MODERN ELECTROMAGNETIC FIELDS

P. Silvester

A concise introduction to the physical origins of the electromagnetic field and to selected methods of solving field problems arising in electrical engineering.
This book provides a detailed description of the major physical concepts of electromagnetic field theory that will furnish a thorough engineering science background for specialized studies in microwave devices, electric machines, and other areas of electromagnetic engineering. To this end, theoretical development is paralleled by the introduction of useful solution methods ranging from analytic conformal transformation to numerical iterative processes, and their application to problems of a simple and practical nature.

**Outstanding Features:**

- Simple numerical methods for solving field problems are introduced at an early stage. Physical comprehension is thereby enhanced without losing sight of the fields problem in complicated mathematics.

- Special relativity is introduced briefly to show the intimate relationship between electric and magnetic fields and to develop Maxwell’s equations as a simple and logical extension of electrostatics.

- Critical bibliographies are included at the end of each chapter. About one-third of the references are from

(Continued on back flap)
THE LIBRARY
THE HARRIS COLLEGE
CORPORATION STREET, PRESTON.

All Books must be Returned to the College Library or Renewed not later than the last date shown below.

<table>
<thead>
<tr>
<th>Date</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 JAN. 1974</td>
<td>14 JAN. 1976</td>
</tr>
<tr>
<td>12 FEB. 1976</td>
<td>15 NOV. 1977</td>
</tr>
<tr>
<td>16 DEC. 1977</td>
<td></td>
</tr>
<tr>
<td>2 JAN. 1978</td>
<td></td>
</tr>
</tbody>
</table>

530.141 SIL
A/C 037041

000 574 969
Modern
Electromagnetic
Fields
PRENTICE-HALL ELECTRICAL ENGINEERING SERIES*
William L. Everitt, Editor

ANNESS Elementary Nonlinear Electronic Circuits
ANNESS Elements of Television Systems
ARMINGTON & VOLZ. An Introduction to Electric Circuit Analysis
BALABANIAN Network Synthesis
BARTON Radar Systems Analysis
BENEDICT Electronics for Engineers and Scientists
BLACKWELL & KOTZEBUE. Semiconductor-Diode Parametric Amplifiers*
BOISVERT, ROBERT, & ROBICHAUD. Signal Flow Graphs and Applications
BURDIC. Radar Signal Analysis
CARLIN & GIORDANO Network Theory: An Introduction to Reciprocal and Nonreciprocal Circuits
CHANG, K. N. Parametric and Tunnel Diodes
CHANG, S. L. Energy Conversion*
CHEN & HAAS Elements of Control Systems Analysis: Classical and Modern Approaches
CHIRLAN Integrated and Active Network Analysis and Synthesis
DAVIS & WEEDE Industrial Electronic Engineering
DECKER Electrical Engineering Materials
DEL TORO Electromechanical Devices for Energy Conversion and Control Systems
DES TORO Principles of Electrical Engineering
DE Plan Linear Active Network Theory
DOWNING Modulation Systems and Noise*
DUNN & BARKER Electrical Measurements Manual
EVANS Experiments in Electronics
EVANS Introduction to Electronics
FICH Transient Analysis in Electrical Engineering
FICH & POTTER Theory of A-C Circuits
FLORES The Logic of Computer Arithmetic*
FOECKE Introduction to Electrical Engineering Science
GENTRY ET AL. Semiconductor Controlled Rectifiers: Principles and Applications of p-n-p-n Devices
GOLDBM Information Theory
GOLSDON Transformation Calculus and Electrical Transients
GOLUMBH ET AL. Digital Communications with Space Applications*
GRAY Digital Computer Engineering*
HERRERO & WILLONEX Synthesis of Filters
HESCHBERGER Principles of Communication Systems
JORDAN Electromagnetic Waves and Radiating Systems
KUO Analysis and Synthesis of Sampled-Data Control Systems*
KUO Automatic Control Systems, 2nd ed.

LE CROISSETTE  Transistors
LEGROS & MARTIN  Transform Calculus for Electrical Engineers
LO, ENDRES, ZAWELS, WALDHÄUER, & CHANG  Transistor Electronics
MALEY & EARLE  The Logic Design of Transistor Digital Computers*
MALEY & SKICO  Modern Digital Computers
MANASSE, EKIS, & GRAY  Modern Transistor Electronics Analysis and Design
MARCUS  Switching Circuits for Engineers, 2nd ed.*
MARTIN  Electronic Circuits
MARTIN  Physical Basis for Electrical Engineering
MARTIN  Ultrahigh Frequency Engineering
MATSCHI  Capacitors, Magnetic Circuits, and Transformers
NIXON  Principles of Automatic Controls
NUSSBAUM  Electromagnetic and Quantum Properties of Materials
NUSSBAUM  Electromagnetic Theory for Engineers and Scientists
NUSSBAUM  Semiconductor Device Physics*
OGATA  State Space Analysis of Control Systems
PARTRIDGE  Principles of Electronic Instruments
PASKUSZ & BUSSELL  Linear Circuit Analysis
PIERUSSCHKA  Principles of Reliability*
POTTER & FICH  Theory of Networks and Lines
PUMPHREY  Electrical Engineering, 2nd ed.
PUMPHREY  Fundamentals of Electrical Engineering, 2nd ed.
RAEMER  Statistical Communication Theory in Radio System Analysis
REEF  Foundation for Electric Network Theory
RIDEOUT  Active Networks
ROBERTS & VANDERSLICE  Ultrahigh Vacuum and Its Applications*
RUBIN & DiFRANCO  Introduction to Radar Detection Theory
RYDER, J. D.  Networks, Lines and Fields, 2nd ed.
SAGE  Optimum Systems Control
SANFORD  Physical Networks
SHEED  Fundamentals of Electromagnetic Waves
SILVESTRE  Modern Electromagnetic Fields
SKRODER & HELM  Circuit Analysis by Laboratory Methods, 2nd ed.
SODOHO  Theory and Application of Ferrites
STOUT  Basic Electrical Measurements, 2nd ed.
STROUM & WARD  Laplace Transformation Solution
THOMSON  Laplace Transformation, 2nd ed.
VAN DER ZIEL  Noise
VAN DER ZIEL  Solid State Physical Electronics
VAN VALKENBURG  Network Analysis, 2nd ed.
WALLMARK & JOHNSON, EDS.,  Field-Effect Transistors: Physics, Technology and Applications
WARFIELD  Introduction to Electronic Analog Computers
WEED & DAVIS  Fundamentals of Electron Devices and Circuits
Modern Electromagnetic Fields

P. SILVESTER
Associate Professor of Electrical Engineering, McGill University

PRENTICE-HALL, INC. ENGLEWOOD CLIFFS, N.J.
This book is aimed at an intermediate-level audience already acquainted with the basic notions of electrical physics, but only beginning to appreciate their relevance and utility in electrical engineering. It has been used successfully by the author in his own junior-year course, as well as in another course taken by Honors Engineering students. As the table of contents will show, the selection of theoretical material is not unconventional, while its level and manner of presentation vary from strictly classical to the frankly experimental. Considerable space is devoted to the development of field theory proper, with wave problems occupying not more than one-quarter of the whole. The intermediate fields course is rarely terminal, and is most often followed by at least one other course with emphasis on wave behavior; an adequate preparation in fields is therefore considered the prime objective here.

Two aspects of this text will be found to deviate significantly from the mainstream of intermediate electromagnetics books. On the purely theoretical side, special relativity is introduced briefly in order to permit the
physical relationships between electricity and magnetism to be exhibited much more clearly than is usual at the undergraduate level. While dyadic or tensor notations are mandatory for a full development of relativistic electrodynamics, it is nevertheless possible to go surprisingly far using nothing more esoteric than matrix algebra! The second, and perhaps more important, novelty concerns the use of numerical methods. Field theory is quite rightly considered a difficult subject by most students, since even very simple problems must be overwhelmed by involved manipulative mathematics, often using strange and unfamiliar functions or notations. By introducing numerical methods at an early stage, it is possible to reduce drastically the manipulative labor intervening between problem statement and solution, enabling many more problems to be solved and strengthening physical understanding very considerably. Several of the more common methods are therefore introduced side by side with analytic techniques. No attempt is made, however, to develop any numerical method beyond its elementary form, for it is surely just as undesirable to obscure the physical problem with elegant numerical analysis as to bury it beneath lengthy development of, for example, properties of certain special functions! The use of computational methods, of course, has also quite significant—and beneficial—side effects on the theoretical development. For example, the method of separation of variables now assumes a primarily physical role, instead of appearing as a neat manipulative trick for solving certain boundary-value problems.

Several time-honored topics will be found absent from this book. Practically nothing is said about the nature and properties of dielectric and magnetic materials, for example. This omission reflects the great and growing importance of materials science, to which at least one undergraduate course is nowadays devoted; little would be accomplished by repeating that subject in dilute form. Similarly, no microwave or other specialized technology is introduced (with the sole exception of waveguides). It is always best to illustrate new viewpoints initially by means of familiar devices—there is plenty of time to learn about new devices once the principles have been understood. Much the same applies to specialized mathematical technique, which is deliberately held at the simplest possible level. The general objective is always development of physical understanding first, and manipulative techniques afterward.

A word about the background expected of the student is in order. He is presumed to be familiar with basic vector analysis, and is expected to have some acquaintance with complex variable theory. The latter is particularly desirable for work on conformal transformation. Of course, numerical solution of problems presupposes a knowledge of elementary Fortran (or equivalent) programming. To avoid excessive difficulties for the beginner, most of the programmes used to solve illustrative problems are grouped in an Appendix, and many of the simpler problems at chapter ends require
the student to modify an already existing programme, rather than to write an altogether new one.

The author is indebted to his colleagues at McGill University, especially Dr. T. J. F. Pavlasek, for their generous and helpful assistance and advice, and to both Margarets for diligence beyond the call of duty. The greatest influence on this book, however, has been that of the most merciless yet kind critics any teacher can have—his students.

P. Silvester
Contents

A Relativistic Preamble, 1

1. The Special Relativity Postulate, 1
2. The Light-Flash Experiment, 3
3. The Lorentz Transformation, 5
4. The Basic Law of Kinematics, 8
   Readings, 9
   Problems, 11

1 Origins and Nature of the Electric Field, 13

1. The Fundamental Postulates, 13
2. The Continuity Principle, 15
3. Flux Density Due to a Point Charge, 16
4. Gauss’s Law, 18
5. Coulomb’s Law, 20
CONTENTS

6. Fields in Material Media, 23
7. Scalar Potential of Conservative Fields, 25
8. General Solution of the Potential Problem, 27
9. Total Energy of a Charge Distribution, 30
10. Properties of Scalar Potentials, 32
11. Fields at Material Boundaries, 35

Readings, 37
Problems, 39

2 Engineering Electrostatics, 41

1. Fields and Potential in a Coaxial Cable, 42
2. Parallel-Wire Lines, 44
3. Conductors at High Voltages, 48
4. The Method of Images, 52
5. Relaxation Solutions, 53
6. Numerical Relaxation, 56
7. Liebmann's Iteration Method, 62
8. A Monte Carlo Method, 65
9. Potential and Induction Coefficients, 66
10. Matrix Solutions to Potential Problems, 70
11. Capacitance and Capacitors, 73
12. Electromechanical Devices, 76

Readings, 78
Problems, 81

3 Mapping of Steady Currents and Fields, 84

1. The Continuity Equation of Current, 84
2. Potential and Charge Distributions, 86
3. Energy and Power, 88
4. Analogue Solutions of Equilibrium Problems, 90
5. Resistive Mesh Analogues, 92
6. Experimental Study of Complex Variables, 93
7. Complex Potentials, 95
8. Conformal Mapping, 98
9. Field of a Long Charged Strip, 101
10. Current Flow Around a Hole, 104
11. The Schwarz-Christoffel Transformation, 106
12. Field at the Edge of a Parallel-Plate Capacitor, 109

Readings, 113
Problems, 114
4 Electrodynamics and Maxwell’s Equations, 116
1. Flux and Field of a Moving Charge, 116
2. Force between Moving Charges, 119
3. Magnetic Field and Flux, 122
4. Ampère’s Law, 125
5. Maxwell's Field Equations, 128
6. Effects of Material Media, 131
7. Fields at Material Boundaries, 132
8. Electromotive Force, 135

Readings, 139
Problems, 141

5 Magnetic Vector Potential and Energy, 143
1. Magnetic Vector Potential, 143
2. Magnetic Moment and Torque, 147
3. Magnetic Field Energy of Current-Carrying Circuits, 149
4. Energy in Terms of the Field Vectors, 151
5. Explicit Calculation of Vector Potential, 153
6. Neumann’s Formula for Inductances, 156
7. Self-Inductance of a Square Loop, 159
8. Neumann’s Formula in Summation Form, 160
9. Direct Inductance Calculations, 163
10. Inductances of Solenoidal Coils, 165

Readings, 167
Problems, 168

6 Magnetostatic Problems in Engineering, 171
1. Fields of Current-Carrying Conductors, 172
2. Finite-Difference Approximations of the Integral Solution, 174
3. Current Parallel to a Permeable Surface, 176
4. Image Solutions for Finite Permeabilities, 179
5. Leakage Field of a Transformer, 182
6. Magnetic Scalar Potential, 187
7. Conformal Transformation of Magnetic Fields, 189
8. Flux Distribution in a Transformer Core, 192
9. Filamentary Source in a Slot or Duct, 194
10. Magnetomotive Force and Reluctance, 199
11. Iterative Solution in Cylindrical Coordinates, 203
12. Inductance Calculation from Field Maps, 206

Readings, 207
Problems, 210
7 Time-Varying Fields in Conductors, 212
   1. Field-Vector Wave Equations, 212
   2. Solution by Separation of Variables, 212
   3. Fields in a Pulse Transformer Core, 217
   4. Separation of Variables: Rectangular Core, 221
   5. Magnetizing Impedance of a Transformer, 224
   6. Skin Effect in Round Wires, 227
   7. Skin Effect in High Frequencies, 230
   8. Electrodynamic Potential and Wave Equations, 233
   9. Direct Integral Solutions: Current in a Flat Bar, 235
  10. Poynting's Theorem, 238

   Readings, 241
   Problems, 243

8 Traveling Electromagnetic Waves, 245
   1. Fields in Free Space, 245
   2. Plane Waves in Unbounded Space, 247
   3. Polarization of Plane Waves, 249
   4. Interference of Plane Waves, 251
   5. Reflection of Plane Waves at Material Boundaries, 253
   6. General Plane Electromagnetic Waves, 255
   7. TEM Waves on Transmission Lines, 258
   8. Calculation of Characteristic Impedances, 260
   9. Waves with Longitudinal Field Components, 263
  10. The Scalar Helmholtz Problem for a Waveguide, 266
  11. TM Waves in a Rectangular Guide, 269
  12. TE Waves in a Rectangular Guide, 271
  13. Numerical Solution of the Helmholtz Equation, 273
  14. Principal Modes of a Ridged Waveguide, 277
  15. The Helmholtz Equation under Conformal Transformation, 283
  16. A High-Power Waveguide, 285

   Readings, 290
   Problems, 292

Appendices, 295

Appendix I: A Note on Units, 295
Appendix II: Examples of Computer Programmes, 300
Appendix III: Some Useful Mathematical Subroutines, 319

Index, 327
Modern
Electromagnetic
Fields
A Relativistic Preamble

The premises underlying electromagnetic field theory are much the same as those on which other engineering disciplines are based. The sole possible exception is the special theory of relativity, a portion of which forms the connecting link between electricity and magnetism. In order to avoid interrupting the development of electromagnetic theory proper later on, the kinematics of special relativity will be examined briefly in this preamble.

1. The Special Relativity Postulate

In mechanics it is postulated that any problem may be formulated in any rectilinear coordinate system. This postulate permits using vector notation for stating the laws of mechanics, for the essence of vector notation lies precisely in the separation of physical entities from the coordinate systems in which they might be expressed. For example, one writes \( f = ma \) to describe the relationship between the physical entities force, mass, and acceleration, without
reference to any specific set of coordinates. Although seldom specifically
spelled out, it is implicitly understood that the coordinate systems referred
to are such as to make Newton’s first law valid; that is, such that a mass
particle placed at some point \( P \) with zero velocity and not subjected to any
forces will always remain at point \( P \). Such coordinate systems and the time
scales that go with them are called \textit{inertial reference frames}. Of course it is
ultimately necessary to choose some one particular reference frame in order
to find the numerical solution of any given problem. This choice of frame is in
principle quite arbitrary, though in practice dictated by the nature of the
problem. Even a simple problem may be made very complicated by choosing
an awkward reference frame.

The special theory of relativity generalizes this principle of equivalence of
reference frames to include not only mechanics but all other branches of
physics. It asserts that \textit{all inertial reference frames are equivalent for the
formulation of all physical laws}. Naturally, no proof of this assertion can
ever be given, for all physical laws can never be known.

It should be carefully noted that this assertion does not deny the posi-
sibility of formulating physical laws in other kinds of coordinate systems; all
that it does say is that they do not retain the same form in others. As a simple
example, let a steel ball be placed on a rotating phonograph turntable, as in
Fig. 1. It is clear that the ball will roll radially outward in response to cen-

![Fig. 1](image)

trifugal and centripetal forces. To analyze the motion of the ball, one might
wish to use polar coordinates attached to the turntable. In this coordinate
system, Newton’s first law is not valid without modification; an object
placed at the point \((r, \theta)\) accelerates in the \(r\)-direction unless restrained by
some force. The modifications required are simple enough, and analysis in
terms of these coordinates is undoubtedly convenient. Nevertheless, this
coordinate system is not part of an inertial reference frame. It is easily seen
that not only rotating but \textit{any} kind of accelerating reference frames are
noninertial. Problems involving noninertial reference frames are considered
in the general theory of relativity and will not be dealt with here.

To develop electromagnetic theory, Cartesian coordinate systems moving
at constant speeds with respect to each other are found to suffice. A mass
particle moving with some constant velocity in any one such reference frame has some other constant velocity in all the others. The coordinate systems to be used are therefore inertial, and special relativity will be found adequate to the task. The historical origin of special relativity lies principally in the work done in optics in the nineteenth century. At that time, two theories of light were current, the wave theory and the corpuscular theory. According to the former, light consists of a wave motion in the ether; the observed velocity of light consequently depends on the properties of the ether and the velocity with which the observer moves through the ether but not on the velocity of the light source. On the other hand, the corpuscular theory requires that the observed velocity depend on both observer and source velocities. This conflict ought to be resolvable by experimental measurements although the necessary experiments are very difficult. They were in fact not successfully performed until nearly the turn of the century. Then, however, it was discovered that the observed velocity was a constant, dependent on neither the source nor the observer!

There exists sufficient direct evidence to convince all but a very few contemporary physicists that, as an experimental fact, the velocity of light is the same in all inertial coordinate systems. Evidently neither the wave nor the corpuscular theory emerges as established without further modification. Just what mechanism underlies this phenomenon, however, is not at all clear without further examination.

2. The Light-Flash Experiment

To try to reach an understanding of the nature of light, it is necessary to examine the form that transformations between inertial coordinate systems assume and then to express the same optical experiment in different reference frames.

Let a force-free particle be placed in a space described by a Cartesian coordinate system. According to Newton’s first two laws, the particle travels in a straight-line path (or it might stand still; a coordinate system in which it moves, however, can always be found). The position of this particle is completely specified by the four coordinate quantities \( x, y, z, t \). Hence three equations in these variables are needed to describe the motion of the particle. It is important to observe that these must be linear equations; if they contain any but the first powers of the coordinates, or of time, then the motion is not uniform and the particle is not force-free. Let the reference frame composed of \( x, y, z, t \) be called \( F \).

In another reference frame \( F' \), also assumed to be inertial and described by \( x', y', z', t' \), exactly the same arguments apply. Thus, extending the same reasoning to \( F'', F''', \ldots \), it is evident that the equations of motion of a force-free particle must be linear in all inertial frames. This is only possible
if the relationship between coordinate quantities in any two reference frames is linear, i.e., if each of the quantities $x', y', z', t'$ can be expressed as a linear combination of $x, y, z, t$. In matrix terms, it must be possible to write

$$
\begin{bmatrix}
    x' \\
    y' \\
    z' \\
    t'
\end{bmatrix} =
\begin{bmatrix}
    T_{11} & T_{12} & T_{13} & T_{14} \\
    T_{21} & T_{22} & T_{23} & T_{24} \\
    T_{31} & T_{32} & T_{33} & T_{34} \\
    T_{41} & T_{42} & T_{43} & T_{44}
\end{bmatrix}
\begin{bmatrix}
    x \\
    y \\
    z \\
    t
\end{bmatrix}
$$

(1)

where all the components $T_{ij}$ of the transformation matrix are independent of the coordinates of either frame of reference.

The transformation matrix for a particular pair of reference frames will now be found. Let both frames $F$ and $F'$ be composed of Cartesian space coordinates and associated time scales. Let the frame $F'$ move at a velocity $v = iv_x$ with respect to $F$, and let the frames be oriented so that at $t = t' = 0$ the coordinate axes coincide, as in Fig. 2. It is clear that since the velocity vector $v$ is along the $x$- and $x'$-axes, $y = y'$ and $z = z'$ always; lines parallel to the $x$-axis are also parallel to the $x'$-axis. All but four of the components of the transformation matrix are thus immediately known:

$$
\begin{bmatrix}
    x' \\
    y' \\
    z' \\
    t'
\end{bmatrix} =
\begin{bmatrix}
    T_{11} & 0 & 0 & T_{14} \\
    0 & 1 & 0 & 0 \\
    0 & 0 & 1 & 0 \\
    T_{41} & 0 & 0 & T_{44}
\end{bmatrix}
\begin{bmatrix}
    x \\
    y \\
    z \\
    t
\end{bmatrix}
$$

(2)

This particularly simple transformation naturally results from the specific choice of coordinates. The remaining four components may be evaluated by considering a simple experiment. At time zero (the same in both frames), let a flashbulb go off at the origin (at that time, the same in both systems). An observer fixed in frame $F$ must perceive the flash as a spherical shell of illumination, whose thickness depends on the duration of the flash, spreading
out from the origin in time; the radius of this shell must be $ct$, where $c$ is the velocity of light. This situation is described mathematically by

$$x^2 + y^2 + z^2 = c^2 t^2$$

(3)

In a similar fashion, an observer fixed in frame $F'$ must also see a spherical shell of illumination propagating from the origin at the speed of light, so that its radius is $ct'$; mathematically,

$$x'^2 + y'^2 + z'^2 = c^2 t'^2$$

(4)

and since $y = y', z = z'$, (3) and (4) may be combined to yield

$$x^2 - c^2 t^2 = x'^2 - c^2 t'^2$$

(5)

To find the four unknown $T_{ij}$, four independent equations in $x$, $t$, $x'$, $t'$ are required. The matrix equation (2) furnishes two, the flashbulb experiment one more. One of the $T$'s can be eliminated by using the definition of the coordinate systems, so the problem may be solved. Elimination of $T_{14}$ may be performed by taking the top line of (2),

$$x' = T_{11} x + T_{14} t$$

and writing it as

$$x' = T_{11} (x + \frac{T_{14} t}{T_{11}})$$

(6)

From the way the reference frames were set up, the origin of frame $F'$ must correspond to a point moving along the $x$-axis in frame $F$. Since the origin of $F'$ is at $x' = 0$ and the velocity of $F'$ relative to $F$ is $v$, substitution in (6) gives

$$x = -\frac{T_{14} t}{T_{11}} = vt$$

so that

$$x' = T_{11} (x - vt)$$

(7)

Equation (7), the last line of (2), and the “experimental” result (5) is then the system of equations that may be solved to yield the remaining three $T$'s.

3. The Lorentz Transformation

In the foregoing, a matrix equation was set up to express the relationship between the coordinates of any two inertial reference frames. It remains to carry out the mathematical solution. To achieve it, Equation (5) may be written

$$t'^2 = \frac{x'^2}{c^2} - \frac{x^2}{c^2} + t^2$$

(8)

Substituting for $x'$ its value from Equation (7) above,

$$t'^2 = \frac{T_{11}}{c^2} x^2 - 2 \frac{T_{11} v}{c^2} x t + \left(1 + \frac{v^2}{c^2} T_{11}\right) t^2$$

(9)
From the fourth line of Equation (2),
\[ t'^a = T_{i4}^a x^2 + 2T_{i4} T_{4t} x t + T_{44} t^2 \]  
(10)

Now Equations (9) and (10) express exactly the same quantity and must represent identically the same polynomial. They must therefore be equal term by term. Equating the coefficients, there are obtained
\[ T_{4i} = \frac{1}{c} \sqrt{T_{ii}^2 - 1}; \quad T_{i4} T_{4i} = -\frac{T_{ii}^2 v^2}{c^2}; \]  
(11)

and \[ T_{44} = \frac{1}{c} \sqrt{v^2 T_{ii}^2 + c^2} \]

These equations are easily solved by multiplying the first and third and equating the result to the second; there obtains
\[ T_{ii} = \frac{1}{\sqrt{1 - (v^2/c^2)}} \]  
(12)

and on substituting,
\[ T_{4i} = -\frac{v}{c^2} \frac{1}{\sqrt{1 - (v^2/c^2)}} \quad \text{and} \quad T_{44} = T_{ii} \]  
(13)

For brevity, write
\[ T_{ii} = T_{44} = \gamma \]  
(14)

so that Equation (2) becomes
\[
\begin{bmatrix}
  x' \\
  y' \\
  z' \\
  ct'
\end{bmatrix} =
\begin{bmatrix}
  \gamma & 0 & 0 & -\frac{v}{c} \gamma \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 1 & 0 \\
  -\frac{v}{c} \gamma & 0 & 0 & \gamma
\end{bmatrix}
\begin{bmatrix}
  x \\
  y \\
  z \\
  ct
\end{bmatrix} \tag{15}
\]

Equation (15) is generally known as the Lorentz transformation and the number \( \gamma \) as the Lorentz factor. It is readily seen that the Lorentz factor is always positive and never smaller than unity. The reason for writing Equation (15) in terms of \( ct \) and \( ct' \), instead of \( t \) and \( t' \) as before, is simply to make the matrices dimensionally homogeneous and permit easy checking of calculations.

The Lorentz transformation has many interesting mathematical and physical properties. Of these, the principal ones to be considered here are
(a) it relates time and distance, showing them not to be independent quantities;
(b) it becomes a simple geometrical coordinate transformation, with time and distance not related, for velocities \( v \) much less than \( c \);
(c) it is a nonsingular (reversible) transformation. As is evident from the
relativity postulate itself, the inverse of the transformation is the same, except for reversal of the sign of \( v \);

(d) it is symmetric in \( x \) and \( ct \) as well as in \( y \) and \( z \).

Some of the kinematic implications of the transformation may now be examined. Let a stick \( L_0 \) units long be laid along the \( x' \)-axis in frame \( F' \), extending from \( x' \) to the point \( (x' + L_0) \). Suppose an observer fixed in frame \( F \) measures the length of the stick as it passes him. Say one end appears to him at point \( x \), the other at \( (x + L) \). By Equation (15), these points are related by

\[
x' = \gamma(x - vt) \quad \text{and} \quad x' + L_0 = \gamma(x + L - vt)
\]

so their difference, on subtracting, is

\[
L = \frac{1}{\gamma}L_0
\]

This equation says that a stick stationary in frame \( F' \) looks shorter by the Lorentz factor in frame \( F \). This phenomenon is called the Lorentz contraction of length and is important in electromagnetic theory.

Suppose next a clock to be fixed at a point \( x', y', z' \) in frame \( F' \). Just as for the stick problem above, one may deduce, using the bottom line of (15), that an interval \( \delta t' \) on this clock corresponds to

\[
\delta t' = \gamma \left( -\frac{v}{c^2} \delta x + \delta t \right)
\]

As the clock is fixed in \( F' \), the top line of (15) says

\[
\delta x' = 0 = \gamma(\delta x - v \delta t)
\]

and combining these,

\[
\delta t = \gamma \delta t'
\]

which means that the observer in \( F \) thinks the clock in \( F' \) loses time. This phenomenon is called the Lorentz time dilation. There exists direct experimental evidence of both the length contraction and the time dilation (e.g., the Michelson–Morley experiment).

It should be observed at this point that neither the length contraction nor the time dilation in any way refers to imperfections in the method of measurement, nor to any curious kind of optical illusion, but expresses the actual state of affairs. Since it seems hard to conceive of a physical stick having several lengths, it might quite naturally be asked which is its real length, perhaps as seen by some "impartial" observer. The essence of special relativity is precisely that there is no impartial observer, and there cannot be one; hence all measured quantities have an equal claim to reality, each in the reference frame in which it was established but in no other. Beauty, in other words, lies strictly in the eye of the beholder!
4. The Basic Law of Kinematics

As kinematics deals with the relative motion of bodies and coordinate frames, its fundamental law is the addition law of relative velocities,

\[ v_{20} = v_{21} + v_{10} \]  

(18)

Under the relativity postulate, this law has to be modified to take account of the contraction and dilation that occur in various reference frames.

To keep the mathematics simple, let there be three reference frames so constructed that their x-axes coincide at time zero, which will be selected so as to have it occur simultaneously in the three systems (i.e., \( t'' = t' = t = 0 \) when the three origins coincide). Let the y- and z-axes of all three systems again be similarly oriented, so that \( y'' = y' = y \) and \( z'' = z' = z \) for all \( x, t \).

Call these frames \( F'' \), \( F' \), and \( F \). Consider three force-free bodies moving with parallel though unequal velocities such that one body is stationary in each of the three frames. (The situation with velocities not parallel is naturally more general, but also more complicated.) Let the velocity of frame \( F'' \) be \( v_{21} \) with respect to frame \( F' \) and \( v_{20} \) with respect to \( F \). Let the velocity of \( F' \) be \( v_{10} \) with respect to \( F \). It is a simple matter to write the Lorentz transformation that relates \( F'' \) to \( F' \):

\[
\begin{bmatrix}
  x'' \\
y'' \\
z'' \\
c't''
\end{bmatrix} = \begin{bmatrix}
  \gamma_{21} & 0 & 0 & -\frac{v_{21}}{c}\gamma_{21} \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 1 & 0 \\
  -\frac{v_{21}}{c}\gamma_{21} & 0 & 0 & \gamma_{21}
\end{bmatrix} \cdot \begin{bmatrix}
  x' \\
y' \\
z' \\
c't'
\end{bmatrix}
\]  

(19)

and \( F' \) to \( F \):

\[
\begin{bmatrix}
x' \\
y' \\
z' \\
c't'
\end{bmatrix} = \begin{bmatrix}
  \gamma_{10} & 0 & 0 & -\frac{v_{10}}{c}\gamma_{10} \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 1 & 0 \\
  -\frac{v_{10}}{c}\gamma_{10} & 0 & 0 & \gamma_{10}
\end{bmatrix} \cdot \begin{bmatrix}
x \\
y \\
z \\
c t
\end{bmatrix}
\]  

(20)

so that, substituting Equation (20) into Equation (19),

\[
\begin{bmatrix}
x'' \\
y'' \\
z'' \\
c't''
\end{bmatrix} = \begin{bmatrix}
  \gamma_{21}\gamma_{10}(1 + \frac{v_{21}v_{10}}{c^2}) & 0 & 0 & -\gamma_{21}\gamma_{10}\frac{v_{21} + v_{10}}{c} \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 1 & 0 \\
  -\gamma_{21}\gamma_{10}\frac{v_{21} + v_{10}}{c} & 0 & 0 & \gamma_{21}\gamma_{10}(1 + \frac{v_{21}v_{10}}{c^2})
\end{bmatrix} \cdot \begin{bmatrix}
x \\
y \\
z \\
c t
\end{bmatrix}
\]  

(21)
But this transformation, relating as it does \( F'' \) to \( F \), must be the same as the
direct transformation

\[
\begin{bmatrix}
x'' \\
y'' \\
z'' \\
ct''
\end{bmatrix} = 
\begin{bmatrix}
\gamma_{20} & 0 & 0 & -\frac{v_{20}}{c}\gamma_{20} \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
-\frac{v_{20}}{c}\gamma_{20} & 0 & 0 & \gamma_{20}
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z \\
ct
\end{bmatrix}
\]  

(22)

Comparing the corresponding elements of the transformation matrices in
(21) and (22), the relationships may be deduced

\[
\frac{v_{20}}{c}\gamma_{20} = \gamma_{21}\gamma_{10}\left(\frac{v_{21} + v_{10}}{c}\right)
\]  

(23)

and

\[
\gamma_{20} = \gamma_{21}\gamma_{10}\left(1 + \frac{v_{21}v_{10}}{c^2}\right)
\]  

(24)

Combining Equations (23) and (24), there results with but a little algebra

\[
v_{20} = \frac{v_{21} + v_{10}}{1 + \frac{v_{21}v_{10}}{c^2}}
\]  

(25)

an equation replacing (18) as the basic law of kinematics. More accurately,
it is easily seen that Equation (18) is an approximation to the result above,
valid at low relative velocities. Equations (23) to (25) may also be rearranged
to yield

\[
v_{21} = \frac{v_{20} - v_{10}}{1 - \frac{v_{20}v_{10}}{c^2}} \quad \text{and} \quad \gamma_{21} = \gamma_{10}\gamma_{20}\left(1 - \frac{v_{10}v_{20}}{c^2}\right)
\]  

(26)

giving the relative velocity between two moving bodies as seen by an observer
stationary in his own reference frame. Equation (25) is usually referred to as
the velocity addition formula.

These kinematic results, few in number but fundamental in their
importance, suffice for the development of electromagnetic theory. For the
theory of relativity, they form only the beginning steps; the interested reader
will find many further unexpected and startling results. The central purpose
here, however, is the development of the electromagnetic theory, not of rela-
tivity, and it is to this main task that the chapters to follow are addressed.

**READINGS**

Although there exist a great many books on special relativity, the majority
is written by physicists for physicists-to-be, and the engineering student
frequently finds himself in the position of the legendary little boy at the public
library ("this book told me a great deal more about penguins than I wanted
to know"). With a little care, however, a suitable selection can be culled from the many volumes available.

A work that deserves to be placed high on any reading list is that by Becker and Sauter (1), a recent revision of Abraham and Becker's long renowned classic work. The chapter in Jackson (2) is primarily directed toward physicists studying electromagnetic theory and while lucidly written, cannot be termed easy; the engineering student should expect to encounter some difficulty after the first four sections. Tralli (3) and Cullwick (4) do not expect quite so much of their readers and yet provide them with a thorough discussion of the subject. The short volume by Rindler (5) delves into special relativity more deeply, but without great mathematical complexity at the start, and may be read with profit. A much easier book to read, but deserving strong recommendation, is that by Katz (6). The few pages in Artley's book (7) are written for an engineering audience but cannot be expected to suffice entirely.

A quite recent addition to the literature, almost in a class by itself, is by Rosser (8). It is written largely for an experimentally minded audience and gives fascinating accounts of many of the classical experiments. Because of his unusually clear exposition, the book deserves strong recommendation.

A great deal of insight into the physical problems involved in special relativity can be gained by spending an evening with one of the many books intended for an intelligent lay readership. The books by Eddington (9) and Schlick (10) are written by authorities on the subject, but for an audience of nonspecialists; they are well worth reading. It is held by some that Einstein has been the best popularizer of relativity, and there can be no doubt that his explanatory book (11) covers the questions most likely to bother the beginner. Another nontechnical, but by no means trivial, alternative in this class is by Bridgman (12).

There exist also numerous books designed for an audience with a less scientific inclination, which tend to avoid use of technical terminology and shun mathematics altogether. A good choice among these is Coleman (13), another Durrell (14). While they undoubtedly provide food for thought, it would be unwise to expect these books to furnish all the material needed to build an understanding of electromagnetism. They should be supplemented by the more detailed treatments enumerated above.

Finally, the ethical (and briefly, some of the religious) implications of a relativistic point of view are examined in a book by Frank (15), an eminent authority on the philosophy of science.


PROBLEMS

1. Suppose that a reference frame \( F' \) moves relative to a frame \( F \) with a velocity \( v \) along the x-axis. When clocks at \( (x'_0, y'_0, z'_0) \) and \( (x_0, y_0, z_0) \) pass each other, they indicate times \( t'_0 \) and \( t_0 \) respectively. Write down the Lorentz transformation for this case.

2. Show that although the apparent length of a moving ruler may vary, the order in which its ends pass an observer is always the same.

3. Show that although the elapsed time between two events differs in different reference frames, the time sequence of the events is always the same.

4. Two events occur at different times, but at the same space point in a given reference frame \( F \). Prove that the time span between events in any other inertial frame is longer than in \( F \) (hence that no frame exists where the two events are simultaneous).

5. Two rulers, each having a length \( l_0 \) at rest, move in opposite directions with
uniform velocities along the $x$-axis. An observer at rest relative to one of the rulers notes that the time interval between instants at which the left and right ends of the rulers pass each other is $\Delta t$. What is the relative velocity of the two rulers? In what order will the ends of the rulers pass each other for observers at rest relative to either of the rulers? for an observer with respect to whom both rulers move with equal velocities in opposite directions?

6. Use the velocity addition formula to prove that no material object can move at a speed greater than $c$.

7. Prove that the Lorentz transformation is indeed reversible by showing that the transformation matrix between frames $F$ and $F'$ is the inverse of the matrix connecting $F'$ and $F$. Show that the transformation matrix is always non-singular.

8. Find the length contraction of an aircraft 150 ft long moving at 600 mph with respect to an observer.

9. Find the amount of time lost by a traveler’s watch in flying from Montreal to Vancouver (2850 mi) at 600 mph.

10. Derive the Lorentz transformation formula without assuming that the velocity of frame $F'$ relative to $F$ is parallel to the $x$-axis. Give the result in a vector form. [Hint: Resolve the position vector $\mathbf{r}$ into the longitudinal and transverse components relative to the direction of the relative velocity $\mathbf{v}$ and use the Lorentz transformation as in Equation (15).]

11. Show that the quantity $r$ defined by

$$r^2 = x^2 + y^2 + z^2 - c^2t^2$$

is invariant under the Lorentz transformation. This quantity corresponds in some ways to distance in Newtonian mechanics and is called *interval*. 
Origins and Nature of the Electric Field

Logical development of any physical theory must start from stated premises. The premises of the theory of electric fields are simple, consisting of nothing more than the addition of a new entity, electric charge, to the mass-energy and space-time of mechanics. Problems encountered in the study of electric fields involve a range of velocities from zero up to the velocity of light, so that Newtonian mechanics is not an adequate tool, and the special theory of relativity must be employed. The basic results of relativistic kinematics have been reviewed in the preamble and need not be reiterated here; only the necessary postulates are recapitulated in the following.

1. The Fundamental Postulates

In order to develop relativistic dynamics it is necessary to agree on three postulates: the existence of space-time, the existence of matter-energy, and finite value of the speed of light in free space.
Space-time, by the first postulate, is an infinite four-dimensional continuum. The four coordinates of space and time are sufficient to describe the location of any event in the physical universe with respect to another. They cannot, however, identify a location with respect to space-time; that would amount to asserting that one part of space-time is somehow distinguishable from the rest. The second postulate asserts that matter-energy exists. As in the case of space-time, the hyphenated word implies that the two quantities are not independent and must be introduced as a single entity. It is supposed that matter-energy is always detectable by some sort of measurement. The third postulate states that the speed of light is finite with respect to some particle of matter (otherwise the word speed means nothing). Since there is no supposition that any one particle of matter-energy is specially privileged, the speed of light must have the same numerical value with respect to any other bit of matter in the universe.

A fourth postulate is necessary to explain some curious experimental facts. In the late eighteenth century, Coulomb established by a series of experiments that an electrified particle of matter always experiences a force when brought into the neighborhood of another. This force may be one of attraction or repulsion, great or small, depending on the distance between particles, the extent of their electrification, and the medium surrounding them. An empirical law relating these variables was deduced by Coulomb from his experimental data. The force exists even when the two particles are placed in a vacuum so that there is no matter surrounding them at all. Forces known in mechanics are either contact forces or d'Alembert forces of acceleration; the electric force is evidently different in kind from either.

No real explanation was offered by Coulomb as to the mechanism by which a particle in one space location could experience forces attributable to another in a different place. He called the apparent transmission of influence through empty space action at a distance, a descriptive but not explanatory term. The theoretical view advanced by many nineteenth-century physicists was that space is not empty but contains a quasi-substance of zero mass density, the ether. An electrically charged particle sets up a strain pattern in the ether, much as a nail driven into wood causes a local compression and a pattern of strains gradually diminishing with distance. Another charge located elsewhere could perceive the existence of this ether strain and become subject to a force resulting from the interaction of its own strain pattern with that of the distant charge. The similarity between the mathematical development of elasticity and electricity that results from this concept makes it very attractive to the analyst.

The ether theory, however, did not satisfy the requirements of optics very well without extensive modification. In the process of modification, the theory became more and more complicated because ever more subtle properties had to be ascribed to the ether to account for all known optical phenomena.
As a result of the Michelson–Morley experiment, it eventually became necessary to endow ether with the “property” of being experimentally undetectable.

In Einstein’s view, only observable physical entities can be real, so an undetectable ether has no real existence at all. A new explanation for the Coulomb experiment then needed to be found. The ether theory had led to a satisfactory development of electricity; the problem now became how to retain the ether strain view of influence, without any ether to strain. One way out of this dilemma is provided by the continuity principle.

2. The Continuity Principle

In order to experience a force attributable to another charged particle, any electrically charged body surely requires intimate contact with some entity related to the distant charge. The principle of continuity, which will be adopted here as explaining the mechanism of the fundamental experiment, asserts that every electrically charged particle possesses a flux which pervades all space, diminishing in density at great distances. For static charge, this view is hardly distinguishable from the ether strain concept; for moving charges, however, important differences arise. The flux is imagined to be tied to the charge and to consist of lines radiating outward from it, much as shown in Fig. 1.1. When moving, every charge carries its flux with it. As physical entities, fluxes are often much easier to visualize than the fluidlike motions of ether strains.

It may well be asked why the ether theory should not be used regardless of its failure in optics. Maxwell, after all, produced a coherent and clear mathematical account of electromagnetic theory based on it. The answer is simply that Maxwell’s development is mathematically but not physically sound. It postulates first the existence of electricity, then the existence of magnetism as a separate entity, and shows that the two interact in certain circumstances. In modern electromagnetism quite the contrary view is taken: magnetism and electricity are one and the same entity viewed in slightly different circumstances. The entire theory consists of the application of the laws of mechanics to electric charges and their fluxes. No other postulates or empirical “laws” are required, further experiment serving only the purpose of verification.

Charge is conventionally imagined to be a fluidlike entity that resides in all matter, with positive and negative charges intermixed in great quantities.
So attractive is this view that there is great temptation to hedge about the concept of flux with reservations, while unhesitatingly accepting charge to be real. On critical examination the reservations are seen to be much better justified in the case of charge than flux. If Einstein's view that observable reality is the only reality is accepted, then only flux is real; no simple experiment can be devised to detect the presence of charge independently of flux. All electrostatic measurements (electrometer and cathode-ray tube experiments, for example) involve the measurement of Coulomb force either directly or inferentially and are therefore really flux measurements. It may even be said that charge is a figment of the imagination, a mere nail in space on which to hang the flux distribution.

The force between two electrified particles in a Coulomb experiment may be of any size and positive or negative. It is therefore necessary to postulate that flux may be strong or weak, positive or negative, and attached to a charge of corresponding magnitude and polarity. This idea is simply expressed by saying that the total flux associated with a particle is proportional to its charge \( q \):

\[
q = K_0 \psi \tag{1-1}
\]

where \( K_0 \) is a constant of proportionality and depends on the units of measurement. In the study of electricity and magnetism, several different systems of units are used. The simplest one for engineering purposes is probably the RMKSC (rationalized meter-kilogram-second-coulomb) system, in which both flux and charge are measured in coulombs, setting the constant \( K_0 \) equal to unity:

\[
q = \psi \tag{1-2}
\]

The reasons for the existence of numerous different systems of units are largely historical. All the development here will be carried out in RMKSC units. However, much engineering work is still done in other systems with which at least a nodding acquaintance must be made by anyone having more than a passing interest in the subject.

3. Flux Density Due to a Point Charge

The idea of flux is necessary to explain the transference of electric influence from one space point to another. Since an object located somewhere in space may be expected to perceive only the flux at its own location, the total flux attached to a distant particle cannot determine the force experienced. Rather, the density of the flux at the point of interest must be the determining factor. Flux density at any point in space is mathematically definable as the amount of flux crossing a unit area normal to the direction of the flux lines. Because the flux density is unlikely to be uniform over a large area, the defini-
Sec. 3 ORIGINS AND NATURE OF THE ELECTRIC FIELD

tion is best stated in differential form as the flux per unit area, taken over a very small area:

\[ D = \frac{d\psi}{dA} \]  

(1-3)

where \( D \) is the flux density and \( dA \) an area element at right angles to the flux lines. Area being a vector quantity, it is convenient to define flux density as a vector quantity also, giving it the direction of the flux lines. Then it will always be true that

\[ \mathbf{D} \cdot d\mathbf{S} = d\psi \]  

(1-4)

where \( d\mathbf{S} \) represents a surface element oriented in any manner whatever, because the scalar product automatically eliminates components of the surface vector not collinear with the flux lines. This relationship is illustrated in Fig. 1.2. Equation (1-4) restated for larger surfaces is

\[ \int_S \mathbf{D} \cdot d\mathbf{S} = \psi \]  

(1-5)

That is, the total flux crossing a surface \( S \) is found as the surface integral over \( S \) of the flux density \( \mathbf{D} \). In the rmksc system, flux is measured in coulombs so the unit of flux density must be the coulomb per square meter.

Let a closed surface be constructed around a single point particle of charge \( q \). All the flux of this particle must intersect the surface. Then the surface integral of flux density must equal the total flux:

\[ \oint \mathbf{D} \cdot d\mathbf{S} = \psi_{\text{total}} \]  

(1-6)

which in turn equals the total charge:

\[ \oint \mathbf{D} \cdot d\mathbf{S} = q \]  

(1-7)

Let it be supposed that the surface is a sphere centered on the charge. By symmetry, the magnitude \( D \) is the same everywhere on the sphere and the vectors \( d\mathbf{S} \) and \( d\mathbf{r} \) coincide. The value of \( D \) can then be found:

\[ \oint_{\text{sphere}} \mathbf{D} \cdot d\mathbf{S} = \oint (D\mathbf{1}_r) \cdot (\mathbf{1}_r d\mathbf{S}) = \oint D dS = D \oint dS = 4\pi r^2 D \]  

(1-8)

Combining Equations (1-7) and (1-8),

\[ \mathbf{D} = \frac{q}{4\pi r^2} \mathbf{1}_r \]  

(1-9)

It must be noted that this equation assumes that there is no flux crossing the surface, except the flux of the point particle in question. As a practical matter, that means any other charges must be far enough away to make their flux contributions negligibly weak.
4. Gauss's Law

Let a system of two charges be considered, one inside and one outside a given simply connected closed surface, as shown in Fig. 1.3. The flux density at any point must be the sum of the flux density vectors $\mathbf{D}_i$ and $\mathbf{D}_o$ of the inside and outside charges:

$$\mathbf{D} = \mathbf{D}_i + \mathbf{D}_o \quad (1-10)$$

The surface integral of $\mathbf{D}$ over the closed surface is therefore

$$\oint \mathbf{D} \cdot d\mathbf{S} = \oint \mathbf{D}_i \cdot d\mathbf{S} + \oint \mathbf{D}_o \cdot d\mathbf{S} \quad (1-11)$$

The two right-hand integrals represent the flux crossing the surface due to the inside and outside particles, respectively:

$$\oint \mathbf{D} \cdot d\mathbf{S} = \oint d\psi_i + \oint d\psi_o \quad (1-12)$$

Since the surface of integration surrounds the inside particle, it must intersect all the flux lines of that particle; each line must originate inside and extend outward. On the other hand, none of the flux of the outside charge originates inside the surface. All the flux $\psi_o$ leaving the surface must therefore enter it at some other point, so that it will be counted twice in the surface integration: once leaving and once entering. The net flux due to the outside particle must therefore be zero, so that

$$\oint_s \mathbf{D} \cdot d\mathbf{S} = \oint_s d\psi_i = q \quad (1-13)$$

This result may be generalized to apply to many charged bodies. Let there be a large number of charged bodies, of which $N$ are inside and $M$ outside a closed surface; let the surface be chosen so that none of the charges is on the
surface itself. (This restriction is necessary because it would otherwise not be clear whether to count such charges as inside or outside.) The flux density at any point in space is

$$D = \sum_{k=1}^{N+M} D_k$$

(1-14)

where $D_k$ is the flux density due to the $k$th charged particle. The total flux may be found by integrating over the closed surface, obtaining

$$\oint_S D \cdot dS = \oint_S \sum_{k=1}^{N} D_k \cdot dS + \oint_S \sum_{k=N+1}^{M+N} D_k \cdot dS$$

(1-15)

where the flux density components have been divided into those associated with charges inside and outside the surface, respectively. Because none of the charges resides in the surface itself, all the flux density components are known to be mathematically continuous everywhere on the surface. The operations of integration and summation may therefore be interchanged. For each charge located outside the surface, the net flux through the surface is known to be zero, so that the total of the contributions of all outside charges is zero. There remain the contributions of the inside charges:

$$\oint_S D \cdot dS = \sum_{k=1}^{N} \oint_S D_k \cdot dS = \sum_{k=1}^{N} \psi_k$$

(1-16)

or simply

$$\oint_S D \cdot dS = Q_{enc}$$

(1-17)

where $Q_{enc}$ is the total charge enclosed by the closed surface of integration. This latter statement is known as the integral form of Gauss's law.

A differential form of Gauss's law is equally useful and is easily developed from (1-17). Let there be a continuous distribution of charges inside the closed surface instead of a set of distinct particles. This situation may be described by specifying the volume density $\rho$ of charge (coulombs per cubic meter) at all interior points. The foregoing development is still applicable if the volume is subdivided into volume elements $dU$, each of which contains a charge of $\rho \ dU$ coulombs. The number $N$ of interior charges is then very great, the magnitude of each charge very small, and the summation of internal charges is best carried out as an integration:

$$Q_{enc} = \int_{U} \rho \ dU$$

(1-18)

so that the integral form of Gauss's law may be converted into

$$\oint_S D \cdot dS = \int_{U} \rho \ dU$$

(1-19)

where $U$ is the volume bounded by the closed surface $S$. Applying the divergence theorem of vector analysis, the left-hand member may be converted into a volume integral also:
\[
\int_V \text{div } \mathbf{D} \, dU = \int_V \rho \, dU
\]  
(1-20)

Since the choice of volume of integration is entirely arbitrary (the equation being true for any volume), the integrands must be equal:
\[
\text{div } \mathbf{D} = \rho
\]  
(1-21)

This result is known as the \textit{differential form of Gauss's law}. The physical significance of both forms is of course the same. Either way, the law states that there are no charges without flux and that no flux line has an end or beginning except where there is a charge.

Gauss's law in either form amounts to a reassertion of the continuity principle. As corollaries, the idea of an electrically neutral universe and the law of conservation of charge may be demonstrated to be implicit in it. Let a set of charges be located in some region \( R \) of space. Let the region \( R \) be bounded by the surface \( S_1 \) and contain a total charge \( Q_R \). Gauss's law in integral form states
\[
\oint_{S_1} \mathbf{D} \cdot d\mathbf{S}_1 = Q_R
\]  
(1-22)

Next, let a surface be constructed that encloses all space except the region \( R \), say \( S_2 \). This second surface must enclose all the charges in the universe, \( Q_{\text{universe}} \), except for those in the region \( R \). The integral form of Gauss's law then requires
\[
\oint_{S_2} \mathbf{D} \cdot d\mathbf{S}_2 = Q_{\text{universe}} - Q_R
\]  
(1-23)

Now it is evident that the two surfaces \( S_1 \) and \( S_2 \) must be identical except that the inside of one is the outside of the other. Therefore
\[
d\mathbf{S}_1 = -d\mathbf{S}_2
\]  
(1-24)

for the outward normal of a surface is conventionally taken to be positive. Combining Equations (1-22) to (1-24), there is obtained
\[
Q_{\text{universe}} = 0
\]  
(1-25)

that is to say, there exists altogether exactly as much positive charge as negative in the universe.

The law of conservation of charge is in turn a corollary of the electrical neutrality of the universe. If it were possible to create charge in any region \( R \), the total amount of charge in the universe would grow and the neutrality principle would be violated.

\section*{5. Coulomb's Law}

The Coulomb experiment indicates that the force experienced by a charge varies with the magnitude and sign of the charge. It also suggests that the influence of another charge must be somehow perceptible. In order to provide
Sec. 5 ORIGINS AND NATURE OF THE ELECTRIC FIELD

a reasonable mechanism for the required transmission of influence through space, the notion of electric flux was introduced. The experiment will now be examined to see whether the flux postulate can in fact adequately account for the forces encountered.

Let two charged particles be placed in space, far away from all other charges. If they are of like polarity, there will exist a repulsive force between them. If the particles are unrestrained, they will accelerate away from each other at a rate determined by their charges and masses. Since there is no \textit{a priori} limitation on either parameter, relativistically important velocities are apt to be encountered; the problem becomes very complicated very quickly. To avoid difficulties, the particles will be restrained from moving by tying one end of a very light, thin, nonconductive thread to each, as in Fig. 1.4.

The force on either particle in equilibrium is the same in magnitude as the force on the other, but opposite in direction. Either force must in fact equal the thread tension. The electrical quantities perceptible at particle \( A \) are the flux density due to the other particle, say \( D_b \), and its own charge \( q_a \); similarly, at particle \( B \), only the flux density \( D_a \) of the first particle, and the charge \( q_b \), can be ascertained. The forces are equal in magnitude, requiring that the mathematical description of the force must have a form that remains unaltered by interchange of the subscripts \( a \) and \( b \), except for reversal of signs. Using Equation (1-9), the flux densities \( D_a \) and \( D_b \) may be written out in detail. They are

\[
D_a = \frac{q_a}{4\pi r^2} \quad D_b = \frac{q_b}{4\pi r^2}
\]

For the physical reasons enumerated, only one flux density and one charge can affect the force on either particle. Only one simple algebraic form fulfills the requirement of subscript interchangeability, the product

\[
D_a q_b = D_b q_a = \frac{(q_a q_b)}{4\pi r^2} \quad (1-26)
\]

Other simple forms (sum, difference, quotient) are seen by trial not to be satisfactory; more complicated forms can always be rewritten as functions of the elementary combinations. Consequently, the force must be a vector quantity in the direction of the thread and of magnitude related to the product \( D_a q_b \). That is,

\[
F = f(D_a q_b) \mathbf{1}_{D_a} \quad (1-27)
\]

The precise nature of the functional relationship must be determined experimentally. Happily, the force in free space turns out to be simply proportional
to the product of charge and flux density. The proportionality is usually written

\[ \mathbf{F} = \frac{D_q q_b}{\varepsilon_0} \quad (1-28) \]

This experiment is the second and final one theoretically required for the development of electromagnetic theory; no further experimental data will be needed to provide material for the argument. Many distinct experiments were necessary, however, to guide the development of the theory, and many more will be needed to suggest new directions as well as to verify correctness of the development and the approximations made.

The force between two point particles is often written out in detail as

\[ \mathbf{F} = \frac{q_a q_b}{4\pi \varepsilon_0 r^2} \mathbf{r} \quad (1-29) \]

In this form there is no clear indication of which force is being calculated and what its direction is, for the vector \( \mathbf{r} \) is in no evident manner related to the force. In contrast, the simpler form (1-28) includes the direction of the force in the flux density vector and is quite unambiguous. The symmetrical form (1-29) will therefore be used with some reserve, and (1-28) will be preferred in further development.

The constant \( \varepsilon_0 \) is written in the denominator for historical reasons. Its numerical value is \( 8.85 \times 10^{-12} \) coulombs squared per joule per meter. The coulomb is a unit that originated in electrochemistry and is inconveniently large for electrostatic force calculations; the microcoulomb is more commonly used.

The ability of an electric flux distribution to produce force on a charged particle is called electric field intensity and denoted by the letter \( \mathbf{E} \). In mathematical terms, this vector is defined as the force per unit charge or the force on a unit test charge:

\[ \mathbf{E} = \frac{\mathbf{F}}{q} \quad (1-30) \]

\( \mathbf{E} \) may be related to \( \mathbf{D} \) using (1-28). Given a set of \( N \) point charges, the test charge \( q \) experiences a force component \( \mathbf{F}_k \) due to each charge \( q_k \). The total force on \( q \) must be the sum of all the \( N \) force components. Thus

\[ \mathbf{F} = \sum_{k=1}^{N} \mathbf{F}_k = \sum_{k=1}^{N} \frac{1}{\varepsilon_0} q \mathbf{D}_k \quad (1-31) \]

But the sum of the individual flux densities is, as previously, the total flux density \( \mathbf{D} \) at the location of the test charge \( q \). The force may therefore be written

\[ \mathbf{F} = \frac{1}{\varepsilon_0} \mathbf{D} q \quad (1-32) \]
Sec. 6 ORIGINS AND NATURE OF THE ELECTRIC FIELD

On comparing with the definition of $E$, (1-30), the relationship between $D$ and $E$ in free space may be stated as

$$D = \varepsilon_0 E$$

From its definition, the field intensity $E$ may be seen to have units of newtons per coulomb in the RMKSC system.

It may well be asked why it should be necessary to bother to introduce separate vectors $D$ and $E$ at all—ought not the electric flux of a charge $q$ be defined so as to make $\varepsilon_0 = 1$ in free space? The advantages gained by so doing are unfortunately small because the simplification won in electrostatics is offset by additional complexities that result in magnetic field problems. In a similar manner, the factor $4\pi$ of Equation (1-29) may be eliminated by introducing it as the constant $K_0$ in Equation (1-1), the continuity postulate. The units of flux and charge will then differ by a factor of $4\pi$ and ought therefore to have different names. The advantages and disadvantages of the resulting possible systems of units have been discussed at length by numerous authors. As already pointed out, the RMKSC system will be followed here, largely because it is employed by the majority of contemporary advanced texts and most periodical articles.

6. Fields in Material Media

Implicit in the continuity postulate is the idea of all matter being filled with large quantities of positive and negative charges, uniformly and homogeneously mixed to produce exact cancellation of external flux. If charged particles are immersed in a material medium, an electric force must tend to move the positive charges in the direction of the flux lines and the negative charges in the opposite direction. If the medium is a conductor, the charges will indeed move until a new equilibrium state is established. On the other hand, in nonconducting media the masses of charge can only be displaced slightly from their normal equilibrium positions. The precise mechanisms by which such a displacement occurs are dealt with in the atomic theory of matter and will not be entered into here. Simply stated, it is imagined that each atom consists of both positive and negative charges arranged in concentric spherical shells. These are thought to shift slightly under the influence of external flux so that their centers no longer coincide and the atoms appear to carry a slight positive charge on one side and a slight negative charge on the other.

Let a rectangular piece of matter be placed in the neighborhood of a charge distribution. For simplicity, the charge distribution will be assumed to be far enough away to make the flux density nearly uniform over the surface of the object and normal to the surface, as in Fig. 1.5. The piece of matter itself is taken to be electrically neutral. If it is a nonconductor, the positive and negative charges in it must shift as already indicated. The shift will not be
discernible inside the slab because the movement of charges away from any point will be compensated by movement of other charges toward it, leaving the total charge densities at their original values. However, at the edges of the material these motions of charges cannot be compensated because there are no charges outside the surface. Hence the faces of the slab acquire an apparent charge as indicated in Fig. 1.5. The thickness of the charge layer must be of the order of atomic dimensions; for macroscopic purposes the layer will be assumed to have infinitesimal thickness. The material is said to have become polarized.

Surface charges resulting from polarization have a flux associated with them, like any other charges. Inside the material, the direction of this "secondary" flux opposes the externally imposed flux, so that the net flux density inside the material is lowered (in comparison with its value in the absence of polarizable material). The net reduction in flux density (i.e., the negative of the secondary flux density) is conventionally called the polarization vector, and denoted by the letter \( \mathbf{P} \). Since it arises from the shift of internal charges by electric forces, its magnitude must be dependent on the electric field intensity \( \mathbf{E} \) inside the material. For a large class of materials (though by no means all) this dependence is happily very nearly describable by a simple proportionality:

\[
\mathbf{P} = \chi \varepsilon_0 \mathbf{E} \tag{1-34}
\]

where \( \chi \) is called the electric susceptibility. The force on a test charge inside the material is still \( \mathbf{F} = q \mathbf{E} \) by definition of electric field intensity. The latter may be calculated from the flux density as usual, remembering that the flux density must now include the polarization vector. Thus \( \mathbf{E} \) may be found from

\[
(\mathbf{D} - \mathbf{P}) = \varepsilon_0 \mathbf{E} \tag{1-35}
\]

The polarization vector may be eliminated by combining Equations (1-34) and (1-35). There results

\[
\mathbf{D} = \varepsilon_0 (1 + \chi) \mathbf{E} \tag{1-36}
\]
which is usually written

\[ \mathbf{D} = e\mathbf{E} \quad \text{where} \quad e = \epsilon' \epsilon_0 = (1 + \chi)e_0 \quad (1-37) \]

The quantity \( e \) is called the permittivity of the material, and the dimensionless multiplier \( \epsilon' \) the relative permittivity. Free space of course has \( \epsilon' = 1 \); extending the terminology of material media, \( e_0 \) is termed the permittivity of free space. A material is described for electrical purposes by specifying either its susceptibility or its relative permittivity; the two are simply related and specifying one determines the other.

The effect of polarization may evidently be accounted for either by adding a term \(-\mathbf{P}\) to flux density or by modifying the value of permittivity. The former technique is usually preferred for investigations into the structure of matter, for it exhibits clearly the different mechanisms involved. The latter course permits mathematical simplification, for the sole distinction between the field equations in regions containing matter, or in free space, becomes a distinction in the numerical value of permittivity. The relative permittivity of air is very nearly unity (about 1.006) at STP; those of commercial insulating materials vary from unity to perhaps 10. Values as high as \(10^3\) are encountered in some barium compounds.

It is assumed in the argument above that the material is nonconductive, that is, that the positive and negative charges in it may be displaced slightly but are not free to move away from their normal locations. In conducting materials, on the other hand, they are free to move about and will do so in response to electric forces. No new equilibrium will be reached as long as such forces exist. When an equilibrium state is eventually established, no such forces can remain, and the field vector, by Equation (1-30), must be zero at all points inside the material. This can only happen when enough charges have been moved to the material surfaces to make the external flux density cancel the flux density due to the surface charges.

### 7. Scalar Potential of Conservative Fields

An electric field always exerts a force on any charged particle. To move a charge into an electric field, work must be done on it; conversely, work may be obtained in permitting the field to push the particle. The field must therefore retain stored energy, to be recalled as work when necessary.

For a simple two-particle situation, the amount of energy involved in a change of relative position is readily calculated. Let a particle of charge \( q_0 \) be placed in space, with another charge \( q \) somewhere it its neighborhood. By Equations (1-9) and (1-33) the electric field of charge \( q_0 \) is given by

\[ \mathbf{E} = \frac{q_0}{4\pi\epsilon_0 r^2} \mathbf{1}_r \quad (1-38) \]
The work done by an external agency in moving a particle against a force \( F \) is, by definition,

\[ W = \int_{r_1}^{r_2} (-F) \cdot dr \]  

(1-39)

Substitution of the force on the moved particle, (1-30), into this general expression gives the work done in moving the charge \( q \) from a distance \( r_1 \) to a distance \( r_2 \) from the charge \( q_0 \):

\[ W = \frac{q_0}{4\pi\epsilon} \left( \frac{1}{r_2} - \frac{1}{r_1} \right) \]  

(1-40)

The work done on the charged particle in permitting it to be pushed back to its starting position is just exactly the negative of the above. It follows that the net amount of work done in moving the particle from its original location to somewhere else and back again is zero.

