int_{r_c}^{r_p} \frac{dr}{\sqrt{\ln\left(\frac{r}{r_c}\right)}} \tag{8.77}$$

If the substitution

$$u = \ln\left(\frac{r}{r_c}\right) \tag{8.78}$$

be made, then the transit time is

$$T = r_c \frac{\sqrt{\ln\left(\frac{r_p}{r_c}\right)}}{v_p} \int_0^{\ln\left(\frac{r_p}{r_c}\right)} \frac{\epsilon^u}{\sqrt{u}}\, du \tag{8.79}$$

This is now in a form which can readily be evaluated by series integration and in which it is apparent that the integral is not infinite. The results may be expressed in the form

$$T = \frac{d}{v_p} B\left(\frac{r_p}{r_c}\right) \tag{8.80}$$

where d is the distance between plate and cathode and $B\left(\frac{r_p}{r_c}\right)$ is a function of the ratio of plate to cathode radius given in Fig. 8.14. It

is seen that when the cathode is inside the transit time is less than that in the plane-electrode case for the same distance and potential.

When there is space charge present in the cylindrical diode, the potential profile is given by

$$V = \left(\frac{r\beta_{rc}^2}{r_p\beta_{cp}^2}\right)^{\frac{3}{4}} V_p \tag{8.81}$$

where r is the radius to any point and β^2 is the function given in Fig. 8.4, the subscripts indicating the distances determining the ratio for which the function is evaluated. The velocity is

$$v = \left(\frac{r\beta_{rc}^2}{r_p\beta_{cp}^2}\right)^{\frac{1}{4}} v_p \tag{8.82}$$

The integral for the transit time is now

$$T = \frac{(r_p\beta_{cp}^2)^{\frac{1}{4}}}{v_p} \int_{r_c}^{r} (r\beta_{rc}^2)^{-\frac{1}{4}}\, dr \tag{8.83}$$

If again the substitution of Eq. (8.78) be made,

$$T = \frac{(r_p\beta_{cp}^2)^{\frac{1}{4}} r_c^{\frac{3}{4}}}{v_p} \int_0^{\ln\left(\frac{r}{r_c}\right)} \left(\frac{\epsilon^u}{\beta_{rc}}\right)^{\frac{3}{4}} du \tag{8.84}$$

This is readily evaluated numerically for small values of $\frac{r}{r_c}$. For large values of $\frac{r}{r_c}$ the form

$$T = \frac{3}{2}\left(\frac{r - r_c}{v}\right) \frac{\frac{d}{du}(\epsilon^u\beta_{rc}^2)}{\epsilon^u - 1} \tag{8.85}$$

is more suitable for computation. The results of the computation can be put into the form[1]

$$T = \frac{d}{v_p} A\left(\frac{r_p}{r_c}\right) \tag{8.86}$$

in which the function $A\left(\frac{r_p}{r_c}\right)$ is that shown in Fig. 8.14. In this figure it is seen that in the cylindrical diode with space charge and an internal cathode the transit time is less than for the corresponding plane electrode but more than for the same case not space-charge-limited. In the

[1] FERRIS, W. R., Input Resistance of Vacuum Tubes as Ultra-high-frequency Amplifiers, *Proc. I.R.E.*, vol. 24, pp. 82–107, January, 1936.

curves of Fig. 8.14 are also included the values for cylindrical diodes with the cathode outside. These are seen to have larger transit times than the plane-electrode diode, which in turn has larger transit times than the cylindrical diode with the cathode inside.

8.11. Summary. The primary effect of space charge in a tube is to make the transmitted current follow a three-halves-power law of plate voltage. In addition, it makes the plate current virtually independent of the filament voltage. Modern tubes are designed so that the emission at rated voltages is more than sufficient to supply the current required

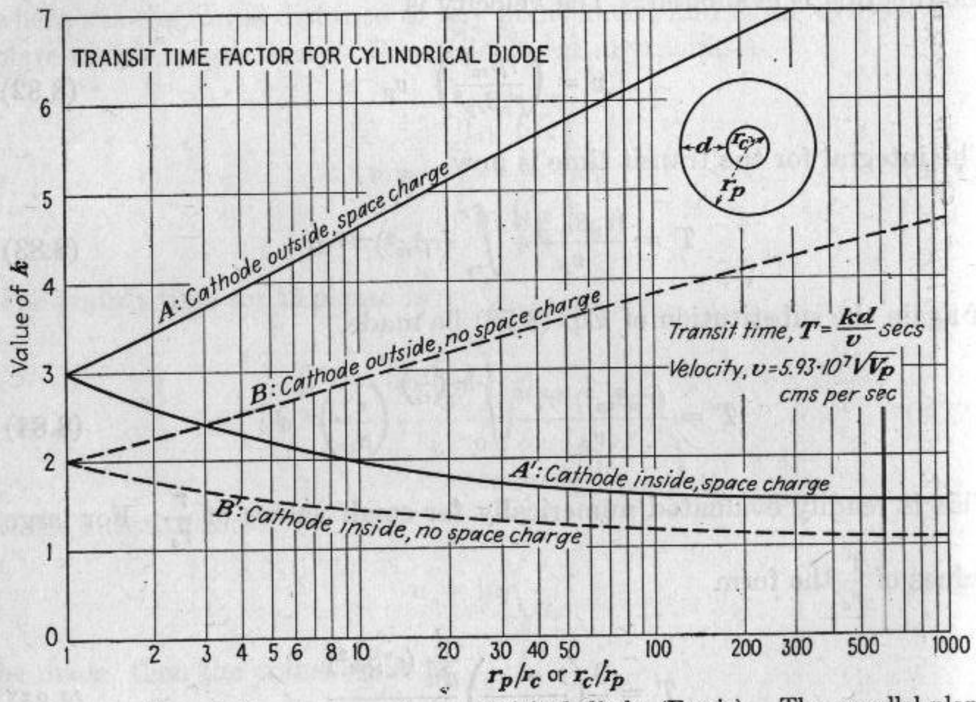


FIG. 8.14.—Transit-time curves of a cylindrical diode (Ferris). The parallel-plane case is given by $r_p/r_c = 1$.

by the Child-Langmuir law. Under these conditions the emission is said to be *space-charge-limited*. The nature of this saturation is shown in Fig. 8.15. Here is shown the variation of the plate current in a diode with cathode heating power for different voltages. If the cathode emission is not very great, a departure from the three-halves-power law of voltage occurs at relatively low voltage. This occurs when the plate is collecting all the current emitted from the cathode and gives rise to what is known as *temperature saturation*. This effect is shown in Fig. 8.16. Here is shown the plate current in a diode as a function of plate voltage for various cathode powers. For low voltages the curve follows the three-halves-power law, and then at some point determined

by the cathode emission the current becomes nearly independent of the plate voltage. The nature of the saturation in this case depends upon the type of emitter. With tungsten as an emitter, the emission is very nearly independent of plate voltage in the saturation range. With thoriated tungsten and even more so with oxide emitters the emission increases slowly with plate voltage in the saturation range. This is because the emission depends upon the gradient at the cathode (Schottky effect). In the case of oxide emitters the increasing gradient through the

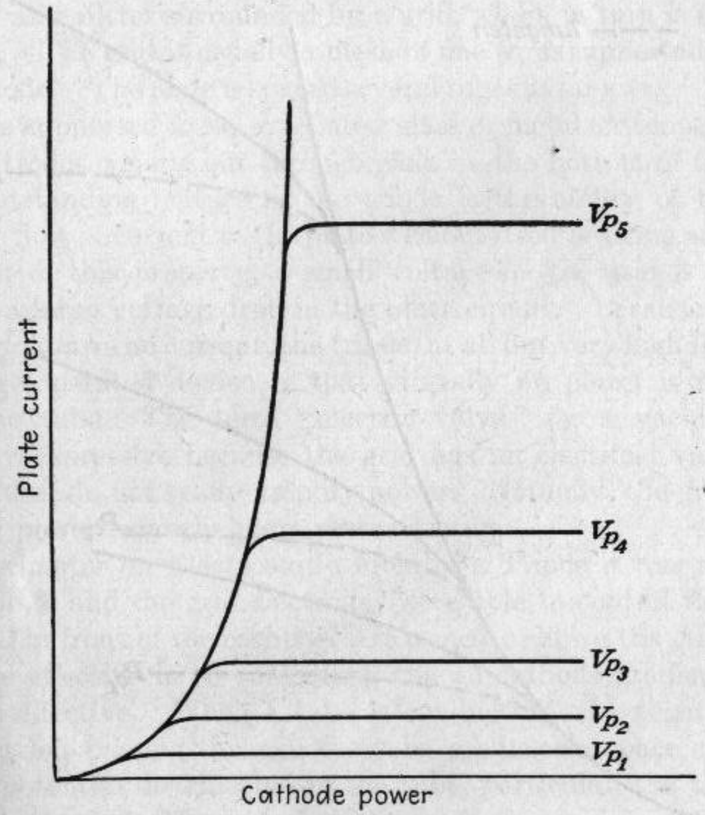


FIG. 8.15.—Diode plate current as a function of cathode power.

emitter increases the liberation of emitting material in addition. In a well-designed tube, temperature-saturation effects will not occur at rated voltages.

In addition to the three-halves-power law of voltage and the saturation effects mentioned above, space charge has the effect of reducing the capacity between electrodes. The capacity between the cathode and plate of a space-charge-saturated diode is three-fifths of that of the cold diode.[1] Further, transit times are in general increased over the diode

[1] LLEWELLYN, F. B., "Electron Inertia Effects," p. 50, Cambridge, London, 1941.

without space charge, being 50 per cent greater in the space-charge-saturated diode with plane electrodes.

Power relations are unchanged. Although the voltage and gradient distribution are different in the presence of space charge and in its absence, the velocity of an electron is always the same relative to the potential, as is required by the energy equation (6.4). The power put

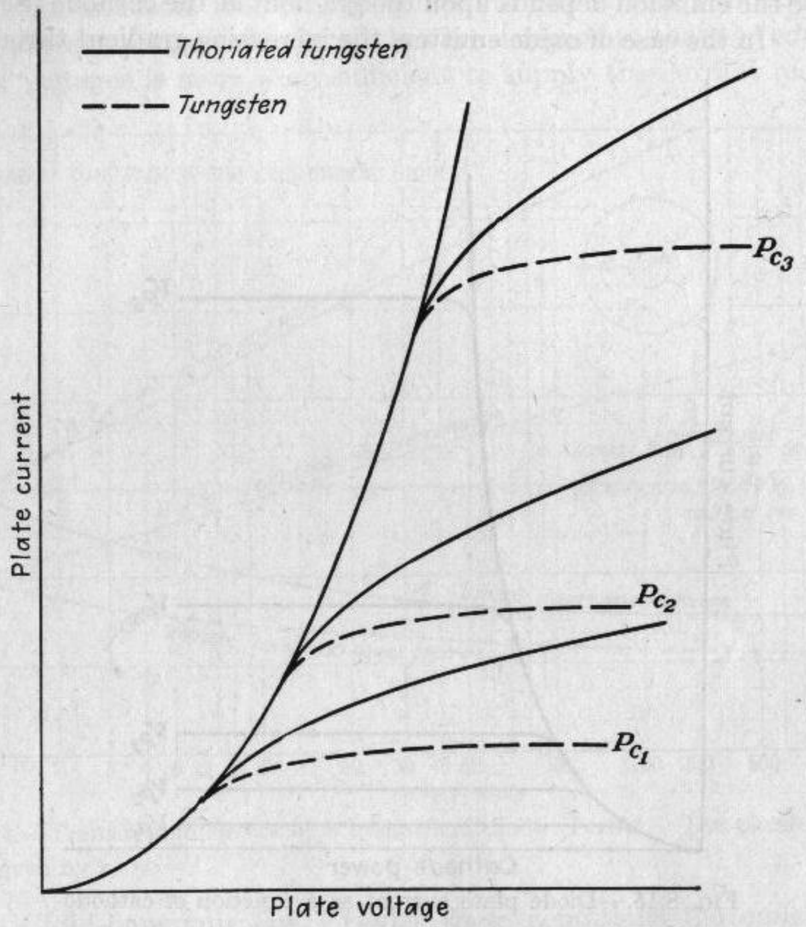


FIG. 8.16.—Diode plate current as a function of plate voltage.

into the tube, as evidenced by the product of the voltage across it and the current through it, appears in the form of heat at the plate. The energy of impact of each electron is $\frac{1}{2}mv^2$, and the number of electrons striking the plate per unit area per second is $\frac{\rho v}{e}$. The product of these is recognized as JV per unit area.

CHAPTER 9

TRIODE CHARACTERISTICS

9.1. Control Action of the Grid. The triode in its commonest form consists of an emitter surrounded by a grid, which in turn is surrounded by a plate. The grid is usually a mesh of fine wires supported quite close to the cathode. The plate is spaced several times as far away. The entire structure is supported in an evacuated glass or metal envelope with leads to the electrodes coming out through glass on the bottom of the tube.

The outstanding feature of the triode is the ability of the grid to control the flow of current to the plate without itself drawing any current. As a result of this property, a small voltage on the grid is capable of producing a large voltage drop in the plate circuit. Because of the fact that the grid draws no current, the triode, at all but very high frequencies, is a voltage-operated device in that virtually no power is required to operate the tube. The term "electric valve" for a vacuum tube is particularly expressive because the grid has an electrical valve action. Vacuum tubes do not really amplify power. Actually, the grid controls the flow of power from the plate power supply.

In the chapter on Electrostatic Field of a Triode it was shown that both the plate and the grid electrodes were able to control the gradient of potential in front of the cathode. It was also shown that the grid was much more effective in so controlling the off-cathode gradient, in fact, μ times as effective. When a tube is conducting, the negative charge of the electrons passing through the tube produces a space charge that alters the potential distribution in the tube, particularly in the vicinity of the cathode, but the control property of the grid is not impaired. The potential distribution between the cathode and grid, for usual combinations of potentials, is now a curve that is concave upward (for plane electrodes) instead of being nearly a straight line. The positive plate potential reaches through the grid and causes the potential on the cathode side of the grid to be positive. Electrons are drawn off the cathode into this region of positive potential and are drawn to the positive plate between the negative-grid wires from whose *immediate* vicinity, however, they are strongly repelled. The negative control grid easily regulates the degree of positiveness of the potential before the cathode. This the grid is able to do primarily because of its greater proximity to the cathode.

9.2. Current-voltage Characteristics of the Triode. The plate current in a triode depends upon both the plate and the grid voltage. It also depends upon the filament voltage, but this is usually held at some suitable fixed value. Hence it is usual to describe the current characteristics of a triode in terms of grid and plate potential alone.

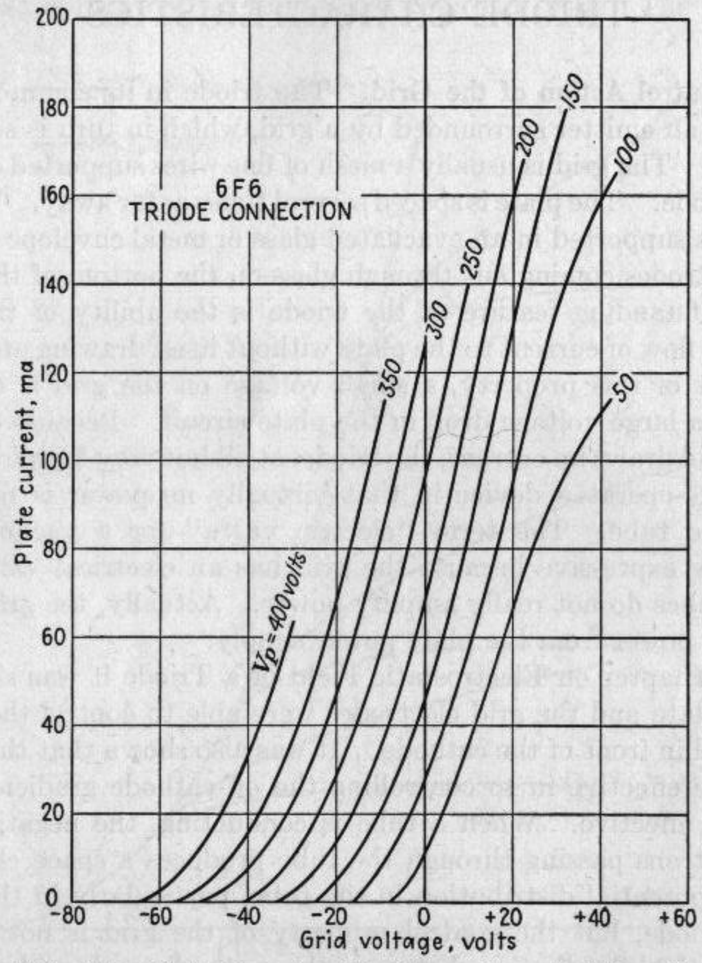


FIG. 9.1.—Plate-current–grid-voltage characteristics of a triode.

Plate-current–Grid-voltage Characteristics. As mentioned in the previous chapter, the space current in a triode is given approximately by

$$I_s = I_g + I_p = G\left(V_g + \frac{V_p}{\mu}\right)^a \tag{9.1}$$

where I_s is space current
I_p is plate current
I_g is grid current

G is perveance
V_g is grid voltage
V_p is plate voltage
μ is amplification factor
a is a constant, approximately $3/2$

When the grid voltage is negative, the entire space current goes to the plate. Even when the grid is positive, the fraction of the space current going to the grid is small so that Eq. (9.1) is a reasonably good approximation for plate current under all conditions except the combination of very small plate voltage and rather large positive grid voltage. An actual

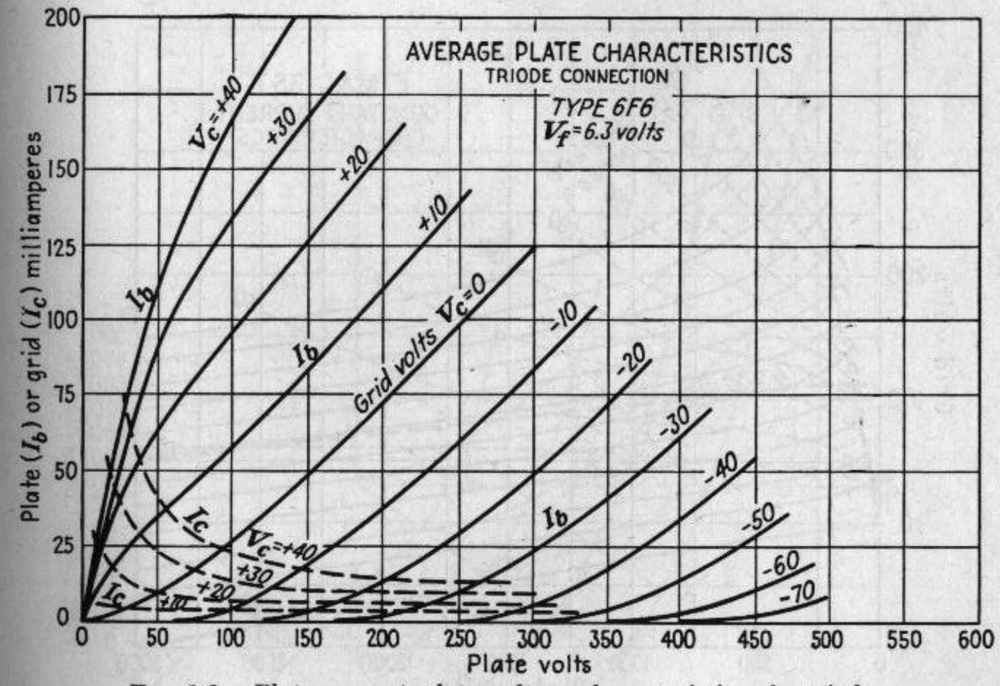


FIG. 9.2.—Plate-current–plate-voltage characteristics of a triode.

plot of plate current as a function of grid voltage is shown in Fig. 9.1. The principal characteristics are quite evident and consistent with expectations. The plate current is seen to increase with both grid and plate voltage but more rapidly with grid voltage. Considering the variation with grid voltage alone, the current increases slowly at first and then more rapidly. In the region of negative grid voltages the curves for the different plate voltages are seen to be of nearly the same shape but merely displaced horizontally. This is consistent with the form of Eq. (9.1). For positive grid voltages the rate of increase of current with grid voltage shows a slight decrease. This is due to two factors. (1) The grid is beginning to draw a fraction of the total, or space, current. (2) There is a tendency for the current to saturate at large plate voltages.

Plate-current–Plate-voltage Characteristics. A typical set of plate-current–plate-voltage characteristics of a triode is shown in Fig. 9.2. The same general properties observed in the plate-current–grid-voltage characteristics are observable here. For negative grid voltages, however, the curves are not similar in this representation. This is because for large negative grid voltages the tube is operating near cutoff and here the amplification factor of the tube is appreciably lower than for grid voltages near zero. The change in shape of the curves for positive grid voltages from concave upward to concave downward is due to the diversion of part of the space current to the grid. The actual space current

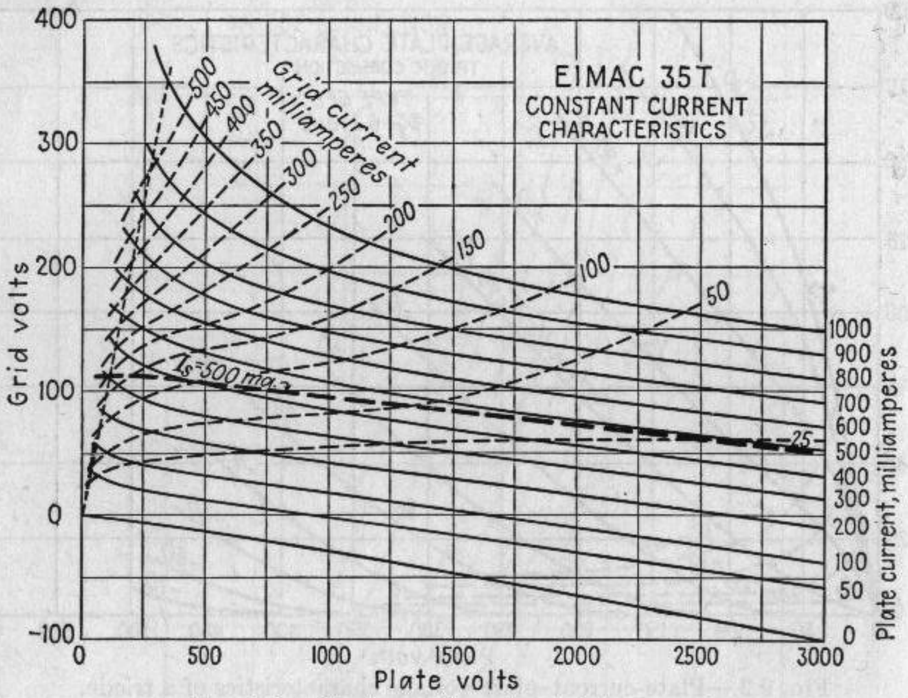


FIG. 9.3.—Contours of constant plate current.

still has an upward curvature with plate voltage, but the fraction of space current taken by the grid decreases as the plate voltage increases.

Contours of Constant Plate Current. In Fig. 9.3 are shown contours of constant plate current plotted along axes of plate voltage and grid voltage. These curves show the combinations of grid and plate potential for which the plate current is constant. Over a large part of their range these contours are parallel straight lines. The significance of this is that the amplification factor of the tube is very nearly constant. For small values of plate voltage and positive grid voltages the curves are curved strongly upward. This is due to the diversion of part of the space current to the grid. If contours of constant space current are plotted

(one is shown dashed in the figure), they are found to be nearly straight lines.

The Plate-current Surface. Inasmuch as plate current is a function of two variables, it may be represented as a surface. The height of this surface above a reference plane is given by the magnitude of the plate current. Position on the reference plane and on the corresponding point on the surface above is given by the plate and grid voltage. A sketch of the plate-current surface is shown in Fig. 9.4. The relation of the surface to the three representations of plate current previously given is evident from the figure. The plate-current–grid-voltage curves

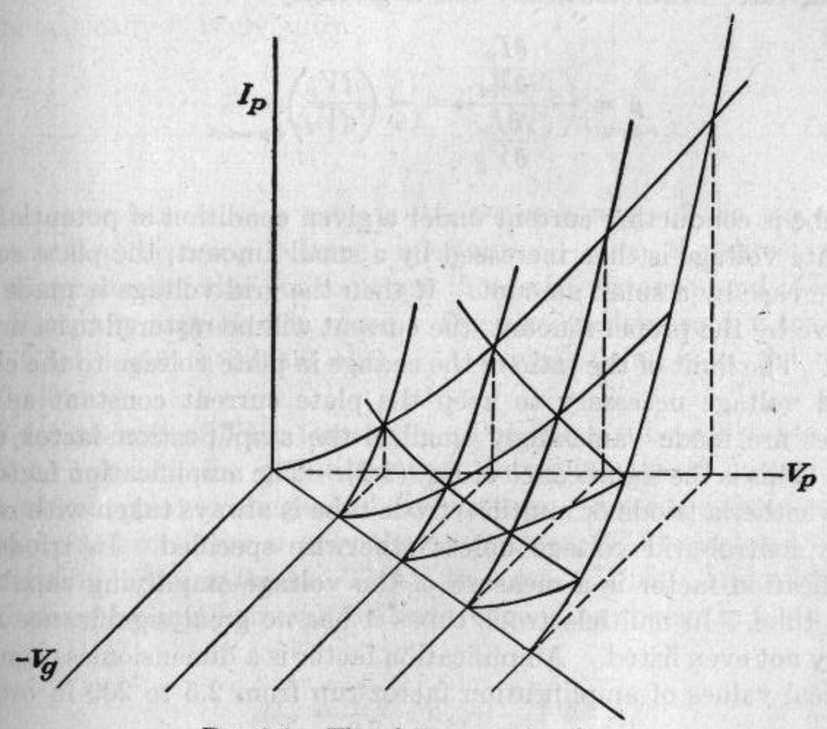


FIG. 9.4.—The plate-current surface.

are the intersections of planes parallel to the plate-current and grid-voltage axes with the surface.

Only a part of the surface is shown, to avoid confusion due to too many lines. The surface becomes a horizontal plateau for large values of current due to voltage saturation, *i.e.*, insufficient emission. Since there can be no plate current for negative plate voltages, the surface turns a corner as it approaches zero values of plate voltage.

9.3. Definition of Triode Constants. *Amplification Factor.* Although the characteristics of a triode are completely specified only by a set of voltage-current curves, an index of the tube's operation is ordinarily

given in terms of the so-called tube "constants." These so-called constants suffice to describe the operation of a tube in the vicinity of a given set of electrode potentials. The three so-called tube constants are the amplification factor, the mutual conductance, and the plate resistance.

The amplification factor is a function of the electrode geometry and has already been given in terms of the dimensions. It has previously been given for a cold tube as the relative effectiveness of the plate and grid potentials in controlling the off-cathode gradient of potential. Another definition and the one generally accepted is that *amplification factor is the relative effectiveness of the plate and grid potentials in controlling the plate current.* Mathematically this is given by

$$\mu = -\frac{\dfrac{\partial I_p}{\partial V_g}}{\dfrac{\partial I_p}{\partial V_p}} = -\left(\frac{dV_p}{dV_g}\right)_{I_p \text{ const}} \tag{9.2}$$

If a tube is conducting current under a given condition of potentials and the plate voltage is then increased by a small amount, the plate current will increase by a small amount. If then the grid voltage is made more negative by the proper amount, the current will be restored to its original value. The limit of the ratio of the change in plate voltage to the change in grid voltage necessary to keep the plate current constant as these changes are made vanishingly small is the amplification factor of the tube. This is the significance of Eq. (9.2). The amplification factor of a tube whether a triode or multielectrode tube is always taken with respect to the control-grid voltage unless otherwise specified. In triodes the amplification factor is a measure of the voltage-amplifying capabilities of the tube. In multielectrode tubes it has no great significance and is usually not even listed. Amplification factor is a dimensionless constant. Practical values of amplification factor run from 2.5 to 200 in ordinary triodes.

Mutual Conductance. The mutual conductance (sometimes called the "transconductance") of a triode has already been referred to in the chapter on Space-charge Effects. *The mutual conductance of a tube is the rate of change of plate current with control-grid voltage.* Mathematically this is given by

$$G_m = \frac{\partial I_p}{\partial V_g} = \left(\frac{dI_p}{dV_g}\right)_{V_p \text{ const}} \tag{9.3}$$

An increase in the grid voltage of a tube effects an increase in the plate current. The limit of the ratio of the changes as the change in grid voltage is made vanishingly small is the mutual conductance. It will

be remembered that the mutual conductance of a triode is directly proportional to the square root of the equivalent voltage of the tube. It is also proportional to the cube root of the space current. As the name implies, the dimensions of this constant are those of conductance. Mutual conductance is usually expressed in units of micromhos or microamperes per volt. Practical values of mutual conductance are between 100 and 10,000 micromhos.

Plate Resistance. Another tube constant that is commonly used is the plate resistance, also known as the "variational plate resistance" or the "dynamic plate resistance." *The plate resistance of a tube is the reciprocal of the rate of change of plate current with plate voltage.* Mathematically it is given by

$$R = \frac{1}{\dfrac{\partial I_p}{\partial V_p}} = \frac{\partial V_p}{\partial I_p} = \left(\frac{dV_p}{dI_p}\right)_{V_g \text{ const}} \tag{9.4}$$

The plate resistance of a tube is the a-c resistance of the plate circuit to a small alternating voltage superimposed upon the direct voltage. The dimensions of this constant are those of resistance, and the magnitude is usually expressed in ohms. The plate resistance of a triode may vary from 1,000 to 50,000 ohms. A typical value is of the order of 5,000 ohms.

Relation between Tube Constants. The three tube constants express relations between the quantities that determine triode operation, *viz.*, plate current, plate voltage, and grid voltage. Since the three constants are expressed in terms of only three variables, it is expected that there is a relation between them. This is the case. If plate and grid voltage are changed by small amounts, the corresponding change in plate current is

$$dI_p = \frac{\partial I_p}{\partial V_g}\,dV_g + \frac{\partial I_p}{\partial V_p}\,dV_p \tag{9.5}$$

or

$$dI_p = G_m\,dV_g + \frac{1}{R_p}\,dV_p \tag{9.6}$$

If the change in plate current is held to zero, then

$$G_m R_p = -\left(\frac{dV_p}{dV_g}\right)_{I_p \text{ const}} = \mu \tag{9.7}$$

This relation, *viz.*, that the product of the mutual conductance and the plate resistance is equal to the amplification factor, holds exactly for any tube for any combination of electrode potentials.

Variation of Tube "Constants." The rather paradoxical heading of this subsection is justified by the fact that the so-called tube constants are not constants at all except approximately so in the vicinity of some operating condition. Actually, the constants may vary considerably over the entire range of voltages and currents in a tube. Of the so-called tube constants, the amplification factor varies the least. This is because it is basically dependent upon the geometrical structure of the tube. If the tube were perfect in that there were no end effects and no asymmetries and if no other electrode but the plate drew current, the amplification factor would not vary at all.

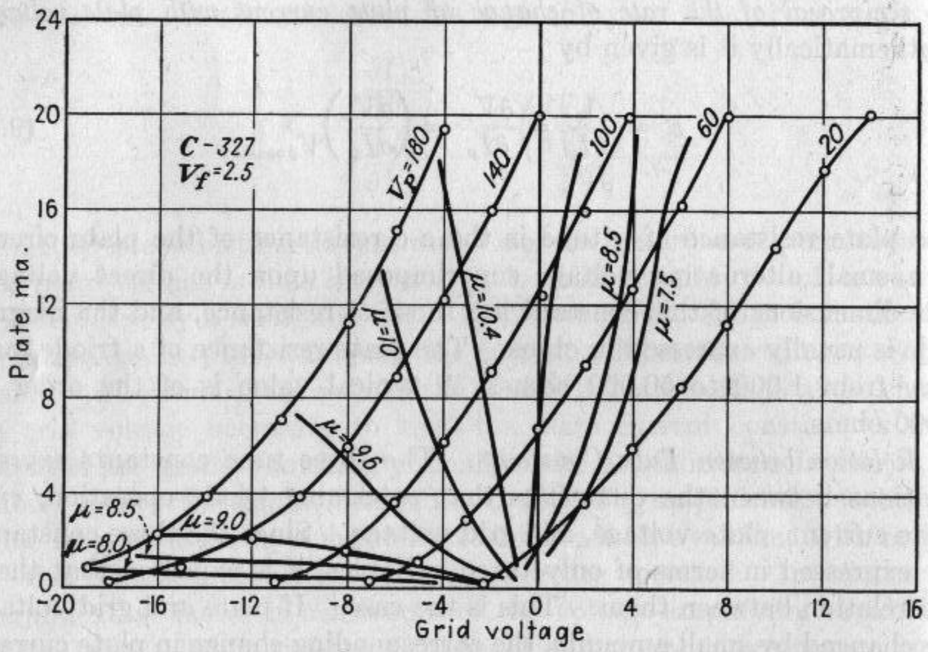


FIG. 9.5.—Contours of constant amplification factor of a triode.

From the definition of amplification factor it is seen that it is given by the negative reciprocal of the slope of the contours of constant plate current in Fig. 9.3. Inspection of the plate-current contours with this in mind reveals that for the most part the amplification factor is fairly constant. It tends to be somewhat low in the vicinity of very low plate currents and even more so in the vicinity of very low plate voltages. A better idea of the nature of the variation of the amplification factor is given in Fig. 9.5, in which there are shown contours of constant amplification factor superimposed upon the plate-current–grid-voltage curves of a triode.[1] The amplification factor is seen to be fairly constant over the entire working range of the tube. Variations in magnitude do not exceed

[1] TERMAN, F. E., and A. L. COOK, Variation in the Amplification Factor of Triodes, *Proc. I.R.E.*, vol. 18, pp. 1044–1047, June, 1930

15 per cent of the mean value. The drop in amplification factor near cutoff is due to end effects in the tube. Because of stray electrostatic fields near the edges of the electrodes, these edges constitute a region of low amplification factor. Thus the actual tube consists of a large electrode area of constant mu in parallel with a small area of much smaller mu. As cutoff is approached, the low-mu portion of the tube cuts off last, giving the effect of a lower amplification factor. The reason why the amplification factor is lower for low plate voltages in the positive-grid region is that here the grid takes a very large portion of the space current. Since, as will be shown later in this chapter, the fraction of the

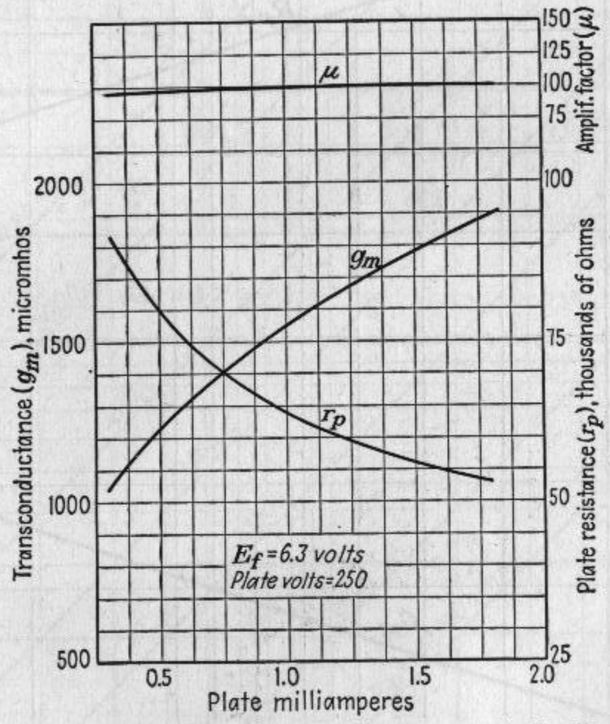


FIG. 9.6.—Variation with plate current of G_m, R_p, and μ of a Type 6F6 as triode.

space current going to the plate increases with plate voltage, a smaller increase in plate voltage relative to a decrease in grid voltage is needed to maintain the plate current constant when the grid voltage is positive than when it is not. This means that the amplification factor is lower for the conditions stated above.

From the definition, mutual conductance is seen to be equal to the slope of the plate-current–grid-voltage curves. Reference to these curves confirms that the mutual conductance is an increasing function of plate current in the region of negative grid voltages.

The mutual conductance of a triode has been shown in the chapter

tion of a number of triodes. The representation is such that the location of the point corresponding to each tube gives the three tube constants. High amplification factors, of the order of 100 or so, may be obtained, but at the expense of a low mutual conductance and a correspondingly high plate resistance.

9.4. Effective Tube Constants of Combinations of Tubes. It is of interest to consider what the effective tube constants are when two or more tubes are connected in parallel. Consider first the case of identical tubes connected in parallel. The effect of this is to double the mutual conductance, halve the plate resistance, and leave the amplification factor unchanged. This is logical, for with two tubes contributing current an increase in grid voltage produces twice as much of an increase in plate current as does a single tube. This explains the doubling of the mutual-conductance value. Since the variations in plate current for a given change in plate voltage are twice as great as for a single tube, the plate resistance is half as great. The amplification factor is unchanged because the product of mutual conductance and plate resistance is the same as for a single tube.

If tubes with different characteristics are connected in parallel, the combination characteristics are still readily determined. The effective mutual conductance is simply the sum of the individual mutual conductances since the plate currents add directly.

$$G_{m\,\text{equiv}} = G_{m1} + G_{m2} + \cdots + G_{mn} \tag{9.8}$$

The equivalent plate resistance is obtained by adding the individual plate resistances as one adds resistances in parallel.

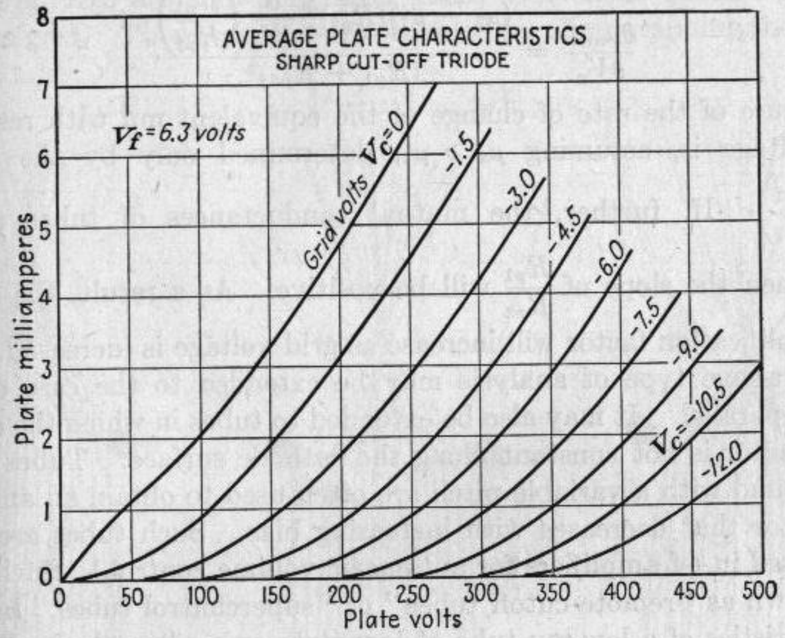
$$\frac{1}{R_{p\,\text{equiv}}} = \frac{1}{R_{p1}} + \frac{1}{R_{p2}} + \cdots + \frac{1}{R_{pn}} \tag{9.9}$$

In other words, the equivalent plate conductance, reciprocal of plate resistance, is the sum of the individual plate conductances.

The equivalent amplification factor is given by the product of the equivalent mutual conductance and the equivalent plate resistance as given by Eqs. (9.8) and (9.9). For the special case of two tubes in parallel the expression for the equivalent amplification factor reduces to

$$\mu_{\text{equiv}} = \frac{\mu_1 R_{p2} + \mu_2 R_{p1}}{R_{p1} + R_{p2}} \tag{9.10}$$

The equivalent amplification factor of tubes in parallel may be higher or lower than one of the individual values but will lie within the extreme values. The equivalent amplification factor will generally decrease as the grid is made more negative. This is because the high-mu tubes



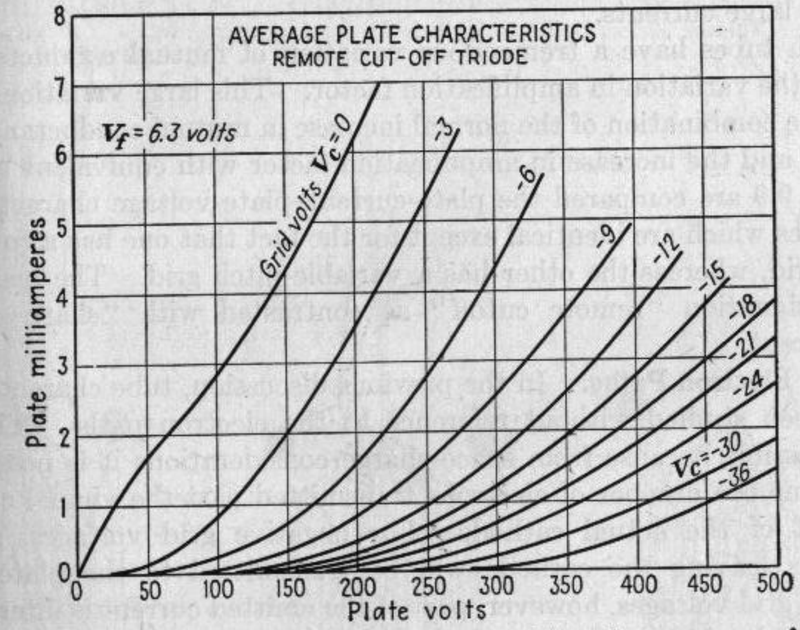


FIG. 9.9.—Comparison of plate-current–plate-voltage curves of the sharp- and remote-cutoff triode of a 6AE6G.

will cut off first, leaving the low-mu tubes to carry current alone. The action may be illustrated in the case of two tubes by studying the rate of change of the equivalent mu with respect to grid voltage. Holding the individual mu's constant in Eq. (9.10),

on Space-charge Effects to increase with the equivalent voltage and with the plate current. The nature of the variation is shown in Fig. 9.6. The variation is quite in accord with expectations as may be shown by plotting the variation of mutual conductance with plate current on logarithmic paper. Such a plot is given in Fig. 9.7 for the same tube. Here it is seen that, since the curve of mutual conductance as a function of plate current is nearly a straight line with a slope of one-third, the

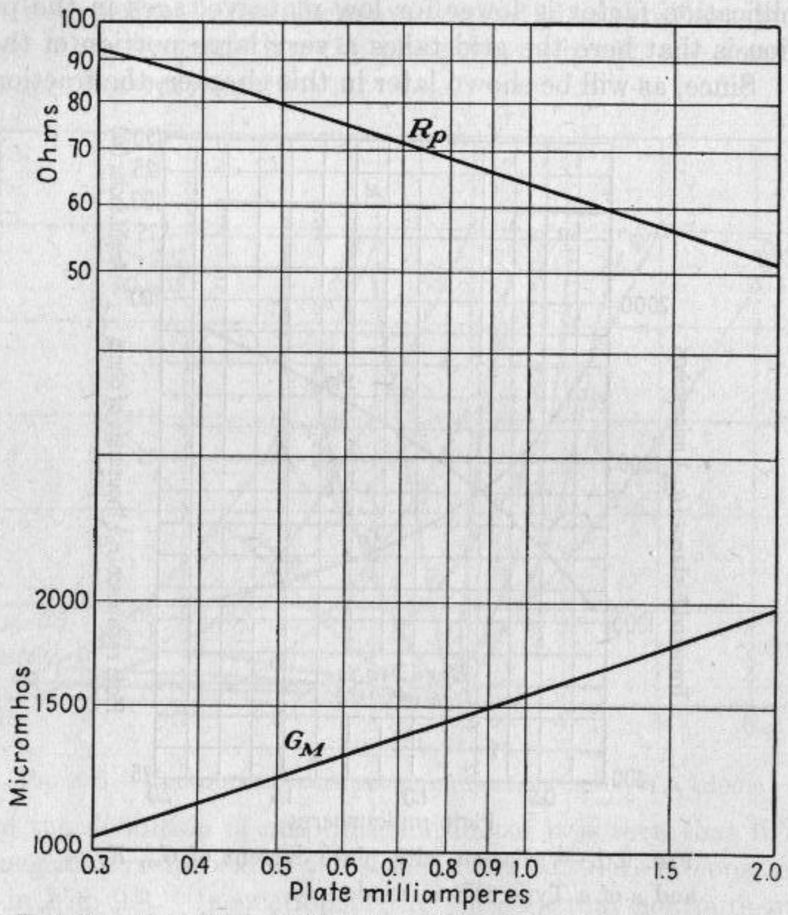


FIG. 9.7.—One-third-power-law variation of mutual conductance and plate resistance with plate current, 6F6 as triode.

mutual conductance follows very closely a one-third-power law of variation with plate current as was predicted in Eq. (8.48). The correspondence between the predicted and actual behavior is best at high currents. At low currents the variation departs somewhat from the one-third-power law because of the reduction in amplification factor.