An important corollary of this last conclusion is that the work that must be furnished to move a charged particle from one place to another is not dependent on the path along which it is moved. If the work were path-dependent, then two paths could be found, one requiring more work than the other. The particle could then be taken from its starting point along the path requiring little work and returned along the path of greater work. On return to the starting point, a quantity of energy would be left over, equal to the difference in the amounts of work along the two paths. By performing the motion along the same contour repeatedly, any desired amount of energy could be manufactured. As that is impossible, the difference in work along any closed path must be zero.

This energy-conservation argument has general applicability and is not restricted to a pair of charges. Given any electric field set up by any collection of charges, the forces on a charged particle must be such as to make energy creation impossible. The work done in moving a charge \( q \) around any closed contour must thus be zero:

\[ W = \oint (-F) \cdot dr = 0 \]  

(1-41)

Using the definition of electric field intensity once again, there results

\[ \oint E \cdot dr = 0 \]  

(1-42)

which may be rewritten by Stokes’ theorem as

\[ \text{curl} \ E = 0 \]  

(1-43)

since the contour around which integration is carried out is entirely arbitrary. Because its curl is zero, the electric field is often called an irrotational field; an equally descriptive term is conservative field. Any conservative force field is readily seen to be irrotational, and vice versa.

Suppose a distribution of charges to be established in space. Very far away from the charge centers, their flux density \( D \) and the associated electric
Sec. 8 ORIGINS AND NATURE OF THE ELECTRIC FIELD

Field \( E \) are negligibly small. Let a particle with very small charge \( \Delta q \) be moved from a very distant point to some point \( P \) where the field has a finite value. The work \( W \) done is given by

\[
W = \int_{\infty}^{P} (-F) \cdot d\mathbf{r} = \Delta q \int_{\infty}^{P} -E \cdot d\mathbf{r}
\]

(1-44)

The charge \( \Delta q \) is taken as very small lest its presence in any way disturb the charges already established in space and the energy associated with them. Only the additional energy of introducing \( \Delta q \) is calculated here. Because the electric field is a conservative field, the quantity

\[
V = \frac{\Delta W}{\Delta q} = -\int_{\infty}^{P} E \cdot d\mathbf{r}
\]

(1-45)

is not dependent on the path of integration but only on the end point \( P \). That is, \( V \) is a scalar position function, an expression of the energy associated with the charge \( \Delta q \) in the field \( E \). It is usually called electric scalar potential or more simply electric potential and measured in volts. It is evident from the foregoing equations that a volt is one joule per coulomb.

The potential at any point in an electric field may be calculated using Equation (1-45). The real usefulness of the notion of potential lies in its scalar nature; in most problems it is easier to deal with scalars than vectors, so that the potential \( V \) is more easily found than the field intensity \( E \). The latter must then be derived from the potential. Recalling that the operations of line integration and forming the gradient are inverse,

\[
\varphi = \int \text{grad } \varphi \cdot d\mathbf{r}
\]

the field intensity may be found from the potential as

\[
E = -\text{grad } V
\]

(1-46)

For this reason, the field intensity is sometimes also called the potential gradient, although it is really the negative of the gradient. The irrotational and hence conservative nature of the electric field is implicit in (1-46). Since curl \( \text{grad } \varphi = 0 \) for any function \( \varphi \), there immediately obtains

\[
\text{curl } E = -\text{curl grad } V = 0
\]

(1-47)

In general, any conservative field (e.g., gravitational fields, velocities of perfect fluids) may be expressed in terms of a scalar potential.

8. General Solution of the Potential Problem

There would be little purpose in defining electric potential if it were always to be calculated from its defining Equation (1-45). Instead, the electric field is to be derived from relatively simple solutions for the potential itself; a method of finding potential directly from a given charge distribution is therefore necessary.
For a single point charge $q_o$, the electric field is given by

$$E = \frac{1}{r} \frac{q_o}{4\pi \epsilon r^2}$$

from which the corresponding potential distribution may be found, in accordance with (1-45), as

$$V = \frac{-q_o}{4\pi \epsilon} \int_{r_0}^{r} \frac{dr}{r^2} = \frac{q_o}{4\pi \epsilon r}$$

It is seen that the potential is symmetrically distributed around the point charge. Surfaces of constant potential form concentric spherical shells centered on the charge, with the potential rising toward the center. A true point charge, it might be noted, leads to an infinite potential.

The potential distribution of a general charge distribution $\rho$ may be found by extending the solution (1-48). Any continuous spatial distribution $\rho(x, y, z)$ may be imagined to be composed of small elements of charge $\rho \, dU$. Each of these must have associated with it a potential distribution similar to (1-48) but of very small magnitude $dV$:

$$dV = \frac{dq}{4\pi \epsilon r} = \frac{\rho \, dU}{4\pi \epsilon r}$$

where the distance $r$ is measured from the point of observation to the location of the volume element $dU$. This distance is of course different for each volume element. The potential $V$ at the point of observation, arising from all the charge distribution $\rho$, may be found by adding up the contributions of all the individual elements. There is found

$$V = \frac{1}{4\pi \epsilon} \int \frac{\rho \, dU}{r}$$

It should be noted that this result relies upon Equation (1-38) for its validity and can only hold wherever the latter is true. That is to say, (1-50) is valid for homogeneous, linear, isotropic media only. Nevertheless, it is a formula of considerable practical utility. The appearance of $\rho$ in the integrand permits the volume of integration to be confined to those portions of space actually containing charges; in many practical problems, this fact makes either analytic or numerical solution of (1-50) relatively easy.

An example of potential calculation from a known charge distribution will help to clarify this point. Let the potential at some arbitrary point $P$ due to a uniformly charged fine wire ring of radius $R$, as shown in Fig. 1.6a, be sought. In view of the rotational symmetry of this problem, it is sufficient to calculate the potential at the point $(r, 0, z)$, for the potential cannot depend on the angle $\theta$. The ring may be imagined to be divided into small volume elements $dU$,

$$dU = aR \, d\theta$$
where \( a \) represents the cross-sectional area of the ring. The distance \( s \) from an arbitrary volume element located at \((R, \theta, 0)\) to the point \((r, 0, z)\) is readily found to be

\[
s = \sqrt{z^2 + R^2 + r^2 - 2Rr \cos \theta}
\]

so that the integral of (1-50) becomes

\[
V = \frac{aR \rho}{4\pi \varepsilon} \int_0^{2\pi} \frac{d\theta}{\sqrt{z^2 + R^2 + r^2 - 2Rr \cos \theta}}
\]

This integral may be evaluated if the substitution

\[
\theta = \pi - 2\varphi
\]
is made so that

\[ V = \frac{a \rho}{2 \pi \epsilon} \sqrt{\frac{R}{r}} k \int_{0}^{\pi/2} \frac{d\varphi}{\sqrt{1 - k^2 \sin^2 \varphi}} \]  

(1-55)

where \( k \) represents the combination of quantities

\[ k = \frac{2 \sqrt{Rr}}{\sqrt{z^2 + (R + r)^2}} \]  

(1-56)

The remaining integral in (1-55) is of a known type. It is called the complete elliptic integral of the first kind to the modulus \( k \) and tabulated in many collections of mathematical tables, where the usual notation is

\[ K(k) = \int_{0}^{\pi/2} \frac{d\varphi}{\sqrt{1 - k^2 \sin^2 \varphi}} \]  

(1-57)

Finally, the desired result is

\[ V = \frac{a \rho}{2 \pi \epsilon} \sqrt{\frac{R}{r}} k K(k) \]  

(1-58)

which may be evaluated numerically using tables or, preferably, by means of a digital computer. A short table of \( K(k) \) appears in Appendix III, along with a computer subroutine ELink(\( z \)) to generate values of \( K(z) \). The equipotential contours resulting from one such evaluation are shown in Fig. 1.6b.

9. Total Energy of a Charge Distribution

The potential \( V \) was defined, Equation (1-45), by considering the energy required to introduce a very small test charge \( \Delta q \) into the field. The total energy associated with the charge distribution cannot be found by simply increasing the size of \( \Delta q \), because the introduction of a significant quantity of charge will alter the field distribution and change the forces on all charges in the distribution. A slight refinement of the argument is necessary to find the total energy \( W \) associated with the field \( \mathbf{E} \) and potential \( V \) of a charge distribution \( \rho \).

From the continuity principle it follows that the flux density \( \mathbf{D} \) of a charge distribution becomes \( h \mathbf{D} \) if all charges in the distribution are multiplied by \( h \), for the total flux must increase by the factor \( h \) but remain distributed in exactly its previous manner. The field strength \( \mathbf{E} \) and potential \( V \) must then become \( h \mathbf{E} \) and \( hV \), respectively. The conservative property of electric fields requires that the total energy of a charge distribution cannot depend on the manner in which the distribution was established. Suppose then that the distribution is established by distributing a small amount of charge in space so as to have the distribution \( h \rho \), where \( h \) is a small number. For a very small \( h \), very nearly no energy is required. The eventual distribution is built up stepwise by adding a tiny amount of charge \( \rho \, dh \) to all parts of the distribu-
tion, then a little more, and so on until \( h = 1 \) and the final distribution has been reached. At each step of this process, the additional charge \( \rho \, dh \) is introduced into a potential \( hV \). The work \( dW \) required to do so is

\[
dW = hV \int (\rho \, dh) \, dU
\]

(1-59)

the volume of integration being all space. The final values \( \rho \) and \( V \) are of course independent of the interim multiplier \( h \). The total work required from start to finish, \( h = 0 \) to \( h = 1 \), is then

\[
W = \int_0^1 h \, dh \int V \rho \, dU = \frac{1}{2} \int V \rho \, dU
\]

(1-60)

As a practical matter, determination of the precise spatial distribution of charges is sometimes more difficult than determination of the field vectors \( \mathbf{D} \) and \( \mathbf{E} \). For such cases the energy may be rewritten in terms of the field vectors. Making use of the differential form of Gauss's law, the energy expression is

\[
W = \frac{1}{2} \int V \, \text{div} \, \mathbf{D} \, dU
\]

(1-61)

This equation may be rewritten, using one of the well-known vector identities, in the form

\[
W = \frac{1}{2} \int \text{div} (VD) \, dU - \frac{1}{2} \int \mathbf{D} \cdot \text{grad} \, V \, dU
\]

(1-62)

Substituting the field intensity for \( \text{grad} \, V \) and applying the divergence theorem to the first member of the right-hand side, the energy becomes

\[
W = \frac{1}{2} \oint V \mathbf{D} \cdot d\mathbf{S} + \frac{1}{2} \int \mathbf{E} \cdot \mathbf{D} \, dU
\]

(1-63)

The integration again has to be carried out over all space in order to include all the energy. The surface of integration for the left-hand integral is then infinitely far away from the center of the charge distribution; by definition, the potential at infinite distance is zero. Hence

\[
W = \frac{1}{2} \int \mathbf{E} \cdot \mathbf{D} \, dU
\]

(1-64)

Just exactly where in space the energy resides is not at all clear from this expression. Because the volume of integration must extend over all space, it does not really matter; upon integration, all the energy will be included anyway. For purposes of visualization, however, it is very convenient to imagine that the energy is stored in the field itself and hence distributed throughout space, with a density of \( \frac{1}{2} \mathbf{D} \cdot \mathbf{E} \) joules per cubic meter. If this viewpoint is taken, the energy in any element of space \( dU \) is simply \( \frac{1}{2} \mathbf{D} \cdot \mathbf{E} \, dU \), and the total energy is the integral of the density over all space. No
contradiction arises with the mathematical development, but a useful physical picture is gained from this definition of energy density.

10. Properties of Scalar Potentials

The relationship between flux density and electric field strength may next be exploited to investigate the laws governing electric potential. Gauss’s law in differential form,

\[ \text{div } \mathbf{D} = \rho \]  
(1-21)

may be used to describe the electric field in the neighborhood of charges by invoking the relationship \( \mathbf{D} = \varepsilon \mathbf{E} \). There results

\[ \text{div } \mathbf{E} = \frac{\rho}{\varepsilon} \]  
(1-65)

The field intensity may next be written in terms of the scalar potential so as to obtain a single equation that embodies the continuity principle (through Gauss’s law) as well as the law of Coulomb (by way of the irrotationality of field intensity). Substituting \( \mathbf{E} = -\text{grad } V \) in (1-65), there obtains

\[ \text{div } \text{grad } V = -\frac{\rho}{\varepsilon} \]  
(1-66)

This equation is called Poisson’s equation. In all portions of space containing no charges, the right-hand member is zero, and Poisson’s equation simplifies to

\[ \text{div } \text{grad } V = 0 \]  
(1-67)

which is called Laplace’s equation.

The maximum and minimum points of the potential \( V \) in any field may be found directly from the Poisson and Laplace equations. At either kind of extremum the space rate of change of potential must be zero, for there cannot be any rise possible beyond a maximum nor any descent beyond a minimum. The gradient of potential must therefore be zero at all extrema. In order to make this possible, all three space derivatives of \( V \) must vanish; in Cartesian coordinates

\[ \frac{\partial V}{\partial x} = \frac{\partial V}{\partial y} = \frac{\partial V}{\partial z} = 0 \]  
(1-68)

Equivalent equations hold true in other coordinate systems. The second derivatives of \( V \) must have positive values at a minimum and negative values at a maximum; zero values indicate a point of inflection. At a true maximum point, all three second derivatives must indicate a maximum. Conversely, a true minimum point requires that all three second derivatives be positive. Then \( \text{div } \text{grad } V \) must be negative or positive, respectively. To summarize,
At potential maxima, \( \text{div grad } V < 0 \)
At potential minima, \( \text{div grad } V > 0 \)

From Poisson's equation it may now be inferred that a true maximum is only possible at a positive charge center, a true minimum at a negative charge center. This general principle may be employed to determine the conditions for stable equilibrium of a charged particle in an electric field. The net force on any particle in equilibrium must be zero. If none but electrical forces are applied to the particle, then the field intensity \( E \) must be zero also; hence \( \text{grad } V = 0 \). For stable equilibrium, any small displacement of the particle must give rise to restoring forces that tend to move the particle back to its equilibrium position. If the particle is assumed to be positively charged, then the restoring forces must arise from an electric field that is directed toward the equilibrium position, as indicated in Fig. 1.7. With the field intensity vector \( E \) directed inward everywhere in the neighborhood of the equilibrium point, the divergence of \( E \) is clearly negative; hence \( \text{div grad } V \) is positive. Poisson's equation states that such a potential distribution can only be associated with a negative charge at the equilibrium position, in contradiction with the initial assumption that the charge in equilibrium is positive. Assuming a negative charge in equilibrium naturally reverses all directions in the entire argument and thus requires that the charge assumed negative must be positive. Only one conclusion is possible: *no charges can find a stable equilibrium position under the influence of electrical forces alone.* This statement is known as Earnshaw's theorem.

An important corollary to Earnshaw's theorem is that mechanical forces must be applied to create an equilibrium state. In conductive bodies (not necessarily perfect conductors), charges are free to move about in response to Coulomb forces; they can only be restrained mechanically at the conductor boundaries. In equilibrium, a conductive body must therefore be electrically neutral except at its surfaces. This conclusion was arrived at by direct physical reasoning earlier; Earnshaw's theorem provides a formalized version of the argument.

Since the potential \( V \) may possess maxima and minima only at charge centers, it must vary in a continuous manner between these extreme values in charge-free regions of space. Hence it must be possible to find regions or surfaces of constant \( V \) in any charge-free portion of space; they are known as equipotential surfaces or regions.

By definition, potential represents the work per unit charge required to bring a test charge from a place of zero potential to its actual location. This
amount of work is clearly the same for all points of an equipotential surface, for all points of the surface have the same potential \( V \). Consequently, zero work is done in transferring a test charge \( \Delta q \) from some point \( A \) to another point \( B \) on the equipotential surface. Mathematically stated,

\[
\int_A^B \mathbf{F} \cdot d\mathbf{r} = 0 \quad (1-70)
\]

If it is supposed that \( A \) and \( B \) are neighboring points, i.e., that they are only a very small distance \( \Delta \mathbf{r} \) apart, then this expression may be simplified to

\[
\mathbf{F} \cdot \Delta \mathbf{r} = 0 \quad (1-71)
\]

Since the two points \( A \) and \( B \) are arbitrary and may be chosen anywhere in the surface, the vector \( \Delta \mathbf{r} \) is only specified as tangential to the equipotential surface. Its actual direction in the surface is arbitrary. By (1-71), the vector \( \mathbf{F} \) must always be at right angles to \( \Delta \mathbf{r} \), and hence normal to the equipotential surface. But in free space (or any isotropic material) the vectors \( \mathbf{F} \), \( \mathbf{E} \), and \( \mathbf{D} \) are collinear and have the direction of flux lines. Hence flux lines always meet equipotential surfaces at right angles.

The same conclusion may be arrived at by starting from Equations (1-46) and (1-65). From vector analysis, grad \( V \) is known to be a vector normal to surfaces \( V = \text{constant} \). The statement that the divergence of this vector is zero (Laplace’s equation) then physically means that lines normal to equipotential surfaces must not begin or end in charge-free space.

The orthogonality property of Equation (1-71) has two very important corollaries, as follows:

1. No two equipotential surfaces are ever allowed to touch. For if a point of contact did exist, the work required to bring a test charge to that point would depend on the path along which it was moved. That, however, contradicts the conservative character of the electric field, which requires that work done be independent of path.

2. No two flux lines are ever allowed to cross. For if a point of intersection did exist, the equipotential surface through that point would have to be normal to both flux lines. The flux lines must therefore be parallel and go through the same point. They are therefore the same flux line.

Both of these corollaries are merely physical interpretations of Laplace’s equation. Laplace’s equation is thus the mathematical equivalent of saying: Find a family of continuous nonintersecting surfaces (the equipotentials) and a family of continuous nonintersecting lines normal to these surfaces (the flux lines).

It is immediately clear that this problem has an infinite number of solutions. In order to make it distinct, boundary conditions must be added to the equation. Physically, these amount to prescribing one or more flux lines that must be members of the family of lines.
11. Fields at Material Boundaries

Mathematically speaking, electrostatic field problems are problems in Poisson's or Laplace's equations, for which suitable boundary values must be specified in each case. The natural boundary conditions of a problem are sometimes known potential values, at other times known charge densities. Wherever conducting media appear in static field problems, charges reside in the material surfaces as a consequence of Earnshaw's theorem. In dielectric solids, charges may appear either through contact with another charged object, in which case they cannot be placed inside the medium, or they may appear as induced charges, which always occupy surface positions. Surface charges consequently play an important role in electrostatics, and a knowledge of how they affect field values at material boundaries is necessary.

Suppose that a surface \( S \) has a total charge \( q \) deposited on it in some continuous fashion. To an elemental portion of the surface, say \( dS \), there corresponds a charge \( dq \). The surface charge density \( \sigma \) will be defined as the quantity of charge per unit surface so that

\[
dq = \sigma \cdot dS
\]

Because \( dS \) is a vector quantity, \( \sigma \) must be a vector also. In the RMKSC system, the unit of surface charge density is the coulomb per square meter.

Consider a body (it does not matter whether it is a conductor or not) with a surface charge density \( \sigma \). A pancake-shaped surface may be constructed around part of the material boundary by taking two like areas, one just inside and one just outside, and connecting them with a thin strip normal to the material as indicated in Fig. 1.8. The result is a closed surface to which Gauss's law may be applied. It may seem at first glance that some of the surface charge will reside in the Gaussian surface also so that the theorem is inapplicable. The amount of charge located at the intersection of surfaces, however, must be infinitesimally small, and whether it is reckoned as being included or excluded makes no difference. The charge inside is given by the surface integral of charge density:

\[
dq = \sigma \cdot dS
\]
\[ q_{\text{enc}} = \oint_S \sigma \cdot dS \]  
\hspace{1cm} (1-73)

The thin bordering strip may be made as narrow as desired; hence the total flux through it arbitrarily small. The surface integral of Gauss’s law then includes merely the top and bottom surfaces \( T \) and \( B \). Equation (1-73) may be written out as

\[ q_{\text{enc}} = \int_T \mathbf{D} \cdot dS + \int_B \mathbf{D} \cdot dS = \int_S \sigma \cdot dS \]  
\hspace{1cm} (1-74)

With the separation between surfaces \( T \) and \( B \) taken as very small, \( T, S, \) and \( B \) have very nearly the same area. Only the normal components of \( \mathbf{D} \) matter in the surface integration. Thus, (1-74) may be simplified to read

\[ D_{\text{normal outside}} - D_{\text{normal inside}} = \sigma \]  
\hspace{1cm} (1-75)

This is the relationship that must be satisfied by the normal components of flux density if the surface \( T \) is thought of as outside surface \( S \); if \( B \) is thought to be outside, the sign of \( \sigma \) must be changed.

The tangential field components may be investigated using the irrotationality property. A path may be constructed of straight-line segments as shown in Fig. 1.9, the long portions of the path being parallel to the surface

![Fig. 1.9](image)

and the short portions normal to it. One of the long portions is taken just to either side of the interface between the two media, i.e., on either side of the boundary. By Stokes’ theorem, the irrotationality of the electric field requires

\[ \oint \mathbf{E} \cdot d\mathbf{r} = 0 \]  
\hspace{1cm} (1-76)

The short portions of the path of integration may be made arbitrarily short, and their contribution to the integral thereby reduced to a negligible amount. The long paths give

\[ \int_{AB} \mathbf{E} \cdot d\mathbf{r} + \int_{CD} \mathbf{E} \cdot d\mathbf{r} = 0 \]  
\hspace{1cm} (1-77)

Only the tangential components of \( \mathbf{E} \) matter in this integration since the two paths are parallel to the surface; normal components are eliminated in the scalar products. If the end paths \( BC \) and \( DA \) are very short, the two long paths are of equal length so that

\[ E_{1\text{ (tang)}} = E_{2\text{ (tang)}} \]  
\hspace{1cm} (1-78)

that is, the tangential field component is unaffected by any charges in the
material surface. An interesting corollary is that any conductor surface must be an equipotential. Inside the conductor, the electric field is always zero; by Equation (1-78) the tangential field outside is also zero. The flux lines are therefore always at right angles to a conductor surface, proving that the surface must be at the same potential throughout.

**READINGS**

The historical and philosophical development of electromagnetism is intimately related to the history of physics as a whole. Fortunately, this aspect of physics is of sufficient intrinsic interest to have occupied the attention of some eminent scientists and engineers, as well as historians. A very clear concise account is given by Loeb (1), whose book incidentally is also a good general reference on introductory electromagnetic theory. More quantitative accounts written for a scientific audience will be found in Whittaker (2) and Lorentz (3), both of whom give clear explanations of the various viewpoints on the nature of charge and flux. The concept of action at a distance is argued eloquently by Maxwell (4). For other references, the excellent annotated bibliography by Scott (5) may be consulted.

Virtually all books on electromagnetic theory include sections on Gauss’s and Coulomb’s laws, although the level varies from book to book. To be recommended are the treatments given by Reitz and Milford (6) and Harnwell (7). Those of Langmuir (8) or Plonsey and Collin (9) are a little more sophisticated, while Rogers (10) or Hayt (11) may furnish instructive background reading.

The subject of dielectric media has not been treated to any degree of depth here for it is presumed that the reader is best advised to seek such information in books on solid-state physics and materials science. However, the relevant chapters in Langmuir (8), Reitz and Milford (6), Plonsey and Collin (9), and Harnwell (7) will provide clear summaries of a potentially difficult subject. Nussbaum (12) devotes a well-written chapter to this topic and works out numerous examples. A profound analysis at a challenging level will be found in Chapter 2 of Landau and Lifshitz (13), one of the classical books of contemporary physics.

Practically every book on this subject devotes at least one chapter to electrostatic fields and electrostatic energy. In addition to the books of Reitz and Milford (6) and Langmuir (8) already referred to, Corson and Lorrain (14) deserve to be especially recommended, as do Slater and Frank (15). Stratton (16) is a more advanced work but may be consulted with profit.

The student seeking to strengthen his comprehension by working suitable problems will find a wide selection in the books already enumerated. In addition, the little volume by Leatham (17) will be found to contain some 80 prob-
lems on electrostatics, most of them well beyond the formula-substitution level. A very comprehensive set of problems, covering all areas of electromagnetism and arranged by topic, has been compiled by Batygin and Toptygin (18); virtually all of them are supplied with either answers or complete solutions.

An exceptionally clear description of the various systems of units used, and the reasons for their existence, will be found in Jackson (19).


Prob. ORIGINS AND NATURE OF THE ELECTRIC FIELD

17. Leathem, J. G., Examples in the mathematical theory of electricity and

18. Batygin, V. V., and Toptygin, I. N., Problems in electrodynamics. New York:
Academic Press, 1964 (Chapter 2).

pp. 611–621).

PROBLEMS

1.1 Find the electric flux density everywhere within and without a uniform spheri-
cal charge distribution of radius \( r_0 \). Plot the result for radial distance up to \( 5r_0 \).

1.2 Two equal and opposite point charges, \( +q \) and \( -q \), are placed a distance
\( r_0 \) apart. Find the potential distribution caused by them. Show that at a
great distance, the potential is dependent on the product \( qr_0 \) but not on \( q 
\) or \( r_0 \) separately.

1.3 Calculate the electric field at a point in space near a long fine wire carrying
a charge of \( q \) coulombs per meter.

1.4 Prove that the field inside an empty sealed tin can is always zero.

1.5 Prove that all conducting material surfaces must be equipotentials in any
electrostatics problem.

1.6 An infinite flat sheet carries a uniform surface charge. Find the electric field
and potential everywhere in its neighborhood. If a second similar but oppo-
sitely charged sheet is placed a distance \( a \) away, again find the field and
potential.

1.7 Write Laplace's equation in spherical coordinates and solve to find the poten-
tial and field outside a uniformly charged spherical shell of radius \( r_0 \).
Compare this solution to the field of a point charge.

1.8 Find the field energy associated with a uniform charge distribution on a
thin spherical shell of radius \( r_0 \).

1.9 Find the charge distribution on a large flat conductive plate that results if
a thin wire ring carrying a charge \( q \) is held in a plane parallel to and a dis-
tance \( h \) removed from the plate. Let the ring have a radius \( R \) and the wire
diameter be \( 2a \). Evaluate the result numerically for \( R = h \).

1.10 Derive an integral expression for the field and potential distributions due
to a volume charge distribution with spherical symmetry.

1.11 Prove that the energy associated with a given charge distribution,

\[
W = \frac{1}{2} \int V \rho \, dU
\]

is not altered if a constant amount \( V_0 \) is added to the potential \( V \) everywhere,
i.e., if the potential reference level is altered.
1.12 If a certain charge distribution causes a potential distribution \( V_1 \) to exist, some other charge distribution causes another potential field \( V_2, \ldots \). Show that the potential distribution

\[
V = V_1 + V_2 + V_3 + \cdots
\]

results if all the charge distributions exist concurrently.

1.13 A point charge of \( \frac{1}{2} q \) coulombs is placed at each vertex of an equilateral triangle. Find the electric flux density vector at each point of a line normal to the triangle and passing through its center. Do the same for a regular hexagon, a regular duodecagon, and finally a circle, in all cases keeping the total charge at \( q \) coulombs.

1.14 Two dielectric media have different relative permittivities. A flux line intersects the boundary between these media, making an angle \( \theta_1 \) with the surface. If there is no surface charge at the interface, find the corresponding angle in the second medium.

1.15 A conducting sphere carrying a charge \( q \) is located at the origin of a Cartesian coordinate system. If the half space defined by positive values of \( x \) is filled with one dielectric medium and the remaining half space with another, find the charge distribution on the sphere.

1.16 It is a well-known fact that an electrostatically charged body (e.g., a glass rod) will attract uncharged bits of paper, dust, etc.; yet Coulomb's law predicts zero force between a charged and an uncharged object. Explain.

1.17 Show that charges free to move will always assume such positions as to minimize the amount of electrostatically stored energy. (Hint: This is a corollary of Earnshaw's theorem and may be proved by investigating the behavior of an added incremental charge \( dq \) in the field of other already established charges.)

1.18 Find the potential distribution inside the space bounded by two similar charged fine wire rings lying on a common axis in parallel planes separated by one ring radius. Plot the equipotential lines. (Such rings are frequently found surrounding hardware parts in high-voltage apparatus.)
The principles of electrostatic fields as developed in Chapter 1 are now applied to a number of engineering problems. There is probably no field of electrical engineering that does not employ dielectric or electrostatic devices of one kind or another; some of these, like the paper dielectric capacitor, are almost universally known, while others are restricted to specialized areas. An attempt is made here to seek devices illustrative of the variety of analytical methods while at the same time choosing problems of fairly general interest.

It may be noted that all these problems are treated on the assumption of static charge distributions, despite the fact that several of the devices dealt with are actually employed in alternating-current engineering. The restrictions, subject to which the electrostatic theory is applicable to time-variant problems, are examined in Chapters 7 and 8 where time-varying fields are considered.
1. Fields and Potential in a Coaxial Cable

Perhaps the simplest real problem in electrostatics is posed by the ordinary two-conductor coaxial cable. Such cables are widely used in both communications and power engineering and have been investigated in considerable detail. As shown in Fig. 2.1, a coaxial cable consists of two concentrically arranged cylindrical conductors separated by a layer of insulating material.

The insulation may be gaseous, liquid, or solid, depending on the size and application of the cable. To calculate the electrical characteristics of such a cable, a knowledge of the electric field distribution in the dielectric material is of prime importance. This problem is best set up in cylindrical coordinates. Between the conductors, there is no charge so that Laplace's equation applies; in cylindrical coordinates it reads

\[
\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial V}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 V}{\partial \theta^2} + \frac{\partial^2 V}{\partial z^2} = 0
\]  

(2-1)

From symmetry, it may be argued that there cannot be any variation in potential with angle \( \theta \); furthermore, if the cable is assumed to be very long and straight, there cannot be any dependence on \( z \). The derivatives with respect to \( \theta \) and \( z \) may therefore be eliminated so that Equation (2-1) becomes

\[
\frac{1}{r} \frac{d}{dr} \left( r \frac{dV}{dr} \right) = 0
\]  

(2-2)

Ordinary rather than partial derivatives appear because \( V = V(r) \) is a function of radius \( r \) only. Since (2-2) is a differential equation of the second order, two boundary conditions are needed to determine the two arbitrary constants that will arise on solution. Such conditions might typically be
potential values at the inner and outer conductors, both of which are equipotential surfaces. Let the potential values be \( V_0 \) on the inner conductor and zero at the outer. The solution of (2-2) is, in general,

\[
V = K_1 \log (K_2 r)
\]

which becomes, on substitution of the boundary values,

\[
V = V_0 \frac{\log (r/r_0)}{\log (a/r_0)}
\]

The potential is not uniformly distributed between the two conductors but, according to Equation (2-4), varies logarithmically with radius. Thus, the variation of potential from point to point is much more gradual near the outer conductor than the inner. Because insulation flashover is dependent on the electric field, not the potential values, it is often of considerable interest to know the field distribution in the interconductor space. The field intensity \( E \) may be found by taking the gradient of potential, with the result

\[
E = -\frac{V_0}{r \log (a/r_0)} \mathbf{1}_r
\]

Electric flux lines have the same direction as the vectors \( \mathbf{D} \) and \( \mathbf{E} \) and must therefore be radially directed. Conversely, equipotential surfaces are given by \( r = \) constant, for the potential does not depend on the other two coordinates. Some of these lines and surfaces are indicated in Fig. 2.2. It should be noted that since the potential varies logarithmically with radius, surfaces equally spaced in potential are not equidistant in space.

On many occasions a knowledge of the charges on the conductors is desirable. It is possible to solve for the charge per unit length on the inner conductor by direct application of Gauss’s law, Equation (1-17), to a cylindrical surface of integration concentric with the conductors and located at some radius \( r \) between them. Making use of (1-33), Gauss’s law may be stated as

\[
\frac{1}{\varepsilon} q = \oint \mathbf{E} \cdot \mathbf{dS}
\]

and applied to this particular problem by introducing \( E \) as given by Equation (2-5). Integrating over the cylindrical surface,

\[
q = \int \frac{-eV_0}{r \log (a/r_0)} r \, d\theta
\]

per unit length. It is sufficient to integrate over the curved surface, for the
ends of a right circular cylindrical surface always have a pure axial surface vector \( \mathbf{dS} \). The dot product of this vector with a radially directed field vector is zero everywhere. Thus, the charge per unit length is

\[
q = -\frac{2\pi e V_0}{\log (a/r_0)}
\]  

(2-8)

The charge on the outer conductor must be \( q \) also, by the law of conservation of charge.

It is frequently convenient to rewrite Equation (2-4) to express the potential at any point in terms of charge rather than reference potential. Substituting (2-8) in (2-4), such an expression is readily found:

\[
V = -\frac{q}{2\pi e} \log \frac{r}{r_0}
\]  

(2-9)

2. Parallel-Wire Lines

Any potential problem leading to Laplace’s equation is very easy to solve if a coordinate system can be found in which the known equipotential surfaces of the problem coincide with coordinate planes. With the usual three coordinate systems, the situations amenable to simple solution are thus (1) infinite parallel planes, (2) coaxial cylinders, and (3) concentric spheres. These are of course solved in Cartesian, cylindrical, and spherical coordinates, respectively. The majority of problems of real interest, however, does not fall in any one of these three categories, so other methods must be found.

One particularly interesting case is that of a pair of long, fine parallel wires. This problem, like many others, can be solved by a superposition technique. The technique is based on the following argument. Suppose there exists some charge distribution \( \rho_1 \), whose attendant potential \( V_1 \) and field \( \mathbf{E}_1 \) are known everywhere in space, and some other charge distribution \( \rho_2 \) whose corresponding potential \( V_2 \) and field \( \mathbf{E}_2 \) are known also. If both charge distributions are made to exist at the same time, the fields are known to be additive so that the total field is

\[
\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2
\]  

(2-10)

But the potential is derived from the field intensity by line integration; and since the integral of a sum equals the sum of the integrals, the potentials superpose also:

\[
V = V_1 + V_2
\]  

(2-11)

This argument may naturally be extended by mathematical induction to any number of separate charge distributions.

In the case at hand, solution for the potential inside a coaxial cable may be made to serve as a starting point. There is in general no reason why
the reference level $V = 0$ should be chosen to be at the outer conductor of a coaxial cable; the expression for potential is unaltered (except for an additive constant) if some point a distance $R$ from the center of the inner conductor is chosen as the zero-potential reference. For the parallel wires shown in Fig. 2.3, the potential at any space point $P$ is then given by

$$V = -\frac{q_1}{2\pi\epsilon} \log \frac{r_1}{R_1} - \frac{q_2}{2\pi\epsilon} \log \frac{r_2}{R_2}$$  

(2-12)

where $R_1, R_2$ denote the distances from the right and left wires to the potential reference point where $V = 0$, and $q_1, q_2$ are the charges on the wires.

If the charges are equal, $q_1 = -q_2 = q$; then (2-12) may be written in the form

$$V = -\frac{q}{2\pi\epsilon} \log \frac{r_1}{r_2} - \frac{q}{2\pi\epsilon} \log \frac{R_2}{R_1}$$  

(2-13)

It should be noted that the second term in this expression is arbitrary since it may be made to equal any convenient number by choosing an appropriate reference point. In particular, if the reference point is chosen anywhere on the $y$-axis, $R_1 = R_2$ and the term vanishes. Then

$$V = -\frac{q}{2\pi\epsilon} \log \frac{r_1}{r_2}$$  

(2-14)

If the potentials of the wires are assumed to be $+V_0$ on the right and $-V_0$ on the left, this expression is readily altered to

$$V = \frac{V_0}{\log (a/b)} \log \left( \frac{r_1}{r_2} \right)$$  

(2-15)

To facilitate its physical interpretation, this solution will be plotted graphically. Equation (2-15) may be rewritten as

$$\frac{r_1}{r_2} = \left( \frac{a}{b} \right)^{V/V_0}$$  

(2-16)
whereupon the equation of an equipotential surface of potential $V_e$ emerges as

$$\frac{r_1}{r_2} = \left(\frac{a}{b}\right)^{V_e/V_e} = k$$

(2-17)

where $k$ may be viewed as a parameter expressing $V_e$. On writing out the distances $r_1$ and $r_2$ in terms of the coordinates $(x, y)$, there is obtained

$$\frac{r_1^2}{r_2^2} = k^2 = \frac{x - (b/2)|^2 + y^2}{x + (b/2)|^2 + y^2}$$

(2-18)

Clearing fractions and completing the square, (2-18) assumes the form

$$\left[ x - \left(\frac{1 + k^2}{1 - k^2}\right) \frac{b}{2} \right]^2 + \frac{y^2}{\left(\frac{2k}{1 - k^2}\right) \frac{b}{2}} = \left(\frac{2k}{1 - k^2}\right) b$$

(2-19)

Equation (2-19) may be recognized as representing a family of circles with centers on the $x$-axis and symmetrically arranged about the $y$-axis. Some members of this family appear in Fig. 2.4.

A complete solution of Laplace’s equation must consist of both the equipotential surfaces and flux lines. Here the equipotentials are cylinders corresponding to the circles in the $x$-$y$ plane; the flux lines must, as always, be given by lines orthogonal to them. All flux lines must begin and end at
charges and must therefore form curves meeting at the points representing the two wires. Elsewhere these curves must not cross or touch, and they must be symmetrically distributed about the \( y \)-axis. All these conditions are satisfied by a family of circles centered on the \( y \)-axis and so located that the angle at which the lines \( r_1 \) and \( r_2 \) meet is constant along each circle. The equations of these circles can be found from differential geometry: the slopes of equipotential lines may be found by differentiation; then the differential equation of a set of curves normal to them constructed and solved. The calculation involved is lengthy and will not be carried out here because a simpler method, based on the theory of functions of a complex variable, will be introduced later to deal with the problem. Some of the resulting circles are shown in Fig. 2.4 also.

It may be seen that the equipotential lines are crowded close together in the region between wires and are relatively far apart elsewhere. The highest field strengths are to be encountered at the wire surfaces nearest each other; unlike the cable problem, this situation does not lead to equal field strength along each equipotential line. The actual field strength may be calculated from the expression for potentials:

\[
E = \frac{-V_0}{\log(a/b)} \left[ j \left( \frac{x + (b/2)}{r_2^2} - \frac{x - (b/2)}{r_1^2} \right) + j \left( \frac{y}{r_2^2} - \frac{y}{r_1^2} \right) \right]
\] (2-20)

As originally set up, the problem is to seek the potential distribution in the neighborhood of a pair of very fine wires. However, it now appears that the restriction to fine wires is unnecessary. No matter what the shape and size of the conductors happens to be, their sole effect on the solution of Laplace's equation in the interconductor space (and this is an important effect!) is that of requiring the conductors themselves to be members of the family of equipotential surfaces. Any other conductors may be substituted without altering the field, provided that their shapes are such as to make them members of the same family. It is evident from Fig. 2.4 that many cylindrical shapes qualify; the solution at hand could easily be modified to cover any of the configurations of Fig. 2.5.

![Fig. 2.5](image_url)
It must be carefully noted that the distance $b/2$ in all cases refers to the separation between the electrical center of a conductor and the origin of the coordinate system, not the mechanical center. The distinction should be evident from Fig. 2.4.

3. Conductors at High Voltages

Dielectric materials respond to applied electric fields by polarizing, as discussed in Chapter 1. In very high fields, an increase in polarization becomes impossible, for the molecular structure of matter places a limit on the relative shift between positive and negative charges in the material. Application of fields in excess of a critical value determined by the nature of the dielectric material results in ionization of the material. Whether the ions eventually recombine into the same or different substances is dependent on the material and its environment. Very complex chains of chemical reactions often result in practice. For example, oxygen at STP ionizes at field strengths of the order of 3–4 megavolts per meter, yielding nascent oxygen. This is an endothermic process and requires energy from the electric fields. Nascent oxygen is chemically highly active and will combine with conductor material, other gases, and various dielectric materials to form oxides; most of it, however, recombines into molecular oxygen and ozone. In turn, the ozone decomposes, forming more oxides and molecular oxygen. The recombination processes are exothermic and result in the production of heat, light, and sound. The light emitted is bluish or reddish, depending on the polarity of the ions, and appears as a luminous halo enveloping the conductor; hence the descriptive name corona is given to the entire process.

Substantial quantities of electrical energy may be required for such chemical processes. For a conductor about 1 cm in diameter, for example, corona loss amounts to about 1 watt per meter at the first appearance of visible corona. As attempts are made to increase the field strength, losses increase rapidly. With a voltage about 25 percent greater than the critical (onset of corona) value applied to an open-wire line, losses rise to about 10 watts per meter for a conductor of 1 cm diameter; at double the critical voltage, about 100 watts per meter are dissipated. The seriousness of these figures may be underlined by converting units: 1 watt per meter equals $\frac{1}{4}$ megawatt per 100 miles of three-phase line.

The requirement of near-absolute freedom from corona, even under the worst conditions of dirt accumulation and icing, places a limitation on the minimum size of high-voltage transmission apparatus. In the case of open-wire or coaxial transmission lines, this limitation is expressed by Equations (2-5) and (2-20) which indicate that the maximum field in the neighborhood of conductors is directly proportional to the line voltage and inversely proportional to the logarithm of the diameter-to-spacing.
ratio, as well as the conductor diameter. In other words, either the conductor diameter must grow proportionally with voltage or the spacing ratio exponentially in order to avoid exceeding a given maximum field strength. The latter choice quickly leads to extreme structure sizes, while the former requires large conductors. The possibility of using hollow conductors of large cross section readily suggests itself, thereby eliminating excess weight while still retaining a large radius. Unfortunately, the weight of ice formed on a conductor varies very roughly as the diameter and the wind pressure varies nearly as the square of the diameter; consequently, a line made of large hollow conductors is still far from ideal. However, both reduction of corona and improvement in terminal characteristics make large conductors extremely attractive. As an acceptable approximation to a hollow conductor, so-called bundle conductors are used, consisting of several strands of wire arranged to resemble approximately the surface of a cylinder (see Fig. 2.6).

It is pictorially evident that the larger the number of conductors in a bundle, the more nearly the bundle will resemble a large conductor. Nevertheless, even so crude an approximation as a two-conductor bundle provides a substantial reduction in field intensity as compared to a single wire.

The field distribution in the neighborhood of bundle conductors is calculable by a technique similar to the solid-wire line, i.e., by superposing solutions for individual wires. However, having already obtained a solution for the two-wire line, it is easier to view the two-conductor bundle line as the superposition of two parallel-wire lines, displaced relative to each other some distance \( c \), as in Fig. 2.7.

The potential \( V \) at any point \( P \) in space will now be calculated. For a
pair of wires, it is of the form

\[ V = K \log \frac{r_1}{r_2} \]  

(2-21)

For the two-conductor bundles it must therefore be of the form

\[ V = K \log \frac{r_1}{r_2} + K \log \frac{s_1}{s_2} = K \log \frac{r_1}{r_2} \frac{s_1}{s_2} \]  

(2-22)

The constant \( K \) must be found from the known boundary potentials. On the left member of the right-hand bundle, \( V = V_0; r_1 = c \) very nearly; \( s_1 = a; r_2 = s_2 = b \) very nearly (it is assumed that \( b \gg c \gg a \)). Substituting in (2-22), there obtains

\[ K = \frac{V_0}{2 \log (\sqrt{ac}/b)} \]  

(2-23)

so that the potential distribution of the two-wire bundle line is given by

\[ V = \frac{V_0}{\log (\sqrt{ac}/b)} \log \frac{r_1}{s_1} \frac{s_1}{s_2} \]  

(2-24)

At distances considerably greater than the bundle spacing \( c, r_1 = s_1, r_2 = s_2 \), and the potential distribution becomes

\[ V = \frac{V_0}{\log (\sqrt{ac}/b)} \log \frac{r_1}{r_2} \]  

(2-25)

that is to say, the twin bundle looks like a single conductor of radius \( \sqrt{ac} \). On the other hand, \( r_2 = s_2 = b \) is an acceptable simplification at distances of the order of bundle spacing. The potential distribution then becomes

\[ V = \frac{V_0}{\log (\sqrt{ac}/b)} \log \frac{r_1}{b} \]  

(2-26)

and the corresponding set of equipotential lines is as shown in Fig. 2.8. These equipotential lines are curves of a form given by \( \log x_1x_2 = \) constant and are called Cassini curves or simply Cassinians.

To explore the corona problem, it may first be noted from Fig. 2.8 that maximum field strength is likely to occur at the outer edges (extremes in the \( x \)-direction) of either bundle. Very near the right wire of this pair, \( s_1 = c \) nearly enough, and the potential distribution is

\[ V = \frac{V_0}{\log (\sqrt{ac}/b)} \log \frac{c}{b} + \frac{V_0}{\log (\sqrt{ac}/b)} \log \frac{s_1}{b} \]  

(2-27)

The field intensity reduction near the wires, as compared to a single conductor, is of the same order of magnitude as the change in the far field. It should be kept in mind, of course, that a fair comparison is not that with a similar single wire but with a single wire of as great weight as the bundle conductor. On this basis, the improvement is not quite so great. Nevertheless, a substantial advantage is evident even for a two-conductor bundle. For present-day high-voltage lines, two-conductor and four-conductor bundles
equivalent to single conductors 1 to 2 ft in diameter are not uncommon; three-conductor bundles are also used. Figure 2.9 illustrates a typical method of actually building such lines.
4. The Method of Images

In addition to the simple two-conductor bundle transmission line of the foregoing section, many other problems may be treated by superposing thin-wire potentials. A different method of extending known solutions was also alluded to in connection with the two-wire line: if some of the equipotential surfaces of a known solution can be found to fit the given equipotentials of a new problem, then the two solutions are mathematically similar and one may be derived from the other. Use may be made of this principle to develop the potential distribution of an eccentric cable from that of a two-wire line, for example, as indicated in Fig. 2.5. A particularly interesting case arises when one or more of the equipotential surfaces in a solution are planes, for they must then be planes of symmetry. A few examples will illustrate the principle.

Let the potential distribution of a single-wire power line with ground return (such lines are widely used in rural power distribution systems around 5–15 kilovolts) be sought. The two given equipotential surfaces are the wire itself, assumed of radius \( a \), and the ground plane, as indicated in Fig. 2.10. In actual fact, the electrical ground plane lies at some depth (a few inches to a few feet) below visible ground; for purposes of the problem, however, the two are assumed to coincide. On examining the equipotential and flux lines of the two-wire line (Fig. 2.4), it is seen that this problem has exactly the same solution if the ground plane is taken to be the halfway equipotential line between the two wires. This situation is often described as replacement of the ground plane by the image of the line. All that remains is to rewrite the potential solution of the two-wire line

\[
V = \frac{V_0}{\log (a/b)} \log \frac{r_1}{r_2}
\]  

(2-15)

in terms of symbols more useful for a single wire above ground:

\[
V = \frac{V_0}{2 \log (a/b)} \log \frac{x^2 + (y - h)^2}{x^2 + (y + h)^2}
\]  

(2-28)

To find the fields around a two-wire line above a ground plane (a very realistic transmission line problem), the line may be thought to be reflected in the plane. Two of the above solutions must then be superposed, with the result
\[ V = \frac{V_0}{2 \log(a/b)} \log \left\{ \frac{[x - (b/2)]^2 + (y - h)^2}{[x + (b/2)]^2 + (y - h)^2} \cdot \frac{[x + (b/2)]^2 + (y + h)^2}{[x - (b/2)]^2 + (y + h)^2} \right\} \]

(2-29)

It is worth noting that the problem of a wire lying near the corner of a duct is solved by the same equation (except perhaps for rewriting in more suitable coordinates) since the three configurations shown in Fig. 2.11 are fully equivalent.

The placement of equipotential surfaces or their equivalent line charges may be argued equally well on mathematical or physical grounds. It is visually simplest to imagine the plane equipotential surfaces as mirrors (for mirrors are optical planes of symmetry) and to try to see where the images of the real charges lie. Because of its similarity to optical image problems, this method of extending solutions is called the method of images. It will be found equally useful in problems involving magnetic fields.

5. Relaxation Solutions

Except for geometrically relatively simple boundary shapes, analytic solution methods often lead to hopeless mathematical complexity. Two-dimensional (and many three-dimensional) problems in Laplace’s and Poisson’s equations may then be solved by numerical or graphical methods. Although there is in principle no objection to using these methods for general three-dimensional problems, the results are necessarily numerical or graphical in form. Considerable difficulty is encountered in their interpretation if they cannot be represented on a single sheet of paper. For solving Laplace’s equation, relaxation methods are particularly suitable. Such solutions are started by guessing at a potential distribution (i.e., at the location of equipotential lines) and then calculating a number of corresponding flux lines to satisfy the requirement of orthogonality. Because the potential distribution is in
general wrong, some of the flux lines will cross or end at points other than charge centers. The flux lines are calculated so as to result in the best possible fit to the wrong equipotentials, using some reasonable criterion for fit. Next, the flux lines are held fixed and a set of equipotentials calculated to give as good a fit as possible. This process of adjusting potential and flux distributions in turn is continued until the errors have decreased to an acceptable level. Relaxation techniques work on the entire field at the same time rather than extrapolating known correct solutions into new areas. Compared to some other numerical methods, their advantage is an absolutely assured convergence, for any error will be ironed out in the process. Their main disadvantage is that the entire solution is produced at the same time, and single bits of information (e.g., field intensity at a particular place) are not available alone, even though very often a few point values may be all that is really desired.

It is possible to obtain relaxation solutions either numerically or by direct graphical procedures.

Working graphically, the entire space considered may be filled with flux and equipotential lines. A qualitative picture of the field is thus obtained; numerical information may be introduced readily (and extracted readily) as shown in the following. Because flux and equipotential lines always cross at right angles, they form more or less rectangular figures. A typical rectangle is shown in Fig. 2.12. The development of its bounding lines into the paper marks out a boxlike space. All the flux crossing through the box enters through the top \( V_1 \) and leaves through the bottom \( V_2 \) equipotential surface; none can cross through the sides, for the sides are defined by flux lines, which may not intersect. At the middle of the box, let the distance between box sides be \( w \). The flux density in the box, assuming unit depth, is then approximately

\[
D = \frac{\Psi}{w}
\]

(2-30)
where $\psi$ is the flux. The accuracy of this approximation may be judged from the variation of box width, for $\psi$ is the same through both top and bottom.

Similarly, the magnitude $E$ may be calculated. In the box, $E = -\text{grad} \ V$ gives, approximately,

$$E = \frac{V_1 - V_2}{h} = \frac{\Delta V}{h}$$  \hspace{1cm} (2-31)

where $h$ is the height of the box in the middle. Now $D = \varepsilon E$; Equations (2-30) and (2-31) may be combined to give

$$\frac{\psi}{w} = \varepsilon \frac{\Delta V}{h}$$

or rewriting,

$$\frac{\psi}{\Delta V} = \varepsilon \frac{w}{h}$$  \hspace{1cm} (2-32)

It should be noted that if the curvilinear figure is a square, so that $w = h$, then

$$\frac{\psi}{\Delta V} = \varepsilon$$  \hspace{1cm} (2-33)

a constant clearly independent of the size of the square.

Figure 2.13 shows a field map drawn so as to include nothing but squares. The flux through any two adjacent squares, say $A$ and $B$, is the same. Since $A$ and $B$ are squares, the ratio of flux to voltage across them must be the same; hence $(V_3 - V_2) = (V_2 - V_1)$. The equipotentials are therefore equally spaced in potential. Similarly, squares $B$ and $C$ are between the same pair of equipotentials. Because the ratio of flux to potential difference is constant for square figures, it may be concluded that the fluxes through the box behind square $B$ and the box behind $C$ must be equal.

To summarize: If surfaces and lines satisfying Laplace's equation are drawn so as to form squares, the resulting lines in the plot mark out equal increments of flux and potential. A complete numerical solution of the problem is then obtainable if the potentials of any two equipotential lines, or the charge at the end of any flux path, is known.

Graphical relaxation is extensively used to plan the layout for numerical solutions. Surprisingly little practice suffices to produce a map accurate to about 5 percent, an accuracy level often entirely adequate for spotting critical regions of a field.

Much useful information on applying the graphical relaxation technique appears in the literature and should be consulted by anyone interested in acquiring familiarity with the method. For the beginner, the following ground rules are offered as a starting point.

**HOW TO RELAX**

1. Necessary equipment consists of stout drawing paper, a soft eraser, and a sharp 2H pencil. Softer pencils smudge on repeated erasure.
2. The human eye is an excellent judge of right angles but is confused by clusters of lines. Squares smaller than about an inch should be avoided.

3. A very few lines (say three or four equipotentials at most) should be established firmly before finer subdivision is attempted.

4. Scratching with short, repeated strokes hides inaccurate work and ultimately must be erased in any case. Lines should always be drawn with single, clear pencil strokes, even if they are wrong.

5. Relaxation is a process of successive improvements on a wrong solution. The temptation to omit drawing a line of uncertain accuracy must be resisted firmly; wrong guesses are a necessary part of the method!

6. Numerical Relaxation

The basic numerical technique for solving field problems, on which most others are based, is merely a numerical adaptation of the graphical relaxation scheme described above. In its original form, it is well suited to hand calculation, for the progress of the whole solution is clearly visible to the computer at all times. This overall view is of considerable importance, for it enables the calculator to eliminate most of the routine arithmetical tedium. At its core, any relaxation method consists of an initial wrong guess followed by some iterative process in which the errors are smoothed out to the vanishing point. Once the general trend of a solution is seen, the human operator can either introduce arithmetical shortcuts or else throw out entire sections of the solution, replacing them with what he thinks is a better guess.

Like most numerical methods, the relaxation technique calculates values of potential at selected points, in contrast to the graphical technique that seeks points corresponding to a given potential. For numerical work, the Poisson differential equation

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = -\frac{\rho}{\epsilon} \quad (2-34)$$

must first be replaced by a finite-difference approximation. To do so, consider the nine points of the x-y plane shown in Fig. 2.14. The derivative of $V$ with respect to $x$ at point $A$ is approximately

$$\left. \frac{\partial V}{\partial x} \right|_A = \frac{V_0 - V_w}{h} \quad (2-35)$$

and at $B$,

$$\left. \frac{\partial V}{\partial x} \right|_B = \frac{V_r - V_0}{h} \quad (2-36)$$

The second derivative at the center point $O$, in turn, is approximately
\[
\frac{\partial^2 V}{\partial x^2} = \frac{1}{h} \left( \frac{\partial V}{\partial x} \bigg|_A - \frac{\partial V}{\partial x} \bigg|_B \right) = \frac{V_W + V_E - 2V_o}{h^2}
\]  

(2-37)

The second derivative with respect to \( y \) may be approximated similarly. The Poisson equation may be approximated by first forming

\[
V_W + V_N + V_E + V_S - 4V_o + \frac{h^2\rho}{\varepsilon} = R_o
\]

(2-38)

not involving the points \( A, B, C, D \) at all. The right-hand side \( R_o \) is called the residual at the point \( O \). If the residual is identically zero, Equation (2-38) represents a finite-difference approximation to Poisson’s equation.

Solution of Laplace’s equation is now possible. The space between boundaries where a solution is desired is ruled off in squares to establish points at which the potential is to be calculated, and a guessed value of potential is assigned to each point. The residuals at all points are then calculated; they clearly correspond to the amounts of charge that must be inserted at these points to yield the guessed potentials. It is evident that Laplace’s equation is only satisfied when all the residuals are zero. The usual method of calculation is to change the potential at the point with the largest residual so as to reduce this residual to a small value. The residuals at the adjacent four points are thereby changed and must in turn be recalculated. Residuals elsewhere, however, remain unaffected. This process of reducing the largest residual is repeated until the residuals are all quite small. As a rough rule, the potential values will contain errors at least one order of magnitude greater than the remaining residuals. Eventually, however, the finite-difference approximation becomes the limiting factor. Greater accuracy is then obtained not by even further reduction of the residuals, but by using this result as the initial guess for restarting the problem with a finer mesh of points. It may be seen that this process is in some ways like a game of chess; a skilled player can frequently see ahead several moves and eliminate large amounts of calculation by judiciously altering several residuals at the same move. If programmed for a digital computer, it is more difficult to arrange such simultaneous alterations of numerous residuals, although changing all the residuals in a row of points (line relaxation) or a field of points (block relaxation) is possible. A good initial guess will also eliminate much work by starting the problem with low initial residuals.

A simple example will serve to illustrate the method. Let it be required to determine the potential distribution in a square duct containing a centrally located round wire, as in Fig. 2.15a. From symmetry, it is evident that only one-eighth of the problem need be worked out. To make the problem as simple as possible, only one grid line distinct from the boundaries will be used for the first solution, so that only two points not on the boundaries need be considered. The potential of the inner wire is taken as +100, that of the duct as zero.
Fig. 2.15
As a first guess, the potential at all nonboundary points will be taken as 50. These values are written below the mesh points; the residuals are calculated and written above the potentials. The large residual (—100) is next reduced to zero by increasing the potential at that point by one-fourth the residual. A residual of —50 is thereby created at the adjacent points (second stage). This residual is in turn reduced by reducing the potential. At the fifth stage, the residuals have become —6 and —2; no benefit will accrue from further refinement of residual values on so coarse a mesh. Rather, the mesh should now be subdivided so as to create more points. Initial guesses for this next set of calculations are shown in Fig. 2.15b. The potentials have all been multiplied by 10 for the more refined stage in calculation in order to avoid working with fractional numbers. Should a plot of the equipotential and flux line be required, it may be constructed freehand by interpolation between tabulated points. Naturally, the finer the mesh the better the interpolation.

To illustrate the application of digital computers to this technique, the slightly more complicated problem of a flat conductor at a potential of +5, centered in a square duct at zero potential, will be considered. This problem and its solution on a fairly coarse mesh are shown in Fig. 2.16.

In hand relaxation, the largest residuals are normally reduced in turn. This is not usually a practical procedure in digital computing since the search to find the largest residual requires, on the average, more time than the recalculation of potential and residual at the point. It is usual to avoid wasting time, and to simplify programming, by simply relaxing the potentials at all field points in turn until the largest residual has dropped to some acceptable level. The computer, it may be seen, will then perform a very much larger number of calculations than a human calculator would in order to reach the same maximum residual. Similarly, to enter guessed potential values individually is usually troublesome; unless a rough functional approximation to the initial guess can be stated, it is best to start with some purely arbitrary value of potentials, for example setting all field points at half the maximum potential value. Again the computing machine will be required to perform many more calculations than its human competitor.

A programme flow chart to solve the problem of Fig. 2.16 is shown in Fig. 2.17. The programme begins with the usual "bookkeeping" operations and then establishes initial values for all potentials. The actual relaxation calculation proceeds in an alternating cycle. Residuals are calculated for all field points and checked against the maximum permissible value. If the solution has not converged, the potential values at all points not on the boundaries are modified so as to reduce each residual to zero in turn. Naturally, the residuals do not remain zero as the calculation proceeds; the programme therefore returns to perform another iteration in the sequence.

An actual programme that solves the problem according to this flow chart
RELAXATION PROBLEM -- FLAT STRIP IN A SQUARE DUCT

Fig. 2.16
appears in Appendix II. Examination of both flow chart and programme will show that many improvements might be made (e.g., recalculation of potentials at points with very small residuals might be omitted) to save computing. As so often happens, improvements to save machine time require expenditure of programming time, and vice versa, and economic considerations will generally dictate the extent of refinement desirable in the programme.

It should again be emphasized that the physical significance of a residual

---

**Fig. 2.17**

---

Reserve a 14 x 14 field for the potential values, and a similar field for residuals. Set potentials at boundary points to given values, all others to a value halfway between.

Set the indicator INDIC to zero

For the next (or first) off-boundary point, calculate the residual value.

Is the residual value greater than maximum allowed?

Yes

Set INDIC to unity

No

Last point in field?

Yes

Add one-quarter the residual value to each of the potential values associated with non-boundary points of the field.

No

Is INDIC zero?

Yes

Print out the resulting answers and the number of iterations required.

---
is that of a charge placed at the appropriate mesh point; the residuals corresponding to any given potential distribution merely describe the charge distribution required to establish this potential distribution. The relaxation method may thus be thought of as a charge-elimination process, in which a set of potential values is altered so as to move extraneous charges out of the field and only leave charges at the boundaries. This interpretation indicates also that residuals at equipotential boundary points must not be reduced to zero. Removal of charges from the entire field will of necessity yield a zero answer for all potentials everywhere!

7. Liebmann's Iteration Method

The relaxation process involves a great deal of computation and is not very fast unless modified constantly to avoid pointless calculations. Such modifications are easy to make in hand calculation but not on automatic computers. For that reason, it is often used where hand calculation is to be employed but not usually programmed for machine use. The most prevalent machine method, first proposed by Liebmann, is an ingenious extension of the relaxation procedure; it generally converges faster and requires substantially less memory space in the computer.

In this method, the entire array of residuals is not calculated, but only one residual at a time. When it is found, the potential at the corresponding point is readjusted immediately. Intuitively, this process should result in faster convergence than relaxation, for each residual to be calculated is not based on a set of potentials obtained on the last cycle, but partly on old, partly on new potentials. It can be shown, in fact (and this does require proof!), that the process is generally convergent; and practice shows the convergence rate to be significantly improved over relaxation.

Throughout the relaxation calculation, potentials are augmented to depress residuals to zero. As successive point potentials are altered, however, the residuals at other points are affected and no longer remain zero. This knowledge leads to another worthwhile improvement of the process. If each point potential is overrelaxed, i.e., altered by too much, then the corresponding residual is depressed beyond zero; but if the amount of overrelaxation is just right, the rise in residual that occurs on recalculation of the neighboring potentials will bring the residual back to zero. The result ought to be a very much improved solution convergence rate. In fact, very substantial acceleration is obtainable by this method. The right amount of overrelaxation depends on the nature of the problem and cannot be stated in general.

To summarize, the Liebmann method consists of scanning the field of points in succession, continuously replacing the potential \( V_{jk} \) by an adjusted value:
\[ V_{jk(\text{new})} = V_{jk(\text{old})} + \frac{\alpha}{4} R_{jk} \]  

(2-39)

where \( \alpha \) is a number called the *overrelaxation factor* and \( R_{jk} \) represents the residual calculated from the potentials currently existing in the array of potential values. It can be shown that the method is always convergent for \( \alpha = 1 \) (i.e., no overrelaxation) and always divergent for \( \alpha = 2 \). Best convergence is obtained for some value between these two limits.

A programme flow chart is given in Fig. 2.18 to illustrate the Liebmann method as applied to the foregoing problem. There is no provision made for an array of residuals since none will be needed. In order to establish whether the solution has converged, it is only necessary to detect the presence of an excessively large residual. This is accomplished by comparing each residual to the maximum permissible value and setting a numerical indicator, initially zero, to unity if the residual is large. After the entire field of points has been scanned, the indicator is examined to see whether a large residual was encountered.

Both the flow chart and the programme listing shown in Appendix II should make it amply clear that the Liebmann iterative method leads to even simpler programming than relaxation in its more primitive form.

Of interest is the variation in solution convergence rate obtained for various values of overrelaxation factor. The solution found by the Liebmann method does of course not differ from Fig. 2.16, but there is a significant variation in the amount of computation required. The accompanying table lists the number of iterations actually needed by the programmes shown in order to depress the largest residual to 0.002.

<table>
<thead>
<tr>
<th>Overrelaxation factor</th>
<th>Number of iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>None (simple relaxation)</td>
<td>155</td>
</tr>
<tr>
<td>1.0</td>
<td>67</td>
</tr>
<tr>
<td>1.1</td>
<td>56</td>
</tr>
<tr>
<td>1.2</td>
<td>46</td>
</tr>
<tr>
<td>1.3</td>
<td>37</td>
</tr>
<tr>
<td>1.4</td>
<td>29</td>
</tr>
<tr>
<td>1.5</td>
<td>22</td>
</tr>
<tr>
<td>1.6</td>
<td>25</td>
</tr>
<tr>
<td>1.7</td>
<td>34</td>
</tr>
<tr>
<td>1.8</td>
<td>47</td>
</tr>
<tr>
<td>1.9</td>
<td>89</td>
</tr>
<tr>
<td>2.0</td>
<td>Divergent</td>
</tr>
</tbody>
</table>

The optimum overrelaxation factor is readily seen to produce very considerable acceleration of solution. Even more spectacular acceleration
Reserve a 14 X 14 field for the potential values only. Set potentials at boundaries to prescribed values, and all others halfway between.

Set indicator INDIC to zero

For the next (or first) mesh point, calculate residual value.

Is the residual value greater than maximum allowed?

Yes

Set INDIC to unity

No

Adjust the potential at this field point in accordance with the residual just calculated, and the prescribed overrelaxation factor.

Is INDIC zero?

Yes

Print out the resulting answers and the number of iterations required.

No

Last point in field?

Fig. 2.18

obtains in fields with many points; for infinitely many points, the optimum overrelaxation factor theoretically yields an improvement by $\sqrt{\infty}$.

It is worth observing that standard pica typewriter spacing, the normal spacing for computer printout, has 10 characters per inch laterally and 6 lines to the inch vertically. Consequently, triple-spaced printout in a format 5
spaces wide is automatically arranged in a half-inch grid. Much time in plotting and evaluating results can be saved if they are placed on the page so as form a geometrically correct grid pattern, on which equipotential lines and the like may be drawn directly.

8. A Monte Carlo Method

Computational techniques making use of a random process of one kind or another and exploiting its probabilistic properties are called Monte Carlo methods. One very simple technique for finding the potential at some point $P$ is based on properties of statistical distributions; it is unfortunately quite slowly convergent and therefore not suitable for use where a complete field map is desired.

Let the space between boundaries in Fig. 2.19 be ruled off in a Cartesian grid, as for the relaxation calculation. Choose at random either of the letters $x$ or $y$ and either of the signs $+$ or $-$, and move from the point $P$ in the direction thus identified (say $+x$) one square. Again choose at random one of the four possible directions and move one square. The path traced out by this sequence of random directions is called a “random walk” or a “drunkard’s walk.” Eventually this walk will reach one of the boundaries; when it does, it will be terminated and the potential $V_1$ at the boundary recorded. The random walk is repeated some number $n$ times, where the number $n$ in general needs to be fairly large. It may be shown that

$$\lim_{n \to \infty} \left( \frac{\sum V_n}{n} \right) = V_p$$

where $V_p$ is the potential at the point $P$. The rate of convergence is proportional to $\sqrt{n}$, so many random walks are necessary if the mesh chosen is fine and reasonable accuracy is required. Manual solution may be carried out easily, using a pair of dice as random number generators, if low accuracy
and coarse mesh size are acceptable. Otherwise, the procedure is readily programmed for a digital computer; but the reservations expressed in connection with relaxation methods are again applicable, especially as concerns rate of convergence. A very important advantage of the Monte Carlo technique, however, is that only the current point on the random walk needs to be stored in the computer memory at every step of the process. In contrast, iterative solutions require storing the whole array of potentials. This very great economy in memory capacity can make the Monte Carlo method attractive where the field problem constitutes part of a much larger programme, and computer storage space is at a premium.

At present, 1 hour of running time on a large digital computer costs about 2 weeks’ salary for a very competent engineer. This price is not likely to decrease greatly in the near future. Quite aside from any gain in knowledge or intellectual satisfaction, the immense economic value of even an approximate analytic solution should be self-evident! Equally, there is an obvious necessity to ensure that the best numerical method is chosen for any given problem.

9. Potential and Induction Coefficients

In Chapter 1, the Poisson equation in free space was solved and the field of a point charge found. This solution was extended to any other charge distribution by considering the general distribution \( \rho(x, y, z) \) to be made up of small elements \( \rho \, dU \), each of which may be treated as a point charge to a good approximation. There resulted, for an infinite homogeneous dielectric medium of permittivity \( \varepsilon \),

\[
V = \frac{1}{4\pi \varepsilon} \int_V \frac{\rho \, dU}{r}
\]  (1-50)

where \( r \) denotes the distance from the element \( dU \) to the point of measurement of \( V \). Of particular interest are distributions that consist of charges placed on discrete conductive bodies. Supposing that there are \( N \) such bodies in space, the potential of the \( k \)th body may be expressed as

\[
V_k = \frac{1}{4\pi \varepsilon} \sum_{j=1}^{N} \frac{\rho_j \, dU_j}{r_{kj}}
\]  (2-41)

where \( r_{kj} \) denotes the distance from any point on the \( k \)th body (all points of that body are at the same potential \( V_k \)) to the element \( \rho_j \, dU_j \) on the \( j \)th body. Equation (2-41) may be simplified substantially by defining a set of numbers \( p_{kj} \) by

\[
p_{kj} = \frac{1}{4\pi \varepsilon} \sum_{j=1}^{N} \frac{\rho_j \, dU_j}{r_{kj}}
\]  (2-42)
These numbers possess an interesting property. If the entire charge distribution is strengthened or weakened by multiplying the charge density everywhere by some number $h$, the numbers $p_{kj}$ are unaffected, as may be verified by inspection of (2-42). In the limit, as $h$ approaches zero, the charges vanish altogether, but each $p_{kj}$ retains its value unaltered. The numbers $p_{kj}$ are thus dependent on the manner in which charge is distributed but not at all on the amount of charge; they are geometrical constants that express the spatial distribution. Physically, the denominator of (2-42) merely represents charge on the $y$th body, while the numerator measures the contribution of that charge to the potential of the $k$th body; thus, the $p_{jk}$ relate potential values to charges of the various bodies. Using (2-42), Equation (2-41) may be rewritten simply as

$$V_k = \sum_{j=1}^{N} p_{kj} q_j$$

(2-43)

where $q_j$ denotes the charge of the $j$th body. All the potentials and charges may be expressed in one matrix equation, of which (2-43) forms the $k$th row:

$$\begin{bmatrix} V_1 \\
\vdots \\
V_N \\
\end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\
\vdots & \vdots & \ddots & \vdots \\
p_{N1} & \cdots & p_{NN} & \end{bmatrix} \begin{bmatrix} q_1 \\
\vdots \\
q_N \\
\end{bmatrix}$$

(2-44)

The $p_{kj}$ are called potential coefficients and physically represent the potential of the $k$th body when unit charge is placed on the $j$th body, all other bodies being uncharged. Another way of stating this is

$$\frac{\partial V_k}{\partial q_j} = p_{kj}$$

(2-45)

The matrix of potential coefficients may be shown to be symmetric by the following argument. The potential of the $k$th body is, by definition, the incremental work done in adding charge to the $k$th body, the charges on all other bodies being held constant:

$$V_k = \frac{\partial W}{\partial q_k}$$

(2-46)

The total field energy $W$ is in general given by

$$W = \frac{1}{2} \int \rho V \, dU$$

(1-60)

which reduces to

$$W = \frac{1}{2} \sum_{j=1}^{N} V_j q_j$$

(2-47)

for a system of conductive bodies at rest since the potential $V_j$ is uniform throughout each body. Substituting this energy expression into (2-46),
\[ V_k = \frac{\partial W}{\partial q_k} = \frac{1}{2} \sum_{j=1}^{N} V_j \frac{\partial q_j}{\partial q_k} + \frac{1}{2} \sum_{j=1}^{N} \frac{\partial V_j}{\partial q_k} q_j \]

\[ = \frac{1}{2} V_k + \frac{1}{2} \sum_{j=1}^{N} p_{jk} q_j \]  

Combining (2-43) and (2-48), it is found that

\[ \sum_{j=1}^{N} (p_{kj} - p_{jk}) q_j = 0 \]

for any arbitrary set of charges \( q_j \). This is only possible if

\[ p_{kj} = p_{jk} \]  

which is the symmetry condition desired.

A simple example of the use of potential coefficients in problem solving may illustrate their physical significance. Let a transmission line with ground return run parallel to a farmer’s fence for some distance (Fig. 2.20). The fence wire will assume some potential different from ground because it is placed in the field of the line; conversely, if the fence is connected to ground (e.g., through the body of a farm animal), an induced charge will move through the ground connection onto the wire. The problem is best analyzed by using the method of images to avoid dealing with an infinite ground plane. In general, the potential of a line charged with \( q \) coulombs per unit length is known to be

\[ V = -\frac{q}{2\pi \epsilon} \log \frac{r}{r_0} \]  

where \( r \) denotes distance from the line and \( r_0 \) is the distance to the reference point \( V = 0 \). In this case, the ground plane is conveniently taken as the
Whenever a charge $+q$ is placed on the line, a corresponding charge $-q$ must appear on its image. The potential at any point is then obtained by superposing the potential contributions of the line and its image (the fence being assumed uncharged). Adapting Equation (2-41), the potential on the line conductor surface is found to be

$$V_{\text{conductor}} = \frac{1}{2\pi e} \left( \log \frac{2h}{R_L} \right) q_{\text{conductor}}$$

(2-50)

where $R_L$ is the conductor radius. The self-potential coefficient of the line is

$$p_{LL} = \frac{1}{2\pi e} \log \left( \frac{2h}{R_L} \right)$$

(2-51)

The potential of the fence for a unit charge on the line is, by a similar development,

$$p_{FL} = \frac{1}{2\pi e} \log \frac{r_2}{r_1}$$

(2-52)

The matrix relationship between potentials and charges is thus found to be

$$\begin{bmatrix} V_L \\ V_F \end{bmatrix} = \begin{bmatrix} \frac{1}{2\pi e} \log \frac{2h}{R_L} & \frac{1}{2\pi e} \log \frac{r_2}{r_1} \\ \frac{1}{2\pi e} \log \frac{r_2}{r_1} & \frac{1}{2\pi e} \log \frac{2h}{R_F} \end{bmatrix} \begin{bmatrix} q_L \\ q_F \end{bmatrix}$$

(2-53)

Any combination of line and fence charges and potentials may be solved for, using this matrix equation. Setting up a matrix of potential coefficients provides a technique for solving much more complicated problems with very little more labor. From (2-53), the potential coefficient matrix for a fence near an unbalanced three-phase, four-wire transmission line, for example, may be written by inspection.

The symmetric matrix $[p_{kj}]$ is in general not singular and may be inverted so as to yield

$$[q] = [c_{jk}][V]$$

(2-54)

The coefficients $c_{jk}$ must also form a symmetric matrix. Physically, they express the charge induced on the $j$th conductor if the $k$th conductor is at unit potential, all other conductors being held at $V = 0$. For this reason, they are called induction coefficients, except for the diagonal elements $c_{jj}$. There is no meaning in charge induced on a conductor by itself, so these diagonal elements are known as capacitance coefficients. In terms of these coefficients, the energy expression (2-47) may be written

$$W = \frac{1}{2} \sum_{j=1}^{N} \sum_{k=1}^{N} c_{jk} V_j V_k$$

(2-55)

a form especially useful because the potentials of a set of conductors are generally easier to determine experimentally than the charges.
10. Matrix Solutions to Potential Problems

A great many practical problems in electrostatics do not involve the determination of potentials from given charge distributions, but they do require finding the charge distribution that corresponds to some given set of potentials. The former problem is not altogether easy, for even in cases where the distribution of charges can be described fairly simply, evaluation of the integral in (1-50) may be difficult. On the other hand, the converse problem is much harder to solve, and analytic solutions in fact only exist for a limited number of cases. Many real problems can only be tackled with success by numerical techniques. A relatively straightforward approximate method is based on the potential coefficient matrix.

Let a prescribed potential distribution be given in the vicinity of a set of bodies. These bodies need not be of any specific shapes or kinds; nor do they need to be necessarily equipotentials. Let the bodies be subdivided into a set of elements in such a fashion that, to an acceptable approximation, the charge distribution and potential values are essentially uniform over each element. These elements may now be regarded as the multiple bodies of the potential coefficient method. Since the charge distribution over each element is assumed uniform, the potential coefficients, in accordance with Equation (2-42), become

\[ p_{kj} = \frac{1}{4\pi\varepsilon} \frac{\rho_j}{\rho_j} \int \frac{dU_j}{r_{kj}} = \frac{1}{4\pi\varepsilon U_j} \int \frac{dU}{r_{kj}} \]  

(2-56)

If the bodies in question are conductive, all charges will appear as surface charges and the integrals in (2-42) become surface integrals. Correspondingly,

\[ p_{kj} = \frac{1}{4\pi\varepsilon S_j} \int \frac{dS}{r_{kj}} \]  

(2-57)

In this latter case, the \( S_j \) are called subareas, and the method of calculation is referred to as the method of subareas.

The two forms (2-56) and (2-57) are fully equivalent. In both cases, the integrals to be evaluated contain known integrands and present no difficulty, at least in principle. Once the potential coefficient matrix is known, the problem of determining the charge distribution is solved by inverting to find the influence coefficient matrix and postmultiplying by the known matrix of potentials, as in Equation (2-54). From the results, it is generally not difficult to judge whether the initial assumptions of uniform charge and potential distribution over each element are tenable. If it appears that they are doubtful, subdivision of the given bodies in another manner may be indicated. Since there is no restriction on the size or shape of the sections, the initial calculation can often be improved upon by choosing a coarser
subdivision in areas where the variation of charge density appears to be small and a fine subdivision where it varies more rapidly.

In a great many practical cases, the problems to be solved have one or more planes or axes of symmetry. Preliminary use of the method of images may in such circumstances reduce the number of elements to be considered very substantially, permitting higher accuracy of solution to be obtained. It is usual for the computing time required for matrix inversion to rise as the cube of the number of subareas; since the potential coefficient matrix is square, the memory space required rises as the square of their number. Typically, inversion of a $15 \times 15$ matrix takes about 10,000 times as long as a single multiplication (on an IBM 7044, this means around 350–400 milliseconds) and requires 225 memory locations. Doubling the fineness of subdivision, it is seen, leads to a memory allocation of 900 spaces and inversion time equivalent to about $10^8$ multiplications. From a practical point of view, presently available computers restrict the maximum matrix order that can be handled conveniently to about 50 or 100; on very large machines, matrices larger than $200 \times 200$ have been inverted. It is evident that any preparatory work that might permit finer subdivision without increase in matrix size is well worth the effort.

To illustrate the method of calculation, let it be desired to find the charge distribution along an isolated metal tube or rod of circular cross section. The tube will first be subdivided into a number of elements, each of which has the form of a short piece of tubing. Each potential coefficient, relating the potential at the $j$th element to the charge at the $k$th, may be evaluated using Equation (2-57). Symmetry dictates that the charge distribution along the tube be similar in both halves; thus, reduction of the number of effective elements by half is possible. As a reasonable approximation, the potential of each tube element may be thought to equal the potential in the center of the element and on the tube center line. The latter is quite easy to calculate. In accordance with Equation (2-57), the potential at the center point of the $j$th element in Fig. 2.21 is determined by

$$p_{jk} = \frac{1}{4\pi \varepsilon S_k} \int \frac{dS}{r_{jk}} + \frac{1}{4\pi \varepsilon S_{k'}} \int \frac{dS}{r_{jk'}}$$

(2-58)

![Fig. 2.21](image-url)
where \( k \) and \( k' \) attached to the integral signs denote integration over the symmetrically placed elements \( k \) and \( k' \), respectively; similarly, distances \( r_{jk} \) and \( r_{jk'} \) are measured from the \( j \)th element to the corresponding elements. The two terms arise from the fact that the charges on elements \( k \) and \( k' \) must be equal at all times, and both must influence \( V_j \).

It should be noted that the distances \( r_{jk} \) always denote physical separation and are scalar numbers, i.e., always positive. Taking the tube to have a radius \( R \) and the individual elements to be of length \( 2\Delta x \), Equation (2-58) reduces to

\[
p_{jk} = \frac{1}{8\pi \varepsilon \Delta x} \left( \int_{|x_k-x_j|+\Delta x}^{|x_k-x_j|+\Delta x} \frac{dx}{\sqrt{R^2 + x^2}} + \int_{|x_k-x_j|+\Delta x}^{|x_k-x_j|+\Delta x} \frac{dx}{\sqrt{R^2 + x^2}} \right) (2-59)
\]

The integration is easily carried out, yielding after substitution of limits

\[
p_{jk} = \frac{1}{8\pi \varepsilon \Delta x} \left( \sinh^{-1} \frac{|x_k-x_j|+\Delta x}{R} - \sinh^{-1} \frac{|x_k-x_j|+\Delta x}{R} + \sinh^{-1} \frac{|x_k+x_j|+\Delta x}{R} - \sinh^{-1} \frac{|x_k+x_j|+\Delta x}{R} \right) (2-60)
\]

Although lengthy for hand calculation, this form is well suited to machine computation, for it is seen on inspection that the self-potential coefficients \( p_{jj} \) are given by the same formula as the off-diagonal terms \( p_{jk} \), facilitating programming. The inverse hyperbolic sine is not normally available as a computer library function. To obtain numerical values, one may employ

\[
\sinh^{-1} x = \log (x + \sqrt{x^2+1}) (2-61)
\]

This equation is valid only for positive \( x \), and its use presupposes attaching the correct sign to the result. A function subroutine \( \text{ARSH}(x) \) that does so is given in Appendix III.

The problem is now ready for computer programming. After the usual bookkeeping operations, the matrix of potential coefficients is constructed using Equations (2-60) and (2-61). It might be noted that since any potential coefficient matrix is of necessity symmetric, only one-half of it need be computed. This matrix is then inverted to yield the influence coefficient matrix, using a library matrix-inversion subroutine. As the problem requires determining the charge distribution, assuming the tube to be at the same potential throughout, all elements of the potential matrix in Equation (2-54) may be taken to be equal to some arbitrary number, say unity. The charges are calculated simply by substitution in (2-54). Figure 2.22 shows the results of such a calculation for a tube 60 cm long and 2 mm in radius, using first a subdivision into 60 elements (i.e., a \( 30 \times 30 \) matrix representation) and, for comparison, the result of a subdivision twice as coarse. The programme used to obtain these results may be found in Appendix II.
The above calculation is based on the two approximations that (a) the continuous charge distribution may be approximated by a stepwise sequence of values, and (b) the potential of each element equals the potential value at its midpoint. For elements shorter than, say, 5 or 10 radii, assumption (b) becomes untenable, especially for elements near the ends of the rod. A more refined calculation can now be carried out, using the results shown in Fig. 2.22 as a basis. Without increasing matrix size, the accuracy can be improved considerably if the rod is modeled not by a set of elements of equal length but by relatively long tubular sections near the middle (where charge density varies but slightly) and ring-shaped sections near the ends. For the latter, potential coefficients may be derived readily from Equation (1-58).

For this particular problem, no exact analytic solution is known. An approximate solution for the self-potential coefficient of a slim rod, however, does exist; it is comforting to note that even the coarser of the two numerical solutions agrees with it to better than 1 percent.

11. Capacitance and Capacitors

A particularly important special case of the problem of multiple charged bodies is the system of two bodies oppositely charged with equal amounts of charge, i.e., with charges $+q$ and $-q$. For such systems, Equation (2-44) becomes

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} +q \\ -q \end{bmatrix}$$

(2-62)
The potential difference $V = V_1 - V_2$ between the two bodies is readily found from (2-62) as

$$V = (p_{11} - 2p_{12} + p_{22})q$$  \hspace{1cm} (2-63)$$
and the energy of the system is, from (2-47),

$$W = \frac{1}{2}Vq$$  \hspace{1cm} (2-64)$$

It is usual to define the capacitance of a two-body system as the geometrical quantity

$$C = \frac{1}{p_{11} - 2p_{12} + p_{22}}$$  \hspace{1cm} (2-65)$$

From Equations (2-63) and (2-64), expressions for the capacitance in terms of electrical quantities may also be found. Thus

$$C = \frac{q}{V} = \frac{\psi}{V}$$  \hspace{1cm} (2-66)$$

$$C = 2\frac{W}{V^2}$$  \hspace{1cm} (2-67)$$

Methods of finding capacitance analytically are generally based on either (2-65) or (2-66), while experimental techniques rely on (2-66) or (2-67).

As a simple example of capacitance calculation from the potential coefficients, let the capacitance from line to fence wire be found for the example of Fig. 2.20. Taking the coefficients $p_{jk}$ as given by Equation (2-53), there obtains on substitution into (2-65)

$$C = \frac{2\pi e}{\log \frac{4h/R_1}{R_2/R_2}}$$  \hspace{1cm} (2-68)$$

Calculation of this capacitance by the "direct" technique, using Equation (2-66), is made quite a bit more laborious by the presence of the ground plane, as is readily verified. Finding capacitance between the conductors of a unit length of coaxial cable, on the other hand, is quite simple by the direct method. The flux density anywhere in the cable is easily calculated by applying Gauss's law to a cylindrical surface concentric with the inner and outer conductors:

$$D = \frac{q}{2\pi r}$$  \hspace{1cm} (2-69)$$

The field intensity that corresponds to that flux is

$$E = \frac{q}{2\pi r}$$  \hspace{1cm} (2-70)$$

This field intensity, however, was already found in terms of the potential difference, Equation (2-5). Equating the two values of field intensity, a relationship between $V_0$ and $q$ is found, and therewith the capacitance of a unit length of cable determined:
\[ C = \frac{2\pi \varepsilon}{\log \left( \frac{r_0}{a} \right)} \]  

(2-71)

This method is general and other problems may be solved by following it: the flux density (or field intensity) is determined by any convenient method, first in terms of potential, then in terms of charge. These two are then equated to find capacitance. Another simple application of this method is in finding the capacitance of a section of unit area in the midregion of a very large pair of parallel plates. Mathematically, this problem may be modeled by a pair of infinite planes located at, say, \( x = -a \) and \( x = +a \). Suppose the potentials of these two plates to be \( -V_0 \) and \( +V_0 \), respectively. In the space between plates, \( V \) may be argued to be independent of \( y \) and \( z \), so that Laplace’s equation becomes simply

\[ \frac{d^2V}{dx^2} = 0 \]  

(2-72)

On integrating twice and introducing the boundary potentials, the solution is found to be

\[ V = V_0 \frac{x}{a} \]  

(2-73)

The field intensity has the value \( E = V_0/a \) throughout the interconductor space, independently of all three coordinates. An expression relating this field to charge is obtained by considering that the charge per unit area \( q \) must equal the surface integral of the flux density vector \( \mathbf{D} \) taken over a unit area very close to one of the plates. But from the above,

\[ \mathbf{D} = \frac{V_0 \varepsilon}{a} \]

so that the capacitance is found to be

\[ C = \frac{\varepsilon}{2a} \]  

(2-74)

per unit area of the two plates.

If the conductors are of a complicated shape, no analytic solution may be possible. The second, or “direct,” method may still be used to find capacitance from a field map constructed by relaxation or iteration methods. Let the map divide the space between conductors into \( N_r \) potential squares; the total potential between conductors must then be

\[ V_{\text{total}} = N_r \Delta V \]  

(2-75)

where \( \Delta V \) is the potential difference across the faces of one square. Similarly, suppose the map divides the total flux into \( N_\psi \) squares; the total flux is known to be

\[ \psi_{\text{total}} = N_\psi \Delta \psi \]  

(2-76)
where $\Delta \psi$ is the flux enclosed by each tube. The capacitance is given by the ratio of total flux to total potential; dividing the above expressions,

$$C = \frac{N\psi}{N_r} \epsilon \text{ per unit length} \quad (2-77)$$

Note that this simple expression depends entirely on having all the map made up of squares; with any other kind of rectangles, greater complexity results.

Finally, it should be observed that neither the International Technical Vocabulary nor any recognized dictionary lists any such word as capacitative; the adjective pertaining to capacitance is capacitive, although some authors seem to have become confused in recent years. (Interestingly, inductive and resistive are not listed either.)

### 12. Electromechanical Devices

In numerous devices, use is made of Coulomb forces to produce mechanical motion or, conversely, to convert mechanical into electrical energy. Because high rates of energy conversion in relatively small volumes give rise to high voltages and attendant insulation problems, electrostatic (i.e., Coulomb-force) machines are much less widely used than electromagnetic (i.e., Lorentz-force) machines for the production of mechanical motion. Two practical electrostatic devices of engineering importance are electrostatic voltmeters, and high-voltage generators.

A simple form of electrostatic voltmeter consists of a jeweled-bearing meter movement supporting one plate of an air-dielectric capacitor. The other plate is fixed to the meter case (see Fig. 2.23). For simplicity in explanation, it will be assumed that the plate separation $D$ is much smaller than the inner and outer radii $R_s$ and $R_t$ of the plates (these are taken to be circular sectors). The plates may then, to a first approximation, be treated as a portion of an infinite parallel-plate capacitor. From Equation (2-74), the capacitance is

$$C = \frac{(R_s^2 - R_t^2)\theta}{2D} \epsilon \quad (2-78)$$

The torque acting on the meter movement may be found by differentiation of stored energy with respect to angle:
\[ T = \frac{\partial W}{\partial \theta} = \left( R_o^2 - R_i^2 \right) \epsilon V^2 \] (2-79)

where Equation (2-67) has been used to evaluate energy. Given a movement with a normal linearly acting spring, the meter deflection will be proportional to torque.

From (2-79) it is at once evident that an electrostatic meter movement will respond to the square of applied voltage. Several interesting points arise from this fact. First, since the square of voltage is always a positive number, the meter is unable to distinguish polarity. Secondly, and for the same reason, it is possible to measure alternating voltages without any auxiliary rectifiers. Third, if the meter is calibrated with direct voltages, it will indicate true root-mean-square value of any waveform. It should be observed that when used to measure direct voltage, such a voltmeter will draw no current. If used to measure alternating or time-varying voltages, it will draw current but no power.

Several variants on this basic movement have been designed. For example, it is possible to leave both plates fixed but attach a dielectric slab to the movement; in this way, no moving electrical connection is needed. The principle of operation is precisely the same, and the torque may be calculated from the stored energy by differentiation. If a scale not calibrated proportionally to volts squared, but in some other manner, is desired, the shape of one or both plates (or of the dielectric) may be altered to produce almost any kind of scale calibration.

A second important practical class of electrostatic devices is the rotating high-voltage generator, occasionally known as a Felici machine. Such machines are built for output voltages in the range of 50 kilovolts to 10 megavolts in sizes from a few watts to a few kilowatts, although there is no reason of principle for a size or voltage restriction. The principle of operation is schematically indicated in Fig. 2.24. In essence, the machine consists of a set of fixed plates between which a set of movable plates is free to rotate. Only one plate of either set is shown in Fig. 2.24. The movable plates are equipped with a commutator, by means of which they are alternately connected to brushes B1 and B2. When the movable plates are fully meshed with the fixed set, their capacitance reaches a maximum value \( C_{\text{max}} \). The sets of plates are then charged through brush B1 to the battery voltage \( V_0 \). As the movable set of

\[ V_0 \]

\[ B1 \]

\[ B2 \]

\[ V_h \]

\[ \text{Fig. 2.24} \]
plates rotates away from the fixed set, the commutator breaks the battery connection so that the charge placed upon the plates remains there. When the movable plates reach their farthest position from the fixed ones, so that the capacitance value has a minimum $C_{\text{min}}$, the two sets of plates are connected through brush $B2$ to a load. Because the charge $Q = CV$ is the same as at the fully meshed position, the load voltage is

$$V_h = \frac{C_{\text{max}}V_0}{C_{\text{min}}}$$

(2-80)

The energy delivered to the load, assuming full discharge, is

$$W_h = \frac{1}{2}QV_h$$

(2-81)

and the energy delivered by the battery

$$W_b = \frac{1}{2}QV_0$$

(2-82)

The difference between these two energies represents the mechanical energy $W_m$ that must be furnished by the shaft:

$$W_m = \frac{C_{\text{max}}}{C_{\text{min}}} - C_{\text{min}}W_0$$

(2-83)

It is evident from this energy balance that the electrostatic machine is a true energy converter, the applied potential difference serving primarily as an excitation source corresponding in many ways to the field winding of a direct-current electromagnetic machine. For a capacitance ratio $(C_{\text{max}}/C_{\text{min}})$ of 20, for example, 95 percent of the output energy is derived from mechanical input.

Electrostatic generators are able to operate at much higher voltages than electromagnetic ones and possess very low weight as well as inertia, since there are heavy iron members neither in the rotor nor in the stator. They are employed in electrostatic dust-precipitation equipment, X-ray machines, and other applications such as electrostatic paint spraying. From the basic principle of operation it is evident that an electrostatic generator can only furnish a fixed maximum amount of charge to its load in each revolution, this amount being $C_{\text{max}}V_0$. Consequently the machine is intrinsically short-circuit-proof; it cannot be harmed electrically because it automatically limits its own maximum terminal current. The total power rating of such machines is clearly dependent on the kind of insulating material employed, and a great deal of active research is in progress to develop particularly suitable liquid and gaseous dielectrics for this application.

**Readings**

Problems dealing with coaxial cables and parallel-wire lines were well known to Maxwell (1), whose entire Treatise, incidentally, is still so fundamental and so widely read as to be currently in print. Among modern books,
those by Boast (2) and by Binns and Lawrenson (3) stand out as particularly lucid on the methods of superposition and images, including various cases of plane and cylindrical symmetry. Boast is undoubtedly the easier to read, while Binns’ and Lawrenson’s book is a good deal more comprehensive. The modern trend to bundle-conductor transmission lines has recently produced a substantial number of papers on that topic; little is to be found in textbooks as yet, but the papers by King (4) and Timasheff (5) make easy reading and do not require specialized knowledge in the area.

There can be no doubt that relaxation methods are among the most useful techniques at present available to the engineer for solving field problems, and indications are that they will become even more important in the future. A detailed explanation of the graphical method was given by Richardson (6), who considered rotationally symmetrical as well as plane problems. Somewhat simpler treatment is given to the method by Weber (7) and Boast (2); both may also be recommended as references on numerical relaxation. McCormick and Salvadori (8) not only discuss the latter method but exhibit several sample computer programmes. The book by Binns and Lawrenson (3) deals with relaxation as well as Monte Carlo calculations. Allen (9) and Thom and Apelt (10) are not difficult to read; both may be recommended to the student. The book by Forsythe and Wasow (11) is more advanced and may serve as a guide to further work. This field, however, is still rapidly developing, so that the most recent advances must be sought in periodicals such as the Computer Journal.

As an example of practical application of relaxation, Metcalf (12) has analyzed rectangular transmission lines quite exhaustively, and in the process he has produced an excellent description of the method. A similar problem, but with rotational symmetry, is solved by Mitra and Salvage (13), who also give a careful explanation of the method. Boers (14) reports on a quite sophisticated computer programme of a similar sort, but with numerous advanced features.

Solution of problems by matrix methods is illustrated in the paper by Carroll and Higgins (15) and applied to a practical design by Chen and Kulikowski (16). An article reporting a similar method of calculation, but perhaps a bit more difficult for the novice to read, is by Cristal (17). The slim rod problem examined above is dealt with by an approximate analytic technique in Jordan’s book (18), which also includes other varied examples of electrostatics problems.

A review article on electrostatic machines by Felici will serve to illustrate the present state of the art (19).

For problems, the student is again referred to Batygin and Toptygin (see Chapter 1), as well as to Myers (20), who provides a comprehensive set of worked examples.


PROBLEMS

2.1 A coaxial cable of the type discussed has its inner conductor at zero potential and its outer conductor at potential $V_0$. Find the electric field strength and potential at a point outside the outer conductor.

2.2 A coaxial cable has a potential difference of $V_0$ volts between its inner and outer conductors. Find the corresponding charge on the inner conductor.

2.3 Given that the highest possible field intensity in air is 3 megavolts per meter, find the minimum radius of an isolated charged sphere at a potential of 500 kilovolts (this may serve to give an idea of the least permissible radius of hardware parts for high voltages).

2.4 Equation (2-19) has the form $(x - h)^2 + y^2 = R^2$. Show that

$$\frac{b}{2} = \sqrt{h^2 - R^2}$$

Use this result to plot accurately the equipotentials in the neighborhood of a pair of conductors 6 cm in diameter, spaced 10 cm center to center.

2.5 Use the preliminary result of Problem 2.4 to show that the centers of two cylindrical conductors of unequal radii $R_1$ and $R_2$ must be located at

$$X_1 = \frac{D^2 + R_1^2 - R_2^2}{2D} \quad \text{and} \quad X_2 = \frac{D^2 - R_1^2 + R_2^2}{2D}$$

where $D$ is the distance between actual geometrical centers of the conductors, and $X_1$, $X_2$ are the distances of the conductor centers from a zero-potential plane between conductors. From these expressions, deduce the geometrical construction required to plot accurately the field of two conductors, with radii 5 cm and 3 cm, respectively, (a) if their centers are separated by 10 cm, and (b) if their centers are separated by 16 mm, i.e., if they form an eccentric cable.

2.6 A particular type of insulating material has a certain breakdown field strength, say $E_b$. Given the outside conductor diameter to be used in constructing a coaxial cable insulated with this material, find the inner conductor diameter that will permit the highest voltage rating for the cable.

2.7 A coaxial cable is to be built using two different dielectrics of equal breakdown field strength but different permittivities. Find the relative thicknesses of the two insulation layers required to maximize the permissible potential difference between conductors. How does the voltage rating compare with that obtained using only one dielectric?

2.8 Write an expression for potential distribution in the vicinity of a two-conductor bundle transmission line in which the conductors in each bundle are placed one above the other (in contrast to lying in a flat plane, as in the case already examined). Give approximate potential distributions in the regions (a) far
away, (b) near the right-hand bundle, and (c) near the upper wire of the right-hand bundle. Sketch some of the equipotentials. Is this arrangement preferable to the flat-plane layout?

2.9 A rectangular transmission line consists of an outer conductor $32 \times 40$ cm and an inner conductor $16 \times 20$ cm, coaxially placed. Solve for the potential distribution in the interconductor space. Use a 2 cm mesh first, then a 1 cm mesh, and compare results. Arrange your computer printout so as to place the mesh potentials in their correct relative positions.

2.10 Solve for the potential at one point $(x, y)$ in the interconductor space of Problem 2.9 by the Monte Carlo method, using a subroutine that provides random numbers between $-1.$ and $+1.$ Arrange for the calculated potential value to be printed out after every two or three random walks, along with an indication of the number of steps in the walk. Use a 2 cm mesh, and compare your result with that obtained by relaxation.

2.11 A high-power radio-frequency transmission line consists of a sheet-metal duct of rectangular cross section $30 \times 40$ cm, and a round sheet-metal tube 14 cm in diameter. By graphical relaxation techniques, find (a) the capacitance of this structure, and (b) the maximum electric field in it, assuming the two conductors to be located coaxially. Also plot charge density on the inner conductor against angular position.

2.12 A round cylindrical wire of radius $R$ is placed halfway between infinite metallic planes separated by a distance $2a$. Solve for the capacitance per unit length, assuming both planes are at zero potential and the wire is at potential $V_0$. Use the method of images, and assume that $R$ is much smaller than $a$. (Hint: The image potentials can be arranged so as to combine into Wallis' product and yield an analytically known answer.)

2.13 In Problem 2.12, assume $R = 1$ cm, $a = 4$ cm. Solve by relaxation to find the electric field strength along a line normal to the flat surfaces and passing through the center of the wire. Find also the capacitance per unit length of this structure and the charge distribution on the infinite plates.

2.14 Three wires have radii $r_1, r_2, r_3$, respectively, and are located at $(0, 0, z), (0, 2, z), (3, 0, z)$. Find the induction and capacitance coefficients between these wires.

2.15 Find the conductor-to-conductor capacitances, and the capacitance to ground, of a three-phase transmission line whose conductors lie in a common horizontal plane.

2.16 Write the matrix of potential coefficients for the three-phase line and fence wire shown in Fig. 2.25. Assume the line conductor potentials to vary sinusoidally in time in a balanced three-phase manner. Find the potential variation at the insulated fence wire. Find also the current carried by the grounding wire if the fence is connected to ground.

2.17 Find the capacitance of a capacitor composed of two concentric spheres.

2.18 An electrostatic voltmeter of the type shown in Fig. 2.23 consists of two plates with inner and outer radii of 1 and 3 cm, respectively. The plates are
2.19 Show that the induction coefficients \( c_{jk} \) are always negative, and the capacitance coefficients \( c_{jj} \) always positive for a physically realizable system.

2.20 Find the charge distribution on a pair of parallel rods 60 cm long and 2 mm in radius, equally but oppositely charged, if their center-to-center separation is 20 cm.

2.21 Repeat Problem 2.20 for a center-to-center spacing of 5 cm. Find the capacitance of the two-rod structure in both cases, and compare with the capacitance of a 60 cm long section of infinitely long lines of similar wire radius and spacing.
Mapping of Steady Currents and Fields

All the problems dealt with so far have involved charges in static equilibrium. The more general case of dynamic equilibrium will be examined next, seeking out problems that require no relativistic corrections in this chapter but including the effects of the Lorentz contraction in the next. The equations governing fields in dynamic equilibrium are found to resemble those underlying electrostatics, and some additional methods for solving problems applicable to either kinds of fields are introduced.

1. The Continuity Equation of Current

In electrostatics, situations were examined in which charges were displaced from their rest positions, but quantitative results were derived only for the steady state when all charges were held immobile. During the displacement process, there is defined everywhere in space not only a charge volume density $\rho$ but also a charge velocity vector $v$. The motion is characterized by these...
two quantities and is governed by a fundamental continuity equation which
derives from the necessity of satisfying the law of charge conservation at all
times. Only relativistically low-velocity cases will be considered in this
chapter. As will be seen in Chapter 4, however, no new terms need be added
at high velocities either.

Let $U$ be a volume through which charges are moving. If the inflow and
outflow do not exactly balance at all times, charge will accumulate inside at
some rate $dQ/dt$. This rate may be found by taking the rate of net charge
inflow at each point on the bounding surface and integrating over the entire
surface. There is obtained

$$\frac{dQ}{dt} = -\int \mathbf{v}_p \cdot \mathbf{dS}$$  \hspace{1cm} (3-1)$$

The closed surface integral may be converted into a volume integral by means
of the divergence theorem:

$$\frac{dQ}{dt} = -\int_V \text{div} (\rho \mathbf{v}) \, dU$$  \hspace{1cm} (3-2)$$

Alternatively, the rate of accumulation may be written in terms of the charge
contained within the volume $U$ as

$$\frac{dQ}{dt} = \int_U \frac{\partial \rho}{\partial t} \, dU$$  \hspace{1cm} (3-3)$$

Eliminating $dQ/dt$ between Equations (3-2) and (3-3), it is found that

$$\int \left( \text{div} (\rho \mathbf{v}) + \frac{\partial \rho}{\partial t} \right) \, dU = 0$$  \hspace{1cm} (3-4)$$

which must hold for any arbitrary volume $U$. It may be concluded that the
integrand must be zero; if it were not, the volume of integration could not
be arbitrary. [Suppose the integrand were positive in some small region
within $U$. A new volume $U'$, containing only that region, could be defined,
violating Equation (3-4).] Thus

$$\text{div} (\rho \mathbf{v}) + \frac{\partial \rho}{\partial t} = 0$$  \hspace{1cm} (3-5)$$

This equation is called the \textit{continuity equation} of electric current, for it
expresses the physical impossibility of a charge-free “hole” in a region of
current flow. Usually the product $\rho \mathbf{v}$ is denoted by

$$\mathbf{J} = \mathbf{v}_p \rho$$  \hspace{1cm} (3-6)$$

and called the \textit{current density vector}. As may be seen from the above develop-
ment of the continuity equation, it measures the rate of motion of charge
across a surface and has the direction of the velocity vector. Its units are
amperes per square meter. In terms of this vector, the continuity equation
(3-5) reads
\[ \text{div } \mathbf{J} + \frac{\partial \rho}{\partial t} = 0 \]  
(3-7)

By using Gauss's law in the differential form, charge density \( \rho \) may be eliminated altogether with the result

\[ \text{div} \left( \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \right) = 0 \]  
(3-8)

An estimate of numerical orders of magnitude of \( \rho \) and \( \mathbf{v} \) is probably helpful. As discussed in connection with electrostatics, all the free charges in a conductor are moved by an applied field, so the charge density \( \rho \) used here must mean not merely the density of extraneously introduced (excess) charges but all the movable charge in the material. In the case of copper, for example, this amounts to two free electrons per atom, or approximately 12 kilocoulombs per cubic centimeter. Consequently, a current density of \( 5 \times 10^6 \) amperes per square meter (a reasonable design figure for many kinds of electrical apparatus) involves velocities of the order of millimeters per second, an order of magnitude relativistically unimportant. Nevertheless, relativistic corrections will be necessary when forces between moving charges are calculated in Chapter 4. Even very small velocities, when combined with the almost unimaginably huge forces between kilocoulomb charges, produce corrections still of significant magnitude.

### 2. Potential and Charge Distributions

Application of an electric field to any conducting material results in movement of charge. The current density that corresponds to a given field strength is dictated by the field strength as well as by the atomic and molecular structure of the material. The microscopic aspects of the conduction process are properly a part of solid-state physics and will not be considered here. It is sufficient for macroscopic purposes to describe each conductive material by stating a functional relationship between \( \mathbf{E} \) and \( \mathbf{J} \). In all isotropic materials, these vectors are collinear, so that their relationship may be written

\[ \mathbf{J} = g \mathbf{E} \]  
(3-9)

where \( g \) is a scalar function. If the material in question is also homogeneous, \( g \) cannot depend on position; and if \( g \) is constant (independent of \( E \)), the material is said to be linear. The function \( g \) is called the conductivity of the material. In the following, only materials of constant conductivity will be considered. Otherwise very great mathematical difficulties arise, obscuring the substance of the argument.

By definition, the potential and field anywhere are related by

\[ \mathbf{E} = -\text{grad } V \]  
(1-46)
From this it follows that $J$ must be normal to equipotential surfaces, for (3-9) requires it to be collinear with the field vector $E$. In the case of a steady current, there can be no accumulation of charge at any point, and the continuity equation becomes simply

$$\text{div } J = 0$$  \hspace{1cm} (3-10)

Combining the last three equations, there obtains

$$\text{div } \text{grad } V = 0$$  \hspace{1cm} (3-11)

That is to say, Laplace's equation is valid in a conducting medium carrying a steady current. Unlike in electrostatics, there will be two sets of lines orthogonal to the equipotentials: the current flow lines and the electric flux lines, i.e., lines of $J$ and $D$. Either may be found from the same field plot and must therefore be geometrically coincident with the other.

Existence of the vectors $D$ and $E$ implies existence of a static charge distribution in space, in addition to the moving charges. To illustrate this point, the very simple example of a round wire carrying a current may be examined qualitatively. Let the two ends of the wire be maintained at potentials $+V_0$ and $-V_0$ by equipotential end pieces as shown in Fig. 3.1. It is readily verified that these boundary values correspond to a uniform distribution of current throughout the wire. If $J$ is a uniform vector, so is $E$. Uniform $E$ in turn implies that all equipotential surfaces are uniformly spaced parallel planes. This requirement is satisfied by the given boundaries, showing the solution $J = \text{uniform}$ to hold in the wire. There can of course be no current outside the wire, so $J = 0$ everywhere outside.

If there were no conducting wire, but only the equipotential end pieces, the electric field distribution (flux and equipotential lines) would be as shown in Fig. 3.2a. Because there is a conductor, however, this field must be altered: lines of flux must coincide with lines of current everywhere within the conductor. $D$ must then also be uniform within the wire, as shown in Fig. 3.2b. Such a distribution of $D$ cannot arise from charges on the two end equipotential surfaces alone; there must be some other charge distribution to cancel the radial component of $D$ everywhere inside. But inside as well as outside the wire, Laplace's equation has been shown valid, so no additional charge
distribution is possible either inside or out. The necessary charges must therefore reside on the wire surfaces as indicated in Fig. 3.2c.

This surface charge must exist if a steady current flow is to be possible at all, for in its absence the necessary differential equations cannot be satisfied. It is this surface charge, furthermore, that sustains the potential distribution required to keep current flowing, even though the wire be twisted, stretched, or contorted in the most complicated manner. For example, the potential and field inside the wire must be radically different from those outside in the case of a wire of uniform diameter that forms a long loop from one terminal of a battery to the other. The midpoint of the wire may well be in a zone of very low external field intensity, but the field inside the wire must still be uniformly high, so as to enable \( \mathbf{J} \) to satisfy the continuity equation.

3. Energy and Power

In moving charges about, work is done. If a steady current is caused to flow, there is continuous charge motion requiring a certain continuous power. Both this power and the total energy may be calculated easily from first principles. The force on a set of charges in a field \( \mathbf{E} \) is, as usual, given by

\[
\mathbf{dF} = \mathbf{E} \, dq,
\]

Supposing the entire charge distribution to be moved a distance \( \mathbf{dr} \), the work \( dW \) is found to be

\[
dW = \mathbf{F} \cdot \mathbf{dr} = \int_\mathbf{r} \rho \mathbf{dr} \cdot \mathbf{E} \, dU
\]

If this amount of work is done in a length of time \( dt \), the power may be calculated:

\[
\frac{dW}{dt} = \int_\mathbf{r} \rho \frac{d\mathbf{r}}{dt} \cdot \mathbf{E} \, dU
\]
But the time rate of distance is velocity, \(\frac{d\mathbf{r}}{dt} = \mathbf{v}\), so that the power may be written

\[
\frac{dW}{dt} = \int_U \mathbf{J} \cdot \mathbf{E} \, dU
\]  

(3-15)

This expression gives the power required to move charges through any arbitrary volume \(U\). If only a very small volume \(\delta U\) is considered, \(\mathbf{J}\) and \(\mathbf{E}\) are essentially constant throughout the integration. As a result, (3-15) may then be simplified to

\[
\mathbf{J} \cdot \mathbf{E} = \frac{1}{\delta U} \frac{dW}{dt}
\]  

(3-16)

The product \(\mathbf{J} \cdot \mathbf{E}\) is called the power density because it represents, by (3-16), the power per unit volume at any point in the field.

A surface-integral formulation of the power is frequently useful and is obtainable from the volume integral of (3-15). Using the definition of electric scalar potential,

\[
\frac{dW}{dt} = \int_U \mathbf{J} \cdot \mathbf{E} \, dU = -\int_U \mathbf{J} \cdot \nabla \mathbf{V} \, dU
\]

\[
= -\int_U \nabla \cdot (\mathbf{VJ}) \, dU + \int_U \mathbf{V} \cdot \nabla \mathbf{J} \, dU
\]

The last term on the right-hand side may be written in terms of the rate of charge accumulation, using the continuity equation (3-5). With the first term converted to a surface integral,

\[
\frac{dW}{dt} = -\oint \mathbf{VJ} \cdot d\mathbf{S} - \int_U \mathbf{V} \frac{\partial \mathbf{J}}{\partial t} \, dU
\]  

(3-17)

In equilibrium current flow (i.e., after the surface charges discussed in the previous section have been established), there is no charge accumulation with time, and the extreme right-hand term of (3-17) vanishes. Thus

\[
\frac{dW}{dt} = -\oint \mathbf{VJ} \cdot d\mathbf{S}
\]  

(3-18)

which is the surface integral sought. It is of particularly great usefulness for problems in which currents enter and leave a conductive body at only a few portions of its surface, for it is then only necessary to evaluate the integral for those portions. An exceptionally important special case is that of a body with only two such surface segments, each forming an equipotential (for example, the wire of Fig. 3.1). Supposing those two surfaces to be at potentials \(V_1\) and \(V_2\), respectively,

\[
\frac{dW}{dt} = -V_1 \oint_{s_1} \mathbf{J} \cdot d\mathbf{S} - V_2 \oint_{s_2} \mathbf{J} \cdot d\mathbf{S}
\]  

(3-19)

It is usual to define current as the rate of charge movement across a given surface \(S\),
\[ i = \int \mathbf{J} \cdot d\mathbf{S} \quad (3-20) \]

Continuity demands that the two surface integrals (i.e., the two currents) of Equation (3-19) be equal in equilibrium current flow. As a result,

\[ \frac{dW}{dt} = (V_1 - V_2)i \quad (3-21) \]

The difference in signs arises from the fact that the outward normal is collinear with the current density at one surface but antiparallel at the other.

Since the current \( i \) depends on the internal field, and hence on the potential difference \( (V_1 - V_2) \), it is usual to define resistance \( R \) of a two-terminal body by

\[ \frac{dW}{dt} = i^2R \quad (3-22) \]

The resistance may be shown to be independent of field and current, using an argument similar to that employed to show capacitance to be a purely geometrical quantity. Because the argument rests in part on the invariance of \( g \) (or \( e \), respectively), it is not valid except for linear, isotropic media; in fact, the resistance or capacitance of a nonlinear, anisotropic body is not independent of the fields.

4. Analogue Solutions of Equilibrium Problems

Electrostatic fields in all source-free regions were shown in Chapter 1 to satisfy Laplace’s equation

\[ \text{div} \ \text{grad} \ V = 0 \quad (1-67) \]

Exactly the same differential equation defines the potential distribution in conductive bodies in current equilibrium. This fact makes it possible to obtain solutions to many difficult electrostatic problems by constructing analogous potential distributions in conductive materials, where direct measurement of potential is readily possible. The results are obtained in the form of a field map and therefore are subject to the same restrictions as solutions found by relaxation techniques: they are numerical in nature and are difficult to record and evaluate in more than two-dimensional form.

One widely used form of analogue is a flat tank with liquid conductor or electrolyte. By using some convenient salt in aqueous solution, the conductivity may be chosen almost at will; it may be made to suit the input impedances of the instruments used, as well as the current and voltage capabilities of the sources. A disadvantage of electrolytic tanks is the tendency for accuracy to be reduced by electrochemical reactions (decomposition, polarization, etc.) at all source electrodes, and numerical results are rarely reliable to better
than 1 percent. However, the essentially three-dimensional nature of a liquid makes it possible to use true three-dimensional analogues where necessary and virtually guarantees homogeneity. For example, tanks containing upward of 10 cubic feet of electrolyte have been used to study the potentials induced in vehicles driven under high-voltage transmission lines in order to find the line clearances necessary at highway and railway crossings.

Equipotential boundaries are generally introduced into tank analogues by forming them out of sheet metal and placing them in the correct locations. For example, the problem of a circular wire in a long square duct may be solved simply by modeling a short section of the wire and duct surfaces and immersing them in a flat tank containing the electrolyte, as in Fig. 3.3. The two metallic surfaces will be equipotentials, provided that their conductivity is much greater than that of the electrolyte. That, however, is easily arranged. Direct measurement of potential at each point is possible with an ordinary electronic voltmeter.

A somewhat less accurate but nevertheless very attractive alternative to using liquid electrolytes is to employ paper coated with a carbon resistance layer. Such paper is commercially produced for use in one particular kind of teleprinter and is therefore readily available. It is possible to introduce equipotentials by clamping electrodes on the coating; a much more reliable technique, however, is to paint them on the paper with a conductive (e.g., silver) paint. The commercially available grade of paper is unfortunately slightly anisotropic (about 1–2 percent) and obtainable in only two conductivities. Since the conductivity of the analogue corresponds to permittivity in electrostatics, solution of problems involving several different permittivities (e.g., design of insulating bushings) cannot be carried out conveniently using conductive paper but must be done with liquid electrolytes.
Differing permittivities can be modeled in a tank by using different electrolytes separated by porous partitions. It is much easier in practice, however, to leave the electrolyte homogeneous and vary the tank depth to alter the effective (sheet) conductivity. Similarly, the myriad engineering problems involving rotational symmetry are readily modeled by simply tilting a flat tank so as to incorporate in the model an increase of peripheral distance proportional to radius.

5. Resistive Mesh Analogues

Analogue solutions of field problems are beset by both systematic and random errors, like any other experimental work. Random errors, unless of large size, are not generally serious since their effect can be reduced (at least in principle) by repeating the experiment a sufficiently large number of times. Systematic errors, however, are not reduced by repetition of measurements and in some cases may even accumulate. Contact potentials and electrode polarization in electrolytes are systematic errors that tend to limit the accuracy of the analogue devices discussed above; and so is the inhomogeneity of resistance paper.

The resistive-sheet analogue methods are essentially experimental devices to produce graphical field maps (although, as usual with graphical methods, numbers may be obtained readily). Accuracy limitations on direct graphical methods were removed when the continuous field was replaced by a discrete set of points and numerical relaxation employed. In a similar way, sheet analogues may be replaced by meshes of resistors, somewhat as depicted in Fig. 3.4a. Because the resistors are manufactured individually, there will be a certain random variation in their values; but if the resistors are carefully mixed before assembly, there will be very little systematic error in the
network. Replacement of the continuous field by discrete resistors does of course introduce an error arising from the finite mesh size. This error, however, is controllable in the sense that its magnitude may be estimated fairly well and that it may be reduced by reducing mesh size. The basic principles of resistance meshes are readily seen to be analogous to conductive-sheet and relaxation methods. For the mesh point \( O \) of Fig. 3.4b, the continuity equation of current demands that

\[
\frac{V_w - V_o}{R} + \frac{V_X - V_o}{R} + \frac{V_E - V_o}{R} + \frac{V_S - V_o}{R} = -i \tag{3-23}
\]

Rewriting,

\[
V_w + V_X + V_E + V_S - 4V_o = -iR \tag{3-24}
\]

On comparison with Equation (2-38), it is seen that (3-24) is precisely the same as the finite-difference approximation used for relaxation. In other words, the resistive mesh analogue is related to the electrolytic tank in exactly the same way as numerical relaxation to graphical field plotting. The basic advantages and disadvantages of one method as against the other are also similar.

One very useful fact is worth mentioning in connection with the mesh of finite size. Each mesh-point potential is determined by the immediately adjacent point potentials, which are in turn determined by others. Each potential value represents in fact an interpolation between the others, and random errors tend to cancel. This cancellation is surprisingly effective. Using resistors of 1 percent manufacturing tolerance, for example, the probable random error (not allowing for error due to finite mesh size) in networks containing of the order of \( 10^3 \) resistors amounts to less than one part in \( 10^4 \), i.e., 0.01 percent.

6. Experimental Study of Complex Variables

The analogue techniques described are useful not only for solving problems in electrostatics but in many other areas as well. In addition to the many physical problems that lead to Laplace's equation, electrolytic tanks are also ideally suited to studying functions of a complex variable. Let \( w(z) \) be a function, \( w = u + jv \), of the complex variable \( z = x + jy \). Its derivatives are

\[
\frac{dw}{dt} = \frac{dw}{dz} \frac{dz}{dx} = \frac{dw}{dz} \quad \text{and} \quad \frac{dw}{dy} = \frac{dw}{dz} \frac{dz}{dy} = j \frac{dw}{dz} \tag{3-25}
\]

On combining these two equations, and separating real and imaginary parts, there obtain the Cauchy–Riemann conditions

\[
\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \tag{3-26}
\]
valid throughout the region of analyticity of \( w(z) \). Second derivatives are easily formed by differentiation of the Cauchy–Riemann equations; combining them,

\[
\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad \text{and} \quad \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0 \quad (3-27)
\]

In vector notation, in the \( x-y \) plane, these may be written

\[
\text{div} \ \text{grad} \ u = 0 \quad \text{and} \quad \text{div} \ \text{grad} \ v = 0 \quad (3-28)
\]

It is seen that the Cauchy–Riemann conditions are equivalent to stating that the real and imaginary parts of the function \( w(z) \) satisfy Laplace’s equation.

As discussed earlier, the geometric meaning of Laplace’s equation, \( \text{div} \ \text{grad} \ u = 0 \), is that the curves \( u = \text{constant} \) never touch and that there must exist another similar family of curves always orthogonal to them. It follows from the above that this second family must be the curves \( v = \text{constant} \). Thus every possible distribution of equipotential and flow lines in an electrolytic tank must correspond to some analytic function \( w(z) \); and conversely, every analytic function must correspond to a particular field distribution. Corresponding to nonanalytic regions or points of the function \( w(z) \), there must be regions or points in the tank at which Laplace’s equation is not valid. It is clear that at those points, Poisson’s equation must hold; i.e., those points are sources. Electrodes placed in a tank therefore directly correspond to singularities of \( w(z) \).

A class of functions of a complex variable of particular interest to electrical engineers is the family of network transfer and immittance functions. Such functions are analytic everywhere except at certain selected points, their poles and zeros. To study network functions by means of an electrolytic tank, the bottom of the tank may be ruled off to represent the complex frequency plane, and poles and zeros introduced as current-carrying wires. Whether a given point is a pole or zero may be represented by its polarity; for example, poles may be taken as positive electrodes and zeros as negative electrodes. It is clear from the continuity equation for electric current that the number of poles and zeros must be equal, for otherwise there will be current “left over”; equally, it is immediately seen that the properties of impedance and admittance functions are exactly inverse, the sole distinction of their representations being polarity!

The singularities that often occur at infinite frequency may be worrisome at first glance but are easily included. The point infinity of a complex plane must be reached by traveling radially away from the origin in any direction, and it must be equidistant along all possible radial lines. This requirement is easily satisfied by a circular double-layer tank in which the point at infinity is placed directly beneath the origin, as in Fig. 3.5. As it is readily accessible electrically, no difficulty attaches to providing poles or zeros at
infinity. It is possible to find coordinate transformations such that the point infinity is mapped into a point in a finite plane. Such transformations, however, also change the shape of coordinate lines near the origin. One very interesting coordinate change, for example, permits logarithmic calibration of the imaginary (real frequency) axis so that frequency-response plots of networks with known pole-zero patterns may be obtained directly.

7. Complex Potentials

If the function \( w(z) \) is analytic, then by Equations (3-27) both its real and imaginary parts satisfy Laplace’s equation. Combining both, the function itself must then also satisfy

\[
\text{div} \ \text{grad} \ w = 0 \tag{3-29}
\]

This equation is formally identical to Laplace’s equation for the potential \( V \) in a stationary electric field; however, \( w \) is a complex number while \( V \) is a pure real. Because of this formal similarity, \( w \) is called the complex potential. That portion \( v \) (either the real or the imaginary part), which represents the equipotential surfaces, is called the potential function, and the other part \( u \), which gives the flux lines, is termed the stream function. These terms originally arose from hydrodynamics, but they are used in electrical engineering as well.

The notion of complex potential is useful in many problems. For example, the potential solution found for the coaxial cable,

\[
V = A \log \frac{r}{r_0} \tag{3-30}
\]

may be generalized to yield the flux lines without any geometrical construction. The potential function \( V \), which is known, is augmented by the stream function \( U \) to give the complex potential; and the radial distance \( r \) is augmented by the angle \( \theta \):

\[
V + jU = A \log \frac{z}{z_0} \left(\frac{e^{j\theta}}{z_0}\right) \tag{3-31}
\]

where \( z = re^{j\theta} \), and \( z_0 = r_0 e^{j\phi} \). Expanding in real and imaginary parts,
The lines of constant stream function (flux lines) are thus directly calculated. To lay out lines so that any two successive ones demarcate equal quantities of flux, it is only necessary to choose constants \( U_1, U_2, U_3, \ldots \) at equal increments: \( U_{n+1} - U_n = \delta U \), where \( \delta U \) is the same for any \( n \). In the particular problem at hand, inspection shows that equal increments in \( U \) correspond to equal increments in \( \theta \).

The above manner of generalization is not arbitrary, and while it may have been simple enough in this special case, more involved procedures must be used in general. In fact, the Cauchy–Riemann differential equations state quite clearly that the stream function must satisfy Laplace’s equation in its own right, as well as provide the resulting complex potential function with its required analytic character. Such functions are in general unique and are known as the conjugate functions to the given potential functions \( V \).

The conjugate function to any given potential function \( V(x, y) \) can always be found by solving the Cauchy–Riemann differential equations (3-26), which determine \( U(x, y) \) uniquely (except for an additive constant, like most other differential equation solutions). The mathematical procedure of doing so, however, often becomes quite involved. In those cases where the potential function itself also needs to be determined, it is generally best to solve for the complex potential function in the first place, for the work involved in doing so is not significantly greater than that of finding \( V \) alone.

As an example of this process, the solution for potentials in the neighborhood of a two-wire transmission line, Equation (2-15), will be rederived with the stream function included in the complex potential solution.

From Equations (3-31) and (3-32), it may be seen that the complex potential in the neighborhood of a single charged wire of small cross section is given, analogously with Equation (2-4), by

\[
V + jU = \frac{V_0}{\log(a/b)} \log \frac{z}{z_0} \quad (3-33)
\]

The value \( V_0 \) is seen to have zero imaginary part always. The field of two wires, one at a potential \( V_0 \), the other at \(-V_0\), may be found by superposing the form (3-33) in the manner earlier employed for real potentials. There results

\[
V + jU = \frac{V_0}{\log(a/b)} \log \frac{z - b/2}{z + b/2} \quad (3-34)
\]

It should be noted that, just as in the case of the real potential solution, the quotient \( a/b \) has only symbolic meaning since the solution is strictly valid only for vanishingly small values of \( a \).

Equation (3-34) may be written with the distances in polar form, so as to give direct geometric meaning to all variables. Using once again the symbolism introduced in Chapter 2,
Sec. 7 MAPPING OF STEADY CURRENTS AND FIELDS

\[ V + jU = \frac{V_0}{\log(a/b)} \log \left( \frac{r_1 \exp(j\theta_1)}{r_2 \exp(j\theta_2)} \right) \]  
\hspace{1cm} (3-35)

or, separating reals and imaginaries,

\[ V + jU = \frac{V_0}{\log(a/b)} \left[ \log \left( \frac{r_1}{r_2} \right) + j(\theta_1 - \theta_2) \right] \]  
\hspace{1cm} (3-36)

It is immediately seen that the real parts return the already known solution of Equation (2-15). On the other hand, the imaginary parts yield the stream function

\[ U = \frac{V_0}{\log(a/b)} (\theta_1 - \theta_2) \]  
\hspace{1cm} (3-37)

From the general discussion in Section 6, it is clear that there is no particular mathematical reason to associate either the real or the imaginary part of the complex potential function with equipotential lines or flux lines; either could be chosen for either purpose, depending on the boundary conditions of the problem to be solved. It follows that the field map will be made up of curvilinear squares, if increments of \( V \) and \( U \) are chosen equally. But from Fig. 2.12 it was previously deduced that any map made up of curvilinear squares (as contrasted to any other rectangular figures) must have all its equipotentials equally spaced. Similarly, its flux lines must mark out equal amounts of intervening flux. Consequently, flux lines marking out equal amounts of flux will be obtained in the present problem if successive values of \( U \) are chosen with equal increments: \( U_{n+1} - U_n = U_n - U_{n-1} = \delta U \), for any \( n \). Each of the chosen values of \( U \) will correspond to some specific value of \( \theta = \theta_1 - \theta_2 \).

The shape and location of these flux lines may be deduced directly from Equation (3-37). From Fig. 2.3, it is easily seen that

\[ \tan \theta_1 = \frac{y}{x - b/2} \quad \text{and} \quad \tan \theta_2 = \frac{y}{x + b/2} \]  
\hspace{1cm} (3-38)

In all cases, flux lines correspond to loci of constant \( \theta = \theta_1 - \theta_2 \) and hence to constant values of \( \tan \theta \). The well-known trigonometric identity

\[ \tan \theta = \frac{\tan \theta_1 - \tan \theta_2}{1 + \tan \theta_1 \tan \theta_2} \]  
\hspace{1cm} (3-39)

becomes on substitution of (3-38)

\[ \tan \theta = \frac{by}{x^2 - b^2/4 + y^2} \]  
\hspace{1cm} (3-40)

On clearing fractions and completing the square, this equation may be rewritten as

\[ x^2 + \left( y - \frac{b}{2} \cot \theta \right)^2 = \left( \frac{1 + \tan \theta}{\tan \theta - \frac{b}{2}} \right)^2 \]  
\hspace{1cm} (3-41)

In this form, the flux lines are readily recognized to form a family of circles,
all centered on the $y$-axis above and below the conductor center line. Several of these circles are shown in Fig. 2.4 also.