Since the product of the mutual conductance and the plate resistance is equal to the amplification factor and since the amplification is almost constant, the plate resistance may be expected to vary with plate current

in a fashion reciprocal to that in which the mutual conductance does. This is seen to be the case in Fig. 9.6. Examining the variation more critically, Fig. 9.7 shows that the plate resistance varies nearly as the negative one-third power of plate current. The plate resistance is the reciprocal of the slope of the plate-current–plate-voltage characteristics of a tube. Reference to Fig. 9.2 shows that the plate resistance decreases with increasing plate current in the negative-grid region.

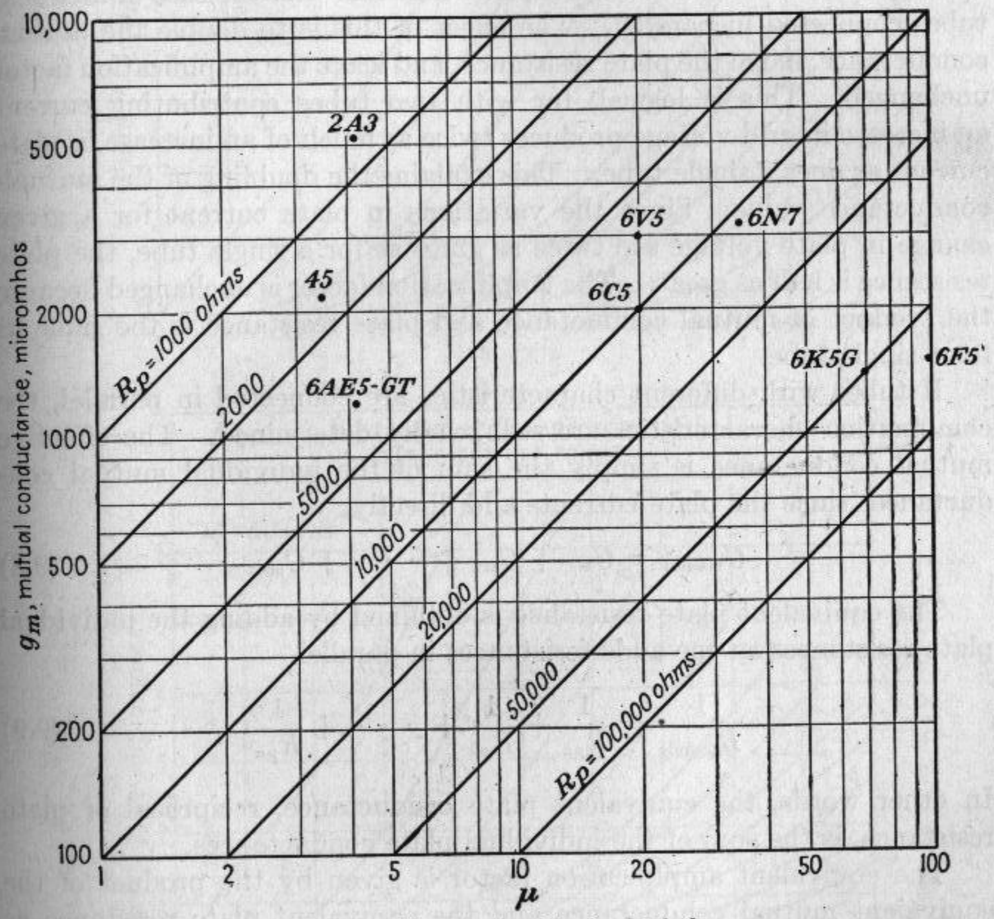


FIG. 9.8.—Constants of typical triodes.

For positive grid voltage the plate resistance may increase again. The plate resistance is lowest at high currents and low plate voltages.

The total range of all the tube constants within a single tube may be considerable. The amplification factor may vary over a range of 20 per cent. The plate resistance and the mutual conductance may vary over a range of three to one. The range of values of the constants encountered from triode to triode is even more considerable. In Fig. 9.8 are shown the tube constants for conditions of recommended opera-

tion of a number of triodes. The representation is such that the location of the point corresponding to each tube gives the three tube constants. High amplification factors, of the order of 100 or so, may be obtained, but at the expense of a low mutual conductance and a correspondingly high plate resistance.

9.4. Effective Tube Constants of Combinations of Tubes. It is of interest to consider what the effective tube constants are when two or more tubes are connected in parallel. Consider first the case of identical tubes connected in parallel. The effect of this is to double the mutual conductance, halve the plate resistance, and leave the amplification factor unchanged. This is logical, for with two tubes contributing current an increase in grid voltage produces twice as much of an increase in plate current as does a single tube. This explains the doubling of the mutual-conductance value. Since the variations in plate current for a given change in plate voltage are twice as great as for a single tube, the plate resistance is half as great. The amplification factor is unchanged because the product of mutual conductance and plate resistance is the same as for a single tube.

If tubes with different characteristics are connected in parallel, the combination characteristics are still readily determined. The effective mutual conductance is simply the sum of the individual mutual conductances since the plate currents add directly.

$$G_{m\,\text{equiv}} = G_{m1} + G_{m2} + \cdots + G_{mn} \tag{9.8}$$

The equivalent plate resistance is obtained by adding the individual plate resistances as one adds resistances in parallel.

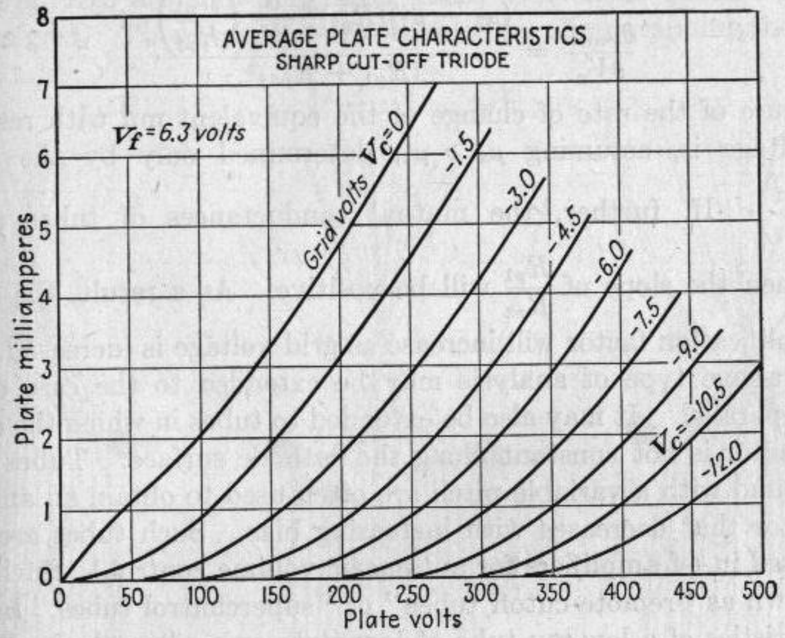
$$\frac{1}{R_{p\,\text{equiv}}} = \frac{1}{R_{p1}} + \frac{1}{R_{p2}} + \cdots + \frac{1}{R_{pn}} \tag{9.9}$$

In other words, the equivalent plate conductance, reciprocal of plate resistance, is the sum of the individual plate conductances.

The equivalent amplification factor is given by the product of the equivalent mutual conductance and the equivalent plate resistance as given by Eqs. (9.8) and (9.9). For the special case of two tubes in parallel the expression for the equivalent amplification factor reduces to

$$\mu_{\text{equiv}} = \frac{\mu_1 R_{p2} + \mu_2 R_{p1}}{R_{p1} + R_{p2}} \tag{9.10}$$

The equivalent amplification factor of tubes in parallel may be higher or lower than one of the individual values but will lie within the extreme values. The equivalent amplification factor will generally decrease as the grid is made more negative. This is because the high-mu tubes



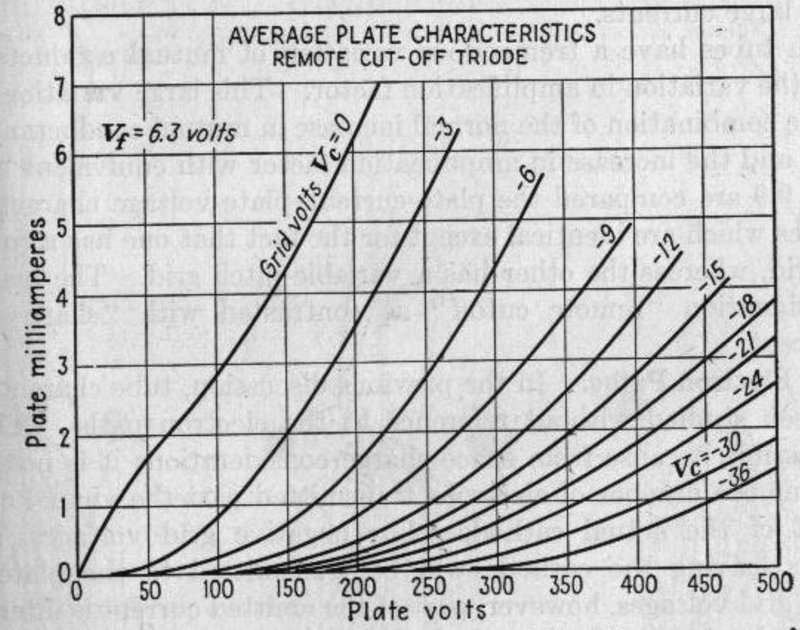


FIG. 9.9.—Comparison of plate-current–plate-voltage curves of the sharp- and remote-cutoff triode of a 6AE6G.

will cut off first, leaving the low-mu tubes to carry current alone. The action may be illustrated in the case of two tubes by studying the rate of change of the equivalent mu with respect to grid voltage. Holding the individual mu's constant in Eq. (9.10),

$$\frac{\partial \mu_{\text{equiv}}}{\partial V_g} = \frac{(\mu_2 - \mu_1)(R_{p2})^2 \dfrac{d}{dV_g}\left(\dfrac{R_{p1}}{R_{p2}}\right)}{(R_{p1} + R_{p2})^2} \tag{9.11}$$

The nature of the rate of change of the equivalent mu with respect to grid voltage is, assuming $\mu_2 > \mu_1$, determined only by the sign of $\dfrac{d}{dV_g}\left(\dfrac{R_{p1}}{R_{p2}}\right)$. If, further, the mutual conductances of tubes are the same, then the slope of $\dfrac{R_{p1}}{R_{p2}}$ will be positive. As a result, the equivalent amplification factor will increase as grid voltage is increased.

The above type of analysis may be extended to the case of more tubes in parallel. It may also be extended to tubes in which the amplification factor is not constant along the cathode surface. Tubes having grids wound with a variable pitch are often used to obtain an amplification factor that decreases with increasing bias. Such tubes are extensively used in r-f amplifiers for automatic volume control.[1] Such tubes, also known as "remote-cutoff tubes" or "supercontrol tubes," have the characteristics of a low-mu tube at low plate currents and of a high-mu tube at large currents.

Such tubes have a tremendous variation of mutual conductance as well as the variation in amplification factor. This large variation results from the combination of the normal increase in mutual conductance with current and the increase in amplification factor with equivalent voltage. In Fig. 9.9 are compared the plate-current–plate-voltage characteristics of triodes which are identical except for the fact that one has a constant-pitch grid, whereas the other has a variable-pitch grid. The reason for the designation "remote cutoff" as contrasted with "sharp cutoff" is apparent.

9.5. Electron Paths. In the previous discussion, tube characteristics have been studied without reference to the electron paths. This has been possible because from space-charge considerations it is possible to determine the number of electrons transmitted past the virtual cathode in front of the actual cathode. For negative grid voltages, all the electrons leaving the cathode will be transmitted to the plate. For positive grid voltages, however, part of the emitted current is intercepted by the grid, and here the actual electron paths are of interest. Electron paths are also of interest in multielectrode tubes, where they have a considerable part in determining the tube characteristics.[2]

[1] BALLANTINE, STUART, and H. A. SNOW, Reduction of Distortion and Cross-talk in Radio Receivers by Means of Variable-mu Tetrodes, *Proc. I.R.E.*, vol. 18, pp. 2102–2127, December, 1930.

[2] THOMPSON, A. C., Electron Beams and Their Application in Low Voltage Devices, *Proc. I.R.E.*, vol. 24, pp. 1276–1297, October, 1936.

In Fig. 9.10 are shown electron paths in a triode operating with a negative grid. These curves were obtained by photographing the motion

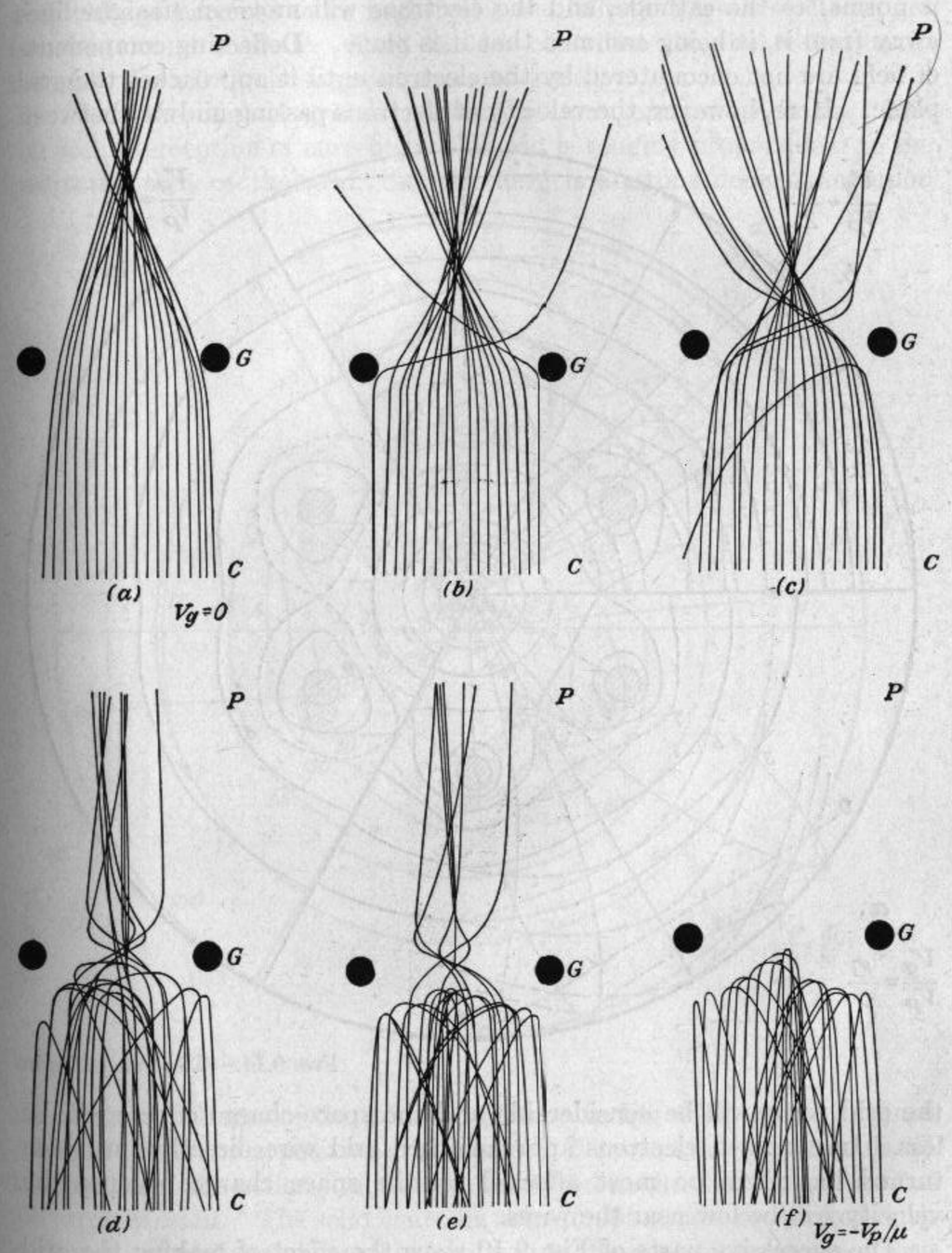


FIG. 9.10.—Electron paths in a negative-grid triode (Kleynen).

of small balls rolled upon a suitably deformed elastic membrane. Such a model of potential takes no account of space-charge effects. It may be expected that in this case the presence or absence of space charge will

make no great difference in the electron paths. The space charge is most pronounced close to the cathode. Here the gradient of potential is normal to the cathode, and the electrons will move in straight lines away from it, it being assumed that it is plane. Deflecting components of field are not encountered by the electron until it approaches the grid plane. Here, however, the velocity of electrons passing midway between

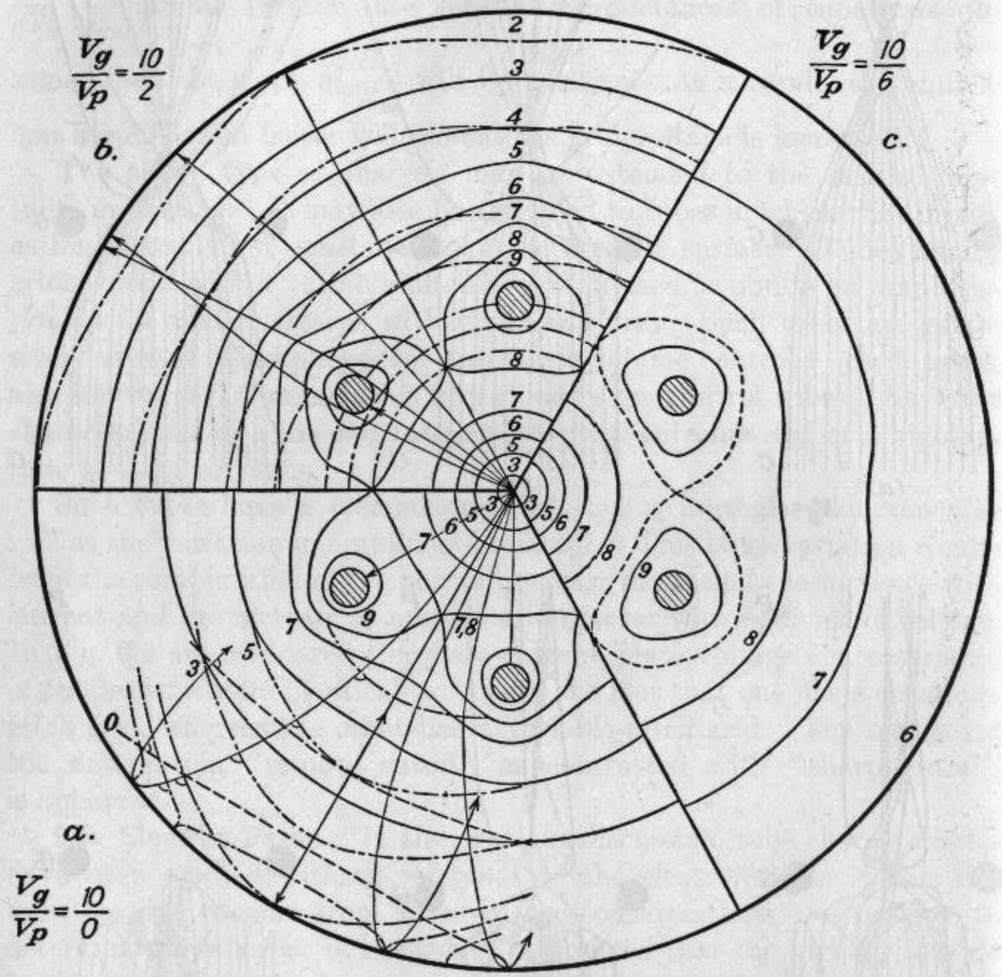


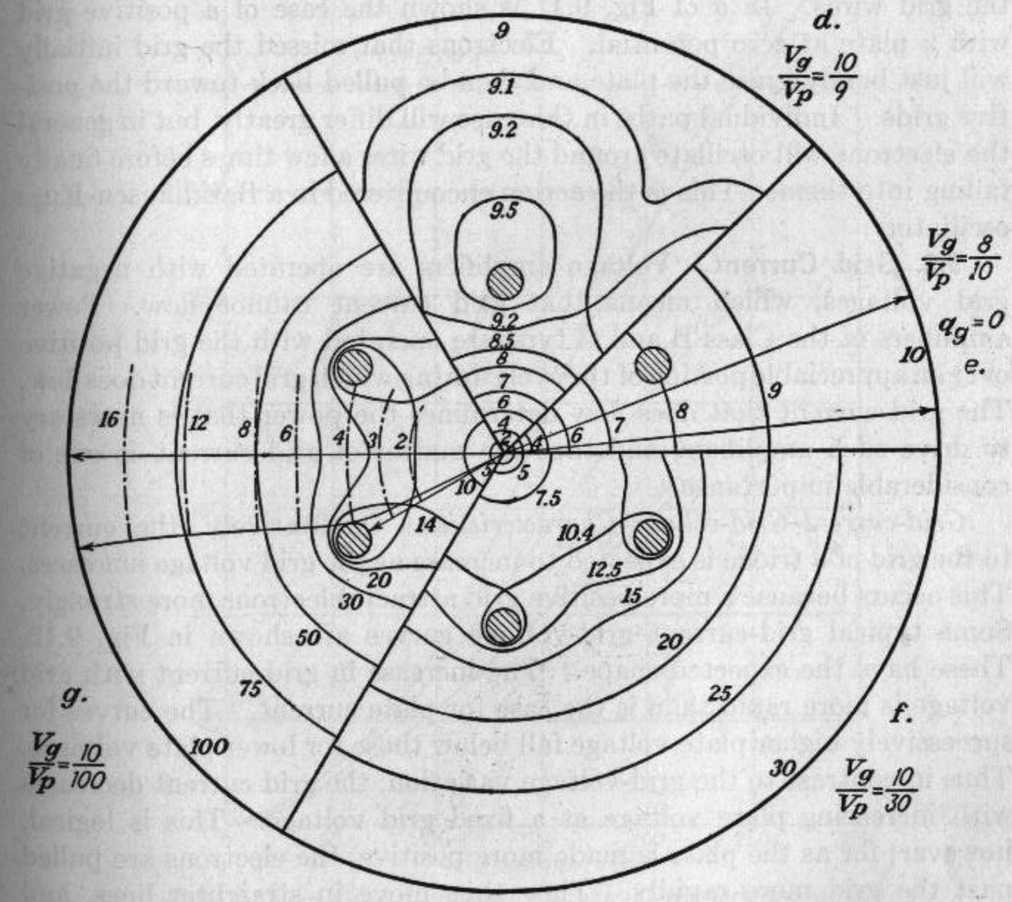
FIG. 9.11.—Electron paths in

the grid wires will be considerable and the space-charge effects will be less. In contrast, electrons approaching a grid wire directly and there turned back will be most affected by the space charge because the velocity will be low near the wires.

The successive parts of Fig. 9.10 show the effect of making the grid more and more negative until cutoff is reached. The sidewise deflecting forces become greater as the grid is made more negative until some of the electrons are turned back. Up to that condition the electrons are passed

through the grid wires in a bunch that is focused beyond the grid plane. Coupled with the condition of more electrons being turned back as the grid voltage becomes more negative is the fact that fewer electrons get by the virtual cathode. Both factors contribute to the reduction in current, though the latter predominates greatly.

When the control grid is positive, it may attract electrons. The normal interception of current by the grid is roughly proportional to the projected area of the grid, though there is a strong dependence upon



a positive-grid triode (Lange).

the relative voltages of the grid and plate. In Fig. 9.11 are shown some electron paths for a positive-grid triode.[1] These paths were calculated by the use of the action function, as described in the chapter on Laws of Electron Motion. The solid contours are equipotentials, the broken-dash contours are surfaces of constant action, and the electron paths are drawn

[1] LANGE, H., Current Division in Triodes and Its Significance in the Determination of Contact Potential, *Zeit. Hochfrequenz*, vol. 31, pp. 105–109, 133–140, 191–196, 1928.

perpendicular to these. When the grid is positive but less than its "natural potential," *i.e.*, that which would give the diode potential distribution in the tube, the action of the potential field causes these electrons, which initially miss the grids, to converge beyond, as was the case with the negative-grid triode. This situation is shown in part *g* of Fig. 9.11. When the grid is more positive than its natural potential and also more positive than the plate, as in *a* and *b* of Fig. 9.11, the potential field has a divergent action and the electrons are pulled into the grid wires. In *a* of Fig. 9.11 is shown the case of a positive grid with a plate at zero potential. Electrons that missed the grid initially will just barely graze the plate and then be pulled back toward the positive grids. Individual paths in this case will differ greatly, but in general the electrons will oscillate around the grid wires a few times before finally falling into them. This is the action encountered in a Barkhausen-Kurz oscillator.

9.6. Grid Current. Voltage amplifiers are operated with negative grid voltages, which means that grid current cannot flow. Power amplifiers of the Class B and C type are operated with the grid positive over an appreciable portion of the cycle during which grid current does flow. The grid current that does flow determines the power that is necessary to drive such amplifiers, and thus the matter of grid current is one of considerable importance.

Grid-current–Grid-voltage Characteristics. Qualitatively, the current to the grid of a triode is expected to increase as the grid voltage increases. This occurs because a more positive grid attracts electrons more strongly. Some typical grid-current–grid-voltage curves are shown in Fig. 9.12. These have the expected shape. The increase in grid current with grid voltage is more rapid than is the case for plate current. The curves for successively higher plate voltage fall below those for lower plate voltages. Thus in contrast to the grid-voltage variation, the grid current decreases with increasing plate voltage at a fixed grid voltage. This is logical, however; for as the plate is made more positive, the electrons are pulled past the grid more rapidly. They thus move in straighter lines, and therefore fewer of them are pulled into the grid. The current characteristics in the presence of secondary emission may be greatly different and will be treated separately later.

Grid-current–Plate-voltage Characteristics. Some typical grid-current–plate-voltage curves are shown in Fig. 9.13. For positive plate voltage the primary grid current decreases with increasing plate voltages as just noted. Curves for high positive grid voltage are shown above those for lower grid voltage. Grid current may flow when the plate voltage is negative, though such an operating condition is rarely encountered

in practice. For this condition all the emitted current is taken by the grid. The grid current drops slightly as the negative plate voltage is made more negative. This is because the space current itself is reduced owing to the reduction in the equivalent voltage in the triode.

Constant-grid-current Contours. Contours of constant grid current are shown along with contours of constant plate current in Fig. 9.3.

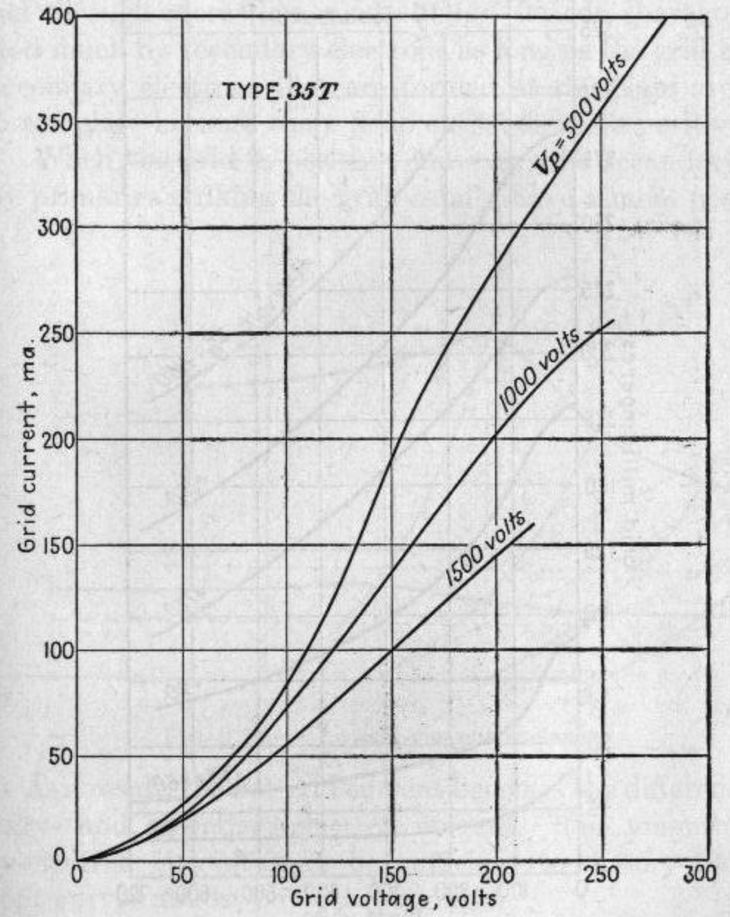


FIG. 9.12.—Grid-current–grid-voltage characteristics of a type 35T triode.

Grid current flows only when the grid is positive. The positive-grid–negative-plate quadrant is not shown because it is of little practical value. In the absence of secondary emission the contours present an orderly appearance. The contours follow no simple law as do the plate-current contours in the negative-grid region. The increase in grid current with grid voltage is much more rapid than the decrease with plate voltage.

The Grid-current Surface. Just as it was possible to draw a surface for the plate current as a function of grid and plate voltage, so is it possible to draw one for grid current. A sketch of such a surface is shown in Fig. 9.14. The previous representations of grid current will be recognized as part of this picture. The grid-current–grid-voltage

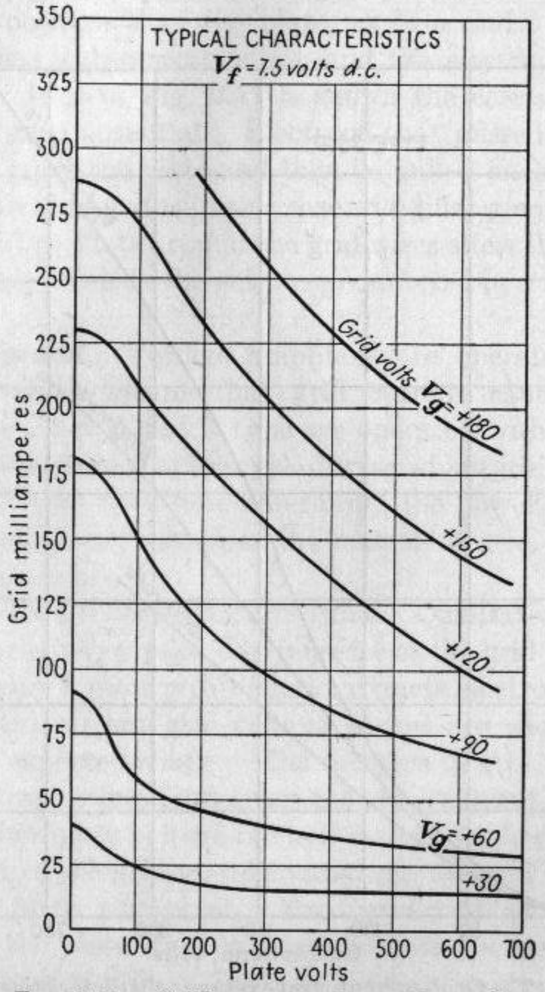


FIG. 9.13.—Grid-current–plate-voltage characteristics of a type 826 triode.

curves are the intersections of the grid-current surface with a plane parallel to the grid-current and grid-voltage axes. The grid-current–plate-voltage curves are intersections of the surface with a plane parallel to the grid-current and plate-voltage axes. The constant-grid-current contours are intersections of the grid-current surface with planes parallel to the grid-voltage and plate-voltage axes.

It is possible to define constants to describe the grid-current action,

but this has no great value. It is of interest to note a few relations, however. The grid-current–plate-voltage transconductance is negative. The equivalent amplification factor for the inverted triode, *i.e.*, one whose grid is positive and whose plate is negative, is the reciprocal of the normal amplification factor of the tube.

Effect of Secondary Electrons. Secondary electrons are created whenever an electrode is struck with primary electrons that have been accelerated through more than a few volts. Triode characteristics are not affected much by secondary electrons as long as the grid is negative, for the secondary electrons that are formed at the plate are attracted back into the plate because there is no electrode more positive for them to go to. When the grid is positive, however, the secondary electrons formed by primaries striking the grid usually have a more positive plate

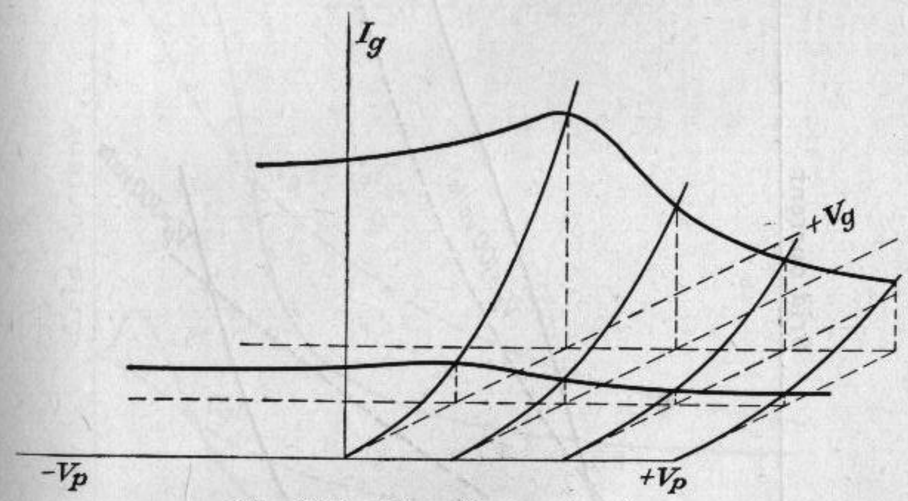


FIG. 9.14.—The grid-current surface.

to go to. As a result, the net grid current becomes the difference between the primary- and secondary-electron current. The magnitude of the secondary-electron current may be sufficient to distort the primary-grid-current curves almost beyond recognition.

When both grid and plate potentials are positive, secondary electrons are formed by primaries striking both. When the plate is more positive than the grid, the secondary electrons from the grid will be attracted to the plate but those formed at the plate will be attracted back into the plate. When the grid is more positive than the plate, the situation is reversed and secondaries from the plate will be attracted to the grid but those created at the grid will be attracted back into the grid itself. The result of this action upon the grid-current–grid-voltage characteristics is shown in Fig. 9.15. In this figure is shown a typical grid-current–grid-voltage curve in the presence of secondary emission

compared with the primary-grid-current curve. For small grid voltages very few secondaries are created, and hence the currents with and without secondary emission are almost equal. As the grid voltage is increased, more secondaries are created and attracted to the plate. The grid current is therefore reduced by the amount of the secondary current to the plate. The grid current may be reduced enough to become negative. As the grid becomes more positive, more secondaries are likely to be

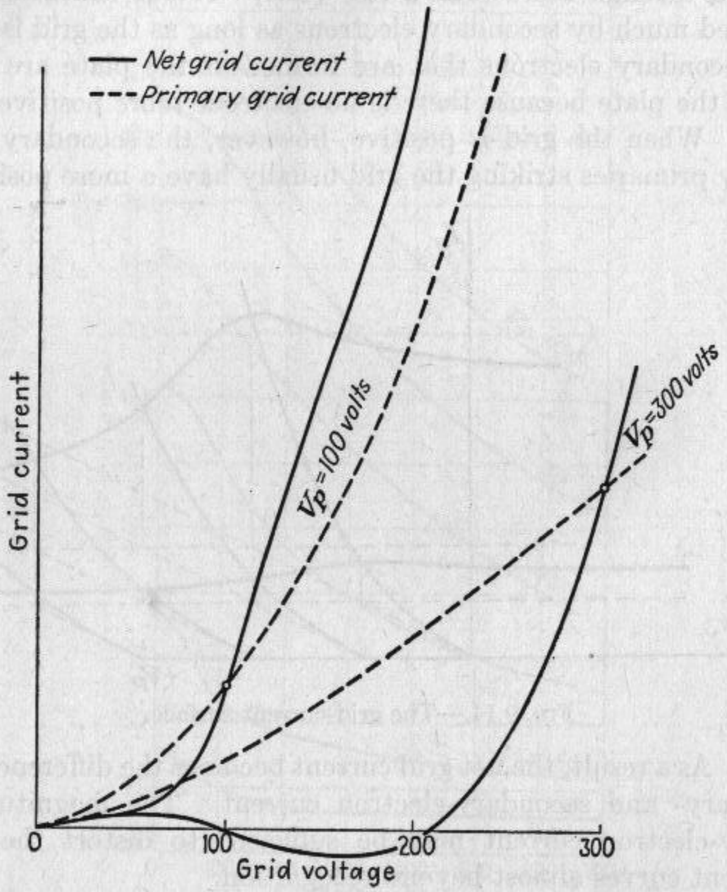


FIG. 9.15.—Influence of secondary emission upon the grid-current–grid-voltage characteristics of a triode.

created but the gradient of potential driving them to the plate becomes smaller, until finally it becomes negative when the grid potential exceeds the plate potential. As this occurs, the primary grid current exceeds the net grid current by less and less until when the grid potential and plate potential are equal the net grid current is nearly equal to the primary current. As the grid voltage is increased still further, the number of secondary electrons created at the grid surface becomes still greater but these electrons are confronted by a negative gradient of potential on all

sides and so are attracted back into the grid. Now, however, secondary electrons liberated from the plate are confronted by a positive gradient of potential that attracts them to the grid. The grid current is now greater than the primary grid current.

An action similar to that described above shows itself on the grid-current–plate-voltage curves. When the plate is less positive than the grid, secondary electrons from the plate are attracted to the grid and hence the actual grid current is greater than the primary grid current.

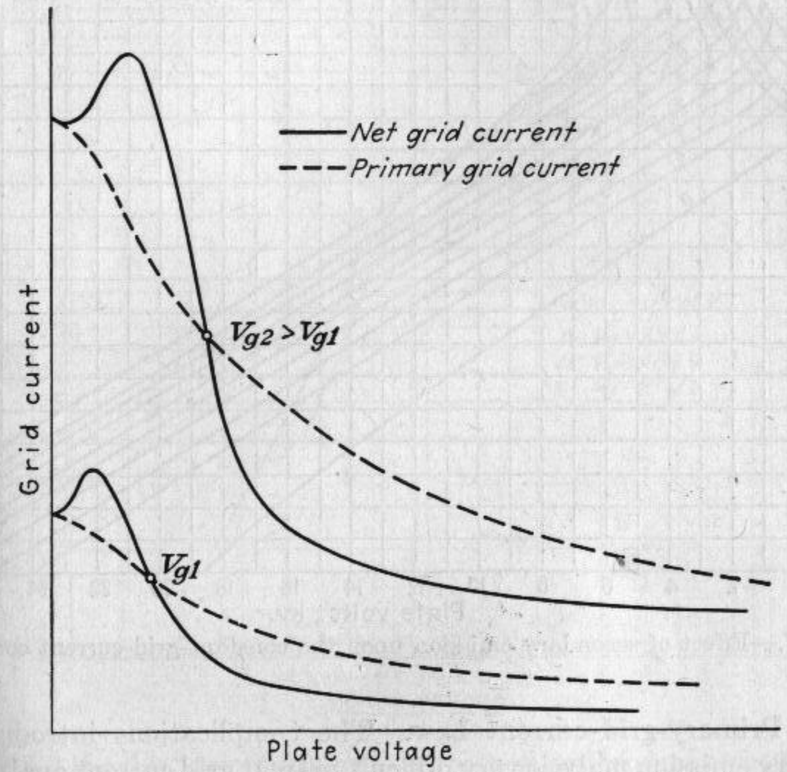


FIG. 9.16.—Influence of secondary emission upon the grid-current–plate-voltage characteristics of a triode.

When the plate voltage is more positive than the grid voltage, secondary electrons from the grid are attracted to the plate and the grid current is less than the primary value. This action is shown in Fig. 9.16. Points of equal grid and plate voltage are crossover points of net and primary grid current. These points are marked by circles.

The effect of secondary emission upon the contours of constant grid current may also be considerable, especially if the secondary emission is great enough to make the grid current negative. In Fig. 9.17 are shown some constant-grid-current contours of a water-cooled tube with a high degree of secondary emission. The effect of secondary emission is to

raise all the large positive-current contours. It is as though a wedge of negative-current contours had been driven under the positive-current contours from the right. Contours of constant plate current are also distorted by secondary emission, though to a lesser degree.

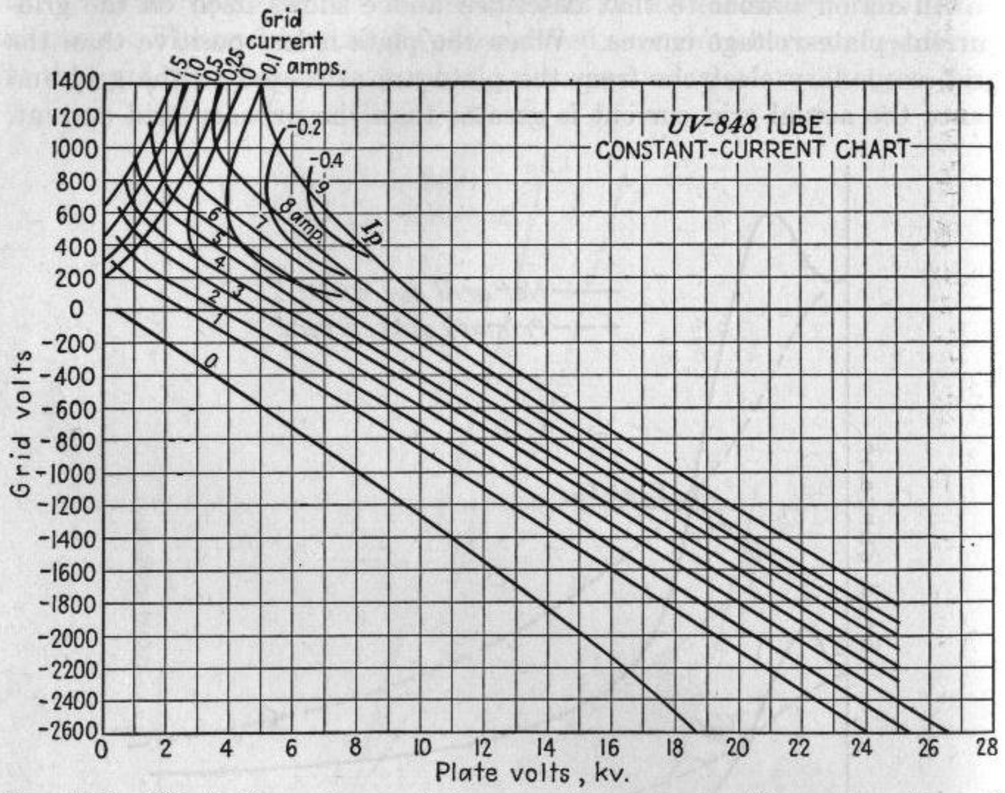


FIG. 9.17.—Effect of secondary emission upon the constant-grid-current contours of a triode.

9.7. Primary-grid-current Law. The complications introduced by secondary emission make it very difficult to treat grid current analytically. A considerable impression can, however, be made upon the subject of primary grid current. The analytical treatment of primary grid current is simplified by the observation made in Sec. 6.5 that, for a given *ratio* of plate to grid voltage, the electron paths within the tube are not altered by a change in the magnitude of these voltages. Since the electron paths are not changed, the division of current between the plate and grid is not changed and hence *the ratio of plate to grid current should be a function of the ratio of plate voltage to grid voltage alone* and be independent of the magnitude of these voltages. Were it not for secondary emission and some other effects such as the change in the position of the virtual cathode, this would be exactly true. Actually, the correspondence with expectations is quite good, as is shown in Fig. 9.18, in which there are

plotted curves of the ratio of currents as a function of the ratio of voltages. The characteristics in this figure are for a small high-mu transmitting triode with tantalum electrodes. Such a tube is relatively free of secondary-emission effects. It is seen that the curves for different potentials superimpose reasonably well. If more curves were given, they would form a bundle within the limits of the curves shown.

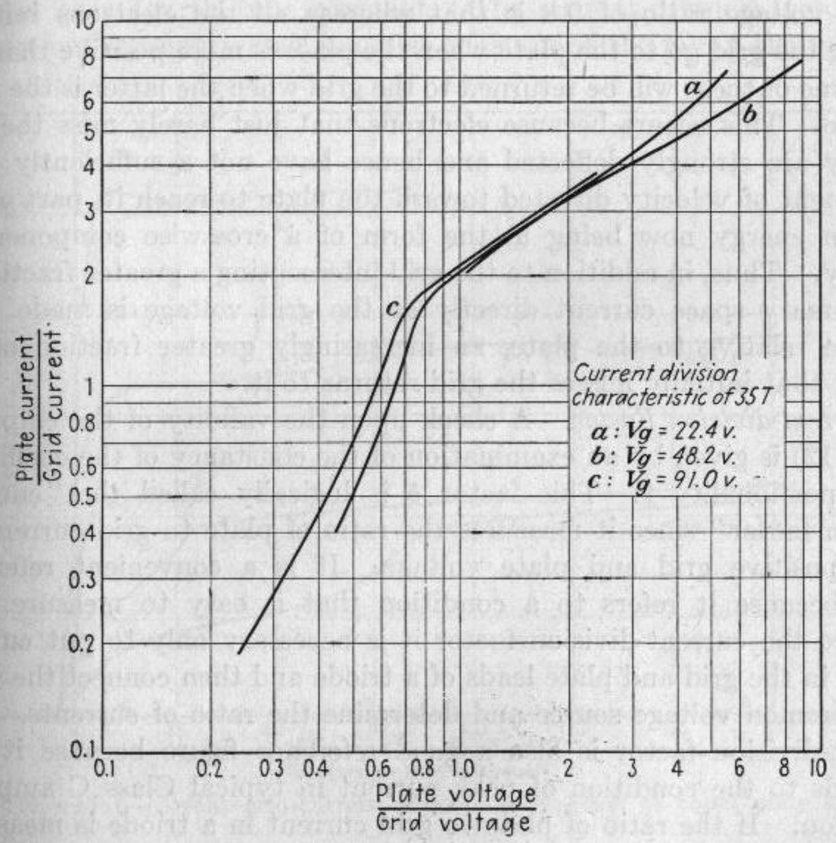


FIG. 9.18.—Electrode-current ratio in a positive-grid triode as a function of electrode-voltage ratio.

Examination of many tubes shows their primary-current-division characteristics to have the general form shown in Fig. 9.18. On such a log-log plot the curves are nearly straight lines with a slope of $\frac{1}{2}$ above a voltage ratio of 0.8 and with a slope of 2 below a voltage ratio of 0.8. Accordingly the primary current division may be expressed by

$$\frac{I_p}{I_g} = \delta \sqrt{\frac{V_p}{V_g}} \qquad \text{for } \frac{V_p}{V_g} > 0.8 \quad (9.12)$$

and

$$\frac{I_p}{I_g} = 1.392\delta \left(\frac{V_p}{V_g}\right)^2 \qquad \text{for } \frac{V_p}{V_g} < 0.8 \quad (9.13)$$

where δ is a constant known as the current-division factor[1–3] and defined as the ratio of plate to grid current for equal positive grid and plate voltages.

Since triodes are seldom operated with the plate less positive than the grid the form of Eq. (9.12) will be of more concern than that of Eq. (9.13). The reason for the change of slope, and hence of exponent, at the voltage ratio of 0.8 is that whereas all the electrons initially missing the grid go to the plate when the plate is more positive than the grid some of these will be returned to the grid when the latter is the more positive. This occurs because electrons that just barely miss the grid initially are strongly deflected and hence have not a sufficiently large component of velocity directed toward the plate to reach it, part of the electron energy now being in the form of a crosswise component of velocity. Thus, in addition to the grid intercepting a greater fraction of the primary space current directly as the grid voltage is made more positive relative to the plate, an increasingly greater fraction of the current that initially misses the grid returns to it.

Current-division Factor. A check upon the validity of the empirical Eq. (9.12) is given by an examination of the constancy of the coefficient of proportionality δ. This factor δ is logically called the "current-division factor" since it measures the ratio of plate to grid current for equal positive grid and plate voltage. It is a convenient reference point because it refers to a condition that is easy to measure. To measure the current-division factor it is necessary only to put current meters in the grid and plate leads of a triode and then connect the leads to a common voltage source and determine the ratio of currents. The current-division factor is also a good reference figure because it corresponds to the condition of peak current in typical Class C amplifier operation. If the ratio of plate to grid current in a triode is measured as a function of equal positive plate and grid voltages, variations of the sort shown in Fig. 9.19 result. For all the triodes shown, the current ratio rises sharply with voltage and then assumes a nearly constant value. The change in the current ratio with low voltages is caused primarily by the change in the position of the virtual cathode in front of the actual cathode. At low voltages the virtual cathode is located a considerable distance out from the actual cathode. As will be shown later, a small

[1] Tank, F., Zur Kentniss der Vorgänge in Elektrodenröhren, Jahr. draht. Tel. u. Tel. vol. 20, p. 80, 1922.

[2] See also Lange, *op. cit.*

[3] Everitt, W. L., and Karl R. Spangenberg, Grid-current Flow as a Factor in the Design of Vacuum-tube Power Amplifiers, *Proc. I.R.E.*, vol. 26, pp. 612–639, May, 1938.

cathode-grid distance leads to a small current-division factor. As the electrode voltages and correspondingly the current are increased, the virtual cathode moves back toward the actual cathode, causing the current ratio first to rise and then quickly to level off. The important observation about Fig. 9.19 is that the ratio of plate to grid current for equal grid and plate voltages is constant enough to make it eligible for a position as a fourth tube constant. The current-division factor in a tube free of secondary emission is as constant as the mu of the tube.

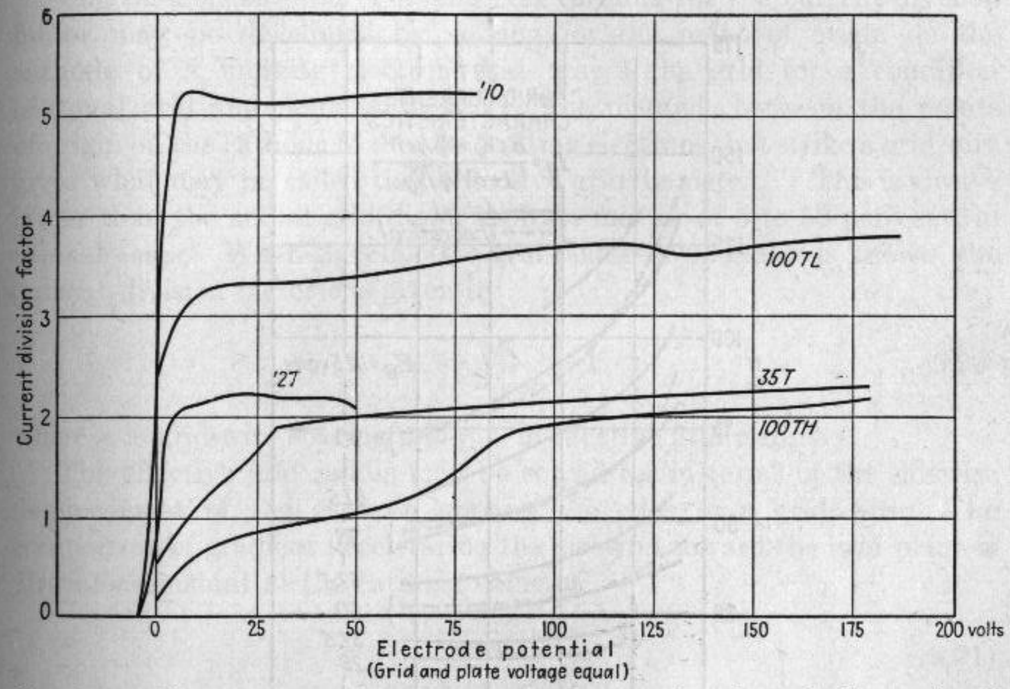


Fig. 9.19.—Plate-current–grid-current ratio as a function of equal plate and grid voltage.

Even when a triode has considerable secondary emission, the measured current ratio for equal positive grid and plate voltages is nearly equal to the primary current ratio because the interchange of secondary electrons between grid and plate is small when their voltages are equal.

Approximate Primary-grid-current Law. Since the total space current in a positive-grid triode is the sum of the grid and plate current,

$$I_s = I_p + I_g \tag{9.14}$$

this can be written

$$I_g = \frac{I_s}{1 + \frac{I_p}{I_g}} \tag{9.15}$$

Then, substituting the relation of Eq. (9.12),

$$I_g = \frac{I_s}{1 + \delta\sqrt{\frac{V_p}{V_g}}} \tag{9.16}$$

The space current itself is given by

$$I_s = G\left(V_g + \frac{V_p}{\mu}\right)^a \tag{9.17}$$

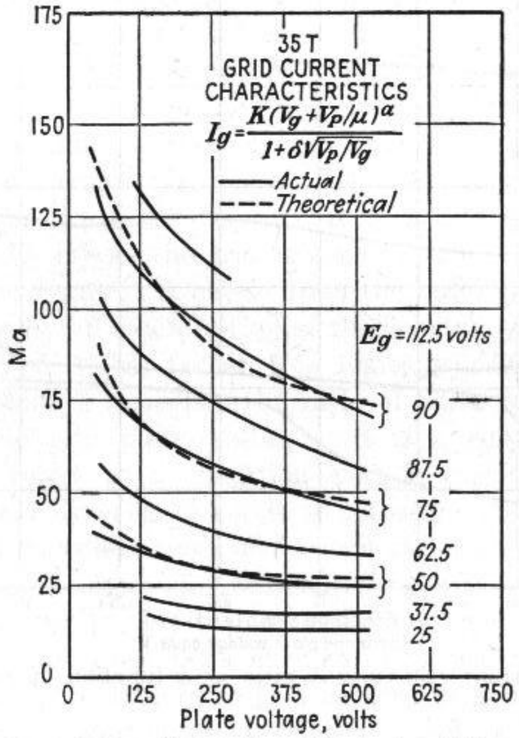


FIG. 9.20.—Comparison of actual and theoretical grid current.

where G is the perveance and a is a constant, approximately $3/2$. The resulting expression for primary grid current is given by

$$I_g = \frac{G\left(V_g + \frac{V_p}{\mu}\right)^a}{1 + \delta\sqrt{\frac{V_p}{V_g}}} \tag{9.18}$$

An idea of the accuracy of this approximation is given by Fig. 9.20, in which actual and theoretical grid-current curves are compared.

The expression for primary plate current for positive grid voltages corresponding to Eq. (9.18) is

$$I_p = \frac{G\left(V_g + \frac{V_p}{\mu}\right)^a}{1 + \frac{1}{\delta\sqrt{\frac{V_p}{V_g}}}} \tag{9.19}$$

Current-division-factor Formula. A formula for the current-division factor may be developed by solving for the point of origin on the cathode of a limiting electron that grazes the grid for a condition of equal grid and plate voltages.[1,2] The distance between the points of origin on the cathode of the two limiting electrons that strike a grid wire gives what may be called the "effective grid diameter." This is always larger than the actual grid diameter by a matter of 5 to 50 per cent, in typical cases. When the effective grid diameter or radius is known, the current-division factor δ is given by

$$\delta = \frac{a}{2r_{g\,\text{eff}}} - 1 \tag{9.20}$$

where a is grid-wire spacing and $r_{g\,\text{eff}}$ is effective grid radius.

The effective grid radius may be solved for in terms of the sidewise displacement of the electron grazing the edge of a grid wire. The component of gradient accelerating the electron toward the grid plane is virtually constant at the cathode value of

$$E_x = \frac{q_c}{\varepsilon_0} \tag{9.21}$$

The component of gradient giving the electron its sidewise deflection is

$$E_y = \frac{q_g \sin\left(\frac{2\pi y}{a}\right)}{2a\varepsilon_0\left[\cosh\left(\frac{2\pi x}{a}\right) - \cos\left(\frac{2\pi y}{a}\right)\right]} \tag{9.22}$$

where x and y are measured from a grid-wire center as in Fig. 7.17, q_c is given by Eq. (7.14a), and q_g is given by Eq. (7.14b). The sidewise deflection of the grazing electron is very nearly that which is obtained

[1] TELLEGEN, B. D. H., De Groote van der Roosterstroom in een Triode, *Physica*, vol. 6, pp. 113–116, March, 1926.

[2] SPANGENBERG, K. R., Current Division in Plane Electrode Triodes, *Proc. I.R.E.*, vol. 28, pp. 226–236, May, 1940.

by assuming that the sidewise force on the grazing electron is the same as that which exists along a line starting on the cathode at a point opposite the edge of a grid wire and passing tangent to the grid wire. This assumed force is correct at the point of contact on the grid, at which point the force is greatest.

Making the small-value approximation for x and taking y as r_g in Eq. (9.22) above,

$$E_y = \frac{q_0 r_g}{2\pi\epsilon_0(x^2 + r_g^2)} \tag{9.23}$$

(Note that this has the correct value when $x = 0$.)

Upon substituting the approximate values of E_x and E_y from the above into the acceleration equations (6.41) and (6.42), eliminating time, and equating grid and plate voltage, the expression for the sidewise displacement of the grazing electron is found to be

$$y_0 = \frac{a\mu}{2\pi(\mu + 1)} \frac{r_g}{2d_{cg}} \ln\left(\frac{4\epsilon d_{cg}}{r_g}\right) \tag{9.24}$$

The effective grid radius is equal to $y_0 + r_g$. When the expression for the effective grid radius is applied to Eq. (9.20), it is found that the current-division factor is

$$\delta = \frac{a}{\frac{a\mu}{\pi(\mu + 1)} \frac{r_g}{2d_{cg}} \ln \frac{4\epsilon d_{cg}}{r_g} + 2r_g} - 1 \tag{9.25}$$

in which a = distance between grid wires
μ = amplification factor
r_g = grid-wire radius
d_{cg} = cathode-grid distance
ϵ = Napierian base, 2.718

The magnitude of the current-division factor is given by the nomographs of Fig. 9.21 and 9.22. In Fig. 9.21 is a nomographic chart from which the effective grid radius is given in terms of the grid-wire spacing, the cathode-grid distance, and the amplification factor. This chart is read by means of two perpendicular lines ruled upon a transparent sheet. The construction cross shown on the chart gives the effective radius of a type 210 tube. The nomograph of Fig. 9.22 is a graphical representation of Eq. (9.20) and gives the current-division factor from the effective grid-wire radius and grid-wire spacing. Examination of Figs. 9.21 and 9.22 shows that the current-division factor increases with both grid-wire spacing and grid-cathode spacing. The current-division factor also increases with amplification factor, but only slightly. Typical

values of the effective grid-wire radius will be 105 to 150 per cent of the actual grid-wire radius.

Current-division Law in the Presence of Secondary Emission. When the analysis that led to the current-division factor is generalized by

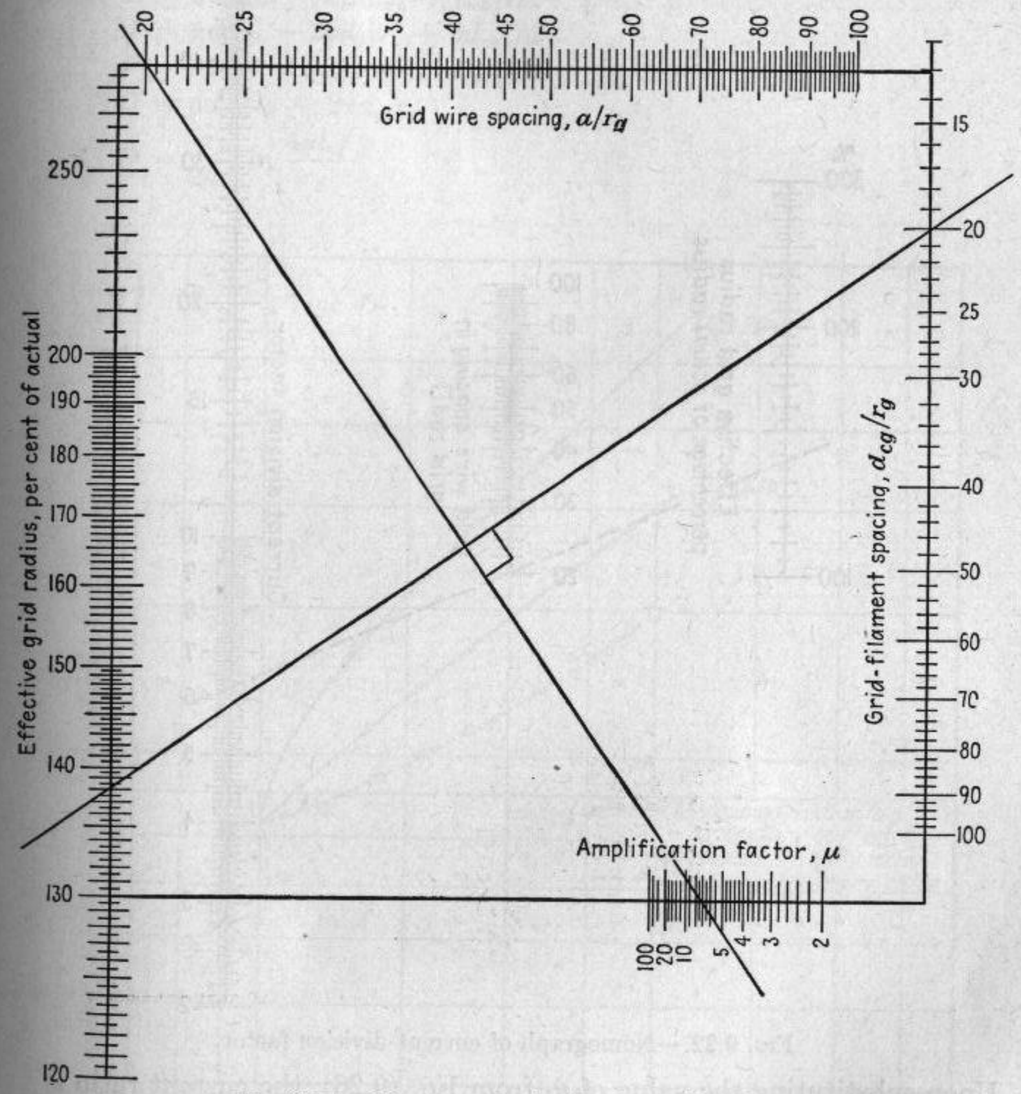


FIG. 9.21.—Nomograph of effective grid radius.

allowing the grid and plate voltage to assume general values, the sidewise deflection of the electron grazing the grid is found to be

$$y_1 = \frac{a\mu[(d_{cg} + d_{gp})V_g - d_{cg}V_p]}{2\pi d_{gp}(V_p + V_g)} \frac{r_g}{2d_{cg}} \ln\left(\frac{4\epsilon d_{cg}}{r_g}\right) \tag{9.26}$$

Equation (9.26) has been arrived at by solving for the sidewise displacement of the electrons grazing the grid as a function of electrode-voltage ratio. The corresponding current ratio is then readily determined. The

electrode-current ratio is

$$\frac{I_p}{I_g} = \frac{a}{2(y_1 + r_g)} - 1 \tag{9.27}$$

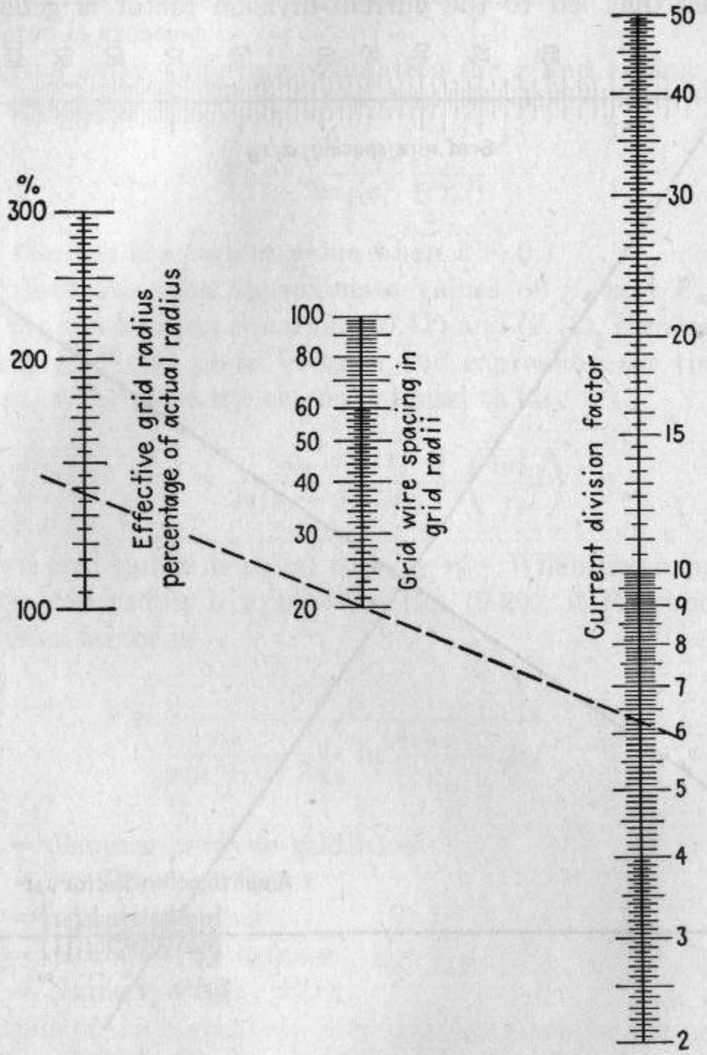


FIG. 9.22.—Nomograph of current-division factor.

Upon substituting the value of y_1 from Eq. (9.26), the current ratio is

$$\frac{I_p}{I_g} = \frac{a}{2\left\{\dfrac{a\mu\left[(d_{cg} + d_{gp}) - d_{cg}\dfrac{V_p}{V_g}\right]D}{2\pi d_{gp}\left(\dfrac{V_p}{V_g} + \mu\right)} + r_g\right\}} - 1^* \tag{9.28a}$$

* A somewhat more accurate formula has since been developed by J. H. L. Jonker and B. D. H. Tellegen, Current to a Positive Grid in Electron Tubes, *Philips Research Reprints*, vol. 1, pp. 13–32, October, 1945.

or

$$\frac{I_p}{I_g} = \frac{LV_g + MV_p}{PV_g + QV_p} \tag{9.28b}$$

where $L = \pi a d_{gp}\mu - a\mu(d_{gp} + d_{cg})D - 2\pi d_{gp} r_g \mu$
$M = \pi a d_{gp} - 2\pi d_{gp} r_g + a d_{cg} D\mu$
$P = a(d_{gp} + d_{cg})\mu + 2\pi d_{gp} r_g \mu$
$Q = a d_{cg} D\mu - 2\pi d_{gp} r_g$
$D = \dfrac{r_g}{2d_{cg}} \ln \dfrac{4\epsilon d_{cg}}{r_g}$

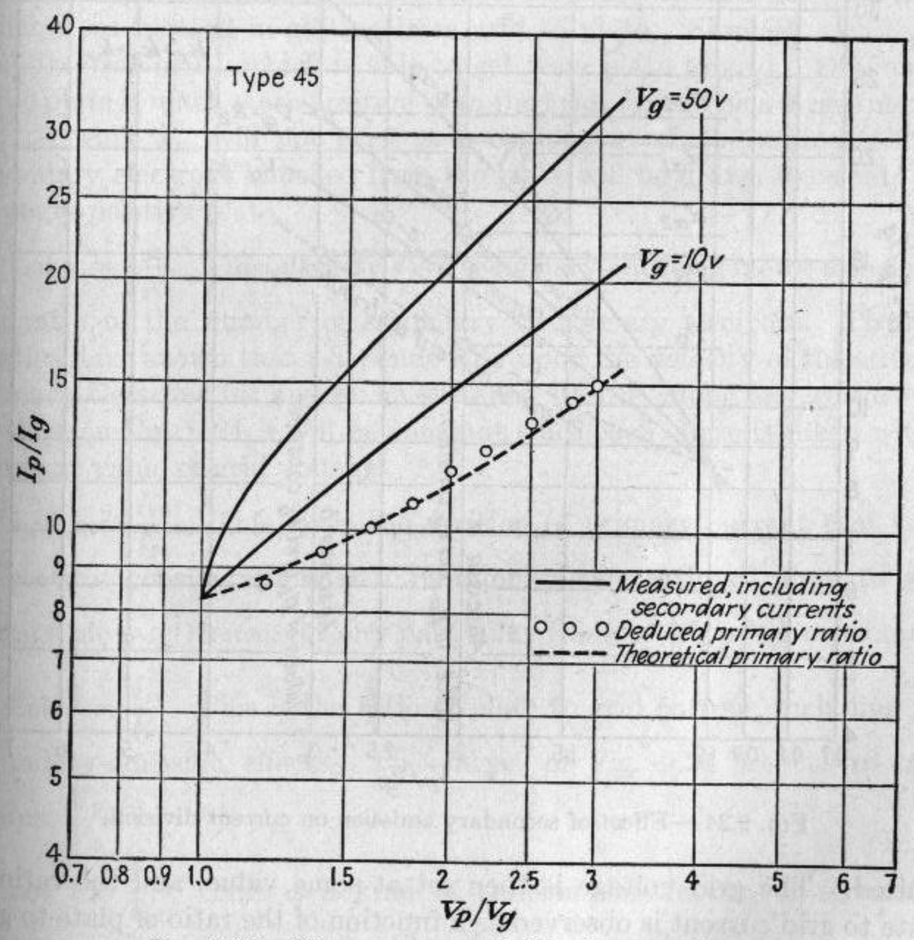


FIG. 9.23.—Primary-current division in a 45 triode.

A plot of Eq. (9.28*a*) in a typical case shows the curve to be concave upward as in the dotted curve of Fig. 9.23. In such a plot, the slope of the true current ratio is between ¼ and ¾ so that the assumption of a one-half-power law when this curve is slightly modified by space-charge and secondary-emission effects is a reasonable one.

To check the correctness of the above equation in actual tubes it is necessary to correct measured curves for the effect of secondary emission, which is always present to a degree. This is done by an extension of methods developed for screen-grid tubes.[1]

The curves from which the deduction of the true primary distribution are made are taken as follows: Filament emission is first reduced to the point where the current is temperature-limited rather than space-charge-

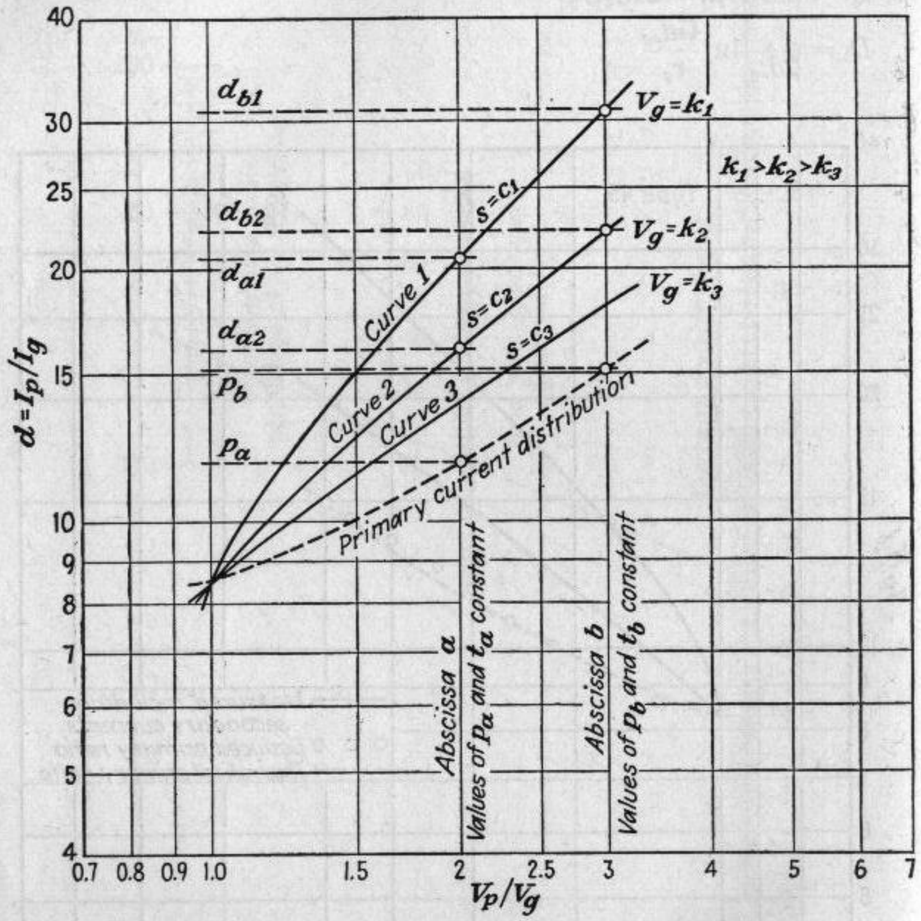


FIG. 9.24.—Effect of secondary emission on current division.

limited. The grid voltage is then set at some value, and the ratio of plate to grid current is observed as a function of the ratio of plate to grid voltage by varying the plate voltage only. The grid voltage is then set at another positive value, and another similar run is made. The two solid curves of Fig. 9.23 were made by this method.

Because of the various factors that have been held constant and

[1] DE LA SABLONIERE, C. J. L., Die Sekundäremission in Schrimgitterrohren, *Hochfreq. u. Audio.*, vol. 41, pp. 195–202, June, 1933.

relations between the various current components a number of relations exist that must be borne in mind. Before summarizing these relations the notation to be used must be indicated in detail. Let I_p and I_g be total plate and grid current, respectively, including secondaries. Let I_{p1} and I_{g1} be those parts of the plate and grid currents which are due to primary electrons, *i.e.*, the primary plate and grid currents. Let I_{p2} and I_{g2} be the currents corresponding to all the secondary electrons that are knocked out of the plate and grid, respectively. This includes not only those secondary electrons which succeed in getting from one electrode to another but also those which are knocked out of one electrode and fall back into that same electrode. Let I_{gp} be that fraction of I_{g2} which does succeed in getting from grid to plate. Similarly, let I_{pg} be that fraction of I_{p2} which is able to get from plate to grid. Obviously, if the plate is much more positive than the grid, I_{gp} will be a large fraction of I_{g2}, while I_{pg} will not exist as a component of I_{p2} because all the secondary electrons knocked from the plate will be drawn back into the strongly positive plate.

Let $s = \dfrac{I_{g2}}{I_{g1}}$. The quantity s is a secondary-emission factor measuring the ratio of the number of secondary to primary electrons. Physical studies have shown that s depends only upon the velocity of the striking primary electrons for any given surface. Hence, along any curve such as those in Fig. 9.24, s will be constant since each curve is taken with a constant value of grid voltage.

Let $p = \dfrac{I_{p1}}{I_{g1}}$. This gives the division of primary current that from theoretical considerations is a function of the ratio of plate to grid voltage alone. Hence, for any particular value of $\dfrac{V_p}{V_g}$, p is a constant.

Let $d = \dfrac{I_p}{I_g}$. This is the ratio of plate to grid current, including the secondary-emission effects. The curves of Fig. 9.24 are curves of d against $\dfrac{V_p}{V_g}$.

Let $t = \dfrac{I_{gp}}{I_{g2}}$. This is a kind of transmission factor for secondary electrons. It measures the fraction of secondaries liberated that succeeds in getting to the plate. Some secondary electrons from the grid have such a low velocity that they are unable to climb the small potential hill between the grid and the plate. De la Sabloniere has assumed that for any value of the abscissa $\dfrac{V_p}{V_g}$ the value of t is constant. That is, for

any value of $\frac{V_p}{V_g}$ the same fraction of the secondary electrons knocked from the grid succeeds in getting to the plate. This is perhaps the only assumption which is questionable. The matter is complicated by the velocity distribution of the secondary electrons, which changes as the striking voltage of the primary electrons changes. For the assumption to be strictly true the velocity-distribution curve of the secondary electrons must expand uniformly as the striking potential of the primary electrons increases. This is not strictly true but for small ranges of primary-electron velocity is approximately so. In the curves of Fig. 9.24 the primary-electron velocities are 10 and 50 volts. It was not found possible to get a good check for velocities of 10 and 200 volts, this being too great a range of primary velocities.

It will be noted further that the space current for each of the experimentally determined curves is approximately constant.

Consider the ratio

$$\frac{I_p}{I_g} = \frac{I_{p1} + I_{gp}}{I_{g1} - I_{gp}} \tag{9.29}$$

Dividing both numerator and denominator by I_{g1} there results

$$\frac{I_p}{I_g} = \frac{\frac{I_{p1}}{I_{g1}} + \frac{I_{gp}}{I_{g1}}}{1 - \frac{I_{gp}}{I_{g1}}} \tag{9.30}$$

But

$$\frac{I_{gp}}{I_{g1}} = \frac{I_{gp}}{I_{g2}} \frac{I_{g2}}{I_{g1}} = ts \tag{9.31}$$

so that the above ratio of net currents can be written as

$$d = \frac{p + ts}{1 - ts} \tag{9.32}$$

Solving this for ts,

$$ts = \frac{d - p}{d + 1} \tag{9.33}$$

Let the various curves of d against $\frac{V_p}{V_g}$ be numbered 1, 2, and so on, as shown in Fig. 9.24. Let the various values of $\frac{V_p}{V_g}$ have letters corresponding to them. Thus the abscissa of $\frac{V_p}{V_g} = 2$ might be lettered a, that of $\frac{V_p}{V_g} = 3$ might be lettered b, and so on. If we consider the four points

formed by the intersection of the upper two curves of Fig. 9.24 and any two abscissas denoted by a and b, then it is possible to write four equations of the form of that last given. These will be

$$t_a s_1 = \frac{d_{a1} - p_a}{d_{a1} + 1} \tag{9.34}$$

and

$$t_a s_2 = \frac{d_{a2} - p_a}{d_{a2} + 1} \tag{9.35}$$

for the intersections of the curves 1 and 2 with the abscissa a and

$$t_b s_1 = \frac{d_{b1} - p_b}{d_{b1} + 1} \tag{9.36}$$

$$t_b s_2 = \frac{d_{b2} - p_b}{d_{b2} + 1} \tag{9.37}$$

for the other intersections. Since t and p are presumed constant for any particular value of $\frac{V_p}{V_g}$, they are given only a lettered subscript.

Dividing the two pairs of equations and equating them gives

$$\frac{d_{a1} - p_a}{d_{a2} - p_a} = \frac{d_{b1} - p_b}{d_{b2} - p_b} \tag{9.38}$$

which is the relation that has been sought. This may be solved for p_b to give

$$p_b = \frac{d_{b1} - d_{b2}\left(\frac{d_{a1} - p_a}{d_{a2} - p_a}\right)}{1 - \left(\frac{d_{a1} - p_a}{d_{a2} - p_a}\right)} \tag{9.39}$$

From this last equation it may be seen that if one point, p_a, on the true primary distribution curve is known, then points at any other abscissa b may be found from a pair of curves giving the net current division in the presence of secondary emission. The above treatment has been given for the case of V_p greater than V_g, but a similar treatment can be applied when this is not so. In this particular instance the primary-current distribution that was taken as known was that corresponding to the condition of the grid being at its "natural potential" relative to the plate. For this case, the electrons move in substantially parallel straight lines from filament to grid and plate, and the ratio of plate to grid current is determined by the ratio of intergrid to grid area. For the 45 tube this ratio of currents is 14.3 when the ratio of voltages is 2.81.

CHAPTER 10

TETRODES

10.1. Types of Tetrode. A tetrode, as its name implies, is a four-electrode tube. The four electrodes are invariably, in the order of their arrangement, the cathode, the control grid, the screen grid, and the plate. There are two types of tetrode. These are the so-called "screen-grid tube" and the "beam-power tube."

The screen-grid tube was the successor to the triode and the predecessor of the pentode, though, as indicated in the chapter on Basic Tube Types, it is now virtually obsolete and seldom used because of unfavorable current-voltage characteristics. The ordinary screen-grid tube has a fine control grid surrounding the emitter, which in turn is surrounded by a coarser screen grid a considerably greater distance out. The screen grid is in turn surrounded by a plate. The intended function of the screen grid was to shield the control grid electrostatically from the plate and so reduce the tendency toward oscillation that existed in r-f amplifiers. The screen grid performed this function, but it also introduced some other characteristics that were not desirable. Specifically, it introduced secondary emission, which distorts the current-voltage characteristics.

The beam-power tube is a special tetrode with aligned control and screen grids. It was the historical successor to the pentode. The pentode was developed to eliminate the secondary-emission action that appears in the screen-grid tube. The beam-power tube was later found capable of doing the same thing without an extra grid if proper attention were paid to grid alignment and to dimensioning.

10.2. Current-voltage Characteristics of the Screen-grid Tube. The screen-grid tube is usually operated with its screen at a fixed direct potential and by-passed with a large condenser to ground so that no alternating components of potential appear on it. The screen grid acts as a shield between the plate and control grid. Electrostatic lines from the plate terminate for the most part on the screen grid. This electrostatic behavior does not interfere with electronic action. An electron stream of varying intensity can still pass between the screen-grid wires.

The current-voltage characteristics of the screen-grid tube are determined by two principal effects that are at work. (1) The relative screen-grid and plate potentials determine how the space current will divide

between these two electrodes. (2) The relative positiveness of plate and screen grid determines how secondary electrons will be interchanged between these two electrodes. In general, the behavior with regard to both factors is similar to that which exists in the positive-grid triode.

As far as space-current effects are concerned, the control grid and screen grid have the principal influence. With respect to first-order effects, it may be said that the screen grid plays the same role in the screen-grid tube as the plate does in the triode. The plate has only a very small influence in modifying space current in the screen-grid tube because of the shielding effect of the screen grid. The space current is given by

$$I_s = G\left(V_1 + \frac{V_2}{\mu_{sg}} + \frac{V_p}{\mu_p}\right)^a \tag{10.1}$$

where V_1 is control-grid potential
V_2 is screen-grid potential
G is perveance
V_p is plate potential
μ_{sg} is equivalent amplification factor of the screen grid, $-\left(\frac{dV_2}{dV_1}\right)_{I_{p\,\text{const}}}$
μ_p is plate amplification factor, $-\left(\frac{dV_p}{dV_1}\right)_{I_{p\,\text{const}}}$
a is a constant, nearly $\frac{3}{2}$

In this expression, μ_{sg} is considerably smaller than μ_p. The equivalent screen-grid amplification factor may be calculated quite accurately from the triode mu formulas by treating the screen grid as though it were the plate. The accuracy of this approximation decreases as the shielding effect of the screen grid decreases. The plate amplification factor may be calculated from some special formulas, which will be developed subsequently. It may be determined approximately by calculating a triode amplification factor, considering the control grid as the cathode, the screen grid as the control grid, and the plate as the plate, and then multiplying this amplification factor by the screen-grid amplification factor. This relation holds because the fictitious amplification factor cited first above measures the screening effect of the screen grid upon the control-grid plane just as the screen-grid amplification factor measures the screening effect of the control grid upon the cathode. The product of these two amplification factors, which are reciprocal screening factors, gives the over-all amplification factor. Thus, if the screen-grid amplification factor were 20 and the triode amplification factor obtained by considering the control grid as the cathode were 10, the plate amplification

factor would be approximately 200. If the cold cathode were at zero potential and all the other electrodes were at the same positive potential, then one-twentieth of the electrostatic flux lines from the cathode would penetrate the control grid into the space beyond (actually, the ratio would be 1 in 21). Of the lines that passed through the control grid, one-tenth would pass on to the plate, and the rest would terminate on the screen grid. The over-all screening effect would be such that only 1 line would reach the plate for every 200 that reached the control grid. The resulting plate amplification factor is 200.

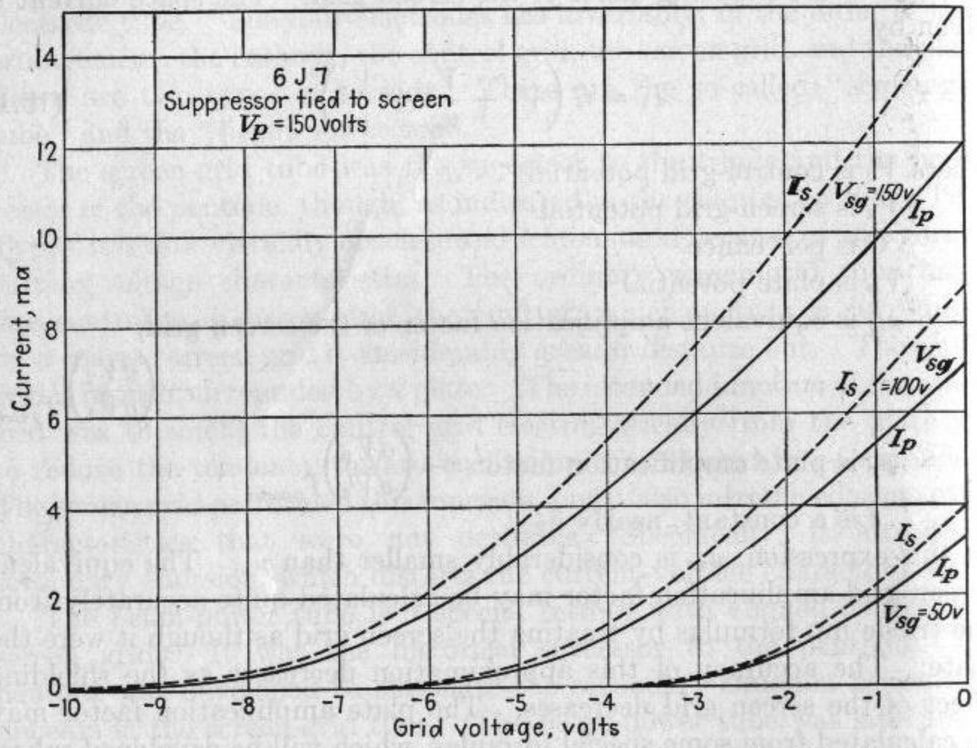


FIG. 10.1.—Plate and space current of a screen-grid tube as a function of control-grid voltage.

The plate-current–control-grid characteristics of the screen-grid tube are almost the same as the triode characteristics that result if the screen grid and plate are connected together. The only difference is that a small part of the space current is taken by the screen grid. Some typical plate-current and space-current characteristics as a function of control-grid voltage are shown in Fig. 10.1. Because of the usually high value of the plate amplification factor the plate potential has only a small effect upon the plate and space current compared with the screen-grid potential. This in turn has much less influence than the control-grid potential.

Plate-current–Plate-voltage Characteristics of the Screen-grid Tube. With a negative control-grid voltage and a positive screen-grid voltage, the plate-current–plate-voltage characteristics of a screen-grid tube have the form shown in Fig. 10.2. The shape of the plate-current curve departs considerably from the shape of the primary plate-current curve because of secondary emission. The probable shape of the primary plate-current curve has been sketched for $V_g = 0$. The primary plate current is not readily measured directly. It is seen to be an increasing fraction of the approximately constant space current I_s. The division of space current between screen grid and plate follows approximately

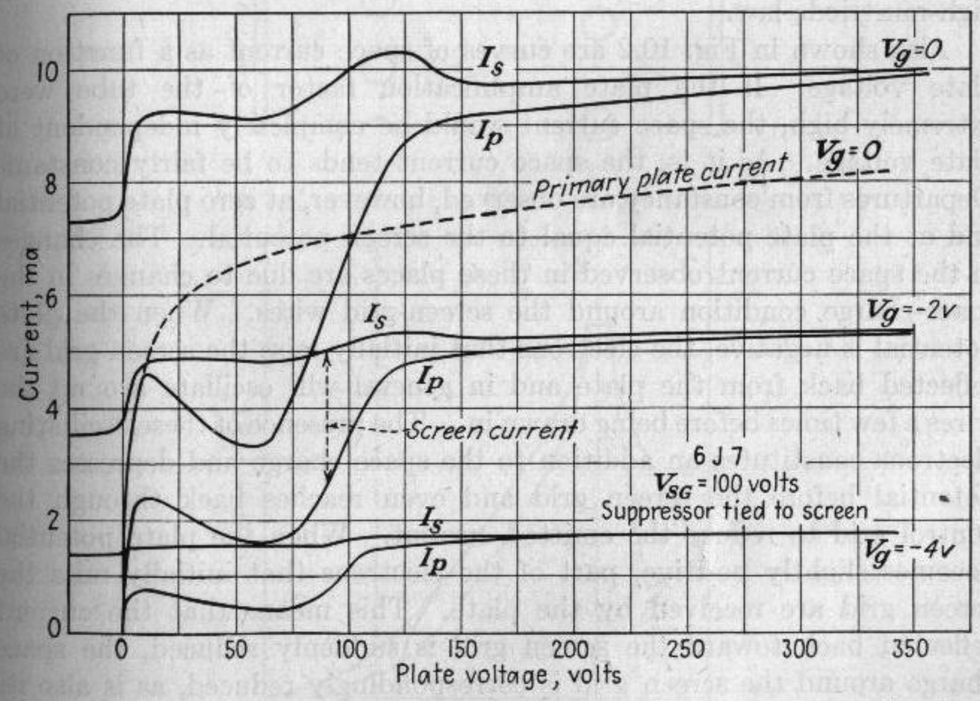


FIG. 10.2.—Plate-current–plate-voltage characteristics of a screen-grid tube.

the same law as does the division of the space current between grid and plate in a triode. When the plate voltage is zero, the plate gets none of the space current. As the plate is made positive, it rapidly acquires a major portion of the space current. When the plate is as positive as the screen grid, it gets a slightly smaller fraction of the total space current than the ratio of the area between the screen-grid wires to the total area of the screen-grid plane. As the plate potential is made still more positive, the plate acquires still more of the space current until at very large voltages the plate is getting nearly all the space current.