8. Conformal Mapping

Generalization and simplification of existing solutions, although useful, are hardly indicative of the power of complex potentials. A much more important aspect emerges when the definition of a complex function is regarded as a coordinate transformation between the $z$-plane and a $w$-plane. For example, Equation (3-31) transforms straight lines parallel to the $u$-axis (in the $w$-plane) into circles in the $z$-plane, as shown in Fig. 3.6, and straight lines parallel to the $v$-axis into straight lines through the origin $z = 0$.

![Diagram showing conformal mapping](image)

Fig. 3.6

If the function $w(z)$ of any general transformation is analytic, the lines $v = \text{constant}$ and $u = \text{constant}$ intersect at right angles in both $w$- and $z$-planes. Any small rectangular figure in the $z$-plane will therefore always map into a rectangular figure (although in general a curvilinear rectangle) in the $w$-plane; for this reason, the transformation is said to be conformal.

A great many two-dimensional field problems can be solved by using some convenient conformal transformation to turn a new problem into one that has already been solved. The key to such solutions lies in the special properties of analytic functions. Suppose that a problem is specified by prescribed equipotentials and source distributions in the $z$-plane; the complex potential function $P(z)$ needs to be found. Suppose also that there exists an analytic function $w = w(z)$ such as to transform the equipotentials and sources of the $z$-plane into a known problem in the $w$-plane. If the complex potential in the $w$-plane is given by $\phi(w)$, the desired solution $P(z)$ is simply
$P(z) = \varphi[w(z)]$, for an analytic function of an analytic function is in turn always analytic.

A simple example will illustrate the principle involved. Let it be desired to determine the potential distribution in the neighborhood of a screen made of many long, parallel, equispaced, fine wires. Such screens are used, for example, in electrostatic precipitation apparatus to charge dust particles carried by a passing air stream. The wires will be taken to have a radius $r_0$ and a potential $V_0$ and to be spaced a distance $a$ apart, as in Fig. 3.7. It may seem at first glance that the problem may be solved by superposing many terms of the type of Equation (2-4). It can, in fact, but the solution hinges on recognizing a convergent infinite series. Solution by conformal transformation is much easier. It is recognized on inspection of Fig. 3.7 that the potential function sought is analytic everywhere except at the points $z = na + j0$, where $n$ is any integer. It must therefore be periodic in $x$, with the period equal to wire spacing $a$. The transformation required would map the periodically spaced point singularities into a point singularity in the $w$-plane; the latter corresponds simply to the solution (3-33). One of the wires of Fig. 3.7 is placed at $x = 0$, so requiring a singularity at the origin. A function that has the required periodic character and maps the origin of one plane into the origin of the other is

$$w = \sin \frac{\pi z}{a} \quad (3-42)$$

As indicated in Fig. 3.8, this transformation makes the points $z = na + j0$ correspond to the origin $w = 0$; in fact, each vertical strip of width $a$ is mapped into the whole $w$-plane. In the $w$-plane the potential solution is, by (3-33),

$$V(w) + jU(w) = K \log w + K' \quad (3-43)$$

Combining (3-42) and (3-43), the solution in the $z$-plane must be

$$V(z) + jU(z) = K \log \sin \frac{\pi z}{a} + K' \quad (3-44)$$

where the constants remain to be determined. They are easily found by separating the real from the imaginary parts and imposing the requirement of

![Fig. 3.7](image-url)
Fig. 3.8

A given potential on the wires:

\[ V(r_0, 0) = V_0 = K \log \sin \frac{\pi r_0}{a} + K' \]  \hspace{1cm} (3-45)

The number \( K' \) physically represents an arbitrary reference potential and may as well be taken as zero. There results

\[ K = \frac{V_0}{\log \sin (\pi r_0/a)} \]  \hspace{1cm} (3-46)

completing the solution.

It should be observed that this method relies on analyticity of the transformation function; (3-44) is a valid solution only because the transformation (3-42) is analytic nearly everywhere. This necessity is physically evident when it is considered that singularities represent Poissonian field region, i.e., charges. A transformation that is not analytic everywhere introduces new singularities and hence new charges where the problem does not have any.
The trick in solving problems by conformal transformation is evidently to choose the right transformation function. Unfortunately, no general method for finding the correct function exists so that the choice must frequently be made on a trial-and-error basis. Familiarity with a number of different transformations is of course very helpful. There exist books that catalogue transformations by their salient characteristics and give graphs of typical potential and stream function contours; these may be regarded as a counterpart in field theory of integral tables or Laplace transform tables.

9. Field of a Long Charged Strip

As a second example of conformal transformation, let the inverse sine function

$$w(z) = K \sin^{-1} z \tag{3-47}$$

be examined. The first question to be answered must of necessity be, Is the function analytic everywhere? By forming derivatives, it is easily shown that the Cauchy–Riemann conditions are satisfied at all finite points, so that the function is indeed analytic and the transformation furnished by (3-47) is conformal. To see what kind of problems might be solvable by using it, corresponding equipotentials and flux lines need to be plotted. This is most easily accomplished by writing

$$z = c \sin w \tag{3-48}$$

or, stating the real and imaginary parts of $z$ and $w$ explicitly,

$$x + jy = c \sin u \cosh v + j c \cos u \sinh v \tag{3-49}$$

or

$$\frac{x}{c \cosh v} = \sin u \quad \frac{y}{c \sinh v} = \cos u \tag{3-50}$$

The latter forms permit easy elimination of either $u$ or $v$ and hence facilitate finding lines of constant $u$ or $v$. Eliminating $u$ and $v$ in turn, the equations are obtained

$$\frac{x^2}{c^2 \cosh^2 v} + \frac{y^2}{c^2 \sinh^2 v} = 1 \tag{3-51}$$

$$\frac{x^2}{c^2 \sin^2 u} - \frac{y^2}{c^2 \cos^2 u} = 1 \tag{3-52}$$

These describe families of ellipses and hyperbolas in the $z$-plane, respectively, with $u$ and $v$ as parameters. That is to say, to any given constant value of $v$, there corresponds an ellipse in the $z$-plane, and to given values of $u$, there correspond hyperbolas. Some of these curves are shown in Fig. 3.9a. It is evident that this transformation may be used, for example, to map the field of a cable made of elliptic conductors into the field of a pair of infinite parallel plates.
MAPPING OF STEADY CURRENTS AND FIELDS

Chap. 3

An interesting problem of some practical importance is that of surface charge distribution on a flat, thin conductor. Such conductors are encountered in many kinds of printed circuits. A flat conductor may be viewed as the limiting ellipse obtained for \( v = 0 \), with semiaxes \( c \) and zero. As seen from Fig. 3.9b, \( 2c \) then represents the conductor width. The transformation effectively opens and stretches out any ellipse, as indicated by the corresponding points \( A-B-B'-D \) in the sketch. In order to obtain a field map, it is only necessary to identify \( v \) as the potential function \( V \) and \( u \) as the stream function \( U \).

While the charge distribution itself is not given explicitly by the potential and stream functions, it is easy to find the flux density near the surface and
infer the charge distribution from Equation (1-75). Since equal increments of the stream function define equal increments of flux, the charge distribution is necessarily given by the rate of variation of the stream function in a tangential direction along the conductor surface. To say the same thing in another way, there cannot be any electric field inside a conductor at equilibrium, so that the surface charge and flux density at the surface are equal:

\[ \sigma = D_{\text{surface}} \]  

This may be written

\[ \sigma = -e \text{ grad } V \]

\[ = -e \mathbf{1}_n \frac{\partial V}{\partial n} \]  

where the letter \( n \) denotes distance measured normal to the conductor surface. However, outside the conductor surface the potential function must be analytic, so that

\[ \frac{\partial V}{\partial n} = \frac{\partial U}{\partial t} \]

where \( t \) denotes distance measured in a tangential direction. Thus, the surface charge distribution may be found from the stream function by

\[ \sigma = -\mathbf{1}_n e \frac{\partial U}{\partial t} \]  

In the particular case at hand, the conductor surfaces lie in the \( x \)-axis, so that the \( x \)-direction is in fact the tangential direction. Combining (3-56) and the first equation of (3-50), the charge density is given by

\[ \sigma = -\mathbf{1}_n e \frac{\partial}{\partial x} \left( K \sin^{-1} \frac{x}{c} \right) \]

\[ = -\mathbf{1}_n \frac{eK}{\sqrt{c^2 - x^2}} \]

where \( K \) is a multiplicative constant associated with the stream function and not yet determined. It can be found from the total charge on the strip. To find the total charge, integrate over the width of the strip (top \textit{and} bottom, since only half the surface charge will reside on one side). There is obtained

\[ q = 2 \int_{-c}^{+c} \sigma \, dx \]

\[ = K \pi e \]  

so that the surface charge is finally

\[ \sigma = \frac{q}{2\pi \sqrt{c^2 - x^2}} \]

\[ = \frac{q}{2\pi \sqrt{c^2 - x^2}} \]

coulombs per square meter of conductor surface area. It is hardly surprising to note that this function has a finite value everywhere, except at the edges of the strip, which have been assumed mathematically sharp!
10. Current Flow around a Hole

The third and final simple example of conformal mapping concerns the transformation

\[ w(z) = \frac{1}{2} \left( \frac{z}{R} + \frac{R}{z} \right) \]  \hspace{1cm} (3-60)

Inspection immediately reveals that this function is analytic everywhere but at the origin of the z-plane, and may therefore be employed in conformal mapping problems. A special property of this transformation, lending it special interest, is that the semicircle \( |z| = R \) in the upper half plane is transformed into a section of the real axis of the w-plane. This may be discovered by substituting \( z = Re^{j\theta} \):

\[ w(Re^{j\theta}) = \frac{1}{2} (e^{j\theta} + e^{-j\theta}) = \cos \theta + j0 \]  \hspace{1cm} (3-61)

All real values of \( z \) outside this circle are mapped into the \( u \)-axis also, so that the \( u \)-axis in its entirety corresponds to the contour in the z-plane shown in Fig. 3.10. That is to say, the upper half w-plane corresponds to the upper half z-plane less a round hole centered on the origin. It should be noted very carefully that this does not imply that the lower half w-plane corresponds to the lower half z-plane plus a semicircle; the transformation of upper half planes does not necessarily furnish any information about the lower half planes. While this may be mathematically surprising, it is physically all but self-evident, as readily seen by the following considerations. The object of transforming the z-plane into the w-plane is to transform a new problem into an old one, e.g., an odd-shaped pair of equipotential lines into a straight parallel pair, for which the field solution is known (the infinite parallel-plate capacitor!). The field between electrodes, however, provides no information about the fields elsewhere.

Contours traced in the z-plane by lines \( u = \text{constant} \) or \( v = \text{constant} \) may be found by solving (3-60) for \( z \) explicitly and obtaining

\[ z = Rw + R\sqrt{w^2 - 1} \]  \hspace{1cm} (3-62)

The resulting field map is shown in Fig. 3.11.
A practical application of this transformation is in determining the current flow lines around a bolt-hole in a heavy-current conductor. The hole may be considered as being represented by the circle \(|z| = R\), so that the lines of constant stream function in Fig. 3.11 give current flow lines (lines of the vector \(\mathbf{J}\)) directly, while the lines \(u = \text{constant}\) represent equipotentials. It is sometimes desired to find the hot-spot temperature in the conductor since this forms one of the performance limitations. In order to determine it, the power density must be determined and heat-flow calculations made to determine temperature rise. The power density is, according to Equation (3-15),

\[
\frac{dW}{dU} = \mathbf{J} \cdot \mathbf{E} = gE^2
\]  

(3-63)

which is readily calculated if \(\mathbf{E}\) is known. This may be found from the transformation equation that led to Fig. 3.11. In the present problem, the real part \(u\) of the complex potential \(w(z)\) has been identified as corresponding to the electric potential. Since \(u\) is a dimensionless quantity, the potential must be taken to be \(V_0u\), where \(V_0\) is the actual potential difference between two successive lines of constant \(u\), i.e., between \(u = U\) and \(u = U - 1\). The electric field is then given by

\[
\mathbf{E} = -V_0 \text{grad} \ u
\]

(3-64)

In order to find the gradient, derivatives of \(u\) in the \(x\)- and \(y\)-directions will now be found. Since \(u\) is the real part of \(w\),
\[
\frac{\partial u}{\partial x} = \frac{1}{2} \frac{\partial}{\partial x} \left( \frac{x}{R} + \frac{xR}{x^2 + y^2} \right)
= \frac{1}{2} \left( \frac{1}{R} - \frac{x^2 - y^2}{(x^2 + y^2)^2} \right) \tag{3-65}
\]

and similarly
\[
\frac{\partial u}{\partial y} = -\frac{1}{2} \left( \frac{2xyR}{(x^2 + y^2)^2} \right) \tag{3-66}
\]

To find the hotspot, i.e., the point of maximum power density, it is necessary to locate the largest magnitude of \( E \). While this could form a lengthy exercise in partial differentiation, it is easiest to determine it by examination of Fig. 3.11, from which it is at once evident that the highest field occurs at the side of the hole, \( z = 0 + jR \). The field components there are, by (3-65) and (3-66),
\[
E_x = \frac{V_0}{R} \quad \text{and} \quad E_y = 0 \tag{3-67}
\]

so that the power density is given by
\[
\frac{dW}{dU} = gE^2 = g\frac{V_0^2}{R^2} \tag{3-68}
\]

It is interesting to note that this result may in some cases place an electrical restriction on the minimum bolt size to be used in fastening conductors, for a small bolt-hole will lead to greater power densities than a large one. Fortunately, the metals generally employed in heavy current work have sufficiently high thermal conductivities to make this a very gentle restriction. This particular transformation, in fact, is of greater interest in connection with magnetic fields, where material properties are not always so convenient.

11. The Schwarz–Christoffel Transformation

For the special class of problems with polygonal boundaries, a general method of finding a transformation does exist. Mathematical development of the equations of transformation may, however, become quite complicated, limiting the analytic usefulness of the method to boundaries composed of at most five or six straight-line segments. Fortunately, a great many engineering problems fall within this restriction. This method, independently discovered by Schwarz and Christoffel, is capable of approximate extension to include polygonal shapes with somewhat rounded as well as mathematically sharp corners, and a large number of applications may be found in engineering literature.

Any polygon is completely specified if all its vertex angles and vertex locations are given. The transformation method consists of inserting into
a general differential equation (called the Schwarz–Christoffel equation) the known vertex angles of the polygonal boundary and then solving the differential equation; information about the size and location of the polygon is inserted in the constants of integration. As the solution of the differential equation, there results an expression that transforms the interior of the polygon into a half plane. Repeated application of the transformation will then map the boundaries into some other polygon, for which solutions are known, thereby establishing a correspondence between a known solution and the problem to be solved.

To develop the method, let there exist two complex variables \( z = x + jy \) and \( t = p + jq \), connected by a functional relationship \( z = z(t) \). The identity

\[
\frac{dz}{dt} = \left( \frac{dz}{dt} \right) dt
\]

may always be written in the form

\[
dz = \zeta e^{i\varphi} dt
\]

since the derivative in (3-69) is a complex number that may be put in polar form. Equation (3-70) describes how a line element \( dt \) is mapped into the \( z \)-plane: it is rotated through an angle \( \varphi \) and magnified by a factor \( \zeta \). As a simple example, let

\[
\frac{dz}{dt} = K(t - a)^{\frac{1}{2}}
\]

where \( K \) and \( a \) are real constants. This transformation rotates a line element \( dt \) through the angle

\[
\varphi = \arg \left( \frac{dz}{dt} \right) = \frac{1}{2} \arg (t - a)
\]

\[
\varphi = \frac{1}{2} \arctan \frac{q}{p - a}
\]

(3-72)

The angle of rotation and the magnification factor are not the same throughout the range of values of \( t \). It is particularly interesting to note how the upper half \( t \)-plane, including the \( p \)-axis, is transformed into the \( z \)-plane. For all \( p > a, \varphi = 0 \). However, for all \( p < a, \varphi = \pi/2 \); as the \( p \)-axis is traversed, the direction of the corresponding line in the \( z \)-plane remains unaltered except for a sudden turn at \( p = a \). Other lines parallel to the \( p \)-axis transform in a somewhat similar manner but do not change direction so suddenly, as may be seen from (3-72) and Fig. 3.12. In fact, lines parallel to the \( p \)-axis correspond to rectangular hyperbolas in the \( z \)-plane.

Rotation through \( \pi/2 \) here arises from the one-half power in Equation (3-71); any other angle of rotation might have been produced had the exponent been chosen to have some other value. The argument of \( (t - a) \) is altered by \( \pi \) as \( t = p + jq \) passes through \( a \), and raising \( (t - a) \) to the \( nth \)
power causes the corresponding line in the $z$-plane to change direction by $n\pi$. More generally, setting

$$\frac{dz}{dt} = K(t - a)^{\alpha/\pi - 1}$$

maps the $p$-axis into two straight-line segments in the $z$-plane, one corresponding to all points $p > a$, the other to $p < a$. The location and scale of these lines are determined by $K$ and $a$, and the angle enclosed between them by $\alpha$. By performing successive rotations at different points along the way, the $p$-axis may be mapped into a sequence of any desired number of straight-line segments. It is only necessary to set

$$\frac{dz}{dt} = K(t - a_1)^{\alpha_1/\pi - 1} (t - a_2)^{\alpha_2/\pi - 1} \cdots (t - a_n)^{\alpha_n/\pi - 1}$$

(3-74)

to establish correspondence between a set of $n + 1$ straight lines in the $z$-plane and the $p$-axis in the $t$-plane. If the line segments in the $z$-plane are regarded as a polygon (not necessarily closed), the interior of the polygon corresponds to the upper half $t$-plane and the numbers $\alpha$ give the interior angles of the polygon. It might be noted that this Schwarz–Christoffel differential equation does not require the change in direction, but rather the interior angle, to be specified. A frequent source of error is eliminated by this means, for interior angles are positive at all finite points and no confusion about sign can arise. It is clearly possible to write the differential equation for any polygonal boundary, however complicated; but integration to find the explicit relationship between $z$ and $t$ may be quite difficult.

In practical problems, boundaries frequently do not close but remain separate. The Schwarz–Christoffel transformation may nevertheless be applied by imagining the open boundaries to form a polygon with at least one vertex infinitely far away; for example, the lines of Fig. 3.12 are often referred to as the sides of a polygon with two vertices (one at the origin, one at infinity). Conversely, it is often desirable to map one of the $z$-plane vertices
into \( t = \infty \). In that case, Equation (3-74) must include one factor \((t - \infty)\), and correspondingly \( K = 0 \) so as to have the expression for \( dz \) remain finite. The product \( K(t - \infty) \) must be a finite constant. As a result, the transformation equation (3-74) may be applied directly, but with only \( n - 1 \) factors, omitting the vertex at infinity. Lest it appear as if information were lost in the process, it ought to be recalled that the sum of interior angles of a polygon with \( n \) vertices is \( \pi(n - 2) \); thus, the last angle is redundant anyway.

By means of the Schwarz-Christoffel equation, the boundaries of the problem to be solved (in general, these might be equipotential surfaces at different potentials) are transformed into the \( p \)-axis. A solution is thereby not yet obtained. It will next be necessary to transform the \( p \)-axis into another polygonal boundary for which the solution of Laplace’s equation is known. For problems with two given equipotential surfaces, this known solution is almost invariably selected to be the field between two infinite parallel plates, for that is the simplest possible two-dimensional polygonal problem.

12. Field at the Edge of a Parallel-Plate Capacitor

The procedure followed in finding suitable transformations by the Schwarz-Christoffel method is best explained by an illustrative example. A reasonably simple problem of considerable practical importance is the field near the edge of a thin conductive plate parallel to an infinite conductive plane. This problem, along with related ones derivable by means of images and superposition, forms an approximation not only to numerous kinds of capacitors, but also to magnetic fringing fields in electric machines and other electromagnetic apparatus. As indicated in Fig. 3.13a, the boundaries may be thought to form a polygon with three vertices (i.e., a triangle), two of the vertices being located at \( z = \infty \). The interior angles at vertices \( B \) and \( C \) are easily seen to be \( 2\pi \) and 0, respectively; the vertex angle at \( A \) must therefore be \(-\pi\) (negative interior angles are possible at infinity, but not at finite points).

To avoid the difficulty of negative angles, vertex \( A \) will be mapped into the point infinity in the \( t \)-plane as well, so that the negative angle is never needed explicitly. The second transformation, between the \( t \)- and \( w \)-planes, involves little difficulty since the angles (with the choice of points \( B \) and \( C \) shown) are 0, \( \pi \), 0 in the \( w \)-plane. Both transformations are readily set up by using (3-74). Since the vertex at \( A \) is at the point infinity, no term corresponding to it is required in the Schwarz-Christoffel equation, and two factors suffice. The first transformation equation is

\[
\frac{dz}{dt} = K_z(t + 1)^2 - 1(t - 0)^{s-1}
\]

\[
= K_z \frac{t + 1}{t}
\]

(3-75)
whence

\[ z = K_z(t + \log t) + K' \]

In a similar manner, the second transformation equation is

\[ \frac{dw}{dt} = K_w(t + 1)^{1-1} (t - 0)^{0-1} \]

\[ = \frac{K_w}{t} \]

so that

\[ w = K_w \log t + K'_w \]

The four constants \( K_z, K_w, K'_z, K'_w \) are as yet undetermined, but they are not arbitrary. This may be seen from geometrical reasoning. Any polygon, open or closed, is completely specified if its internal angles and vertices are given. Since the vertices \( B \) and \( C \) are specified in Fig. 3.13 and the angles have been included in the Schwarz–Christoffel equation, the four constants have been included implicitly and may be found by substituting the vertex points into the transformation equations.
A difficulty which arises in this problem and is quite common in applications of the Schwarz–Christoffel transformation is that both sides of Equations (3-76) and (3-78) become infinite when point $C$ is substituted, preventing determination of the constants. This occurs simply because the infinitely far points $C$ of the $z$- and $w$-planes have been transformed into a finite point in the $t$-plane. However, the constants may be determined by a simple artifice. The change in $z$ between two points $a$ and $b$, one on either side of the vertex, may be written as

$$\Delta z = \int_a^b dz$$  \hspace{1cm} (3-79)

or, by Equation (3-75), as

$$\Delta z = \int_{a'}^{b'} K_t \frac{t + 1}{t} dt$$  \hspace{1cm} (3-80)

where $a'$ and $b'$ are the corresponding points in the $t$-plane. If the two $z$-plane points $a$ and $b$ are selected as shown in Fig. 3.14a, the integral of (3-80) cannot be evaluated directly since its integrand is not analytic at $t = 0$. However, the path of integration may be taken entirely within the region of analyticity by constructing a small circular path of radius $\zeta$ around the singularity as shown and then making $\zeta$ vanishingly small. Equation (3-80) may be rewritten with the integral in polar form, $t = \zeta e^{j\phi}$:

$$\Delta z = \lim_{\zeta \to 0} K_z \int_\pi^0 \frac{\zeta e^{j\phi} + 1}{\zeta e^{j\phi}} j\zeta e^{j\phi} d\phi$$  \hspace{1cm} (3-81)

which is easily evaluated. From Fig. 3.13, the corresponding change in the $z$-plane is simply $\Delta z = -jh$, so that $K_z$ may be evaluated:

$$-jh = jK_z \lim_{\zeta \to 0} \int_\pi^0 (\zeta e^{j\phi} + 1)d\phi$$

$$K_z = \frac{h}{\pi}$$  \hspace{1cm} (3-82)

The same artifice, with the same substitution $t = \zeta e^{j\phi}$, may be employed for the $w$–$t$ transformation, yielding upon integration around the point $t = 0$
\[ K_w = j \frac{U_0}{\pi} \]  

(3-83)

The remaining two constants \( K' \) and \( K'' \) may be evaluated by substitution of the coordinates of point \( B \) (Fig. 3.13). This point is finite in all three planes, and no difficulty arises. There result the two complete transformation equations

\[ z = \frac{h}{\pi} (1 + t + \log t) \]  

(3-84)

and

\[ w = j \frac{U_0}{\pi} \log t + u_0 \]  

(3-85)

These may be combined, eliminating the variable \( t \) altogether, to give

\[ z = \frac{h}{\pi} \left[ 1 + j \pi \frac{w - u_0}{u_0} + \exp \left( j \pi \frac{w - u_0}{u_0} \right) \right] \]  

(3-86)

as the transformation that maps the parallel-line field of the \( w \)-plane into the field required by the boundaries in the \( z \)-plane, as shown in Fig. 3.15. In principle, the Schwarz–Christoffel method may be used to solve any problem involving polygonal boundaries. Integration of Equation (3-74), and determination of constants, becomes very difficult if the polygon has many corners, or if the corner angles are not integer multiples of a right angle. Nevertheless, this technique has long provided one of the most powerful tools for solution of difficult field problems. Its great importance for the future lies in the possibility of integrating the Schwarz–Christoffel equation numerically, thus removing the restrictions due to difficulties in integration and permitting routine solution of virtually any problem with polygonal boundaries.

This field mapping technique, like several others, is perhaps of even
greater interest in connection with magnetic fields than purely electrostatic ones. It will not be pursued further at this point but will be examined again in Chapter 6, treating both magnetic and electric problems concurrently.

**READINGS**

The subject of current flow in conductors is primarily of interest from the solid-state physics point of view and is generally only very briefly treated by books on field theory. Carter (1) gives a clear explanation, as does Arzeliès (2), though both are quite short. A more elementary treatment may be found in Elliott (3).

Analogue methods of solving field problems have a long history and a large literature. A very comprehensive work on this subject, well written and exceptionally easy to read, is by Karplus (4). This book will also furnish many further literature references. An interesting description of an actual working apparatus, with applications, is given by Diggle and Hartill (5). A survey paper of the mesh analogue technique, with further bibliographic references, is that by Liebmann (6). The paper by Okoshi (7) is a recent report of application to practical problems.

Investigation of functions of complex variables in an electrolytic tank is described, and applied to circuit theory problems, by Boothroyd, Cherry, and Makar (8). Various applications of electrolytic tanks, as well as other conductive analogues and numerical relaxation techniques, may be found in the recent book by Vitkovitch (9).

A great many sources may be consulted for information on conformal transformations. The sections in Bewley (10) and Ramo, Whinnery, and Van Duzer (11) are brief but succinct. A very clear physical explanation is furnished of the Schwarz–Christoffel transformation in the little book by Walker (12). More detailed mathematical treatments are those in Binns and Lawrenson (13) and Gibbs (14). Many applications involving relatively simple mathematics are listed in the monograph by Rothe, Ollendorff, and Pohlhausen (15).

A catalogue of many transformations, with graphical representations and very brief descriptions, is that by Kober (16).


**PROBLEMS**

3.1 Show that there always exists a vector such that its curl is given by the current density $\mathbf{J} + \mathbf{D}$. Is this vector unique?

3.2 Show clearly that the resistance of a resistor made of linear isotropic material is not dependent on the current or field.

3.3 By graphical relaxation, find the equipotential and stream lines of current flowing around a right-angle bend in a square conductor. Hence calculate the equivalent resistance of a right-angle bend.

3.4 Equation (2–38) and again Equation (3–24) give an approximation to the Poisson equation, using finite differences. By making a Taylor series expansion of the potential function $V$, find the magnitude of the largest neglected term in these equations. Hence obtain an accuracy estimate for the finite-difference approximations. (*Hint: This is the classical error-estimate technique, and it may be found in texts on numerical analysis.*)

3.5 Prove that a surface charge exists at the interface between two conductive media if a current flows across the interface.
3.6 Show that the function \( w = \log \tan z \) may be used to represent the wire and planes of Problem 2.12, assuming the wire to be thin. Use this result to solve the latter problem.

3.7 Show that the function \( w = \log (z^n - R^n) \) may be used to solve the problem of \( n \) parallel wires symmetrically arranged on the periphery of a circle of radius \( R \). Show that this reduces to the known result for a two-wire situation.

3.8 Find the capacitance of a parallel-plate capacitor made of thin circular plates 10 cm in diameter separated by an air space of 2 mm.

3.9 By superposing thin-wire complex potentials, find the equipotentials and flux lines near a bundle conductor composed of four subconductors set in a square. Evaluate the potential and stream functions at a sufficient number of points to plot this field reasonably accurately.

3.10 The transformation \( u = x, v = py \), where \( p \) is a real constant, maps concentric circles into ellipses. Could it be used as an alternative to Equation (3-47), to solve for the electric field in a cable of elliptic cross section?

3.11 Draw a clear circuit diagram of the resistive network required for solution of Problem 2.9 by analogue means.

3.12 Draw a clear circuit diagram, giving all necessary values of components, to solve Problem 2.11 by means of a resistive analogue network.

3.13 Prove that equipotential and flux lines of an electrostatic fields problem may be simulated in a resistive mesh analogue by rows of short-circuited and open-circuited resistors, respectively.

3.14 Sketch lines of constant \( u \) and constant \( v \) in the \( z \)-plane for the transformation
\[
w(z) = \frac{1}{2} (z/R + a^2 R/z)
\]

3.15 Show that the Cauchy–Riemann conditions require flux and equipotential lines to cross at right angles.

3.16 Find the capacitance of a cable consisting of a flat strip of width 2 cm inside an elliptic sheath with major semiaxis 2 cm and focal separation 2 cm.

3.17 Write the Schwarz–Christoffel differential equation for the transformation of the shapes shown in Fig. 3.16 into the upper half plane. Can the constant \( K \) be determined in each case without actually performing the required integration?
Electrodynamics and Maxwell’s Equations

The development of Coulomb’s law in Chapter 1 was based on the assumption that all charges under consideration were fixed in space. In order to deal with the forces between moving charges, a generalized form of Coulomb’s law, in which the charges are permitted to move, becomes necessary. Such a form is developed in this chapter. There appear new forces, termed magnetic forces, in addition to the expected electrostatic ones. Detailed consideration of charge motions in all cases is rendered unnecessary by defining a new field, the magnetic field, in terms of these forces. The interrelationship between the electrostatic and magnetic fields resulting from any given set of moving charges is concisely expressed in four differential equations first formulated in their modern form by Maxwell.

1. Flux and Field of a Moving Charge

The flux density of a single moving charge will first be calculated as seen by a stationary observer, starting from the continuity principle in much the same
way as in electrostatics. The law of conservation of charge, a corollary of the continuity principle, implies that the charge of a moving particle must be independent of the velocity of motion; for were a charge to grow or shrink with velocity while its corresponding opposite charge is held immobile, the neutrality of the universe would no longer hold true.

Independence of velocity also implies invariance with respect to changes of inertial coordinate frames. Let a charge be measured in two reference frames, \( F \) and \( F' \), as being \( q \) and \( q' \) coulombs, respectively. If the relative velocity of the two frames is \( v \), either (or neither) of the frames may be made the rest frame by requiring the charge to move at an appropriate velocity. Now \( q \) and \( q' \) are independent of velocity and refer to the same charged object. It remains to conclude that \( q = q' \), proving the assertion that a charge is the same in all inertial frames.

The continuity postulate asserts that charge and flux are always equal (or, in some systems of units, related by an invariant proportionality constant). The invariant character of charge must then be accompanied by invariance of flux with respect to changes of coordinate frames, or with respect to motion within a given frame.

In contrast to charge and flux, flux density is a vector quantity and not invariant. That its value depends on velocity is clear from its definition as the amount of flux per unit area; flux is invariant but area is not. The transformation rule for flux density may be deduced from Gauss's law. Over some portion of a Gaussian surface surrounding a moving charge, as in Fig. 4.1, the flux \( d\psi' \) through an elemental area \( dS' \) is given by

\[
d\psi' = D' \cdot dS'
\]

in the rest frame \( F' \) of the charge. This surface is assumed to be stationary with respect to the charge. In another frame \( F \), through which the charge and

![Fig. 4.1](image-url)
surface move with velocity \( \mathbf{v} = iv_z \), the corresponding flux is

\[
d\psi = \mathbf{D} \cdot d\mathbf{S}
\]  \hspace{1cm} (4-2)

or

\[
d\psi = D_x dS_x + D_y dS_y + D_z dS_z
\]  \hspace{1cm} (4-3)

The magnitudes of the three area elements may be written as

\[
dS_x = dy \, dz \quad dS_y = dx \, dz \quad dS_z = dx \, dy
\]  \hspace{1cm} (4-4)

Because the primed frame \( F' \) is moving with a velocity \( \mathbf{v} \) with respect to the unprimed frame \( F \), any length measured in the \( x \)-direction in frame \( F \) must contract by an appropriate Lorentz factor \( \gamma \). Distances measured in the \( y \)- and \( z \)-directions, however, are unaffected. Thus

\[
\begin{bmatrix}
    dx' \\
    dy' \\
    dz'
\end{bmatrix}
= \begin{bmatrix}
    \gamma & 0 & 0 \\
    0 & 1 & 0 \\
    0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    dx \\
    dy \\
    dz
\end{bmatrix}
\]  \hspace{1cm} (4-5)

so the area elements become

\[
\begin{bmatrix}
    dS'_x \\
    dS'_y \\
    dS'_z
\end{bmatrix}
= \begin{bmatrix}
    1 & 0 & 0 \\
    0 & \gamma & 0 \\
    0 & 0 & \gamma
\end{bmatrix}
\begin{bmatrix}
    dS_x \\
    dS_y \\
    dS_z
\end{bmatrix}
\]  \hspace{1cm} (4-6)

In order to have the flux be invariant, \( d\psi = d\psi' \), the components of flux density must then be altered by the Lorentz factor also. Equation (4-2) demands that the transformation matrices of area and flux density be inverse. Hence

\[
\begin{bmatrix}
    D_x \\
    D_y \\
    D_z
\end{bmatrix}
= \begin{bmatrix}
    1 & 0 & 0 \\
    0 & \gamma & 0 \\
    0 & 0 & \gamma
\end{bmatrix}
\begin{bmatrix}
    D'_x \\
    D'_y \\
    D'_z
\end{bmatrix}
\]  \hspace{1cm} (4-7)

It must be noted that (4-6) and (4-7) do not transform the area and flux density from any arbitrary frame to any other; the primed frame is the rest frame of the charge and is not arbitrary. Neither is the unprimed frame, which must be the rest frame of the point of measurement. Since the special theory of relativity denies the existence of any specially privileged coordinate frame, the properties of space are argued to be the same in any inertial frame. Hence the permittivity of free space is invariant under the Lorentz transformation, and the electric field must follow exactly the same law of transformation as the flux density:

\[
\begin{bmatrix}
    E_x \\
    E_y \\
    E_z
\end{bmatrix}
= \begin{bmatrix}
    1 & 0 & 0 \\
    0 & \gamma & 0 \\
    0 & 0 & \gamma
\end{bmatrix}
\begin{bmatrix}
    E'_x \\
    E'_y \\
    E'_z
\end{bmatrix}
\]  \hspace{1cm} (4-8)

The charge \( q \) is independent of the coordinate system. The force on a charged particle, given by

\[
\mathbf{F} = q \mathbf{E}
\]
must therefore obey a transformation law exactly similar to (4-7) for flux density and (4-8) for field strength. The force, as measured in the rest frame of the charge, is simply the Coulomb force already dealt with at some length. Behavior of charges in relative motion will now be investigated, applying this transformation to the electric force on a charged particle.

2. Force between Moving Charges

The field expression of Equation (4-8) may be used to correct Coulomb's law to allow for motion of the source charge. In most engineering problems, however, it is not at all convenient to have results expressed in a coordinate frame attached to one or the other of the charges. Especially in situations involving a multiplicity of moving charges (e.g., a system of current-carrying wires), it is useful to refer all the forces to an "impartial" general observer in whose rest frame all the charges are in motion. The necessary transformation laws are easily found by considering a system of two charges in motion with respect to each other as well as with respect to an observer.

In order to keep the mathematical development simple, the charges and the observer are at first assumed to have unequal but parallel velocities, directed along the \(x\)-axes of all three coordinate frames. (See Fig. 4.2.) The

Field at a point \(P_2\) stationary in frame \(F''\), due to a charge \(q_1\) located at point \(P_1\) and stationary in frame \(F'\), will be calculated, and the result given in terms of a third frame \(F\). It is assumed that the velocity of \(F''\) relative to \(F\) is \(v_2\), the velocity of \(F'\) relative to \(F\) is \(v_1\), and that \(v_{21}\) denotes the relative velocity of \(F''\) with respect to \(F'\).
The field of \( q_x \) at point \( P_2 \), referred to \( F'' \), is obtained directly by application of Equation (4-8):

\[
\begin{bmatrix}
E''_{2x} \\
E''_{2y} \\
E''_{2z}
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 0 \\
0 & \gamma_{12} & 0 \\
0 & 0 & \gamma_{12}
\end{bmatrix}
\begin{bmatrix}
E'_{2x} \\
E'_{2y} \\
E'_{2z}
\end{bmatrix}
\tag{4-9}
\]

Here the numerical and literal subscripts identify, respectively, the point at which \( E \) is measured and the component under consideration, while the primes indicate in which frame the measurement is carried out. Exactly the same field \( E'' \) (referred to frame \( F'' \)) could be created at \( P_2 \) by some system of charges stationary in frame \( F \). If that were done, the resulting field would be related to the field measured in \( F \) by

\[
\begin{bmatrix}
E''_{2x} \\
E''_{2y} \\
E''_{2z}
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 0 \\
0 & \gamma_2 & 0 \\
0 & 0 & \gamma_2
\end{bmatrix}
\begin{bmatrix}
E_{2x} \\
E_{2y} \\
E_{2z}
\end{bmatrix}
\tag{4-10}
\]

The thought experiment of placing charges in frame \( F \) is necessary here because Equation (4-8) is valid only if the field components on the right side are expressed in the rest frame of the source charges and those on the left are given in the rest frame of the point of measurement. It would not be correct to interpret (4-8) simply as a transformation law for the components of \( E \) between any two frames of reference!

The fields \( E''_2 \) of Equations (4-9) and (4-10) are required to be identically the same. Eliminating \( E''_2 \), there results the relationship between \( E_2 \) and \( E'_2 \)

\[
\begin{bmatrix}
E_{2x} \\
E_{2y} \\
E_{2z}
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 0 \\
0 & \gamma_{12}/\gamma_2 & 0 \\
0 & 0 & \gamma_{12}/\gamma_2
\end{bmatrix}
\begin{bmatrix}
E'_{2x} \\
E'_{2y} \\
E'_{2z}
\end{bmatrix}
\tag{4-11}
\]

This equation describes the electric field required in frame \( F \) to produce the same field in \( F'' \) as the charge \( q_x \). From Equation (26), the Lorentz factor \( \gamma_{12} \) is known to be expressible as

\[
\gamma_{12} = \gamma_1 \gamma_2 \left( 1 - \frac{v_1 v_2}{c^2} \right)
\tag{4-12}
\]

so that (4-11) may be rewritten as

\[
\begin{bmatrix}
E_{2x} \\
E_{2y} \\
E_{2z}
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 0 \\
0 & \gamma_1 & 0 \\
0 & 0 & \gamma_1
\end{bmatrix}
\cdot
\begin{bmatrix}
E'_{2x} \\
E'_{2y} \\
E'_{2z}
\end{bmatrix}
- \frac{v_1 v_2}{c^2} \gamma_1
\begin{bmatrix}
0 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
E'_{2x} \\
E'_{2y} \\
E'_{2z}
\end{bmatrix}
\tag{4-13}
\]

On comparison with Equation (4-8), it is seen that the first term on the right-hand side of (4-13) represents the relativistically corrected electrostatic field, as before; but the second term is altogether new. This new term is called the electrodynamic field because it depends on both the motion of the charge \( q_x \)
and the motion of the point of measurement. In the rest frame of either, the electrodynamic field is clearly zero. Combining Equations (4-8) and (4-13), the dynamic field \( \mathbf{E}_d \) may be written in terms of the static field \( \mathbf{E}_s \), with all quantities expressed in the same coordinate frame \( F \), as

\[
\begin{bmatrix}
E_{2x} \\
E_{2y} \\
E_{2z, d}
\end{bmatrix} = -\frac{v_1 v_2}{c^2} \begin{bmatrix}
0 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
E_{2x} \\
E_{2y} \\
E_{2z, s}
\end{bmatrix}
\]  
(4-14)

It will be noted that the dynamic field has no component in the direction of motion because the Lorentz contraction only affects length components in that direction. Furthermore, since \( v_1 \) and \( v_2 \) can never exceed \( c \), the dynamic field components must always be smaller than the corresponding static ones, regardless of the actual velocity values.

The more general transformation involving nonparallel velocities may be inferred from the above. The electrodynamic field has been seen to arise from two simultaneous Lorentz contractions, one in the direction of \( v_2 \) and one in the direction of \( v_{12} \). From Equation (4-8), it is evident that field components collinear with velocity are not affected by the Lorentz contraction; an electrodynamic field component in the direction of motion can never appear. The dynamic field must therefore always be expressible as the vector product \( (v_2 \times \mathbf{V}) \) where \( \mathbf{V} \) is some as yet undetermined vector. The nature of \( \mathbf{V} \) may be clarified by writing out the simple parallel-velocity case explicitly in vector form. Without loss of generality, it may be assumed that the charge \( q_1 \) and point \( P_2 \) lie in the plane \( z = z' = z'' = 0 \), as in Fig. 4.3, and are moving in the \( x \)-direction. From Equation (4-14), the dynamic field at point \( P_2 \) is

\[
\mathbf{E}_d = \mathbf{j} \frac{v_2 v_1}{c^2} D_1 \sin \alpha \frac{1}{\varepsilon_0}
\]  
(4-15)

Rewriting this expression in the form \( (v_2 \times \mathbf{V}) \) as

\[
\mathbf{E}_d = \frac{1}{\varepsilon_0 c^2} \mathbf{v}_2 \times (k v_1 D_1 \sin \alpha)
\]  
(4-16)

the term in parentheses becomes recognizable as a vector product:

\[
\mathbf{E}_d = \frac{1}{\varepsilon_0 c^2} \mathbf{v}_2 \times (\mathbf{v}_1 \times \mathbf{D}_1)
\]  
(4-17)

Equation (4-17) could also be arrived at by purely algebraic means, starting from (4-5) but with one set of coordinates rotated so that the coordinate axes
of the different frames never coincide. While this is in principle an attractive course, the mathematical development is lengthier and does not at all points lend itself to direct physical interpretation.

It should be noted that all quantities in Equation (4-17) are expressed in the coordinate frame \( F \) of the "impartial" observer. In that frame, the force on a test charge \( q_2 \) placed at the point \( P_2 \) may now be written, using the definition of electric field \( \mathbf{F} = q \mathbf{E} \), as

\[
\mathbf{F}_2 = q_2 (\mathbf{E}_2 + \mathbf{E}_d) = q_2 \frac{\mathbf{D}_2}{\varepsilon_0} + q_2 \frac{1}{c^2 \varepsilon_0} \mathbf{v}_2 \times (\mathbf{v}_1 \times \mathbf{D}_1) \quad (4-18)
\]

This is the result sought, an expression clearly indicating the force between two moving charges as seen in an arbitrary reference frame. It is to be emphasized that the dynamic force contribution is a purely relativistic phenomenon, i.e., entirely dependent on the observer's frame of reference. No observer attached to a charge can ever detect the presence of this force because in the rest frame of either charge, one of the velocities becomes zero and the electrodynamic force vanishes. On the other hand, the electrostatic force contribution is seen to be subject to relativistic correction (\( \mathbf{D}_1 \) is velocity-variant) but cannot be made to vanish by appropriate choice of coordinate frame.

3. Magnetic Field and Flux

The word electrodynamic is not conventionally employed to describe the motional force component of Equation (4-18), the more common adjective being magnetic. It is also usual to eliminate the necessity of writing \( \varepsilon_0 c^2 \) repeatedly by introducing the invariant constant

\[
\mu_0 = \frac{1}{\varepsilon_0 c^2} \quad (4-19)
\]

This constant is called the permeability of free space; its value is \( 4\pi \times 10^{-7} \) RMKSC units. The magnetic force on charge \( q_2 \) may then be written

\[
\mathbf{F}_{2m} = q_2 \mu_0 \mathbf{v}_2 \times (\mathbf{v}_1 \times \mathbf{D}_1)
\]

A further simplification is effected by defining a new vector

\[
\mathbf{H} = \mathbf{v} \times \mathbf{D} \quad (4-20)
\]

with whose use the force equation is further simplified to

\[
\mathbf{F}_{2m} = q_2 \mathbf{v}_2 \times \mu_0 \mathbf{H}_1 \quad (4-21)
\]

The new vector \( \mathbf{H} \) is called the magnetic field vector. From its definition it is seen that the magnetic field of any charge in its own rest frame is zero, for the velocity \( \mathbf{v} \) of Equation (4-20) is the velocity of a charge, relative to the observer in whose rest frame the field \( \mathbf{H} \) is expressed. The magnetic field of a single
moving charge may be visualized without great difficulty. Its direction must always be at right angles to both the line of motion and to its own electric field. The latter vector is directed radially outward from the charge. Hence the magnetic field vector must lie in a spherical surface centered on the charge, as well as in a plane orthogonal to the velocity \( v \). If a spherical surface surrounding the charge is imagined as a globe moving in the north-south direction, \( H \) lies in, and points along, parallels of constant latitude. At high velocities, the Lorentz contraction distorts the imaginary spherical surface into an oblate spheroid but does not affect the lateral symmetry of fields about the line of motion. The qualitative nature of the magnetic field remains unaltered, as shown in Fig. 4.4. Lines of the magnetic field vector evidently describe closed curves; they can have no beginning or end, lest symmetry be destroyed. In fact, these field lines must be circles centered on the line of motion of the charge. \( H \) is, in other words, a solenoidal vector. It should be emphasized that while the magnetic force of Equation (4-21) depends on the motion of two charges relative to an observer for its existence, the magnetic field is defined in terms of only one charge and its motion; it is therefore perfectly proper to speak of the magnetic field of a single moving charge.

If the charged particle in question is caused to move at a velocity approaching \( c \), the Lorentz contraction in the direction of motion will clearly assume major importance. Without delving into the mathematical details, it is immediately clear from Equations (4-8) and (4-13) that the electric field at right angles to velocity \( v \) grows larger and larger, assuming an ever increasing importance compared to the axial field. The total field of a moving charge thus crowds into an essentially planar configuration as the charge accelerates. For velocities very near \( c \), practically all the electric field—and by Equation (4-20), practically all the magnetic field—will lie in a plane normal to \( v \) and centered on the charge.
The solenoidal property of the magnetic field is also easily deduced from Equation (4-20) by a purely mathematical operation. Let the divergence of \( \mathbf{H} \) be sought,

\[
\text{div} \, \mathbf{H} = \text{div} \, (\mathbf{v} \times \mathbf{D}) \tag{4-22}
\]

This may be expanded by a well-known theorem of vector analysis as

\[
\text{div} \, (\mathbf{v} \times \mathbf{D}) = \mathbf{D} \cdot \text{curl} \, \mathbf{v} - \mathbf{v} \cdot \text{curl} \, \mathbf{D} \tag{4-23}
\]

The electric flux density vector is known to be irrotational in the rest frame of the charge with which it is associated (conservative property of the static electric field!), so the second term on the right-hand side vanishes. In order for special relativity to be applicable, motions must be kept uniform so as to avoid noninertial rest frames; hence \( \mathbf{v} \) must be irrotational also. It may be concluded that

\[
\text{div} \, \mathbf{H} = 0 \tag{4-24}
\]

and, by the divergence theorem,

\[
\oint \mathbf{H} \cdot d\mathbf{S} = 0 \tag{4-25}
\]

This development holds strictly true for any irrotational motion, \( \text{curl} \, \mathbf{v} = 0 \); it is not restricted to rectilinear motions only. For more general motions, a proof with less restrictive premises is required. These mathematical statements confirm the result of physical argument: there are no sources or sinks of the magnetic field.

In a manner analogous to electric field and flux density, magnetic flux density is defined as a vector related to the magnetic field by

\[
\mathbf{B} = \mu_0 \mathbf{H} \tag{4-26}
\]

and the magnetic flux, as the surface integral of the flux density:

\[
\phi = \int \mathbf{B} \cdot d\mathbf{S} \tag{4-27}
\]

A result analogous to Gauss's theorem in electrostatics may now be stated. The total flux out of a region of space may be found by taking the surface of integration in Equation (4-27) as closed. But by Equations (4-25) and (4-26), the magnetic flux density is a solenoidal vector, so the total flux emanating from any closed surface must be zero,

\[
\oint \mathbf{B} \cdot d\mathbf{S} = 0 \tag{4-28}
\]

Both magnetic vectors \( \mathbf{B} \) and \( \mathbf{H} \) are solenoidal. Their curls, however, are readily seen not to be zero. The fundamental mathematical difference between the electric and magnetic fields is that the electrostatic field is irrotational but divergent, while the magnetic field is nondivergent but rotational. Physically, the corresponding difference is that the electrostatic field is built
around point sources of flux, while the magnetic field lacks such sources so that the field lines close upon themselves.

In terms of the magnetic flux density vector, the electrodynamic field of Equation (4-17) may be written simply as

$$ E_d = v \times B $$

(4-29)

and the force equation (4-18) as

$$ F = q(E_s + v \times B) $$

(4-30)

The subscripts have been dropped since the meaning of the individual symbols is quite unambiguous. The force component arising from the magnetic field is frequently called the Lorentz force, and Equation (4-30) is referred to as the Lorentz force equation. Expressed in the above form, it is directly useful in charged particle ballistic studies. The beam electrons in a television picture tube, for example, are accelerated toward the screen by an electrostatic force and deflected to form a picture by magnetic fields created by coils placed around the neck of the tube. In many other applications, this form of Equation (4-30) is inconvenient because charge velocities associated with currents in conductors are rarely explicitly known. It will be necessary next to rewrite (4-30) in terms of current values rather than charges and their velocities, in order to serve the majority of engineering purposes.

Note should be taken of the fact that the charge $q$, on which the Lorentz force

$$ F = qv \times B $$

(4-31)

is exerted, does not need to be moving at a relativistically high velocity $v$. To experience a significant force, it is merely necessary to place it in a strong field $B$. It is this magnetic field, not the Lorentz force, that arises from relativistic contraction.

4. Ampère's Law

As noted, (4-31) expresses the force on a particle whose velocity needs to be known and is often not a convenient one to employ. Very frequently the individual charges $q$ exist only as portions of charge distributions, and their velocities $v$ can only be stated by inference. To make the above formulas more directly applicable, it is useful to rewrite them in terms of currents rather than charges.

Any slim conductor carrying a current $i$ may be thought of as subdivided into short portions of length $dl$ so that each portion contains a quantity $dq$ of free charges. If a time $dt$ is taken by the charge $dq$ to move a distance $dl$, then

$$ v dq = \frac{dl}{dt} dq = \frac{dq}{dt} dl = i dl $$

(4-32)
If the conductors involved in a problem are not fine enough, or the charge
distributions along them not nearly uniform enough, to use this relationship,
the cross-sectional area of each conductor may be subdivided into squares
dS at right angles to the current flow. The current crossing such a square is

$$i = J \cdot dS$$  \hspace{1cm} (4-33)

Substituting this expression into (4-32),

$$i\, dl = (J \cdot dS)\, dl = J(dS \cdot dl) = J\, dU$$  \hspace{1cm} (4-34)

where \(dU\) is the volume element \(dS \cdot dl\). There results the set of equivalent
forms

$$v\, dq = i\, dl = J\, dU$$  \hspace{1cm} (4-35)

The definition of magnetic field, Equation (4-20), may now be written in
terms of \(i\, dl\) instead of \(v\, dq\), using (4-35). A certain amount of mathematical
difficulty is encountered in doing so, unfortunately. In (4-20), all quantities
are referred to the rest frame of a single observer with respect to whom the
charge \(q\) is moving at a velocity \(v\). Thus, \(D\) must be given in his reference
frame, not in the rest frame of the charge. To find \(D\), it is necessary to write,
according to Equation (1-9),

$$dD' = \frac{dq}{4\pi(r')^2}$$

in the rest frame of the charge \(dq\), subsequently transform \(dD'\) into the ob-
server's frame by (4-7), and insert the result into (4-20) to yield an expression
for \(dH\). The equations that result are considerably complicated by the trans-
formation of \(dD\). Fortunately, however, it is very seldom necessary to perform
this transformation in engineering practice, for the charge velocities encoun-
tered in nearly all ordinary problems are tiny compared to the velocity of
light. In fact, it is so usual to set \(r = r'\) and \(\gamma = 1\) that many textbooks on
electromagnetic theory do not at any stage make any distinction between the
different reference frames. It may offhand seem that setting \(\gamma = 1\) de-
sroys the Lorentz force altogether, in view of the manner of its derivation.
On the other hand, Equation (4-31) does not confirm this expectation. The
reason lies in the fact that very low average charge velocities in metallic
conductors are coupled with enormous quantities of charge; although \(v\) is
small, the product \(qv\) remains of significant size. A numerical example may do
much to explain the situation. Let a pair of parallel wires stretch through
space, each 1 mm² in cross-sectional area, with a spacing of 1 cm center to
center. Suppose these wires to be of circular cross section (AWG No. 17
wire, very nearly), and suppose them to be carrying a current of 6 amperes
to a load; the wire-to-wire voltage will be supposed to be 2200 volts. [The
voltage chosen is just a bit short of the breakdown limit of air-insulated line,
see Equation (2-20), and the current the maximum permissible for convec-
tion-cooled line.] The excess charge on each wire will be about 0.04 micro-
coulomb per meter, as a result of the impressed voltage. This is the value of charge for calculations concerning electrostatic forces. The current flow, on the other hand, may be imagined as the process of moving all the free electrons down the wire at a speed sufficient to move 6 coulombs every second (definition of an ampere). Clearly what is being moved is not the 0.04 micro-coulombs per meter, which is merely the excess charge of either wire as compared to its uncharged (electrically neutral) state, but rather all the free charges in the wire. In copper, the conduction electrons make up the enormous charge quantity of 12 kilocoulombs per meter of wire! To deliver a current of 6 amperes, the necessary velocity is seen to be only 0.5 mm/sec, about 11 orders of magnitude less than the velocity of light.

It should be clear from this simple example that most engineering problems fall into two broad classes: those dealing with motions of charged particles, for which relativistic correction is necessary and which are best handled by direct application of the Lorentz force equation; and those involving current flow in conductive materials, where charge velocities are very low. In problems falling into the latter category, very slight error will be involved in setting \( \mathbf{dD}' = \mathbf{dD} \), or the equivalent \( \gamma = 1 \). If this approximation is made, Equation (4-20) may be written

\[
\mathbf{dH} = \mathbf{v} \times \mathbf{1_r} = \frac{1}{4\pi r^2} \mathbf{dl} \times \mathbf{1_r} \tag{4-36}
\]

Expressed in terms of the current, the field now becomes

\[
\mathbf{H} = \frac{i}{4\pi} \int \frac{\mathbf{dl} \times \mathbf{1_r}}{r^2} \tag{4-37}
\]

The force, from (4-31), may similarly be written first in differential form

\[
\mathbf{dF} = dq \mathbf{v} \times \mathbf{B} = i \mathbf{dl} \times \mathbf{B} \tag{4-38}
\]

and then integrated to give

\[
\mathbf{F} = i \int \mathbf{dl} \times \mathbf{B} \tag{4-39}
\]

or, in full,

\[
\mathbf{F} = \frac{i_1 i_2}{4\pi} \mu_0 \int \frac{\mathbf{dl}_1 \times (\mathbf{dl}_2 \times \mathbf{1_r})}{r^2} \tag{4-40}
\]

Equation (4-40) gives the magnetic force between two conductor segments carrying different currents. It forms the basis for all calculations of forces in systems of current-carrying conductors and therefore assumes a position of obvious importance in the analysis of all electromagnetic machinery. The usual name by which this equation is known is Ampère's law.

In problems dealing with distributed currents (e.g., electromagnetic molten metal pumps), it is sometimes more convenient to make the second substitution of Equation (4-35), \( \mathbf{v} dq = \mathbf{J} dU \), obtaining
The vector product $\mathbf{J} \times \mathbf{B}$ may be thought to represent the force on a volume element $dU$ of the current-carrying medium. It is therefore often called the volume force density.

5. Maxwell's Field Equations

The study of electromagnetic fields—that is to say, of the electric field, the magnetic field, and their interdependence—may proceed entirely on the basis of the foregoing definitions. The fact that relativistic corrections are rarely necessary in engineering work, however, permits rearrangement of the equations governing the interdependence of fields in such a manner as to make their use at low velocities more convenient. Such a formulation of the inter-relationship of the field vectors, in four differential equations, was arrived at by Maxwell in 1865.

In general, any vector quantity is specified (up to an arbitrary additive constant) if its curl and divergence are known. Between the five vectors $\mathbf{E}$, $\mathbf{D}$, $\mathbf{B}$, $\mathbf{H}$, $\mathbf{J}$, there are three so-called constitutive relations,

$$\mathbf{D} = \varepsilon \mathbf{E} \quad (1-33)$$
$$\mathbf{B} = \mu \mathbf{H} \quad (4-26)$$
$$\mathbf{J} = \sigma \mathbf{E} \quad (3-9)$$

so that knowledge of the divergence and curl of one electric and one magnetic field vector will suffice to determine all. These curls and divergences will now be sought.

The curl of the magnetic field intensity vector is found from its definition,

$$\text{curl } \mathbf{H} = \text{curl } (\mathbf{v} \times \mathbf{D}) \quad (4-42)$$

Expanding the curl of the cross product by one of the well-known vector identities,

$$\text{curl } \mathbf{H} = \mathbf{v} \text{ div } \mathbf{D} - \mathbf{D} \text{ div } \mathbf{v} + (\mathbf{D} \cdot \nabla)\mathbf{v} - (\mathbf{v} \cdot \nabla)\mathbf{D} \quad (4-43)$$

The first term on the right of Equation (4-43) may be recognized as

$$\mathbf{v} \text{ div } \mathbf{D} = \mathbf{v} \rho = \mathbf{J} \quad (4-44)$$

The second and third terms involve differentiation of the velocity vector, which is assumed uniform throughout. They are therefore zero ex hypothesi. There remains

$$\text{curl } \mathbf{H} = \mathbf{J} - (\mathbf{v} \cdot \nabla)\mathbf{D} \quad (4-45)$$

By exactly the same expansion, but substituting $\mathbf{B}$ for $\mathbf{D}$, there is obtained

$$\text{curl } (\mathbf{v} \times \mathbf{B}) = -(\mathbf{v} \cdot \nabla)\mathbf{B} \quad (4-46)$$
There is no term here corresponding to the current density $J$ in Equation (4-45), for $\text{div } B = 0$. Equation (4-46) may be rewritten as

$$\text{curl } (v \times B) = -v(1_v \cdot \nabla)B$$  \hspace{1cm} (4-47)

The bracketed differential operator may be recognized as the directional derivative operator of vector calculus, indicating differentiation in the direction of motion. If distance measured in that direction is denoted by $s$, then

$$(1_v \cdot \nabla)B = \frac{\partial B}{\partial s}$$  \hspace{1cm} (4-48)

The velocity is related to distance by

$$v = \frac{ds}{dt} \; 1_v$$  \hspace{1cm} (4-49)

as always. Equation (4-47) may now be used to find the curl of the electric field. The electric field in general consists of a static and a dynamic component,

$$\text{curl } E = \text{curl } E_s + \text{curl } E_d$$  \hspace{1cm} (4-50)

of which the first is known to be identically zero from electrostatics. Using the relationship $E_d = v \times B$, where $v$ denotes the velocity of the point of measurement through the rest frame of the field $B$, and combining Equations (4-47) through (4-50),

$$\text{curl } E = -\frac{\partial B}{\partial t}$$  \hspace{1cm} (4-51)

In a similar manner to (4-48),

$$(1_v \cdot \nabla)D = \frac{\partial D}{\partial s}$$  \hspace{1cm} (4-52)

On comparing Equations (4-20) and (4-29), an important distinction should be noted: in the former, the velocity $v$ measures the rate of motion of the field past the point of measurement, but in the latter, the motion of the point through the field. For this reason, (4-48) and (4-52) lead to different signs. Thus

$$(-v)(1_v \cdot \nabla)D = \frac{\partial D}{\partial t}$$  \hspace{1cm} (4-53)

and, from Equation (4-45),

$$\text{curl } H = J + \frac{\partial D}{\partial t}$$  \hspace{1cm} (4-54)

It is interesting to note that the identity

$$\text{div } \text{curl } H = 0$$

immediately yields the continuity equation of electric currents,

$$\text{div } (J + \frac{\partial D}{\partial t}) = 0$$  \hspace{1cm} (3-8)
which was earlier arrived at purely on the basis of the continuity postulate.

The equations obtained above give the curls of the electric and magnetic field intensity vectors. Divergences of the corresponding flux density vectors are already known. To summarize,

\[
\text{curl } \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (1) \quad \text{curl } \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \quad (2) \quad \text{(MD)}
\]

\[
\text{div } \mathbf{D} = \rho \quad (3) \quad \text{div } \mathbf{B} = 0 \quad (4)
\]

These four differential equations are known as Maxwell’s equations in their differential form. They evidently sum up all the theoretical development so far. For example, the continuity equation is embodied in two of them; and the system of Maxwell’s four equations could easily be replaced by two, plus the continuity requirement.

The synthesis of all previous development in a single set of equations could clearly not be possible without the term \(\frac{\partial \mathbf{D}}{\partial t}\). The discovery of this term and its introduction into electromagnetic theory represent the essence of Maxwell’s contribution, for once this term is accepted, the entire structure is consistent and can be expressed in the form above. In many books, \(\frac{\partial \mathbf{D}}{\partial t}\) is called the Maxwell displacement current density, or simply the displacement current density. This name arises from the ether theory of electromagnetism and cannot be ascribed much more than historical significance. The reasoning that led Maxwell to this name is primarily based on the electrical behavior of dielectric materials. Since the apparent shift of effective charge centers in polarizable materials is dependent on the electric field, \(\mathbf{D}\) may be thought to measure the displacement of charge centers. The time derivative of \(\mathbf{D}\) then measures the rate at which displacement of charges takes place, hence the name. However, it is not at all evident how anything can be “displaced” in a vacuum, although the continuity principle is demonstrably just as valid in vacua as in material media. In the ether theory, of course, \(\mathbf{D}\) is thought to express the polarization of the ether.