The difference between the primary plate-current curves and the actual plate-current curves is obviously due to secondary-emission effects. The effects are the same as in the positive-grid triode. When

the plate is less positive than the screen grid, secondary electrons liberated at the plate surface are attracted to the screen grid, thus reducing the plate current. This accounts for the pronounced dip in the plate-current characteristic. When the plate potential is equal to the screen-grid potential, the interchange of secondary electrons between plate and screen grid nearly balances and the actual plate current is nearly equal to the primary plate current. As the plate becomes more positive, it collects secondary electrons that are liberated from the screen grid, and, as a result, the actual plate current exceeds the primary plate current. The variation of plate current with control-grid voltage follows the high-mu-triode law.

Also shown in Fig. 10.2 are curves of space current as a function of plate voltage. If the plate amplification factor of the tube were extremely high, the space current would be completely independent of plate voltage. As it is, the space current tends to be fairly constant. Departures from constancy are observed, however, at zero plate potential and at the plate potential equal to the screen potential. The changes in the space current observed in these places are due to changes in the space-charge condition around the screen-grid wires. When the plate potential is negative, the electrons that initially miss the screen grid are reflected back from the plate and in general will oscillate around the wires a few times before being drawn in. The presence of these oscillating electrons constitutes an addition to the space charge and depresses the potential before the screen grid and even reaches back through the control grid to reduce the emitted current. When the plate potential becomes slightly positive, part of the electrons that initially miss the screen grid are received by the plate. This means that the current reflected back toward the screen grid is suddenly reduced, the space charge around the screen grid is correspondingly reduced, as is also its depressing effect upon the potential before the cathode, and as a result the emitted current suddenly increases. The nature of the change in the potential distribution within the screen-grid tube as the plate potential is changed from negative to positive is sketched in Fig. 10.3. The dotted lines in this figure show potential profiles for a negative plate potential, while the solid lines show potential profiles for a positive plate potential. The manner in which the plate potential controls the off-cathode gradient through the medium of the oscillating space charge about the screen grid may also be seen.

When the plate potential becomes more positive than the screen-grid potential, there is a change from a condition of partial reflection of electrons from the plate to one of no reflection, for all electrons reach the plate, no matter how strongly deflected by the screen grid. Here the

space charge around the screen grid is again suddenly reduced, and the space current increases. Thus the space current is influenced most by the condition of current transmission to the plate and is hardly affected by secondary emission.

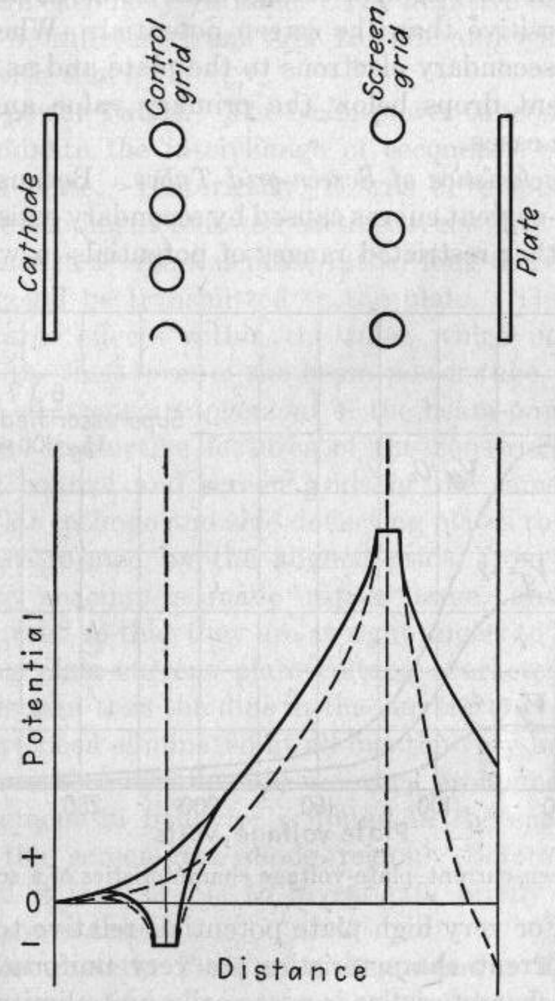


FIG. 10.3.—Potential distributions within a screen-grid tube for negative and positive plate potentials.

Screen-current–Plate-voltage Characteristics of the Screen-grid Tube. The screen-grid current is the difference between the space current and the plate current in Fig. 10.2. This difference is plotted as screen current in Fig. 10.4 as a function of plate voltage. The screen-current–plate-voltage curves are like the positive-grid-current–plate-voltage curves of a triode. Exactly the same factors enter into its composition. The primary distribution is such that the screen current decreases uniformly

with increasing plate voltage exactly as is the case for the triode grid. When the plate voltage is less positive than the screen-grid voltage, the screen grid acquires current from the plate and hence rises above the primary-current value. If the secondary emission is sufficient, the net screen-grid current will rise with voltage until it falls as the plate potential becomes more positive than the screen potential. When this happens, the screen loses secondary electrons to the plate and as a result the net screen-grid current drops below the primary value and may even go negative in some cases.

General Characteristics of Screen-grid Tubes. Because of the distortions in the plate-current curves caused by secondary emission, the screen-grid tube has rather restricted ranges of potentials in which it operates

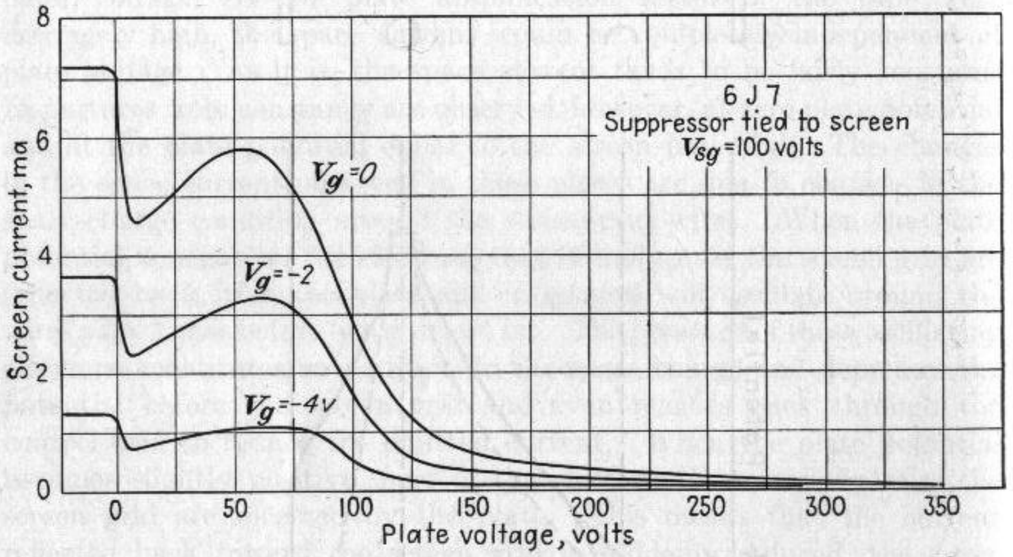


FIG. 10.4.—Screen-current–plate-voltage characteristics of a screen-grid tube.

satisfactorily. For very high plate potentials relative to the screen-grid potential, the current characteristics are very uniform. The range of uniform current characteristics is necessarily quite limited. The screen-grid potential must be relatively high to draw sufficient current. The plate potential must be at least this positive to avoid secondary-emission distortions and yet cannot be too much more positive because then the plate dissipation becomes excessive. In this operating region the plate resistance of the tube is very high. The amplification factor is also high, but the mutual conductance is of the same order as in a triode.

Use is sometimes made of the negative plate-resistance characteristic that the screen-grid tube displays at low plate potentials. It will be recalled that the plate resistance of a tube is given by the reciprocal of the slope of the plate-current–plate-voltage characteristic. Hence

the plate resistance is negative whenever the slope of this characteristic is negative. The negative resistance that can be realized from a screen-grid tube has a limited amplitude of current and voltage to which it can be subjected. It is further not very stable because secondary-emission characteristics are extremely variable. The negative resistance that can be realized will be different from tube to tube and will even change in the same tube with time.

10.3. Beam-power Tubes. The beam-power tube is a special tetrode designed to eliminate the interchange of secondary electrons between screen grid and plate. Historically, it was developed later than the pentode. Its development followed the discovery that when the screen-grid–plate distance in a tube was made rather long there was a maximum current which could be transmitted to the plate. This led to a study of the space-charge effects within the tube, which in turn led to the development of the final form of the beam-power tube.

The internal electrode arrangement of the beam-power tube is shown in Fig. 2.6. The distinctive features of the construction of this tube are the aligned control and screen grids of the same pitch. This is coupled with a flat cathode and side deflecting plates to keep the current sheets, which are formed by the aligned grids, from spreading. The screen-grid–plate spacing is made rather large, and the successive electrodes are curved so that they are at right angles to the electron flow.

The resulting plate-current–plate-voltage characteristics are shown in Fig. 2.7. It is seen that the dips in the current curves due to secondary emission have been eliminated at all but the very lowest control-grid voltages, and even here the dips are not very pronounced. The reason for this improvement in behavior is found in the space-charge effects that occur in the screen-grid–anode region. Before examining this subject in detail it is desirable to investigate briefly the electrostatic field of a beam-power tube.

10.4. The Electrostatic Field of a Beam-power Tube. The same general methods that have been described in the chapter on Triode Characteristics can be applied to multielectrode tubes in some cases. For tube structures in which the grid wires have a regular pattern the method of conformal transformations is easily applied. This is the case for the tetrode with aligned grids, the structure of the beam-power tube, which will be treated here by an extension of the method employed with the triode.

The line-charge configuration of Fig. 10.5*a* gives rise to the configuration of Fig. 10.5*b* upon application of the transformation $W = \frac{a}{2\pi} \ln Z$. The relation between the parts in the two planes is apparent from the

previous study of triodes. The small circle about the origin in the Z plane goes into a cathode line in the W plane. The control-grid-wire circle at (1,0) in the Z plane goes into the series of equally spaced control-grid wires in the W plane. The screen-grid wire at $(\epsilon^{\frac{2\pi d_{gs}}{a}}, 0)$ in the Z plane goes into the line of screen-grid wires in the W plane. A large plate circle about the origin in the Z plane goes into the plate line in the W plane.

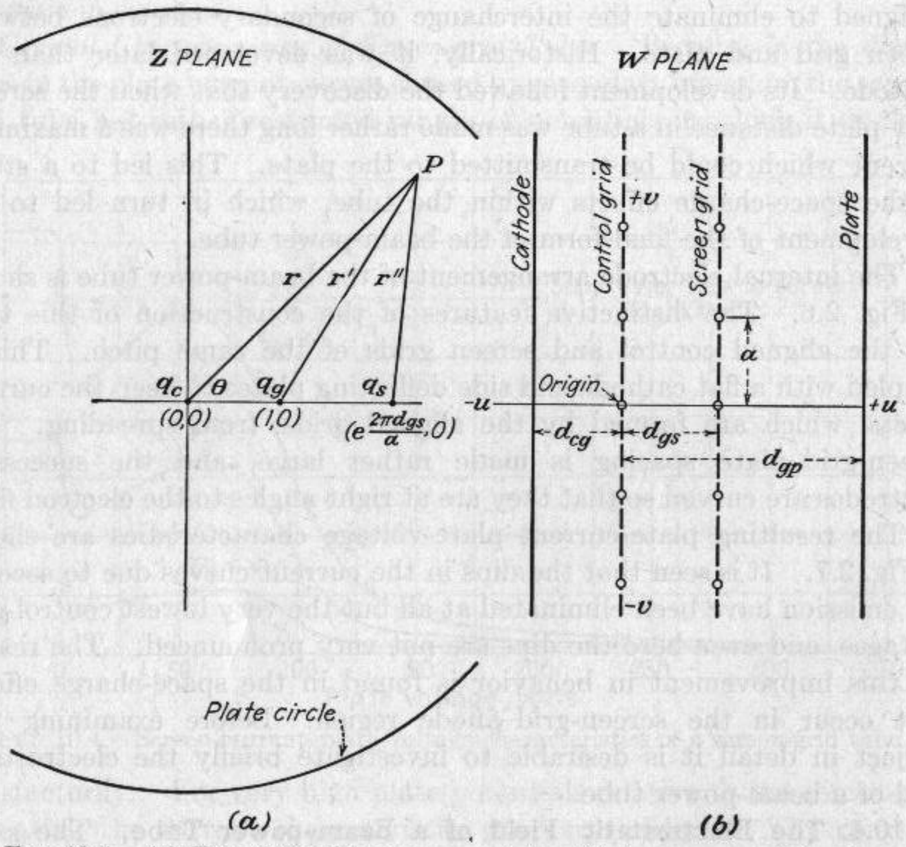


FIG. 10.5.—(*a*) Transformed beam-power tube, (*b*) beam-power-tube electrode arrangement.

It is necessary only to write an expression for the potential at any point P in the Z plane, transform it by the logarithmic transformation, and then evaluate the electrode potentials in terms of the charges and potentials. This is the procedure that was used for the triode, though the form of the resulting expressions may be expected to be more complicated because of the introduction of another electrode. In the treatment that follows the small-grid-wire approximations will be made.

The potential at any point P in the Z plane is given by

$$V = -\frac{q_c}{2\pi\epsilon_0}\ln r - \frac{q_g}{2\pi\epsilon_0}\ln r' - \frac{q_s}{2\pi\epsilon_0}\ln r'' + C \qquad (10.2)$$

but

$$(r')^2 = (1 + r^2 - 2r\cos\theta) \qquad (10.3)$$

and

$$(r'')^2 = (\epsilon^{\frac{4\pi d_{gs}}{a}} + r^2 - 2\epsilon^{\frac{2\pi d_{gs}}{a}} r\cos\theta) \qquad (10.4)$$

Substitution of Eqs. (10.3) and (10.4) into Eq. (10.2) and application of the logarithmic-transformation coordinate equations gives

$$V = -\frac{u}{a\epsilon_0}q_c - \frac{q_g}{4\pi\epsilon_0}\ln\left(1 - 2\epsilon^{\frac{2\pi u}{a}}\cos\frac{2\pi v}{a} + \epsilon^{\frac{4\pi u}{a}}\right) - \frac{q_s}{4\pi\epsilon_0}\ln\left(\epsilon^{\frac{4\pi d_{gs}}{a}} - 2\epsilon^{\frac{2\pi(u+d_{gs})}{a}}\cos\frac{2\pi v}{a} + \epsilon^{\frac{4\pi u}{a}}\right) + C \qquad (10.5)$$

To determine the electrode potentials in terms of the charges and dimensions, let $u = -d_{cg}$, $v = 0$, where $d_{cg} \gg a$, and set the cathode potential equal to zero. This gives an expression for the constant C that can be put into subsequent expressions.

$$C = \frac{1}{a\epsilon_0}(-q_c d_{cg} + q_s d_{gs}) \qquad (10.6)$$

To find the control-grid potential let $u = 0$, $v = r_g$, where $r_g < \frac{a}{20}$. Then

$$V_g = -\frac{d_{cg}q_c}{a\epsilon_0} - \frac{q_g}{2\pi\epsilon_0}\ln\frac{2\pi r_g}{a} \qquad (10.7)$$

To find the screen potential let $u = d_{gs}$, $v = r_s$ where $d_{gs} > a$. Then

$$V_s = -\frac{d_{sg} + d_{cg}}{a\epsilon_0}q_c - \frac{d_{gs}}{a\epsilon_0}q_g - \frac{q_s}{2\pi\epsilon_0}\ln\frac{2\pi r_s}{a} \qquad (10.8)$$

To find the plate potential let $u = d_{gp}$, $v = 0$, where $d_{gp} \gg 2a$. Then

$$V_p = -\frac{d_{gp} + d_{cg}}{a\epsilon_0}q_c - \frac{d_{gp}}{a\epsilon_0}q_g - \frac{d_{gp} - d_{gs}}{a\epsilon_0}q_s \qquad (10.9)$$

The last three equations give three electrode potentials in terms of three charges. The system can, of course, be solved for the charges in terms of the potentials. Solving for the cathode charge,

$$q_c = \frac{1}{\Delta a^2\epsilon_0^2}\left\{\left[d_{gs}d_{sp} - \frac{a}{2\pi}d_{gp}\ln\left(\frac{2\pi r_s}{a}\right)\right]V_g - V_s\frac{a}{2\pi}\ln\left(\frac{2\pi r_g}{a}\right)d_{sp} + V_p\left(\frac{a}{2\pi}\right)^2\ln\frac{2\pi r_g}{a}\ln\frac{2\pi r_s}{a}\right\} \qquad (10.10)$$

where Δ is the determinant of the coefficients of the q's in Eqs. (10.7), (10.8), and (10.9). From Eq. 10.10 the grid-plate amplification factor is given by the ratio of the coefficients of V_g and V_p as

$$\mu_p = \frac{d_{gs}d_{sp} - \frac{ad_{gp}}{2\pi}\ln\left(\frac{2\pi r_s}{a}\right)}{\left(\frac{a}{2\pi}\right)^2 \ln\left(\frac{2\pi r_g}{a}\right)\ln\left(\frac{2\pi r_s}{a}\right)} \tag{10.11}$$

This expression is accurate to within a few per cent provided that the spacings between the various electrode planes are all greater than the grid-wire spacing and provided that the screening fraction of the grids (ratio of grid-wire diameter to grid-wire spacing) is less than 0.1.

It should be pointed out that the amplification factor derived above gives the relative effectiveness of the control grid and plate in controlling the *total space current* and not the plate current so that the above constant will not correspond exactly with that given in the tube manuals. That given in the tube manuals gives the relative effectiveness of the control grid and plate in controlling the *plate current*, and this depends upon the factor of Eq. (10.11) and also upon the way the space current divides between screen and plate. However, since the current-division function of a beam-power tube does not vary greatly with electrode potentials, the above expression for amplification factor is accurate enough for most purposes.

10.5. Space-charge Effects in the Screen-grid–Anode Region of Beam-power Tubes. In Sec. 8.9 of the chapter on Space-charge Effects it was shown that the effect of initial velocities in a diode was to create a virtual cathode between the actual cathode and the plate. Similarly, in tetrodes of proper design it is possible to get a virtual cathode or potential minimum between the screen grid and plate. If a satisfactory potential minimum can be achieved, it will suppress secondary electrons from the plate and do away with the need for a suppressor grid. Such a tetrode is the beam-power tube. It is in many respects superior to the conventional pentode.

In order that space-charge effects be appreciable, it is necessary that there be a very nearly parallel flow of electrons. This is not the case in the ordinary tetrode, for the use of control and screen grids with different pitches breaks up the electron flow. It is, however, possible to get what is nearly a parallel flow in a tetrode by making the control and screen grid have equal pitch and aligning the grid wires so that electrons which pass through the spaces of the control grid will also pass through the spaces in the screen grid. In Fig. 10.6 are shown some typical electron

paths in a beam-power tube.[1] The paths give a sufficiently close approximation to a parallel flow in the screen-grid–plate region so that observed tube characteristics correlate well with theoretical properties deduced from this assumption.

Assuming a parallel flow of electrons starting at a high positive potential at the screen, a number of different potential distributions are possible depending upon the plate potential and the magnitude of the

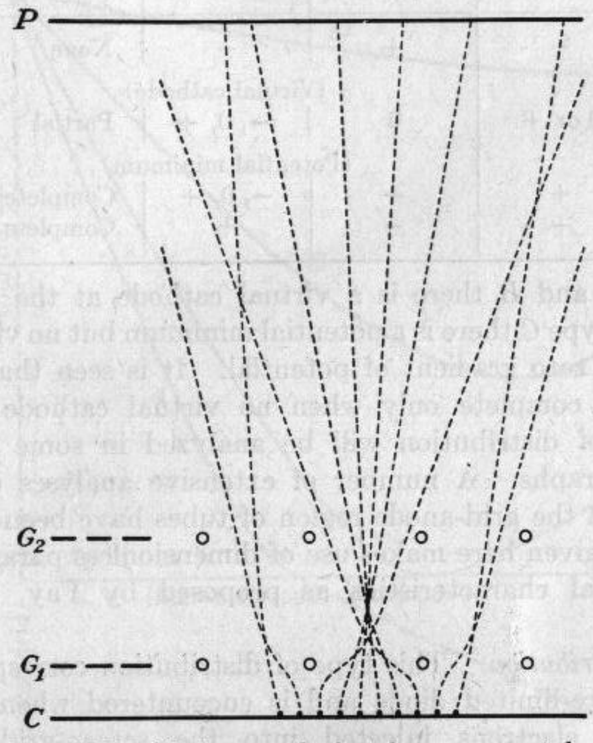


FIG. 10.6.—Electron paths in a beam-power tube. (*Jonker.*)

current injected into the screen-grid–anode region. Associated operating conditions are correspondingly different. The types of distributions encountered are shown in Fig. 10.7. The characteristics associated with these various distributions are best listed in tabular form. They are essentially determined by the sign of the constant which appears after the first integration of Poisson's equation as in Eq. (8.5) which may be written

$$\left(\frac{dV}{dx}\right)^2 = \frac{16J}{9a^2}(V^{1/2} + C_1) \tag{10.12}$$

[1] From JONKER, J. H. L., Pentode and Tetrode Output Valves, *Wireless Engr.*, vol. 26, [No. 189], pp. 274–286.

where $J = \frac{a^2V^{3/2}}{d^2}$ and a^2 has the value 2.335×10^{-6} amperes per volt$^{3/2}$. The properties of the distributions as determined by the sign of the constant C_1 are as follows:

Type	V_p	C_1	$\frac{dV}{dx}$	Current transmission to plate	V
A	−	+	−	None	+, 0, (−)
			(Virtual cathode)		
B	0 or +	0	−, 0, +	Partial	+, 0, +
			(Potential minimum)		
C	+	−	−, 0, +	Complete	+
D	+	−	+	Complete	+

In types A and B there is a virtual cathode at the point of zero potential. In type C there is a potential minimum but no virtual cathode at the point of zero gradient of potential. It is seen that the current transmission is complete only when no virtual cathode exists. The various types of distribution will be analyzed in some detail in the following paragraphs. A number of extensive analyses of the space-charge effects in the grid-anode region of tubes have been published.[1–5] The treatment given here makes use of dimensionless parameters giving rise to universal characteristics as proposed by Fay, Samuel, and Shockley.

Type A Distribution. This type of distribution corresponds to that of a temperature-limited diode and is encountered when the plate is negative. The electrons injected into the screen-grid–plate space encounter a retarding field and are thus slowed down until they finally reach a zero velocity at some point before the plate, reverse, and return

[1] HARRIES, J. H. O., The Anode to Accelerating Electrode Space in Thermionic Valves, *Wireless Engr.*, vol. 13, pp. 190–199, April, 1936.

[2] PLATO, G., W. KLEEN, and H. ROTHE, The Space Charge Equations for Electrons with Initial Velocity, Part I, *Zeit. für Phys.*, vol. 101 [No. 5], pp. 509–520, 1936.

[3] KLEEN, W., and H. ROTHE, The Space Charge Equations for Electrons with Initial Velocity, Part II, *Zeit. für Phys.*, vol. 104 [Nos. 11, 12], pp. 711–723, 1937.

[4] SALZBERG, B., and A. V. HAEFF, Effects of Space Charge in the Grid-anode Region of Vacuum Tubes, *RCA Rev.*, vol. 2, pp. 336–374, January, 1938. Excellent discussion of dynamic characteristics.

[5] FAY, C. E., A. L. SAMUEL, and W. SHOCKLEY, On the Theory of Space Charge between Parallel Plane Electrodes, *Bell Sys. Tech. Jour.*, vol. 17, pp. 49–79, January, 1938.

to the screen. There is thus a virtual cathode at the point of zero potential. The potential distribution from the virtual cathode to the negative plate is linear.

The equations for the relations between potential, distance, and

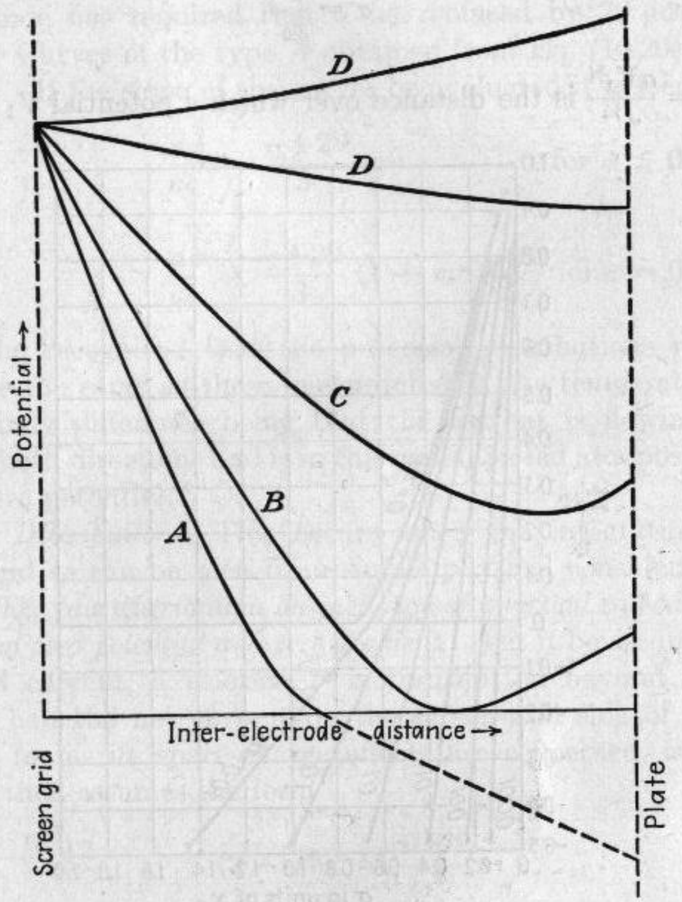


FIG. 10.7.—Types of potential distributions in the screen-grid–plate region of a beam-power tube.

current are obtained by letting the constant in Eq. (10.12) assume the positive value of $m^{1/2}V_1^{1/2}$, where m is related to the slope of the potential-distribution curve as will be shown and V_1 is the screen potential. The differential equation then has the form

$$\left(\frac{dV}{dx}\right)^2 = \frac{16(2J)}{9a^2}(V^{1/2} + m^{1/2}V_1^{1/2}) \tag{10.13}$$

or

$$\frac{dV}{dx} = \frac{-4 \cdot 2^{1/2}J^{1/2}}{3a}(V^{1/2} + m^{1/2}V_1^{1/2})^{1/2} \tag{10.14}$$

Let the following changes in variables be introduced:

$$\phi = \frac{V}{V_1} \tag{10.15}$$

$$\sigma = \frac{x}{x_0} \tag{10.16}$$

where $x_0 = \frac{aV_1^{3/4}}{J^{1/2}}$ is the distance over which a potential V_1 will produce

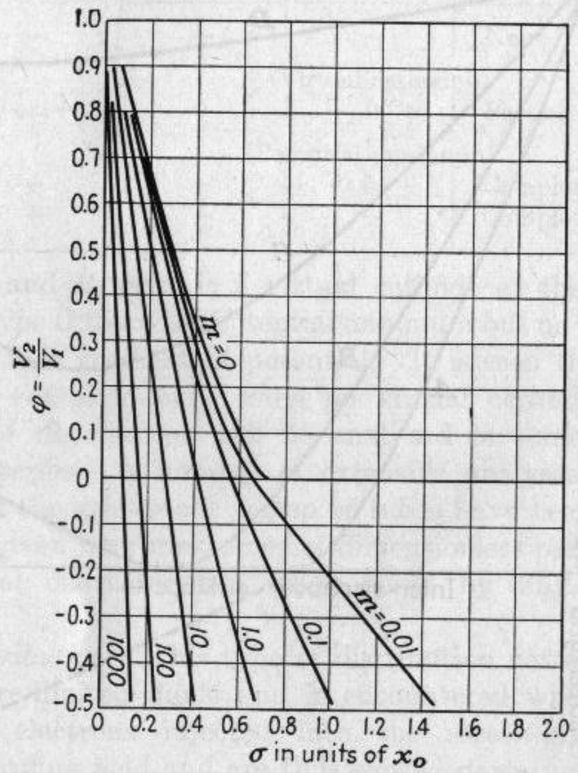


FIG. 10.8.—Potential-distribution curves of the type *A*.

a current density J in a space-charge-saturated diode. The Eq. (10.14) becomes

$$\frac{d\phi}{d\sigma} = \frac{-4 \cdot 2^{1/2}}{3} (\phi^{1/2} + m^{1/2})^{1/2} \tag{10.17}$$

This may be written as

$$\frac{d\phi}{(\phi^{1/2} + m^{1/2})^{1/2}} = \frac{-4 \cdot 2^{1/2}}{3} d\sigma \tag{10.18}$$

for convenience of integration. Let $d\phi = 2\phi^{1/2}\, d(\phi^{1/2})$, and then, upon integration,

$$-(\phi^{1/2} - 2m^{1/2})(\phi^{1/2} + m^{1/2})^{1/2} = 2^{1/2}\sigma + C_2 \tag{10.19}$$

Since $\phi = 1$ when $\sigma = 0$,

$$2^{1/2}\sigma = (1 - 2m^{1/2})(1 + m^{1/2})^{1/2} - (\phi^{1/2} - 2m^{1/2})(\phi^{1/2} + m^{1/2})^{1/2} \tag{10.20}$$

The factor $2^{1/2}$ has been introduced because the reversal of current at the virtual cathode has required that J be replaced by $2J$ in the above derivation. Curves of the type *A* obtained from Eq. (10.20) are shown in Fig. 10.8. If the slope of the curves be evaluated it is seen that

$$\frac{d\phi}{d\sigma} = \frac{-4 \cdot 2^{1/2}}{3} m^{1/4} \qquad \text{for } \phi \leqq 0 \tag{10.21}$$

and

$$\frac{d\phi}{d\sigma} = \frac{-4 \cdot 2^{1/2}}{3} (1 + m^{1/2})^{1/2} \qquad \text{for } \sigma = 0 \tag{10.22}$$

It will be recognized that the potential distributions resulting in this case are the same as those encountered in the temperature-limited diode, the only difference being that the current is flowing in equal amounts in both directions and is in this case injected at a positive rather than at a zero potential.

Type B Distribution. This occurs when the integration constant C_1 is zero and as can be seen from simple physical considerations *gives rise to a Child's law distribution on each side of a virtual cathode that exists at the point of zero potential and zero gradient.* Let it be assumed that, of the injected current, a fraction P is transmitted beyond the virtual cathode. Then the net current on the screen-grid side of the virtual cathode, as far as its space-charge effects are concerned, is $(2 - P)J$. Child's law then assumes the form

$$(2 - P)J = \frac{a^2 V^{3/2}}{{x_1}^2} \tag{10.23}$$

but since $J = \frac{a^2 {V_1}^{3/2}}{{x_0}^2}$, then in terms of the factors ϕ and σ

$${\sigma_1}^2 = \frac{\phi^{3/2}}{(2 - P)} \tag{10.24}$$

Since the actual potential factor is $\phi = 1$ when σ is zero and the potential decreases with increasing σ, the relation must be put into the form

$$\sigma_L = \frac{1 - \phi^{3/4}}{(2 - P)^{1/2}} \tag{10.25}$$

in which the subscript L indicates that the relation holds to the left of the virtual cathode for values of ϕ between 0 and 1 and σ is measured from the point of current injection.

To the right of the virtual cathode the current density is PJ so that

$$PJ = \frac{a^2 V^{3/2}}{x^2} \tag{10.26}$$

where x is now measured from the virtual cathode. Combining this with the Child's law relation,

$$\sigma^2 = \frac{\phi^{3/2}}{P^{1/2}} \tag{10.27}$$

Actually, the potential factor is zero when σ has a value of $\frac{1}{(2-P)^{1/2}}$ as may be seen from Eq. (10.25), so that the desired relation is

$$\sigma_R = \frac{1}{(2-P)^{1/2}} + \frac{\phi^{3/4}}{P^{1/2}} \tag{10.28}$$

where the subscript R indicates that the expression holds only to the right of the virtual cathode and σ is again measured from the point of current injection. Curves of the type B as determined from Eqs. (10.25) and (10.26) are shown in Fig. 10.9. It will be recognized that these are all three-halves-power-law curves drawn with different scales from both sides of the virtual cathode.

A curve of considerable importance in the family (Fig. 10.9) is the limiting curve that gives the maximum value of ϕ for a fixed value of σ to the right of the virtual cathode. This is an envelope to the family of type B curves. If the expression of Eq. (10.28) be solved for ϕ and maximized with respect to P, there results

$$\sigma = \frac{2}{(2-P)^{3/2}} \tag{10.29}$$

When this is substituted in Eq. (10.28), there is obtained

$$P = \frac{2\phi^{1/2}}{1+\phi^{1/2}} \tag{10.30}$$

The factor P can be eliminated between Eqs. (10.29) and (10.30) to give the relation between ϕ and σ.

$$\sigma = (1+\phi^{1/2})^{3/2} 2^{-1/2} \tag{10.31}$$

This curve is shown dotted in Fig. 10.9. Equation (10.30) tells what the maximum transmitted current for any plate potential is. If the attempt is made to increase the transmitted current beyond this value, the distribution will jump from a type B to a type C or D distribution.

Type C Distribution. This type of distribution is characterized by the existence of a potential minimum that is not at zero potential. The distributions are obtained by letting $C_1 = -(\alpha V_1)^{1/2}$ in Eq. (10.12).

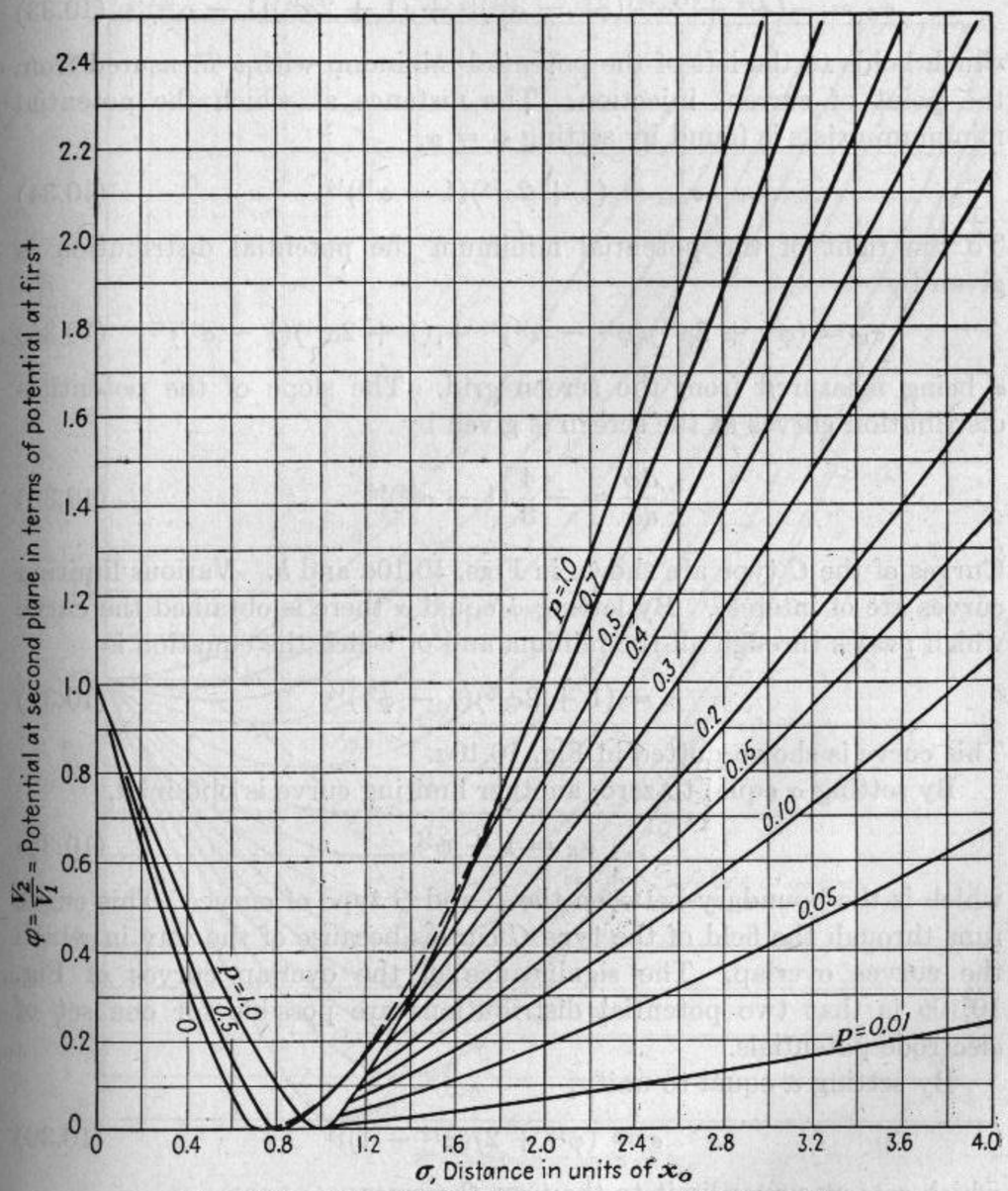


FIG. 10.9.—Potential-distribution curves of the type B.

This gives a positive value of V, equal to αV_1, when $\frac{dV}{dx}$ equals zero. Integration of Eq. (10.12) with the above value of the constant gives

$$\sigma = \pm(\phi^{1/2} + 2\alpha^{1/2})(\phi^{1/2} - \alpha^{1/2})^{1/2} + C_2 \tag{10.32}$$

in which the negative sign goes with a negative slope at the screen and the positive sign gives a distribution of the type D. Since ϕ is unity when σ is zero,

$$\sigma_L = -(\phi^{1/2} + 2\alpha^{1/2})(\phi^{1/2} - \alpha^{1/2})^{1/2} + (1 + 2\alpha^{1/2})(1 - \alpha^{1/2})^{1/2} \quad (10.33)$$

which holds to the left of the potential minimum with σ measured from the point of current injection. The distance at which the potential minimum exists is found by setting $\phi = \alpha$.

$$\sigma_{\min} = (1 + 2\alpha^{1/2})(1 - \alpha^{1/2})^{1/2} \quad (10.34)$$

To the right of the potential minimum the potential distribution is given by

$$\sigma_R = (\phi^{1/2} + 2\alpha^{1/2})(\phi^{1/2} - \alpha^{1/2})^{1/2} + (1 + 2\alpha^{1/2})(1 - \alpha^{1/2})^{1/2} \quad (10.35)$$

σ being measured from the screen grid. The slope of the potential-distribution curves at the screen is given by

$$\frac{d\phi}{d\sigma} = -\frac{4}{3}(1 - \alpha^{1/2})^{1/2} \quad (10.36)$$

Curves of the C type are shown in Figs. 10.10a and b. Various limiting curves are of interest. By letting ϕ equal α there is obtained the curve which passes through all the minima and of which the equation is

$$\sigma = (1 + 2\phi^{1/2})(1 - \phi^{1/2})^{1/2} \quad (10.37)$$

This curve is shown dotted in Fig. 10.10a.

By setting α equal to zero, another limiting curve is obtained,

$$\sigma_R = 1 + \phi^{3/4} \quad (10.38)$$

which is the boundary between the B and C type of curve. This curve runs through the field of the type C curves because of the way in which the curves overlap. The significance of the overlap curves of Fig. 10.10b is that two potential distributions are possible for one set of electrode potentials.

By setting α equal to unity,

$$\sigma = (\phi^{1/2} + 2)(\phi^{1/2} - 1)^{1/2} \quad (10.39)$$

which sets an upper limit to the type C curves.

Another limiting curve is obtained by making σ_R a maximum with respect to α and holding ϕ constant. This gives

$$\alpha = \phi(1 + \phi^{1/2})^{-2} \quad (10.40)$$

and

$$\sigma = (1 + \phi^{1/2})^{3/2} \quad (10.41a)$$

In these expressions σ and ϕ are coordinates of a lower-limit envelope that is tangent to the type C curves. The parameter α determines which curve is tangent to the envelope at the point in question. Type C dis-

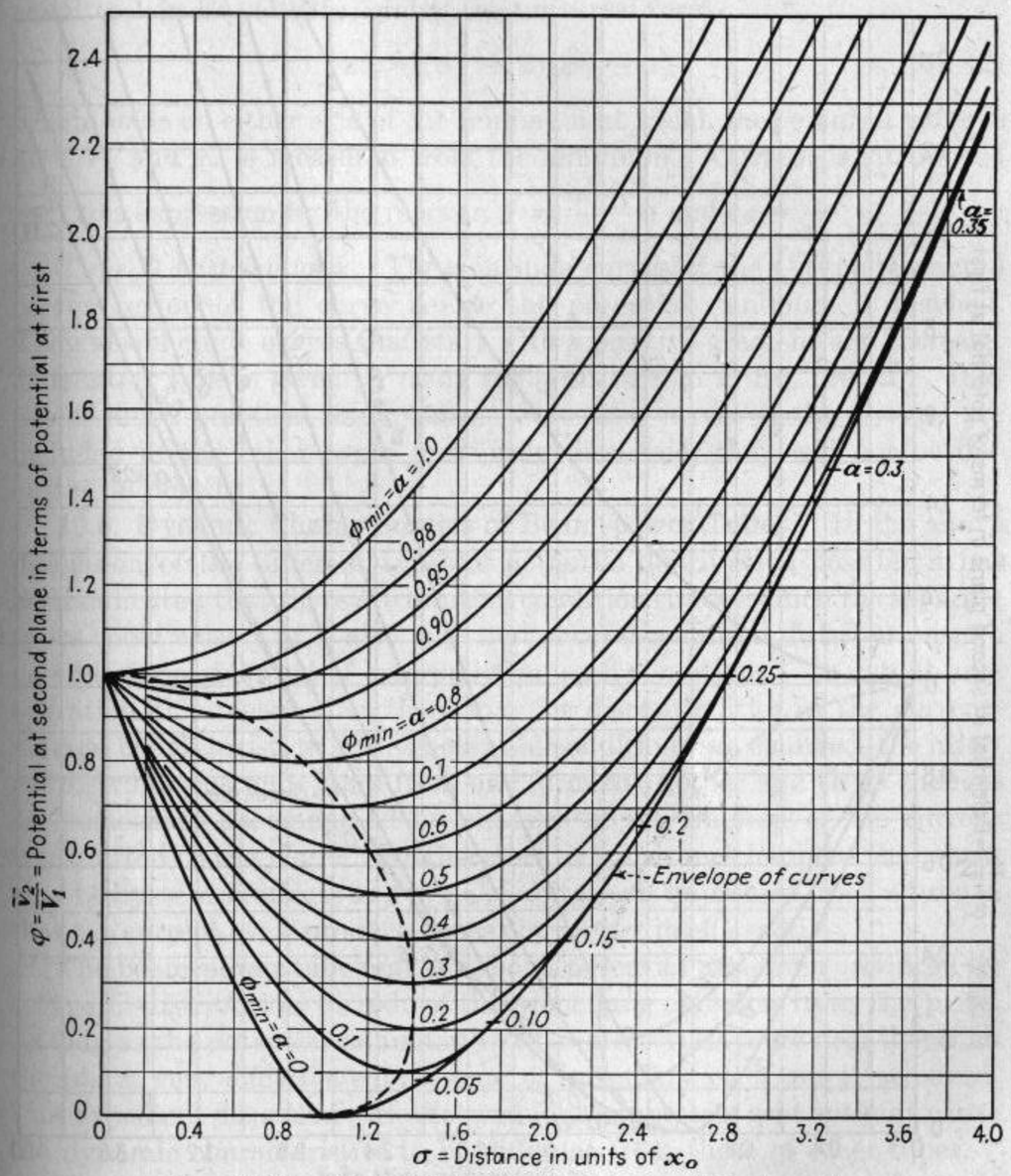


FIG. 10.10a.—Potential-distribution curves of the type C.

tributions cannot exist beyond this condition. If current or voltage is changed beyond this boundary, the distribution jumps to a type B curve.

From Eq. (10.41a) is obtained the expression that gives the maximum current that can be transmitted between electrodes at potentials V_1

and V_2 separated by a distance x. The limiting current density is

$$J_{max} = \frac{2.335 \times 10^{-6}(V_1^{1/2} + V_2^{1/2})^3}{x^2} \quad \text{amperes per unit area} \quad (10.41b)$$

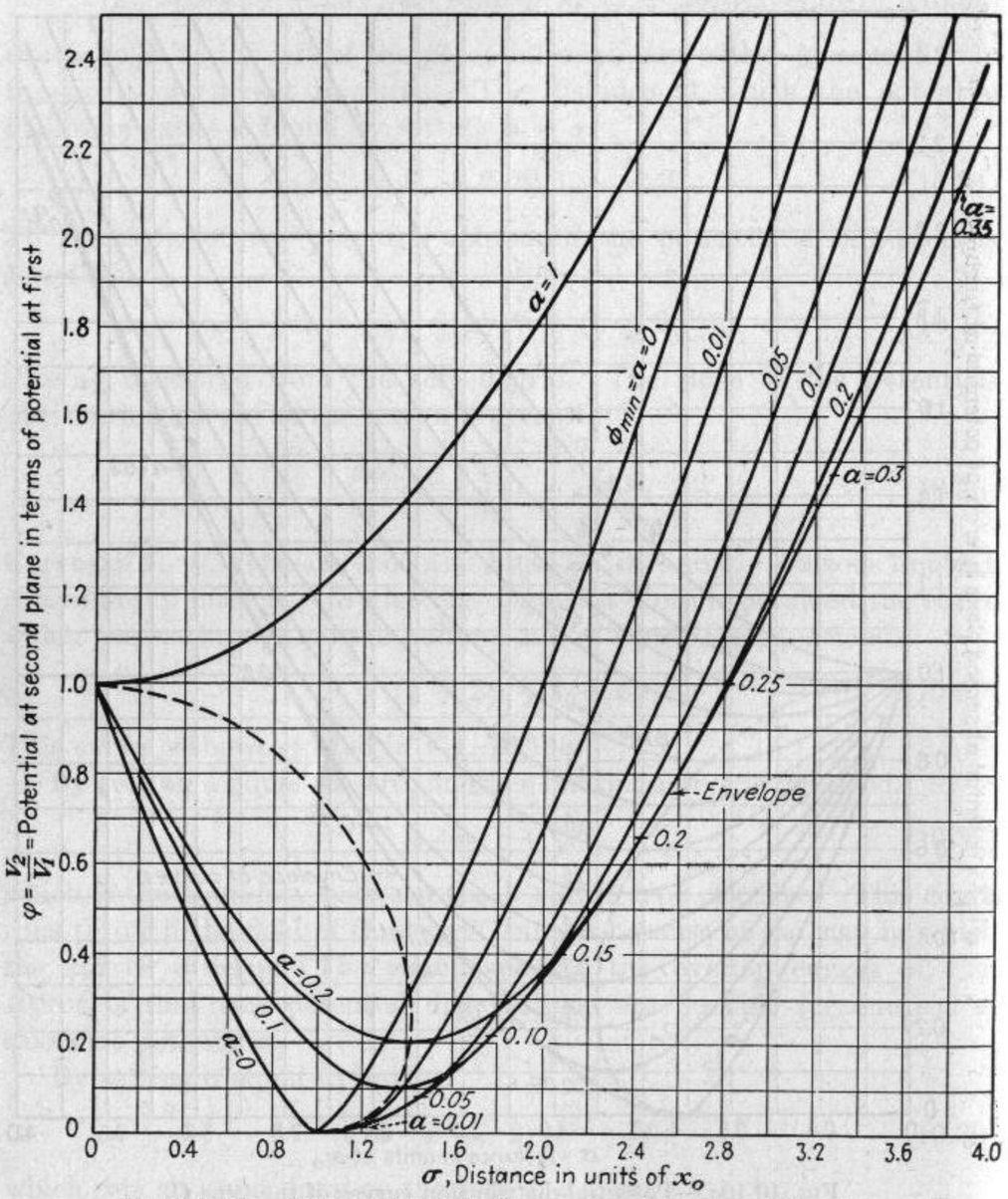


FIG. 10.10*b*.—Potential-distribution curves of the type *C* overlap.

If the injected current exceeds this amount, the potential distribution will jump from a type *C* to a type *B* distribution, with an attendant reduction in transmitted current.

Examination of the curves of Fig. 10.10 and their equations shows

that the curves are all of the same form but differ only in scale and position. Hence the upper limiting curve for which α equals 1 is a universal curve showing how the potential varies on either side of the potential minimum. The universal form of this curve is obtained by setting α equal to 1 in Eq. (10.35), giving the universal form

$$\sigma_u = (\phi^{1/2} + 2)(\phi^{1/2} - 1)^{1/2} \quad (10.42)$$

which holds on either side of the minimum at which the potential value is now V_1 and σ_u is measured from the minimum. Current is introduced into this expression by the relation $J = \frac{a^2 V_1^{3/2}}{x_{u^2}}$, and $\sigma_u = \frac{x}{x_u}$.

Type D Distributions. These include curves of the *C* type restricted to the region of the curve before the potential minimum is reached. They also include curves that start with a positive gradient and increase. This latter type is given by using the positive sign in Eq. (10.32). It is not of much practical importance. Since the curves of the *D* type are included in the other types previously discussed, they will not be discussed in detail.

10.6. Dynamic Characteristics of Beam-power Tubes. In the above discussion of the different types of potential distribution possible it has been indicated that there are limiting conditions under which the separate types could exist. It is also true that several potential distributions are possible for a given set of externally imposed conditions. In actual tube operation this means that there may be discontinuities in the current-voltage characteristics; for as potential conditions are changed, the internal distributions may jump from one form to another and these changes are sometimes accompanied by changes in the fraction of the current transmitted to the plate. Furthermore, it sometimes happens that there may be hysteresis effects as a cycle of voltages is impressed upon a tube in that the current cycle produced does not retrace itself exactly.

The beam-power tube makes use of a potential minimum produced by a type *C* distribution to reduce the secondary emission from the plate. As long as the potential minimum is 20 or more volts more negative than the plate, very effective suppression of secondary electrons is achieved. This expedient dispenses with the need for a suppressor grid but may cause the dynamic characteristics to be different from those of other tubes.

Two of the most important dynamic characteristics will be discussed in a qualitative fashion. Quantitative analyses have been given,[1] but these are somewhat limited in value in that the ideal conditions of parallel electron flow cannot be realized exactly in any actual tube.

Injected Current Varied, Potentials Constant. One case that is of

[1] SALZBERG and HAEFF, *loc. cit.*

considerable interest is that in which the screen and plate of a beam-power tube are maintained at the *same* potential and the current injected into the screen-grid–anode region is increased from zero to a large value and then decreased. The changes in the potential distribution encountered in the tube are shown in Fig. 10.11. Initially, for no injected current, the potential distribution from screen grid to plate is a straight horizontal line as shown at *a*. As the injected current is increased, the potential-distribution curve is depressed, assuming the form of the type *C* distributions as shown at *b*, in this case symmetrical with respect to a potential minimum at the center. As the injected current is further

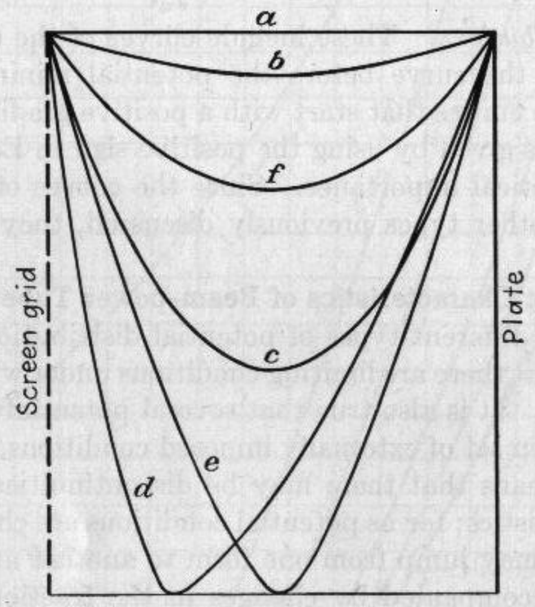


FIG. 10.11.—Potential distributions in a beam-power tube as the injected current is varied.

increased, the potential-distribution curve is depressed still further, maintaining its symmetry. The physical reasons for the action are quite apparent. As the current is increased, the space charge is increased, which reduces the potential, which decreases the electron velocity, which increases the space charge still further and thus depresses the potential still more. Thus an increase in injected current starts a cycle of action that is very sensitive to changes in current, so much so that an equilibrium may not always be reached. This occurs *in this case* when the potential curve has been moved about three-quarters of the way down to a zero potential, as at *c*, at which point any further increase in current causes the potential curve to drop as far as it will go because of the instability in the sequence of actions described above. The potential

curve can drop only to zero, at which point a virtual cathode is formed halfway between grid and plate. When this occurs, part of the current, which has previously all been going to the plate, is turned back toward the grid, thus increasing the space charge or effective current to the left of the virtual cathode. Child's law demands that when the current is increased without a change in potential the distance must be decreased so that the virtual cathode moves toward the screen, finally coming to rest at some position, as shown at *d*. Thus the distribution changes immediately from that at *c* to that at *d*, with an abrupt reduction in current to the plate. Any further increase in injected current increases the space charge and current on the screen side of the virtual cathode and causes it to move closer to the screen, with a further reduction in transmitted current.

If now the current is decreased, the sequence of operations will not be exactly the same, for the initial conditions are different. For a given set of voltages and current, two potential distributions may be possible but only one can exist at a time, of course, and the physical choice between the possible distributions is determined by the order in which limiting conditions are established. If the current is decreased, the virtual cathode moves toward the plate, an increasing fraction of the current going to the plate. Finally the virtual cathode reaches the mid-point, all the current going to the plate. The virtual cathode is now "saturated." The potential field and electron paths for such a condition are shown in Fig. 10.12. The type *B* distribution cannot exist with any smaller injected current, and thus a further decrease causes the potential distribution to jump to the type *C* distribution, jumping from the distribution at *e* to that at *f*. Any further decrease in current now maintains the same type of symmetrical distribution, the potential minimum rising until finally it is flat, with no current. Because of the fact that there is a maximum value of plate current for any set of screen and plate voltages in a beam-power tube, difficulties may sometimes be encountered with pulsed operation.

Although the above discussion has been given for *equal screen and plate voltages*, the same sort of behavior results if the electrode potentials are not the same. In general, specific limiting conditions will be different for different cases.

The associated current behavior is shown in Fig. 10.13, which shows the relation between the plate current and the injected current. As the injected current is increased, at first all the current is transmitted to the plate, giving the straight-line characteristic shown. When the potential distribution jumps from *c* to *d*, the plate current suddenly drops in value and then decreases uniformly, as shown, as the injected

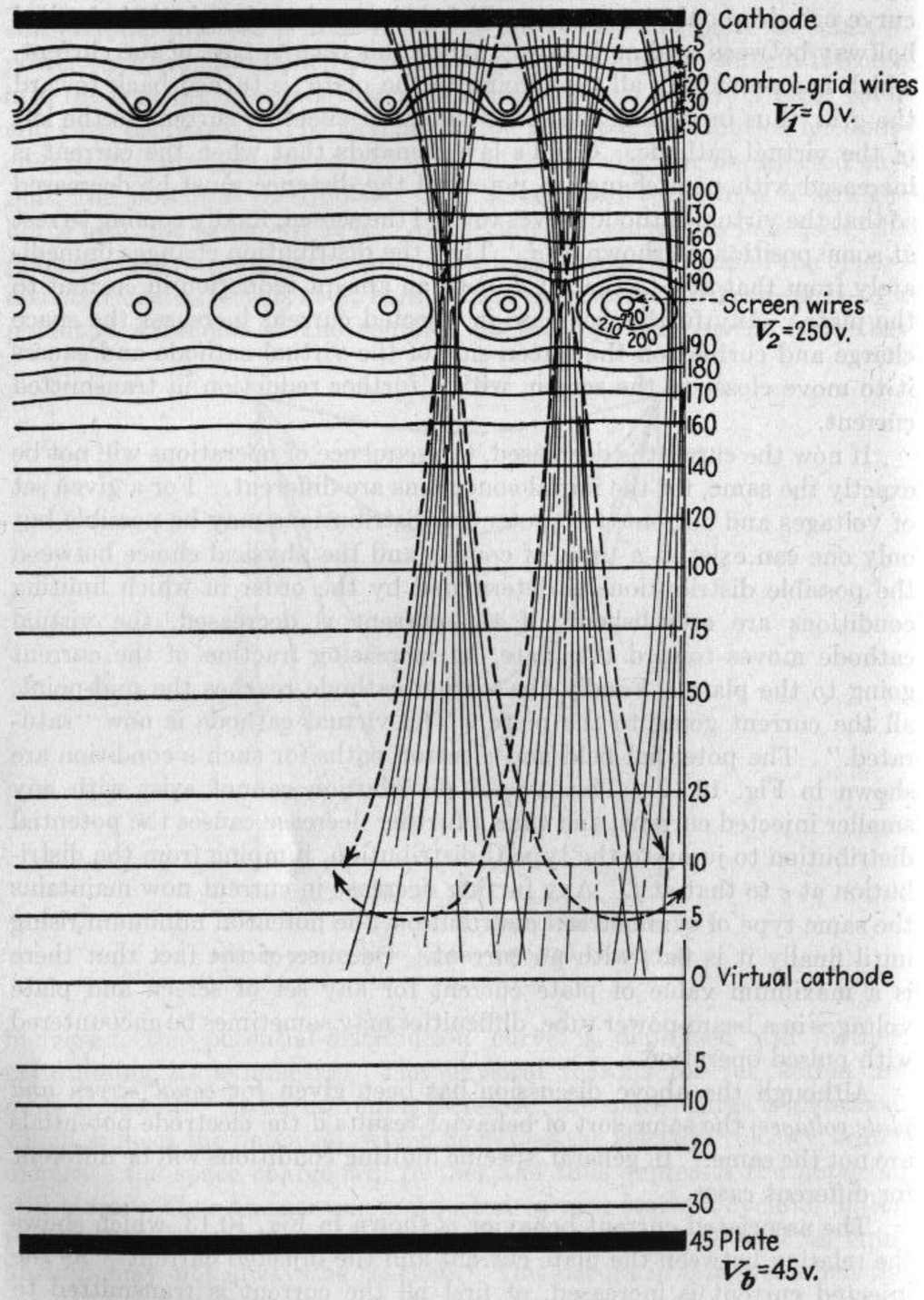


FIG. 10.12.—Potential field of a beam-power tube with a saturated virtual cathode (*Schade.*)

current is increased. When the injected current is then decreased, the transmitted current is increased uniformly until the distribution shown at *e* in Fig. 10.11 is reached. In this case the jump in distribution from *e* to *f* produces no change in current, though it will be observed that the highest current reached on the retrace of the cycle is less than that obtained as the injected current was increased. All the injected current now goes to the plate again; and as the injected current is decreased further, the plate current decreases correspondingly, moving down the straight-line portion of the curve of Fig. 10.13. The portion at the extreme right of the current characteristic is seen to exhibit a negative

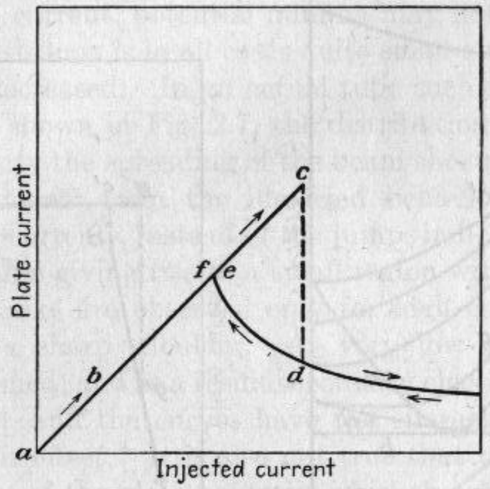


FIG. 10.13.—Transmitted current in a beam-power tube as a function of injected current.

transconductance since the injected current is a continuous function of the control-grid voltage. Oscillators have been built utilizing this property. The characteristic shown in Fig. 10.13 can actually be observed on an oscilloscope if a beam-power tube is connected so that the vertical deflection is proportional to plate current and the horizontal deflection proportional to space current as the control-grid voltage is varied sinusoidally.

Plate Potential Varied, Screen Potential and Injected Current Constant. Another case of operation which is of particular importance is that which occurs when the plate potential alone is varied. Consider the case in which the injected current is quite high, corresponding to a positive control-grid voltage on a beam-power tube. Starting with a negative plate potential, the potential distribution is of the type *A* (temperature-limited), as shown at *a* in Fig. 10.14. When the plate potential reaches a value of zero, a distribution of the type *B* (space-charge-limited) exists

as shown at b. Then, as the potential is further increased, a three-halves-power-law distribution holds on each side of the virtual cathode, which moves toward the plate. Finally, a limiting type B curve is reached, and even though all the current is not being transmitted to the plate the distribution jumps from that at c to that at c', giving a type C (potential minimum) distribution with a complete instead of a partial transmission of current. As the potential is further increased, the potential minimum moves toward the screen as the curve moves up. Then, as the cycle is reversed and the potential is decreased, the curve moves

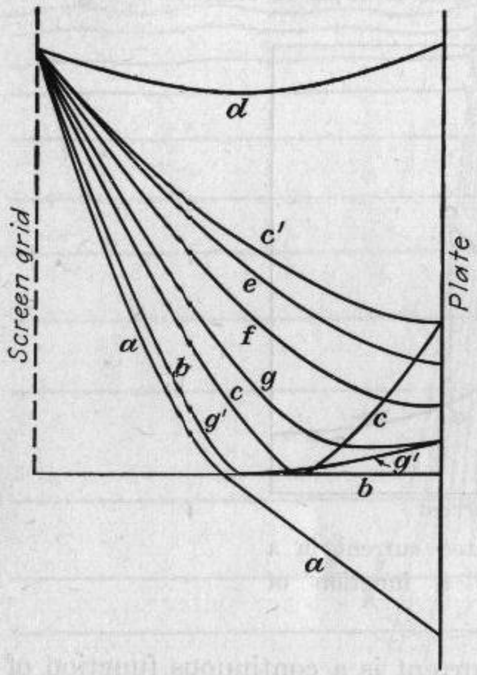


FIG. 10.14.—Potential distributions in a beam-power tube as plate voltage is varied.

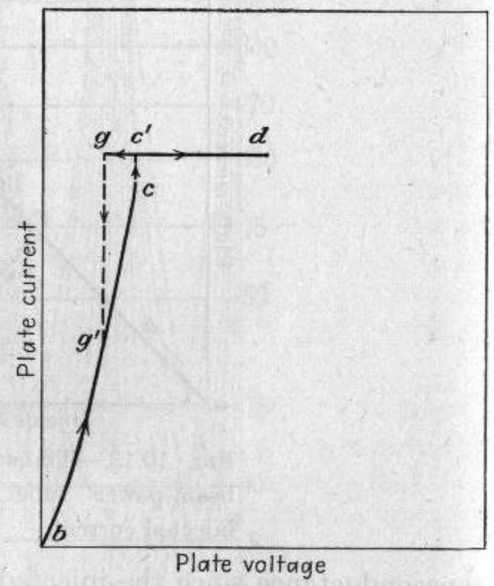


FIG. 10.15.—Plate current in an ideal beam-power tube as a function of plate voltage.

down through the stages indicated by e and f. At g there is reached a limiting curve of the type C, and the distribution jumps to that at g', giving a partial transmission of current. From this the virtual cathode moves toward the screen as the potential and plate current decrease to zero. The curve a is obtained again as the plate voltage is made negative.

The corresponding current behavior is shown in Fig. 10.15, in which is shown the variation of plate current with plate potential. Plate current begins to flow at b and continues to increase until c because of the partial transmission of current. If the position of the virtual cathode remained fixed, the current would increase as the three-halves power of the plate potential in this region. Since it moves toward the plate, the

current increases faster than the three-halves power, giving a very steep characteristic and accounting for the sharp shoulder of the plate-current–plate-voltage curves of the beam-power tube. At c the current jumps to that at c', at which the transmission is complete and the current remains constant for further increases in potential. When the cycle is reversed, the current remains constant as the potential is decreased until g is reached, at which it drops to the value on the curve between b and c. It will be noted that the jump in current on the retrace occurs at a lower value of potential than when the potential is increasing.

The type of behavior described above occurs for other values of current, though the effects are most pronounced when the current is high. At lower values of current, potential minima may not be formed. The area of the hysteresis loop is in all cases quite small and becomes smaller as the current is decreased. In an actual tube such as the 6L6, having the characteristics shown in Fig. 2.7, the distribution of initial velocities and more particularly the spreading of the beam sheets cause the current characteristic to depart from the idealized behavior indicated here.[1] Sometimes at high currents, instead of the jumps indicated, the curve will run through a small s giving rise to a small region with a negative plate resistance. The kinks are observed only for high currents. At lower currents there is a sharp shoulder. At very low currents, potential minima are not formed; and as a result secondary electrons from the plate are not suppressed, and the curves have the characteristic depressions associated with this effect. It is also not true that the current is completely independent of the plate potential when the current transmission is complete, for the cathode region is not completely shielded from the plate.

[1] SCHADE, O. H., Beam Power Tubes, *Proc. I.R.E.*, vol. 26, pp. 137–181, February, 1938.

CHAPTER 11

PENTODES

11.1. Electrode Arrangement in a Pentode. The pentode, as its name implies, is a five-electrode tube. The five electrodes, in order, are cathode, control grid, screen grid, suppressor grid, and plate. The effects of all these electrodes except the suppressor grid have been studied.

The suppressor grid was added to the screen-grid tube to eliminate the exchange of secondary electrons between screen grid and plate. It is invariably a coarse-mesh grid placed between the screen grid and plate and operated at cathode potential. At this potential it is able to suppress secondary electrons by causing a deep dip in the potential between screen grid and plate while at the same time its coarse mesh allows electrons to pass on through it to the plate. The potential profiles of a pentode are shown in Fig. 11.1. From these it is seen that both the plate and the grid present negative gradients of potential to secondary electrons created at their surface. This eliminates the exchange of secondary electrons between screen grid and plate and results in current-voltage characteristics which are almost exactly those which would occur in a perfect screen-grid tube having no secondary emission.

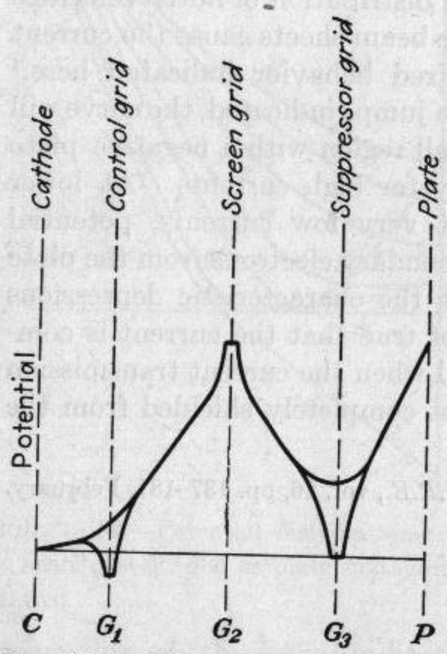


FIG. 11.1.—Potential profiles in a pentode.

Of all the various types of vacuum tube, the pentode is probably the one in most extensive use. For voltage amplification whether at audio or radio frequencies it is the invariable choice. It is used at audio frequencies because a higher gain per stage can be realized than with a triode. It is used at radio frequencies because the extremely low control-grid-to-plate capacity virtually eliminates the possibility of regeneration. Even as a power tube, it finds considerable use because its control-grid current and hence the power necessary to drive it are lower than for the corresponding triode.

11.2. Current-voltage Characteristics of the Pentode. As stated above, the suppression of secondary electrons in the pentode gives it the characteristics of a screen-grid tube that is free of secondary emission. The addition of the extra grid increases the extent to which the control grid is shielded from the plate and results in a somewhat higher amplification factor and a somewhat higher plate resistance.

Plate-current–Control-grid Voltage Characteristics. The plate-current–control-grid characteristics of a pentode are similar to those of a tetrode and not greatly different from those of a triode. The plate current is

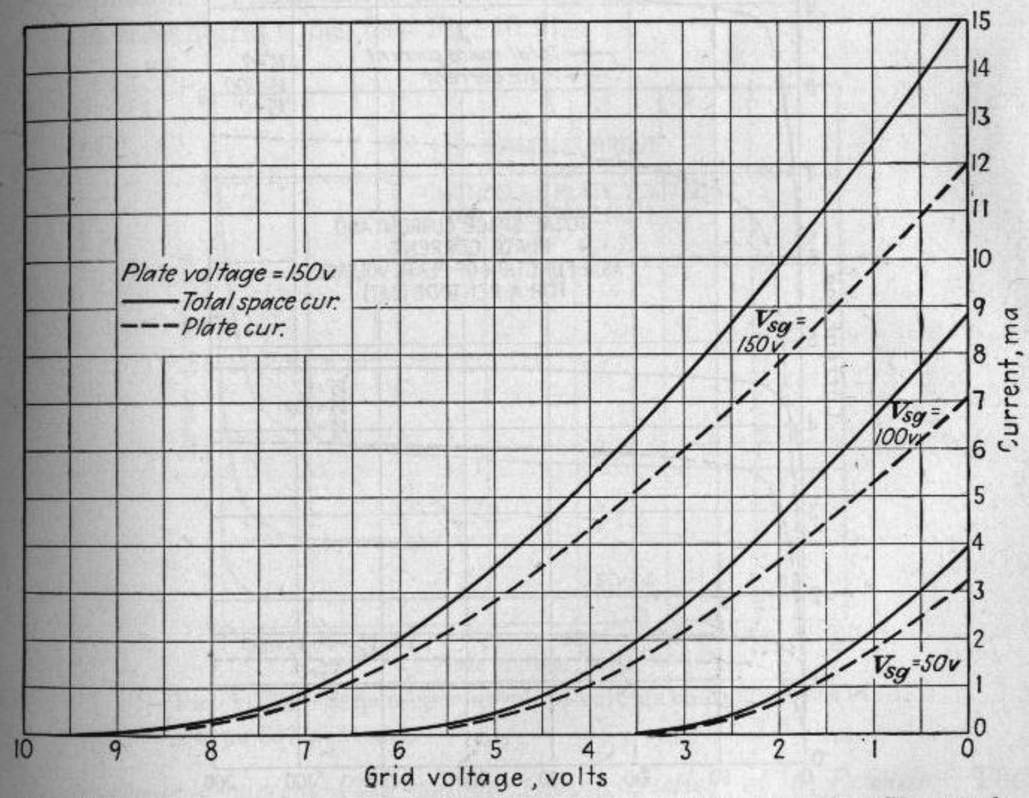


FIG. 11.2.—Plate-current–control-grid voltage characteristics of a type 6J7 pentode.

most easily influenced by the control grid, less so by the screen grid, and hardly at all by the plate. Some typical plate-current–control-grid voltage characteristics of a pentode are shown in Fig. 11.2. Here there is shown a group of curves for different screen-grid voltages. These curves are almost identical with the corresponding curves in a triode. If a group of curves for different plate voltages were shown, they would be very closely grouped and, for the same screen-grid potential, would have the same cutoff potential.

Plate-current–Plate-voltage Characteristics of a Pentode. The plate-current–plate-voltage characteristics of a pentode are shown in Fig.

11.3*a*. In a well-designed pentode the plate current varies only slightly with plate potential for all plate voltages greater than 50 per cent of the screen-grid voltage. Below this value of voltage the current falls rapidly to zero. The magitude of the space current in a pentode is determined almost entirely by the control-grid and screen-grid potentials. The plate potential determines only what fraction of the space current is

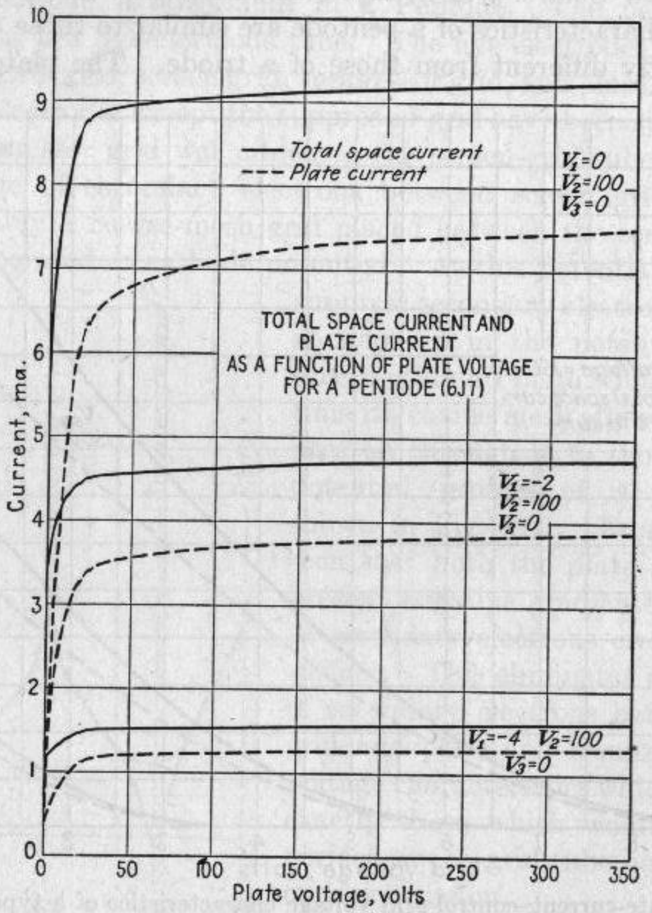


FIG. 11.3*a*.—Plate-current–plate-voltage characteristics of a type 6J7 pentode.

transmitted to the plate. It does, of course, have a second-order influence upon the plate current, with the result that the plate current rises slowly as the plate voltage is increased, but this rise is even slower than in the screen-grid tube, where there is only one shielding grid between the plate and the control grid. As a result of the action of the plate voltage in determining the fraction of the space current that is transmitted to the plate, all the plate-current–plate-voltage curves are similar in shape and differ only in scale.

Space-current–Plate-voltage Characteristics of the Pentode. Also shown in Fig. 11.3*a* are the space-current–plate-voltage characteristics of a pentode. The space current is even more constant with plate voltage than is the plate current. The only departure from near constancy occurs near zero plate potential. Here the space current increases by about 20 to 40 per cent as the plate potential is increased to about half of the screen-grid potential. This increase in space current occurs because there is a change in the space-charge condition around the screen grid as the condition of reflection of electrons from the plate changes to one of transmission. The action is exactly the same as that which was encountered in screen-grid tubes (see Fig. 10.2).

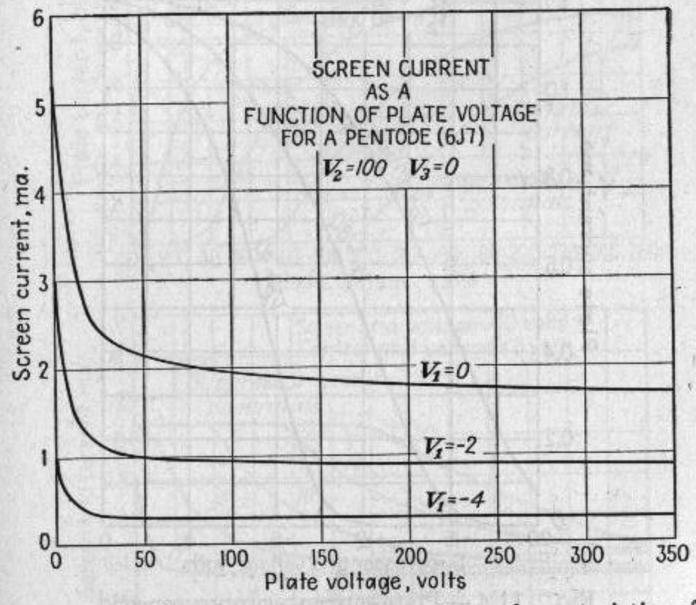


FIG. 11.3*b*.—Screen-current–plate-voltage characteristics of a type 6J7 pentode.

Screen-grid-current–Plate-voltage Characteristics of a Pentode. The screen-grid current in a pentode is the difference between the space current and the plate current, provided that the other grids in the tube are drawing no current. The nature of the screen-grid-current variation with plate voltage is shown in Fig. 11.3*b*. The screen-grid current is seen to have a uniformly decreasing characteristic with plate voltage. The screen current will generally lie between one-fifth and one-third of the plate current at large plate voltages. The effectiveness of the suppressor grid in suppressing secondary electrons is so complete that the screen current rarely shows even a trace of distortion due to this cause.

Suppressor-grid Effects. With small pentodes such as are used for voltage amplification the suppressor grid is operated at cathode potential

and is not used to influence the plate current. In power-output pentodes, however, the suppressor grid may be used as an active electrode. It can be used to modulate the plate current in amplifiers or oscillators.[1] Use is here made of the fact that the suppressor grid is able to control the fraction of the current transmitted past the plane of the screen grid that goes on to the plate. When the suppressor grid is at a low potential relative to plate and screen grid, as it usually is, it can sort out the electrons having a large component of energy directed toward the plate from those which, because of deflection on passing close to a screen-grid wire, have a lower component of plate-directed energy. Some typical

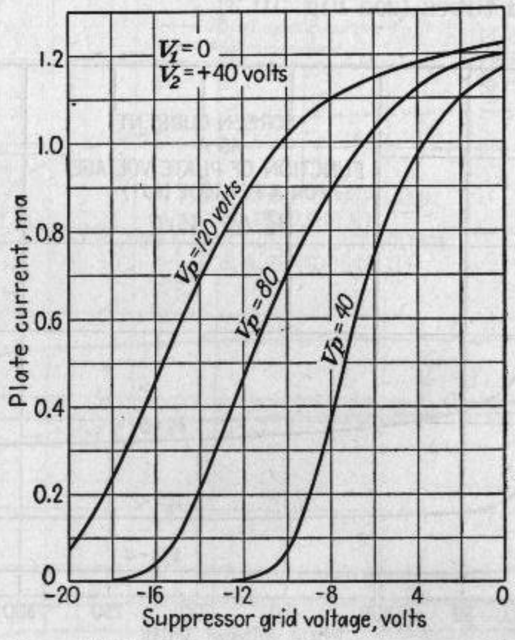


FIG. 11.4.—Plate-current–suppressor-grid characteristics of a type 6J7 pentode.

curves showing the current transmitted to the plate as a function of suppressor-grid voltage are shown in Fig. 11.4. It is seen that the plate current is moderately sensitive to suppressor-grid voltage and that the suppressor grid is readily capable of completely cutting off the plate current. The action of the suppressor grid in controlling the plate current is a combination of its action as a velocity sorter and a direct control on the gradient of potential before the suppressor grid. At voltages near zero the first action predominates. As the suppressor grid is made more negative, the second action becomes predominant. As the suppressor grid approaches cutoff, there is a strong tendency for a virtual

[1] GREEN, C. B., Suppressor Grid Modulation, *Bell Lab. Rec.*, vol. 17, pp. 41–44, October, 1938.

cathode to form before the suppressor grid owing to the space charge of the approaching electrons, which have been reduced to a very low velocity. The suppressor grid then has an action very similar to that of the control grid in front of an actual cathode, as in a triode.

An alternative representation of the effect of the suppressor grid is shown in Fig. 11.5. Here the control-grid action of the suppressor grid is evident at large negative values of suppressor-grid potential. Power

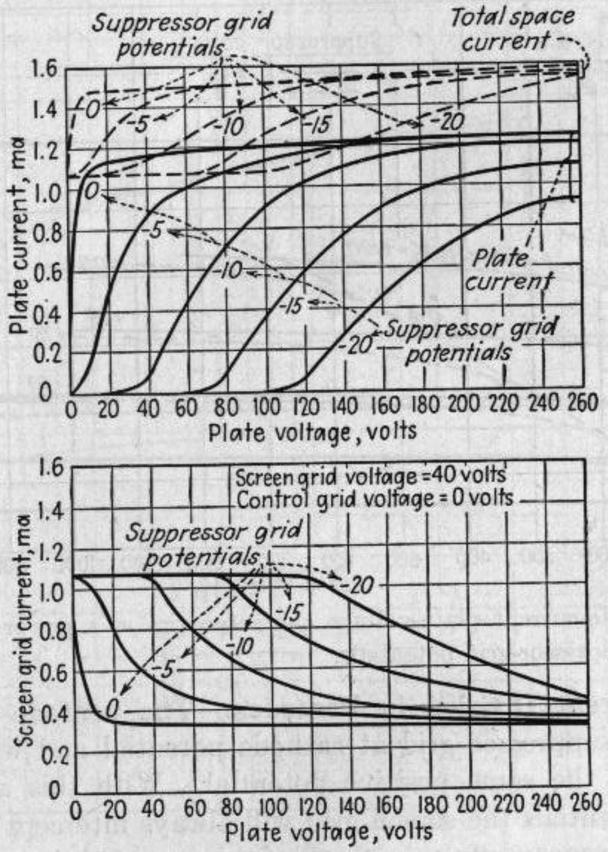


FIG. 11.5.—Effect of suppressor-grid voltage upon the plate and screen-grid-current characteristics of a pentode.

pentodes are frequently operated with slightly positive suppressor grids in order to get a sharper shoulder on the plate-current–plate-voltage characteristics. Some typical power-pentode characteristics are shown in Fig. 11.6. The use of a positive suppressor-grid potential is seen to give a considerable sharpness to the shoulder of the characteristics. The reason why this is necessary in the power tubes is twofold. (1) The current densities involved are greater, increasing the tendency for a virtual cathode to form in front of the suppressor grid and hence requiring a more positive value of suppressor-grid voltage to pass the major

portion of the space current on to the plate. (2) The screen-grid pitch is made relatively large to reduce the total current and power taken by the screen. As this is done, the deflection imparted to the electrons passing through the screen grid is increased and hence a smaller fraction of them have enough plate-directed velocity to pass through the positive potential spaces between the suppressor-grid wires.

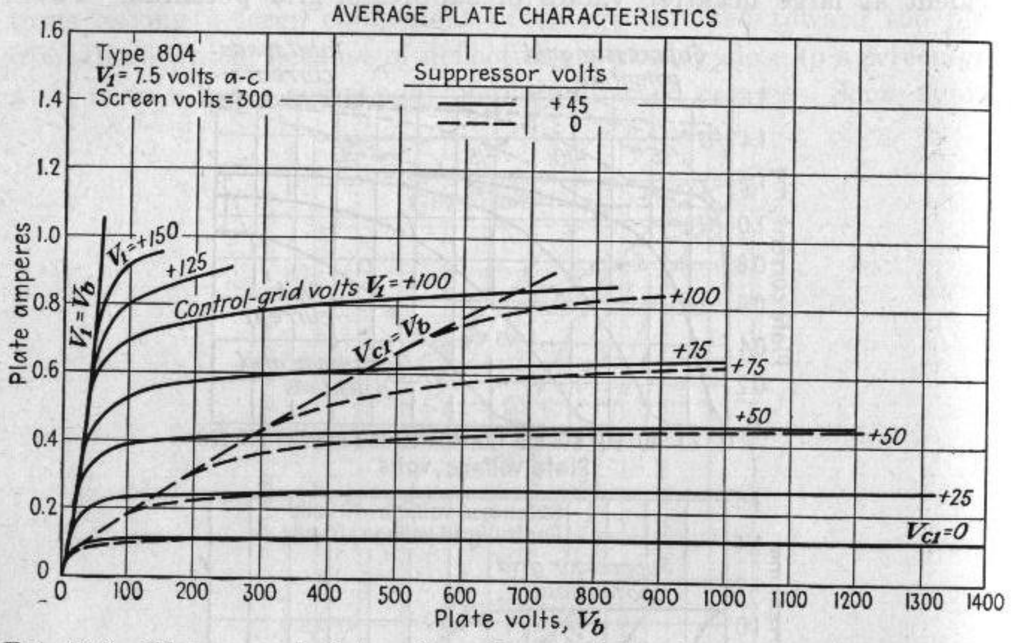


FIG. 11.6.—Plate-current–plate-voltage characteristics of a power pentode for zero and positive suppressor-grid potentials.

11.3. Current Division in Pentodes. The pentode is often operated with a suppressor grid at cathode potential and with screen grid and plate at the same positive potential. With this arrangement of electrode potentials the screen grid will always intercept an appreciable fraction of the space current. Some of this current is intercepted directly on the first passage of the electrons. Some of the screen-grid current consists of electrons which were so strongly deflected by the screen-grid wires that they did not have enough plate-directed velocity to pass between the suppressor-grid wires. *All* electrons that fail to penetrate the suppressor grid upon their first attempt may be expected to return to the screen grid.

The fraction of the total current transmitted to the plate is expected to be a function of the ratio of plate to screen-grid voltages. So also is the ratio of plate to screen-grid current. In Fig. 11.7 are shown curves of these current ratios as a function of the electrode-voltage ratio. Both current ratios are seen to vary rather slowly with the electrode-voltage

ratio. The ratio of plate to space current in the vicinity of equal plate and screen-grid voltages varies something like the one-tenth power of the ratio of plate to screen-grid voltage. The ratio of plate to screen current varies approximately as the one-fifth power of the ratio of plate to screen-grid voltage. No simple theoretical analysis is available to give the current-division law directly in either case.

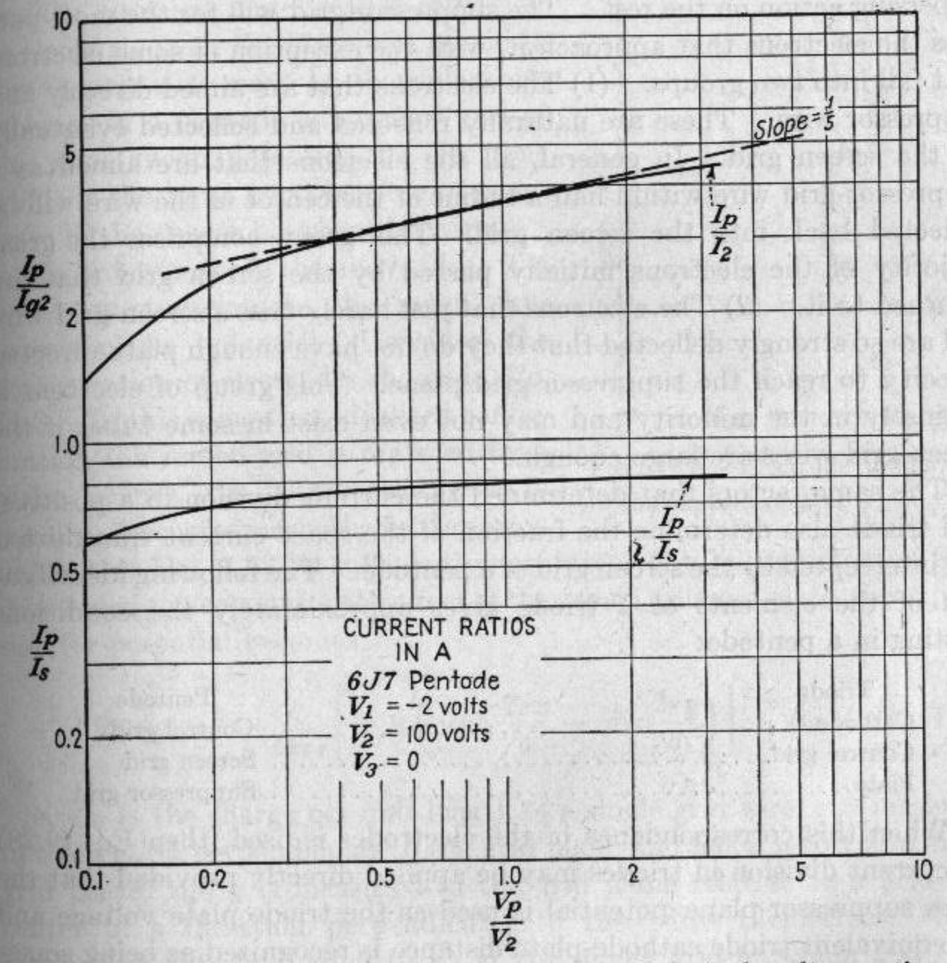


FIG. 11.7.—Pentode current ratios as a function of electrode-voltage ratios.