As a matter of convenience in calculation, integral equations are occasionally preferred to differential equations. Equally important, integral expressions refer to field distributions in entire regions of space, making them often easier to interpret physically than differential equations, which only refer to the immediate neighborhood of a point. The Maxwell curl equations may be transformed into integral equations by applying Stokes’ theorem. Taking integrals over some convenient simply connected surface,

\[
\int \text{curl } \mathbf{E} \cdot d\mathbf{S} = \oint \mathbf{E} \cdot d\mathbf{r} = -\int \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} = -\frac{\partial \phi}{\partial t} \quad (4-55)
\]

\[
\int \text{curl } \mathbf{H} \cdot d\mathbf{S} = \oint \mathbf{H} \cdot d\mathbf{r} = \int (\mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}) \cdot d\mathbf{S} = i + \frac{\partial \psi}{\partial t} \quad (4-56)
\]

The divergence equations are already known in their integral formulations. Summarizing,
\[ \oint \mathbf{E} \cdot d\mathbf{r} = -\frac{\partial \phi}{\partial t} \quad (1) \quad \oint \mathbf{H} \cdot d\mathbf{r} = i + \frac{\partial \psi}{\partial t} \quad (2) \] 
\[ \oint \mathbf{D} \cdot d\mathbf{S} = Q \quad (3) \quad \oint \mathbf{B} \cdot d\mathbf{S} = 0 \quad (4) \]

These four equations are known as *Maxwell's equations in their integral form*. Taken together with the constitutive relations, the four equations in either integral or differential form specify any electromagnetic field completely. They are sufficiently important that *every student of electromagnetic theory should commit Equations (MD) and (MI) to memory.*

### 6. Effects of Material Media

Electrostatic fields in material media were not found to differ fundamentally from their counterparts in free space. In view of its electric origins, it might reasonably be expected that behavior of the magnetic field would follow a similar pattern. In dielectric materials, charge centers in individual material particles shift, causing polarization, which is taken into account by altering the value of permittivity. The behavior of magnetic materials is somewhat more complicated in its details, for every portion of matter is constituted of a great many elementary particles, many of which are electrically charged and in relative motion in both their orbital and spinning motions. Without delving into the microstructure of materials, it is not possible to say what particular materials will exhibit which kind of behavior; it is, however, reasonable to expect that there will exist at least some class of materials in which the various elementary particle motions are such as to produce a macroscopic effect. This effect will only arise if the internal currents (charged particle motions) are not purely random and will be externally equivalent to currents circulating within the material. In magnetic fields, such materials react by aligning their internal currents in accordance with the externally impressed field. The effect of material magnetization is dependent on the external field, analogously to electric polarization. In the ferromagnetic materials commonly used in engineering, the magnetization aids the external field; it is usual to describe it in terms of a *magnetization vector* \( \mathbf{M} \),

\[ \mathbf{M} = \kappa \mathbf{H} \quad (4-57) \]

where \( \kappa \) is called the magnetic susceptibility of the material. The flux density in the material is the result of the combined action of the external field and the magnetization of the material,

\[ \mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M}) \quad (4-58) \]

This manner of describing the fields is very convenient for studies concerned with the structure and properties of the material, for it clearly exhibits the separate effects of external and internal contributions to the flux density.
vector. On the other hand, engineering design calculations are frequently not concerned with the separate contributions, but only with the total effect. For such calculations, Equations (4-57) and (4-58) may be combined to eliminate $\mathbf{M}$, giving

$$\mathbf{B} = \mu_0 (1 + \kappa) \mathbf{H} \tag{4-59}$$

The quantity

$$\mu' = (1 + \kappa) \tag{4-60}$$

is called the relative permeability of the material (relative to free space, for which $\mu' = 1$). For brevity, it is usual to write

$$\mu = \mu_0 \mu' \tag{4-61}$$

where $\mu$ is called the permeability of the material. The close similarity of this notation to that used for electric fields is evident. It should be pointed out, however, that magnetic field calculations are substantially complicated by the frequent nonlinearity of $\mu$; in contrast to dielectric materials which generally have constant permittivity, the permeability of most magnetic materials is strongly field-dependent. Also, there exist materials with relative permeability smaller than unity, while no dielectrics are known whose relative permittivity is less than one. The structure of magnetic materials is generally more involved than that of dielectrics, and it forms an important field of materials science. For the student of electromagnetic fields, however, it is not necessary to delve into the microstructure of materials; if their macroscopic properties, that is, their permittivity and permeability, can be ascertained, enough is known to permit solution of Maxwell’s equations.

7. Fields at Material Boundaries

Any electromagnetic field problem may be solved simply by finding solution functions that satisfy Maxwell’s equations. Since these are partial differential equations, there exists an infinite number of possible solutions; which solution is correct for the particular problem at hand is determined by the additional requirement that certain conditions must be satisfied at all the boundaries that separate various regions containing distinct media. These boundary conditions are themselves consequences of the fundamental electromagnetic laws and may be discovered by examining the behavior of the field vectors in the neighborhood of material boundaries. The three constitutive relations describe the relationships of field vectors to each other; since the material properties are presumed to be known if any attempt at a solution is to be made at all, there remain two entirely independent field vectors, one electric and one magnetic, at each boundary. Because either of these vectors may be resolved into a normal and tangential component, four distinct conditions must be found in order to specify the field behavior completely.
Considering the normal components first, it is immediately noted that the equation

$$\oint \mathbf{D} \cdot d\mathbf{S} = Q$$  \hspace{1cm} (MI-3)

remains unchanged from electrostatics. In Chapter 1 it was applied to a thin, pancake-shaped volume that just encloses the material interface; there results the condition

$$D_{n1} - D_{n2} = \sigma$$  \hspace{1cm} (1-75)

where the subscript $n$ denotes the normal component, and the numerical subscripts identify two different media. Exactly the same argument may be repeated to find the behavior of the normal component of the magnetic flux density vector $\mathbf{B}$ by applying the equation

$$\oint \mathbf{B} \cdot d\mathbf{S} = 0$$  \hspace{1cm} (MI-4)

Since there cannot be any such thing as a magnetic surface charge, there emerges the condition

$$B_{n1} - B_{n2} = 0$$  \hspace{1cm} (4-62)

The tangential vector components may be examined in a similar manner. Let a rectangular path be laid out, as in Chapter 1, with its long sides parallel to the material interface and its short sides normal to it. Let the short sides be just long enough to penetrate the interface and the long sides have a length $w$. If the electric field strength is integrated along this path, the Maxwell equation

$$\oint \mathbf{E} \cdot d\mathbf{r} = -\phi$$  \hspace{1cm} (MI-1)

reduces essentially to the sum of integrals along the long sides; the short sides are of infinitesimal length and their contribution may be neglected. It follows that

$$E_{t1} - E_{t2} = -\frac{1}{w} \phi_{\text{enclosed}}$$  \hspace{1cm} (4-63)

exactly as in electrostatics, Equation (1-78), except that the flux term now appears. Since the rectangular path is of infinitesimal area, however, the enclosed flux is very small. It may in fact be made arbitrarily small by taking the rectangle sufficiently narrow. The right-hand member of Equation (4-63) then approaches zero, with the result

$$E_{t1} - E_{t2} = 0$$  \hspace{1cm} (4-64)

exactly as in the electrostatic case.

For the tangential component of the magnetic field vector, a similar argument may be applied. Again integrating around a path that just encloses a section of the surface between two media, the Maxwell equation
\[ \oint H \cdot dr = i + \frac{\partial \psi}{\partial t} \] (MI-2)
gives
\[ H_{l1} - H_{l2} = \frac{1}{w} \left( \frac{\partial \psi}{\partial t} + i \right)_{\text{enclosed}} \] (4-65)

Once again the amount of flux enclosed may be made arbitrarily small by taking the rectangular path narrow enough, and the first term on the right-hand side of Equation (4-65) vanishes. The current term vanishes also in materials with finite conductivity, but it remains if one of the media is a perfect conductor. Perfect conductors are not only of interest in superconductivity studies, but they form a very attractive simplifying assumption in many other problems; it would be unwise to dismiss them without further examination.

In a superconductor, the internal field intensity \( E \) must at all times be zero, for the current density in the material cannot be infinite throughout a volume of finite size; if it were, the associated energy would have to be infinite, as will shortly be seen. As a matter of fact, restricting \( E \) to zero value precludes the possibility of transferring any energy into the interior of the medium and hence prohibits nonzero values of current density in the interior altogether. All the current that flows along a perfect conductor must consequently flow in an infinitesimally thin sheet along its surface, so that the current term in Equation (4-65) cannot be affected by the choice of integration path. The current density \( J \) is clearly infinite, but only over an infinitesimal thickness. Any mathematical inconvenience is avoidable by defining \textit{surface current density} or \textit{current sheet density} as the amount of current per unit width of surface:

\[ J_s = \frac{i}{w} \mathbf{l}_f \] (4-66)
a vector quantity having the direction of current flow. The boundary condition for tangential magnetic field may hence be written

\[ H_{l1} - H_{l2} = J_{so} \] (4-67)

where the subscript \( o \) denotes the component orthogonal to the selected direction of the tangential component of \( H \).

Such a situation can actually occur only in superconductors. However, the analysis of many other situations is greatly simplified if real current distributions are replaced by approximately equivalent current sheets. The currents in waveguide walls, for example, flow in thin enough layers to warrant modeling the waveguide by perfectly conducting walls and current sheets. In electric machines, current distributions that approximate to thin sheets are commonly created by distributing current-carrying conductors over the surfaces of machine members. An induction machine, for example, is conveniently viewed as two concentric cylindrical current sheets, embedded
in ferromagnetic material and separated by an air gap. For purposes of approximate analysis, machine windings are often imagined to be compressed into a very thin sheet, and the air gap is taken as wider than its actual size to allow for slotting of the machine members.

8. Electromotive Force

Many difficult problems in electrostatics were considerably simplified by defining the potential function $V$, a scalar quantity, as the line integral of the electric field. In the general case involving moving charges, however, the electric field vector is not irrotational, and a similar definition of potential is not possible because the line integral of $E$ does not have a unique value. The electrodynamic field, in other words, is not conservative. A potential-like scalar function may nevertheless be defined even for nonconservative vector fields; but it will not have a unique value. Its value will be related both to the field $E$ and to the nature of the contour of integration, so that the function becomes a description not only of the field but also of the contour. Such pseudopotentials are of especially wide use in network theory and are called electromotive forces. For a given closed contour $C$, the electromotive force $U$ is defined by

$$ U = \oint_C E \cdot dr $$

(4-68)

The physical significance of this quantity may be seen from consideration of a simple problem. Let an infinitely long straight conductor, carrying a current $i$, be moving at right angles to its axis with a velocity $v$. Let a closed contour $PQRS$ be made up of four straight sides parallel and at right angles to the wire, as in Fig. 4.5. The emf around this contour will be given by

$$ U = \int_{PQ} E \cdot dr + \int_{RS} E \cdot dr $$

(4-69)

No contribution accrues to the integral from the sides $QR$ and $SP$, for there

Fig. 4.5
the direction of \( \mathbf{E} \) is at right angles to the line of integration. Along the other two sides, \( \mathbf{E} = \mathbf{B} \times \mathbf{v} \) is in the direction of the contour. It can be shown that the field values required are
\[
\mathbf{E}_{PQ} = \frac{\mu i}{2\pi r_1} v \mathbf{j}, \quad \mathbf{E}_{RS} = \frac{\mu i}{2\pi r_2} v \mathbf{j} \tag{4-70}
\]
Combining these, the value of \( U \) is given by
\[
U = \frac{\mu i w}{2\pi} \left( \frac{1}{r_1} - \frac{1}{r_2} \right) v \tag{4-71}
\]
If the contour is physically made up of conductive material, as for example a wire loop, the existence of a field intensity \( \mathbf{E} \) requires a current density \( \mathbf{J} \) to exist; a current will flow around the loop. Were there no moving magnetic field, the same total current and the same current distribution could be achieved in the loop by creating the internal electric field in some other way. One suitable method might be to break the loop at points \( A-B \), as shown in Fig. 4.6, and fix \( A \) and \( B \) at two different electric potentials (say by means of a chemical battery). If the current is to be the same as before, the field distribution in the conductor must be similar also; the line integral of \( \mathbf{E} \) around the contour must then be unaltered. Consequently, the required potential difference is
\[
U = V_A - V_B \tag{4-72}
\]
Simply stated, the electromotive force of a contour is numerically equal to the electric potential difference needed to create the same current distribution in the contour when no changing magnetic flux is present. This interpretation may also be arrived at by purely mathematical means, by combining Equations (4-68) and (M1-1) to give
\[
U = -\frac{\partial \phi}{\partial t} \tag{4-73}
\]
an equation occasionally called Faraday’s law.
It might be expected from the formal similarity of the magnetic and electric line integral equations (MI-1) and (MI-2) that a magnetic quantity dual to electromotive force could be introduced. It can; and it will be defined in connection with static magnetic fields in Chapter 6.

9. Kirchhoff’s Circuit Laws

The two most widely applied equations in electrical engineering are probably the Kirchhoff circuit laws. Because they are of interest in themselves, as well as capable of illuminating the significance of the Maxwell equations, it will now be shown that they are in fact nothing more than approximate representations of the field equations, specialized for use in network problems.

Let a problem be considered in which a complete field solution is not desired but only a knowledge of the potentials at a set of points. Let the configuration of current-carrying conducting bodies be such that currents flow only along essentially filamentary paths. Potential values will be required at all points where the filamentary currents meet. These points shall be called nodes.

The equation of continuity demands that at each and every point in a field

$$\text{div } \mathbf{J} + \frac{\partial \rho}{\partial t} = 0$$  \hspace{1cm} (3-7)

Imagining a typical node to be enclosed within a small closed surface, (3-7) may be written, using the divergence theorem,

$$\oint \mathbf{J} \cdot d\mathbf{S} + \frac{\partial}{\partial t} \int \rho \, dU = 0$$  \hspace{1cm} (4-74)

The second term in this equation merely represents the rate of charge variation within the surface of integration. A nonzero derivative of enclosed charge implies accumulation of stored charge at a physically very small junction. From electrostatics, this is known not to be a practically tenable situation except for very small stored charges. Thus, unless the rate of variation is very rapid, there results the approximation

$$\oint \mathbf{J} \cdot d\mathbf{S} = 0$$  \hspace{1cm} (4-75)

Now the current density must be zero everywhere around a node except where the filamentary currents flow; hence (4-75) may be written as

$$\sum_k i_k = 0$$  \hspace{1cm} (4-76)

a statement known as Kirchhoff’s current law. It is evident from the above
that it (a) arises from the magnetic curl equation, \( \text{curl } H = J + \dot{D} \), and (b) is dependent on the assumption \( \dot{D} \ll J \). The latter is a good approximation at low frequencies but grows worse as rates of time variation increase; even a very tiny charge can give rise to an appreciable displacement current term \( \dot{D} \) if it is varied quickly enough!

Just as the current law arises from the magnetic curl equation, so the Kirchhoff voltage law is a special case of the electric curl equation. In integral form, the latter may be written

\[
\oint E \cdot dr = -\phi
\]  

(MI-1)

Consider first the simple situation in which no changing magnetic flux exists. The right side of (MI-1) is then zero, the electric field vector irrotational, and the usual scalar potential \( V \) may be defined. Only the values of potential at the nodes are of interest. Thus, the line integral may be first rewritten as the sum of line integrals from node to node,

\[
\int_{1}^{2} E \cdot dr + \int_{2}^{3} E \cdot dr + \ldots + \int_{n}^{1} E \cdot dr = 0
\]  

(4-77)

whereupon each term may be recognized as merely the scalar potential difference between the end points of integration. This sum of integrals may be written

\[
\sum_{k=1}^{n-1} V_{(k)(k+1)} + V_{n1} = 0
\]  

(4-78)

a statement known as Kirchhoff's voltage law. The restriction of no changing magnetic field may next be removed by cutting the contour (as in Fig. 4.6) and placing a source equal to the emf into this cut. An additional node is created by the cutting operation. The summation of node-to-node potential differences then includes one more term, the extra term representing the loop emf. This process is illustrated by Fig. 4.7. On the left is shown a contour encompassing five distinct nodes but not linking any time-variant flux. Equation (4-78) reads, for this contour,

\[
V_{12} + V_{23} + V_{34} + V_{45} + V_{51} = 0
\]  

(4-79)
The contour in the right half of Fig. 4.7 is similar, except that it does link a time-varying magnetic flux. A cut is arbitrarily located just to the right of node 1; a sixth node is thereby created. The first and sixth nodes are taken very close together, so that little if any electrostatic potential difference appears between them. There must exist, however, a potential discontinuity between these two nodes, representing the electromotive force $U$ of the contour. On adding it to the potential differences in (4-79), there results

$$V_{12} + V_{23} + V_{34} + V_{45} + V_{56} - U = 0 \quad (4-80)$$

It is not necessary to locate the cut next to one of the existing nodes, although this is frequently convenient; it could be placed at any point within the contour. In such cases, there may result two new nodes. No fundamental change occurs in Equation (4-80), however, since the effect is merely to partition one of the existing potential-difference terms without altering the sum of its parts.

**READINGS**

Numerous books present a relativistic derivation of the magnetic field, though difficulties arise in some with the varying notations employed as well as with the extent to which familiarity with special relativity is presumed to exist. Carter (1) presents a derivation in many respects similar to that shown above and may be consulted with profit, as may the more advanced book by Page and Adams (2). A closely related approach may be found in Elliott (3) or Tralli (4), who first derive a general force transformation and then apply it directly to Coulomb’s law. Particularly, the former book is to be recommended. Many advanced texts employ a tensor notation, which tends to confuse the uninitiated at first. However, the presentation of Moon and Spencer (5) is sufficiently well written to prevent excessive difficulty from arising. A self-contained expository article so clear and well written as to merit recommendation over many of the established textbooks is that by Elliott (6).

The basic physical ideas of relativistic electrodynamics are clearly though briefly explained in an article by Page (7). Tripp (8) is more specifically concerned with forces and the mechanism underlying them.

It is possible, as an alternative to relativistic kinematics, to accept Ampère’s law as an experimental fact, and derive Maxwell’s equations thence. Many authors, especially of the older books, do this. One very good and readable example is furnished by Jordan (9). An excellent modern version of this approach may be found in Corson and Lorrain (10). Reitz and Milford (11) also give a good discussion of the Ampère and Biot–Savart laws in their classical formulation.

As in the case of dielectrics, the reader is referred to books on materials science for information on the nature of magnetic materials. Short explana-
tory descriptions of material properties may, however, be found in numerous field theory books. A good, readable chapter is presented by Nussbaum (12), as well as Corson and Lorrain (10) or Harnwell (13). The chapter in Elliott’s book (3) spans nearly a hundred pages, and must be considered more demanding. The great standard treatise on ferromagnetic substances is Bozorth (14); a very comprehensive bibliography will also be found there. A much less ambitious work in the same area is that of Brailsford (15).

The relationship between network theory and field theory is not very carefully explored by most authors. A welcome exception is Holt (16). Very careful consideration is given to this problem by King (17); the treatment of Ramo, Whinnery, and Van Duzer (18) is quite brief, but lucid. Jordan (9) also devotes a dozen or so pages to the subject, stressing in particular the sinusoidal steady state.

Finally, a delightfully clear exposition of Faraday’s law, electromagnetic force, potential difference, and the difficulties these may cause is furnished by Bewley (19).


PROBLEMS

4.1 A charged particle moves along a straight-line path. Assuming its charge to be 1 microcoulomb and its velocity 1 m/sec, find the curves of equal value of $H$ in a plane containing the line of motion. Plot them to scale.

4.2 Two beam electrons in a cathode-ray tube travel toward the screen with equal and parallel velocities. Assuming that they have been accelerated by a potential difference of 5000 volts acting uniformly over a distance of 25 mm, find the force exerted by either electron on the other.

4.3 The electrons of Problem 4.2 are traveling at the same velocity and hence are stationary with respect to each other. Their forces should therefore be purely electrostatic. Are they?

4.4 From Equation (4-37), deduce the magnetic field in the neighborhood of a straight, cylindrical, current-carrying wire.

4.5 From Equation (4-41), find the force between two parallel current-carrying wires.

4.6 A molten metal trough is 20 cm wide and 2 cm deep. A magnetic field is set up such that a flux density normal to the metal surface of 0.01 tesla exists. Assuming the conductivity of the metal to be $2 \times 10^6$ mho per meter, find the potential necessary between carbon electrodes at the sides of the trough in order to move metal along the trough up a 20° incline. What current per meter of trough is needed? How much power is used? Assume the metal to have a specific gravity of 7.

4.7 A circular glass container has its outer rim and a center pin made of highly conductive metal; between these two electrodes, a potential difference is maintained, with the center pin positive with respect to the rim. The container is filled with a conductive solution, so that a radial current flows from
the center outward, and placed in a uniform vertical magnetic field. Find the velocity of rotation of the liquid, assuming viscosity to be negligible.

4.8 Expand Maxwell's equations in Cartesian coordinates. Derive three scalar equations from each of the vector equations.

4.9 Expand Maxwell's equations in cylindrical coordinates. Is it always possible to resolve the vector equations into three scalar equations each?

4.10 A two-wire power line with a wire spacing of \( d_1 \) meters runs parallel to a two-wire telephone line with a wire spacing of \( d_2 \) meters. The distance between the lines is \( d_o \) meters, and the power line carries a current of \( I \cos \omega t \) amperes. Find the strength of the interfering signal in the telephone line for a parallel run of \( k \) kilometers of line. Suggest simple means of reducing the interference on an existing line.

4.11 In certain geophysical measurements, microvolt signals (d-c to low audio frequencies) must be carried several hundred feet from detection apparatus to measuring instruments. It is to be expected that the leadwires will pick up 60 Hz noise from stray electric and magnetic fields. The manufacturer of NoHum-4 cable claims that his cable is for all practical purposes noiseproof. Upon investigation, NoHum-4 turns out to consist of four wires arranged in a square, with diagonally opposite leads connected in parallel at both ends (Fig. 4.8). Prove the manufacturer's claim either valid or invalid.

![Fig. 4.8](image_url)

4.12 Investigate the suitability of coaxial cable for the purposes of Problem 4.11.

4.13 Prove that the normal derivative of the tangential component of the magnetic field is zero at the surface of an ideal conductor.

4.14 A rectangular loop of fine wire carries a current \( i \). The loop is placed between two infinite equipotential planes so that its longer sides are parallel to the planes, and it is moved in the direction of these sides with a uniform velocity \( v \). Show that there exists a torque on the loop, even though the loop carries no net charge.
The central purpose of all electrical engineering is the control and use of electromagnetic energy, whether for its own sake or as an encoding medium for information. The energy stored in the electric field was investigated in Chapter 1, and its dissipation by electric currents in Chapter 3. Addition of magnetic to electric fields, however, creates the possibility of storing energy in magnetic fields as well. In this chapter, a few mechanical and mathematical properties of the magnetic field are investigated as a preliminary step, and the laws governing storage and movement of energy in the electromagnetic field then established.

1. Magnetic Vector Potential

In electrostatics, much laborious vector calculation is avoided by defining an electric scalar potential. The definition of potential presupposes the field to be single-valued and irrotational everywhere. A similar method can un-
Fortunately not be adopted for magnetic fields, for their curl is in general nonzero. However, substantial simplification of magnetic field calculations can nevertheless be achieved by exploiting the vector identity

$$\text{div curl } \mathbf{W} = 0$$

valid for any differentiable vector \( \mathbf{W} \). The Maxwell divergence equation \( \text{div } \mathbf{B} = 0 \) will always be satisfied if a vector \( \mathbf{A} \) is defined such that

$$\mathbf{B} = \text{curl } \mathbf{A} \quad (5-1)$$

This differential equation does not uniquely define \( \mathbf{A} \), however; any irrotational vector \( \mathbf{S} \) may be added to it without altering \( \mathbf{B} \), since curl \( \mathbf{S} = 0 \) so that curl \( (\mathbf{A} + \mathbf{S}) = \text{curl } \mathbf{A} \). Another restriction is needed to specify \( \mathbf{A} \) completely. Arbitrarily, the so-called Coulomb convention

$$\text{div } \mathbf{A} = 0 \quad (5-2)$$

will be adopted here. Obviously this choice is not the only one possible; a number of authors, as a matter of fact, have for various good reasons chosen to specify the divergence of \( \mathbf{A} \) to have some other value. Whatever choice may be made, \( \mathbf{A} \) is uniquely specified by its curl and divergence, except for an arbitrary additive constant.

The implicit definition of \( \mathbf{A} \) above may be put in explicit form. Taking curls, there is first obtained from Equation (5-1)

$$\text{curl } \mathbf{B} = \text{curl curl } \mathbf{A} \quad (5-3)$$

so that

$$\text{curl } \mathbf{B} = \nabla^2 \mathbf{A} \quad (5-4)$$

But

$$\text{curl } \mathbf{B} = \mu \text{ curl } \mathbf{H} = -\mu \left( \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \right)$$

Combining with Equation (5-4),

$$\nabla^2 \mathbf{A} = -\mu \left( \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \right) \quad (5-5)$$

This result resembles the Poisson equation of electrostatics, except for the fact that on the left-hand side, a vector Laplacian operator operates on a vector instead of on a scalar potential. Because of this formal similarity to the electrostatic potential \( V \), the name usually given to \( \mathbf{A} \) is magnetic vector potential.

The field in the neighborhood of a long wire of radius \( a \), carrying a current \( i \), may serve to illustrate the physical meaning of magnetic vector potential. From symmetry it may be argued that the magnetic field intensity \( \mathbf{H} \) can only depend on distance from the wire but not on angular or longitudinal
position (note the similarity to the treatment of electrostatically charged wires!). The value of $H$ at a distance $r$ from the wire may be calculated by taking a circular path of integration centered on the wire, as in Fig. 5.1, and applying the Maxwell equation

$$\oint H \cdot dr = i + \frac{\partial \psi}{\partial t}$$  \hspace{1cm} \text{(MI-2)}$$

Assuming the current to be steady, there is no time variation of electric flux; furthermore, the magnitude of $H$ is constant and its direction coincident with $dr$ throughout the integration. Thus, outside the wire

$$H = \frac{i}{2\pi r_0} \quad r_o \geq a$$  \hspace{1cm} \text{(5-6)}$$

and the flux density, as a consequence, is

$$B = \frac{\mu i}{2\pi r_0}$$  \hspace{1cm} \text{(5-7)}$$

Inside the wire these equations give the wrong result because a path of integration that does not lie entirely outside the wire cannot enclose the whole current $i$. Assuming the current density to be uniform throughout the wire cross section, the current enclosed by a circular path inside the wire, $r_o \leq a$, is given by

$$i_{\text{enclosed}} = \left(\frac{r_o^2}{a^2}\right)i$$  \hspace{1cm} \text{(5-8)}$$

so that the value of $H$ is given by

$$H = \frac{r_o i}{2\pi a^2} \quad r_o \leq a$$  \hspace{1cm} \text{(5-9)}$$

The variation of $H$ and $B$ with radius $r_o$ is shown in Fig. 5.2.
To find the vector potential near the wire and relate it to the calculated $B$ and $H$, it will first be noted that the integration of $A$ over a closed path produces simply the magnetic flux:

$$\oint A \cdot dr = \int \text{curl} A \cdot dS = \int B \cdot dS = \phi$$  \hspace{1cm} (5-10)

Let a rectangular path $PQRS$ be laid out, with the sides $PS$ and $QR$ at right angles to the wire and $RS$ parallel to the wire, as in Fig. 5.3. From the vector Poisson equation (5-5) it is clear that $A$ only has a component in the direction of the current density vector. Along $PQRS$, it is then normal to the path along two sides and parallel to it along the other two. Letting this rectangular path be the contour of integration, there obtains

$$\oint A \cdot dr = \int_{PQ} A \cdot dr + \int_{QR} A \cdot dr + \int_{RS} A \cdot dr + \int_{SP} A \cdot dr$$  \hspace{1cm} (5-11)

The value of $A$ is now readily found by combining this equation with the
expression for flux density. There obtains

\[ A = \frac{\mu I}{4\pi a^2} \quad \text{inside the wire} \quad (5-12) \]

\[ = \frac{\mu I}{2\pi} \left( \frac{1}{2} + \log \frac{r}{a} \right) \quad \text{outside} \quad (5-13) \]

The variation of \( A \) with radial distance is shown in Fig. 5.4. It will be noted that the magnetic vector potential, unlike the electric scalar potential, increases logarithmically with distance from a line source. This is hardly surprising in view of its physical interpretation as a measure of the flux between the current-carrying conductor and the place of measurement!

2. Magnetic Moment and Torque

Energy stored in electric fields was calculated in Chapter 1 by examining the forces necessary to move a small test charge from one point in the field to another. Investigation of magnetic field energy is possible along similar lines, but with a unit element of current replacing the test charge. In order to satisfy the principle of charge conservation, all currents must ultimately flow in closed paths, requiring a unit current loop as a standard test element. In this section, the torque on a very small current-carrying loop will be found. This result will subsequently be generalized and used to find the forces, torques, and energy of any volume distribution of currents. To assure that the results represent purely magnetic phenomena, all the current-carrying paths will initially be assumed to be electrically neutral. This assumption will later be removed, permitting electric and magnetic fields to participate in energy storage simultaneously.

The entire procedure to be followed, it should be pointed out, is open to one serious objection in principle. The existence and properties of the magnetic field as developed in Chapter 4 are based on the special theory of relativity, which is valid for all inertial reference frames. Currents that flow in closed paths, however, must represent charge motion along curved contours, possibly involving angular acceleration, and may not strictly fall within the scope of the special theory of relativity. In order to use the results of Chapter 4, the discussion must therefore be limited to charge velocities so low that angular accelerations assume negligibly small values; otherwise, redevelopment on the basis of the general theory of relativity (in which accelerating reference frames are considered) is indicated. Fortunately, the low-velocity assumption is well satisfied by all conduction currents in conductive materials, as well as by elementary particle motions in applications at present within the scope of electrical engineering.

Let a small, plane, current-carrying loop be considered, as shown in Fig. 5.5. Such a loop may be thought to consist of a sequence of current elements
i \, dl \text{ laid end-to-end. If the loop is immersed in a uniform magnetic field } H, 
the force on one of the current elements is

\[ dF = \mu_0 \, i \, dl \times H = i \, dl \times B \]  \hspace{1cm} (5-14)

and the torque about the origin is

\[dT = r \times dF = r \times (i \, dl \times B) \]  \hspace{1cm} (5-15)

The total torque on the entire loop is then

\[ T = \oint r \times (i \, dl \times B) \]  \hspace{1cm} (5-16)

Two well-known vector identities,

\[ A \times (B \times C) = B \times (A \times C) - C \times (A \times B) \]

and

\[ A \times (B \times C) = B (A \cdot C) - C (A \cdot B) \]

may be employed to simplify Equation (5-16). By using these in the order shown, there is obtained

\[ T = i \oint r (dl \cdot B) - i \oint B (r \cdot dl) - i \oint B \times (r \times dl) \]  \hspace{1cm} (5-17)

The middle integral on the right-hand side may be converted by Stokes’ theorem to a surface integral:

\[ \oint B (r \cdot dl) = B \oint r \cdot dl = B \int \text{curl} \, r \cdot dS = 0 \]  \hspace{1cm} (5-18)

The first integral on the right-hand side of (5-17) may be rewritten as

\[ \oint r (dl \cdot B) = 1_r \oint r (B \cdot dl) \]  \hspace{1cm} (5-19)

and the inequalities set up

\[ r_{\text{minimum}} \oint B \cdot dl \leq \oint r (B \cdot dl) \leq r_{\text{maximum}} \oint B \cdot dl \]  \hspace{1cm} (5-20)

The line integral of flux density in either inequality may be expressed in terms
of the field as
\[ \oint B \cdot dl = \mu \oint H \cdot dl \]  
(5-21)

By Maxwell's equations, this becomes in turn
\[ \oint B \cdot dl = \mu \left( i_{\text{enclosed}} + \frac{\partial \psi}{\partial t} \right) \]
that is to say,
\[ \oint B \cdot dl = 0 \]  
(5-22)

since no current of any kind flows across the surface of the loop. Hence the inequalities (5-20) become
\[ 0 \leq \oint r(B \cdot dl) \leq 0 \]  
(5-23)
so that the first integral on the right-hand side of (5-17) must vanish. There remains
\[ T = -i \oint B \times (r \times dl) \]  
(5-24)

which, since \( B \) is assumed uniform over the small loop, may be rewritten
\[ T = -iB \times S \]  
(5-25)
where \( S \) is the surface vector of the loop. It is usual to define the magnetic moment of a small loop by the equation
\[ \mathcal{M} = Si \]  
(5-26)
so that the torque on the loop is finally
\[ T = \mathcal{M} \times B \]  
(5-27)

For nonuniform fields \( B \), this development is nearly true if the loop is small, for \( B \) must vary in a continuous fashion. Hence \( B \) is nearly uniform over the area of the loop. This approximation may be made exact by adjusting the loop current and area so as to make the area vanishingly small, while keeping its magnetic moment constant. Such a limiting process results in an elementary magnetic dipole; if the magnetic moment is chosen to have the value unity, the result is called a unit magnetic dipole.

3. Magnetic Field Energy of Current-Carrying Circuits

Any combination of current-carrying loops of finite size may be built up out of a large number of elementary magnetic dipoles, much as surface or volume distributions of charge can be constructed out of elementary charges. For example, let the large loop shown in Fig. 5.6 be replaced by a set of meshes as shown. No alteration in the spatial distribution of current results, as can be seen from the sketch, if each of the meshes carries a current equal to the loop current; for the currents in all the mesh sides cancel, with the exception
of the outer edges. However, the meshes may be approximated by magnetic dipoles; the smaller the meshes are taken, the more nearly exact the approximation. A current-carrying loop, it is seen, may thus always be replaced by a dipole sheet. By a similar argument, volume distributions of current may be replaced by volume distributions of dipoles.

The energy associated with a current distribution may be found by calculating the work necessary to establish the equivalent dipole distribution. Each dipole may be imagined to be brought to its final location from infinitely far away, just as electric charge distributions are imagined to be built up piecemeal. Since the magnetic moment \( \mathcal{M} \) of a dipole (unlike electric charge) is a vector quantity, its direction as well as its placement must be correct, and each dipole must be turned to have correct orientation once in its proper position. The work done in turning a dipole against its torque through an angle \( \delta \theta \) is

\[
\delta W = T \delta \theta \tag{5-28}
\]

which may be evaluated as

\[
W = |\mathcal{M} \times \mathbf{B}| \delta \theta = i \Delta S B \sin \theta \delta \theta \tag{5-29}
\]

where \( \Delta S \) is the surface area of the elementary dipole mesh. In turning through an angle \( \theta \), the field energy acquired by the dipole is found by integrating:

\[
\Delta W = i \int \Delta S B \sin \theta \, d\theta
\]

which gives

\[
\Delta W = i(\Delta S B \cos \theta) \tag{5-30}
\]

\[
= i \phi_d \tag{5-31}
\]

where \( \phi_d \) is the magnetic flux linking the small mesh equivalent to a dipole, and \( i \) is the mesh current.

From Equation (5-30), it is seen that a dipole possesses an invariant minimum amount of energy—which might as well be taken to be zero—when the dipole moment and flux density vectors are at right angles. It is possible to move a dipole from a region of no magnetic flux to a region of high flux density, without doing any work, as long as this orthogonality condition is maintained. Consequently, any magnetic field distribution may be built up by moving dipoles to their proper locations from infinitely far away...
without doing any work in the process and then turning each dipole to have correct orientation. This is in direct contrast to electric fields, where any elementary charge may be rotated through an arbitrary angle without doing work, but translational motions in general involve expenditure of energy.

The energy of a small dipole placed in an external field \( \mathbf{B} \) may be written, from Equation (5-31), as

\[
W_{\text{dipole}} = \mathbf{B} \cdot \mathbf{S} i
\]

or

\[
W_{\text{dipole}} = \mathbf{B} \cdot \mathbf{M}
\]

(5-32)

This relationship leads some authors to define unit flux density \( \mathbf{B} \) as that in which unit energy is required to rotate a dipole of unit moment through a right angle.

A current distribution of finite size must necessarily be replaced by a large number of equivalent dipoles or meshes. The energy of the entire distribution may always be found by summing the individual energies, each calculated by finding the energy required to realign every dipole in turn. However, this process is very tedious because the reorientation of each dipole affects the flux \( \phi_k \) of all the others. The energy is much more simply calculated in a manner similar to that employed in electrostatics: all the dipoles, represented by little loops, are imagined to be assembled in their correct locations and orientations, but with the loop currents all held at zero. Clearly, no energy is required to do so. All the currents are then raised toward their final values. The fluxes all must rise in proportion to the currents. At some stage in the process, the dipole fluxes will have values \( h \phi_k \) and the currents \( hi_k \), where \( h \) is some number between zero and unity. On raising the currents by the amount \( i_k dh \), an energy increase \( dW \) takes place. From (5-31),

\[
dW = \sum_k h \phi_k i_k dh = h dh \sum_k \phi_k i_k
\]

(5-33)

The total energy may now be found by integration over all values of \( h \):

\[
W = \int_0^1 h dh \sum_k \phi_k i_k
\]

\[
W = \frac{1}{2} \sum_k \phi_k i_k
\]

(5-34)

A similarity with the corresponding expression for electrostatic field energy is apparent.

4. Energy in Terms of the Field Vectors

The energy expression (5-34) is often inconvenient for actual calculation because its use demands that all the actual currents in a physical situation be replaced by imaginary dipole meshes. In some special circumstances this
requirement can be met easily; more generally, however, it is easier to carry out calculations in terms of the field vectors than imaginary dipole distributions. To obtain an expression for the magnetic field energy in terms of vector quantities descriptive of the field, the magnetic vector potential may be introduced in Equation (5-34) by using (5-10):

\[ W = \frac{1}{2} \sum_k A_k \cdot dl_k 
\]

(5-35)

The product \( dl_k i_k \) may be replaced by \( J_k dU_k \), yielding

\[ W = \frac{1}{2} \sum_k A_k \cdot J_k dU_k \]

(5-36)

where the integration is carried out over each current-carrying region in the field. Exactly the same meaning is conveyed, but with simplified notation, by writing

\[ W = \frac{1}{2} \int A \cdot J dU \]

(5-37)

with the integration understood to extend over all space. Since the integrand is zero in all portions of space not included in Equation (5-36), no additional contribution accrues from extending the region of integration beyond that occupied by currents.

All the foregoing argument has been based on the assumption that no electrostatic fields need be considered. If that assumption is discarded, the same line of development may be followed, but with two significant points of difference: the energy in the field is no longer purely magnetic but consists of a magnetic and an electric component, \( W = W_{\text{magnetic}} + W_{\text{electric}} \); and the currents considered are no longer purely conduction currents but may contain a displacement component as well. As indicated by the Maxwell equations (MD-2) and (MI-2), conduction current density \( J \) is magnetically indistinguishable from displacement current density \( D \). In the presence of both kinds of current density, the energy expression (5-37) must therefore read

\[ W_{\text{magnetic}} = \frac{1}{2} \int A \cdot \left( J + \frac{\partial D}{\partial t} \right) dU \]

(5-38)

The magnetically stored energy may also be expressed in terms of the vectors \( B \) and \( H \). To do so, it is convenient to rewrite Equation (5-38) as

\[ W_{\text{magnetic}} = \frac{1}{2} \int A \cdot \text{curl} H dU \]

(5-39)

and use the vector identity

\[ \text{div} (H \times A) = A \cdot \text{curl} H - H \cdot \text{curl} A \]
to obtain

\[ W_{\text{magnetic}} = \frac{1}{2} \int \text{div} (\mathbf{H} \times \mathbf{A}) \, dU + \frac{1}{2} \int \mathbf{H} \cdot \text{curl} \, \mathbf{A} \, dU \]

\[ = \frac{1}{2} \oint \mathbf{H} \times \mathbf{A} \, dS + \frac{1}{2} \int \mathbf{H} \cdot \mathbf{B} \, dU \]  \hspace{1cm}(5-40)

Here the surface of integration for the closed surface integral must extend over an infinitely large surface bounding all space. At a great distance from all currents, the field intensity \( \mathbf{H} \) diminishes rapidly, while the vector potential \( \mathbf{A} \) approaches at most a constant value. The surface integral for an infinitely large surface of integration may thus be shown to be zero, so that

\[ W_{\text{magnetic}} = \frac{1}{2} \int \mathbf{B} \cdot \mathbf{H} \, dU \]  \hspace{1cm}(5-41)

The similarity of this argument to the corresponding development for electric fields is apparent at once. As in the case of the electric field, it is often convenient to imagine magnetically stored energy to be distributed throughout space with the density \( \frac{1}{2} \mathbf{B} \cdot \mathbf{H} \) joules per cubic meter, and to speak of \( \frac{1}{2} \mathbf{B} \cdot \mathbf{H} \) as the energy density of the magnetic field.

5. Explicit Calculation of Vector Potential

The vector Poisson equation (5-5) may be solved formally in a manner analogous to the electrostatic scalar equation (1-66). In fact, in Cartesian coordinates, (5-5) is nothing more than a combination of three scalar equations, one for each spatial component of the vector potential, for in this coordinate system there is no interdependence between components. Written out in detail,

\[ \nabla^2 A_x = -\mu \left( J_x + \frac{\partial D_x}{\partial t} \right) \]

\[ \nabla^2 A_y = -\mu \left( J_y + \frac{\partial D_y}{\partial t} \right) \]  \hspace{1cm}(5-42)

\[ \nabla^2 A_z = -\mu \left( J_z + \frac{\partial D_z}{\partial t} \right) \]

appear as the three independent scalar equations expressed by (5-5). Each of these component equations is formally identical to the Poisson equation of electrostatics and may be solved in exactly the same manner. Any general current distribution may be viewed as an assembly of current elements \( (\mathbf{J} + \mathbf{D}) \, dU \), just as a charge distribution may be thought to be composed of charge elements \( \rho \, dU \). The potential due to one current element at a distance \( r \), Fig. 5.7, has the x-component

\[ dA_x = \frac{\mu}{4\pi r} \left( J_x + \frac{\partial D_x}{\partial t} \right) \, dU \]  \hspace{1cm}(5-43)
and similar \( y \)- and \( z \)-components. All three components in combination may be expressed by the vector equation

\[
d\mathbf{A} = \frac{\mu}{4\pi} \mathbf{J} \frac{dU}{dt} + \frac{\mu}{4\pi} \frac{\partial \mathbf{D}}{\partial t} \frac{dU}{dt}
\]  

(5-44)

The total vector potential at any point \( P \) in space may be found as the sum of the potentials due to each current element:

\[
\mathbf{A} = \frac{\mu}{4\pi} \int \left( \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \right) \frac{dU}{r}
\]  

(5-45)

the integration being extended over all space. This equation, just exactly like its electrostatic counterpart (1-50), is valid for current distributions in a homogeneous, unbounded space.

The formal solution (5-45) suffers from a disadvantage common to many integral formulas: it frequently leads to integrations difficult to carry out. The seemingly simple problem of an infinitely long wire, for example, leads to substantial mathematical complication if direct solution for the vector potential is attempted.

A problem involving minimal mathematical difficulty, while physically useful, is finding the field surrounding the current-carrying circular loop of Fig. 5.8, where the loop radius \( R \) is assumed to be much greater than the wire radius \( a \). With the loop assumed to lie in the plane \( z = 0 \) of a cylindrical coordinate system, the current density in the loop is a purely tangential vector. From (5-45), it may be concluded that the vector potential at any point \((r, \theta, z)\) is also a purely tangential vector whose magnitude depends on the coordinates \((r, z)\) but not on the angular position \(\theta\). It will therefore suffice to calculate \( A_{\theta}(r, 0, z) \):

\[
A_{\theta} = \frac{\mu}{4\pi} \int_{L} i \cos \theta \frac{dl}{s}
\]  

(5-46)

The distance \( s \) from the point \((R, \theta, 0)\) to the point \((r, 0, z)\) is given by

\[
s = \sqrt{z^2 + r^2 + R^2 - 2rR \cos \theta}
\]  

(5-47)
so the vector potential is found on integrating

\[ A_\theta = \frac{\mu i}{4\pi} R \int_0^{2\pi} \frac{\cos \theta \, d\theta}{\sqrt{z^2 + r^2 + R^2 - 2rR \cos \theta}} \]  \hspace{1cm} (5-48)

This integral is not one of the elementary forms found in tables. It may be evaluated by making the substitutions

\[ k^2 = \frac{4Rr}{z^2 + (R-r)^2} \quad \text{and} \quad \theta = \pi - 2\phi \]  \hspace{1cm} (5-49)

whereupon Equation (5-48) may be transformed into

\[ A_\theta = \frac{\mu i}{2\pi} \sqrt{\frac{R}{r}} \left( \frac{2 - k^2}{k} \right) K(k) - \frac{\mu i}{k\pi} \sqrt{\frac{R}{r}} E(k) \]  \hspace{1cm} (5-50)

where \( K(k) \) and \( E(k) \) are the definite integrals

\[ K(k) = \int_0^{\pi/2} \frac{d\phi}{\sqrt{1 - k^2 \sin^2 \phi}} \]  \hspace{1cm} (5-51)

\[ E(k) = \int_0^{\pi/2} \sqrt{1 - k^2 \sin^2 \phi} \, d\phi \]

called the complete elliptic integrals of the first and second kinds, respectively. No general algebraic expressions can be given for the corresponding indefinite integrals; the definite integrals, however, are of sufficient interest to form part of the standard corpus of tabulated transcendental functions (along with the trigonometric, logarithmic, exponential, and several other families).
A short table of these elliptic integrals appears in Appendix III, together with the combination

\[ h(k) = \left( \frac{2}{k} - k \right) K(k) - \frac{2}{k} E(k) \]  

(5-52)

Using this latter combined function, the vector potential of a loop becomes simply

\[ A_r = \frac{\mu i}{2\pi} \sqrt{\frac{R}{r}} h(k) \]  

(5-53)

Since the flux density \( B \) is finite everywhere, Equations (5-1) and (5-10) require the vector potential to have zero value everywhere along the \( z \)-axis, as Equation (5-53) indeed gives in the limit. This reference position is implicitly introduced by adopting the symmetrical form of Equation (5-46). Interpreting the physical significance of \( A \) in accordance with Equation (5-10), in a manner similar to the infinite wire problem, the vector potential here measures the amount of flux through a sector bounded by an arc of unit length. A number of vector equipotentials, corresponding to flux lines, are shown in Fig. 5.9.

6. Neumann’s Formula for Inductances

In engineering applications, current distributions made up of slim paths or loops are frequently encountered. It will now be shown that the energy in the magnetic field caused by such a distribution may be expressed as a combination of the path currents and certain geometrical constants. All the currents
will be assumed to flow in a finite (or at least enumerably infinite) number of distinct paths, each of which may be decomposed into a sequence of current elements \( i \, dl \). The total energy of such a distribution is given by

\[
W = \frac{1}{2} \int \mathbf{A} \cdot \mathbf{J} \, dU
\]  

(5-37)

which may be written as

\[
W = \frac{1}{2} \sum_{m=1}^{N} \int_{m} A \cdot dl \, i_{m}
\]

(5-54)

the summation being carried out for the \( N \) closed contours comprising the current distribution. Each contour current \( i_{n} \) contributes to the magnetic vector potential; thus, by (5-45),

\[
A = \frac{\mu}{4\pi} \sum_{n=1}^{N} \int_{n} \frac{i_{n} \, dl_{n}}{r}
\]

(5-55)

so that the total energy may be written as the double sum

\[
W = \sum_{m=1}^{N} \sum_{n=1}^{N} \frac{\mu}{8\pi} i_{m} i_{n} \int \int \frac{dl_{m} \cdot dl_{n}}{r}
\]

(5-56)

Each term in the double summation is seen to consist of a product of currents \( i_{m} i_{n} \) and a purely geometric integral expression not dependent on any electromagnetic quantity. It is usual to define geometrical coefficients called inductances by

\[
L_{mn} = \frac{\mu}{4\pi} \int \int \frac{dl_{m} \cdot dl_{n}}{r}
\]

(5-57)

The expression (5-57) is known as Neumann's formula. It is immediately evident from it that \( L_{mn} = L_{nm} \). Careful note should be given to the fact that this formula relates to closed contours only since integration over two closed paths is required.

As an example of the use of Neumann's formula, let the mutual inductance between two circular loop conductors, as shown in Fig. 5.10, be required. Applying Equation (5-57), the required inductance is

\[
L_{12} = \frac{\mu}{4\pi} \int_{0}^{2\pi} R_{1} \int_{0}^{2\pi} R_{2} \frac{\cos (\theta_{2} - \theta_{1}) \, d\theta_{2}}{\sqrt{R_{1}^{2} + R_{2}^{2} - 2R_{1}R_{2} \cos (\theta_{2} - \theta_{1})}}
\]

(5-58)

This double integral may be put in the form of two elliptic integrals by means of the substitutions (5-49). After some algebraic manipulation, there is obtained

\[
L_{12} = \mu \sqrt{R_{1}R_{2}} h(k)
\]

(5-59)

where the function \( h(k) \) is as defined in Equation (5-52), but with the substitutions \( R = R_{1} \) and \( r = R_{2} \).
To find the self-inductance of a loop using Neumann's formula, the two contours of integration are made the same; there results immediately

$$L_{\text{self}} = \mu R h(1)$$

(5-60)

for it is readily verified that \(k = 1\) for \(R_1 = R_2, z = 0\). On consulting the table in Appendix III, however, it is seen that \(h(1) = \infty\), so that the self-inductance of a circular loop is infinite! The physical absurdity of this result derives from the implicit assumption that a finite current \(i\) can flow in an infinitesimally thin wire. From Equation (5-7) it is seen that the flux density near the surface of such a wire approaches infinity; hence not the result, but the problem, lacks physical meaning. A more realistic approximation is arrived at by initially neglecting the flux inside the wire and performing the first contour integration (following the mean current path) along the wire center line, \(r = R\); but taking the second integration along the inner surface of the wire, \(r = R - a\), where \(a\) is the wire radius. The result then is

$$L = \mu R h \left(1 - \frac{a^2}{8R^3}\right)$$

which for small \(a\), is approximated by

$$L = \mu R \left(\log \frac{8R}{a} - 2\right)$$

(5-61)

A correction to allow for the internal flux may be made, assuming the field distribution inside the wire to follow Equation (5-9).
7. Self-Inductance of a Square Loop

As a second example of inductance calculations using Neumann’s formula, the self-inductance of a square loop made of fine wire, shown diagrammatically in Fig. 5.11, will be calculated. Let the loop have a side length $a$ and a wire radius $r_0$. The inductance may be calculated by applying Neumann’s formula and evaluating the closed contour integrals in sections. The scalar product of vector line elements in (5-57) may first be converted into a product of lengths and the cosine of the included angle, yielding

$$L_{mn} = \frac{\mu}{4\pi} \int \oint \frac{dl_m dl_n \cos \theta}{r}$$

(5-62)

Because the wire loop is square, it is noted that the included angle $\theta$ only assumes values $0, \frac{1}{2}\pi, \pi$, so that $\cos \theta$ may be $+1, 0, -1$. In other words, nonvanishing contributions to the total integral are only obtained where the two integrations are carried out over parallel sides of the square. There are four pairs of parallel sides, so that

$$L = \frac{\mu}{\pi} \int_A^B \int_A^B \frac{dx_1 dx_2}{r} - \frac{\mu}{\pi} \int_A^B \int_C^D \frac{dx_1 dx_2}{r}$$

(5-63)

The mathematical aspect of this problem is evidently to evaluate the integrals, both special cases of the general form (which may be found in tables)

$$\frac{1}{2a} \int_0^a \int_0^a \frac{dx \, dy}{\sqrt{b^2 + (x - y)^2}} = \log \left( \frac{a}{b} + \sqrt{\frac{a^2}{b^2} + 1} \right) - \sqrt{\frac{b^2}{a^2} + 1} + \frac{b}{a}$$

(5-64)

As in the case of the round wire loop, it is again seen that the integrals diverge
if both contours of integration are the same. The reason is once again the implicit requirement that a finite current must flow in an infinitesimally thin wire, which is not physically meaningful. The way out is again to perform the second integration along a contour with the highest value of magnetic field actually physically reached, i.e., along the wire surface (see Fig. 5.2). In this case, the separation \( b \) of contours in Equation (5-64) assumes values equal to the wire radius \( r_0 \) for the first integral of (5-63), and to the loop breadth \( a \) for the second integral. There results the inductance value

\[
L = \frac{2\mu a}{\pi} \left( \log \frac{p + \sqrt{1 + p^2}}{1 + \sqrt{2}} + \frac{1}{p} - 1 + \sqrt{2} - \frac{1}{p} \sqrt{1 + p^2} \right)
\]

(5-65)

where \( p = a/r_0 \), and the stored energy associated with flux inside the wire itself has been neglected. Its contribution may be assessed to a reasonable degree of accuracy, as in the case of a round loop, by assuming the internal flux to be distributed symmetrically, just as in an isolated infinitely long wire.

Equation (5-65) is in any case valid only for reasonably large \( p \). In many practical cases, the wire diameter is considerably smaller than the loop dimensions; with the resulting large values of \( p \), the further simplification may be made

\[
L = \frac{2\mu a}{\pi} \left( \log \frac{2p}{1 + \sqrt{2}} - 2 + \sqrt{2} \right)
\]

(5-66)

The mutual and self-inductances of other figures made up of rectilinear segments are calculable by a similar process; because they are of some practical importance, tables of calculated inductances exist for the lower-order regular polygons. It should be clear that the calculation of inductances for any figures made up of straight-line segments, however, may be accomplished by writing the iterated integrations of Neumann's formula in segments and evaluating the individual integrals. Except for a number of special cases, the integrations turn out to be difficult. It is useful, therefore, to examine next a more general numerical formulation of the problem, which is not subject to the same restrictions. The existing tables will then furnish a way of checking the accuracy and reliability of numerical evaluation.

8. Neumann's Formula in Summation Form

The iterated contour integrations of Neumann's formula may be viewed as limits of iterated summations. If this view is adopted, an approximate method of evaluating the formula immediately suggests itself: instead of requiring the actual limit of summations, one might rest content with finite summations taken so as to assure a suitable number of accurate significant figures. To
achieve this, some convenient number of points on either contour may be specified, and the path between each pair of adjacent points assumed straight. In effect, the arbitrarily shaped contours are thereby replaced by general polygonal figures composed of a number of linear segments. Let one of the contours be subdivided into \( N \) segments and the other into \( M \). Let the segment lengths \( \Delta s_k \) be chosen so short that no significant error will accrue from assuming them to be straight. The mutual inductance between contours 1 and 2 may then be written approximately as

\[
L_{12} = \frac{\mu_0}{4\pi} \sum_{j=1}^{N} \sum_{k=1}^{M} \frac{\Delta s_j \cdot \Delta s_k}{r_{jk}}
\]  
(5-67)

where \( r_{jk} \) represents the distance between midpoints of the two segments. The summations are understood to cover the entire closed contours. This approximation is clearly only valid if the contours do not intersect, so that the midpoint-to-midpoint distance always has a magnitude greater than the segment lengths. Should it be desired to calculate self-inductance of an odd-shaped loop, this restriction is naturally quite inconvenient. It may be removed, however, by using (5-67) as it stands, except for the terms that give rise to \( j = k \), so that \( r_{jk} = 0 \); instead of this incorrect value, a closer approximation, readily obtainable from (5-64), is then used.

Each term of (5-67) involves the scalar product of two line segment vectors. If the summations are to be evaluated by hand, it will probably be found easiest to rewrite this dot product as the product of the segment lengths and the cosine of the included angle. Except in very unusual cases, however, the summations will in practice be carried out on a digital computer. Since a cosine subroutine typically requires about 12 to 15 times as much computer time as a single multiplication, significant savings in time are effected (and programming simplified as well) by writing

\[
\Delta s_j \cdot \Delta s_k = (\Delta x_j \Delta x_k) + (\Delta y_j \Delta y_k) + (\Delta z_j \Delta z_k)
\]

which requires only three multiplications. (Additions typically require very little time.) The denominator of each term in (5-67) may be evaluated similarly. If the segment midpoints are denoted by \((x_j, y_j, z_j)\) and \((x_k, y_k, z_k)\), \( r_{jk} \) is given by

\[
r_{jk} = \sqrt{(x_j - x_k)^2 + (y_j - y_k)^2 + (z_j - z_k)^2}
\]

(5-69)

Thus, evaluation of (5-67) may be accomplished if the end points of each segment are known; the midpoints are easily calculated from end point coordinates when required.

To illustrate the method, a programme to evaluate mutual inductances between triangular contours has been written and appears in Appendix II. A flow chart for this computation is shown in Fig. 5.12. This programme first reads from data cards the coordinates of the three vertices of either triangle. In order to simplify programming, it is convenient to operate be-
between four rather than three vertices, the coordinates of the fourth and first being identical; this artifice obviates any necessity for writing special instructions for the final contour segment that bridges the first and last points. For this reason, the data include an extra vertex. The data cards also specify how many segments each side of the triangle is to be subdivided into, by giving the segment numbers that fall at the vertices. Once the data have been read, each triangle side is subdivided into the required number of equal segments. Segment end point coordinate values are stored in suitable arrays; the vertices are stored at their specified locations, and the intervening array elements subsequently filled in equal steps. For example, the data might specify the first and second vertices of a triangle to fall at (0, 0, 0) and (10, 0, 20) and require these to be the first and eleventh end points. The intervening points would then be calculated as (1, 0, 2), (2, 0, 4), and so on.
Sec. 9  MAGNETIC VECTOR POTENTIAL AND ENERGY  163

With the end points of all segments known, summation may proceed. For each segment pair, the dot product (5-68) is found and divided by the midpoint distance (5-69). The quotients are accumulated into the sum required by (5-67) and eventually multiplied by constant factors, as required.

To assess the accuracy obtainable with this method, the mutual inductance between two equilateral triangular contours, parallel to each other and separated by a distance equal to half the side length, has been calculated. With either triangle subdivided into 72 equal segments, the result agrees with tables to the four tabulated significant figures. With either triangle subdivided into 36 segments, an error of about 0.1 percent is found. The accuracy is naturally dependent on the relative size and separation of the contours, and such numbers can only be taken as indicative guides, not in any sense as accuracy guarantees.

The programme shown in Appendix II is far from economical in terms of computer time and needs to be refined if extended use is to be made of it, e.g., in constructing tables. For example, the dot product evaluation is performed separately for each term in the summations; yet for triangular contours, there exist only nine possible distinct segment dot products. Significant savings will result from programme modification to eliminate these unnecessary recalculations. While there is little point in worrying about a few seconds more or less in a programme that is to be used only once, repeated use of an uneconomic programme can swallow up machine time amazingly fast.

9. Direct Inductance Calculations

On occasion, it is relatively easy to determine the magnetic field surrounding current-carrying contours by some means other than direct evaluation of Equation (5-45). For example, superposition of already known solutions may often produce new ones. In such cases, Neumann’s formula is at times unattractive in comparison with simpler techniques for calculating inductance. One such method is quite similar to the methods used to determine capacitance in electrostatics. It is based on the following argument. Let Equations (5-56) and (5-57) be combined to give

$$W = \frac{1}{2} \sum_m \sum_n L_{mn} i_m i_n \quad (5-70)$$

Comparing with Equation (5-34), it is seen that this is only possible for any arbitrary set of currents if

$$\phi_n = \sum_m L_{mn} i_m \quad (5-71)$$

On differentiation with respect to one of the currents, say $i_k$, there obtains

$$\frac{\partial \phi_n}{\partial i_k} = L_{nk} \quad (5-72)$$
Since the permeability \( \mu \) has been assumed constant, the inductance is independent of current. Hence the partial derivative of Equation (5-72) must also be constant, and it is given by

\[
L_{nk} = \frac{\phi_n}{i_k}
\]  

(5-73)

if all currents other than \( i_k \) are assumed to have been set to zero. Interpreted physically, this equation states that the inductance \( L_{nk} \) represents the amount of flux per unit current linking the \( n \)th contour, if none but the \( k \)th current is made to flow.

As an example of direct inductance calculation, let the self-inductance per unit length of a very long parallel-wire transmission line be found. (This problem is also not difficult to handle by Neumann's formula, as may readily be verified.) Let the line be composed of two wires of radius \( a \), with a wire-to-wire spacing \( h \) much larger than the wire radius (Fig. 5.13). At a point \( P \) not inside either wire, the magnetic vector potential is readily found by superposing the component vector potentials due to the individual wires, as given by Equation (5-13). There results

\[
A = \frac{\mu i}{2\pi} \left[ \left( \frac{1}{2} + \log \frac{r_1}{a} \right) - \left( \frac{1}{2} + \log \frac{r_2}{a} \right) \right]
\]

(5-74)

or, simplifying,

\[
A = \frac{\mu i}{2\pi} \log \frac{r_1}{r_2}
\]

(5-75)

The direction of the vector \( A \) is everywhere coincident with the direction of the wires themselves. To find the self-inductance of a unit length of line, a rectangle of unit length in the line direction may be laid out, as shown dotted in Fig. 5.13. The flux linking a unit length of line may then be calcu-
lated by
\[
\phi = \int A \cdot dr
\]
\[
= \int_{BC} A \cdot dr + \int_{DA} A \cdot dr
\]
(5-10)
since the integral has zero value along sides AB and CD, the vectors A and dr being orthogonal. The contributions of paths BC and DA are equal, and one readily finds
\[
\phi = \frac{\mu i}{\pi} \log \frac{h}{a}
\]
(5-76)
It should be noted that once again the path of integration has been taken along the surface, rather than along the center, of either wire. From a mathematical point of view, this is necessary because expression (5-13) for the vector potential is only valid outside a current-carrying wire; physically, it amounts to ignoring all the flux that crosses through the wire. The result (5-76) is therefore approximate, the inaccuracy depending on the relative sizes of h and a. A more refined calculation may readily be performed by adding the flux so far neglected; it may be found by using the expression for vector potential inside a wire, Equation (5-12). Inductance is subsequently found from (5-76) by dividing by the current:
\[
L = \frac{\mu}{\pi} \log \frac{h}{a}
\]
(5-77)
The mathematical steps of this process, it may be seen, are in principle similar to the iterated integrations of Neumann's formula. The advantage lies in direct physical interpretation of each step, with the possibility of introducing approximations step by step as needed, thereby avoiding mathematical entanglement.

10. Inductances of Solenoidal Coils

The evaluation of Neumann's formula, or use of the direct method, is quite difficult in many cases of practical interest, for evaluation of the necessary integral expressions often becomes very complicated. A very powerful calculation method for such cases is provided by Equation (5-70), which permits the inductance of a complicated system of conductors to be found from the more readily computed inductance values of its individual parts. A simple example is furnished by a multiturn solenoidal coil, for which the direct evaluation of inductance is very complicated indeed, the paths of integration being multiturn spirals. However, such a coil may be thought of as nearly equivalent to a series of flat circular turns, as indicated in Fig. 5.14. This approximation is fairly good as long as the coil pitch is reasonably
close, so that the current flow is essentially circumferential. Because all the turns are connected in series, they must all carry equal current. The general expression for stored energy,

\[ W = \frac{1}{2} \sum_n \sum_m L_{mn} i_m i_n \]  

(5-70)

then becomes

\[ W = \frac{1}{2} i^2 \sum_m \sum_n L_{mn} \]  

(5-78)

It is immediately recognized on comparing the two equations above that the total self-inductance of the multiturn solenoid is given by

\[ L = \sum_n \sum_m L_{mn} \]  

(5-79)

Using Equations (5-59) and (5-61), this sum is readily evaluated; but unless the number of turns is quite small, the sums are long and evaluation by hand calculation not to be recommended. In evaluation by digital computer, each mutual term \((m \neq n)\) requires subroutine calls to find values of the elliptic integrals \(K(k)\) and \(E(k)\). It is desirable to minimize the number of times these subroutines are employed, lest computation time become excessive. Because mutual inductances are known to be reciprocal, \(L_{mn} = L_{nm}\), the number of terms may first of all be halved by rewriting Equation (5-79) in the form

\[ L = NL_{\text{self}} + 2 \left( \sum_{n=2}^{N} L_{1n} + \sum_{n=3}^{N} L_{2n} + \cdots + L_{(N-1),N} \right) \]  

(5-80)

where \(N\) represents the total number of turns. The total number of times that evaluation of Equation (5-59) is required has been reduced thereby, but it is still proportional to \(N^2\). If the coil pitch is uniform, the axial distance between the \(m\)th and \(n\)th turns is \(|m - n| p\), where \(p\) represents the distance between successive turns. The mutual inductance between any two turns, however, depends on the turn radii and this separation; hence all terms \(L_{mn}\) for which \(|m - n|\) is the same must be equal; e.g., \(L_{12} = L_{23} = L_{87}\), etc. The first term in each summation of (5-80) has \(|m - n| = 1\), the second term \(|m - n| = 2\), and so on. Hence it is possible to rewrite the latter equation, simplifying computation once again, as

\[ L = NL_{11} + 2(N-1)L_{12} + 2(N-2)L_{13} + \cdots + 2L_{1N} \]  

(5-81)

\[ = NL_{11} + 2 \sum_{n=2}^{N} (N + 1 - n)L_{1n} \]

This expression, finally, only requires evaluation of as many mutual inductances as there are turns. It is therefore entirely practical for calculating inductance values of real coils with numerous turns.
It might be observed also that the development above is not in any way restricted to coils of circular cross section, but may equally well be applied to coils of other shapes; all that is required is that calculation of turn-to-turn inductances be possible. There is not even any necessity for an analytic expression for the latter since a subroutine to evaluate them numerically as required can be written using a summation approximation to Neumann's formula.