Just as it was possible to define a current-division factor for positive-grid triodes, so is it possible to define a current-divison factor for pentodes. Let

$$\frac{I_p}{I_2} = \mathfrak{D}\left(\frac{V_p}{V_2}\right)^m \tag{11.1}$$

Here $\mathfrak{D}$ is a current-division factor that measures the ratio of plate to screen-grid current when the plate and screen grid have the same voltage, other grids being presumed to draw no current. This pentode current-

division factor may be determined quite closely from the considerations that applied to the positive-grid triode.

Figure 11.11 shows the general nature of the potential field and electron paths within a pentode. The control grid has a slight focusing action upon the electrons that pass between its wires. The screen grid intercepts a fraction of the current that passes through its plane and has a dispersing action on the rest. The suppressor grid will for the most part pass the electrons that approach it, with the exception of some electrons that fall into two groups. (1) The electrons that are aimed directly at a suppressor wire. These are naturally reflected and collected eventually by the screen grid. In general, all the electrons that are aimed at a suppressor-grid wire within half a radius of the center of the wire will be deflected back into the screen grid. This group comprises the great majority of the electrons initially passed by the screen grid that are returned to it. (2) The electrons that just barely miss a screen-grid wire and are so strongly deflected that they do not have enough plate-directed velocity to reach the suppressor-grid plane. This group of electrons is distinctly in the minority and may not even exist in some tubes if the screen-grid wires are large enough.

The same factors that determined the current division in a positive-grid triode also determine the fraction of the space current transmitted and intercepted by the screen grid of a pentode. The following identification of the elements of a triode gives approximately the conditions existing in a pentode:

Triode	Pentode
Cathode	Control grid
Control grid	Screen grid
Plate	Suppressor grid

When this correspondence of the electrodes is used, then Eq. (9.28) for current division in triodes may be applied directly provided that the mean suppressor-plane potential is used as the triode plate voltage and the equivalent triode cathode-plate distance is recognized as being somewhat larger than the pentode control-grid–screen-grid distance. When these considerations are applied, a transmission factor for the screen grid is determined. It is then necessary only to correct this for the additional electrons reflected from elastic collisions with the suppressor-grid wires.

For the condition of equal plate and screen-grid voltages prescribed for the pentode current-division factor, the mean suppressor-plane potential is a very small fraction of the plate and screen-grid potential. In terms of the equivalent-triode current division this means that interest is centered on the region of very small voltage ratios, far to the left on the

true primary-current-division curve of Fig. 9.18. The value of the current ratio will not be a great deal less than the triode current-division factor because the current-ratio curve is concave upward. The true primary-division curve rather than the measured triode division law may be expected to apply in the case of the pentode because there is no modification of the electrode currents by secondary emission and because reflection of electrons from the plate of a triode at low potentials may be considerable, whereas in a pentode suppressor grid of the same mean potential the electrons will, for the most part, be able to penetrate in the spaces between the wires.

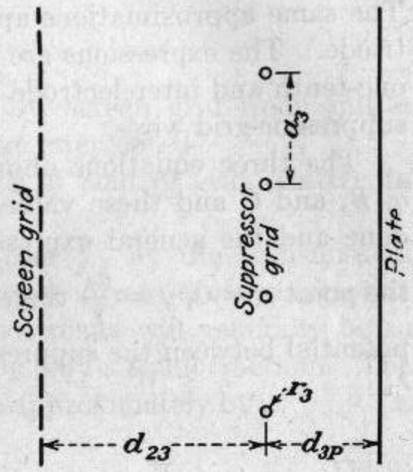


FIG. 11.8.—Electrode dimensions in the suppressor-grid region.

In order to apply the above ideas it is necessary to evaluate the mean suppressor-plane potential. Consider the configuration of electrodes shown in Fig. 11.8. In this somewhat idealized configuration the screen grid is replaced by a plane. Since the suppressor grid is usually of a coarse mesh, the field approximation of Eq. (7.72) is sufficiently accurate. If a linear potential term is included to account for the effect of the screen grid and plate, the expression for potential becomes

$$V(x,y) = -\frac{q}{4\pi\varepsilon_0}\left[\ln 2\left(\cosh\frac{2\pi x}{a} - \cos\frac{2\pi y}{a}\right)\right] + Bx + C \quad (11.2)$$

where q is the charge per unit length of a single grid wire, x is measured in a direction perpendicular to the grid plane and has a zero value in the grid plane, and y is measured in the grid plane relative to a grid-wire center in a direction perpendicular to the wires themselves. Upon setting the potential equal to V_2 at the screen grid at which $x = -d_{23}$ and $y = 0$, then approximately

$$V_2 = -\frac{q}{2\varepsilon_0}\frac{d_{23}}{a_3} - Bd_{23} + C \quad (11.3)$$

Upon setting the potential equal to zero at the suppressor at which $x = 0$ and $y = r_3$, then approximately

$$0 = -\frac{q}{2\pi\varepsilon_0}\ln\left[2\sin\left(\frac{\pi r_3}{a_3}\right)\right] + C \quad (11.4)$$

Upon setting the plate potential also equal to V_2 at $x = d_{3p}$, $y = 0$, there results, approximately,

$$V_2 = -\frac{q}{2\epsilon_0}\frac{d_{3p}}{a_3} + Bd_{3p} + C \tag{11.5}$$

The same approximations apply above as do in the case of the low-mu triode. The expressions are valid only for screening fractions less than one-tenth and interelectrode spacings greater than the distance between suppressor-grid wires.

The three equations above may be solved for the three unknowns q, B, and C and these values substituted in Eq. (11.2). When this is done and the general expression for potential resulting is restricted to the point $x = 0$, $y = \frac{a_3}{2}$, there is obtained an expression for the maximum potential between the suppressor-grid wires,

$$V_{3\,\max} = \frac{V_2 \ln\left[\sin\left(\frac{\pi r_3}{a_3}\right)\right]}{\ln 2\left[\sin\left(\frac{\pi r_3}{a_3}\right)\right] - \frac{2\pi}{a_3}\frac{d_{23}d_{3p}}{d_{2p}}} \tag{11.6}$$

The *mean* value of the suppressor-grid-plane potential is just half this.

$$V_{3\,\mathrm{av}} = \frac{V_2 \ln\left[\sin\left(\frac{\pi r_3}{a_3}\right)\right]}{2\left\{\ln\left[2\sin\left(\frac{\pi r_3}{a_3}\right)\right] - \frac{2\pi}{a_3}\frac{d_{23}d_{3p}}{d_{2p}}\right\}} \tag{11.7}$$

When this mean value of the suppressor-plane potential is known, Eq. (9.28*a*), which is repeated here, can be applied with the electrode correspondence previously mentioned to obtain the ratio of the current transmitted by the screen grid to the current intercepted on the initial passage.

$$\frac{I_p}{I_g} = \frac{a}{2\left\{\dfrac{a\mu\left[(d_{cg} + d_{gp}) - d_{cg}\dfrac{V_p}{V_g}\right]D}{2\pi d_{gp}\left(\dfrac{V_p}{V_g} + \mu\right)} + r_g\right\}} - 1 \tag{9.28a}$$

Upon substituting equivalent factors to suit the pentode problem this expression becomes

$$T_2 = \frac{a_2}{2\left\{\dfrac{a_2\mu_{23}\left[(d_{12} + d_{23}) - d_{12}\dfrac{V_{3\,\mathrm{av}}}{V_2}\right]D}{2\pi d_{23}\left(\dfrac{V_{3\,\mathrm{av}}}{V_2} + \mu_{23}\right)} + r_2\right\}} - 1 \tag{11.8}$$

where $D = \frac{r_2}{2d_{12}} \ln \frac{4\epsilon d_{12}}{r_2}$

T_2 = ratio of current transmitted by screen grid upon initial passage of electrons to current intercepted

μ_{23} = geometrical amplification factor of control grid relative to screen grid

It is necessary only to modify the factor T_2 by the transmission factor of the suppressor grid to obtain the screen current-division factor $\mathfrak{D}$.

The transmission through the suppressor plane will generally be so large that it will affect the over-all result by only a small fraction. The transmission of the suppressor plane is given approximately by

$$T_3 = \frac{a_3}{r_3} - 1 \tag{11.9}$$

since membrane-model studies show that only those electrons aimed at the center *half* of a suppressor-grid wire will strike it. The over-all current ratio is then given approximately by

$$\frac{1}{\mathfrak{D}} = \frac{1}{T_2} + \frac{1}{T_3} \tag{11.10}$$

Example: Consider the 6J7 pentode, which has the following dimensions:

d_{c1} =	8.65 mils	a_1 =	21.1 mils	r_1 =	1.1 mils
d_{12} =	56.8 mils	a_2 =	15.9 mils	r_2 =	1.25 mils
d_{23} =	109.5 mils	a_3 =	50 mils	r_3 =	2.25 mils
d_{3p} =	110 mils				

$$\frac{V_{3\,\mathrm{av}}}{V_2} = \frac{\ln\left[\sin\left(\frac{\pi \times 2.25}{50}\right)\right]}{2\left\{\ln\left[2\sin\left(\frac{\pi \times 2.25}{50}\right)\right] - \frac{2\pi}{50}\frac{109.5 \times 110}{219.5}\right\}}$$

$= 0.12$ from Eq. (11.7)

$\mu_{23} = 56$ from Fig. 7.11

for

$$\frac{d_{23}}{a_2} = \frac{109.5}{15.9} = 6.89$$

and

$$S = \frac{2r_2}{a_2} = \frac{2.5}{15.9} = 0.1572$$

$$T_2 = \frac{15.9}{2\left\{\frac{15.9 \times 56[(56.8 + 109.5) - 56.8 \times 0.12]0.0683}{2\pi \times 109.5(0.12 + 56)} + 1.25\right\}} - 1$$
$$= 4.3 \qquad \text{from Eq. (11.8)}$$

Note that this value of current transmission is only slightly less than the δ factor of 4.63 calculated from Eq. (9.25), considering the screen grid of the pentode as the control grid of a triode.

$$T_3 = \frac{50}{2.25} - 1 = 21.2 \qquad \text{from Eq. (11.9)}$$

$$\mathfrak{D} = \frac{1}{\frac{1}{4.3} + \frac{1}{21.2}} = 3.57 \qquad \text{from Eq. (11.10)}$$

This calculated value of $\mathfrak{D}$ agrees well with a measured value of 3.65. The agreement is, in fact, better than there is any reason to expect in view of the fact that the 6J7 does not have a plane-electrode structure at all but has a circular cathode and suppressor grid, elliptical control and screen grids, and a plane plate.

11.4. Amplification Factor of a Pentode. The design of a pentode presents a rather complex problem. Relatively little has been published on this important subject. Most of the design equations exist in the private notebooks of a few workers in the field and are largely empirical modifications of simple theoretical relations. In this and subsequent sections there is given a sketch of the factors involved in the determination of pentode tube constants. The results that are given can serve only as a rough guide to the fundamental relations and should not be taken as anything more than approximate design equations.

The amplification-factor formulas of a pentode may be expected to be considerably more complicated than those of a triode for two reasons. (1) There are three grids instead of one. (2) The division of current between the various electrodes is a function of the relative electrode potentials. If the amplification factor is calculated from electrostatic considerations as was done for the triode, there results an expression that gives the relative influence of the plate and control grid in keeping the *space current* constant. This is not the true amplification factor but what will be referred to as the "electrostatic amplification factor" since it gives the relative influence of the plate and control grid in controlling the off-cathode gradient of potential in a cold tube (or space current in a hot tube). The true amplification factor is a modification of this value that gives the relative influence of plate and control grid in controlling the *plate current*. The subject will be treated by first studying the field in a pentode, deducing the electrostatic amplification factor from it, and then modifying this to obtain the true amplification factor.

Electrostatic Field of a Pentode. The method of conformal transformations is not readily applied to tubes having several grids of different pitches. It is, however, possible to construct the field from the expression for the potential due to a single row of grid wires given in Eq. (7.72),

$$V(x,y) = -\frac{q}{4\pi\varepsilon_0} \ln\left[2\left(\cosh\frac{2\pi x}{a} - \cos\frac{2\pi y}{a}\right)\right] + C \qquad (7.72)$$

where the wires are spaced a distance a apart upon the y axis and q is the charge per unit length of wire. If three terms like the function in Eq. (7.72) are combined properly with a linearly varying component of potential, the resultant expression is a satisfactory representation of the field of a pentode.

First consider some of the properties of Eq. (7.72). In the first place the constant has the value

$$C = V_g + \frac{q}{2\pi\varepsilon_0} \ln\left(\frac{2\pi r_g}{a}\right) \qquad (11.11)$$

where V_g is the potential of the isolated grid and r_g is the radius of the grid wires. Near the grid wires the equipotential contours are circles concentric with the grid wire. In this vicinity the potential is given approximately by

$$V(x,y) = V_g - \frac{q}{2\pi\varepsilon_0} \ln\left(\frac{r}{r_g}\right) \qquad (11.12)$$

At a considerable distance from the grid the equipotential contours are straight lines parallel to the grid-wire plane whose approximate potential is given by

$$V(x) = V_g + \frac{q}{2\pi\varepsilon_0} \ln\left(\frac{2\pi r_g}{a}\right) \mp \left(\frac{qx}{2a\varepsilon_0}\right) \qquad (11.13)$$

where the upper sign is associated with potentials to the right of the grid plane and the lower with potentials to the left. It is seen that the potential varies linearly with distance from the grid-wire plane, just as it would from a plane with a surface-charge density of $\frac{q}{2a}$. The second term above gives the difference between the actual grid-wire potential and the *equivalent* potential of the grid plane, found by extending the straight-line portions of the potential profiles back to an intersection, as shown in Fig. 11.9. The depth of the fillet about the grid wires is given by letting

$x = 0$ and $y = \frac{a}{2}$ in Eq. (7.72). This substitution gives for the maximum deviation from grid-wire potential in the grid plane

$$V_g - V\left(0, \frac{a}{2}\right) = -\frac{q}{2\pi\varepsilon_0} \ln\left(\frac{\pi r_g}{a}\right) \tag{11.14}$$

It is therefore always true that the potential difference between the grid wire and the equivalent potential of the grid plane is 0.693 (= ln 2) of the

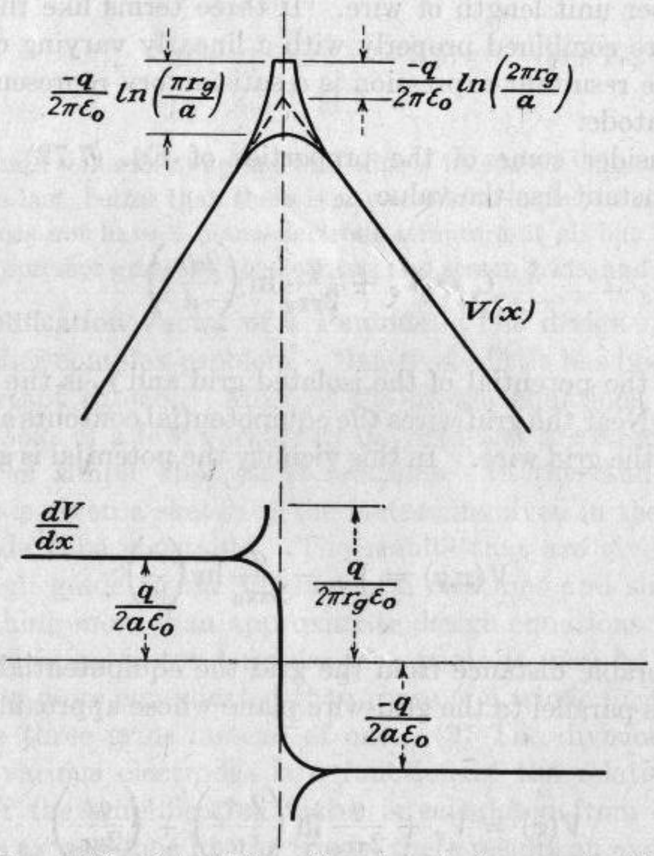


FIG. 11.9.—Potential profiles of a single row of grid wires.

maximum difference of potential encountered in the grid plane. Also shown in Fig. 11.9 is the gradient of potential about a row of grid wires.

Provided that the distance between electrodes is greater than the distance between grid wires, which is a good approximation with the exception that the suppressor-grid–plate distance is often less than the suppressor-grid pitch, the field of a pentode is given by the sum of three expressions like the right-hand side of Eq. (7.72) plus a linear component of field plus a constant.

$$V(x,y) = \sum_{n=1}^{3} -\frac{q_n}{2\pi\varepsilon_0} \ln\left\{2\left[\cosh 2\pi\left(\frac{x - d_{cn}}{a_n}\right) - \cos 2\pi\left(\frac{y - b_n}{a_n}\right)\right]\right\} - \frac{q_c x}{\varepsilon_0} + C \tag{11.15}$$

in which the subscript n assumes the values 1,2,3 to correspond to each of the three grids, d_{cn} is the distance from the cathode to the nth grid, b_n gives the relative location of the grid wires along a reference line normal to the cathode, a_n is the pitch of the nth grid, q_c is the cathode charge per unit area, and q_n is the charge per unit length of grid wire on the nth grid.

Upon making the usual approximations for large values of x and small values of y, it is possible to write the equations relating the potential at each electrode to the electrode charges. These equations are

$$V_c = 0 = -\frac{q_1 d_{c1}}{a_1\varepsilon_0} - \frac{q_2 d_{c2}}{a_2\varepsilon_0} - \frac{q_3 d_{c3}}{a_3\varepsilon_0} + 0 + C \tag{11.16}$$

$$V_1 = -\frac{q_1}{2\pi\varepsilon_0} \ln\left(2 \sin\frac{\pi r_1}{a_1}\right) - \frac{q_2 d_{12}}{a_2\varepsilon_0} - \frac{q_3 d_{13}}{a_3\varepsilon_0} - \frac{q_c d_{c1}}{\varepsilon_0} + C \tag{11.17}$$

$$V_2 = -\frac{q_1 d_{12}}{a_1\varepsilon_0} - \frac{q_2}{2\pi\varepsilon_0} \ln\left(2 \sin\frac{\pi r_2}{a_2}\right) - \frac{q_3 d_{23}}{a_3\varepsilon_0} - \frac{q_c d_{c2}}{\varepsilon_0} + C \tag{11.18}$$

$$V_3 = -\frac{q_1 d_{13}}{a_1\varepsilon_0} - \frac{q_2 d_{23}}{a_2\varepsilon_0} - \frac{q_3}{2\pi\varepsilon_0} \ln\left(2 \sin\frac{\pi r_3}{a_3}\right) - \frac{q_c d_{c3}}{\varepsilon_0} + C \tag{11.19}$$

$$V_p = -\frac{q_1 d_{1p}}{a_1\varepsilon_0} - \frac{q_2 d_{2p}}{a_2\varepsilon_0} - \frac{q_3 d_{3p}}{a_3\varepsilon_0} - \frac{q_c d_{cp}}{\varepsilon_0} + C \tag{11.20}$$

The above expressions may be solved for the charges in terms of the electrode potentials and then applied to Eq. (11.15) for the potential field. This process is somewhat involved, however; for ordinary purposes a simpler procedure that yields results accurate enough for most purposes is recommended. This simplified procedure consists in sketching the potential profiles and then correcting the originally assumed values. Ordinarily only one correction is necessary.

The simplified procedure for determining potential fields in pentodes is applied as follows: Ordinarily a complete plot of the potential field is not required, and potential profiles are sufficient. The procedure first calls for a sketch of the potential profiles within the tube. For convenience the profiles will be drawn through a wire of each grid and midway between the grid wires and the segments of such profiles joined. Actually, there may be no actual straight line in the tube that goes through a wire of each grid, but this makes no difference. Such a sketch is shown

in Fig. 11.10. The specified electrode potentials are marked, fillets of reasonable size are attached, and the fillets are then joined by straight lines. The next step in the procedure is to draw a curve of the gradient of potential between electrodes as taken from the slope of the straight-line portions of the profiles joining the electrodes. Such a curve of gradients is also shown in Fig. 11.10. If the gradients are taken from the profiles in units of volts per meter, then the gradients will have the values indicated on the figure in terms of the charges. From the four values of the gradient between the electrodes, the four electrode charges in units of coulombs per square meter can be calculated. When the electrode charges are known, then the factor $-\frac{q_n}{2\pi\epsilon_0}\ln\left(\frac{2\pi r_n}{a_n}\right)$ applied to each grid, as shown in Fig. 11.9, to see how good the original guess on the size of the fillet of potential around the wires was. Generally, the original guesses are not exact, and some values of electrode potentials different from those desired are found to fit the straight-line portions of the potential profiles. Correcting these values is a simple matter, and usually the first correction will be close enough to the exact one for ordinary purposes.

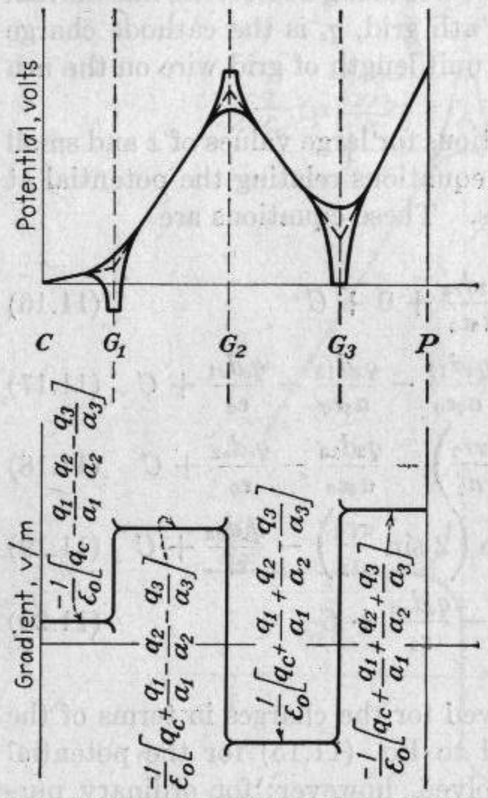
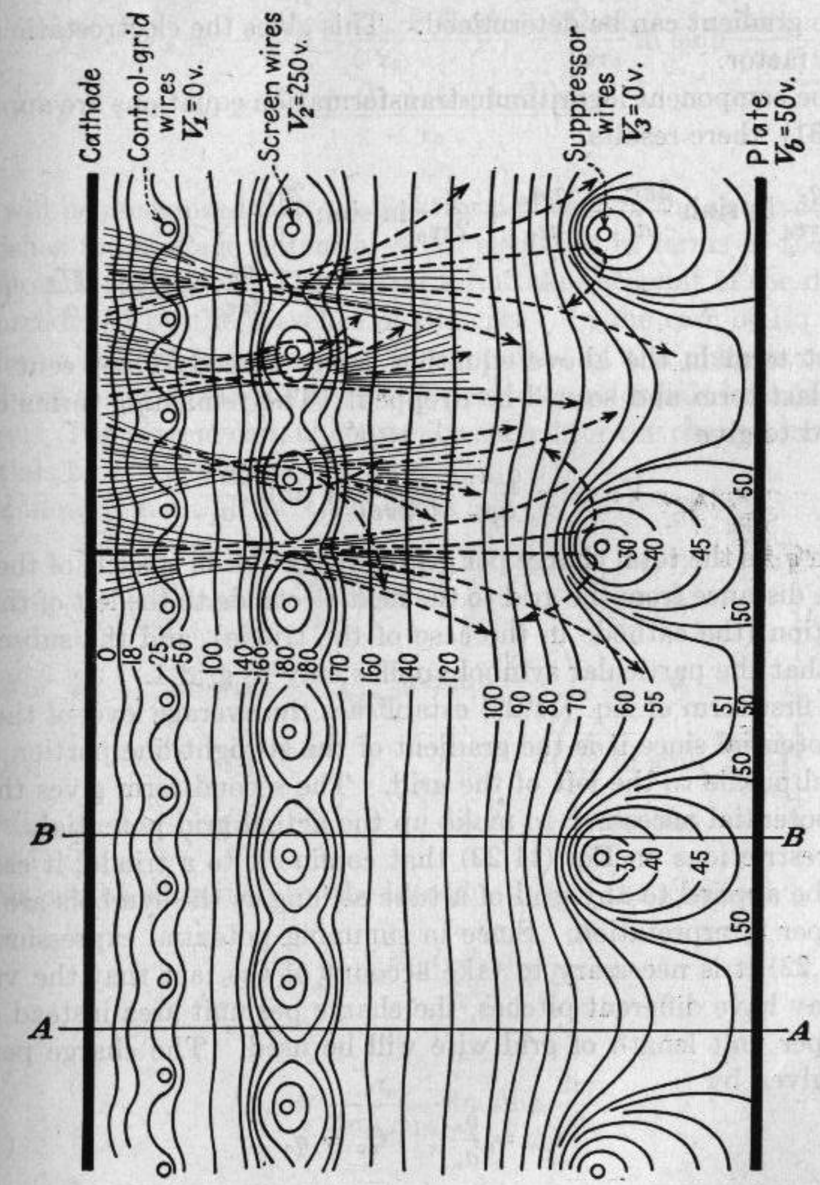


FIG. 11.10.—Potentials and gradients in a pentode.

A plot of the complete field within a pentode is shown in Fig. 11.11. The figure shows some typical electrode dimensions and electrode potentials and gradients in a pentode.

Electrostatic Amplification Factor of a Pentode.[1] From work on the triode it is possible to find an expression giving the potential of a grid in terms of its charge and the charge to the left of it. By combining such expressions for all the grids of a pentode there is obtained an expression for the cathode charge in terms of the electrode potentials. From this the relative effectiveness of the various electrodes in controlling the off-

[1] See also Dow, W. G., Equivalent Electrostatic Circuits for Vacuum Tubes, *Proc. I.R.E.*, vol. 28, pp. 548–556, December, 1940.



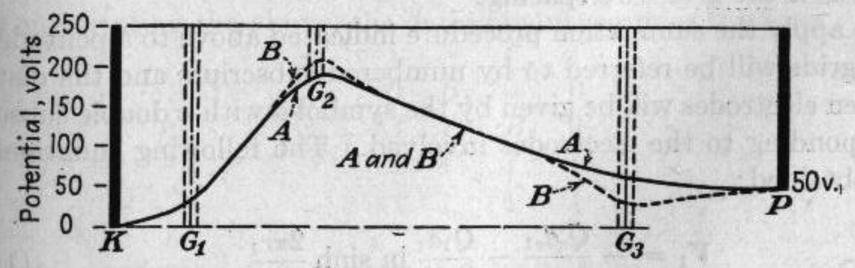


FIG. 11.11.—Electrostatic field and electron paths in a pentode. (*Schade.*)

cathode gradient can be determined. This gives the electrostatic amplification factor.

If the component logarithmic-transformation equations are applied to Eq. (7.31), there results

$$V_g = \frac{q_c}{2\pi\varepsilon_0} \ln \sinh \frac{2\pi r_g}{a} - \frac{q_c d_c}{a\varepsilon_0} + \frac{q_p}{2\pi\varepsilon_0} \ln \sinh \frac{2\pi r_g}{a} - \frac{q_p}{2\pi\varepsilon_0} \ln \cosh \frac{2\pi r_g}{a} \quad (11.21)$$

The last term in the above equation rarely exceeds 1 per cent of the second last term and so will be dropped. The remaining terms can be arranged to give

$$V_1 = -\frac{q_L d_L}{a_1 \varepsilon_0} - \frac{q_1}{2\pi\varepsilon_0} \ln \sinh \frac{2\pi r_1}{a_1} \quad (11.22)$$

in which q_L is the total charge per grid-wire section to the left of the grid, d_L is the distance from the grid to the next electrode to the left of the grid in question (the cathode in the case of the triode), and the subscript 1 means that the particular symbol applies only to grid 1.

The first term of Eq. (11.22) establishes the average level of the grid-plane potential since it is the gradient of the straight-line portion of the potential profile to the left of the grid. The second term gives the rest of the potential necessary to make up the actual grid potential. There are no restrictions on Eq. (11.22) that confine it to a triode; it can just as well be applied to any grid of a tube as long as the symbols are given the proper interpretation. Since in summing potential expressions like Eq. (11.22) it is necessary to take account of the fact that the various grids may have different pitches, the charge per unit area instead of the charge per unit length of grid wire will be used. The charge per unit area is given by

$$Q_n = \frac{q_n}{a_n} \qquad Q_c = q_c \quad (11.23)$$

where a_n is the grid-wire spacing.

To apply the summation procedure indicated above to a pentode, the three grids will be referred to by numbered subscripts and the distance between electrodes will be given by the symbol d with a double subscript, corresponding to the electrodes involved. The following equations are thus obtained:

$$V_1 = -\frac{Q_c d_{c1}}{\varepsilon_0} - \frac{Q_1 a_1}{2\pi\varepsilon_0} \ln \sinh \frac{2\pi r_1}{a_1} \quad (11.24)$$

$$V_2 - V_1 = -\frac{(Q_c + Q_1) d_{12}}{\varepsilon_0} - \frac{Q_2 a_2}{2\pi\varepsilon_0} \ln \sinh \frac{2\pi r_2}{a_2} \quad (11.25)$$

$$V_3 - V_2 = -\frac{(Q_c + Q_1 + Q_2) d_{23}}{\varepsilon_0} - \frac{Q_3 a_3}{2\pi\varepsilon_0} \ln \sinh \frac{2\pi r_3}{a_3} \quad (11.26)$$

$$V_p - V_3 = -\frac{(Q_c + Q_1 + Q_2 + Q_3) d_{3p}}{\varepsilon_0} \quad (11.27)$$

It will be recognized that the first term of each of the above equations establishes the average potential of an electrode in terms of the charge and distance to its left. The second term takes account of the deformation introduced by the presence of the grid. In the case of Eq. (11.27) there is no second term because the potential of the plate is constant. The above set of equations gives the electrode potentials in terms of the charges. The system can, of course, be solved for the charges in terms of potentials by Cramer's rule.

The determinant of the coefficients is

$$\Delta = \begin{vmatrix} d_{c1} & \left(\frac{a_1}{2\pi} \ln \sinh \frac{2\pi r_1}{a_1}\right) & 0 & 0 \\ d_{12} & d_{12} & \left(\frac{a_2}{2\pi} \ln \sinh \frac{2\pi r_2}{a_2}\right) & 0 \\ d_{23} & d_{23} & d_{23} & \left(\frac{a_3}{2\pi} \ln \sinh \frac{2\pi r_3}{a_3}\right) \\ d_{3p} & d_{3p} & d_{3p} & d_{3p} \end{vmatrix} \quad (11.28)$$

This can be simplified by the introduction of the symbols

$$K = \frac{d_{c1} d_{12} d_{23} d_{3p}}{\varepsilon_0^4} \quad (11.29)$$

and

$$B_n = \frac{a_n}{2\pi d_{(n-1)n}} \ln \sinh \frac{2\pi r_n}{a_n} \quad (11.30)$$

when $n = 1$, $n - 1 = c$.

After substituting the above, the determinant assumes the form

$$\Delta = K \begin{vmatrix} 1 & B_1 & 0 & 0 \\ 1 & 1 & B_2 & 0 \\ 1 & 1 & 1 & B_3 \\ 1 & 1 & 1 & 1 \end{vmatrix} \quad (11.31)$$

The value of this determinant has the simple form

$$\Delta = K(1 - B_1)(1 - B_2)(1 - B_3) \quad (11.32)$$

The cathode charge is given by

$$Q_c = V_1\left[\frac{(1-B_2)(1-B_3)}{d_{c1}} + \frac{B_1(1-B_3)}{d_{12}}\right] + V_2\left[-\frac{B_1(1-B_3)}{d_{12}} - \frac{B_1B_2}{d_{23}}\right] + V_3\left[\frac{B_1B_2}{d_{23}} + \frac{B_1B_2B_3}{d_{3p}}\right] + V_p\left[-\frac{B_1B_2B_3}{d_{3p}}\right] \quad (11.33)$$

Reference to Eq. (7.35) shows that the electrostatic amplification factor is given by the ratio of the coefficients of the first and last terms.

$$\mu_{1p} = \frac{\dfrac{(1-B_2)(1-B_3)}{d_{c1}} + \dfrac{B_1(1-B_3)}{d_{12}}}{-\dfrac{B_1B_2B_3}{d_{3p}}} \quad (11.34)$$

Substitution of the values of the various B's shows that the above expression for amplification factor is independent of d_{c1}. In general, the electrostatic amplification factor of the pentode will have a value approximately equal to that given by the product of three triode mu's calculated by considering the plate, the suppressor grid, and the screen grid as plates of a triode and the next two eleetrodes in order toward the cathode as grid and cathode.[1]

Also available from Eq. (11.33) is the electrostatic amplification factor giving the relative effectiveness of the control grid and screen grid in controlling space current. It is

$$\mu_{12} = \frac{\dfrac{(1-B_2)(1-B_3)}{d_{c1}} + \dfrac{B_1(1-B_3)}{d_{12}}}{-\dfrac{B_1(1-B_3)}{d_{12}} - \dfrac{B_1B_2}{d_{23}}} \quad (11.35)$$

The above expressions for amplification factor have some small inherent inaccuracies due to the fact that the filleting of potentials about some of the grids has been neglected.The inaccuracy is probably not more than a few per cent. More accurate expressions for electrostatic amplification factor may be derived from Eqs. (11.16) to (11.20), but these are so cumbersome as to be almost totally useless.

True Amplification Factor of a Pentode. The true amplification factor of a pentode must take into account not only the electrostatic action

[1] THOMPSON, B. J., Space Current Flow in Vacuum Tube Structures, *Proc. I.R.E.*, vol. 31, pp. 485–491, September, 1943.

within the tube but also the division of current between the electrodes. Let the true amplification factor of the pentode be defined by

$$\mu_p = -\left(\frac{dV_p}{dV_1}\right)_{I_p \text{ const}} = -\left(\frac{\dfrac{\partial I_p}{\partial V_1}}{\dfrac{\partial I_p}{\partial V_p}}\right) \quad (11.36)$$

In deriving the true amplification factor, use will be made of the following relation,

$$I_s = G\left(V_1 + \frac{V_2}{\mu_{12}} + \frac{V_3}{\mu_{13}} + \frac{V_p}{\mu_{1p}}\right)^n \quad (11.37)$$

where n is a constant of approximate value 1.5, the μ's are electrostatic amplification factors measuring the relative effectiveness of the electrode in question and the control grid in controlling the space current, G is perveance of value $\dfrac{2.335 \times 10^{-6}}{d^2}$ amperes per unit area per volt^n, and d is the equivalent control-grid–cathode spacing as calculated from Eq. (8.45) but with the screen grid considered as the triode plate.

Let the ratio of plate to screen-grid current be given by

$$\frac{I_p}{I_2} = g\left(\frac{V_p}{V_2}\right) \quad (11.38)$$

and let the functional relation be indicated subsequently by the symbol g.

The ratio of plate to space current is given by

$$\frac{I_p}{I_s} = \frac{I_p}{I_2 + I_p} = \frac{\dfrac{I_p}{I_2}}{1 + \dfrac{I_p}{I_2}} = \frac{g}{1+g} \quad (11.39)$$

and hence

$$I_p = I_s\frac{I_p}{I_s} = G(V_{eq})^n\frac{g}{1+g} \quad (11.40)$$

where V_{eq} is the equivalent voltage $V_1 + \dfrac{V_2}{\mu_{12}} + \dfrac{V_3}{\mu_{13}} + \dfrac{V_p}{\mu_{1p}}$.

The partial derivatives that enter into the determination of the true amplification factor can now be evaluated.

$$\frac{\partial I_p}{\partial V_1} = nG(V_{eq})^{n-1}\frac{g}{1+g} \quad (11.41)$$

$$\frac{\partial I_p}{\partial V_p} = \frac{nG}{\mu_{1p}}(V_{eq})^{n-1}\frac{g}{1+g} + G(V_{eq})^n\frac{\dfrac{dg}{dV_p}}{(1+g)^2} \quad (11.42)$$

Upon combining these last two expressions, the resulting expression for the true amplification factor is

$$\mu_p = \frac{\mu_{1p}}{1 + \dfrac{\mu_{1p} V_{eq} \dfrac{dg}{dV_p}}{ng(1+g)}} \tag{11.43}$$

If the ratio of plate to screen-grid current is assumed to have the form previously given

$$g = \frac{I_p}{I_2} = \mathfrak{D}\left(\frac{V_p}{V_2}\right)^m \tag{11.1}$$

then

$$\mu_p = \frac{\mu_{1p}}{1 + \dfrac{\mu_{1p} m V_{eq}}{nV_p\left[1 + \mathfrak{D}\left(\dfrac{V_p}{V_2}\right)^m\right]}} \tag{11.44}$$

From this it is seen that the true amplification factor is less than the electrostatic amplification factor by a considerable factor.

11.5. Transconductance of a Pentode. The transconductance of a pentode is readily obtained. Let

$$G_m = G_{1p} = \frac{\partial I_p}{\partial V_1} \tag{11.45}$$

but

$$I_p = I_s \frac{I_p}{I_s} \tag{11.46}$$

and so

$$G_{1p} = \frac{\partial I_s}{\partial V_1} \frac{I_p}{I_s} \tag{11.47}$$

since the ratio of plate to space current is independent of control-grid voltage. The quantity $\dfrac{\partial I_s}{\partial V_1}$ may be designated as G_{1s} and may be obtained from the triode mutual-conductance formula [Eq. (8.47)] by considering the screen grid of the pentode as the plate of a triode. Equation (11.47) states that the control-grid–plate transconductance of a pentode is equal to the triode mutual conductance of the first three electrodes, reduced by the ratio of the plate current to the screen current.

11.6. Plate Resistance of a Pentode. The plate resistance of a pentode is the reciprocal of the plate conductance defined by

$$\frac{1}{R_p} = G_p = \frac{\partial I_p}{\partial V_p} \tag{11.48}$$

Utilizing Eq. (11.40),

$$G_p = \frac{\partial I_s}{\partial V_p} \frac{g}{1+g} + \frac{I_s \dfrac{dg}{dV_p}}{(1+g)^2} \tag{11.49}$$

The first term in this last expression results from the variation of the emitted current with plate voltage. The second term results from the change in plate current occasioned by the variation in current division with plate voltage. When the current-division function of the pentode has the form of Eq. (11.1), the above expression reduces to

$$G_p = \frac{G_{1s}}{\mu_{1p}} \frac{I_p}{I_s} + \frac{I_s \mathfrak{D} m \left(\dfrac{V_p}{V_2}\right)^{m-1}}{V_2\left[1 + \left(\dfrac{V_p}{V_2}\right)^m\right]^2} \tag{11.50}$$

11.7. Design Considerations. In a pentode the ordinary constants are readily made to assume satisfactory values. Prime interest is centered in the transconductance. Here the same considerations apply as in the triode, and no greater difficulty is encountered. The amplification factor and plate resistance are naturally high and require no particular attention. Thus interest is focused upon some of the other characteristics of the tube that affect its operation. These other characteristics are

1. Suppression of secondary electrons.
2. Sharpness of the shoulder of the plate-current–plate-voltage characteristic.
3. Plate-current to screen-grid-current ratio.

The above factors are controlled by some factors that have not appeared before in this study of vacuum-tube design. These are

1. Shape of the potential field before the plate.
2. Electron deflection by the grids.

In previous considerations of tube characteristics the principal concern has been with the potential field and with the space-charge flow. *In the pentode, in addition, the electron paths are critical.*

The suppression of secondary electrons from the plate is not a difficult problem, though some attention must be paid to the electrode dimensions. The critical factors are the pitch of the suppressor grid and its distance from the plate. In Fig. 11.12 is shown the effect of different suppressor-grid pitches upon the retarding potential offered to secondary electrons created at the plate as a function of plate voltage as calculated from Eq.

(11.6). The minimum potential referred to in the figure is the minimum of potential on a line normal to the plate passing midway between grid wires. As the suppressor-grid pitch is decreased, this minimum potential is decreased and the retarding potential offered to the secondaries thus increased. The effects are linear with the various potentials involved, provided that space-charge effects do not distort the potentials appreciably. In Fig. 11.13 is shown the effect of putting the suppressor grid in different positions between the screen grid and plate, as calculated from Eq. (11.6). This changes both the magnitude and the rate of change of the retarding potential. For suppressing plate secondaries it is desirable to have a fixed retarding potential. This cannot be realized, and therefore the arrangement of electrodes that gives the most retarding potential at *low* plate potentials is desired. This means that a small suppressor-grid–plate distance is indicated. So also is a small suppressor-grid pitch. Some other considerations, as we shall see, limit the degree to which the suppressor pitch can be reduced, but both the above factors should be considered to ensure secondary suppression.[1]

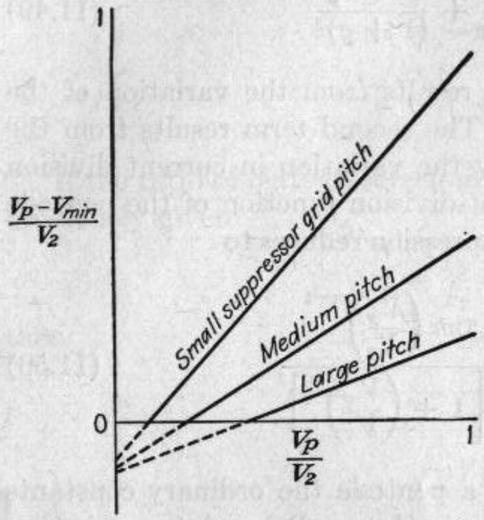


FIG. 11.12.—Retarding increment of potential before the plate of a pentode as a function of plate voltage for various suppressor-grid pitches.

The other new factor in pentode design, *viz.*, the deflection of electrons by the grids, is probably more important than the secondary suppression, since the latter is usually achieved without too much difficulty. The electron deflection by the grids will affect strongly both the sharpness of the plate-current–plate-voltage characteristic and the ratio of plate to screen-grid current. The more the electrons are deflected by the grids, the fewer will reach the plate at ordinary plate voltages and the less rapid will be the increase of plate current with plate voltage.

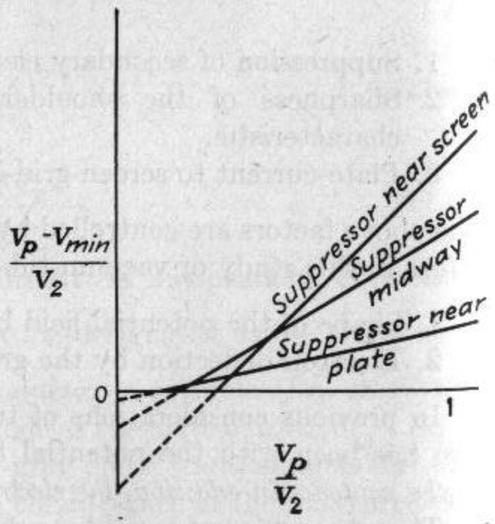


FIG. 11.13.—Retarding increment of potential before the plate of a pentode as a function of plate voltage for various suppressor-grid locations.

[1] JONKER, J. H. L., Pentode and Tetrode Output Valves, Parts I, II, *Wireless Engr.*, vol. 16, pp. 274–286, 344–349, July, 1939.

An exact study of electron deflection by the grid wires is rather difficult, but an excellent approximate analysis of the deflections can be made by using the fact that the potential field between the grid wires acts like a cylindrical lens and may have the effect of either focusing or dispersing the electrons which pass between them. Thus in Fig. 9.10*a*

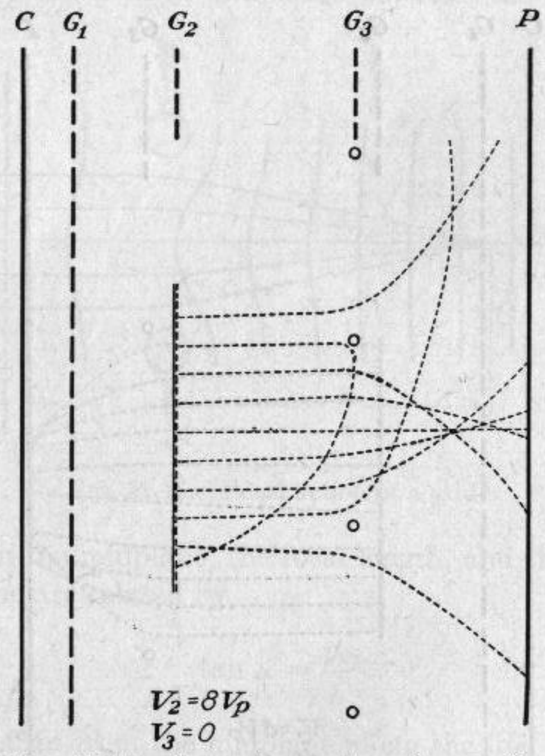


FIG. 11.14.—Scattering action of a small-pitch suppressor grid far from the plate of a pentode.

it is seen that the control grid gives the focusing action of a convergent lens. This lens, however, has some very pronounced aberrations. The focal length for parts of the lens near the grid wires is less than for the center of the space. In the language of optics, the lens has a positive spherical aberration. In Fig. 10.6 the focusing action of the control grid may again be seen. It is also seen that the screen grid exhibits the characteristics of a divergent lens. In Figs. 11.14 and 11.15 it is seen that the suppressor grid has a convergent-lens action. The nature of the lens caused by the potential field between grid wires depends upon the gradient of potential on the two sides of the grid and is, to first order, independent of such things as the grid radius and pitch.

The focal length of the lenses formed by the grid wires will be shown in the chapter on Electrostatic Electron Optics to be given by

$$f = \frac{2V_n}{\left(\frac{dV}{dx}\right)_R - \left(\frac{dV}{dx}\right)_L} \tag{11.51}$$

where V_n is the potential midway between the wires of grid n, $\left(\frac{dV}{dx}\right)_R$ is

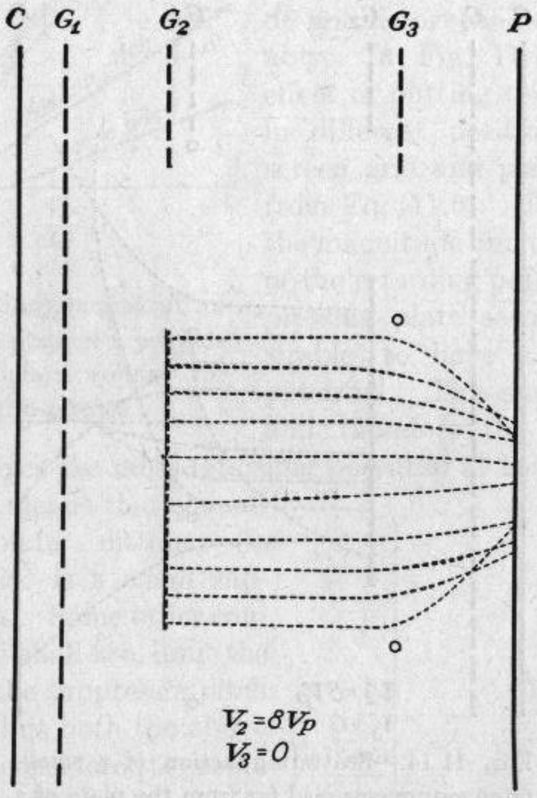


FIG. 11.15.—Focusing action of a large-pitch suppressor grid close to the plate of a pentode.

the gradient of potential to the right of the grid, and $\left(\frac{dV}{dx}\right)_L$ is the gradient of potential profile to the left of the grid plane. It is seen that when the gradient of potential increases upon passing through the grid, f is positive and the lens is convergent. When the opposite is true, the lens is divergent. The focal distance as given by the above formula will be modified somewhat in actual cases when the gradient of potential is not zero on the side where the focus occurs, for the electron trajectories

will tend to be sections of parabolic curves instead of straight lines. The formula does, however, give a fairly exact indication of the principal effects and the correct value of sidewise components of velocity.

The nature of the focusing of a set of grid wires is shown schematically in Fig. 11.16. An electron passing midway between grid wires suffers no sidewise deflection whatever. As the distance from the midplane increases, the electrons receive more and more deflection, by a linear law, so that they all cross over at the same point. This holds true almost exactly for the center half of the space between the grid wires. The

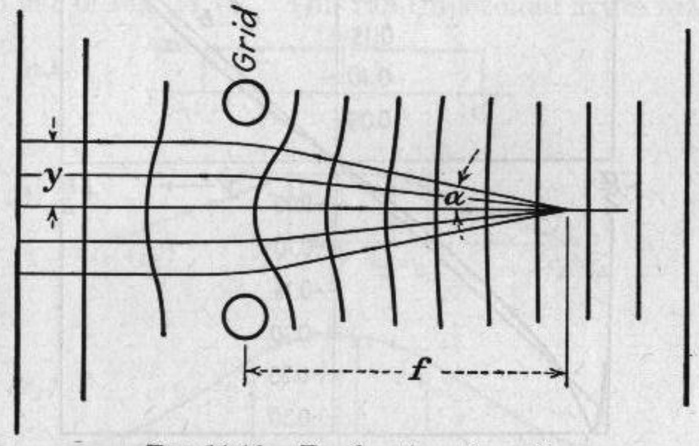


FIG. 11.16.—Focal action of a grid.

initial offset from the midplane, the focal length, and the tangent of the angle of deflection are related by

$$\tan \alpha = \frac{y_0}{f} \tag{11.52}$$

where y_0 is the offset from the midplane along the line of the grid wires of the electron's initial position, f is the focal length, and α is the angle of deflection. In terms of velocity components,

$$\frac{y_0}{f} = \frac{v_y}{v_x} \cong \frac{v_y}{v} \tag{11.53}$$

where v_x and v_y are the forward and sidewise components of velocity possessed by the electron shortly after passing through the grid plane and v is the total velocity.

The experimental agreement between this formula and the actual behavior is quite good, as shown in Fig. 11.17. Here is shown the actual deflection as measured on an elastic membrane (curve a) and the deflection calculated on the assumption of a constant grid-plane potential equal to the mean potential of the grid plane (curve b). The measured

deflection is seen to become greater than the linear value as the electrons first approach the grid wires. This is due to lens aberration. The deflection then decreases. This is because, as may be seen in Fig. 9.10, those electrons which come very close to the grid wires are so strongly deflected that they come within the influence of the next grid wire and suffer a deflection in the opposite direction.

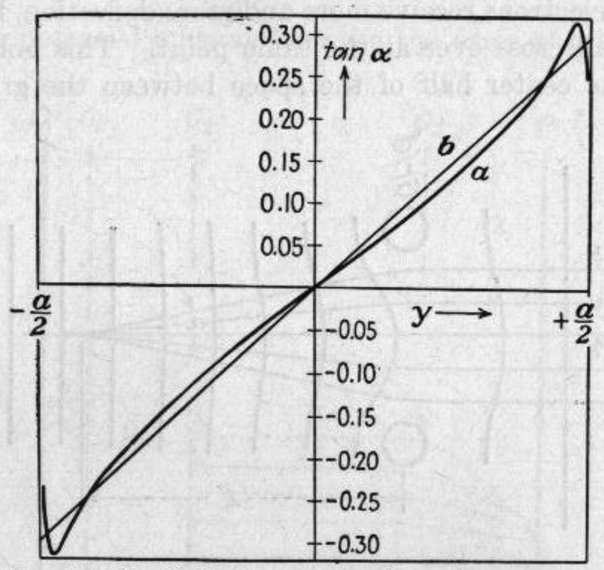


FIG. 11.17.—Deflection angle of an electron as a function of the offset from the mid-point between grid wires.

Introducing the value of f from Eq. (11.51) and the expression for velocity in terms of potential,

$$v_{yn} = \frac{2.96 \times 10^5}{\sqrt{V_n}} y_0 \left[\left(\frac{dV}{dx} \right)_R - \left(\frac{dV}{dx} \right)_L \right] \tag{11.54}$$

in which v_{yn} is the sidewise component of velocity acquired, in meters per second, from the nth grid, V_n is the mean grid potential of the nth grid, in volts, and $\left(\frac{dV}{dx}\right)_R$ and $\left(\frac{dV}{dx}\right)_L$ are the gradients of potential to the right and left of the nth grid plane, respectively, in volts per meter. The corresponding expressions for the three grids of a pentode become

$$v_{y1} = \frac{2.96 \times 10^5}{\sqrt{V_1}} y_0 \left(\frac{V_2 - V_1}{d_{12}} - \frac{V_1}{d_{c1}} \right) \tag{11.55}$$

$$v_{y2} = \frac{2.96 \times 10^5}{\sqrt{V_2}} y_0 \left(\frac{V_3 - V_2}{d_{23}} - \frac{V_2 - V_1}{d_{12}} \right) \tag{11.56}$$

$$v_{y3} = \frac{2.96 \times 10^5}{\sqrt{V_3}} y_0 \left(\frac{V_p - V_3}{d_{3p}} - \frac{V_3 - V_2}{d_{23}} \right) \tag{11.57}$$

in which the sidewise components of velocity are in meters per second, grid potentials are understood to be mean grid-plane potentials in volts, and distances are in meters.

The sidewise components of velocity are additive in the form shown in Fig. 11.18. Curve a of this figure shows the distribution of the sidewise components of velocity after passing through one grid. The sidewise components of velocity are uniformly distributed between plus and minus v_1 where this is the maximum component of this tangential velocity. After passing through two grids the distribution of velocities has the form shown in b of Fig. 11.18. This is a trapezoidal figure with velocities

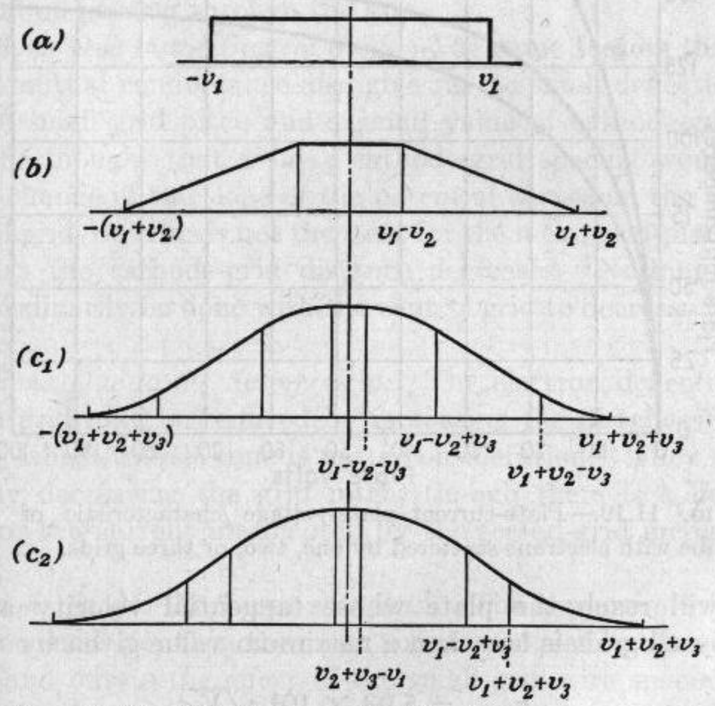


FIG. 11.18.—Distribution of sidewise component of velocity in a beam of electrons scattered by one, two, or three grids.

reaching from $v_1 + v_2$ to the negative of the same quantity. In this particular figure it has been assumed that v_2 is less than v_1. The distribution given is arrived at by sliding a rectangular aperture of width $2v_2$ and of the same height as the rectangle of part a of the figure over the rectangle of part a and plotting the exposed area as a function of the displacement of the aperture. Upon repeating the process with an aperture of width $2v_3$, the distributions of c_1 and c_2 result. The distribution of c_1 results from the assumption that v_3 is less than $v_1 - v_2$, whereas the distribution of c_2 results when v_3 is greater than $v_1 - v_2$.[1] Both

[1] JONKER, J. L. H., Electron Trajectories in Multigrid Valves, *Philips Tech. Rev.*, vol. 5, pp. 131–139, May, 1940.

these distributions consist of parabolic and straight-line sections with the same over-all span. The difference is that the distribution of c_1 has a straight-line center section, while that of c_2 has a parabolic center section of large curvature. Both these distributions already show approximately the form of a Gauss error curve, which they would obtain from the random deflection of a large number of grids.

If the distribution of tangential velocities as given above is known, the plate-current–plate-voltage characteristic can be calculated. It is necessary only to remember that at every plate voltage V_p only those

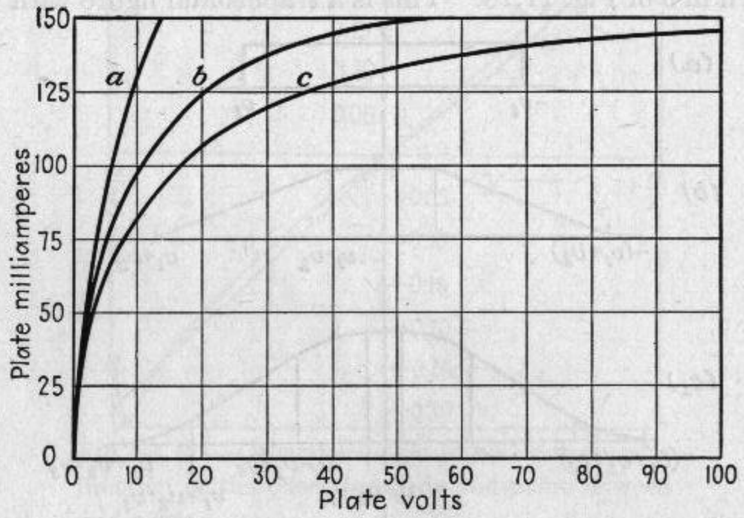


FIG. 11.19.—Plate-current–plate-voltage characteristic of a tube with electrons scattered by one, two, or three grids.

electrons will reach the plate whose tangential velocity after being deflected by all grids is less than a maximum value given by

$$v_{y\,\max} = 5.93 \times 10^5 \sqrt{V_p} \tag{11.58}$$

This follows from the fact that an electron reaching the plate will have the value of velocity given by Eq. (11.58), and if all this velocity exists in a sidewise component then the electron will graze the plate and fall back through more positive spaces in the suppressor grid. If the plate-current–plate-voltage characteristic be calculated on this basis, the curves of Fig. 11.19 result. The curves shown are for the corresponding velocity diagrams of Fig. 11.18. Curve *a* is similar to that which occurs in a beam-power tube. This is a tetrode with aligned grids in which the net effect of deflection by two grids is not very different from that of one grid as seen in Fig. 10.6. This effect undoubtedly contributes to the sharpness of the shoulder of the beam-power tube. The effect of two grids is shown at *b*. This would correspond to the curve of an ordinary tetrode free from secondary emission. The curve *c* is a typical pentode characteristic resulting from the random action of three grids. It is evident that it is necessary to keep the total sidewise velocity component of the electrons low in order to achieve a sharp shoulder to the characteristic. Since all three grids contribute to this in approximately the same amount, it is necessary to study the effect of each of the grids to see what can be done to reduce the resulting sidewise component of velocity. Examination of Eq. (11.54) shows that, in general, the sidewise component of velocity introduced by grid deflection may be reduced by either reducing the grid pitch or decreasing the change in the gradient of potential on passing through the grid.

Deflections Due to the Control Grid. The same factors that give rise to a large mutual conductance also give rise to small deflection. These factors are small grid pitch and a small value of cathode-grid spacing. It might be thought that a small cathode-grid spacing would give rise to a large change in the slope of the potential curves on the two sides of the control grid, but this is not the case, for the mean grid-plane potential increases as the cathode-grid distance decreases. Nothing much can therefore ordinarily be done with the control grid to decrease the electron deflection.

Deflections Due to the Screen Grid. The electron deflections due to the screen grid may be reduced by increasing the distances d_{12} and d_{23} in cases in which transit time is not a consideration. They can also be reduced by decreasing the grid pitch, though there is a limit to this method, for the current intercepted by the screen grid increases as this is done.

Deflections Due to the Suppressor Grid. The suppressor pitch cannot be made too small, for then the mean suppressor-plane potential becomes too small and offsets the effect of the small grid-wire spacing as far as electron deflections are concerned. It is, however, possible to make the suppressor-grid–plate distance quite small, with resulting improvement in the deflection characteristics. This has the added advantage, as is apparent in Fig. 11.15, that the current is concentrated in front of the plate, giving rise to considerable space charge, which aids in increasing the retarding potential presented to the secondary electrons originating at the plate. The bad deflection situation that results from the use of a large suppressor-grid–plate distance is shown in Fig. 11.14.

By making use of the possibilities indicated in the above outline it is possible to make pentode tubes with shoulders of the plate-current–plate-voltage characteristic nearly as sharp as those of the beam-power tube.

CHAPTER 12

NOISE IN VACUUM TUBES

12.1. Noise as a Limiting Factor in the Ultimate Sensitivity of Electronic Devices. Vacuum-tube amplifiers making use of triodes and pentodes are capable of giving extremely large amplification of power and voltage. In fact, it may be said that an amplification of any desired magnitude may be achieved by the use of vacuum tubes. At first glance this would seem to imply that arbitrarily small signals could be detected. Ultimately, however, it is found that there is a limit determined by the noise generated by the random motion of electrons at the input of the circuit. Any signal whose level is appreciably less than that of the electron noise will be masked by it. The order of the equivalent voltage of the electron noise is extremely small, of the order of millimicrovolts, but many electronic devices have enough amplification to bring this up to a detectable level.

Electron noise shows up in both resistors and in vacuum tubes. Even in a passive resistor, the molecular and the electronic agitation is evident with sufficient amplification. Here the noise is referred to as "thermal-agitation noise." In vacuum tubes the random emission and fluctuation of space-charge-limited currents contribute a similar noise. In temperature-limited tubes this noise is called "shot noise" and is due to random emission. In space-charge-limited emission tubes the noise is much less and is called "reduced shot noise." Both types of noise are characterized by a uniform distribution of energy over the frequency spectrum. Depending upon the application, the noise from either the tubes or the resistors at the input of an electronic device may predominate.

Needless to say, resistor and tube noise is an exclusive concern of electronic devices. No other type of device can have sufficient sensitivity to be limited by random electron motion. Resistor and tube noise set the ultimate sensitivity of high-gain amplifiers, receivers, phototube input circuits, and television pickup tubes. Although resistor and tube noise can never be avoided, much can be done by circuit design and selection of tubes to approach the minimum attainable noise.

Although the formulas for various types of electron noise and their application are quite simple, their derivation is dependent upon some

intimate aspects of thermodynamics and the kinetic theory of gases that have not been developed in this book. For this reason, only the basis of the development will be given, and emphasis will be upon the interpretation and application of the formulas.

12.2. Noise in Resistors. Noise in resistors is due to the random motion of electrons within them. The noise energy is proportional to the resistance, the absolute temperature, and the frequency band width over which the noise is observed and is independent of the material of which the resistor is made. The noise energy increases with absolute temperature because the molecular agitation is proportional to this. The noise is probably made up of extremely short and sharp pulses resulting from the impact of the electrons with the molecules. These pulses are probably so short and sharp that they are made up of a continuous distribution of frequency components of equal amplitude up to the highest frequencies in use today. As a result, the noise energy is uniformly distributed over the useful r-f spectrum.

The mean-square thermal-agitation noise voltage $\bar{e}^2$ across the terminals of a resistor R at an absolute temperature T, associated with a frequency band width B, is

$$\bar{e}^2 = 4kT_rRB \tag{12.1}$$

where e is rms value of the noise voltage, volts
k is Boltzmann's constant, 1.3805×10^{-23} watt-second per °K
R is resistance, ohms
T_r is room temperature, °K (°C + 273)
B is frequency band width, cycles per sec[1,2]

If room temperature is taken as 290°K (63°F), the expression for the rms noise voltage becomes

$$e_{\text{rms}} = 126.0\sqrt{RB} \quad \text{micromicrovolts} \tag{12.2}$$

A nomograph of this equation is shown in Fig. 12.1. For the sample construction line shown, the rms value of the noise voltage across a 1,000-ohm resistor in a frequency band of 10 mc is 12.6 microvolts.

The effect of thermal-agitation noise can be expressed either as an emf in series with the resistor considered noiseless or as a constant-current generator in parallel with the resistor considered noiseless. This follows from application of Thévénin's and Norton's theorems. The two equivalent circuits of a noisy resistor are shown in Fig. 12.2. The circle in the figure indicates a zero-impedance constant-voltage generator

[1] JOHNSON, J. B., Thermal Agitation of Electricity in Conductors, *Phys. Rev.*, vol. 32, pp. 97–109, July, 1928.

[2] NYQUIST, H., Thermal Agitation of Electric Charge in Conductors, *Phys. Rev.*, vol. 32, pp. 110–113, July, 1928.

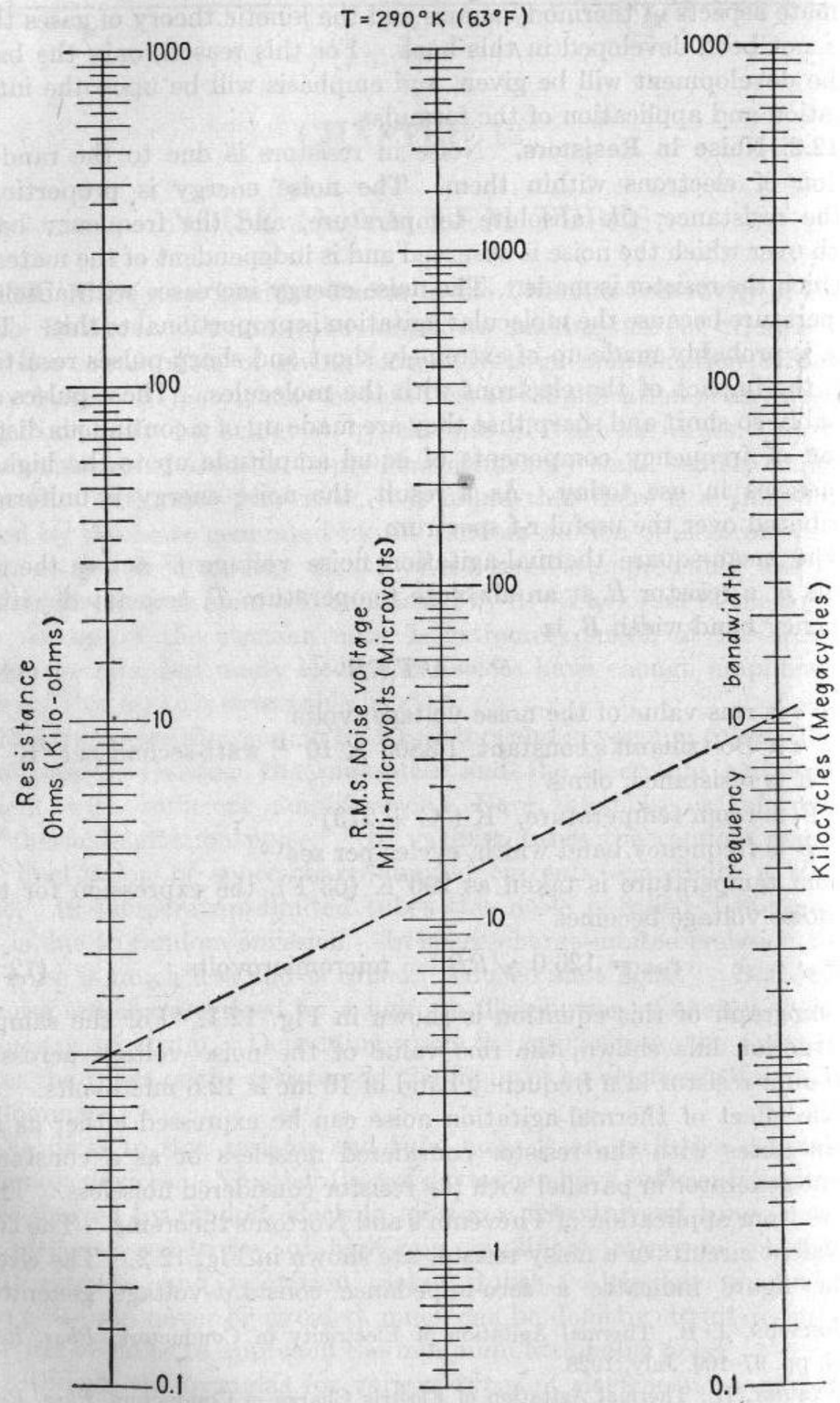


FIG. 12.1.—Nomographic chart of rms noise voltage across a resistor.

of the value given by Eq. (12.2). The square in the figure indicates an infinite-impedance constant-current generator whose output is

$$i_{\text{rms}} = 126.0\sqrt{\frac{B}{R}} \quad \text{micromicroamperes} \tag{12.3}$$

at room temperature. This is obtained by dividing the expression for the rms noise voltage by the resistance. The general form of the above equation is

$$i_{\text{rms}} = \sqrt{\frac{4kTB}{R}} \quad \text{amperes} \tag{12.4}$$

where all the symbols have their previous significance. A nomograph of the rms noise-current equivalent of a noisy resistor to fit the right circuit of

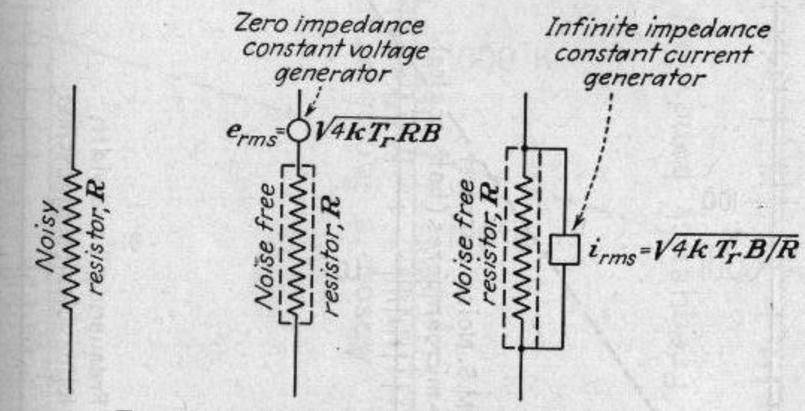


FIG. 12.2.—Equivalent circuits of a noisy resistor.

Fig. 12.2 is given in Fig. 12.3. For the sample construction line shown, the equivalent rms noise current of a 1,000-ohm resistor is 4 milli-microamperes for a frequency band width of 1 mc.

The noise power associated with thermal-agitation noise in a resistor is i^2R, or

$$N = 4kTB \quad \text{watts} \tag{12.5}$$

Note that this is independent of the value of resistance. The *available* noise power that can be obtained from a resistor by perfect matching is just one-fourth of this value. This follows immediately from maximizing the power obtainable from the middle circuit of Fig. 12.2 by varying the resistance load on such a generator. Maximum output power is obtained when the load resistance equals the generator resistance and has a value of $\frac{e^2}{4R}$. Thus the *available* noise power is

$$N_a = kTB \quad \text{watts} \tag{12.6}$$

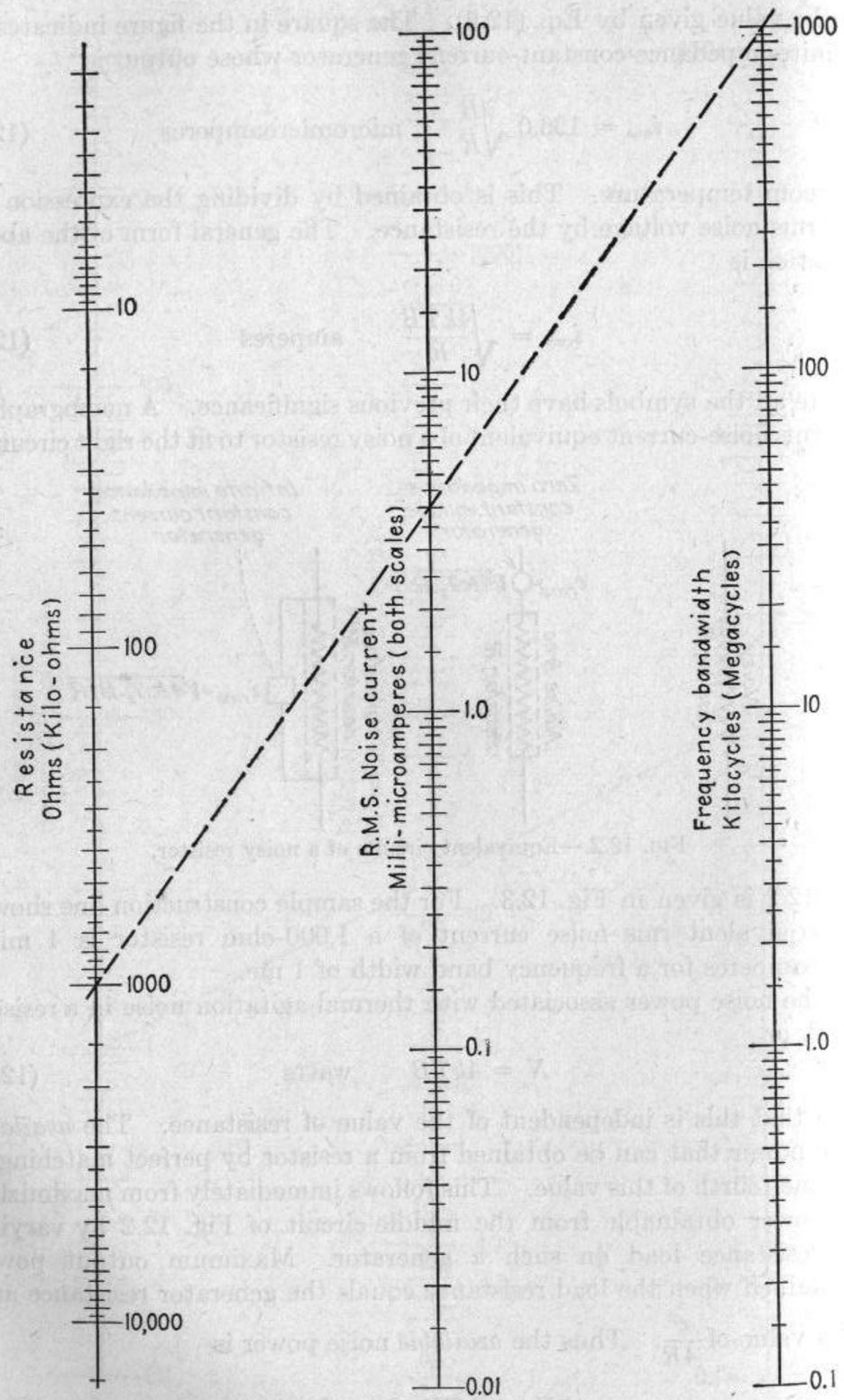


FIG. 12.3.—Nomographic chart of equivalent rms current through a noisy resistor.

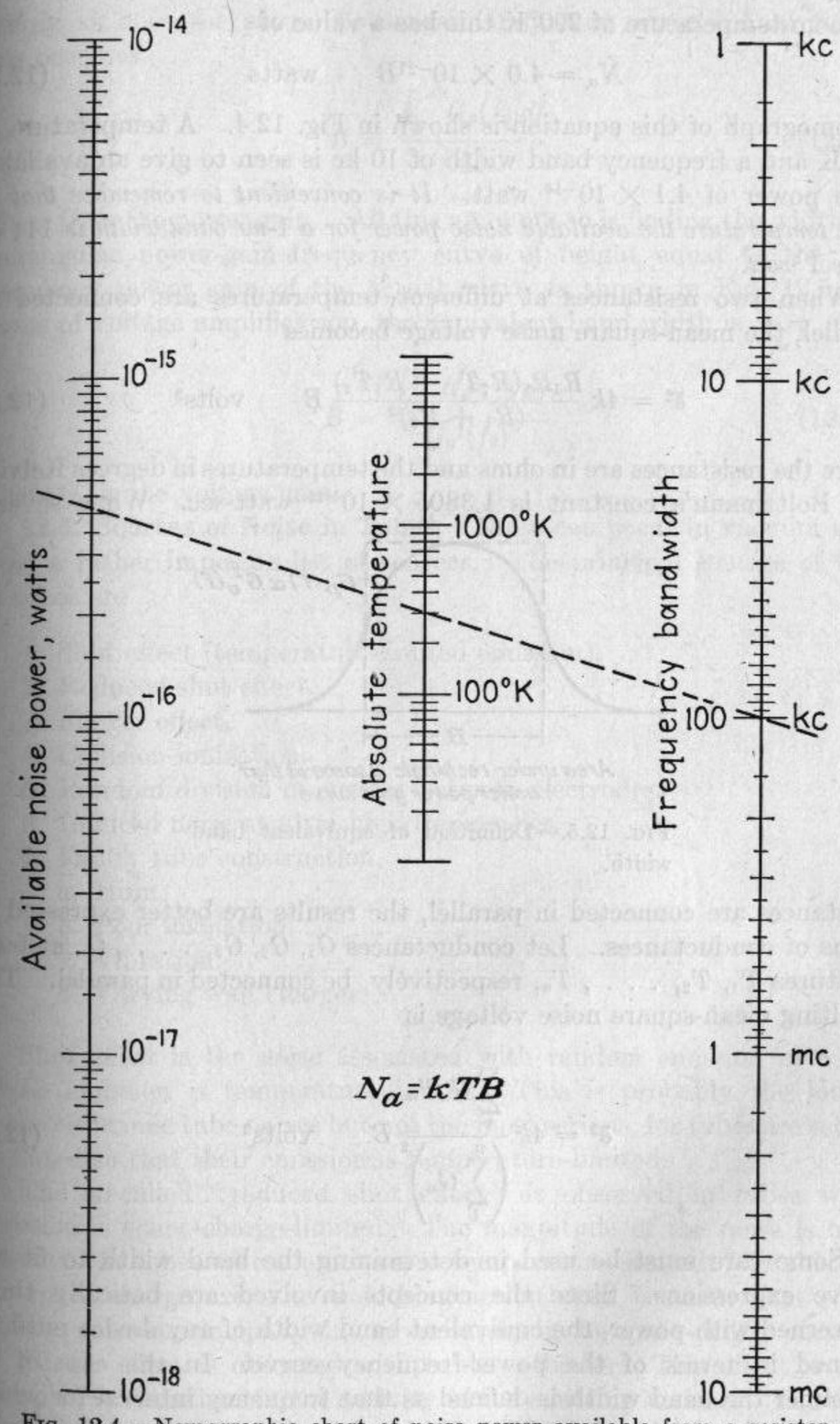


FIG. 12.4.—Nomographic chart of noise power available from a resistor.

At room temperature of 290°K this has a value of

$$N_a = 4.0 \times 10^{-21} B \qquad \text{watts} \tag{12.7}$$

A nomograph of this equation is shown in Fig. 12.4. A temperature of 300°K and a frequency band width of 10 kc is seen to give an available noise power of 4.1×10^{-16} watt. *It is convenient to remember that at room temperature the available noise power for a* 1*-mc band width is* 144 *db below* 1 *watt.*

When two resistances at different temperatures are connected in parallel, the mean-square noise voltage becomes

$$\bar{e}^2 = 4k \frac{R_1 R_2 (R_2 T_1 + R_1 T_2)}{(R_1 + R_2)^2} B \qquad \text{volts}^2 \tag{12.8}$$

where the resistances are in ohms and the temperatures in degrees Kelvin, and Boltzmann's constant is 1.3805×10^{-23} watt-sec. When several

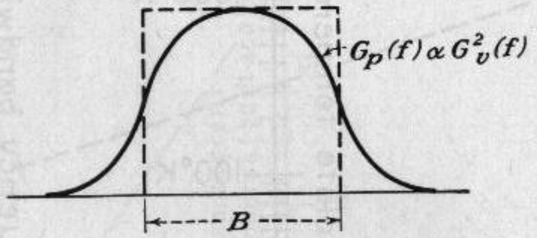


FIG. 12.5.—Definition of equivalent band width.

resistances are connected in parallel, the results are better expressed in terms of conductances. Let conductances $G_1, G_2, G_3, \ldots, G_n$ at temperatures $T_1, T_2, \ldots, T_n$, respectively, be connected in parallel. The resulting mean-square noise voltage is

$$\bar{e}^2 = 4k \frac{\sum_1^n G_n T_n}{\left(\sum_1^n G_n\right)^2} B \qquad \text{volts}^2 \tag{12.9}$$

Some care must be used in determining the band width to fit the above expressions. Since the concepts involved are basically those concerned with power, the equivalent band width of any device must be defined in terms of the power-frequency curve. In the case of an amplifier the band width is defined as that frequency interval for which a power gain equal to the gain at mid-band would transmit the same noise

energy as does the actual power-gain–frequency curve. Analytically, this becomes

$$B = \frac{\int_0^\infty G_p(f)\, df}{G_p(f_0)} \tag{12.10a}$$

where G_p is the power gain. All this amounts to is finding the width of a rectangular power-gain–frequency curve of height equal to the mid-frequency power gain of the actual curve as shown in Fig. 12.5. In terms of voltage amplification, the equivalent band width is

$$B = \frac{\int_0^\infty G_v^2(f)\, df}{G_v^2(f_0)} \tag{12.10b}$$

where G_v is the voltage gain.

12.3. Sources of Noise in Tubes. Noise can occur in vacuum tubes from a rather imposing list of sources. The principal sources of noise in tubes are

1. Shot effect (temperature-limited emission).
2. Reduced shot effect.
3. Flicker effect.
4. Collision ionization.
5. Random division of current between electrodes.
6. Induced noise at ultra-high frequencies.
7. Faulty tube construction.
   a. Hum.
   b. Poor insulation.
   c. Vibration.
   d. Varying wall charges.

Shot effect is the noise associated with random emission in a tube whose emission is temperature-limited. This is probably the loudest of the electronic tube noises but not the most serious, for tubes are seldom operated so that their emission is temperature-limited.

The so-called "reduced shot effect" is observed in tubes whose emission is space-charge-limited. The magnitude of the noise is much less than in tubes whose emission is temperature-limited. In this case the space charge exerts a smoothing action upon the true shot effect, and the noise is principally due to variations in the space-charge currents.

Flicker effect is observed with oxide cathodes. This effect is associated with variations in the activity of the emitting surface. The effect is much more noisy than the true shot effect for temperature-limited

emission. When the emission is space-charge-limited, the magnitude of the noise is greatly reduced.

Noise in tubes is raised by the presence of an appreciable amount of gas. This is due to the fact that gas molecules are ionized by collision with emitted electrons and the positive ions formed subsequently liberate little bursts of electrons as they penetrate the virtual cathode in front of the emitting surface. Gas noise is inappreciable unless the positive-ion gas current is more than a few hundredths of a microampere. Such gas noise appears mostly below 10 mc.

Random division of current between electrodes contributes to the noise of multielectrode tubes and makes pentodes three to five times as noisy as the same tube connected as a triode. It may be said that multielectrode tubes will always be noisier than triodes because of this additional factor contributing to the tube noise.

The ultra-high-frequency components of the random fluctuations of space charge in a tube will induce voltages in the grid circuit, which in turn will react back upon the space-charge flow. This effect is important only for frequency components above 30 mc.

Noise due to faulty tube construction is always present to a degree. If the filament is not sufficiently noninductive, hum will result. If insulation is poor at any point in the tube, there will be leakage currents, which will generally create noise because of nonconstant leakage resistance. Vibration may be a factor in an electromechanical feedback circuit. Dirt on the glass inside of a tube may give rise to varying wall charges, which will influence the tube current in a noisy manner. All these items can, however, with sufficient care in construction be held to a very low level.

Items 1, 2, 4, 5, and 6 listed above can never be removed entirely. They are, however, subject to an analysis that shows how their effects may be minimized. These items will be the subject of the subsequent sections. It has been found in most cases that it is convenient to express tube-noise effects in terms of equivalent noisy resistors. These resistors in turn are considered to have internal-noise emfs.

12.4. Shot Noise in Diodes with Temperature-limited Emission. Noise in diodes was probably the first form of tube noise ever detected. It is generally referred to as shot noise but also as "*Schrot* noise" and "Schottky noise," after the scientist who first analyzed the effect.[1,2] The noise is due to the random emission and arrival of electrons

[1] SCHOTTKY, W., Spontaneous Current Fluctuations in Electron Streams, *Ann. Physik*, vol. 57, pp. 541–567, Dec. 20, 1918.

[2] See also FRY, T. C., The Theory of the Schroteffekt, *Jour. Franklin Inst.*, vol. 199, February, 1925.

at the plate. It cannot be explained in terms of individual electron emission or arrival alone, for if the electrons were emitted at a uniform rate the lowest frequency component of noise would be above the highest frequency that vacuum tubes can handle. Thus a current of 1 ma corresponds to a flow of approximately 10^{16} electrons per sec. If these did flow at a uniform rate, there would be no noise components below 10^{16} cycles per sec. The electron stream evidently exhibits rather pronounced variations in density caused by the electrons arriving in groups. The mean square of the fluctuation components of current is

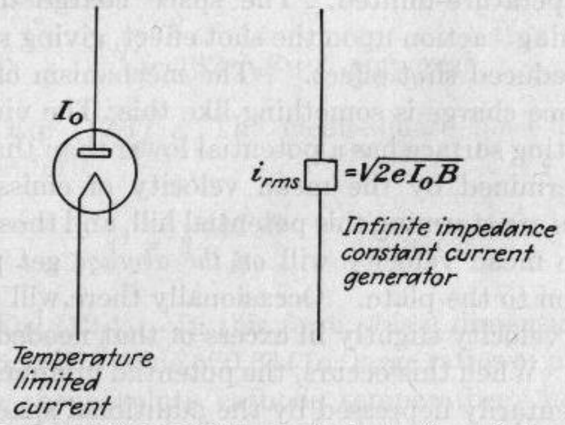


FIG. 12.6.—Constant-current-generator equivalent of a diode with temperature-limited current.

found to depend only upon the magnitude of the emitted current and the frequency band width

where e is electronic charge, 1.6020×10^{-19} coulomb

$$\bar{i}^2 = 2eI_0B \qquad \text{amperes}^2 \tag{12.11}$$

I_0 is emission current, amperes
B is band width, cycles per sec

This expression may more conveniently be written

$$\bar{i}^2 = 3.2041 \times 10^{-19} I_0 B \qquad \text{amperes}^2 \tag{12.12}$$

If the current from a diode with temperature-limited emission is put through a resistor R, the diode effectively puts a noise power of value $\bar{i}^2R$ into the resistor.

The above relations have been verified experimentally and they are reproducible to a high degree of accuracy. This property makes the diode with temperature-limited emission valuable as a standard noise source for such purposes as receiver and amplifier sensitivity measurement.

The diode with temperature-limited emission acts like a constant-

current generator as far as noise energy is concerned and will put a noise current given by the above equations through a resistor of any size. The equivalent circuit of the diode with temperature-limited emission is shown in Fig. 12.6.

12.5. Reduced Shot Effect in Diodes with Space-charge-limited Emission. In diodes in which the emission is space-charge-limited, the shot noise is much less than in the same diode passing the same current when its emission is temperature-limited.[1–4] The noise power is of the order of 10 per cent of that encountered for the same current when the emission is temperature-limited. The space charge thus has a very definite "smoothing" action upon the shot effect, giving rise to what may be called the "reduced shot effect." The mechanism of the smoothing action of the space charge is something like this: The virtual cathode in front of the emitting surface has a potential lower than that of the emitter by a value determined by the mean velocity of emission. Electrons with all velocities are storming this potential hill, and those with velocities greater than the mean velocity will *on the average* get past the virtual cathode and go on to the plate. Occasionally there will come a group of electrons with a velocity slightly in excess of that needed to get past the virtual cathode. When this occurs, the potential minimum at the virtual cathode is momentarily depressed by the additional space charge and as a result a few electrons that normally would have got past the potential minimum fail to do so and are returned to the emitter. This means that for every burst of electrons which might give rise to noise there is a compensating current set up in the opposite direction which tends to cancel the noise produced by the burst. The net result is an over-all reduction in noise that is considerable. The resulting noise levels are, however, high enough still to be of concern in the design of electronic equipment.

By considering the action of each increment of the velocities encountered in the process of emission some fairly good theoretical expressions for the reduced shot noise may be obtained. If the ratio of the noise

[1] RACK, A. J., Effect of Space Charge and Transit Time on the Shot Noise in Diodes, *Bell Sys. Tech. Jour.*, vol. 17, pp. 1–28, October, 1938.

[2] NORTH, D. O., Fluctuations in Space-charge-limited Currents at Moderately High Frequencies, *RCA Rev.*, vol. 4, Part II, pp. 441–473, April, 1940; vol. 5, pp. 244–260, July, 1940.

[3] WILLIAMS, F. C., Fluctuations of Space Charge Limited Currents in Diodes, *Jour. I.E.E.*, vol. 89, Part III, pp. 219–229, December, 1941.