**READINGS**

Most electromagnetics texts include some development of the magnetic field energy and vector potential, although often the analogy with electrostatics is resorted to and the treatment is very brief. A reasonably detailed treatment is that by Carter (1), while Elliott (2), Reitz and Milford (3), or Artley (4) may be consulted for short explanations. Initial difficulties in interpreting and visualizing the vector potential may be alleviated by the article by McRae (5). Extensive use is made of the vector potential for inductance and field calculations by King (6), while Boast (7) also describes a broad range of problems, particularly mathematically simple ones, and will undoubtedly be found the easier to read. The chapter in Javid and Brown (8) deals with capacitance and inductance calculations side by side and thereby achieves an attractive unification of methods. A clear development of the properties of magnetic fields, including some examples of inductance calculations, may be found in Tralli (9), and a quite comprehensive treatment is in Jackson (10). Hague's account (11) is one of the most pleasant to read and quite detailed, but the use of cgs units may prove inconvenient to readers not accustomed to the extra factors $4\pi$!

The theoretical notion of self-inductance is closely examined as to its physical significance, and regarded as a special case of mutual inductance, in a paper by Graneau (12).

Many of the difficulties encountered in working with vector potentials are resolvable by examining a few physical problems closely. Several examples designed to exhibit parallelism between electric and magnetic potential problems are solved in detail by Haus and Penhune (13), while Batygin and Toptygin (14) once again provide a wealth of problems, but with considerably briefer solutions that at times are reduced to a mere answer.

Finally, an introductory treatment of elliptic functions may be found in the little book by Bowman (15), while comprehensive tables of calculated inductance values are in Grover (16).

5.2 Find the magnetic vector potential as a function of radius for an infinitely long iron tube carrying direct current in the z-direction. Plot $A$ and $H$ as functions of the radius for some reasonable range of radii.

5.3 Plot the magnetic field of two parallel two-conductor bundles such as illustrated in Fig. 2.7.

5.4 Show that the mutual inductance between two parallel-wire lines (as in Fig. 5.13) is zero if the lines cross at right angles.
5.4 Find an expression for the self-inductance of a rectangular plane wire loop with side lengths $a$ and $b$ meters.

5.5 Modify the Neumann’s formula programme of Appendix II so as to make it applicable to general contours (e.g., read in as data cards). Use it to find the mutual inductance of two identical fine wire rings and compare with the known correct result (5-59).

5.6 Modify the Neumann’s formula programme of Appendix II so as to make it applicable to self-inductance calculations. Check it by evaluating the self-inductance of a circular loop.

5.7 By means of the series expansions of $K(k)$ and $E(k)$, show that for small moduli $k$,

$$h(k) = \frac{\pi}{16} k^3 + \cdots$$

and use this to prove that the magnetic field of a very small current-carrying loop is given by

$$A = \frac{\mu}{4\pi s^3} M \times s$$

where $s$ is the distance from the loop to the point of measurement.

5.8 A solenoidal coil is made of round wire 0.050 in. in diameter, with insulation of negligible thickness, by winding 100 turns on a cylindrical form 1.75 in. in diameter. Calculate the inductance of this coil by

(a) neglecting the internal flux of each turn of wire;

(b) including the internal flux, on the not entirely well-founded assumption that it is symmetrically distributed in each wire.

5.9 A coil is made of round wire 0.047 in. in diameter, with negligible insulation thickness, by winding 40 turns in a flat spiral. The innermost turn has a diameter of 2.00 in., and the outermost 5.75 in. Find the inductance of this coil.

5.10 Prove that the ratio of inductances of two coils is equal to the square of the ratio of their turns, provided that all dimensions of the coils are the same.

5.11 Prove that the ratio of inductances of two coils is equal to the ratio of their linear dimensions, provided they have the same shape and number of turns.

5.12 In electronic apparatus, multiple coils wound on a common form are frequently employed. Find the mutual inductance of two coils, either of 20 turns and 1 cm long, wound on a cylindrical form of 15 mm diameter, as a function of distance between coils. Assume the coils to be wound without spacing between turns.

5.13 Two current-carrying loops are placed in each other’s neighborhood. Show that they are in general subject both to a force and a torque. Find expressions for both, assuming the mutual and self-inductances to be calculable for any spatial orientation. (Hint: Use the principle of virtual work.)

5.14 Find the axial force between the coils of Problem 5.12, using the results of Problem 5.13.
5.15 Two triangular wire loops are placed in space so that the vertices of one triangle are (0, 0, 0), (8, 0, 0), and (0, 8, 0) and the vertices of the other are (0, 0, 6), (0, 12, 0), and (6, 0, 0). Find the mutual inductance between these loops, assuming the dimensions to be given in inches.

5.16 Prove that the tangential component of $A$ is continuous across an interface between media.

5.17 Derive an expression for the mutual inductance of two similar plane square loops located in parallel planes above each other.

5.18 Use the results of Problems 5.4 and 5.17 to find the self-inductance of a 50-turn winding that occupies a 1-in. length of square 1-in. winding former.
Magnetostatic Problems in Engineering

A very great many practical devices used in electrical engineering and technology rely on energy storage in a magnetic field; for example, transformers, rotating machines, magnetic recording devices, and other apparatus involving ferromagnetic materials. Numerous practical field problems arising in the analysis and design of these and other devices can frequently be solved to a reasonable accuracy by treating the field as two-dimensional. Others, although not really two-dimensional at all, are also often analyzed on this basis simply because magnetic fields are essentially vector rather than scalar in nature, and grave difficulties are encountered in all but the simplest problems if all three vector components must be depicted in three-dimensional space. Some of the methods introduced in Chapter 2 in connection with two-dimensional electrostatic fields will be extended and applied to magnetic fields in this chapter. Because the mathematical treatment of either field resembles the other strongly, the extensions may in turn be employed to solve more complicated electrostatic problems, and so forth.
1. Fields of Current-Carrying Conductors

In Chapter 5 the magnetic vector potential resulting from current flowing along a long fine wire was investigated at some length. The solution there obtained may be generalized so as to yield the magnetic field of any other conductor as well, provided the current density is reasonably uniformly distributed over the conductor cross section. No matter what its shape, any long straight conductor may be regarded as a collection of parallel filaments, each carrying some fraction of the total current. Such filaments are in the limit equivalent to fine round wires, and their total fields may be built up of the contributions resulting from all the individual filamentary currents.

For a single round wire, the magnetic vector potential was found in Chapter 5 to be

\[ A = \frac{\mu i}{2\pi} \left( \frac{1}{2} + \log \frac{r}{a} \right) \quad (5-13) \]

everywhere outside the wire. If each component filament of a massive conductor is imagined to carry some current \( di = J \cdot dS \), its magnetic vector potential contribution must be

\[ dA = \frac{\mu J}{2\pi} \log r \, dS + \frac{\mu}{2\pi} \left( \frac{1}{2} - \log a \right) \, di \quad (6-1) \]

Both \( J \) and \( A \) are treated as scalar quantities here, for both only possess longitudinal components along the conductor. The term on the extreme right of Equation (6-1) is immediately seen to be a constant independent of \( r \). As is usual with potential functions, any arbitrary uniform vector may be added to \( A \) without altering the fields, for in the differentiation \( B = \nabla \times A \), the uniform vector vanishes. Thus, the inconvenient constant term may be annihilated by adding an appropriate reference level to \( dA \), i.e., by assuming some return path for the current \( di \), and there is obtained

\[ dA = \frac{\mu J}{2\pi} \log r \, dS \quad (6-2) \]

as the potential contribution of a single filament. Needless to say, it is only determinate up to an additive uniform quantity.

The magnetic field pattern of any current-carrying conductor may now be built up by adding up potential contributions of the form (6-2):

\[ A = \frac{\mu}{2\pi} \iint J \log \sqrt{(x - \xi)^2 + (y - \eta)^2} \, d\xi \, d\eta + A_0 \quad (6-3) \]

The integration is of course carried out over the entire current-carrying cross-sectional area of the conductor, which is imagined to extend infinitely far in the \( z \)-direction. \( P: (x, y) \) denotes the point at which \( A \) is being measured, and \( A_0 \) is an arbitrary additive quantity. It is interesting to observe that
Equation (6-3) could also, with a small amount of additional mathematical difficulty, be derived directly as the two-dimensional form of (5-45).

An example of practical application of (6-3) is readily given. Let it be desired to find the magnetic field surrounding a strip conductor 2a units wide and b units thick, where b is very small, as shown in Fig. 6.1. With the axes oriented as shown, integration in the y-direction follows immediately; also, since \( J \) is uniform, it may be removed from under the integral sign. It remains to evaluate

\[
A(x, y) = \frac{\mu J b}{2\pi} \int_{-a}^{a} \log \sqrt{(x - \xi)^2 + y^2} \, d\xi
\]  

This integral is tabulated. An explicit solution is readily found by using the indefinite integral

\[
\int \log (x^2 + k^2) \, dx = x \log (x^2 + k^2) - 2x + 2k \arctan \frac{x}{k}
\]  

and substituting limits as appropriate. The result is finally

\[
A(x, y) = \frac{-\mu J b}{4\pi} \left[ (x + a) \log \frac{(x + a)^2 + y^2}{a^2} - (x - a) \log \frac{(x - a)^2 + y^2}{a^2} \right.
\]

\[
+ 2y \left( \arctan \frac{x + a}{y} - \arctan \frac{x - a}{y} \right) \]

where constants have been adjusted so as to yield zero potential at the origin. Equation (6-6) is not at all immediately comprehensible physically, and it is
necessary to evaluate the field at a number of points and plot it in order to appreciate its structure. Such a plot appears in Fig. 6.2. As would be expected from the known properties of magnetic fields, flux lines are continuous and curve continuously, except where their direction is sharply altered by the presence of a surface current density—a direct corroboration of Equation (4-67).

2. Finite-Difference Approximations of the Integral Solution

Although expressions of the type encountered above are analytically integrable for a relatively large number of cases (e.g., round and rectangular conductors), the results tend to become quite lengthy and unwieldy to handle. Cases in which irregular shapes are encountered or the current density is not uniform arise not infrequently in practice and lend themselves poorly to purely analytic treatment. In such instances, it is convenient to approximate the integral expression (6-3) by a finite-difference summation. As a matter of convenience, let the finite increments $\Delta x$ and $\Delta y$ be chosen all equal, as indicated for one particular conductor in Fig. 6.3. Equation (6-3) may be written

$$ A(x, y) = \frac{\mu J}{2\pi} \sum_n \int \int_n \log \sqrt{(x - \xi_n)^2 + (y - \eta_n)^2} \, d\xi_n \, d\eta_n $$

(6-7)

where each integration is carried out over one of the area elements $(\Delta x \Delta y) = h^2$. No approximation of any kind is involved in this reformulation.

If the point $P: (x, y)$ is outside the $n$th area element, then it may be argued that the terms $(x - \xi_n)$ and $(y - \eta_n)$ do not change much in integration over the $n$th element and may be regarded as essentially constant. The integral of Equation (6-7) may then be approximated by

$$ \int \int_n \log \sqrt{(x - \xi_n)^2 + (y - \eta_n)^2} \, d\xi_n \, d\eta_n \approx h^2 \log \sqrt{(x - x_n)^2 + (y - y_n)^2} $$

(6-8)

where $(x_n, y_n)$ is the center point of the $n$th square, as shown in Fig. 6.3. Obviously this approximation becomes less and less accurate as $P: (x, y)$ is placed nearer and nearer $(x_n, y_n)$, although the error is in fact surprisingly
small; for \((x - x_n) = h\), for example, it amounts to only 0.65 percent! However, if point \(P\) is placed within the \(n\)th square, the error grows without bound, and a better approximation clearly must be found. In fact it is not difficult to evaluate this special case analytically. Instead of the midpoint value, the average value of the contribution to \(A(x_n, y_n)\) may be inserted in the summation. This average contribution due to current in the same square may be calculated by repeated application of Equation (6-5), yielding

\[
\frac{1}{h^2} \int_0^h \int_0^h \int_0^h \log \sqrt{(x - \xi)^2 + (y - \eta)^2} \, d\xi \, d\eta \, dx \, dy = h^2 \log (0.44705 \ h)
\]

(6-9)

With the understanding that the average term (6-9) is to be inserted instead of approximation (6-8) wherever applicable, Equation (6-7) may be written as

\[
A(x, y) = \frac{\mu I}{2\pi N} \sum_{n=1}^{N} \log \sqrt{(x - x_n)^2 + (y - y_n)^2}
\]

(6-10)

where \(N\) is the total number of squares, and \(I = NJ\) represents the total current. This expression can readily be evaluated by means of a digital computer, given values of \((x, y)\) and all \((x_n, y_n)\). The latter set of points may be regarded as a description of the conductor, for it specifies the distribution of conductive material throughout the \(x-y\) plane.

Equation (6-10) lends itself ill to practical computation since it involves \(N\) uses of both the logarithm and square root subroutines, equivalent in computing time to about 25–30 multiplications, for each point \((x, y)\). A great deal of needless calculation may be eliminated by rewriting the sum of logarithms as the equivalent logarithm of a product:

\[
A(x, y) = \frac{\mu I}{4\pi N} \log \prod_{n=1}^{N} [(x - x_n)^2 + (y - y_n)^2]
\]

(6-11)

A roughly tenfold saving in computation results. It is necessary to proceed with some caution in evaluating (6-11), however. For even a reasonably small problem, say \(N = 40\), computer arithmetic overflows may result if the individual factors in (6-11) are of the order of 5 or 10; and conversely, arithmetic underflows are likely if the terms are less than unity. It will therefore pay to have the programme examine the product from time to time as it is formed, and to take suitable remedial action if an overflow or underflow seems likely. For example, some predetermined large factor (e.g., \(e^{25}\) or \(10^{10}\)) may be divided out whenever the product becomes large, the appropriate amount being later added to the logarithm.

An interesting geometrical notion inherent in (6-11) is encountered in numerous books on two-dimensional fields. If desired, (6-11) may be rewritten, defining a new quantity \(D\) implicitly, as

\[
A(x, y) = \frac{\mu I}{2\pi} \log D
\]

(6-12)
$D$ is called the *geometric mean distance* between $P: (x, y)$ and the cross-sectional shape defined by the points $(x_n, y_n)$, because it is the geometric mean of the distances of all the elements from $P$:

$$D = \frac{2N}{\sqrt{\prod_{n} [(x - x_n)^2 + (y - y_n)^2]}} \quad (6-13)$$

An example of such finite-difference approximate calculation is furnished by the hollow rectangular conductor of Fig. 6.4. As in many other cases, symmetry of the problem permits reduction of the amount of calculation required at the expense of some slight complication of the computer programme. Thus, instead of evaluating the required distances one at a time, it is profitable to take the elements four at a time, as indicated. Similarly, only the field distribution in one quadrant need be calculated.

A computer programme to evaluate the vector potential over the first quadrant of a symmetrical distribution appears in Appendix II. This programme has been used to solve the problem posed in Fig. 6.4, with the results shown in Fig. 6.5. The programme itself is relatively general and permits solution of other symmetrical problems since the conductor description is not built into the programme indexing but is read from data cards. A safeguard against arithmetic overflows, but not against underflows, is provided; it might be of interest to note that in evaluating the field shown in Fig. 6.5, products of the order of $10^{138}$ are encountered!

**3. Current Parallel to a Permeable Surface**

Because the integral solution (5-45) of the vector Poisson equation (5-5) is only valid in a large homogeneous space of uniform permeability, no attempt was made to include magnetic materials of any sort in the foregoing calculations. The solutions found for a homogeneous space will next be extended to inhomogeneous problems by the method of images.

Let a current-carrying conductor be placed parallel to, and a distance $h$ away from, an infinite plane surface separating two media of differing permeabilities $\mu_1$ and $\mu_2$, as in Fig. 6.6. It will be supposed that both permeabilities are constant. Except in the special case of equal permeabilities, the solution (5-45) is no longer valid for it cannot satisfy the boundary
### Vector Potential in First Quadrant

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>0.11</td>
<td>0.11</td>
</tr>
<tr>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>0.16</td>
<td>0.16</td>
</tr>
<tr>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>0.19</td>
<td>0.19</td>
</tr>
<tr>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td>0.22</td>
<td>0.22</td>
</tr>
<tr>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>0.26</td>
<td>0.26</td>
</tr>
<tr>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>0.28</td>
<td>0.28</td>
</tr>
<tr>
<td>0.29</td>
<td>0.29</td>
</tr>
<tr>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>0.32</td>
<td>0.32</td>
</tr>
<tr>
<td>0.33</td>
<td>0.33</td>
</tr>
<tr>
<td>0.34</td>
<td>0.34</td>
</tr>
<tr>
<td>0.35</td>
<td>0.35</td>
</tr>
<tr>
<td>0.36</td>
<td>0.36</td>
</tr>
<tr>
<td>0.37</td>
<td>0.37</td>
</tr>
<tr>
<td>0.38</td>
<td>0.38</td>
</tr>
<tr>
<td>0.39</td>
<td>0.39</td>
</tr>
<tr>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>0.41</td>
<td>0.41</td>
</tr>
<tr>
<td>0.42</td>
<td>0.42</td>
</tr>
<tr>
<td>0.43</td>
<td>0.43</td>
</tr>
<tr>
<td>0.44</td>
<td>0.44</td>
</tr>
<tr>
<td>0.45</td>
<td>0.45</td>
</tr>
<tr>
<td>0.46</td>
<td>0.46</td>
</tr>
<tr>
<td>0.47</td>
<td>0.47</td>
</tr>
<tr>
<td>0.48</td>
<td>0.48</td>
</tr>
<tr>
<td>0.49</td>
<td>0.49</td>
</tr>
<tr>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>0.51</td>
<td>0.51</td>
</tr>
<tr>
<td>0.52</td>
<td>0.52</td>
</tr>
<tr>
<td>0.53</td>
<td>0.53</td>
</tr>
<tr>
<td>0.54</td>
<td>0.54</td>
</tr>
<tr>
<td>0.55</td>
<td>0.55</td>
</tr>
<tr>
<td>0.56</td>
<td>0.56</td>
</tr>
<tr>
<td>0.57</td>
<td>0.57</td>
</tr>
<tr>
<td>0.58</td>
<td>0.58</td>
</tr>
<tr>
<td>0.59</td>
<td>0.59</td>
</tr>
<tr>
<td>0.60</td>
<td>0.60</td>
</tr>
<tr>
<td>0.61</td>
<td>0.61</td>
</tr>
<tr>
<td>0.62</td>
<td>0.62</td>
</tr>
<tr>
<td>0.63</td>
<td>0.63</td>
</tr>
<tr>
<td>0.64</td>
<td>0.64</td>
</tr>
<tr>
<td>0.65</td>
<td>0.65</td>
</tr>
<tr>
<td>0.66</td>
<td>0.66</td>
</tr>
<tr>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>0.68</td>
<td>0.68</td>
</tr>
<tr>
<td>0.69</td>
<td>0.69</td>
</tr>
<tr>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>0.71</td>
<td>0.71</td>
</tr>
<tr>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td>0.73</td>
<td>0.73</td>
</tr>
<tr>
<td>0.74</td>
<td>0.74</td>
</tr>
<tr>
<td>0.75</td>
<td>0.75</td>
</tr>
<tr>
<td>0.76</td>
<td>0.76</td>
</tr>
<tr>
<td>0.77</td>
<td>0.77</td>
</tr>
<tr>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td>0.79</td>
<td>0.79</td>
</tr>
<tr>
<td>0.80</td>
<td>0.80</td>
</tr>
<tr>
<td>0.81</td>
<td>0.81</td>
</tr>
<tr>
<td>0.82</td>
<td>0.82</td>
</tr>
<tr>
<td>0.83</td>
<td>0.83</td>
</tr>
<tr>
<td>0.84</td>
<td>0.84</td>
</tr>
<tr>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>0.88</td>
<td>0.88</td>
</tr>
<tr>
<td>0.89</td>
<td>0.89</td>
</tr>
<tr>
<td>0.90</td>
<td>0.90</td>
</tr>
<tr>
<td>0.91</td>
<td>0.91</td>
</tr>
<tr>
<td>0.92</td>
<td>0.92</td>
</tr>
<tr>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>0.99</td>
<td>0.99</td>
</tr>
</tbody>
</table>

**Scale Factor** = 0.40011E-07
conditions

\[ H_{t1} = H_{t2} \quad B_{n1} = B_{n2} \]

for the tangential and normal components. It is readily appreciated that the actual solution must in fact be discontinuous at the interface; the tangential component of flux density must suddenly change at the interface, and so must the normal component of magnetic field. The magnitude of this jump is in either case proportional to the ratio of permeabilities.

In the limiting case of infinite permeability \( \mu_2 \), a particularly simple solution is obtained. If any magnetic field whatsoever exists in the lower medium, an infinite flux density must result from \( \mathbf{B} = \mu_2 \mathbf{H} \). Since this is prevented by the requirement that the flux density vector have zero divergence, there can exist at most a zero field. But the tangential component of field must be continuous across the interface between media; there can consequently not exist any tangential component of \( \mathbf{H} \) in the upper medium near the boundary. As a result, flux lines in the upper medium must meet the interface at right angles.

A solution valid in the upper medium will have been achieved if a function is found that satisfies the vector Poisson equation (5-5) throughout the upper medium, as well as the boundary conditions at the interface. Since both the current and the interface are normal to the \( x-y \) plane in Fig. 6.6, only a \( z \)-directed component of vector potential exists, and attention need only be directed to

\[ \nabla^2 A = -\mu_1 J \quad (6-14) \]

Subscripts on \( A \) and \( J \) have been omitted since only one component of each exists. This equation and the boundary conditions are satisfied by superposing two single-wire solutions, one corresponding to the actual current-carrying wire, the other to its image in the boundary plane:

\[ A = \frac{\mu_1 I}{2\pi} \log \frac{\sqrt{x^2 + (y - h)^2} \sqrt{x^2 + (y + h)^2}}{a^2} \quad (6-15) \]

As with any other potential solution, any arbitrary constant may of course
be added to (6-15) without altering its character. It is readily verified by
differentiation that this solution does indeed satisfy the problem. However,
it is instructive to check this physically. The addition of an image does not
affect Equation (6-14), for the image potential is Laplacian throughout the
upper region; at the surface, on the other hand, the tangential field compo-
nents arising from the wire and its image must be of equal magnitude, but
opposite in direction; they therefore cancel. The resulting field pattern is iden-
tical to the electric field found in the case of a two-wire bundle conductor, as in
Fig. 2.8. Needless to say, this solution is only valid in the upper medium, the
whole effect of the lower medium having been replaced by the image current.
For the lower medium, a separate solution will in general have to be con-
structed. In the limiting case of infinite permeability, however, there is no
need to do so; it is known a priori that \( \mathbf{H} = 0 \) throughout.

The opposite limiting case, where the lower medium has a much lower
permeability than the upper, is similarly treated. For \( \mu_2 \) approaching zero
(i.e., much smaller than \( \mu_1 \)), a very great field is required to make even a
small flux density exist in the lower medium. Since the normal component
of flux density must be continuous across the interface, it may be concluded
that only a very small (in the limit, zero) normal component of \( \mathbf{B} \) is to be
permitted. It may be seen that the required image current is of opposite
polarity and the field pattern much like the electrostatic field of two oppo-
sitely charged wires (Fig. 2.4).

4. Image Solutions for Finite Permeabilities

The general problem with both \( \mu_1 \) and \( \mu_2 \) of finite value may be solved by
superposition of a partial image on the solution for a homogeneous medium.
It is clear from the discussion above that this partial image will consist of a
current in the image position, but of strength \( K_i \), where \( K \) is some multiplier
between +1 and −1. The resulting solution will, of course, be valid only
for the upper medium. In the lower medium there are in reality no currents
at all, and the magnetic vector potential must be governed by Laplace’s
equation:

\[
\nabla^2 A = 0
\]

(6-16)

Any solution valid in the lower medium must satisfy this equation as well
as the boundary conditions. From the point of view of an observer in the
lower medium, the only permissible point for an image is in the upper
medium, i.e., coincident with the original current itself. No image current
may exist in the lower medium, for at the image position, (6-16) would not
be valid. A partial image coincident with the original current itself amounts,
in effect, to a current of modified value, say \( Mi \). These currents and images
are shown in Fig. 6.7.
The vector potential for the upper half space will be given by the two currents shown in Fig. 6.7a, and the vector potential in the lower half space will be determined by the modified current shown in Fig. 6.7b. At the boundary, the two solutions must match. The tangential component of field resulting from the two currents of Fig. 6.7a, $H_{at}$, is given by

$$H_{at} = H_{ot}(1 - K)$$  \hspace{1cm} (6-17)$$

where $H_{ot}$ represents the tangential field that would exist, were there no discontinuity in permeability and hence no image at all. Similarly, the tangential field $H_{tb}$ arising from the modified current in Fig. 6.7b is given by

$$H_{tb} = MH_{ot}$$  \hspace{1cm} (6-18)$$

Since these two expressions are both valid at the interface, they must give the same result. Thus $H_{at} = H_{tb}$, and

$$M = 1 - K$$  \hspace{1cm} (6-19)$$

The second boundary condition to be satisfied relates to normal components of flux density. Using subscripts similar to those above, Fig. 6.7a gives

$$B_{an} = \mu_1 H_{on}(1 + K)$$  \hspace{1cm} (6-20)$$

and Fig. 6.7b yields

$$B_{bn} = \mu_2 H_{on}M$$  \hspace{1cm} (6-21)$$

Combining (6-20) and (6-21), there is obtained

$$M = \frac{\mu_1}{\mu_2}(1 + K)$$  \hspace{1cm} (6-22)$$

which may be solved simultaneously with (6-19) to give

$$M = \frac{2\mu_1}{\mu_2 + \mu_1} \quad \text{and} \quad K = \frac{\mu_2 - \mu_1}{\mu_2 + \mu_1}$$  \hspace{1cm} (6-23)$$

These values of $K$ and $M$ may now be used to find explicit expressions for the magnetic vector potential above and below the interface. The potential solutions are made up simply by superposing terms of the form
(5-5). Arbitrarily choosing \((0, 0)\) as the potential reference point, the solution valid in the upper half plane is

\[ A = \frac{\mu_1 i}{2\pi} \log \frac{\sqrt{x^2 + (y - h)^2}}{h} + \frac{\mu_2 - \mu_1}{\mu_2 + \mu_1} \log \frac{\sqrt{x^2 + (y + h)^2}}{h} \]  

(6-24)

and in the lower half plane

\[ A = \frac{\mu_1 \mu_2 i}{\pi (\mu_1 + \mu_2)} \log \frac{\sqrt{x^2 + (y - h)^2}}{h} \]  

(6-25)

It is easily verified that these solutions do indeed reduce to the limiting cases discussed above, when one or the other of the material permeabilities is much larger than the other. Conversely, the image term in (6-24) vanishes, and (6-25) reduces to (6-24) when the two permeabilities approach equality.

As discussed in Chapter 5, contours of constant magnetic vector potential correspond to flux lines in a two-dimensional field. To obtain a plot of flux lines, it is only necessary to substitute coordinate values into the solutions for \(A\) and trace out lines of equal magnetic vector potential. Such a plot is shown in Fig. 6.8, for a 10:1 ratio of permeabilities.

The vast majority of engineering problems concerned with magnetic materials deals with iron, or other ferromagnetics, where relative permeabilities far in excess of 10 are encountered. Examination of Equations (6-24) and (6-25) shows, however, that the equipotentials of Fig. 6.8 are not very greatly affected by further increases beyond 10. Even with this comparatively low relative permeability, the flux lines meet the interface very nearly at
right angles; and potential values inside the iron differ by only about 10 percent from those obtained with an infinitely permeable material.

5. Leakage Field of a Transformer

A large number of realistic magnetic field problems may, as discussed above, be analyzed on the assumption that iron is infinitely permeable compared to air. However, even with this simplifying assumption, real problems in magnetics tend to be more difficult than electrostatic ones. Their complexity arises both from the intrinsic vector nature of currents, as contrasted to charges, and from the fact that current densities often appear in problems as volume distributions, while charges very frequently may be assumed to reside only in surface layers. Purely analytic solution techniques consequently lead to very highly idealized solutions to a relatively small number of problems, and a large proportion of practical problems must be solved by analogue, graphical, or numerical methods. Even so, it is often necessary to resort to crude approximations. Chief of these is undoubtedly the practical necessity of constraining the field to two dimensions, lest it become necessary to keep track of all three components of vector potential in a three-dimensional field; to do so over a set of only $10 \times 10 \times 10$ field points requires 3000 computer memory locations!

One practical case where a true three-dimensional problem can be treated in two dimensions, while still producing accuracies of the order of 10 percent, is the leakage field of a transformer. Although the leakage flux in fact forms a true three-dimensional distribution, a majority of it follows paths largely made up of iron, and it closes upon itself through the relatively short air paths in the core window rather than around the outside of the core. Figure 6.9 illustrates the core and coils of a two-winding transformer, imagined for purposes of analysis to extend infinitely far into the paper. This approximate representation can be put to practical use by calculating the leakage inductance per unit depth from it and allowing some empirically determined extra length (of the order of 10 or 20 percent) to account for the leakage flux that actually spreads around the ends of the core stack and the outside of the core.

Since the core material is assumed to be infinitely permeable, flux lines must meet iron surfaces at right angles everywhere. Furthermore, the assumption of infinite permeability is equivalent to neglecting magnetizing current of the transformer, so that primary and secondary ampere-turns are always exactly equal, but oppositely directed. It is usual to design transformers with both windings of the same conductor material, and with equal current densities in both windings, for in this manner minimum resistive losses are encountered. Consequently the primary and secondary windings are, from a field point of view, quite indistinguishable and it will suffice to
find a field solution for only one-eighth of the window area, as indicated in Fig. 6.9.

The analysis itself is best performed by Liebmann iteration. Throughout the winding cross section, Poisson’s equation

\[ \nabla^2 A = -\mu_0 J \] (6-26)

applies and may be treated as a scalar equation since both the current density and the vector potential only possess components normal to the plane of the paper. In the insulation space, the field must satisfy Laplace’s equation, i.e., (6-26) with zero current density. To specify the problem completely, boundary conditions are required. Along the symmetry line between primary and secondary windings, \( A \) must possess a constant value, for symmetry requires this position to be occupied by a flux line; this arbitrary constant value may, to enhance symmetry further, be taken as zero. The remaining field boundaries are made up of a symmetry plane requiring orthogonally crossing flux lines, and two iron faces which the flux lines must meet at right angles. This kind of orthogonality condition may be described mathematically by

\[ \frac{\partial A}{\partial n} = 0 \text{ at surfaces} \] (6-27)

where \( n \) denotes distance measured at right angles to the surface. To summarize, the field problem here is specified by Poisson’s equation (6-26) and both equipotential and normal-derivative boundary conditions. These suffice to determine a unique solution.
Coding this problem for computer solution is not more difficult than the electrostatics problems of Chapter 2. As in the latter, a finite-difference Poisson equation is first made up by constructing an approximation for the Laplacian operator. With the notations of Fig. 2.14, the expression for residual, corresponding to Equation (2-38), becomes

\[ R_0 = \frac{A_W + A_N + A_E + A_S - 4A_0}{h^2} + \mu_0 J \]  

(6-28)

In field regions containing no currents, (6-28) of course applies with \( J = 0 \). Equation (6-28) is a valid approximation to (6-26) if the residual \( R_0 \) is zero. The solution procedure here is, as earlier, to assume a set of values of \( A \) quite arbitrarily and iterate until the residuals have decreased to some acceptably small values. The boundary condition along the equipotential edge is introduced by initially setting the edge potentials to zero and subsequently not disturbing them. The normal-gradient boundaries, however, require special treatment. A zero normal derivative is approximated by imagining that an image point outside the field boundary possesses a potential value exactly equal to the inner one, as indicated in Fig. 6.10 where the heavy line represents the actual field boundary. In view of the enforced equality of potential values within and without the boundary, Equation (6-28) here assumes the special form

\[ \frac{A_N + A_S + 2A_W - 4A_0}{h^2} + \mu_0 J = R_0 \]  

(6-29)

While there is nothing inherently difficult about handling this type of boundary condition, the necessity does arise for writing special arithmetic statements into the computer programme for points along the field edges. The result is a lengthening of the programme as compared to purely equipotential boundaries. An alternative to writing special statements is to border the normal-gradient boundary with an extra set of field points, which are not disturbed during iteration but whose potentials are set equal to the
corresponding interior points after each iteration cycle. The labor saved in writing arithmetic statements, however, is often again lost in coding the special routine to set field point values, as well as avoiding difficulties at inside corners where one exterior point must serve as the image point for two interior points.

The distinction between Laplacian and Poissonian field points is most easily taken care of by taking all points to be Poissonian and including in the programme an array of current density values corresponding to the field points; Laplacian points will merely have zero current values assigned to them. An outline flow chart for such a programme is shown in Fig. 6.11, and the results of a typical calculation in Fig. 6.12. In the latter figure, the current-carrying area is outlined with a broken line, the remaining field points representing insulation space. The computer programme used to produce Fig. 6.12 is included in Appendix II. It may be noted that the overrelaxation factor used in this problem is substantially higher, and the

![Flow Chart](image-url)

---

Fig. 6.11
LEAKAGE FLUX PATTERN OF TWO-WINDING TRANSFORMER

<table>
<thead>
<tr>
<th>No. of Turns Squared</th>
<th>Leakage Inductance (Microhenries)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.09203</td>
</tr>
<tr>
<td>0.04</td>
<td>0.00046</td>
</tr>
</tbody>
</table>

LEAKAGE INDUCTANCE PER METRE DEPTH = 0.09203 MICROHENRIES
* (NO. OF TURNS Squared)

LARGEST RESIDUAL = 0.00046 AFTER 92 ITERATIONS

Fig. 6.12
convergence a good deal slower, than in earlier problems that involved equipotential boundaries, as a rule. This slower convergence is typical of normal-gradient boundary conditions and also of Poissonian fields.

In practice, the field map is not sufficient in and of itself, but rather serves as a basis for calculating leakage inductance. The general expression for inductance,

$$L = \frac{1}{i^2} \int \mathbf{A} \cdot \mathbf{J} \, dU$$  \hspace{1cm} (6-30)

readily derived from the stored energy expressions of Chapter 5, may be expressed in approximate finite-difference form as

$$L = \frac{w}{i^2} \sum_k A_k J_k h^2$$  \hspace{1cm} (6-31)

where $w$ represents the depth of the transformer core stack, and $h$ is, as usual, the distance increment in the finite-difference approximation. To avoid explicit appearance of the current, the equation expressing total ampere-turns,

$$ni = \sum_k h^2 J_k$$  \hspace{1cm} (6-32)

may be inserted into (6-31). Since uniform current density exists throughout the winding, the latter equation then assumes the computationally very convenient form

$$L = n^2 \frac{w}{N^3 (Jh^2)^2} \sum_{k=1}^{N} A_k J_k h^2$$ \hspace{1cm} (6-33)

$N$ represents the total number of Poissonian field points and $n$ the number of turns in the winding. Recognizing that $Nh^2$ is merely the cross-sectional area of the current-carrying winding, it is interesting to note that the inductance be written

$$L = w \left( \frac{\sum A_k}{Nni} \right) n^2$$ \hspace{1cm} (6-34)

In view of the fact that $A$ at any point is proportional to the current density, and hence to the product $ni$, inductance is once again seen to be a geometric property of the transformer and independent of the currents and fields themselves.

6. Magnetic Scalar Potential

The magnetic field calculations carried out above may be considered typical of two-dimensional magnetic problems, in which the magnetic vector potential $\mathbf{A}$ possesses only one component. This component $A$ obeys Laplace's
equation throughout the current-free portions of space and plays essentially the same role as the stream function did in electrostatic problems; i.e., it describes the paths taken by flux lines. In accordance with the Cauchy–Riemann conditions on an analytic function, there must exist yet another function which also obeys Laplace's equation in the same regions; for if the function $A(x, y)$ is considered as the real part of an analytic function, the corresponding imaginary part must also satisfy Laplace's equation. It should be evident that this conjugate function will play a role analogous to the scalar potential function of electrostatics; while the magnetic vector potential describes flux line paths, equal value contours of the new function will yield lines at right angles to the flux paths. Such lines are called magnetic scalar equipotentials, and the function is termed the magnetic scalar potential function. This nomenclature is derived almost entirely from analogies with electrostatics and is unfortunately in some respects misleading.

The scalar magnetic potential function is readily expressible in terms of $A$ and the flux density $B$. Let there exist some analytic function whose real part is $A$ and imaginary part $F$:

$$ W(x, y) = A(x, y) + jF(x, y) \quad (6-35) $$

The magnetic vector potential $A$ was defined implicitly in terms of its curl,

$$ B = \text{curl} \ A \quad (5-1) $$

In the present case, $A$ possesses only a $z$-component, and the vector $B$ consequently lies in the $x$-$y$ plane. By (5-1), it is given by

$$ B = i \frac{\partial A}{\partial y} - j \frac{\partial A}{\partial x} \quad (6-36) $$

In accordance with the Cauchy–Riemann conditions (3-26), the derivatives of $A$ may be replaced by the equivalent derivatives of $F$:

$$ B = -i \frac{\partial F}{\partial y} - j \frac{\partial F}{\partial x} \quad (6-37) $$

(Care should be taken to distinguish between the unit of imaginaries $j$ and the unit vector in the $y$-direction $\mathbf{j}$.) Equation (6-37) may be recognized as stating that $B$ is given by the plane gradient of $F$:

$$ B = -\text{grad} \ F \quad (6-38) $$

permitting explicit solution for the function $F$ in terms of $B$ or $A$ as

$$ F = -\int B \cdot d\mathbf{r} = -\int \text{curl} \ A \cdot d\mathbf{r} \quad (6-39) $$

In linear media, the magnetic scalar potential $\Omega$ is conventionally defined as the conjugate function $F$ divided by the material permeability:

$$ \Omega = \frac{1}{\mu} F = -\int \mathbf{H} \cdot d\mathbf{r} \quad (6-40) $$
It follows directly that the scalar potential \( \Omega \) must satisfy Laplace’s equation
\[
\text{div} \ \text{grad} \ \Omega = 0
\] (6-41)

In analogy with the electric field, Equation (6-40) is occasionally interpreted to mean that \( \Omega \) measures the work done on a unit magnetic pole in moving it against the field \( \mathbf{H} \). The objection to this interpretation is simply that there do not exist any points where the magnetic field is nondivergent; i.e., there do not exist any magnetic poles. Not much physical significance can be attached to work done in moving something that cannot exist at all!

Equation (6-41), it should be reemphasized, was derived on the assumption that the magnetic vector potential satisfies Laplace’s equation within the region under consideration. This requires that the region be free of currents of every sort, \( (\mathbf{J} + \mathbf{D}) = 0 \). Conduction currents \( \mathbf{J} \) are invariably contained by conductive material boundaries and in many problems only appear in well-defined, isolated regions. The displacement current \( \mathbf{D} \), on the other hand, may exist in any medium except a superconductor. In view of the time-derivative nature of this term, its importance rises with rising frequency; there exists some limiting rate of time variation beyond which \( \mathbf{D} \) becomes significantly large in almost any problem so that the region examined may no longer be considered current-free. For this reason, it is sometimes said that the magnetic scalar potential is only defined at low frequencies.

At any interface between media with widely differing permeabilities (e.g., air and iron), flux lines in the low-permeability medium are very nearly normal to the interface. It may then be concluded that a medium of very high permeability forms essentially a magnetic scalar equipotential surface, much as a conductive medium forms an electrostatic equipotential surface. This similarity between electric and magnetic field problems may be exploited to transfer solution methods from one area to the other. In particular, conductive analogue techniques and conformal mapping are widely used for solution of magnetic field problems. For nonlinear problems, however, relaxation and iteration are virtually the only useful methods of significant accuracy.

7. Conformal Transformation of Magnetic Fields

A great many practical magnetics problems require the calculation of fields in air near iron boundaries. In such cases, the magnetic scalar potential method is often useful, for the high relative permeability of normal iron (rarely below \( 10^3 \), and more usually higher than \( 10^9 \)) permits the assumption to be made that all iron surfaces are scalar equipotentials. While this does not often lead to useful conclusions about the field within the iron itself, the
field in the surrounding air can sometimes be calculated with high accuracy by conformal mapping.

To illustrate the method, the field pattern of a magnetic tape recording head will be considered. Since tapes are normally made of nonmagnetic plastic materials carrying a very light coating of iron oxide, the tape itself may well be ignored in a first analysis. Ordinarily, a recording head consists of a signal coil exciting an air-gapped iron core. As the tape moves past the air gap, a greater or lesser amount of remanent magnetism is "written" onto the tape, in accordance with the strength of the air-gap field. The crucial region is that near the air gap; a sketch of it appears in Fig. 6.13a. The field to be investigated is clearly symmetrical about the air-gap center line so only one-half of it, as indicated in Fig. 6.13b, need be analyzed. From examining Fig. 6.13a, it is clear that the problem may be dealt with as an open polygon with three vertices, marked A, B, C in Fig. 6.13b. To solve, the boundaries of Fig. 6.13b may be transformed into the upper half plane shown in Fig. 6.14a; subsequently a transformation can be found to map Fig. 6.14a into Fig. 6.14b, whose solution is known. The internal angles at B and C are $\frac{\pi}{2}$ and zero, respectively; thus, the Schwarz–Christoffel differential equation (3-74) becomes

$$\frac{dz}{dt} = K_s(t - 0)^{-1/2}(t - 1)^{3/2-1}$$

(6-42)

$$= K_s \sqrt{\frac{t - 1}{t}}$$

(6-43)
and the transformation, on integrating, turns out to be

\[ z = 2K_z(\sqrt{t - 1} - \tan^{-1}\sqrt{t - 1}) + K'z \]  

(6-44)
The transformation between \( w \)- and \( t \)-planes is still

\[ w = K_w \log t + K'_w \]  

(3-78)
The last two equations completely determine the required transformation, except for determination of the constants. These may be found by means of a technique just like that previously employed. The change \( z = jh \) that occurs at point \( C \) in the \( z \)-plane may be used to determine \( K_z \), for it must be obtainable by the integration \( z = \int dz \) performed either in the \( z \)- or \( t \)-plane. Thus

\[ jh = -\lim_{\zeta \to 0} K_z \int_0^\pi \frac{\sqrt{\zeta e^{j\phi} - 1}}{\zeta e^{j\phi}} \, d\phi \]

(6-45)

where the substitution \( t = \zeta e^{j\phi} \) has been made in much the same manner as in Chapter 3.

By substituting the coordinates of corresponding finite points \( B \), coefficient \( K'_z \) is easily shown to be zero; no difficulty arises here, for point \( B \) is a finite point in both planes. The \( w-t \) transformation may be evaluated exactly as before, obtaining

\[ w = -\frac{j\mu_0}{\pi} \log t \]  

(6-46)

Fig. 6.15
so that the required transformation between $z$- and $w$-planes finally may be written as

$$z = j\frac{2h}{\pi} \left[ \sqrt{\exp \left( \frac{jw}{u_0} \pi \right)} - 1 \right] - \tan^{-1} \sqrt{\exp \left( \frac{jw}{u_0} \pi \right)} - 1]$$  \hspace{1cm} (6-47)

The field map obtained is as shown in Fig. 6.15.

8. Flux Distribution in a Transformer Core

A slightly more difficult problem in the Schwarz–Christoffel transformation is the calculation of flux paths around the corner of an ordinary transformer core. Normal transformers are constructed of laminations, each insulated from its neighbors; and essentially all the flux in any one lamination is confined to that lamination, so that the problem may be treated as two-dimensional to a very high accuracy. In this case the problem to be considered deals with a flux that exists, for all practical purposes, entirely within iron. The iron boundaries therefore act not as equipotential lines, but rather as impermeable walls that may not be crossed by flux lines. That is to say, they must belong to the family of flux lines themselves.

As indicated in Fig. 6.16a, only one corner of the transformer lamination, which in actual fact, of course, has four essentially similar corners, will be treated. It may be presumed that the other corners are sufficiently far away to influence the flux distribution insignificantly. Thus, the Schwarz–Christoffel problem is that of a four-cornered polygon $ABCD$, with interior angles $\frac{3\pi}{2}$, $0$, $\frac{\pi}{2}$, $0$, respectively. The two limiting flux lines $DAB$ and $DCB$ are to
be transformed into the negative and positive real semiaxes of the $t$-plane, as in Fig. 6.16b. The Schwarz–Christoffel differential equation for this case reads

\[
\frac{dz}{dt} = K_z(t + 1)^{1/2}(t - 0)^{-1}(t - 1)^{-1/2}
\]

(6-48)

\[
= K_z \frac{1}{t} \sqrt{\frac{t + 1}{t - 1}}
\]

(6-49)

This expression is directly integrable and yields*

\[
z = K_z \left[ \log (t + \sqrt{t^2 - 1}) + \log t \right] + K'_z
\]

(6-50)

The constants $K_z$ and $K'_z$ must next be evaluated. As in previous examples, this can be accomplished by calculating the change $\Delta z = h$ in going from $(0, y)$ to $(h, y)$ in the $z$-plane, using the Schwarz–Christoffel differential equation (6-48). Making the usual substitution $t = \zeta e^{i\phi}$ to convert to polar form,

\[
\Delta z = h = K_z \lim_{t \to -\infty} \int_0^\pi \frac{\sqrt{\zeta e^{i\phi} + 1}}{\sqrt{\zeta e^{i\phi} - 1}} j d\phi
\]

(6-51)

so that

\[
K_z = \frac{h}{\pi}
\]

(6-52)

On substituting the coordinates of point $C$ in the two coordinate planes and taking the principal value $\log (1) = 0$,

\[
K'_z = 0
\]

(6-53)

whence finally

\[
z = \frac{h}{\pi} \left[ \log (t + \sqrt{t^2 - 1}) + \log t \right]
\]

(6-54)

The transformation between $w$- and $t$-planes is exactly the same as in the last example, with the same points required to correspond. In consequence, the transformation equation is again (6-46).

By combining (6-46) with (6-54), an explicit expression for $z$ in terms of $w$ may be obtained, as was done for the recording head problem. Especially if any of the numerical evaluation is to be done by digital computer, however, it will be found at least equally convenient to leave the transformation equations separate. In fact, it may pay to write out the two transformation equations as a set of several simpler equations, taking real and imaginary parts of the various transcendental functions one at a time. The reason for this is merely that a sequence of numerous simple equations is generally easier to programme for computation (and less likely to lead to programming errors) than a few complicated expressions.

A sketch of the field described by Equations (6-46) and (6-54) appears in Fig. 6.17. Naturally, the same sketch, with field and equipotential lines interchanged, could also be used to calculate the resistance of a right-angle bend in a conductor, the electric field between two bent electrodes, or any other mathematically equivalent situation.

9. Filamentary Source in a Slot or Duct

The final example to be given of the Schwarz–Christoffel transformation involves not only an equipotential boundary but a source external to it as well, thereby necessitating an extra step in the procedure. Let a current-carrying wire of small diameter be placed in a deep slot cut into ferromagnetic material, as indicated in Fig. 6.18. As in previous problems, the iron will
be assumed to be infinitely permeable so that its surface may be treated as a magnetic scalar equipotential. If the slot sides are transformed into the p-axis of a t-plane, the source located at \( z = jb \) will be mapped into another source at \( t = jh \). There remains the problem (in the t-plane) of a line source above an infinite equipotential plane—a problem which can be solved by further conformal transformation.

The Schwarz–Christoffel equation corresponding to Fig. 6.18 is

\[
\frac{dz}{dt} = K_s(t - 1)^{-1/2}(t + 1)^{-1/2} \quad \frac{K_s}{\sqrt{t^2 - 1}}
\]

This expression is directly integrable and may be found in tables:

\[
z = K_s \text{ arc sinh } t + K'_s
\]

Direct evaluation of the constants is possible by substitution of the points B and C. Substituting C and observing that \( \text{arc sinh } 1 = 0 \),

\[
K'_s = a \quad (6-57)
\]

and on substituting B, it is found that

\[
K_s = \frac{2a}{\pi} \quad (6-58)
\]

since \( \text{arc sinh } (-1) = j\pi \). The resulting transformation is

\[
z = \frac{2a}{\pi} \text{ arc sinh } t + a \quad (6-59)
\]

To find the transformed source location \( t = jh \), \( z = jb \) may be substituted in (6-59); there results

\[
jb = \frac{j2a}{\pi} \text{ arc sinh } (jh) + a
\]

so that

\[
h = \sinh \frac{\pi b}{2a} \quad (6-60)
\]

A solution to the problem could now be written in terms of the \( z \)-plane coordinates, if a solution in the \( t \)-plane were known. As extensively discussed above, the flux lines must all meet the infinitely permeable iron face in the \( t \)-plane orthogonally. This requirement is satisfied by placing an image current at \( p = -jh \), thereby relating the solution to that for a two-wire line, which has already been explored at some length. The solution is found by transforming the wire and image of the \( t \)-plane into a single wire in a \( w' \)-plane, by

\[
w' = (t - jh)(t + jh) \quad (6-61)
\]
and finally into the $w$-plane, just as for all previous problems, by

$$w = \log w'$$  \hspace{1cm} (6-62)

These transformations are illustrated in Fig. 6.19. It is at this point easy to solve a second related problem, that of an electrostatically charged wire in a conductive metal duct; the solution is not the same, for this time the image charge has opposite polarity. The appropriate transformation corresponding to (6-61) is

$$w'' = \frac{t - jh}{t + jh}$$  \hspace{1cm} (6-63)

followed, as previously, by

$$w = \log w''$$

These transformations are also illustrated in Fig. 6.19. Both sequences of expressions are readily programmed for numerical evaluation, and the resulting plots are shown in Figs. 6.20 and 6.21, respectively. Although these figures show both flux and equipotential lines, only the numerical printout of the magnetic vector equipotential values is shown in Fig. 6.20 and only the electric scalar potential values are given in Fig. 6.21.
MAGNETOSTATIC PROBLEMS IN ENGINEERING

MAGNETIC FIELD IN IRON DUCT

Fig. 6.20
Electric Field in Metal Duct

![Electric Field Diagram](image-url)

Fig. 6.21
10. Magnetomotive Force and Reluctance

Many practical electromagnetic devices employ iron cores linked by current-carrying conductors. In numerous calculations involving such devices, primary interest centers not on the actual distribution of fields within the iron, but on the electrical terminal behavior of the windings, i.e., on the various winding inductances. Because the majority of methods for solving field problems is concerned with producing flux plots, it is of more than passing interest to discover how inductance is related to flux distribution within the core.

By Equation (5-73), the inductance \( L_{jk} \) between any two contours is given by

\[
L_{jk} = \frac{\phi_j}{i_k} \quad (5-73)
\]

where \( \phi_j \) represents the flux linked by the \( j \)th contour and \( i_k \) the current in the \( k \)th contour, assumed to be the only one energized. Let these contours be formed by wires wound on a magnetic core, as in Fig. 6.22. Clearly the total core flux is threaded through each contour several times; if \( N_j \) represents the number of turns of wire in the \( j \)th winding, then

\[
L_{jk} = N_j \frac{\phi}{i_k} \quad (6-64)
\]

where \( \phi \) is the total flux in the core. This flux is easily found from a field map obtained by any convenient method.

To relate the current \( i_k \) to the field map, it is convenient to define a quantity \( \mathcal{F} \) called magnetomotive force, dual to electromotive force, by

\[
\mathcal{F} = \int \mathbf{H} \cdot d\mathbf{r} \quad (6-65)
\]

It is worth noting that this equation is the magnetic counterpart of (4-68) and subject to quite similar limitations. In any simply connected Laplacian region, the integral in (6-65) must be zero identically, for the magnetic field is conservative in regions not containing currents. If, however, the contour of integration encloses current, as happens in Fig. 6.22 for any contour around the whole core, then \( \mathcal{F} \) is no longer zero but equals the total current crossing the surface spanned by the contour. This current is clearly equal to \( N_k i_k \), for the current \( i_k \) is assumed to be the only one flowing. If other currents should
happen to flow as well, then
\[ \mathcal{F} = \sum_j N_j j_j \]  
(6-66)

Again, the duality with the corresponding electric quantities, Equation (4-81), is worth observing. Physically, (6-66) may be interpreted as saying that total mmf equals the sum of magnetic scalar potential differences across successive sections of the core.

In terms of mmf and flux, the inductance \( L_{jk} \) may now be written
\[
L_{jk} = N_j N_k \frac{\phi_j}{f_{jk}}
\]
(6-67)

The quotient of core flux and mmf is independent of current in any linear medium, for it merely represents the ratio of numbers of field map squares across and along the flux lines. (A parallelism with capacitance calculation from field maps should be evident.) In nonlinear media, on the other hand, the field distribution alters with field strength; the ratio of total flux to mmf consequently changes and so does the inductance. This ratio clearly characterizes any given magnetic path and describes, via (6-67), its appearance as presented to an external winding. It is conventional to write
\[ \mathcal{F} = \mathcal{R} \phi \]  
(6-68)

and to call \( \mathcal{R} \) the reluctance of the magnetic path in question. Equation (6-68) is termed the magnetic Ohm's law by some authors. The similarity of this method of inductance calculation to the calculation of electric resistance of odd-shaped conductors is evident. In fact, the reluctance of a transformer lamination and its resistance to circumferential current flow are calculable from exactly the same field map, their sole difference arising in the multiplying material constants. In terms of the flux map, reluctance may be written
\[ \mathcal{R} = \frac{1}{\mu} \frac{n_c}{n_t} \frac{1}{p} \]  
(6-69)

where \( n_c \) is the number of map squares around the core circumferentially (following flux lines) and \( n_t \) the number of squares counted transversely; \( p \) represents depth into the paper.

Equation (6-67), incidentally, provides a quick way of obtaining rough estimates of inductance. If it is assumed that flux is uniformly distributed across the cross-sectional area of the core, Equation (6-69) may be written
\[ \mathcal{R} = \frac{1}{\mu} \frac{P}{w_p} = \frac{1}{\mu} \frac{P}{A} \]  
(6-70)

Here \( P \) denotes path length around the core (somewhere about the center
line, the exact position matters little since the entire calculation is approximate in any case), \( w \) the width, and \( A = wp \) the cross-sectional area. Thus, an approximation to the inductance is provided by

\[
L_{jk} \approx N_j N_k \frac{A \mu}{P}
\]  

(6-71)

A simple example of reluctance and inductance calculation may serve to illustrate the point. For the toroid with square section of Fig. 6.23, it is easily shown that the flux density distribution is radially logarithmic but uniform in the axial direction. However, for approximate calculations it may suffice to assume uniform flux density across the entire cross section. In that case, the mean path length \( P \) may be taken to equal the mean circumferential distance, \( P = 2\pi(R - \frac{1}{2}w) \), and the effective cross-sectional area to equal the geometric value, \( A = wZ \). The inductance of the \( N \) turn winding is given approximately by Equation (6-71):

\[
L = N^2 \frac{\mu wZ}{2\pi(R - w/2)}
\]  

(6-72)

How large an error this simplified equation contains, depends primarily on the core shape. Clearly a core of relatively large radius and small cross section will not lead to a large error, for the assumption of uniform flux density is not far from truth; on the other hand, a relatively large area coupled with a comparatively short path length will lead to inaccurate results. It should be noted carefully that these remarks are applicable to the approximate equations (6-70) to (6-72); the reluctance as defined by (6-68) or (6-69) is valid regardless of shape, for no geometrical approximations are involved in the definition.

It is readily seen that the notion of reluctance is a direct analogue to resistance in electric circuits and subject to similar practical constraints. Wherever the current (or flux) distribution is known, an accurate value of reluctance (or resistance) may be found; but where a field solution is not known, approximate values can still be guessed at. An example for which no simple analytic solution can be obtained is provided by a form of inductor widely employed in electronic apparatus. As illustrated in Fig. 6.24, a bobbin-wound coil is placed in a cylindrical "pot" of magnetic material and fitted with a "lid," so that the entire winding is encased inside the core. Typically, this core will be made of a ferrite or powdered-iron material, with a relative
permeability in the range 100–2500, so that little if any flux exists outside it. An inductance estimate for this winding may be obtained by assuming some reasonably probable mean flux path through each portion of the core and calculating approximate reluctance values for the individual portions. An estimate will be arrived at whose accuracy is naturally highly dependent on the skill and experience of the estimator.

If the pot core is regarded as divided into geometrically simple sections, as in Fig. 6.24, the flux through each section must be equal to the flux through any other. Consequently the total mmf must be simply the sum of the individual contributions, for by (6-68),

\[ \mathcal{F} = R_1 \phi + R_2 \phi + \cdots \]
\[ = \phi \sum_k R_k \]  

(6-73)

That is to say, reluctances in magnetic circuits are additive in the same manner as resistances in electric circuits. The individual terms for the pot core are easily evaluated. For the inner post, assuming uniform flux density distribution,

\[ R_1 = \frac{h}{\pi r^2 \mu} \]  

(6-74)

Similarly, for the outer sleeve section,

\[ R_2 = \frac{h}{2\pi Rw \mu} \]  

(6-75)

In the bottom and lid sections, a reasonable estimate is arrived at by assuming uniform flux distribution in the vertical and logarithmic distribution in the radial direction. Then

\[ R_3 = \frac{2}{\mu d} \log \frac{R}{r} \]  

(6-76)
the factor 2 having been introduced to allow for both top and bottom. Adding, the total reluctance is
\[
\mathcal{R} = \frac{1}{\mu} \left[ \frac{h}{r} + \frac{1}{2Rw} + \frac{2}{d} \log \frac{R}{r} \right]
\]
from which the winding inductance is easily determined by
\[
L = \frac{N^2}{\mathcal{R}}
\]
in accordance with Equations (6-67) and (6-71).

11. Iterative Solution in Cylindrical Coordinates

While the cylindrical inductor of the foregoing section is not readily amenable to direct analytic treatment, it can quite satisfactorily be handled by the iteration method. It does differ from problems already treated by this technique, however, in possessing rotational rather than translational symmetry, so that it cannot be solved in Cartesians but must be treated in a cylindrical coordinate system. Clearly the vector \( \mathbf{A} \) here possesses only a component in the tangential direction so that in a plane \( \theta = \) constant, it is sufficient to solve a scalar Laplace's equation. Because no variation in \( A_\theta \) occurs with angle, no angular derivative appears, and Laplace's equation in cylindrical coordinates becomes
\[
\frac{\partial^2 A}{\partial r^2} + \frac{1}{r} \frac{\partial A}{\partial r} + \frac{\partial^2 A}{\partial z^2} - \frac{1}{r^2} A = 0
\]

A finite-difference approximation to the Laplacian operator may be found by the same method as employed in Chapter 2. Let a set of adjacent points, as in Fig. 6.25, be considered, noting that the points radially interior (I) and exterior (E) to the center point are distinctly labeled. In the \( z \)-direction, just as in Cartesians,
\[
\frac{\partial^2 A}{\partial z^2} \approx \frac{A_N + A_S - 2A_0}{h^2}
\]
Similarly, in the \( r \)-direction,
\[
\frac{\partial^2 A}{\partial r^2} \approx \frac{A_E + A_I - 2A_0}{h^2}
\]
Finally, the first derivative term in (6-79) may be approximated by
\[
\frac{1}{r} \frac{\partial A}{\partial r} \approx \frac{1}{r} \frac{A_E - A_I}{2h}
\]
\[\text{Fig. 6.25}\]
As in previous problems, the solution technique itself will again consist of setting the (approximate) Laplacian of the potential equal to a residual at each point and operating iteratively until the residuals have decreased to a value considered acceptably close to zero. In accordance with (6-79), the residual will here be defined by

\[ R_o = A_N + A_S + \left( 1 + \frac{h}{2r} \right) A_E + \left( 1 - \frac{h}{2r} \right) A_I - 4 \left( 1 + \frac{h^2}{r^2} \right) A_0 \quad (6-83) \]

The iterative technique used in previous examples may now be employed, except at points along the center line \( r = 0 \), where the residual is not defined by (6-83). For small radii, \( A_E \) and \( A_I \) approach each other and their difference tends to zero, while the radius also tends to zero, yielding the indeterminate form 0/0. A special set of equations for points along the center line may be obtained by application of L'Hospital's rule, which states that the limit of an indeterminate quotient is equal to the limit obtained by dividing derivatives of numerator and denominator. In the present case,

\[ \lim_{r \to 0} \left( \frac{1}{r} \frac{\partial A}{\partial r} \right) = \lim_{r \to 0} \frac{\partial^2 A / \partial r^2}{1} = \frac{\partial^2 A}{\partial r^2} \quad (6-84) \]

so that along the line \( r = 0 \) the residual may be defined by

\[ R_o = A_N + A_S + 4A_E - 6A_0 \quad (6-85) \]

Because of rotational symmetry, \( A_E = A_I \) at all points \( O \) along the center line, accounting for the lack of \( A_I \) in Equation (6-85).

In the present problem, it is in any case physically necessary to assume zero vector potential along the core center line, so that the necessity of writing special equations happily does not arise. If it is assumed that the material permeability is high, then the flux line, or lines, corresponding to \( A = 0 \) must follow the outside surface of the core, while another vector equipotential surface must coincide with the interior surface. The boundaries defining the field are therewith the inner and outer equipotential surfaces, and a symmetry plane. The potential value assigned to the inner core surface is not known; but for purposes of field pattern determination, it may be chosen to have any arbitrary value. A solution obtained on this basis is shown in Fig. 6.26, while the programme employed appears in Appendix II. The programme outline is very similar to those employed previously, with one possibly significant alteration. Equation (6-83) requires two divisions and three multiplications to be performed for each residual, thereby using up more computing time than a corresponding solution in Cartesian coordinates. Of these operations, the divisions are repetitive, always forming the same quotients \( h/2r \) over and over; hence a worthwhile saving in time is achieved by calculating an array of radial multipliers before starting the solution itself, then using them as required.
Fig. 6.26

LARGEST RESIDUAL AFTER 19 ITERATIONS = 0.0002!
12. Inductance Calculation from Field Maps

From the iterative calculation above, there results a field map that describes the relative locations of flux lines but does not yet give any information about the core reluctance, for the map is based on an arbitrarily chosen potential value. In order to obtain further information, it is necessary to know what current must flow in the windings to establish this potential value. This current may be found by employing

\[ \oint \mathbf{H} \cdot d\mathbf{s} = i + \psi \]  

(MI-2)

where the magnetic field value must be found by numerically differentiating the magnetic vector potential:

\[ i = \frac{1}{\mu} \oint \mathbf{curl} \mathbf{A} \cdot d\mathbf{s} \]  

(6-86)

Since there exists only a tangential component of \( \mathbf{A} \), this general expression becomes, in cylindrical coordinates,

\[ i = \frac{1}{\mu} \oint \left[ -1 \frac{\partial A}{\partial z} + 1 \frac{1}{r} \frac{\partial}{\partial r} (rA) \right] \cdot d\mathbf{s} \]  

(6-87)

The derivatives must be replaced by differences, and the integration by a summation, in order to enable calculation to be carried out numerically. As indicated in Fig. 6.27, a contour of integration is first chosen. There is no particular reason to pick any one specific contour since all contours within the core ought to lead to the same result. However, the error caused by finite-difference approximation will be different for different contours, so that in practice the integration is often carried out several times along different contours in order to obtain some idea of the error involved. Along the horizontal section of the line of integration, only the first term of the integrand in Equation (6-87) survives; it may be approximated by

\[ -\frac{\partial A}{\partial z} \, ds \approx A \left( z_0 + \frac{h}{2} \right) - A \left( z_0 - \frac{h}{2} \right) \]  

(6-88)

Similarly, along the vertical sections the approximation may be made

\[ \frac{1}{r} \frac{\partial}{\partial r} (rA) \, ds \approx \frac{1}{r_0} \left[ (r_0 + \frac{h}{2})A \left( r_0 + \frac{h}{2} \right) - \left( r_0 - \frac{h}{2} \right)A \left( r_0 - \frac{h}{2} \right) \right] \]  

(6-89)

In Equations (6-88) and (6-89), the increment in both directions is presumed to be \( h \), and the line of integration is taken as lying exactly halfway between mesh points, as indicated in Fig. 6.28. Instead of integrating continuously, of course, the line integral is formed approximately by adding terms as given by (6-88) and (6-89) above. It may be seen that at corners on the path of integration, inner points such as \( P_1 \) of Fig. 6.28 will be used twice, while
outer points such as $P_2$ will not be employed in the calculation at all. The calculations described are programmed to apply to the pot-core inductor problem in the same programme as the field map, and appear in Appendix II.

Once the current required to establish the assumed potentials is known, the inductance calculation is readily accomplished. It is only necessary to find the total flux by

$$\phi = \oint \mathbf{A} \cdot d\mathbf{s}$$

and find the inductance itself from Equation (6-67). The line integration to find flux actually requires practically no calculation if the contour of integration is chosen as a circular path surrounding the inner post of the core, for $A$ is constant throughout and the integration reduces to merely multiplying by $2\pi r$.

It should hardly be necessary to point out that the general principle employed here is not restricted to calculations carried out in cylindrical coordinates; given a field map of vector equipotential lines, it is always possible to determine the associated current or mmf by numerical differentiation and line integration. However, the precise formulas, by which derivatives are to be calculated from potential values, are dependent on the coordinate system.

**READINGS**

Static and slowly varying magnetic fields form the basis for numerous electrical and electromagnetic devices, lending the subject a considerable practical importance and resulting in an extensive literature. Some introductory notions of the subject will be found in nearly every book on field theory; but for detailed treatments, more specialized books must be consulted. Binns and Lawrenson (1) is undoubtedly the basic reference for this work. Hague (2) deserves to be highly recommended both for his general discus-
sion of magnetic fields and specifically for his lucid treatment of image methods and superposition; however, his use of cgs units may occasionally leave the reader wondering where the extra $4\pi$ came from! Probably the definitive modern treatise of the subject is by Ollendorff (3); because it is largely mathematical, most of the material should be accessible to the English-speaking reader despite the language barrier. Fields near iron boundaries are discussed at some length by Moullin (4) and applied to electric machine theory in another, more specialized, book (5). Numerous other applications are described by Roters (6).

Methods based on the notion of geometric mean distance are introduced and applied to some relatively simple problems by Boast (7). As indicated above, such methods are especially powerful when coupled with numerical techniques; the first paper to discuss them in this modern context is probably that of House and Whidden (8).

Books dealing with conformal transformation were listed at the end of Chapter 3 and will not be reiterated here. Applied examples of the method will be found described in the papers by El-Hakim (9) and Dvoracek (10). The analytic limitations of this technique may be overcome, at least for polygonal boundaries, by seeking the required constants in the Schwarz–Christoffel equation numerically, using a relatively simple technique developed by Binns (11) and then integrating numerically also.

Because boundary shapes encountered in practice are often quite complicated, numerical methods play an important role in magnetic field calculations. Pirktl (12) discusses their use in design in fairly broad terms and furnishes several interesting examples. The importance of relaxation and related techniques has been realized for some time, especially as concerns graphical methods; Weber (13) or Boast (7) may be consulted for details. The numerical relaxation technique was applied to magnetic problems in the early paper by Motz and Worthy (14) and in the much more recent one by Mamak and Laithwaite (15). Where nonlinearities must be accounted for, point iterative methods are the only useful ones known to date; their use in nonlinear analysis in terms of the scalar potential is illustrated by Trutt, Erdélyi, and Jackson (16), and in terms of the vector potential by Ahamed and Erdélyi (17). An application of iterative methods to cylindrically symmetrical problems in electrostatics is reported by Binns and Randall (18).

The use of images can eliminate a great deal of useless labor in both numerical and analytic solution of problems. A paper by Hammond (19) examines the similarities and differences between electric and magnetic image methods and clarifies several potentially troublesome points.


PROBLEMS

6.1 Prove that $\nabla^2 H = 0$ in all current-free regions of a magnetic field.

6.2 Solve for and plot the flux lines outside a plane iron block of finite permeability, with a current-carrying conductor parallel to the iron surface but embedded in the iron a distance $h$.

6.3 Use the partial image method to solve for the electric field near a flat strip conductor glued to a dielectric block (as in printed circuits).

6.4 Two wires parallel to a block of iron carry equal but opposite currents. Assuming the wire diameter to be $a$, the wire spacing $b$, and their separation from the iron face $h$, find the force exerted on the iron. (Hint: Use the principle of virtual work.)

6.5 By superposing solutions of the type of Equation (6-6), plot the magnetic field of a very thin aluminum angle conductor.

6.6 Find the magnetic field pattern of a pair of parallel conductors, carrying direct currents in opposite directions, spaced 10 cm center to center. The conductors are rectangular, 3 cm wide and 10 cm high, and are mounted side by side.

6.7 Find the inductance per unit length of the conductors in Problem 6.6.

6.8 Repeat the leakage inductance calculation for a transformer, but with twice the insulation thickness surrounding the windings. What conclusions can be drawn about the leakage inductance variation?

6.9 Use conformal mapping to solve for the magnetic field surrounding a thin current-carrying wire at $P$: $(a, b)$ if a thin, highly permeable iron sheet lies in the $x$-axis between $x = -c$ and $x = +c$. Plot the field for $c = 6, a = 5, b = 7$.

6.10 Prove Equation (6-41) true.

6.11 Find accurately the reluctance of a ring core of $1 \times 1$ cm square cross section and 5 cm inner diameter, of relative permeability 1000. Find also the inductance of a 300-turn winding placed on this core. Compare with the result obtained from the approximate formula (6-72).

6.12 Repeat Problem 6.11 for a core $3 \times 3$ cm in section and 3 cm inner diameter.

6.13 Find the reluctance of a rectangular core of $1 \times 1$ cm section and $4 \times 5$ cm outside dimensions. Solve by iteration, using a programme similar to (or a modification of) that of Problem 2.9.

6.14 Check the results of Problem 6.13 against the answers obtained by applying the conformal mapping of a square corner (as in the text) to each corner of the core individually.

6.15 Extend the conformal mapping of the square corner to apply to the situation where the strips either side of the corner are not of equal width.
6.16 Find the paths traced by flux lines in the neighborhood of a constriction in their path, as shown in Fig. 6.29.

![Fig. 6.29](image)

6.17 Let the surfaces shown in Fig. 6.29 represent either electric or magnetic equipotentials. Use the results of Problem 6.16 to redesign the sharp corners, curving them so as to have field strength nowhere higher than double its average value in the narrow part.

6.18 Repeat the iterative calculation of flux in a pot core, using a mesh twice as fine. Print out points corresponding to those shown in Fig. 6.26, in order to assess the error involved in the rather crude solution shown here. Also calculate the resulting inductance.
Time-Varying Fields in Conductors

All the field problems examined so far have been restricted to be essentially time-invariant; that is, current and charge distributions have been assumed stationary, so that the time-derivative terms in Maxwell's equations become zero. The coupling between electric and magnetic fluxes expressed by the field equations lies precisely in these time derivatives. It has consequently been possible to analyze the electric and magnetic fields separately. In the general case of time-variant fields, however, it is not possible to do so, and other methods of solution must be resorted to. Some of the most basic techniques for analyzing time-variant fields will be developed and briefly illustrated in this chapter.

1. Field-Vector Wave Equations

To solve any electromagnetic field problem is ultimately to find a solution of Maxwell's equations in either the differential or the integral form; the
solution must satisfy either form of the equations. This is unfortunately a formidable mathematical problem: four partial differential equations and their boundary conditions must be solved as a simultaneous system. The best way of going about this task in general is to eliminate variables until a single partial differential equation in only one field vector is obtained; often it is possible to find a solution of this equation directly.

In the following discussion, no assumption is made to restrict charge or current distributions in any way. However, for mathematical convenience, all media will be assumed linear. Thus, the constitutive relations

\[ \mathbf{D} = \varepsilon \mathbf{E} \quad \mathbf{J} = g \mathbf{E} \quad \mathbf{B} = \mu \mathbf{H} \]

will be assumed to hold, with \( \mu, g, \varepsilon \) constant. They may be used together with the Maxwell equation

\[ \text{curl} \, \mathbf{H} = \mathbf{J} + \dot{\mathbf{D}} \quad (\text{MD}-2) \]

to yield the equation, reduced to two variables,

\[ \text{curl} \, \mathbf{H} = g \mathbf{E} + \varepsilon \ddot{\mathbf{E}} \quad (7-1) \]

The other Maxwell curl equation may be used to eliminate the electric field vector. Taking the curl of both sides of (7-1),

\[ \text{curl} \, \text{curl} \, \mathbf{H} = g \text{ curl} \, \mathbf{E} + \varepsilon \frac{\partial}{\partial t} \text{ curl} \, \mathbf{E} \quad (7-2) \]

Now \( \text{curl} \, \text{curl} \, \mathbf{H} = \text{ grad } \text{ div } \mathbf{H} - \nabla^2 \mathbf{H} \) for any vector \( \mathbf{H} \) (one of the fundamental identities of vector analysis). The magnetic field vector \( \mathbf{H} \), furthermore, is always solenoidal; (7-2) may therefore be written

\[ -\nabla^2 \mathbf{H} = g \text{ curl} \, \mathbf{E} + \varepsilon \frac{\partial}{\partial t} \text{ curl} \, \mathbf{E} \quad (7-3) \]

Finally, the magnetic constitutive relation may be brought into play, and (7-3) altered to

\[ \nabla^2 \mathbf{H} = \mu g \frac{\partial \mathbf{H}}{\partial t} + \mu \varepsilon \frac{\partial^2 \mathbf{H}}{\partial t^2} \quad (7-4) \]

using the fact that \( \text{ curl } \mathbf{E} = -\dot{\mathbf{B}} \). This resulting equation is called the general homogeneous wave equation, for reasons which will become apparent in due course. It is important to note that a similar equation is satisfied by \( \mathbf{B} \); this is easily shown by dividing (7-4) by \( \mu \).

A second wave equation may be obtained by starting with the other Maxwell curl equation (MD-1) and again taking curls of both sides:

\[ \text{curl} \, \text{curl} \, \mathbf{E} = -\text{curl} \, \dot{\mathbf{B}} \]

\[ = -\mu \text{ curl} \, \dot{\mathbf{H}} \]

This expression may now have \( \mathbf{H} \) eliminated by substitution of (MD-2):
curl \, \text{curl} \, \mathbf{E} = \mu (\mathbf{J} + \mathbf{D}) \quad (7-5)

Finally, the three constitutive relations may again be used to yield

\[
\text{grad} \, \text{div} \, \mathbf{E} - \nabla^2 \mathbf{E} = -\mu \varepsilon \frac{\partial \mathbf{E}}{\partial t} - \mu \varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (7-6)
\]

Unlike the magnetic field, the electric field vector is not solenoidal, and the first term on the left does not automatically disappear. It may, however, be written, using the differential form of Gauss's law, in terms of the volume charge density \( \rho \). There is obtained

\[
\nabla^2 \mathbf{E} = \mu \varepsilon \frac{\partial \mathbf{E}}{\partial t} + \mu \varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} + \frac{1}{\varepsilon} \text{grad} \, \rho \quad (7-7)
\]

This equation is known as the general inhomogeneous wave equation. It is readily seen that the two forms (7-4) and (7-7) are similar in all portions of space that contain a uniform charge distribution, or no charges at all, so that the gradient term on the right of (7-7) vanishes. As in the magnetic case, both the field and flux density vectors satisfy a similar equation; this may be shown by multiplying (7-7) by \( \varepsilon \).

In many problems it is possible to omit one or more terms of the general wave equations (7-4) and (7-7). For example, in studying the behavior of electromagnetic fields in free space, all terms on the right-hand side vanish except the second time derivatives. Then

\[
\nabla^2 \mathbf{H} = \mu \varepsilon \frac{\partial^2 \mathbf{H}}{\partial t^2} \quad (7-8)
\]

\[
\nabla^2 \mathbf{E} = \mu \varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (7-9)
\]

Equations (7-8) and (7-9) are called the free-space wave equations, or the lossless wave equations. These simplified wave equations are much easier to solve than the general forms and are studied for that reason.

Another useful simplification occurs in conductive media, where the conduction current \( \mathbf{J} \) is generally much larger than the displacement current \( \mathbf{D} \). Instead of (7-1), the simplified form may then be used:

\[
\text{curl} \, \mathbf{H} = g \mathbf{E} \quad (7-10)
\]

The electric field vector may be eliminated by taking curls, just as in developing Equation (7-2):

\[
\text{curl} \, \text{curl} \, \mathbf{H} = g \, \text{curl} \, \mathbf{E} \quad (7-11)
\]

so that there is obtained, instead of (7-4),

\[
\nabla^2 \mathbf{H} = \mu g \frac{\partial \mathbf{H}}{\partial t} \quad (7-12)
\]

A corresponding development may be carried out for the electric field vector, resulting in
These equations are of a form known as the diffusion equation because of its formal similarity with equations arising in heat diffusion problems. They are useful for determining the current and field distributions in conductive bodies, as will be seen in the examples below. It should be noted that not only \( \mathbf{B} \) and \( \mathbf{D} \) but also \( \mathbf{J} \) obey a similar diffusion equation, obtainable by applying the constitutive relations to (7-12) and (7-13).

2. Solution by Separation of Variables

The partial differential equations (7-4) and (7-7), and their counterparts in the other field quantities, are solvable in many interesting cases by the process of separating variables, i.e., by separating the problem mathematically into two distinct parts, one of which involves time only and the other, space coordinates only. The principle of this method is easiest explained with reference to the homogeneous wave equation (7-4). Let it be assumed that the vector \( \mathbf{H} \) can always be written in the form

\[
\mathbf{H} = S(x, y, z)T(t)
\]

(7-14)

where \( S \) is a time-invariant vector, and \( T \) a space-invariant time function. If this assumed expression is substituted into (7-4), there results

\[
T \nabla^2 S = \mu g ST' + \mu \varepsilon ST''
\]

(7-15)

which may be rewritten as

\[
\nabla^2 S = S \left( \frac{\mu g T' + \mu \varepsilon T''}{T} \right)
\]

(7-16)

It will be observed that the left-hand side of (7-16) is dependent on the space coordinates only, and the expression in parentheses on the right-hand side on time only. Since this equation is to be valid for any combination of space and time variable values, the latter expression involving \( T \) must in fact be independent of time \( t \); and since it cannot vary with the space coordinates, it must be a constant.

A formalized version of the above argument may serve to indicate the limits of this method, as well as to convince the skeptical. On differentiating (7-16) with respect to time, there results

\[
\frac{\partial}{\partial t} (\nabla^2 S) = \frac{\partial S}{\partial t} f(t) + \frac{df}{dt} S
\]

(7-17)

where

\[
f(t) = \frac{\mu g T' + \mu \varepsilon T''}{T}
\]
for brevity. If the time and space functions are entirely independent, then
differentiation of $S$ with respect to time yields zero, and (7-17) becomes

$$\frac{df}{dt} = 0$$

proving that $f(t)$ is in fact a constant. If there exists some defined relation-
ship between space coordinates and time, on the other hand, (7-18) is not
valid. This situation arises in problems involving moving media, where the
velocity and/or acceleration of the medium prescribe an interdependence
between position coordinates and time. Solution methods of other kinds
must be applied to such problems. In the rest of this chapter, however,
moving media will be disregarded altogether so that (7-18) is always true.
Integrating it with respect to time, there is obtained

$$f(t) = -k^2$$

Here $-k^2$ is an arbitrary constant, usually taken in this form in order to
avoid negative signs in the resulting equation in $S$, which is called the
vector Helmholtz equation:

$$(\nabla^2 + k^2)S = 0$$

It will be observed that this equation does not in any way involve time.
The time variations inherent in (7-4) have rather been separated from its
space-dependence and must be given by the time function part of (7-16):

$$\mu \varepsilon T''' + \mu \sigma T' + k^2 T = 0$$

Equations (7-20) and (7-21) are mathematically independent but not phys-
ically unrelated since the same constant $k$ appears in both. Equation (7-21)
is a simple second-order ordinary differential equation and causes no
difficulties. On the other hand, (7-20) is a vector partial differential equation
and may or may not be readily solvable. It is sometimes possible to solve
it by employing the separation process once again. To do so, it is necessary
to write out the vector Helmholtz equation as three scalar equations in the
components of $S$ and apply the separation method to each one. For example,
in Cartesian coordinates, (7-20) is resolvable into its scalar constituents

$$\nabla^2 S_x = -k^2 S_x$$

$$\nabla^2 S_y = -k^2 S_y$$

$$\nabla^2 S_z = -k^2 S_z$$

each of which may be treated in turn by assuming three independent solu-
tions of the form

$$S_x = X(x) Y(y) Z(z)$$

where $X, Y, Z$ are each a function of one coordinate only. Substitution
into (7-22) will then produce three simultaneous ordinary differential equations. These can sometimes be separated and individually solved if it is possible to specify a suitable set of boundary values. The latter must correspond exactly to the boundary conditions imposed upon the partial differential equation (7-4). In general, it is only possible to find solutions if the problem boundaries coincide with coordinate surfaces, so that the boundary conditions can be applied to the separated equations individually. In order to solve the Helmholtz equation by separation of variables, in other words, it is imperative to formulate the problem in an appropriate coordinate system from the start. Unfortunately, separation of the three simultaneous differential equations into independent ones, for which solution methods are available, is only possible for a limited number of coordinate systems. Since these do include Cartesian and cylindrical coordinates, however, many problems of practical interest are solvable by the separation technique.

Many engineering problems are concerned with functions that vary sinusoidally in time. The time variable separation in such cases can be performed directly, and a value of \( k \) obtained immediately, by assuming that the time function \( T \) is a complex exponential, \( T(t) = e^{j\omega t} \). With this substitution, Equation (7-21) becomes

\[
-o^2 \mu \epsilon + j\omega \mu g + k^2 = 0
\]

This equation is readily solved for the separation constant \(-k^2\), obtaining

\[
k = \sqrt{o^2 \mu \epsilon - j\omega \mu g}
\]

The number \( k \) is usually called the wave number. In general it is a complex quantity. With the wave number substituted, the vector Helmholtz equation becomes

\[
(\nabla^2 + o^2 \mu \epsilon - j\omega \mu g)\mathbf{S} = 0
\]

The methods of solution of this equation are of course no different from those applicable to (7-20) since the equation is in fact the same; only the labeling of constants has been altered.

### 3. Fields in a Pulse Transformer Core

A simple example of the separation process is furnished in finding the fields in the laminated core of a transformer when its winding current is suddenly altered. Because each core lamination links the entire winding, but no other currents, the fields in all the laminations will be alike; analysis of only one will then suffice. For simplicity, the laminations will be assumed to be very broad and long compared with their thickness, so that field variations need only be considered in the direction across the metal sheet (the \( x \)-direction
of Fig. 7.1). Coordinate axes may be chosen so that the magnetic field in the metal points in the y-direction:

\[ H = jH_y(x, t) \]  

(7-27)

In the metal laminations, assumed of conductivity \( g \) and permeability \( \mu \), the conduction current density \( J \) by far outweighs the displacement current density \( D \). Consequently, the problem is one of solving the diffusion equation, subject to some suitable boundary conditions. As the simplest case, let the turn-off transient of a pulse be considered. The current flowing in the transformer winding may be assumed to have had a fixed value for some time so that there exists a corresponding magnetic field in the iron, say \( H_0 \). When the winding current suddenly stops, the magnetic field at the surface of each iron lamination must fall to zero instantly, for there no longer exists any net external magnetomotive force. The total flux, however, cannot fall to zero immediately, for energy stored in the transformer core cannot be removed instantaneously. Since there is no longer any current in the winding, a magnetic field (different from the surface field) will temporarily be sustained inside the iron by currents flowing in the lamination itself. Thus, the boundary conditions may be stated as

\[ H(\pm a, t) = 0 \quad H(x, 0) = H_0 \]  

(7-28)

The mathematical problem is therewith formulated; it consists of the diffusion equation (7-12) and the boundary conditions above. Separating variables, the time equation (7-21) becomes
\[ \mu g T' + k^2 T = 0 \]  
(7-29)

and the space equation (7-20),

\[ \frac{d^2 S}{dx^2} = -k^2 S \]  
(7-30)

These equations are seen to be satisfied by

\[ T = \exp \left( -\frac{k^2}{\mu g} t \right) \quad \text{and} \quad S = K \cos kx + K' \sin kx \]  
(7-31)

Examining Fig. 7.1 and the boundary conditions (7-28), it is clear that \( K' = 0 \) because of symmetry. At all times \( t \geq 0 \), (7-28) requires the field at the edges of the lamination, \( x = \pm a \), to be zero. This is only possible if \( k \) is a number such that \( K \cos kx \) has zeros there; consequently

\[ k = \frac{n\pi}{2a} \]  
(7-32)

where \( n \) is any odd positive integer. This choice satisfies the requirement that the field must always be zero at the surfaces; but it does not accord with the condition of initially uniform field throughout the lamination. However, Equation (7-32) provides the possibility of an infinite number of solutions, one for each integer \( n \); any linear combination of these must in turn be a solution also. A linear combination of harmonically related cosine functions is merely a simple Fourier series, which can be made to converge to a constant \( H_0 \) in the interval \( -a \leq x \leq a \) by choosing the coefficients to be

\[ K_n = \frac{4 H_0}{\pi n} \sin \frac{n\pi}{2} \]  
(7-33)

These coefficients are zero for all even \( n \) and have alternating signs for odd \( n \).

Each term in the Fourier series must consist not only of a space distribution \( K_n \cos kx \) but also of a multiplying time function, as required by (7-14). These functions, given by (7-31), are all exponential functions possessing different time constants \( p_n \):

\[ p_n = \frac{n^2 \pi^2}{4a^2 \mu g} \]  
(7-34)

The whole solution is thus

\[ H = j \sum_{n=1}^{\infty} K_n \cos \frac{n\pi x}{2a} \exp (-p_n t) \]  
(7-35)

A corresponding solution for the switch-on transient is readily obtained. Boundary conditions must be similar to the above, except for interchange of initial and terminal states; the process of transition from one to the other is described by

\[ H = j \sum_{n=1}^{\infty} K_n \cos \frac{n\pi x}{2a} [1 - \exp(-p_n t)] \]  
(7-36)
The currents required to sustain these fields are readily found. Since displacement currents have been assumed small, \( J = \text{curl} \, \mathbf{H} \). Because \( \mathbf{H} \) only has one nonzero component,

\[
J = k \frac{\partial H_y}{\partial x} = kJ
\]  
(7-37)

and hence, for the turn-off transient,

\[
J = k \sum_{n=1}^{\infty} \frac{-2H_0}{a} \sin \frac{n\pi}{2} \sin \frac{n\pi x}{2a} \exp(-p_n t)
\]  
(7-38)

Currents in the turn-on transient are similar, except for alteration of the time-dependent multiplier. It is clear that there corresponds one current density term to every term in the field expression and that every current term is independent of all other terms. This latter fact is easily verified by suppressing one term in the field expression (7-35); the corresponding current density term is automatically suppressed also, but all others remain unaffected. Physically, this may be interpreted to mean that there exists a set of mutually independent currents, each with an accompanying magnetic field distribution. All these distributions are such as to share neither net flux linkages nor any net resistive drop; if they did, they would not be independent. One such mode of distribution is shown in Fig. 7.2. Because the various field distributions share no net flux linkages, i.e., are magnetically uncoupled, these modes of distribution are often called normal modes or orthogonal modes; these words may be thought to express the notion that current-carrying coils must be placed orthogonally to each other in order to obtain zero magnetic coupling.

![Fig. 7.2](image-url)
4. Separation of Variables: Rectangular Core

Much the same technique as above may also be employed in two-dimensional problems. For example, the change in magnetic field in a rectangular relay core, in response to a sudden change in coil current, may be calculated by the process of separating variables. As seen from Fig. 7.3, it is again possible to set up the problem in a Cartesian coordinate system in such a manner as to have the z-axis coincide with the direction of the magnetic field vector,

\[ \mathbf{H} = k \mathbf{H} \]

so that only one component need be solved for. Once again the case of field decay, rather than growth, will be solved first. The boundary conditions are similar to the one-dimensional case: \( H(x, y, t) \) may be taken as zero on the boundaries and as \( H_0 \) at all interior points initially. Thus

\[
\begin{align*}
H(±a, b, t) &= 0 \\
H(a, ±b, t) &= 0 \\
H(x, y, 0) &= H_0
\end{align*}
\]  

(7-39)

Separation of time from space variables has already been accomplished in Equations (7-21) and (7-22); the time differential equation that governs the response of the rectangular core is identical to that obtained for flat laminations and has already been found to be satisfied by

\[ T = \exp \left( \frac{-k^2}{\mu \sigma} t \right) \]  

(7-31)

The space variation, however, remains to be determined. It may be found following the method indicated in Equation (7-23). In the present case, the field may be assumed not to depend on \( z \) if the rectangular core is much longer than its lateral dimensions; thus, it is sufficient to let

\[ S = X(x)Y(y) \]  

(7-40)
The scalar Helmholtz equations (7-22) become simply

\[ \frac{\partial^2 S}{\partial x^2} + \frac{\partial^2 S}{\partial y^2} = -k^2 S \]  

(7-41)

and, upon substitution of (7-40),

\[ \frac{X''}{X} + \frac{Y''}{Y} = -k^2 \]  

(7-42)

The two terms on the left-hand side of this equation are independent functions of two independent variables. If (7-42) is to be an equation at all, they must in fact be constants, for otherwise it would be possible to give either term an arbitrary value by selecting new values of \( x \) or \( y \). Letting these constant values be represented by \( -\alpha^2 \) and \( -\beta^2 \), subject to the restriction

\[ k^2 = \alpha^2 + \beta^2 \]  

(7-43)

there are obtained the ordinary differential equations

\[ X'' = -\alpha^2 X \quad \text{and} \quad Y'' = -\beta^2 Y \]  

(7-44)

whose solutions are

\[ X = K_x \cos \alpha x \]
\[ Y = K_y \cos \beta y \]  

(7-45)

There cannot exist, in this case, solutions of the forms \( \sin \alpha x \), \( \sin \beta y \), for reasons of symmetry. Consequently the space variation of the solutions is given by

\[ S = K \cos \alpha x \cos \beta y \]  

(7-46)

and a solution of the whole problem by

\[ H = K \cos \alpha x \cos \beta y \exp \left( \frac{-\alpha^2 + \beta^2}{\mu g} \right) \]  

(7-47)

This solution, however, does not satisfy the boundary conditions (7-39) without further modification. The condition of zero field at the core boundaries is readily satisfied by restricting \( \alpha \) and \( \beta \) to the values

\[ \alpha = \frac{m\pi}{2a} \quad \text{and} \quad \beta = \frac{n\pi}{2b} \]  

(7-48)

where \( m, n \) are any odd numbers. On the other hand, the requirement of a uniform field \( H_0 \) throughout the core at \( t = 0 \) cannot be met by a single solution of the form (7-47), which represents a field varying in both intensity and direction everywhere. As in the case of the flat transformer lamination, a composite solution is required. This must be in the form of a double Fourier series, representing a superposition of numerous solutions of the above type. At time \( t = 0 \), this total solution must equal \( H_0 \):

\[ H_0 = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} K_{mn} \cos \frac{m\pi}{2a} x \cos \frac{n\pi}{2b} y \]  

(7-49)
On multiplying by the appropriate cosine functions and integrating over the cross-sectional area of the core, the series coefficients are obtained as

$$K_{mn} = \frac{16H_0}{\pi^2 mn} \sin \frac{m\pi}{2} \sin \frac{n\pi}{2}$$

(7-50)

As in the one-dimensional case, symmetry requires that there exist coefficients only for odd $m$, $n$, and these alternate in sign. The field variation resulting from the sudden turning-off of coil current may now be obtained by combining the partial solutions with their corresponding time functions,

$$H = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} K_{mn} \cos \left(\frac{m\pi}{2a}x\right) \cos \left(\frac{n\pi}{2b}y\right) \exp(-p_{mn}t)$$

(7-51)

where

$$p_{mn} = \frac{1}{\mu g} \left(\frac{m^2}{a^2} + \frac{n^2}{b^2}\right) \frac{\pi^2}{4}$$

(7-52)

It will once again be noted that the time constants become smaller and smaller as $m$, $n$ increase. The individual terms of (7-51) represent independent partial fields, each distributed in a multiply-humped manner and each decaying at its own rate. It is interesting to note that each of the field patterns divides the cross-sectional area of the core into rectangles, within which that field pattern is repeated. The whole field, given by the sum of all the individual terms, however, shows no evidence of such patterning.

The existence of a field in the core in the absence of external currents implies that currents must flow within the core itself. These currents are easily identified, for they are directly derivable from the now known field distribution. Since displacement currents have in any event been neglected, $\text{curl } H = J$. In view of the existence of only one component of $H$,

$$J = i \frac{\partial H_z}{\partial y} - j \frac{\partial H_z}{\partial x}$$

(7-53)

These differentiations may be carried out in (7-51) term by term. It is evident that there corresponds one circulating current term to each field term, and vice versa; since the field terms are independent, the currents are in fact precisely those required to sustain each field term independently. Naturally, the rate of time decrease of each field distribution and its associated currents must be the same. In fact, on differentiation there is obtained

$$J_{mn} = K_{mn} \exp(-p_{mn}t) \left(-i \frac{n\pi}{2b} \cos \frac{m\pi}{2a}x \sin \frac{n\pi}{2b}y + j \frac{m\pi}{2a} \sin \frac{m\pi}{2a}x \cos \frac{n\pi}{2b}y\right)$$

(7-54)

Finally, the turn-on transient solution, in which the relay is actuated by a current suddenly appearing in its winding, is similar to the above. However, the boundary conditions become reversed in their time sequence. It is easily shown that the solutions are in all details similar, except that exponential terms of the form $e^{-pt}$ are replaced by $(1 - e^{-pt})$. 
5. Magnetizing Impedance of a Transformer

Both the above examples relate to the application of separation of variables to transient problems, in which the time differential equation (7-21) has exponential solutions. In the large class of time-variant problems involving sinusoidal excitations, on the other hand, the assumption of sinusoidally time-varying solutions was seen to lead to the complex wave number

\[ k = \sqrt{\frac{\omega^2 \mu e - j \omega \mu g}{2}} \]  

(7-25)

Here \( k \) is entirely determinate, in contrast to the transient problems where it had to be found from boundary conditions. As a result, sinusoidal problems may often be solved in explicit terms, involving no series expansions of unknown functional descriptions. On the other hand, the results are obtained in complex form and occasionally cause some algebraic difficulties in evaluation.

To illustrate, let the transformer lamination of Fig. 7.1 again be considered, this time assuming a sinusoidal magnetizing current to flow in the windings. Let it be desired to find the manner in which the internal fields vary, as well as the impedance presented to the external coil terminals. It may again be assumed that displacement currents are much smaller than conduction currents; the wave number is then given by

\[ k = \sqrt{-j \omega \mu g} = (j - 1)\sqrt{\frac{\omega \mu g}{2}} \]  

(7-55)

The spatial variation is determined, as before, by the equation

\[ \frac{d^2 S}{dx^2} = -k^2 S \]  

(7-30)

but it must be kept in mind that \( k \) now represents a complex number. Since the time variation of the solution is required to be sinusoidal, there is no longer any requirement for a boundary condition in time; in space, however, some specification is needed. In fact, the surface value of \( H \) is again directly dependent on the transformer winding current, so that the boundary condition may be stated:

\[ H(\pm a, t) = H_0 e^{j\omega t} \]  

(7-56)

The entire problem is thereby specified. The solution is very simple; it is only necessary to recognize that symmetry requires (7-30) to yield

\[ S = K \cos kx \]  

(7-57)

an explicit solution whose single constant \( K \) is readily evaluated. At the surfaces of the lamination, (7-57) requires

\[ S(\pm a) = K \cos ka \]  

(7-58)
so that the complete solution is

\[ H = H_0 \frac{\cos kx}{\cos ka} e^{jat} \] (7-59)

To make practical use of this result, the identity

\[ \cos (j - 1)px = \cosh px \cos px + j \sinh px \sin px \] (7-60)

may be used to resolve the complex functions of (7-59) into functions of real arguments, for which tables and computer subroutines are readily available. Setting

\[ p = \sqrt{\frac{\omega \mu g}{2}} \] (7-61)

there is obtained

\[ H = H_0 \frac{\cos px \cosh px + j \sin px \sinh px}{\cos pa \cosh pa + j \sin pa \sinh pa} e^{jat} \] (7-62)

Making use of trigonometric identities, the magnitude of the field may be written more simply as

\[ \left| \frac{H}{H_0} \right| = \sqrt{\frac{\cosh 2px + \cos 2px}{\cosh 2pa + \cos 2pa}} \] (7-63)

and its phase angle \( \theta \) defined, relative to the phase at the surfaces, by

\[ \tan \theta = \tanh px \tan px - \tanh pa \tan pa \] (7-64)

It is now possible to calculate the magnetizing impedance of the energizing winding from these expressions for the field. A much more direct solution, however, is obtained by starting with Equation (7-59). Taking the magnetizing current of the transformer to be \( I_0 e^{jat} \), the line integral equation (MI-2) yields

\[ H_0 = \frac{n}{s} I_0 \] (7-65)

where \( n \) is the number of turns in the magnetizing winding and \( s \) is the mean magnetic path length around the transformer core. To find the emf associated with this winding, it is first necessary to calculate the core flux. If \( w \) represents the width of the laminations in the \( z \)-direction and \( N \) is the number of laminations, then

\[ \phi = \mu \int \mathbf{H} \cdot d\mathbf{S} \]

\[ = \frac{wN H_0}{\cos ka} \mu \int_{-a}^{+a} \cos kx \, dx \, e^{jat} \] (7-66)

\[ = \frac{2wN I_0 a}{ska} \mu \tan ka \, e^{jat} \]

From this expression, the winding emf \( U \) is readily deduced by differentiation with respect to time and multiplication by the turns number \( n \):
\[ U = j\omega \mu n^2 I_0 \tan ka e^{i\omega t} \]  
(7-67)

where \( A = 2awN \) represents the cross-sectional area of all the laminations taken together, i.e., the effective area of the whole transformer core. Dividing the winding emf by magnetizing current, the magnetizing impedance is now obtained:

\[ Z_m = j\omega n^2 \frac{A}{s} \frac{\mu \tan ka}{ka} \]  
(7-68)

It is readily recognized that the product \( \mu n^2 (A/s) = L_{ac} \) is simply the low-frequency inductance of the magnetizing winding, that is, inductance as calculated by the magnetostatic methods of Chapter 6 without regard to eddy currents in the core. The actual magnetizing impedance is thus given as the product of the winding reactance without any eddy currents taken into account times a complex correction factor:

\[ Z_m = j\omega L_{ac} \frac{\tan ka}{ka} \]  
(7-69)

This expression is not easily evaluated numerically unless tables of tangents of complex arguments happen to be available. These are comparatively rare, and it is a good idea to reduce the magnetizing impedance to an expression in terms of readily accessible functions. Combining Equations (7-55) and (7-61) and eliminating \( k \) in (7-69), the magnetizing impedance may be written

\[ Z_m = \frac{j}{(j - 1)pa} \omega L_{ac} \frac{\sin (j - 1)pa}{\cos (j - 1)pa} \]  
(7-70)

The latter expression is readily converted to a suitable form. Multiplying and dividing by \( \cos (j + 1)pa \) and converting trigonometric functions to hyperbolic functions,

\[ Z_m = \frac{-\sqrt{2}}{j - 1} \frac{\omega L_{ac}}{\sqrt{2} pa} \left( \frac{\sinh 2pa + j \sin 2pa}{\cosh 2pa + \cos 2pa} \right) \]  
(7-71)

\[ Z_m = \frac{\sqrt{2} L_{ac} \mu g a^2}{\omega} \frac{\sinh 2pa + j \sin 2pa}{\cosh 2pa + \cos 2pa} e^{i\pi/4} \]  
(7-72)

Several interesting points emerge from Equation (7-72). It is to be noted, first of all, that the low-frequency behavior (i.e., \( pa \) small) is essentially independent of eddy currents, the magnetizing impedance being very nearly a pure reactance. At very high frequencies, on the other hand, the phase angle approaches 45°, and the magnitude of impedance rises as the square root of frequency. Perhaps the most curious behavior of all occurs in an intermediate range of frequencies. As \( pa \) increases, \( \sinh 2pa \) rises monotonically, while \( \sin 2pa \) oscillates between +1 and -1. In consequence, the angle of \( Z_m \) does not swing from its initial 90° (pure reactance) value to 45° monotonically, but approaches the high-frequency asymptote in an
oscillatory manner. This oscillation is not difficult to observe experimentally, though it is not of major practical importance; since the term \( \sinh pa \) grows rather quickly for \( pa \geq \pi/2 \), the angle of \( Z_m \) initially only rises to about 46.6°, and falls to 44.9° in the next semiperiod. It should be noted also that the trigonometric and hyperbolic functions involved here have \( pa \) as their argument and hence vary with the square root of frequency, not the frequency itself.

Equations (7-59) and (7-72) involve neither the material properties nor the lamination thickness individually but always the combined form (7-61). It may be concluded that no external measurement can be used to distinguish between these variables and that all the above comments regarding behavior at various frequencies may be equally well applied to varying thicknesses, conductivities, or permeabilities. However, the entire analysis is based on constant material parameters throughout the transformer core and may, strictly speaking, not be applied to real sheet steel at all. Nevertheless, these results do serve to give a good qualitative indication of what is to be expected in practice, especially in power and audio transformers that are normally operated below saturation flux levels.

In order to refine this analysis so as to apply to real iron, it is necessary to solve the corresponding nonlinear problem. There arises the additional complication of hysteresis, which cannot be treated as a purely additive phenomenon (although, in the absence of more precise information, it often is). The reason for this is simply that in a nonlinear material, permeability is field-dependent, so that any field change attributable to hysteretic lag will in turn affect the eddy current distribution, and vice versa. It is usual industrial practice to calculate the eddy current loss on the assumption of a constant average permeability, add to this the hysteresis loss obtained by measuring the area of a d-c hysteresis loop, and compare the resulting figures with experimentally determined losses. The difference, which amounts to more than the predicted loss in many modern materials, is called the anomalous loss, and the ratio of real loss to predicted loss the anomaly factor. It is at the present time not yet clearly understood whether the loss anomaly arises solely from nonlinear coupling between eddy current and hysteresis losses, or whether failure to consider the inhomogeneity of the magnetic domain structure of iron is partly responsible.

6. Skin Effect in Round Wires

Another problem of great practical importance, with strong mathematical similarity to the foregoing, is that of alternating current distribution in a conductor. At very low frequencies, currents flow along conductors so as to satisfy Laplace’s equation both inside and outside the conductor, as
discussed in Chapter 3. At higher frequencies, however, an increasing proportion of the total current flows near the conductor surface, and a lesser amount inside. This outward crowding of current is commonly termed skin effect. The general problem of skin effect in conductors is not solvable in terms of the classical transcendental functions and must be treated numerically. However, the very important practical case of a cylindrical wire lends itself to relatively simple analytic solution.

Let an infinitely long, round wire be centered on the z-axis of a cylindrical coordinate system, as in Fig. 7.4. Once again neglecting displacement current in comparison with conduction current, it may be assumed that the current density vector \( J = kJ_0 e^{jat} \).

This problem differs from the previous one in only two respects, a mathematically trivial variable alteration (\( J \) instead of \( H \)) and the more significant change in coordinate system. In the present case, the Helmholtz equation corresponding to (7-20)

\[
(\nabla^2 + k^2)J = 0
\]

becomes, in the cylindrical coordinate system,

\[
\frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial J}{\partial r} \right) + k^2 J = 0
\]  

(7-73)

Derivatives with respect to \( \theta \) and \( z \) do not appear because of symmetry and assumed infinite length. Equation (7-73) may be rewritten as

\[
\frac{d^2 J}{dr^2} + \frac{1}{r} \frac{dJ}{dr} + k^2 J = 0
\]

(7-74)

This equation is known as Bessel’s equation of order zero. It is in essence another version of the equation of simple harmonic motion, and would have exponentially damped sines and cosines as its solutions, were it not for the fact that the damping term is inversely proportional to radius. In fact, the solution functions are somewhat similar to damped sines and cosines, but they do not possess a fixed period. They are called Bessel functions of the first and second kinds and may be found explicitly, solving (7-74) by the infinite series method. Much as for the trigonometric functions, there result defining series for the Bessel functions, which may then be evaluated numerically to any desired number of significant figures. Tables of these functions may be found in most of the usual collections since they form an important part of the classical collection of transcendental functions.

In the conventional notation, the solution applicable to (7-74), subject to the implicit requirement of finite values everywhere, is

\[
J = KJ_0(kr)
\]

(7-75)
It is necessary to distinguish carefully between \( J \), the current density, and \( J_0(x) \), the Bessel function of the first kind and order zero. Tables of the Bessel functions are easily obtained and the problem would now be solved (except for determining \( K \)), were the wave number \( k \) a pure real. Since it is not, it is conventional to write

\[
k = (j - 1)p
\]

where \( p \) is defined by Equation (7-61). Rewriting, Equation (7-75) becomes

\[
J = KJ_0((j - 1)p)
\]

It is to be noted that the phase angle of \( k \) is always \( \frac{3}{4} \pi \), independently of the problem variables. This fact permits use to be made of the so-called Kelvin functions, defined by

\[
\text{ber}(x) + j \, \text{bei}(x) = J_0(\sqrt{-j} \, x)
\]

which represent simply the real and imaginary parts of the Bessel function. The symbols \( \text{ber} \) and \( \text{bei} \) stand for "Bessel-real" and "Bessel-imaginary," respectively; these functions, however, are named after Lord Kelvin, who first computed their values. The Kelvin functions are tabulated also, and subroutines for their computation are readily available.

To complete solution of the problem, it is necessary to find the coefficient \( K \) in Equation (7-75). This could be done most directly by integrating \( J \) over the wire cross-sectional area, thereby obtaining an expression for \( K \) in terms of the total current. A somewhat subtler mathematical approach, however, avoids integration of the Kelvin functions while still making use of the same physical information. At the wire surface, the magnetic field must be

\[
H = \mathbf{1}_\theta \frac{I}{2\pi R}
\]

where \( I \) is the total current in the wire. But from Maxwell's equations,

\[
\text{curl} \, J = g \, \text{curl} \, E = -j \omega \mu g H \mathbf{1}_\theta
\]

Now \( J \) only possesses an axial component, which varies only with the radius. Thus

\[
\text{curl} \, J = -\mathbf{1}_\theta \frac{dJ}{dr}
\]

which, combined with (7-78) and (7-79), yields

\[
\frac{dJ}{dr} = \frac{j \omega \mu g I}{2\pi R}
\]

Substitution of the Kelvin functions now yields derivatives, which are available from tables, instead of integrals:
\[ K = j \frac{\sqrt{\omega \mu g I}}{2\pi R} \frac{1}{\text{ber}'(mr) + \text{bei}'(mr)} \]  

(7-82)

where primes denote differentiation, and \( m = p\sqrt{2} \). The whole result is finally

\[ J = \frac{jmI}{2\pi R} \cdot \frac{\text{ber}(mr) + j\text{bei}(mr)}{\text{ber}'(mR) + j\text{bei}'(mR)} \]  

(7-83)

The current density given by this expression is always greatest at the wire surface and smaller (or equal to the surface value) inside.

A point worth noting about this last problem is that its solution did not at any point employ any data about the values of fields at the wire boundary but made use of the total current instead. That is to say, the whole problem consisted of the differential equation to be solved and a requirement on the surface integral of the solution; but no specification existed about boundary values. Thus, it is not really a boundary value problem at all, but belongs to a different class; it is sometimes called an integral value problem. There are many others of its kind in electromagnetic theory. In particular, all skin effect problems belong to this class, as do many other diffusion problems. Unless a special symmetry property is found to simplify the mathematical task (as for example in the above), no solution to the field differential equations can be obtained. Instead, it becomes profitable to formulate the problem in terms of an integral equation which may be solved by some convenient technique, most frequently numerical. A simple example of this process will be given in Section 9 below.

7. Skin Effect at High Frequencies

From the round-wire solution above, it may be seen that current flows along a wire reasonably uniformly at low frequencies but tends to crowd more and more toward the outer conductor extremities as frequency is raised (or as conductivity is raised). At high frequencies, the solution given by (7-83) becomes quite inconvenient to use, for the Kelvin functions are tabulated only up to some limited value of the argument, and evaluation of the formal solution by the use of tables is thereby restricted. Evaluation by digital computer (which effectively amounts to extending the tables) is hindered by the fact that the infinite series defining the Kelvin functions, or series derived from these, converge slowly for large arguments. Since the current in any case only occupies a thin layer near the conductor surface, however, it is possible to obtain an alternative solution that does not suffer from the same defects. If the frequency is sufficiently high for the effective thickness of the current layer to be small in comparison to the radius of curvature of the conductor, it is reasonable to reformulate the problem in
Cartesian coordinates with the origin taken at the surface, as indicated in Fig. 7.5. The surface then coincides, nearly enough, with the surface \( z = 0 \), and the variation of current density in the \( x \)-direction may be neglected. In the \( y \)-direction, there is no variation in \( J \) at all so that Equation (7-74) may be replaced, to a good degree of approximation, by the simpler equation

\[
\frac{d^2J}{dz^2} + k^2J = 0 \quad (7-84)
\]

Comparison shows that Equations (7-74) and (7-84) are equivalent if the conductor radius is taken to be very large; in fact, (7-84) may be obtained either by starting the problem anew in Cartesian coordinates or else by shifting the origin in (7-74) to the surface by a variable substitution and seeking a solution valid only for large radii. The solution is well-known. While its trigonometric form, as in Equation (7-59), served well with the boundary conditions (7-56), its exponential form is better suited to the present problem:

\[
J = J_0 e^{j(kz \pm \omega t)} \quad (7-85)
\]

Here \( J_0 \) represents the current density at the surface itself. In terms of the real number \( p \) defined by Equation (7-61), \( J \) may be written alternatively as

\[
J = J_0 e^{-pz} e^{j(pz \pm \omega t)} \quad (7-86)
\]

Clearly, not only the magnitude but also the phase of current density varies with depth. In view of the result (7-83), this conclusion is hardly unexpected.

Equation (7-86) is particularly useful at high frequencies where \( p \) is sufficiently large to depress the current density to very low values quite near the surface, for all currents may then be regarded as surface currents in the sense of Equation (4-66). Because this approximation is of considerable utility in the analysis of high-frequency devices, it is worth examining a little more closely.

It will frequently be attempted to replace the exponential current distribution of Equation (7-86) with a single sheet current at the surface. The external magnetic field will be unchanged if the total current beneath the surface is not altered, so that replacement may be effected by making the sheet current equal to the total current in the exponential distribution at each instant. The time-function current, given by the real part of (7-86),

\[
J = J_0 e^{-pz} \cos (pz \pm \omega t)
\]

will then be replaced by a sheet current density \( J_s \) equal at each instant to

\[
J_s = J_0 \int_0^\infty e^{-pz} \cos (pz \pm \omega t) \, dz \quad (7-87)
\]
which is readily evaluated to give

\[ J_s = \frac{J_0}{2\pi} \cos \omega t \]  

(7-88)

The magnetic field outside the conductor will remain unaltered by this replacement. A second important quantity which it is necessary to keep invariant, however, is the total power dissipated in the conductor. Taking the time average over one cycle, the power actually dissipated in a unit of surface is given by

\[ \dot{W}_{av} = \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{g} \int_0^\infty J_0^2 e^{-2pz} \cos^2 (pz \pm \omega t) \, dz \, d\omega t \]

or, by the usual trigonometric identity,

\[ \dot{W}_{av} = \frac{1}{4\pi} \int_0^{2\pi} \frac{1}{g} \int_0^\infty J_0^2 e^{-2pz} [1 + \cos 2(pz \pm \omega t)] \, dz \, d\omega t \]  

(7-89)

The time-varying term in the latter integral clearly has zero average over a full cycle and does not contribute to the average power at all. Hence the power expression (7-89) may be written

\[ \dot{W}_{av} = \frac{1}{2g} \int_0^\infty J_0^2 e^{-2pz} \, dz \]  

(7-90)

But this value is exactly what would be obtained were it to be assumed that the total current, as given by (7-88), flows with uniform distribution in a layer of thickness \(1/p\). For this reason, \(\delta = 1/p\) is called the skin depth or depth of penetration; for all external purposes, the entire current distribution behaves as if it were uniformly distributed and had uniform phase throughout a layer of depth \(\delta\). This appearance is schematically indicated in Fig. 7.6. It may be seen that the stipulation of large radius means, in terms of this sketch, that the depth of penetration is small enough to prevent

![Fig. 7.6](image-url)
the exponential distribution from overlapping to any significant extent with that penetrating into the conductor from the other side.

The skin depth varies inversely with the square root of frequency, as well as with the material properties; it will be observed that either a very high frequency or a very high conductivity will lead to small skin depth. At 60 Hz, the depth of penetration in copper is about 8.5 mm and in aluminum about 12 mm. This relatively shallow depth implies that large conductors are not really used very efficiently, and makes multiple conductor arrangements desirable for high-current work (e.g., bundle conductors for long-distance transmission). On the other hand, at microwave frequencies, say 1 GHz, \( \delta = 2 \times 10^{-6} \) meters approximately, so that initial analyses of waveguides and similar devices may be carried out on the assumption that they are made of perfectly conductive materials carrying only surface currents. A subsequent correction may be made to allow for the losses that must occur in real conductors.

8. Electrodynamic Potential and Wave Equations

The examples above were all treated by what might be termed direct methods, i.e., by solving the pertinent equations for the field quantities directly. Many time-varying problems give rise to great difficulties if such solutions are attempted but yield readily to methods relying on potentials, from which the fields may subsequently be derived. The electric scalar and magnetic vector potentials of earlier chapters are again found to be useful, although they need to be modified a little in order to obtain the greatest possible simplification.

The magnetic vector potential \( \mathbf{A} \) will still be defined by the basic relationship

\[
\mathbf{B} = \text{curl} \mathbf{A} \quad (5-1)
\]

Since Maxwell's equations relate the electric field to \( \dot{\mathbf{B}} \), a direct relationship between \( \mathbf{E} \) and \( \mathbf{A} \) may conveniently be obtained. From Equation (MD-1)

\[
\text{curl} \mathbf{E} = -\dot{\mathbf{B}} \quad (\text{MD-1})
\]

so

\[
\text{curl} \mathbf{E} = -\frac{\partial}{\partial t} \text{curl} \mathbf{A} \quad (7-91)
\]

It is tempting, but untrue, to conclude from this equation that \( \mathbf{E} = -\dot{\mathbf{A}} \); any irrotational vector could be added to \( \mathbf{A} \) or \( \mathbf{E} \) without altering (7-91). Since any irrotational vector can always be written as the gradient of some scalar function or other, (7-91) does require that

\[
\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \text{grad} \, \nu \quad (7-92)
\]
For static fields, the time-derivative term of (7-92) vanishes. The resulting electric field should be the same as that given by Equation (1-46); hence \( V \) must be, as already suggested by the choice of symbol, the electric scalar potential. Thus Equation (7-92) may be viewed as a generalization of (1-46) to time-varying fields.

With an expression for \( E \) in terms of the potentials known, two wave equations may be derived for the potentials similar to those for the fields themselves. Taking curls on both sides of Equation (5-1),

\[
\text{curl curl } A = \mu (J + \dot{D})
\]

The left-hand term may be rewritten, making use of the same vector identity as in proceeding from (7-2). There results

\[
\nabla^2 A = \text{grad div } A - \mu J - \mu \dot{D}
\]  

(7-93)

The displacement-current term in this equation may be eliminated by inserting (7-92):

\[
\nabla^2 A = \text{grad div } A - \mu J - \mu \dot{D} + \mu \varepsilon \frac{\partial^2 A}{\partial t^2}
\]

(7-94)

This equation is not at all convenient as it stands and is not much improved by assuming \( \text{div } A = 0 \). Since it is in any case permissible to choose the divergence of \( A \) arbitrarily, another choice is better. If the Lorentz condition

\[
\text{div } A = -\mu \varepsilon \frac{\partial V}{\partial t}
\]

(7-95)

is adopted, (7-94) reduces to the relatively simple form

\[
\nabla^2 A - \frac{\partial^2 A}{\partial (ct)^2} + \mu J = 0
\]

(7-96)

a wave equation quite similar to (7-7) or (7-9). A similar wave equation may be set up for the scalar electric potential. On substituting into Gauss's law,

\[
\text{div } D = \rho
\]

(MD-3)

the field expression (7-92), and rearranging slightly, there is obtained

\[
\nabla^2 V - \frac{\partial^2 V}{\partial (ct)^2} + \frac{\rho}{\varepsilon} = 0
\]

(7-97)

a result perfectly symmetric with (7-96). These two wave equations form the basis of many problems dealing with the generation and radiation of electromagnetic waves in free space, for they conveniently group all field sources (currents and charges) into distinct terms.

A striking symmetry appears in the Lorentz convention, Equation (7-95), if the three space coordinates and time are thought to form a single complex
four-dimensional entity \((x, y, z, jct)\). The natural generalization of the operator \(\nabla\) of three-dimensional vector analysis is then the operator

\[
\square = \left[ \frac{\partial}{\partial x} \frac{\partial}{\partial y} \frac{\partial}{\partial z} \frac{\partial}{\partial (jct)} \right]
\]

The three spatial components of the magnetic vector potential and the scalar electric potential may similarly be considered to form the four components of a four-dimensional complex potential \(K\), expressible as a column matrix

\[
K = \begin{bmatrix} A_x & A_y & A_z & j \frac{V}{c} \end{bmatrix}_{\text{transp}}
\]

In terms of these quantities, the Lorentz condition appears as a natural generalization of the Coulomb convention adopted previously for static and quasi-static fields:

\[
\square \cdot K = 0 \quad (7-98)
\]

A similar four-component notation may be used for the wave equations, as well as other equations of electromagnetic theory, compressing them into a brief and elegant form. Such terse notations are essential in advanced relativistic electrodynamics, for they permit stating the transformation properties of the various field vectors in a compact manner. However, they will not be employed further here, lest a conflict in notation obscure the substance of the main argument.

9. Direct Integral Solutions: Current in a Flat Bar

The physical significance of Equation (7-92), which forms the cornerstone for the entire development of the electrodynamic potentials, may perhaps best be illustrated by a simple problem leading to an integral equation. Let there exist a conductor of some uniform cross section, as sketched in Fig. 7.7, carrying current in the \(z\)-direction. It will initially be assumed that the electric potential difference between points 1 meter apart along the conductor is \(E_0\) volts, i.e., \(\text{grad} V = -E_0\); no significant variation in potential is assumed to exist in the \(x\)- or \(y\)-directions. If direct current flows in the bar, the electric field along the bar is clearly uniform and time-invariant, and the current density uniformly distributed. However, if the potential difference is time-variant, say in a sinusoidal fashion, the electric field does not equal \(E_0\), for there must exist a time-variant magnetic vector potential and both terms on the
right-hand side of (7-92) must be taken into account. The current density distribution, though still purely \( z \)-directed, must also be nonuniform:

\[
J = -g \text{ grad } V - g \frac{\partial \mathbf{A}}{\partial t} \tag{7-99}
\]

In this example, grad \( V \) clearly represents the electric field contribution arising from the applied potential difference. The added dynamic term accounts for the electric field induced in any given section of the conductor by currents flowing in the whole conductor (and possibly in other neighboring conductors as well). Because the magnetic vector potential does not have the same value everywhere in the bar, this added dynamic term will have different values at different points, leading to a nonuniform current distribution.

To solve for the current density in the bar, the vector potential \( \mathbf{A} \) may be eliminated from Equation (7-99) by substituting the general expression (6-3):

\[
A = \frac{\mu}{2\pi} \int J \log r \, dS \tag{6-3}
\]

There results upon substitution

\[
J = -\frac{\mu g}{2\pi} \int \frac{\partial J}{\partial t} \log \sqrt{(x - \xi)^2 + (y - \eta)^2} \, d\xi \, d\eta + gE_0 \tag{7-100}
\]

This expression assumes somewhat simpler form if all currents and fields are regarded as sinusoidally varying, so that \( J \) and \( E_0 \) may be regarded as phasor quantities. In that case,

\[
J(x, y) = -\frac{j_0 \mu g}{2\pi} \int J(\xi, \eta) \log \sqrt{(x - \xi)^2 + (y - \eta)^2} \, d\xi \, d\eta + gE_0 \tag{7-101}
\]

This equation is of a general type known as an integral equation, for the unknown function \( J(x, y) \) appears in it both explicitly and under a definite integral sign; there is no way of rearranging the equation so as to solve it by a direct means. It will be noted that the integral equation, unlike a differential equation, requires no boundary conditions but is complete in itself. On the other hand, while (7-100) is an integral equation in the space variables, it is a differential equation in time. It therefore requires for solution an added initial condition in time, though no boundary conditions in the space variables are needed.

A solution to (7-101) in analytic form is not known for any kind of area of integration except a circle. In that case, however, the physical problem is merely to find the skin effect in a round wire, which has already been solved as a differential equation in Section 6. Numerical solution is possible by a method very similar to that employed for static problems in Chapter 6; the integral equation may be approximated by a matrix equation, and the latter solved by inversion. The surface of integration (i.e., the conductor
cross-sectional area) is imagined to be divided into sections, which may be unequal in area. If these sections are sufficiently small, the current density may be thought nearly enough uniform over each section to be regarded as a constant in the integration over that section. Equation (7-101) may then be written approximately as

$$J_m = -\frac{j \omega \mu g}{2\pi} \sum_n J_n \int \int_n \log \sqrt{(x_m - \xi)^2 + (y_m - \eta)^2} \, d\xi \, d\eta + gE_0$$

(7-102)

where the subscripts \(m, n\) refer to the \(m\)th and \(n\)th sections of the cross-sectional area, respectively. As discussed in connection with Equation (6-10), the logarithmic term may be replaced by the logarithm of the distance between the midpoints of the \(m\)th and \(n\)th area sections, except in the case \(m = n\). In particular, if the subdivision is into squares of equal sides \(h\), then the geometric mean distance in the case \(m = n\) becomes \(0.44705h\). Again with the stipulation that the latter value is inserted for \(m = n\), (7-102) may be written

$$J_m = -\frac{j \omega \mu g}{2\pi} h^2 \sum_n J_n \log D_{mn} + gE_0$$

(7-103)

which may be regarded as one line taken out of the matrix equation

$$(U + jX)J = G$$

(7-104)

Here \(J\) represents the column matrix of current densities in successive squares, \(G\) represents a similar column matrix with each element equal to \(gE_0\), \(U\) is the unit matrix, and \(X\) is a matrix whose elements are given by

$$X_{mn} = \frac{\omega \mu g}{2\pi} h^2 \log D_{mn}$$

Since \(J\) is a phasor (complex) matrix, (7-104) is in any case a complex matrix equation. It may be solved very simply by inverting the "impedance" matrix that premultiplies \(J\) and postmultiplying the result by \(G\).

Use of complex arithmetic may sometimes prove inconvenient on the available computer. A simple artifice may then be employed, at the price of doubling matrix size. The current density matrix \(J\) may be replaced by a column matrix twice its size, but pure real, consisting of the real parts of \(J_1, \ldots, J_N\) in its first \(N\) places and the corresponding imaginary parts in the remainder. Similarly, \(G\) may be expanded into a double-size column matrix listing first all the real parts (all equal to \(gE_0\)) and then the corresponding imaginary parts (all zero). The matrix equation (7-104) may then be written in expanded, but purely real, form as

$$\begin{bmatrix} U & -X \end{bmatrix} \begin{bmatrix} J_{\text{real}} \end{bmatrix} = \begin{bmatrix} G \end{bmatrix}$$

(7-105)

which is readily handled by any real inversion subroutine.
To illustrate the practical execution of such a calculation, the specific problem of a flat bar (as in Fig. 7.7) 5 mm thick and 100 mm wide, made of copper and carrying a current at 60 Hz, has been solved by subdividing it into 20 squares, \( h = 5 \text{ mm} \). An obvious saving may be realized here by taking advantage of symmetry and dealing with pairs of symmetrically placed elements, just exactly as was done in Chapter 2 for the problem of charge distribution on a rod. The calculations for solving the problem are all explicit, not iterative. The \( X \)-matrix is first built-up and then incorporated into a double-sized matrix. The latter is inverted and postmultiplied by \( G \) to solve the problem. The complete programme appears in Appendix II, and numerical results are shown in Fig. 7.8.

10. Poynting’s Theorem

In order to trace the amount of power or energy involved in time-varying fields of the foregoing examples, it was necessary each time to relate the field quantities to emf’s and currents so as to permit calculation of power as the voltage-current product. This procedure can hardly be termed convenient, especially in investigations of fields not clearly related to any specific network configuration. Happily, it is possible to write equations of energy balance directly in terms of field quantities themselves, using a theorem proposed by Poynting. This theorem states that the power flow into any given closed region of space can be calculated from a closed surface integral taken over the bounding surface of this region. If \( W_s \) represents the energy stored in the volume \( U \) bounded by surface \( S \), \( \dot{W}_s \) represents the rate of
energy removal from the volume, and $\dot{W}_r$ stands for the power loss in conductive material enclosed within, then

$$\dot{W}_o + \dot{W}_s + \dot{W}_r = 0 \quad (7-106)$$

This equation amounts simply to a statement of the law of conservation of energy. The only term in (7-106) for which an explicit expression is known with certainty in all cases is $\dot{W}_r$:

$$\dot{W}_r = \int \mathbf{J} \cdot \mathbf{E} \, dU = -\dot{W}_o - \dot{W}_s \quad (3-15)$$

The object now is to find an expression for the output power $\dot{W}_o$. This is readily accomplished by transforming (3-15). First of all, by substituting $\mathbf{J} = \text{curl} \, \mathbf{H} - \dot{\mathbf{D}}$, there is found

$$\dot{W}_o + \dot{W}_s = -\int \mathbf{E} \cdot \text{curl} \, \mathbf{H} \, dU + \int \mathbf{E} \cdot \dot{\mathbf{D}} \, dU \quad (7-107)$$

Subsequently, the vector identity

$$\text{div} (\mathbf{E} \times \mathbf{H}) = \mathbf{H} \cdot \text{curl} \, \mathbf{E} - \mathbf{E} \cdot \text{curl} \, \mathbf{H} \quad (7-108)$$

may be used to change (7-107) into

$$\dot{W}_o + \dot{W}_s = \int \text{div} (\mathbf{E} \times \mathbf{H}) \, dU - \int \mathbf{H} \cdot \text{curl} \, \mathbf{E} \, dU + \int \mathbf{E} \cdot \dot{\mathbf{D}} \, dU \quad (7-109)$$

Finally, curl $\mathbf{E}$ may be removed by substituting the time rate of magnetic flux density:

$$\dot{W}_o + \dot{W}_s = \int \text{div} (\mathbf{E} \times \mathbf{H}) \, dU - \int \mathbf{H} \cdot \dot{\mathbf{B}} \, dU + \int \mathbf{E} \cdot \dot{\mathbf{D}} \, dU \quad (7-110)$$

The physical significance of the first term in this equation may be discerned by considering some simple direct-current circuit experiment. Clearly $\dot{\mathbf{D}} = \dot{\mathbf{B}} = 0$ for steady fields, so that only the first right-hand term of (7-110) survives. It can next be argued without difficulty that absence of time-varying fields implies a zero rate of energy accumulation in any given volume. Suppose the rate of accumulation were nonzero, say positive. Then it must always remain positive in a time-invariant problem, and an infinite amount of energy must eventually accumulate. Since a similar argument applies to negative rates, one may conclude

$$\dot{W}_o = \int \text{div} (\mathbf{E} \times \mathbf{H}) \, dU \quad (7-111)$$

It is usual to define a new vector, called Poynting's vector, by

$$\mathbf{N} = \mathbf{E} \times \mathbf{H} \quad (7-112)$$

After substituting $\mathbf{N}$ in (7-111) and applying the divergence theorem to
convert the volume integral into a closed surface integral, there finally emerges

$$\dot{W}_s = \int \mathbf{N} \cdot d\mathbf{S}$$  \hspace{1cm} (7-113)

This statement is known as Poynting's theorem. It is normally interpreted as stating that the power outflow rate from any given volume equals the closed surface integral of the Poynting vector.