[4] BELL, D. A., Fluctuations in Space Charge Limited Currents, *Jour. I.E.E.*, vol. 89, Part III, pp. 207–212, December, 1942.

See also SCHOTTKY, W., *Wiss. Veröffent. Siemens-Werken*, vol. 14, p. 15, 1937.

power of a diode passing a given current with and without space-charge limitation of current be designated by Γ^2, then, for a plane-electrode diode with a large ratio of plate voltage V_p to average emission voltage (velocity equivalent) V_e,

$$\Gamma^2 = \frac{9V_e}{V_p} \tag{12.13}$$

seems to be functionally correct.[1,2] The mean-square noise current can be expressed by an equation similar to that for the diode with temperature-limited emission as

$$\bar{i}^2 = \Gamma^2 2eI_0B \qquad \text{amperes}^2 \tag{12.14}$$

by analogy to Eq. (12.11). The mean-square noise current can also be written in the form

$$\bar{i}^2 = \theta \frac{4kT_cB}{R_1} \qquad \text{amperes}^2 \tag{12.15}$$

by analogy to Eq. (12.4). In this form, θ is a dimensionless parameter that has an asymptotic value of 0.644 for large ratios of plate to minimum potential, T_c is the absolute cathode temperature, and R_1 is the a-c diode resistance. Theoretically, the parameter θ is within a fraction of a per cent of the asymptotic value as long as the plate current is less than 80 per cent of the emission current and the plate voltage is greater than 2 volts for normal oxide operating temperatures.[3–5] Experimentally determined values of diode noise are 50 per cent higher than predicted by the theoretical expressions of Eqs. (12.13) and (12.15), so that θ assumes a value of unity for diodes. *The significance of Eq.* (12.15) *is that the noise power from a diode whose emission is space-charge-limited is the same as that from a resistor at the cathode temperature equal to the a-c diode resistance.*[6] The equivalent circuit for this case is given in Fig. 12.7.

[1] WILLIAMS, *op. cit.*

[2] BELL, *op. cit.*

[3] PEARSON, G. L., Shot Effect and Thermal Agitation in an Electron Current Limited by Space Charge, *Physics*, vol. 6, pp. 6–9, January, 1935.

[4] RACK, *op. cit.*

[5] NORTH, *op. cit.*

[6] Combination of Eqs. (12.14) and (12.15) and the use of the fact that $R_1 = \frac{2}{3}\frac{V_p}{I_0}$ for the diode with space charge suggests that Γ^2 has the value $(0.644)\frac{3kT_c}{eV_p}$, which is consistant with the observed relation of Eq. (12.13).

12.6. Reduced Shot Effect in Triodes with Space-charge-limited Current. Triodes, too, are noisy. In fact, the noise in triodes, as pointed out before, may be a limiting factor in the sensitivity of an electronic device. The noise is due to the effects just observed in diodes and shows

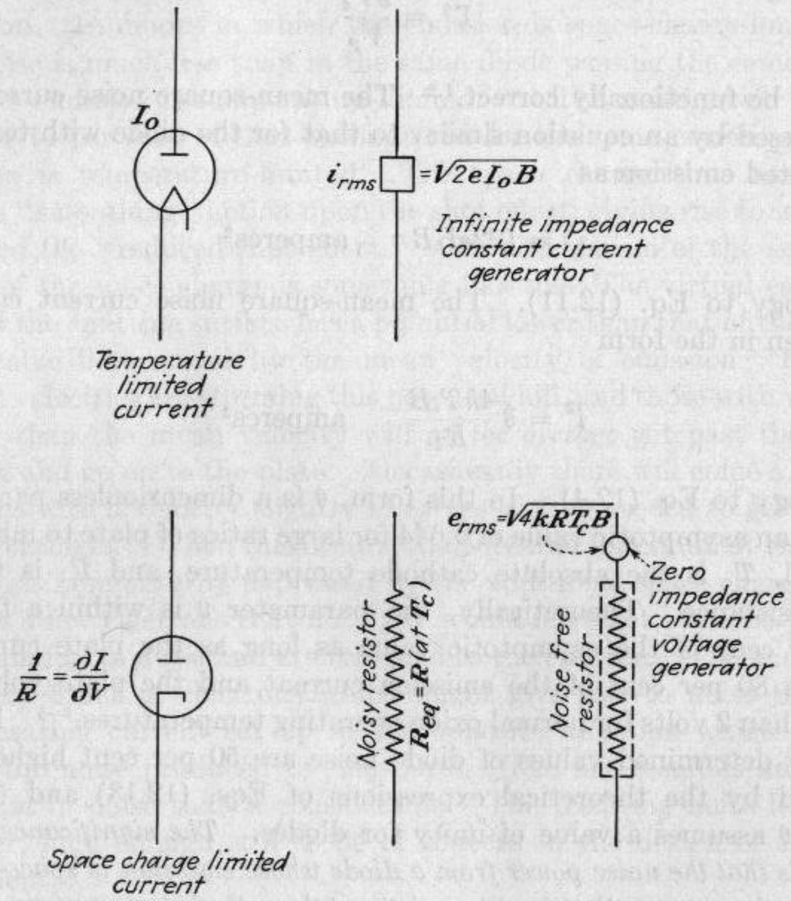


FIG. 12.7.—Resistor equivalent to a diode with space-charge-limited current.

as a fluctuation in plate current. It is convenient to interpret the noise in a triode as being due to a noisy resistor in series with the grid of the tube considered free from noise. The value of this noisy resistor in series with the grid is

$$R_{eq} = \frac{\theta T_c G}{G_m T_r G_m} \quad \text{ohms} \tag{12.16}$$

where θ is the effective cathode-temperature ratio of approximate value two-thirds, G_m is the mutual conductance of the triode, G is the a-c conductance of the diode equivalent of the triode, T_c is absolute cathode temperature, and T_r is absolute room temperature. The ratio of equivalent-diode conductance to triode mutual conductance is the inverse-square ratio of equivalent-diode spacing to the cathode-grid spacing as discussed in Chaps. 7 and 8. For the plane-electrode diode, from Eq. (8.45), this ratio is

$$\frac{G}{G_m} = \left[1 + \frac{1}{\mu}\left(\frac{d_{cp}}{d_{cg}}\right)^{4/3}\right]^{3/2} \tag{12.17}$$

The corresponding ratio for cylindrical triodes is difficult to express

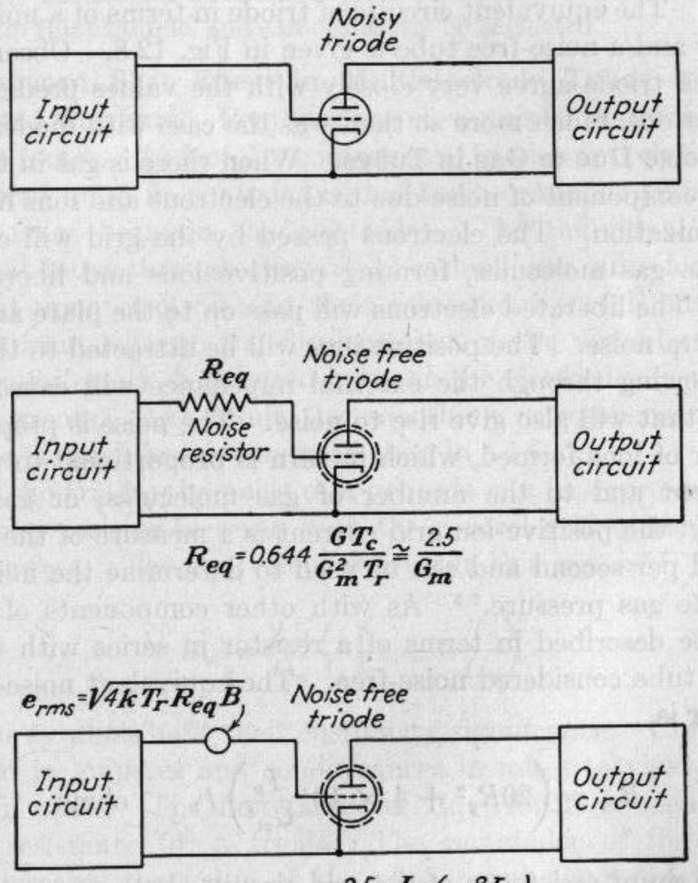


FIG. 12.8.—Equivalent circuit of a noisy triode.

exactly but is given approximately by

$$\frac{G}{G_m} = 1 + \frac{1}{\mu}\left[1 + \frac{2}{3}\ln\frac{r_p}{r_g}\right] \tag{12.18}$$

The value of this ratio will generally lie between 1 and 2. Assuming that a typical value of the conductance ratio is 1.25 and that the cathode

temperature is 3.33 times the room temperature, Eq. (12.16) reduces to

$$R_{eq} = \frac{2.5}{G_m} \quad \text{ohms} \tag{12.19}$$

which is sufficiently accurate for a good many purposes.[1] The value of equivalent resistance thus given is that resistance which if inserted in series with the grid of the triode considered noise-free would cause as much noise current in the plate circuit of the tube as does the reduced shot effect. The equivalent circuits of triode in terms of a noise-generating resistor and a noise-free tube is given in Fig. 12.8. Observed values of noise in a triode agree very closely with the values predicted by the above equations, much more so than was the case with diodes.

12.7. Noise Due to Gas in Tubes. When there is gas in tubes, there is an extra component of noise due to the electrons and ions liberated by collision ionization. The electrons passed by the grid will collide with some of the gas molecules, forming positive ions and liberating more electrons. The liberated electrons will pass on to the plate and give rise to some extra noise. The positive ions will be attracted to the negative grid and flowing through the external impedance will cause a voltage in the grid that will also give rise to noise. The noise is proportional to the number of ions formed, which in turn is proportional to the normal space current and to the number of gas molecules, or gas pressure. Fortunately, the positive-ion grid current is a measure of the number of ions formed per second and can be used to determine the noise without knowing the gas pressure.[2,3] As with other components of noise, the noise can be described in terms of a resistor in series with the control grid of the tube considered noise-free. The equivalent noise-generating grid resistor is

$$R_{eq} = \left(20R_g^2 + 4 \times 10^4 \frac{I_p}{G_m^3}\right) I_1 \quad \text{ohms} \tag{12.20}$$

where R_g is shunt resistance of the grid circuit, ohms
G_m is mutual conductance, mhos
I_p is plate current, amperes
I_1 is control-grid current, amperes

[1] HARRIS, W. A., Space Charge Limited Current Fluctuations in Vacuum Tube Amplifiers and Input Systems, *RCA Rev.*, vol. 5, pp. 505–524, April, 1941; vol. 6, pp. 114–124, July, 1941.

[2] *Ibid.*

[3] THOMPSON, B. J., and D. O. NORTH, Fluctuations in Space-charge-limited Currents Caused by Collision Ionization, *RCA Rev.*, vol. 5, pp. 371–388, January, 1941.

The first term in this expression is due to the flow of positive-ion current through the external grid impedance. The second term is due to the electrons liberated upon ionization that are attracted to the plate.

As an example, consider the case of a gassy tube for which the positive-ion control-grid current is 0.01 microampere. Let the mutual conductance of the tube be $5{,}000 \times 10^{-6}$ mho, the plate current 1 ma, and the shunt resistance of the grid circuit 100,000 ohms. Then the first term of Eq. (12.20) contributes a noise resistance of 2,000 ohms, and the second term contributes a noise resistance of 3.20 ohms. The second term of Eq. (12.20) is usually much smaller than the first term, as in this example, and can ordinarily be neglected.

12.8. Reduced Shot Effect in Multielectrode Tubes with Space-charge-limited Currents. Pentodes are even noisier than triodes—by a considerable factor. In fact, by comparison, triodes are relatively quiet. The additional noise in pentodes is due to the random division of the fluctuation noise between the electrodes. The individual groups of electrons that burst through the virtual cathode are quite local in their impingement upon electrodes in their subsequent travel, but the compensating currents due to the displacement of the virtual cathode are more or less uniformly distributed between the electrodes in the ratio of the direct currents. As a result, the smoothing, or compensating, action in a pentode is much less pronounced than in a triode.

As with the triode, the noise of a pentode can be expressed as being due to a noisy resistor in series with the control grid of the tube considered noise-free. The equivalent resistor in this case is[1,2]

$$R_{eq} = \frac{2.5}{G_m} \frac{I_p}{I_s} \left(1 + 8 \frac{I_2}{G_m}\right) \tag{12.21}$$

where all the symbols have their customary significance. Currents must be expressed in amperes and conductances in mhos to yield equivalent resistance in ohms. The first factor of Eq. (12.21) is seen to be the equivalent resistance for a triode. The remainder of the expression generally increases the value of the triode equivalent resistance by a factor of three to five. For screen currents much smaller than the plate current the noise in both the plate and screen circuits is approximately equal to the true shot effect for a current equal to the screen current. The value of the mean-square noise current in the plate circuit is readily obtained from this by means of Eq. (12.12) and converted to an equivalent noisy resistor *in the plate circuit* by means of the nomograph of Fig. 12.3.

[1] NORTH, D. O., Fluctuations in Space Charge Limited Currents in Multi-collectors, *RCA Rev.*, vol. 5, pp. 244–260, October, 1940.

[2] HARRIS, *op. cit.*

12.9. Noise in Mixer Tubes. Mixers are also noisy. The factors that contribute to the noise of triodes and pentodes also contribute to mixer noise. In a mixer a large voltage from the local oscillator is applied to the tube so that the current and the mutual conductance swing over a large value. Since the noise of tubes is known as a function of current and mutual conductance, it is not too difficult to evaluate the mixer noise.

In a mixer it is the noise in the intermediate-frequency band that is of importance. The noise in this band will vary periodically over the local oscillator cycle as the mutual conductance and current vary with the voltage applied at the local oscillator frequency. The total intermediate-frequency noise can be obtained by summing the plate noise over the local oscillator cycle.

$$\bar{i}_{i-f}^2 = \frac{1}{2\pi}\int_0^{2\pi} \bar{i}_{pn}^2\, d(\omega t) \tag{12.22}$$

where $\bar{i}_{pn}^2$ is the mean-square noise current in the plate at any instant t and ω is 2π times the local oscillator frequency. It is convenient (except in the case of the diode mixer) to express the mixer noise in terms of an equivalent noise resistance in series with the control grid of the tube considered as noise-free. Thus

$$\bar{e}_n^2 = \frac{\bar{i}_{i-f}^2}{G_c^2} \tag{12.23}$$

where G_c is the conversion transconductance of the tube. Corresponding to this value of input noise, the equivalent input noisy resistor is

$$R_{eq} = \frac{\bar{e}_n^2}{4kT_rB} \tag{12.24}$$

$$R_{eq} = \frac{\bar{i}_{i-f}^2}{4kT_rG_c^2B} \tag{12.25}$$

Upon applying the above ideas there are obtained the results shown in Fig. 12.9 for a fictitious pentode tube connected in accordance with the four most common mixer connections.[1] In this figure, G_0 is the maximum value which the mutual conductance assumes over the local oscillator cycle, generally that corresponding to zero grid voltage. The quantity G_x is the maximum value of the screen-plate transconductance. The quantity I_x is the maximum value of plate current when the signal is

[1] HEROLD, E. W., Superheterodyne Converter Considerations in Television Receivers, *RCA Rev.*, vol. 4, pp. 324–337, January, 1940.

See also the summarizing articles by HEROLD, E. W., and L. MALTER, Some Aspects of Radio Reception at Ultra-high Frequency, *Proc. I.R.E.*, vol. 31, pp. 423–438, 491–510, 567–582.

injected into the screen circuit and the local oscillator into the control-grid circuit.

It is seen that the triode is the best converter and the pentode with signal and local oscillator applied to the control grid is the next best. The others are too noisy for high-sensitivity applications.

TYPE OF FIRST DETECTOR	TYPICAL CHARACTERISTICS	APPROX. CONVERSION TRANSCOND.	APPROX. EQUIV. NOISE RESIST.	EQUIVALENT NOISE MICROVOLTS ON GRID*
Sig. + osc.	G_o, G_m, V_g / I_o, I_b, V_g	$0.28G_o$	$\frac{\Pi}{G_o}$	8
Sig. + osc.	G_o, $0.8G_o$, $G_{m1\text{-}p}$, V_1 / I_o, $0.8I_o$, I_b, V_1	$0.23G_o$	$\frac{15+21\frac{I_o}{G_o}}{G_o}$	16
Sig. Osc.	G_o, $0.5G_o$, $G_{m1\text{-}p}$, V_3 / I_o, $0.5I_o$, I_b, V_3	$0.14G_o$	$\frac{120\frac{I_o}{G_o}}{G_o}$	32
Sig. Osc.	G_x, $G_{m3\text{-}p}$, V_1 / I_x, I_b, V_1	$0.28G_x$	$\frac{57\frac{I_x}{G_x}}{G_x}$	—

* Assuming $G_o = 15 \times 10^{-3}$ mhos, $I_o = 30 \times 10^{-3}$ amps, $\Delta f = 4$ mc

FIG. 12.9.—Comparison of fictitious modulators presumed to have similar cathode and first-grid structures. (*Herold.*)

When the values required in the tabulation of Fig. 12.9 are not available, the following formulas will be found sufficiently accurate for most purposes:

For triode mixers,

$$R_{eq} = \frac{4}{G_c} \tag{12.26}$$

or

$$R_{eq} = \frac{2.5\bar{G}_m}{G_c^2} \tag{12.27}$$

For pentode mixers

$$R_{eq} = \frac{I_p}{I_s}\left(\frac{4}{G_c} + \frac{20I_2}{G_c^2}\right) \tag{12.28}$$

or

$$R_{eq} = \frac{I_p}{I_s}\left(\frac{2.5\bar{G}_m}{G_c^2} + \frac{20I_2}{G_c^2}\right) \tag{12.29}$$

where currents must be expressed in amperes and conductances in mhos. $\bar{G}_m$ is the mean transconductance over the local oscillator cycle, and G_c is the conversion transconductance of the local oscillator. In addition, the following rules of thumb may be applied,

$$G_c \text{ (as converter)} = \tfrac{1}{4}G_m \text{ (as amplifier)} \tag{12.30}$$
$$I_p \text{ (as converter)} = \tfrac{1}{4}I_p \text{ (as amplifier)} \tag{12.31}$$
$$I_2 \text{ (as converter)} = \tfrac{1}{4}I_2 \text{ (as amplifier)} \tag{12.32}$$

and, for pentodes only,

$$R_{eq} \text{ (as converter)} = 4R_{eq} \text{ (as amplifier)} \tag{12.33}$$

in which the values "as amplifier" refer to the peak of the local oscillator cycle.[1]

12.10. Noise Induced at Ultra-high Frequencies by Random Emission. At ultra-high frequencies there is a conductive component to the input admittance of a tube. The finite transit time of the electrons makes it possible for the grid to transfer energy to the electrons as they are accelerated.[2] There is a separate component of noise associated with this effect.[3,4] The high-frequency components in the tube noise induce currents on the grid, which in turn influence the electron current, thus giving rise to an extra component of noise. The equivalent input noise conductance is found to be the same as the input conductance, but at approximately five times room temperature. This extra component of noise may be represented as a constant-current generator across the electronic component of grid conductance across the input circuit. To a first order of approximation the induced noise is independent of the normal noise component, which may be added in series with the grid as

[1] HARRIS, *op. cit.*

See also articles by HEROLD, *Proc. I.R.E.*, for some typical values.

[2] FERRIS, W. R., Input Resistance of Tubes as Ultra-high Frequency Amplifiers, *Proc. I.R.E.*, vol. 24, pp. 82–107, January, 1936.

[3] NORTH, D. O., and W. R. FERRIS, Fluctuations Induced in Vacuum Tube Grids at High Frequencies, *Proc. I.R.E.*, vol. 29, pp. 49–50, February, 1941.

[4] BALLANTINE, STUART, Schrot Effect in High Frequency Circuits, *Jour. Franklin Inst.*, vol. 206, pp. 159–168 August, 1928.

has been done before. The equivalent circuit for this effect is shown in Fig. 12.10. At ultra-high frequencies there is a component of grid-input conductance due to feedback through the cathode-lead inductance, as well as the component due to electron-transit-time effects. Both components vary as the square of frequency and thus are hard to separate. The feedback component has noise associated with it too, but as a resistor at room temperature.

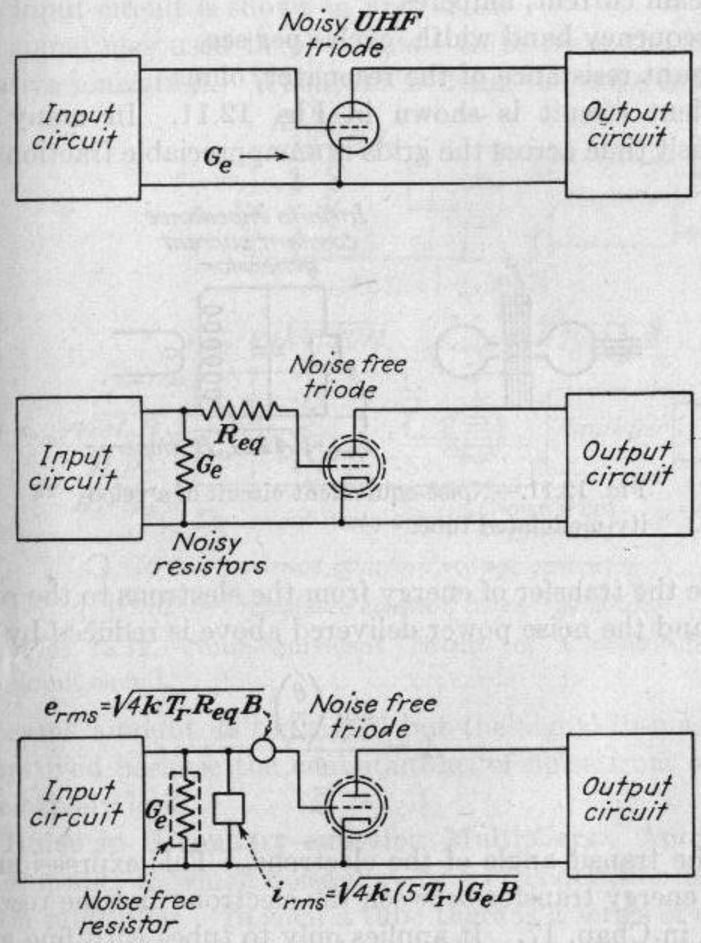


FIG. 12.10.—Equivalent circuit for induced noise in a tube at ultra-high frequencies.

In general, the noise of amplifiers does not change much with feedback. This is because the amplification and input impedance are ordinarily changed by the same factor.

12.11. Noise in Velocity-modulation Tubes. In a velocity-modulation tube a beam of electrons passes through the two grids of a resonator, between which there appears a high shunt resistance. Noise is generated, but there is evidently very little space-charge smoothing action in

this case. This is expected from the fact that there is no virtual cathode between the grids of the resonator. As a result, the noise is nearly the full shot noise of the beam, and therefore the noise power delivered to the resonator is given closely by

$$N = 2eI_0BR_{sh} \quad \text{watts} \tag{12.34}$$

where e is charge of the electron, 1.6020×10^{-19} coulomb
I_0 is beam current, amperes
B is frequency band width, cycles per sec
R_{sh} is shunt resistance of the resonator, ohms

The equivalent circuit is shown in Fig. 12.11. In many cases the electron transit time across the grids is an appreciable fraction of a cycle,

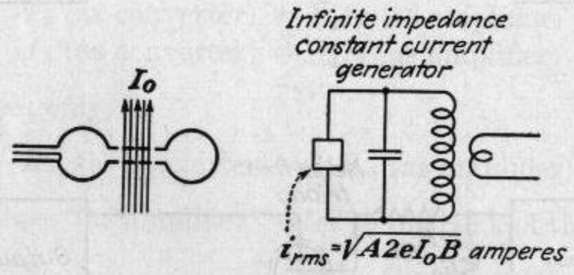


FIG. 12.11.—Noise-equivalent circuit of a velocity-modulated tube.

in which case the transfer of energy from the electrons to the resonator is not perfect and the noise power delivered above is reduced by the factor

$$A = \frac{\sin\left(\frac{\theta}{2}\right)}{\frac{\theta}{2}} \tag{12.35}$$

where θ is the transit angle of the electrons. This expression gives the efficiency of energy transfer between the electrons and the resonator and is developed in Chap. 17. It applies only to tubes with fine grids.

12.12. Noise in Phototubes. High-vacuum phototubes produce the same noise as do hot-cathode tubes with temperature-limited emission. This is true shot effect, giving rise to a mean-square fluctuation current of the value

$$\bar{i}_n^2 = 2eI_0B = 3.2040 \times 10^{-19}I_0B \quad \text{amperes}^2 \tag{12.36}$$

This is the same value of current as is produced by a resistance at room temperature of value

$$R_{eq} = \tfrac{1}{2}\tfrac{0}{I_0} \quad \text{ohms} \tag{12.37}$$

The noise may be represented by a constant-current generator in parallel with the tube considered noise-free. The noise from the tube will often be of the same order as that produced by the large-value resistor used to develop voltage. Thus, if a 10-megohm resistor is used to develop voltage, the noise from the tube will be the same as the noise from the resistor when the current is 0.005 microampere. At higher currents the phototube noise predominates. The equivalent circuit of a typical phototube input circuit is shown in Fig. 12.12.

Gas is sometimes used in phototubes to increase the plate current by cumulative ionization. When this is done, the noise is increased by

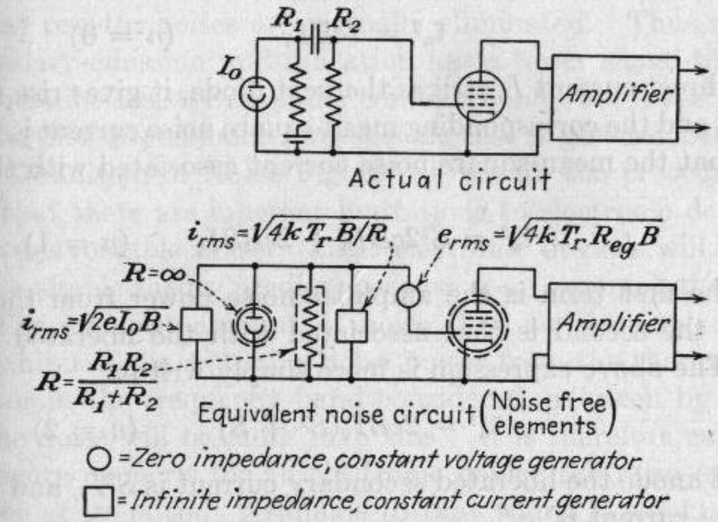


FIG. 12.12.—Noise-equivalent circuit of a phototube input circuit.

about the same amount as the signal but the signal-to-noise ratio is in general improved because the contributions of noise from other sources becomes relatively less.

12.13. Noise in Secondary-emission Multipliers. Another type of electronic amplifier in which noise may be a consideration is the secondary-emission multiplier. In such a tube there is a series of electrodes at successively higher potentials, each coated with a material that emits a large ratio of secondary to primary electrons. Electrons that strike the first anode give rise to S times as many electrons, which are attracted to the second anode, where they give rise to S times as many as the striking electrons, or S^2 times as many as struck the first anode. After n such impacts the current is S^n times as great as it was originally, S being the ratio of secondary to primary electrons—a number that can be made as high as 9 or 10. It might be thought that with such a system tremendous amplifications could be obtained. True, they can, but no improvement in signal-to-noise ratio can be achieved.

Experimental results indicate that the following laws govern the noise associated with secondary emission in such a device:

1. Secondary emission from any anode follows the shot-effect law, $\bar{i}_n^2 = 2eBI_0$, where I_0 is the emitted secondary current.
2. Shot noise from any anode is multiplied by subsequent stages in the same way as the signal is.[1]

Consider the action in a few successive stages of secondary-emission amplification. Let the current from a first anode be I_0; then the mean-square noise current associated with this is

$$\bar{i}_n^2 = 2eBI_0 \qquad (n = 0) \qquad (12.38)$$

When the direct current I_0 strikes the next diode, it gives rise to a direct current SI_0 and the corresponding mean-square noise current is multiplied by S^2, so that the mean-square noise current associated with the current SI_0 is

$$\bar{i}_n^2 = S^2 2eBI_0 + 2eBSI_0 \qquad (n = 1) \qquad (12.39)$$

in which the first term is the amplifier noise power from the previous anode and the second is that associated with the liberated secondary current. The above expression is more simply written

$$\bar{i}_n^2 = 2eBI_0(S^2 + S) \qquad (n = 2) \qquad (12.40)$$

At the next anode the liberated secondary current is S^2I_0, and the mean-square noise current is

$$\bar{i}_n^2 = S^2 2eBI_0(S^2 + S) + 2eBS^2I_0 \quad (n = 2) \qquad (12.41)$$

which is equal to

$$\bar{i}_n^2 = 2eBI_0(S^4 + S^3 + S^2) \qquad (n = 2) \qquad (12.42)$$

Extension of this process to n stages yields

$$\bar{i}_n^2 = 2eBI_0 \frac{S^n(S^{n+1} - 1)}{S - 1} \qquad (n = n) \qquad (12.43)$$

This means that the ratio of output to input noise power is

$$\frac{\text{Noise power out}}{\text{Noise power in}} = \frac{S^n(S^{n+1} - 1)}{S - 1} \qquad (12.44)$$

Correspondingly, the ratio of signal power out to signal power in is

$$\frac{\text{Signal power out}}{\text{Signal power in}} = S^{2n} \qquad (12.45)$$

[1] ZWORYKIN, V. K., G. A. MORTON, and L. MALTER, The Secondary Emission Multiplier, *Proc. I.R.E.*, vol. 24, pp. 351–375, March, 1936.

Upon taking the quotient of the last two equations, the relative change in the ratio of signal to noise power is

$$\frac{\text{Signal-power-to-noise ratio in}}{\text{Signal-power-to-noise ratio out}} = \frac{S^n(S^{n+1} - 1)}{S^{2n}(S - 1)} \qquad (12.46)$$

This ratio is slightly greater than 1 but approaches this value as the secondary-emission ratio S is increased. This simply means that signal and noise are amplified about the same amount in a secondary-emission multiplier, and as a result there is no gain on the signal-to-noise ratio. There is, however, an advantage to using secondary-emission multiplication in that resistor noises are virtually eliminated. Thus a phototube with secondary-emission multiplication has a lower signal-to-noise ratio than a phototube-resistor-amplifier combination at low levels of illumination. A further discussion of this specific case is given in Sec. 19.20.

12.14. Definition of Noise Figure. From all the preceding sections it is seen that there are inherent limitations to electronic devices determined by unavoidable noise. Most electronic devices will, in fact, be noisier than simple theory predicts because of an accumulation of various effects. The smallest amount of noise that an electronic device can possibly exhibit is the available noise power from the thermal agitation of a resistor in the frequency band considered, as given by Eq. (12.6). Usually the noise will be more than this. It is therefore convenient to use as a figure of merit for an electronic device the ratio of the actual noise power at its output terminals to that which it would have if the noise were limited to the minimum noise from thermal agitation. This figure of merit is called the *noise figure* of the device. Basically, the *noise figure is an excess noise ratio.*

A rigorous definition of noise figure involves a consideration of the gain of the device and the available input noise power and the output power. The gain of a device, invariably a four-terminal network, is defined as the ratio of the available signal power at the output to the available signal power at the output of the signal generator,

$$\text{Power gain} = \frac{S_{\text{out}}}{S_{\text{in}}} = G \qquad (12.47)$$

This definition of gain is independent of the output impedance of the device, but it does depend upon the impedance of the signal generator, which is taken as the nominal input impedance of the device. In many applications there is no difficulty in using a signal-generator impedance that is equal to the input impedance of the device under consideration, but in some applications, such as extremely broad-band devices, it is extremely difficult to maintain a constant input impedance. Hence

it is logical that the figure of merit should include nonconstancy of the input impedance. The available noise power between two terminals is defined as the noise power that would be absorbed by a matched output circuit. The available input noise power is simply that given from the Johnson noise formula of Eq. (12.6),

$$N_{\text{in}} = kTB \tag{12.48}$$

The noise figure is defined in terms of the factor of most importance in the ultimate sensitivity of an electronic device, the ratio of output signal to noise power. The maximum value this can have is the ratio of available input signal to noise power if there are no other noise sources in the device and if all impedances are properly matched. Some four-terminal networks that consist of passive elements only, say a transformer or a transmission line, have no noise sources present in them, but electronic devices always have some extra sources of noise. *The noise figure F of the device is defined as the ratio of the available signal-to-noise ratio at the signal-generator (input) terminals to the available signal-to-noise ratio at the output terminals*[1,2]

$$F = \frac{\dfrac{S_{\text{in}}}{kTB}}{\dfrac{S_{\text{out}}}{N_{\text{out}}}} \tag{12.49}$$

The noise figure of an electronic device is always greater than unity. The reference temperature is invariably taken as 290°K (63°F).

Some rearrangements of Eq. (12.49) are useful. Upon utilizing the power-gain definition of Eq. (12.47) the noise figure may be written

$$F = \frac{N_{\text{out}}}{GkTB} \tag{12.50}$$

From this the available noise output power is

$$N_{\text{out}} = FGkTB \quad \text{watts} \tag{12.51}$$

The available output noise due only to noise sources in the network is

$$N_{\text{out}} - GkTB = (F - 1)GkTB \quad \text{watts} \tag{12.52}$$

Example: It is desired to calculate the noise figure of a 250-ohm-input 30-mc intermediate-frequency amplifier having a band width of 2 mc and using 6AC7

[1] FRIIS, H. T., Noise Figures of Radio Receivers, *Proc. I.R.E.*, vol. 32, pp. 419–422, July, 1944.

[2] NORTH, D. O., Absolute Sensitivity of Radio Receivers, *RCA Rev.*, vol. 6, pp. 332–343, January, 1942.

pentodes. Coupling between stages is provided with two inductively coupled tuned circuits. Let the tube constants be

$$\begin{aligned} I_p &= 10 \times 10^{-3} && \text{ampere} \\ I_2 &= 2.5 \times 10^{-3} && \text{ampere} \\ G_m &= 9 \times 10^{-3} && \text{mho} \end{aligned}$$

Then, from Eq. (12.21), the equivalent noisy resistor in series with the control grid is

$$\begin{aligned} R_{\text{eq}} &= \frac{2.5}{9 \times 10^{-3}} \frac{10 \times 10^{-3}}{12.5 \times 10^{-3}} \left(1 + \frac{8 \times 2.5 \times 10^{-3}}{9 \times 10^{-3}}\right) \\ &= 715 \quad \text{ohms} \end{aligned}$$

From Fig. 12.1 this corresponds to an rms noise voltage of 4.85 microvolts. Since the noise is developed across the input impedance in series with the equivalent noise resistor, while the signal is developed only across the input impedance, the noise figure is

$$F = \frac{715 + 250}{250} = 3.86$$

In cases in which the stage amplification is low the effect of the noise from the following stage must be considered. In this case consider that the coupling network is a unity-ratio impedance transformer with an input impedance of 50,000 ohms. If the second stage is identical to the first in operating characteristics, then there is a noise voltage of 4.85 microvolts developed in series with its grid. This voltage corresponds to a current of 97.0 micromicroamperes through the plate load of the first stage. Upon referring this current back to the control grid of the first stage, it corresponds to a voltage of $\frac{97 \times 10^{-12}}{9 \times 10^{-3}}$, or 0.01075 microvolt. This is small compared with the 4.85 microvolts due to the first tube and so can be neglected in this case.

Noise Figure for Two Networks in Cascade. Usually the sources contributing to the excess noise in an electronic device are principally in the input circuit of the device. However, when the first stage of amplification has insufficient gain, the noise sources effectively located in the input circuit of the second stage of amplification contribute appreciably to the over-all noise as well. Noise from subsequent stages is generally so small compared with the amplified noise from the earlier extra sources that it may be neglected. When the condition cited above is the case, then the following relations hold: Let the first and second stages of amplification be designated by subscripts 1 and 2, respectively. Let over-all characteristics be designated by the subscript 12. The over-all power gain G_{12} is equal to the product of the gains of the first and second stages, G_1G_2. Let the band width be that of the over-all charac-

teristic as defined by Eq. (12.9). The available noise power at the output of the first stage of amplification is

$$N_1 = F_1 G_1 kTB \quad \text{watts} \tag{12.53}$$

from Eq. (12.51). The noise power in the output of the second stage due only to sources in that stage is, from Eq. (12.52),

$$N_2 - G_2 kTB = (F_2 - 1)G_2 kTB \tag{12.54}$$

The total available noise power at the output of the second stage is G_2 times the quantity in Eq. (12.53) plus the quantity in Eq. (12.54),

$$N_{12} = G_2 F_1 G_1 kTB + (F_2 - 1)G_2 kTB \tag{12.55}$$

or

$$N_{12} = \left[F_1 + \left(\frac{F_2 - 1}{G_1}\right)\right] G_1 G_2 kTB \tag{12.56}$$

Now, upon applying the general formula [Eq. (12.51)] to the over-all situation,

$$N_{12} = F_{12} G_{12} kTB \tag{12.57}$$

The values of available output noise over-all, from Eqs. (12.56) and (12.57), must be equal; therefore

$$F_{12} = F_1 + \frac{F_2 - 1}{G_1} \tag{12.58}$$

This is the important relation that has been sought. It gives the over-all noise figure of two amplifiers in terms of the separate noise figures and the gain of the first amplifier. Needless to say, the expression is not limited to amplifiers but may be applied to mixers and networks in general.

Sometimes noise figures of circuit elements are expressed in terms of equivalent temperatures, simply the room temperature multiplied by the noise figure. The symbol t is often used to designate the ratio of the actual noise to the available thermal-agitation noise of Eq. (12.6). This is most frequently done with passive elements that produce noise, such as crystal detectors. The noise figure of a crystal input receiver is

$$F = L_x(t_x - 1 + F_a) \tag{12.59}$$

where L_x is the conversion loss of the crystal detector corresponding to the reciprocal of the conversion gain of a tube mixer, t_x is the equivalent temperature ratio of the crystal, and F_a is the noise figure of the amplifier into which the output of the crystal detector is fed. If the crystal were noiseless, t_x would equal unity.

The design of electronic equipment with regard to obtaining a low noise figure is a rather complex problem, which will only be touched upon here.[1] The problem is a combination one, involving consideration of circuits and noise sources. It is possible in some cases that the minimum noise figure will be obtained with a condition of mismatched impedances owing to the fact that a mismatch will reduce the noise more than it will the signal. With receivers it is found practical to use r-f amplification before the mixer only up to a certain frequency. This frequency is of the order of 600 mc at this time. Beyond this frequency the noise figure of amplifiers is so great that the signal-to-noise ratio is increased rather than reduced upon amplification. There is also an upper frequency limit at which vacuum-tube mixers are practical. At present, this limit occurs at about 1,000 mc. It is quite possible that these frequency limits will be extended with time. The limits occur because the effective mutual conductance of all tubes decreases with increasing frequency owing to transit-time effects, thus raising the noise. The ultra-high-frequency induced noise effect also contributes to increasing noise.

12.15. Measurement of Noise and Noise Figure. It is not difficult to measure the noise and noise figure of an electronic device when its gain is sufficiently high so that an appreciable output noise power is obtained. Square-law devices such as thermocouples are preferred in noise measurements because the output indication is directly proportional to power. Crystal detectors may also be used at low levels of power. If the rectified crystal current is kept below a few microamperes, the crystal is almost certain to be a square-law device and as such is extremely sensitive. In this case the rectified crystal current is proportional to the input power. If a satisfactory power-output indicating device is used, the noise figure of a device can be measured by simply introducing signal input power until the output indication from the noise alone is doubled. The input power is then equal to that generated by the internal noise sources, and the noise figure is given by

$$F = \frac{S_{in}}{kTB} \tag{12.60}$$

where S_{in} is the signal power in. For any other adjustment of signal input power the noise figure is given directly from the defining relation of Eq. (12.49), it being remembered that the output-power indication is the sum of the noise and signal power out.

Diodes operating with temperature-limited emission may be used as a standard source of noise. Such diodes preferably have either a tungsten

[1] See the survey articles by HEROLD and MALTER, *op. cit.*

or a thoriated tungsten filament. It is difficult to keep the emitted current from an oxide-emitting surface constant under temperature-limited conditions of emission. The noise, of course, is due to shot effect, with the mean-square noise current given by Eq. (12.12). Diodes may successfully be operated as a standard noise source up to 100 mc. Beyond that frequency the impedance transformation introduced by the leads cannot be determined very accurately. Undoubtedly, special diodes can be built for noise measurements at higher frequencies. A standard noise source for measurements of the noise figure of an intermediate-frequency amplifier would consist of the diode in a shielded can with leads brought in through properly by-passed chokes. Across the diode there can be placed a coil that tunes the capacity of the diode to the center of the band of interest. The tuned circuit thus formed should be shunted by the nominal input resistance of the amplifier to be tested. Output leads are then brought from across the tuned circuit to the amplifier under test. Let the nominal input resistance of the amplifier be R; then the noise power delivered to the resistor of value R shunted across the diode is $2eI_0RB$, and the available power into the receiver is one-fourth of this. Let the diode current I_0 be adjusted until the normal noise output power of the receiver is doubled when the standard diode noise is connected to its input. Under these conditions the noise figure, from Eq. (12.60), is given by $\frac{eI_0RB}{2kTB}$; or since $\frac{e}{2kT}$ has the numerical value of 20, the noise figure is given simply by

$$F = 20I_0R \tag{12.61}$$

where I_0 is the diode current in amperes and R is the nominal resistance of the receiver in ohms.

12.16. Typical Tube-noise Values. In the table on the next page are given some typical operating conditions and associated noise values of representative triodes, pentodes, and mixers.

TABLE VI
TUBE-NOISE VALUES*

Type	Application	Voltages			Currents			Trans-conductance, micromhos	Noise-equivalent resistance		Noise-equivalent input, voltage** microvolts
		Plate volts	Screen volts	Bias volts	Plate ma	Screen ma	Cathode ma		Calculated ohms	Measured ohms	
6SK7	Pentode amplifier	250	100	−3	9.2	2.4	11.6	2,000	10,500	9,400–11,500	0.94
6SJ7	Pentode amplifier	250	100	−3	3	0.8	3.8	1,650	5,800	5,800	0.70
6SG7	Pentode amplifier	250	125	−1	11.8	4.4	16.2	4,700	3,300	.....	0.53
6AC7/1852	Pentode amplifier	300	150	−2	10	2.5	12.5	9,000	720	600–760	0.25
956	Pentode amplifier	250	100	−3	5.5	1.8	7.3	1,800	9,400	.....	0.90
1T4	Pentode amplifier	90	45	0	2.0	0.65	2.65	750	20,000	.....	1.3
6SA7	Frequency converter	250	100	0	3.4	8.0	11.9	450§	240,000	210,000	4.5
6K8	Frequency converter	250	100	−3	2.5	6.0	8.5‡	350§	290,000	.....	4.9
1R5	Frequency converter	90	45	0	0.8	1.8	2.75	250§	170,000	.....	3.8
6L7	Pentagrid mixer	250	100	−3	2.4	7.1	9.5	375§	255,000	210,000	4.6
6J5	Triode amplifier	250	...	−8	9.0	...	.....	2,600	960	1,250	0.28
955	Triode amplifier	180	...	−5	4.5	...	.....	2,000	1,250	.....	0.32
6AC7/1852	Triode amplifier	150	150	−2	...	...	12.5	11,200	220	200	0.14
6AC7/1852	Pentode mixer	300	150	−1†	5.2	1.3	6.5	3,400§	2,750	3,000	0.48
6SG7	Pentode mixer	250	125	−1†	3.0	1.1	4.1	1,180†	13,000	.....	1.0
956	Pentode mixer	250	100	−1†	2.3	0.8	3.1	650†	33,000	.....	1.7
6J5	Triode mixer	100	...	−1†	2.1	...	.....	620§	6,500	.....	0.74
6AC7/1852	Triode mixer	150	150	−1†	...	...	6.5	4,200§	950	.....	0.28
955	Triode mixer	150	...	−1†	2.8	...	.....	660§	6,100	.....	0.72

* Reproduced from HARRIS, W. A., Space Charge Limited Current Fluctuations in Vacuum Tube Amplifiers and Input Systems, *RCA Rev.*, vol. 5, pp. 505–524, 6, pp. 114–124, April, July, 1941.

† At peak of oscillator cycle. ‡ Hexode section only. Triode section takes its current from a separate part of the cathode.

§ Conversion transconductance value. ** For effective bandwith of 5,000 cycles.