It is immediately evident from the definition of the Poynting vector that no electromagnetic energy can be transported at all by an electric field acting alone, nor by a magnetic field acting alone. Both are necessary for movement of energy to take place. A corollary of particular interest to the designer of circuit components is that there cannot exist any such thing as a pure inductor or a pure capacitor, in the sense of a device for storing energy solely in a magnetic or an electric field. True, the steady-state storage may take place in only one or the other kind of field, but in any time-variant situation whatever, both fields must enter the picture so that at least some of the total energy must appear in either form.

To illustrate this point, consider again the long round wire of Fig. 7.4. At the wire surface there must exist a longitudinal electric field \( \mathbf{E} = \mathbf{J}/g \). Outside the wire the magnetic field must be tangential, as given by Equation (7-78). The Poynting vector is therefore purely radial:

$$\mathbf{N} = -1 \cdot \frac{I}{2\pi R} \mathbf{g}$$  \hspace{1cm} (7-114)

In the direct-current case the current density is uniform across the entire wire cross section, and (7-114) becomes

$$\mathbf{N} = -1 \cdot \frac{I}{2\pi RG}$$  \hspace{1cm} (7-115)

where \( R \) is the wire radius and \( G \) the conductance per unit length. The electric field must have a continuous longitudinal component not only at the wire surface itself but farther away as well, otherwise no provision would exist for flow of energy into the wire.

The Poynting vector has units of power per area and, as a consequence of Equation (7-113), it is often interpreted to mean the power flow density per unit area. There are of course situations where this interpretation can be applied rigorously (e.g., where \( \mathbf{N} \) is uniform over the entire surface of integration), but it does not have general validity. To appreciate this fact, let Fig. 7.9 be considered. In this figure, a rectangular space is defined by the poles of a ferrite (nonconductive) permanent magnet and two metal plates connected to a chemical battery. There is at all times a vertical magnetic field \( \mathbf{H} \) and a horizontal electric field \( \mathbf{E} \) in that space. If the power-flow-per-unit-area view of Poynting's vector is insisted upon, then the
conclusion is inescapable that an everlasting stream of energy flows through the boxlike space.

In a similar manner, there exists a strong temptation to interpret the two right-hand terms of (7-110) as indicating the rate of magnetic and electric energy increase, respectively. There is no doubt that this viewpoint can in some cases facilitate visualization, especially if the view is adopted that $\frac{1}{2}(E \cdot D)$ represents electric energy density in space and $\frac{1}{2}(B \cdot H)$ represents magnetic energy density. The latter views, however, are a matter of faith rather than proof, as already discussed in Chapters 1 and 5. Care must be taken to use them in their proper role as interpretive tools rather than as established facts on which to base theoretical development.

**READINGS**

A derivation of the wave equations and their subsequent examination in some detail forms the starting point for nearly every book on electromagnetic waves. Some of the best examples are Fano, Chu, and Adler (1), Harrington (2), and Moon and Spencer (3). A somewhat different treatment is furnished by Javid and Brown (4).

Separation of variables forms one of the most fundamental techniques of solving partial differential equations. Concise descriptions of the method may be found in most books. A work devoted almost entirely to this method is by Moon and Spencer (5). Because this technique requires problem boundaries to bear a simple relationship to the coordinates in which solution is attempted, the same authors have also published an extensive catalogue of known coordinate systems and the appearance of vector operators in them (6).

The relay core (or lamination) transient problem is solved by variable separation in Rüdenberg’s book (7); a modal interpretation is given by Silvester (8). Eddy currents in the sinusoidal steady state are considered by Carter (9) as well as Langmuir (10). Because both skin effect and eddy currents are diffusion-equation phenomena, they are treated jointly by Bewley (11), Moullin (12), and Myers (13). Williams and Woodford (14) also describe skin effect in round wires quite comprehensively. Solution of the current distribution problem by matrix approximation to the basic
integral equation is described, and several interesting illustrative examples
given, by Schaffer and Banderet (15); a more general technique valid for
other than sinusoidal currents is based on finding the eigenvectors of the
approximating matrix (16).

No doubt the best explanation of Poynting’s theorem is given by Pro-
fessor Poynting himself (17). The curious interpretations that can some-
times arise, however, are examined, and no fewer than nine distinct Poynting
vectors discovered by Hammond (18).

1. Fano, R. M., Chu, L. J., and Adler, R. B., Electromagnetic fields, energy and
4. Javid, M., and Brown, P. M., Field analysis and electromagnetics. New York:
7. Rüdenberg, R., Transient performance of electric power systems. New York:
8. Silvester, P., Eddy-current modes in linear solid-iron bars. Proc. IEE, 112,
9. Carter, G., The electromagnetic field in its engineering aspects. New York:
Longmans, Green, 1954 (Chapter 11).
1961.
11. Bewley, L. V., Two-dimensional fields in electrical engineering. New York:
Dover, 1963 (pp. 90–100).
York: Oxford Univ. Press, 1955 (pp. 238–250).
Butterworths, 1958 (pp. 127–132).
Macmillan, 1957 (pp. 22–36).
15. Schaffer, G., and Banderet, P., Skin effect in heavy-current conductor bars. 


**PROBLEMS**

7.1 Develop the diffusion equation in $E$ directly from the Maxwell differential equations.

7.2 Show that the components of the vector space functions $S$ (Equation 7-22) must be describable by trigonometric and hyperbolic functions if the vector Helmholtz equation is separable in Cartesian coordinates.

7.3 Prove Equation (7-36) valid.

7.4 Evaluate Equation (7-35) numerically at 21 points across the lamination (including the two edges and the center) at $t = 0$, and at reasonably spaced times up to 3 times the dominant time constant. Use the results to sketch reasonably accurate curves showing the field variation.

7.5 Repeat the calculations of Problem 7.4 for the eddy-current density, using Equation (7-38).

7.6 Prove the independence of the terms in Equation (7-35) by showing the total stored energy at any time to equal the sum of all the energies attributable to the individual modes of field distribution.

7.7 Prove the independence of the terms in Equation (7-38) by showing the total power at any time to be the sum of the powers attributable to the individual current distributions acting alone.

7.8 Repeat Problem 7.6 for Equation (7-51).

7.9 Evaluate the terms $m = 1$, $n = 3$ and $m = 3$, $n = 1$ of Equation (7-49) numerically at a sufficient number of points to be able to sketch contour lines of constant field.

7.10 Prove that Equation (7-35) is a special case of (7-51).

7.11 For a transformer lamination 0.014 in. thick, of 4 per cent silicon steel (resistivity 0.6 microhm-m, relative permeability 2000), evaluate magnitude and phase of field strength at 21 equally spaced points across the thickness of the lamination, including center and edges. Carry out this calculation for frequencies of 25, 60, 400, 1200, 4000, and 8000 Hz. Sketch the results graphically.

7.12 Use Equation (7-72) to calculate the magnetizing impedance and the core loss of a transformer whose core is composed of 10 in.$^3$ of laminations such as described in Problem 7.11. (This is about the amount of iron in a normal 200 watt, 60 Hz transformer.) Plot impedance and loss as functions of frequency.
7.13 Recommend the proper thickness of transformer laminations to be used at 400 Hz (common in aircraft systems) if the steel of Problem 7.11 is specified, and no-load losses are not to exceed 20 watts per kilogram. (Steel has a specific gravity of 7.8.)

7.14 Use the Kelvin function subroutines \( \text{ber}(x) \), \( \text{bei}(x) \), \( \text{berp}(x) \), \( \text{beip}(x) \) to evaluate the current density in a round wire, for values of \( mR \) ranging from 0 up to 8. Calculate the corresponding frequencies, assuming the wire to be made of copper with (a) diameter 1 mm, and (b) diameter 1 in.

7.15 The internal impedance per meter of a wire may be defined as the ratio of surface field \( E \) to the total current \( I \). Justify this definition and use it to produce a curve of relative wire resistance against the parameter \( mR \). What is the resistance of a 1 mm diameter copper wire at 1000 Hz, as compared to its d-c value?

7.16 Assuming that \( \varepsilon = \varepsilon_0 \) for copper (a precise measurement is difficult), find the frequency at which \( J = \dot{D} \) in a copper wire subjected to a longitudinal electric field \( E_0 \cos \omega t \). Comment on the validity of the assumption \( \dot{D} = 0 \) made in the diffusion equation problems of this chapter.

7.17 Derive a method for finding the a-c resistance of a conductor of any cross section and use it to evaluate the resistance of the flat conductor of Fig. 7.7. How does it compare with the d-c resistance?

7.18 Modify the current distribution programme so as to be applicable to a general cross section and use it to calculate the current distribution at 60 Hz of the rectangular aluminum conductor shown in Fig. 6.4.

7.19 Plot the variation in magnitude and phase angle of the magnetizing impedance of a transformer core, using the data given in Problem 7.11. What phase angles are to be expected in practical cases?

7.20 From the Equation (7–114), derive an expression for the alternating-current resistance of a round wire.
Traveling Electromagnetic Waves

In the last chapter, the basic equations governing time-variant fields were derived. Solutions were sought only in a few situations where conduction currents alone had any significance. A very large class of practically important problems, however, deals with precisely the opposite situation of fields in essentially pure dielectric media, where displacement current is dominant. Within this category fall practically all wave propagation and radiation phenomena, and it is this area which will next be examined.

1. Fields in Free Space

In a dielectric medium, or in free space, the behavior of electromagnetic energy is governed by the lossless wave equation,

\[ \nabla^2 H = \mu e \frac{\partial^2 H}{\partial t^2} \]
to which must be attached a sufficient number of suitable boundary conditions
to guarantee a unique solution. The simplest conceivable case is merely that
of no boundaries at all. In this case, a nonzero solution to Maxwell’s equations
is still possible, as will be shown. Because no conductive material is assumed
to exist anywhere in the space being considered, no damping term appears
in the time differential equation (7-21). This fact suggests consideration of
sinusoidally time-varying solutions \( T = e^{i\omega t} \), for in the absence of any
time-dependent boundary conditions, such solutions are the natural (homogeneous)
solutions of the time differential equation. The wave number in this
case must be

\[
k = \omega \sqrt{\mu \varepsilon}
\]

(8-1)

so that the space-function vector Helmholtz equation must read

\[
(\nabla^2 + \omega^2 \mu \varepsilon)S = 0
\]

(8-2)

If there exists any solution for this equation at all (aside from the trivial
solution \( S = 0 \)), there must exist infinitely many; for once any one solution
is known, new ones may be generated from it by coordinate translations and
rotations. Any linear combination of these newly generated solutions must in
turn be a solution, for there are no boundary conditions specified to exclude
any particular solution as impossible. Some set of additional conditions may
thus be laid down in order to specify a single solution, or set of solutions,
in the first place. Arbitrarily, it will here be assumed that

\[
S = jS(z)
\]

(8-3)

There is no \textit{a priori} guarantee that such a solution can be found; whether it
can exist at all will develop as part of the solution itself. Under this assumption,
the vector Helmholtz equation reduces to the simple scalar equation

\[
\frac{d^2 S}{dz^2} = -\omega^2 \mu \varepsilon S
\]

(8-4)

This equation is formally identical to the time differential equation and is
clearly satisfied by solutions of the same type:

\[
S = H_0 \exp \left( j\omega \frac{z}{c} \right)
\]

(8-5)

i.e., sine and/or cosine terms in \( z \). It should be noted that \( c = 1/\sqrt{\mu \varepsilon} \), as in
Chapter 4, has been inserted.

Combining space and time functions, there results the possibility of having
solutions of the type

\[
H = jH_0 \exp \left[ j\omega \sqrt{\mu \varepsilon} (z \pm ct) \right]
\]

(8-6)

A physical interpretation of these solutions may be obtained most easily
if they are written in real time-function form. Taking the real part of (8-6),
for example,
\[ H = jH_o \cos \frac{\omega}{c} (z - ct) \]  
\hspace{1cm} (8-7)

represents a sinusoidal space distribution of magnetic fields, traveling in the 
z-direction with velocity \( c \), as indicated in Fig. 8.1. Correspondingly, the 
solution with \( (z + ct) \) yields a sinusoidal distribution traveling in the negative 
z-direction. It should be noted carefully that the direction of the vector \( H \) is 
not the same as the direction of motion, but rather at right angles to it; this 
is a direct consequence of assumption (8-3), which defined the space variation 
of \( S \). Other kinds of solutions are obtained by starting with different assumptions regarding \( S \).

![Fig. 8.1](image)

### 2. Plane Waves in Unbounded Space

The solution (8-7) is not possible in isolation; in order for the field \( H \) described by it to exist, there must also exist the other fields required by 
Maxwell's equations. These may be discovered by successive differentiations 
and integrations of the expression for the magnetic field. Of primary interest 
is the electric field \( E \), since the magnetic flux density \( B \) and the electric flux 
density \( D \) are directly obtainable from \( H \) and \( E \), respectively. Because there 
are no conduction currents, the Maxwell magnetic curl equation becomes 
\[ \text{curl } H = \dot{D} \]  
\hspace{1cm} (8-8)

from which it may be deduced that 
\[ \dot{D} = -ij \frac{\partial H_y}{\partial z} \]

that is, 
\[ \dot{D} = iH_o \frac{\omega}{c} \sin \frac{\omega}{c} (z - ct) \]  
\hspace{1cm} (8-9)
Integrating with respect to time and dividing by the permittivity,

\[ E = iH_0 \sqrt{\frac{\mu}{\varepsilon}} \cos \frac{\omega}{c} (z - ct) \]  

(8-10)

In other words, there exists a clearly determinate electric field which must always accompany the traveling magnetic field if the latter is to be permitted to exist at all. The two fields must be in time phase always and everywhere; for any given value of \( z \), their law of time variation is exactly the same. Their spatial directions, on the other hand, are at right angles to each other; a magnetic field in the \( y \)-direction requires an electric field in the \( x \)-direction, as indicated by Fig. 8.2. This situation is often summarized briefly by the phrase “in time phase and space quadrature.”

It is of interest to note that the relative magnitudes of electric and magnetic field are determined by the properties of the medium in which the fields propagate (in the present case, of free space). Combining Equations (8-7) and (8-10), there is obtained the ratio of magnitudes

\[ \zeta_0 = \frac{E}{H} = \sqrt{\frac{\mu}{\varepsilon}} \]  

(8-11)

This ratio is not dependent on time at all since \( E \) and \( H \) are in time phase. As is readily verified, its dimensions are ohms; it is an intrinsic property of the medium and is therefore termed the intrinsic impedance of the medium. Since the permittivity and permeability of any real lossless medium are pure reals, the intrinsic impedance, it might be noted, is also a pure real number.

It is worthwhile to observe the physical makeup of this pair of fields. At some point \( z_0 \) in space, say \( z_0 = \theta_0 / \omega \sqrt{\mu \varepsilon} \), there exist the fields

\[ E = i \zeta_0 H_0 \cos (\omega t - \theta_0) \]
\[ H = j H_0 \cos (\omega t - \theta_0) \]  

(8-12)

These are simple time-harmonic vector fields, in time phase and space quadrature. On the other hand, at any fixed time \( t_0 \),

\[ E = i \zeta_0 H_0 \cos \frac{\omega}{c} (z - z_0) \]
\[ H = j H_0 \cos \frac{\omega}{c} (z - z_0) \]  

(8-13)

where \( z_0 = c t_0 \). Equations (8-13) describe two fields sinusoidally distributed in space, as depicted by Fig. 8.2. Any given pair of electric and magnetic field values may be described by a fixed value of

\[ \frac{\omega}{c} (z - ct) = \varphi \]
e.g., \( H = H_0 \cos \phi \). Evidently this value is constant all along a plane normal to the \( z \)-axis. Such a plane is termed a \textit{surface of constant phase} since \( \phi \) has a fixed value everywhere on it. As time goes on, Equations (8-7) and (8-10) require these planes to move at a velocity \( c \) along the \( z \)-axis, for the surface described by some particular value \( \phi_0 \) is given by

\[
\phi_0 = \frac{\omega}{c} (z - ct)
\]

which can only be maintained constant if \( z \) increases as \( t \) increases. Because the solution here described results in translational motion of sinusoidally distributed values of \( E \) and \( H \), with plane surfaces of constant phase, it is generally called a \textit{plane wave}.

The sinusoidal distribution of electric and magnetic fields in a plane wave as in Equations (8-13) leads to a sinusoidal distribution of energy as well. Wherever the field values reach a maximum, energy density reaches a peak; wherever the fields are zero, so is stored energy. As the plane wave propagates, it must move energy along. The energy flow is readily computable by Poynting’s theorem. At any point in space, the Poynting vector is

\[
N = E \times H
\]

\[
= kH_0^2 \sqrt{\frac{\mu}{\varepsilon}} \cos^2 \frac{\omega}{c} (z - ct)
\]

watts per square meter. On rewriting, the spatial distribution of energy is seen to be sinusoidal also:

\[
N = k \frac{H_0^2}{2} \sqrt{\frac{\mu}{\varepsilon}} \left[ 1 + \cos 2 \frac{\omega}{c} (z - ct) \right]
\]

The power flow, it should be noted, is never negative but varies over a range of positive values. The flow velocity, furthermore, is equal to the velocity of light; in view of the fact that the \( E \) and \( H \) vectors propagate with a velocity \( c \), this is hardly surprising.

An essential point to observe is that the directions of \( E \) and \( H \) are not arbitrary but must be related by the requirement of orthogonality as well as the direction of the Poynting vector. Equation (8-16) describes an energy distribution traveling in the direction of the Poynting vector. Were the direction of field travel to be reversed by letting the velocity of propagation be \(-c\), an additional minus sign would arise in the differentiation (8-9), reversing the sense of \( E \) and hence of \( N \).

3. Polarization of Plane Waves

The above solution to the unbounded wave equation is only one of a large variety of possible solutions, having been arrived at by assuming the magnetic field to posses only one component \( H_y \), which depended only on \( z \). For this
traveling wave, the accompanying electric fields must necessarily be directed in the \(x\)-direction. However, these directions are in themselves quite arbitrary; rotation of the coordinate system about the \(z\)-axis through some angle \(\theta\), for example, will retain the orthogonality of \(E\) and \(H\) vectors and preserve the direction of propagation; but it will cause both field vectors to possess both \(x\)- and \(y\)-components, as shown in Fig. 8.3. It is usual to call such waves \textit{linearly polarized}, for either field vector lies in the same line of any plane \(z = \text{constant}\) (i.e., when viewed in the direction of propagation, the \(E\) and \(H\) vectors everywhere stand in a line). Usually the case \(H = jH_y\) is termed a \textit{vertically polarized} wave, and the case \(H = iH_x\) \textit{horizontally polarized}. These terms are naturally physical rather than mathematical and refer to directions relative to an imagined horizontal plane \(y = \text{constant}\). All other waves, possessing both vertical and horizontal components, are termed \textit{obliquely polarized}. A typical obliquely polarized wave might be

\[
\begin{align*}
E &= iE_x \cos \frac{\omega}{c} (z - ct) + jE_y \cos \frac{\omega}{c} (z - ct) \\
H &= iH_x \cos \frac{\omega}{c} (z - ct) + jH_y \cos \frac{\omega}{c} (z - ct)
\end{align*}
\]  
(8-17)

It is readily shown that the individual components of this wave, and hence the whole wave, obey the relationship \(E = \zeta_0 H\), as before.

A different and very interesting case arises when the \(x\)- and \(y\)-components of the waves are not in time phase with each other, i.e., when the whole wave is composed of one vertically and one horizontally polarized wave, the two being in time quadrature:

\[
\begin{align*}
E &= iE_x \cos \frac{\omega}{c} (z - ct) + jE_y \sin \frac{\omega}{c} (z - ct) \\
H &= -iH_x \sin \frac{\omega}{c} (z - ct) + jH_y \cos \frac{\omega}{c} (z - ct)
\end{align*}
\]  
(8-18)

The horizontally polarized wave evidently lags the vertically polarized one in time by 90°. The sum of these waves, however, behaves differently from either one taken singly. For example, there exists no plane where either field vector has a zero value, for the quarter-period shift must occur both in time and space. The horizontally polarized wave reaches a maximum wherever the vertically polarized one has a zero, and vice versa.

The magnitude of field in such a wave is interesting to calculate. For the electric field, there is obtained
\[ E = \sqrt{\frac{E_x^2 \cos^2 \frac{\omega}{c}(z - ct) + E_y^2 \sin^2 \frac{\omega}{c}(z - ct)}{\frac{1}{2}} + \frac{1}{2} (E_y^2 - E_x^2) \cos 2\frac{\omega}{c}(z - ct)} \]

which may be recognized as a parametric representation of an ellipse:

\[ E = \sqrt{\frac{E_x^2 + E_y^2}{2}} \sqrt{1 + \frac{E_y^2 - E_x^2}{E_x^2 + E_y^2} \cos 2\frac{\omega}{c}(z - ct)} \]

At different times (or, what amounts to the same thing in a traveling wave, in different places), the value of this field differs both in magnitude and direction; viewed along the direction of propagation, the tips of successive vectors \( \mathbf{E} \) trace out an ellipse in any plane \( z = \) constant. If \( E_x = E_y \), a circular locus results. Such waves are therefore known as \textit{elliptically polarized} and \textit{circularly polarized} waves, respectively. However, these names are in a sense misleading. Successive positions of the tip of the field vector do not in fact trace out ellipses or circles in space; while the necessary rotation to describe an elliptic shape takes place, the wave also moves forward so that the loci of the field vectors \( \mathbf{E} \) and \( \mathbf{H} \) are actually elliptical or circular spirals about the axis of propagation. At any fixed space point, the field vectors rotate in the \( x-y \) plane with an angular velocity \( \omega \). According to their direction of rotation, they are termed either \textit{right-handed} or \textit{left-handed}.

While only the electric field (8-20) has been calculated in detail, the magnetic field is readily deduced from it: its magnitude must be given by the electric field divided by the intrinsic impedance of the medium, while its direction must be at right angles to the electric field so as to preserve the direction of wave propagation.

\section*{4. Interference of Plane Waves}

It is interesting to note that all the variously polarized waves of the above section are arrived at by much the same superposition process as used in Chapter 2 to find electrostatic fields of many complicated configurations of charges. The component elements in this case are not static fluxes but plane waves of the simplest kind, all shifted or rotated with respect to each other in space or differently phased in time. Only a few simple possibilities have so far been considered, the discussion having been restricted entirely to waves traveling in the same direction. Many more different kinds of waves can be constructed if this restriction is removed. As a matter of fact, any kind of traveling wave may in principle be built up of elementary plane waves, just as any charge distribution can be constructed of point charges. Whether this physical view is also mathematically fruitful is just as much dependent on the particular problem at hand in waves as it was in electrostatics.
The simplest possible case of waves propagating in different directions is that of two simple plane waves, both traveling along the z-axis but one directed to the right, the other left. Taking either to be of the type described in Equation (8-7), the total magnetic field of the two waves is

\[ \mathbf{H} = jH_1 \cos \beta(z - ct) + jH_2 \cos \beta(z + ct) \] (8-21)

where \( \beta = \omega / c \) has been written for brevity. The first wave in (8-21) is evidently directed to the right, the second to the left. Correspondingly, the electric field must be

\[ \mathbf{E} = i\xi_0 H_1 \cos \beta(z - ct) - i\xi_0 H_2 \cos \beta(z + ct) \] (8-22)

The second vector term, it should be noted carefully, has a direction opposite to the first. This direction is not at all arbitrary. Maxwell's curl equations require that the electric and magnetic fields be derivable from each other, as in Equations (8-8) to (8-10); the signs are implicitly determined thereby.

By employing the usual trigonometric identities, the fields of (8-21) and (8-22) may be written as

\[ \mathbf{H} = j(H_1 - H_2) \cos \beta(z - ct) + j2H_2 \cos \beta z \cos \omega t \] (8-23)

and

\[ \mathbf{E} = i\xi_0 (H_1 - H_2) \cos \beta(z - ct) + i\xi_0 2H_2 \sin \beta z \sin \omega t \] (8-24)

Inspection of the rewritten equations shows that the resultant of two oppositely directed traveling waves consists of one component traveling to the right and a second component which does not travel at all. This possibly somewhat surprising conclusion is readily verified by checking the behavior of \( \cos \beta z \cos \omega t \) for different values of time and space. Any zero of this function, for example, remains fixed in either space or time. This field component is sinusoidally distributed in space in a time-invariant manner and is therefore called a standing wave; it is readily recognized that its mathematical description does not differ from those of standing waves encountered in acoustics or mechanics. It will be noted that the magnitudes of \( \mathbf{E} \) and \( \mathbf{H} \) in the standing wave are still related by (8-11); i.e., their ratio is the intrinsic impedance \( \xi_0 \). Similarly, the traveling components form a true traveling wave, with the electric and magnetic fields in the same ratio. It is often said that the two oppositely traveling waves have interfered with each other to produce the standing wave.

The two waves initially postulated in (8-12) propagate in opposite directions and therefore have oppositely directed Poynting vectors, as indicated in Fig. 8.4. The resultant traveling wave of (8-23) possesses a right-directed Poynting vector, as it should, but naturally not of a magnitude equal to the right-traveling elementary wave component. A most interesting behavior is required of the standing wave, however; this wave pulsates in time but does not propagate. The energy associated with it obviously cannot simply vanish
when the fields reach zero and reappear from nowhere as the fields grow again. Inspection of (8-23) and (8-24) clarifies this situation: the electric and magnetic fields of the standing wave are in time and space quadrature so that the electric field reaches a maximum when the magnetic field is zero, and vice versa. The stored energy is thus not altered but merely handed back and forth between the two fields. Since the Poynting vector must be $z$-directed, energy has to flow in the $z$-direction in the process. This requirement is also provided for in the field equations by a quarter-wavelength shift of the electric and magnetic fields along the $z$-axis. Thus, the energy is shuttled back and forth in the $z$-direction as it is converted from electric to magnetic form, and back again.

5. Reflection of Plane Waves at Material Boundaries

To clarify the general discussion of the preceding section, two specific physical situations will be examined in which interfering oppositely traveling waves occur.

Suppose the half space $z > 0$ to be filled with an infinitely conductive medium. According to the general principles set forth in Chapter 3, the electric field in such a medium must always be zero. No energy transfer can take place anywhere within this medium, for the Poynting vector is identically zero everywhere, and hence its closed surface integral is zero also. A right-traveling wave assumed to exist in the half space $z < 0$ carries energy toward this medium; but the energy is not permitted to penetrate into it. Only one possibility exists: the energy carried to the right must be carried back again by a reflected wave.

The nature of this reflected wave is readily deduced by imposing such boundary conditions as must prevail at the surface separating the half spaces. The coefficients $H_1$ and $H_2$ will be determined thereby and a complete solution obtained.
In Chapter 4 it was shown that the tangential component of electric field must be continuous at all material surfaces,

\[ E_{t1} - E_{t2} = 0 \]  

(4-64)

In the present case, \( E_{t2} = 0 \), for no electric field is permitted in the perfect conductor. Consequently \( E_{t1} = 0 \) also; there cannot be any electric field at the interface on either side. Inspection of Fig. 8.4 shows that this condition is satisfied either if no electric field exists anywhere or else if \( E_1 = -E_2 \), i.e., if the two oppositely traveling waves are equal in magnitude. This latter conclusion agrees well with the reasoning based on energy flow: all the energy must be reflected, so the magnitude of the reflected wave must equal that of the incident wave. The incident and reflected magnetic field vectors must clearly be equal also. Equations (8-23) and (8-24) become

\[ H = j2H_0 \cos \beta z \cos \omega t \]  

(8-25)

\[ E = i2\zeta_0 H_0 \sin \beta z \sin \omega t \]  

(8-26)

It ought to be noted that the magnetic field reaches a maximum value at the surface \( z = 0 \), where the electric field is always zero. By Equation (4-67), this fact implies a surface current flowing in the infinitely conductive medium of \( 2H_0 \) amperes per meter in a direction coincident with \( E_1 \).

A somewhat more general case of reflection occurs at a surface between two dielectrics, where no surface current can be conducted and perfect reflection cannot occur. It is physically evident that the incident energy must now be divided between a partial reflected wave and a transmitted wave that penetrates the interface into the second medium. There are consequently three waves to be matched at the interface, as indicated in Fig. 8.5: the incident wave, with associated Poynting vector \( \mathbf{N}_1 \); the reflected component \( \mathbf{N}_2 \); and the transmitted wave \( \mathbf{N}_t \). Thus the solution will consist of an incident and a reflected wave (i.e., a propagated and a standing wave) on the left and
a single propagated wave on the right. (Note the similarity of the partial reflection to the partial images of Chapter 6, Sections 3 and 4!)

The electric field must be continuous at the interface, just as in the preceding case. Since there can be no surface current in a dielectric, the general condition for tangential magnetic field components, (4-67), is specialized to

\[ H_{t1} - H_{t2} = 0 \]  \hspace{1cm} (8-27)

That is to say, the tangential magnetic field components must be continuous also. In terms of the vectors defined by Fig. 8.5,

\[ H_1 + H_2 = H_t \]  \hspace{1cm} (8-28)
\[ E_1 - E_2 = E_t \]  \hspace{1cm} (8-29)

These two equations are sufficient to solve for all the field components in terms of \( H_t \), for the properties of the two media determine their intrinsic impedances \( \zeta_A \) and \( \zeta_B \) (left and right, respectively). The fields must always satisfy

\[ E_1 = \zeta_A H_1 \]  \hspace{1cm} (8-30)
\[ E_2 = \zeta_A H_2 \]  \hspace{1cm} (8-31)
\[ E_t = \zeta_B H_t \]  \hspace{1cm} (8-32)

These five equations may be combined to yield

\[ E_2 = -\Gamma E_1 \]  \hspace{1cm} (8-33)
\[ H_2 = -\Gamma H_1 \]  \hspace{1cm} (8-34)

where

\[ \Gamma = \frac{\zeta_B - \zeta_A}{\zeta_B + \zeta_A} \]  \hspace{1cm} (8-35)

is called the reflection coefficient. It evidently expresses the relationship between the incident and reflected waves. A transmission coefficient may be similarly defined, but it is only rarely encountered.

In reflection from a perfectly conductive surface, the reflected magnetic field tended to augment the incident field. This is not always true of reflection from dielectric surfaces. Equation (8-35), in fact, shows that reflections may be either positive or negative, subject to the sole restriction that the reflection coefficient can never exceed unity, and hence no reflected wave can be greater than the incident wave causing it. From a consideration of the energies associated with incident and reflected waves, this fact ought to have been obvious in the first place!

6. General Plane Electromagnetic Waves

The foregoing simple solutions of the wave equation represent part of an important special group of electromagnetic traveling waves. Waves of this kind are termed transverse electromagnetic (sometimes abbreviated TEM)
because the $\mathbf{E}$ and $\mathbf{H}$ vectors associated with them possess only components transverse to the direction of propagation. While Poynting's theorem guarantees that transverse components must exist (otherwise no energy can be propagated), it is in general possible to have longitudinal as well as transverse vector components in wave equation solutions.

In general, the wave equation is a vector differential equation in four variables and possesses a great many possible solutions of diverse physical nature and mathematical form. The solutions to be examined here will be restricted to those obtainable by at least partial separation of variables, that is, such that the wave equation in some vector $\mathbf{Q}$ (which might be any one of $\mathbf{E}$, $\mathbf{H}$, $\mathbf{D}$, or $\mathbf{B}$) is satisfied by a solution of the form

$$\mathbf{Q} = \mathbf{P}(x, y)Z(z)T(t)$$  \hspace{1cm} (8-36)

i.e., solutions in which the vector space functions $\mathbf{S}$ of Equation (7-21) can be separated into a function $Z$ of $z$ only, and a vector quantity $\mathbf{P}$ not dependent on $z$ at all: $\mathbf{S}(x, y, z) = \mathbf{P}(x, y)Z(z)$. The vector quantity $\mathbf{P}$ in general still possesses components in all three directions, but each of these will depend only on the space variables $(x, y)$ in the plane normal to the direction of wave propagation. As a result, $\mathbf{P}(x, y)$ cannot change as the wave propagates; the vector field described by it simply moves in the direction of propagation, but it is not altered as it moves. An easily visualized analogous hydrodynamic situation is that of laminar flow along a pipe. Any particular fluid mass identified by a colored dye splash will propagate along the pipe, but its spatial distribution, and its relationship to other fluid masses around it, will remain unaltered.

Assuming (8-36) to represent possible solutions, the Helmholtz equation becomes

$$\left(\nabla^2 + k^2\right)\mathbf{P}Z = 0$$  \hspace{1cm} (8-37)

which may be rearranged, yielding

$$\nabla^2 \mathbf{P} + \left(\frac{Z''}{Z} + k^2\right)\mathbf{P} = 0$$  \hspace{1cm} (8-38)

Inspection shows that only the portion in parentheses is dependent on $z$. By the usual separation argument, then, it must equal a constant. Thus, set

$$Z'' = \gamma^2 Z$$  \hspace{1cm} (8-39)

This leads immediately to solutions of the form

$$Z = e^{\pm \gamma z}$$  \hspace{1cm} (8-40)

and reduces the Helmholtz equation (8-37) to

$$\left(\nabla^2 + k^2 + \gamma^2\right)\mathbf{P} = 0$$  \hspace{1cm} (8-41)

This is another Helmholtz equation, of course, but with the significant simplification that the unknown vector function $\mathbf{P}$, although in general composed
of three components, is a function only of the two space variables \((x, y)\). The practical importance of this fact is that solution by some technique becomes possible in all cases. Especially in problems with complicated boundary shapes, where analytic solutions are very often impossible to achieve, two-dimensional variation of the components of \(\mathbf{P}\) brings the problem within the range of numerical iterative techniques. In the case of sinusoidally varying fields, the solutions for electric or magnetic field must now read

\[
\mathbf{Q} = \mathbf{P} e^{\gamma x + j\omega t} \tag{8-42}
\]

The dimensional nature of \(\mathbf{P}\) depends naturally on the field for which a solution is being sought; the basic mathematical feature of its dependence on only \(x\) and \(y\) remains.

As already shown, the constant \(k^2\) is in general dependent on frequency and medium parameters and may be real or complex. Consequently, \(\gamma\) may be complex also. It is usual to set

\[
\gamma = \alpha + j\beta \tag{8-43}
\]

so that (8-42) reads

\[
\mathbf{Q} = \mathbf{P} e^{\alpha x} e^{j(\beta z + \omega t)} \tag{8-44}
\]

It should be noted that a positive sign attached to \(\alpha\) corresponds to a positive sign with \(\beta\), and refers to a wave propagating in the negative \(z\)-direction. A positive \(\alpha\) thus represents a wave attenuated in the process of being propagated; \(\alpha\) is therefore known as the attenuation constant, while \(\beta\) is called the phase constant. Their combination \(\gamma\) is termed the propagation constant, although it is not really constant at all but depends on the various data entering the particular problem at hand. From (8-44), it is readily seen that any given wave phase (e.g., a zero) is propagated so that \((\beta z \pm \omega t)\) remains constant. Its velocity of propagation, often called the phase velocity, is then

\[
v_p = \frac{$$}{\beta} \tag{8-45}
\]

In solving for the possible constants \(k\) of the relay core or lamination in Chapter 7, infinitely many possible values were discovered to satisfy the Helmholtz equation, each corresponding to a different distribution of field. Similarly, (8-44) generally possesses many possible solutions, each representing a different field distribution and yielding a different value of \((k^2 + \gamma^2)\). However, the constant \(k^2\) is usually known, since it is a function [as in Equation (7-25)] of the frequency and material properties. Consequently, each of the many possible solutions possesses a different propagation constant and represents not only a distinctive field distribution but also denotes propagation with a distinctive velocity. Since \(k\) is a function of frequency, the velocity is clearly a function of frequency also.

In the special case of a simple unbounded plane wave examined previously, the propagation constant is evidently
According to (8-45), the propagation velocity is then that of light, the same at all frequencies.

7. TEM Waves on Transmission Lines

A very important practical application of the foregoing principles occurs in transmission lines and cables. It will now be shown that such structures are capable of propagating a plane wave traveling at the velocity of light and that such a wave must be accompanied by boundary currents and potentials in the conductors that make up the line. The usual technique of guessing at a solution and showing it valid will be followed. If the kind of wave sought is to be possible at all, it must be obtainable as the solution of the two Helmholtz equations

\[(\nabla^2 + k^2 + \gamma^2)E(x, y) = 0\]
\[(\nabla^2 + k^2 + \gamma^2)H(x, y) = 0\]  

Equations (8-47) simplify to

\[\nabla^2 E(x, y) = 0\]
\[\nabla^2 H(x, y) = 0\]  

It can easily be shown that these two equations describe electrostatic and magnetostatic fields, exactly as discussed in Chapters 2 and 6. The proof is most easily constructed by starting from the general electrostatic basis

\[\text{curl } E = 0\]  

which must hold true if the field in question is to be describable by the potential function \(V\). There naturally follows

\[\text{curl curl } E = 0\]  

which may, however, be rewritten as

\[\text{grad div } E - \nabla^2 E = 0\]  

Now in any charge-free space, the electric field must be nondivergent. There only remains

\[\nabla^2 E = 0\]  

In other words, a field in which the electrostatic potential satisfies Laplace's equation must also be one in which (8-51) is valid, and vice versa. A similar proof may be constructed for the magnetic field. On examining Equations (8-48), one is forced to conclude that the fields associated with a plane travel-
ing on a transmission system must satisfy Laplace's equation in a plane normal to the direction of wave propagation. (If they satisfied Laplace's equation in all three coordinates, then the fields would have to be entirely static and could not travel at all.) The boundary conditions to which the traveling wave is subject are similar to those encountered for static fields. Thus, the problem of TEM wave propagation in a cable or line is readily solved by solving for the plane electrostatic and magnetostatic fields in the same structure. If solution is achieved in complex potentials, furthermore, only one field solution will suffice, for the potential and stream functions of the electrostatic problem coincide with the stream and potential functions of the magnetostatic problem. For example, a plane wave propagating in a coaxial cable must have an associated electric field obtainable by substitution of (2-5) into (8-44):

\[ E(r, \theta, z, t) = \mathbf{1}_e \frac{-V}{r \log \left( \frac{a}{b} \right)} e^{\pm j(kz - \omega t)} \]  

(8-52)

where \( V \) is the potential difference between outer and inner conductors, assumed to have radii \( a \) and \( b \), respectively. The corresponding magnetic field distribution follows from Equation (5-6), again in combination with (8-44):

\[ H(r, \theta, z, t) = \mathbf{1}_h \frac{i}{2\pi r} e^{\pm j(kz - \omega t)} \]  

(8-53)

Inspection will verify that these fields are exactly static in the \( (r, \theta) \) plane and propagate in the \( z \)-direction in the laminar-flow or TEM manner.

Equations (2-5) and (5-6), and hence the expressions for traveling waves (8-52) and (8-53), require that a potential difference \( V \) exist between the conductors of the coaxial cable and that a current \( i \) flow in either conductor. For a single traveling wave, these two quantities may not be in an arbitrary ratio. This is readily shown by dividing the electric and magnetic field magnitudes of Equations (8-52) and (8-53):

\[ \frac{E}{H} = \frac{2\pi}{\log \left( \frac{a}{b} \right)} \frac{V}{i} = \zeta_0 \]  

(8-54)

The ratio of potential difference to cable current is then

\[ \frac{V}{i} = \frac{\zeta_0}{2\pi} \log \left( \frac{a}{b} \right) \]  

(8-55)

This quantity is independent of the fields themselves and varies only with the materials and size of the cable. Because it is a number characterizing the cable and has dimensions of impedance, it is called the characteristic impedance of the cable or transmission line. The usual symbol for characteristic impedance is \( Z_0 \).

Of course it is well-known that the potential difference and current in a transmission line are not always in the ratio \( Z_0 \). It should be kept clearly in mind that the entire discussion above relates to a single wave only. Just as an arbitrary boundary condition imposed on traveling waves in free space
leads to reflections and standing waves, so on a transmission line, \( E \) and \( H \) at any point represent the sums and differences of the fields associated with possibly quite numerous incident and reflected waves. Each wave of \( E \) and \( H \) naturally is required to carry along corresponding boundary charges and currents and thereby contribute to \( V \) and \( i \).

A second word of caution is in order. The validity of a TEM wave solution in any cable or transmission line was produced above by assuming such a solution to exist and finding the boundary potentials and currents necessary to make it so. Nothing in that development, however, precludes the existence of other modes of propagation. In fact, there exist infinitely many modes (i.e., infinitely many possible kinds of independent field distributions) in any given transmission system, each with its own requirements as to boundary conditions and each with its own phase velocity. The only conclusions that may be drawn with certainty from the foregoing are that (a) a TEM wave is possible if accompanied by waves of \( V \) and \( i \), and (b) because solutions to Laplace’s equation are unique, no other TEM mode can exist; all other modes must contain longitudinal components of \( E \), \( H \), or both.

8. Calculation of Characteristic Impedances

The behavior of transverse electromagnetic waves along any transmission line or cable is basically governed by only two numerical parameters, the propagation constant \( \gamma \) and characteristic impedance \( Z_0 \). Methods for calculating these parameters from the known geometrical and material data of a line therefore assume great importance. If analytic expressions for the fields can be obtained, the defining equation (8-55) for characteristic impedance may of course be used directly. In the great majority of practical cases, however, no direct analytic answers are available, and field calculations will be made graphically or numerically. It is necessary to develop such expressions for \( Z_0 \) as will lend themselves to use with numerical solutions.

The defining equation (8-55) may be rewritten

\[
Z_0 = \frac{V}{i} = \frac{\int E \cdot dr}{\int H \cdot dr} \tag{8-56}
\]

where the paths of integration are as shown in Fig. 8.6. Suppose that a map of the electric and magnetic fields is available and has been plotted so as to form curvilinear squares. It is then known that the potential difference across each square is the same as across any other (Chapter 2, Section 5). Consequently, (8-56) may be written as

\[
Z_0 = \frac{n_e}{n_m} \frac{\Delta V}{\Delta \Omega} \tag{8-57}
\]
where \( n_e \) is the number of electric potential squares, and \( n_m \) is the number of magnetic scalar potential squares. \( \Delta V \) and \( \Delta \Omega \) are the potential increments per square. Consider now any map square of side \( h \). Taking \( E \) and \( H \) as the field magnitudes inside the square,

\[
\frac{\Delta V}{\Delta \Omega} = \frac{E h}{H h} = \frac{E}{H} = \zeta_0
\]  

Then (8-58) becomes

\[
Z_0 = \frac{n_e}{n_m} \zeta_0
\]  

Given a field map, it is thus only necessary to count squares \( n_e \) between the conductors and \( n_m \) around either conductor in order to determine \( Z_0 \). In the general case, two field maps will be required, one for the electric and one for the magnetic field. The two fields may not be identical; charges in the electric field map will reside almost entirely on conductor surfaces, while currents are distributed throughout the conductors. In good conductors, however, sufficiently pronounced skin effect occurs even at comparatively low frequencies to permit the assumption that currents flow only along conductor surfaces, so that in the case of a homogenous dielectric, the two field maps are identical. This assumption is further reinforced by the fact that, as seen in Chapter 5, the internal flux linkages of a conductor contribute relatively little to total inductance.

Alternatively, (8-56) may be rewritten, using the relationship \( E = \zeta_0 H \) to eliminate \( H \) from the denominator:

\[
Z_0 = \frac{\oint E \cdot dr}{\oint \mathbf{1}_r E \cdot dr} \zeta_0
\]  

The vector \( \mathbf{1}_r E \) is normal to \( E \) since it follows the electric equipotential lines on the field map. Using the definition \( E = -\nabla V \),

\[
\mathbf{1}_r E = i \frac{\partial V}{\partial y} - j \frac{\partial V}{\partial x}
\]  

Fig. 8.6
and exchanging derivatives of the potential and stream functions $V$ and $U$ as prescribed by the Cauchy–Riemann conditions,

$$1 \mathbf{n} \cdot \mathbf{E} = - \nabla U$$

(8-62)

It should be noted that the gradient operations are in both cases two-dimensional, for plane fields are being considered.

Combining (8-60) and (8-62), there results

$$Z_0 = \frac{\int \nabla V \cdot d\mathbf{r}}{\int \nabla U \cdot d\mathbf{r}} = \frac{V_{\text{total}}}{U_{\text{total}}}$$

(8-63)

This expression is directly usable for all problems in which a field map is obtained by direct mapping techniques, i.e., conformal transformation, electrolytic tank analogue, or graphical relaxation. To use it for determining characteristic impedance where the field solution has been obtained as a discrete set of regularly spaced values, as in resistance-network or numerical iteration methods, replacement of the integrals by summations is indicated. This procedure is very similar to the line integration used to determine total flux in the magnetic core problem of Chapter 6, Section 12. At any point on the contour of integration, let the vector $\mathbf{dn}$ be normal to the contour: $\mathbf{dn} \cdot d\mathbf{r} = 0$. Then

$$\int \nabla U \cdot d\mathbf{r} = \oint \frac{\partial U}{\partial r} dr$$

(8-64)

and by the Cauchy–Riemann conditions,

$$\int \nabla U \cdot d\mathbf{r} = \oint \frac{\partial V}{\partial n} dr$$

(8-65)

The latter integral is readily evaluated from a set of point values of $V$. As indicated in Fig. 8.7, the contour may be chosen to fall between mesh points. The normal derivative at the point $P$ is then approximately

$$\frac{\partial V}{\partial n} = \frac{V_B - V_A}{h}$$

(8-66)
Replacing the contour integration by addition, the very simple form

$$\oint \frac{\partial V}{\partial n} \, dr = \sum_{\text{contour } \text{BEH} \ldots \text{B}} V_k - \sum_{\text{contour } \text{AUG} \ldots \text{A}} V_k$$

(8-67)

is found since the mesh width \( h \) cancels. It should be observed, however, that interior corner points such as \( C \) appear twice and exterior corner points such as \( E \) not at all in the respective summations.

Equations (8-67) and (8-63) permit direct evaluation of characteristic impedance from a field map of the scalar potential \( V \) only, without any need for explicit knowledge of the conjugate potential function. For the strip conductor in square duct of Fig. 2.16, for example, addition of potentials, as required by Equation (8-67), gives \( Z_0 = 0.56 \zeta_0 \).

The simple methods outlined above, it should be noted carefully, are based on the rather sweeping assumptions of both strong skin effect and homogeneous dielectric, assumptions which are not in all cases valid. More general derivations are possible and will be found in books on wave propagation in transmission lines. Any calculation of this kind, however, is ultimately nothing more than a solution of two problems in Laplace’s equation, one for the magnetic and one for the electric field. Since solution methods for such problems have been considered in some detail in Chapters 2, 3 and 6, they will not be reiterated here.

As an alternative possibility, the integral solution technique discussed in Chapters 2 and 6 is also useful for characteristic impedance calculations. The structure to be analyzed is subdivided into a large number of parallel strips, each of which is imagined to constitute a separate conductor in much the manner of the individual wires of a bundle transmission line. A potential coefficient matrix relating these subconductors is then constructed and inverted to find the influence coefficient matrix. Charges on the individual subconductors are then readily found. To find the electric field properties of the entire line, it is only necessary to determine the total charge corresponding to an arbitrarily chosen potential difference by summing the subconductor charges. Correspondingly, the magnetic properties of the structure are found in the manner used in Chapter 6 for calculating the magnetic field of a current-carrying conductor. Subdividing the conductors again (not necessarily in the same manner), each section is imagined to constitute a current-carrying subconductor. A magnetic field map, and hence the magnetic behavior of the whole line, may then be calculated by summing the contributions of the individual sections.

9. Waves with Longitudinal Field Components

The waves that transport electromagnetic energy along transmission lines are almost invariably of the TEM type. As shown in the foregoing discussion, such a wave is supported by two necessary boundary conditions: a set of
charges is needed to sustain the necessary electric field, and a set of currents to provide the accompanying magnetic field. Either field is of the static type in any plane normal to the direction of propagation.

A second possible mechanism for propagating guided waves exists. If the center conductor of a coaxial cable is removed, no current can flow along the cable and no transverse magnetic field can be sustained in it. TEM waves can then no longer exist. If, however, the electric field in the cable is permitted to have a longitudinal as well as a transverse component, there will exist a rate of change of electric flux $\frac{D}{\dot{D}}$ in the longitudinal direction. From the magnetic field point of view, there is no distinction between $\mathbf{J}$ and $\mathbf{D}$, so a magnetic field can once again exist. The current density $\mathbf{J}$ can sustain a magnetic field at any frequency; $\mathbf{D}$, on the other hand, is time-dependent, and a wave can propagate only if its frequency is high enough to cause a sufficiently large time derivative of electric flux. For this reason, waves of this sort can only be propagated along any given structure above some critical frequency dependent on its size, shape, and materials. Because the magnetic field is still purely transverse, although the electric field must have both transverse and longitudinal components, such waves are called transverse magnetic or TM waves.

A sketch of possible fields in a rectangular waveguide (the remaining outer conductor of a rectangular coaxial cable) is shown in Fig. 8.8. The magnetic field, it may be seen, is substantially similar to what it might be in the presence of an inner conductor, but the electric flux lines are quite different. It is not possible for the electric field to be radial outward (as it would
be in a cable) everywhere, for there is no inner conductor to carry charges. Every flux line directed outward to meet the guide wall must originate at a charge, which can in turn only be on the guide wall. Thus the electric flux must be essentially axial in some parts of the guide and essentially transverse in others. Since the magnetic field is now supported entirely by electric flux lines, not by current in a center conductor, it can only attain appreciable values in those sections of the guide where the electric field is essentially axial and must vanish at places where only a transverse electric field exists. As indicated in Fig. 8.8, one clear result is that the electric and magnetic fields are not only nonuniformly distributed along the length of the guide but alternate in direction from point to point, unlike in the cable or transmission line.

A third possibility exists for propagating waves: the magnetic field may possess a longitudinal component, while the electric field does not. Waves of this kind are of considerable importance in microwave technology and are termed transverse electric or TE waves. A longitudinal component of magnetic field can only exist if a current flows transversely across the waveguide. Naturally, this current cannot be a conduction current \( J \), for no conductive material normally exists in the interior; only a displacement current \( \dot{D} \) is permitted. A simple sketch of such fields is shown in Fig. 8.9. Both in this case and in the TM mode, the continuity equation requires that the displacement current \( \dot{D} \) be replaced in the conductive waveguide walls by a conduction current \( J \). Consequently, the currents flow, as required, in closed paths across the waveguide (as displacement current) and around the wall surfaces to the other side (in the form of conduction current). As in the TM mode, the displacement current is a time-dependent quantity and can support a wave only if the frequency is sufficiently high to yield a great enough rate of change of electric flux.
To summarize, waves can be propagated along a pipe without an inner conductor if a longitudinal component of either electric or magnetic field is permitted and if the frequency is high enough. This propagation is possible because at high enough frequencies the displacement current $\mathbf{D}$ can support the magnetic field alone, without any inner conductor to carry a conduction current.

Not unexpectedly, the conditions under which TM and TE modes are possible, as well as the details of the associated electric and magnetic fields, may be found by solving the vector Helmholtz equation subject to suitable boundary conditions. This can be accomplished by relating field quantities to a plane potential function, as was done in order to solve for fields in two-conductor lines and cables.

The existence of two separate conductors, it should be added, merely implies that the TEM mode of propagation is possible, but it in no way excludes other modes. An ordinary coaxial cable is able to sustain TE or TM waves as well as the TEM variety. The phase velocities of the various modes are not equal, however, and in most practical transmission problems, wave guiding structures are so chosen that the frequency of operation falls below the cutoff frequency of all modes but one. If this precaution is not observed, the total energy to be transmitted may be split between various different kinds of waves. These will not arrive at the receiving end of the guide concurrently, for each mode possesses a distinctive propagation velocity different from all the others. The consequent interference of a signal with itself at the receiving end is undesirable in communication systems, and often considerable effort is devoted to suppressing unwanted modes.

10. *The Scalar Helmholtz Problem for a Waveguide*

The general formulation of the vector Helmholtz equation (7-20), which is very difficult to solve for nearly any practical boundary conditions, was reduced to two-dimensional space-dependence in Equations (8-47):

$$ (\nabla^2 + k^2 + \gamma^2)E(x, y) = 0 $$  \hspace{1cm} (8-47a)

with a similar equation for the magnetic field $\mathbf{H}$. While it is still a vector differential equation, (8-47a) can be decomposed into three scalar equations in Cartesian coordinates. Writing $u^2 = k^2 + \gamma^2$ for brevity,

$$ (\nabla^2 + u^2)E_x = 0 $$  \hspace{1cm} (8-68a)

$$ (\nabla^2 + u^2)E_y = 0 $$  \hspace{1cm} (8-68b)

$$ (\nabla^2 + u^2)E_z = 0 $$  \hspace{1cm} (8-68c)

Although each component of $\mathbf{E}$ depends only on $x$ and $y$, the vector itself still possesses all three components. These three equations require simul-
taneous solution, which is not generally easy to obtain. If an infinitely long waveguide of perfectly conducting wall material and perfect dielectric is assumed, however, then one of these equations can be made independent of the others and solved as a scalar equation. Solutions for the others may then be deduced from it; in effect, one component of $E$ is employed as a potential function from which the other two can be derived.

As argued in detail in Section 5 above, the electric field is not permitted any tangential component at a perfectly conductive surface. The longitudinal component of $E$ must therefore always be zero at all points along the walls of an infinitely long, uniform waveguide. Thus it is possible to solve (8-68c) subject to the boundary condition that, at all points along wall,

$$E_z = 0$$

Equations (8-68c) and (8-69) constitute a two-dimensional scalar boundary value problem of exactly the same type as the rectangular core problem of Chapter 7, Section 4, and may be solved in just the same manner.

If the magnetic field is being sought, it becomes necessary to solve

$$(\nabla^2 + u^2)H_x = 0$$

(8-70a)

$$(\nabla^2 + u^2)H_y = 0$$

(8-70b)

$$(\nabla^2 + u^2)H_z = 0$$

(8-70c)

subject to the same boundary condition (8-69). Since the boundary condition and differential equations are not in terms of the same vector, it is first necessary to find a condition on $H$ explicitly. The Maxwell equation $\text{curl } H = \varepsilon \text{E}$ indicates that $\text{curl } H$ cannot possess a tangential component along the guide walls, by virtue of (8-69). Since both $H$ and $E$ are to be functions of $(x, y)$ only, it is always possible to write the tangential component of $\text{curl } H$ as

$$\left(\text{curl } H\right)_{\text{tang}} = \frac{\partial H_z}{\partial n}$$

(8-71)

where $n$ denotes distance measured in the $x$-$y$ plane in a direction normal to the guide wall. It remains to be concluded that

$$\frac{\partial H_z}{\partial n} = 0$$

(8-72)

The boundary value problem in $H_z$ therefore consists of Equations (8-70c) and (8-72).

Either boundary value problem is automatically satisfied by fields of the TEM mode, for in that mode, $E_z = H_z = 0$. Whether to solve the boundary value problem in $E_z$ or $H_z$ in other cases depends on whether the TE or TM type of waves are sought, the other equation in each case being satisfied by the absence of a longitudinal field. Thus the longitudinal components of either $E$ or $H$ can be solved for explicitly. It remains to find the remaining components from them. This can be accomplished via Maxwell’s curl equations.
Noting that both the electric and magnetic fields are of the form \( P(x, y, z) \exp(\pm \gamma z + j\omega t) \) so that differentiations with respect to \( z \) and \( t \) can be carried out immediately, the electric curl equation (MD-1) may be written
\[
\frac{\partial E_z}{\partial y} \pm j\gamma E_y = j\omega \mu H_z
\]
\[
\pm j\gamma E_x - \frac{\partial E_z}{\partial x} = -j\omega \mu H_y
\]
\[
\frac{\partial E_y}{\partial x} - \frac{\partial E_z}{\partial y} = -j\omega \mu H_z
\]

(8-73)

For brevity, the first two members of this set may be written as
\[
H_t = \frac{j}{\omega \mu} k \times (\pm j\gamma E_t + \nabla_t E_z)
\]

(8-74)

where the subscript \( t \) on all vectors and operators denotes that \( z \)-components are to be discarded, e.g.,
\[
\nabla_t = i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y}
\]

By following exactly the same steps starting from the magnetic curl equation, there is obtained
\[
E_t = \frac{j}{\omega \varepsilon} k \times (\pm j\gamma H_t + \nabla_t H_z)
\]

(8-75)

The vectors \( E_t \) and \( H_t \) are precisely the quantities being sought since whichever of \( E_z \) or \( H_z \) is not identically zero will have been found by solving the appropriate boundary value problem. Simultaneous solution of (8-74) and (8-75) will therefore provide a complete field solution. To accomplish it, Equation (8-74) may be substituted into (8-75), thereby eliminating \( H_t \). Recognizing that \( k \times (k \times V_t) = -V_t \) for any vector \( V \), the result may be simplified to
\[
E_t = \frac{\pm j}{u^2} \nabla_t E_z + \frac{j \omega \mu}{u^2} k \times \nabla_t H_z
\]

(8-76)

and similarly,
\[
H_t = \frac{\mp j}{u^2} \nabla_t H_z - \frac{j \omega \varepsilon}{u^2} k \times \nabla_t E_z
\]

(8-77)

These solutions are quite general in that they permit either \( E_z \) or \( H_z \), or both, to assume nonzero values. In addition to the pure TE and pure TM modes of propagation, any mixture of the two may thus be calculated. Such mixtures, however, do not represent new kinds of waves but are merely additive superpositions of the TE and TM types. Consideration of only the two basic kinds of waves therefore suffices to cover all possible cases.
11. TM Waves in a Rectangular Guide

While there is no intention to treat guided wave propagation exhaustively here, a simple illustration will undoubtedly do much to clarify the principles developed in the foregoing. The most widely employed actual waveguides are rectangular, and the most easily handled boundary conditions those of the Dirichlet type (i.e., prescribed values along the boundary); an easy example, both practical and instructive, is thus furnished by the boundary value problem

\[(\nabla^2 + k^2 + \gamma^2)E_z = 0\]  
\[E_z(\pm a, y) = 0\]  
\[E_z(x, \pm b) = 0\]  

Equations (8-78) are merely (8-69) written for the specific geometric configuration of Fig. 8.10. It will be assumed that \(H_z = 0\).

Except for an alteration of symbols, this Helmholtz problem may be recognized as identical to the rectangular magnetic core problem of Fig. 7.3. The solution for \(E_z\) may therefore be copied directly from Equations (7-46) to (7-48):

\[E_z = E_{mn} \cos\left(\frac{m\pi x}{2a}\right) \cos\left(\frac{n\pi y}{2b}\right) \exp(\pm \gamma z + j\omega t)\]  

As in the magnetic core problem, \(m, n\) are any two integers which satisfy

\[k^2 + \gamma^2 = \left(\frac{m\pi}{2a}\right)^2 + \left(\frac{n\pi}{2b}\right)^2\]  

The propagation constant \(\gamma\) is readily solved for. Remembering that \(k^2 = \omega^2 \mu \varepsilon\) in a lossless medium,

\[\gamma = \sqrt{\left(\frac{m\pi}{2a}\right)^2 + \left(\frac{n\pi}{2b}\right)^2 - \omega^2 \mu \varepsilon}\]  

Equations (8-79) and (8-81a) illustrate what was anticipated in a general way before. There must exist a doubly infinite set of possible solutions, one for each combination \((m, n)\), just as there existed a doubly infinite set of independent flux distribution modes in the relay core; for that matter, the mathematical similarity of the two problems indicates that the field distributions in any given cross-sectional plane will be similar. Any one or several of these solutions may exist concurrently. Each mode in the guide, however, represents a traveling wave which propagates in the longitudinal direction with some phase velocity \(v_{mn}\); inspection of (8-81a) shows the phase velocity
to be different for every combination \((m, n)\). As a matter of fact, the phase velocity is seen to depend also on the frequency \(\omega\).

It is apparent from (8-81a) that the propagation constant may assume real or imaginary values, depending on the combination of \(m, n\), and \(\omega\). Referring to Equation (8-44), it is seen that a purely real propagation constant indicates a wave being attenuated, while a purely imaginary value of \(\gamma\) indicates unattenuated propagation. The existence of a minimum frequency for successful TM wave propagation emerged above from the general physical nature of guided waves, but no direct method of calculating this frequency was evident. It is now clear that this cutoff frequency lies where the propagation constant changes from real to imaginary values. Its value is deduced from (8-81a) as

\[
\omega_c = \frac{\pi c}{2} \sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2}
\]  

(8-81b)

The lowest possible value of cutoff frequency is obtained for \(m = n = 1\). Taking as an example \(a = 2\) cm, \(b = 1\) cm, there is obtained \(\omega_c = 5.27 \times 10^8\) radians per second. Clearly, waveguides of any reasonable size represent practical energy transmission devices only at quite high frequencies. They are in practical use over a very wide spectrum and hence made in a large variety of sizes. Some typical sections of rectangular waveguide are illustrated in Fig. 8.11; of these, the largest has \(2a = 13\) cm, \(2b = 6.5\) cm.

The various possible solutions for different combinations of \(m\) and \(n\) all possess different patterns of field distribution. By substitution of (8-79)
in (8-76) and (8-77), transverse fields in the guide are readily found. The electric field components are

\[ E_x = \frac{\gamma}{u^2} E_{mn} \frac{m\pi}{2a} \sin \frac{m\pi x}{2a} \cos \frac{n\pi y}{2b} \exp (\gamma z + j\omega t) \]  

\[ E_y = \frac{\gamma}{u^2} E_{mn} \frac{n\pi}{2b} \cos \frac{m\pi x}{2a} \sin \frac{n\pi y}{2b} \exp (\gamma z + j\omega t) \]  

The magnetic field components are similar in form:

\[ H_z = E_{mn} \frac{j\omega \epsilon}{u^2} \frac{n\pi}{2b} \cos \frac{m\pi x}{2a} \sin \frac{n\pi y}{2b} \exp (\gamma z + j\omega t) \]  

\[ H_y = -E_{mn} \frac{j\omega \epsilon}{u^2} \frac{m\pi}{2a} \sin \frac{m\pi x}{2a} \cos \frac{n\pi y}{2b} \exp (\gamma z + j\omega t) \]  

The ± signs with each propagation constant have been omitted in order to keep sign changes distinct. Both \( E \) and \( H \) are periodic in space across the guide, the number of semiperiods being given by \( m \) and \( n \). Sets of magnetic field lines plotted for several of the low-order modes are shown in Fig. 8.12. Note that even as well as odd values of \( m, n \) are admissible. In contrast to the rectangular bar problem, there is no symmetry requirement imposed by Equations (8-78) to exclude even values.

![Fig. 8.12](image)

**12. TE Waves in a Rectangular Guide**

The transverse electric waves may be solved for by following a very similar mathematical procedure. The boundary value problem to be solved consists
of the Helmholtz equation
\[(\nabla^2 + c^2)H_z = 0\]  
subject to the boundary condition that, at guide walls,
\[\frac{\partial H_z}{\partial n} = 0\]

Comparing this problem with that of the TM waves, it is immediately clear that the solutions are once again going to be given by trigonometric functions. This time, however, the solutions must have zero normal derivatives, not zero value, at the surfaces. Zero derivatives occur at the peaks of sinusoidal waves, so that redefining the origin of coordinates as in Fig. 8.13 makes the solution for TE waves mathematically similar to that for TM waves. The solution, which is easily verified by substitution, is given by

\[H_z = H_{mn} \cos \frac{m\pi x}{2a} \cos \frac{n\pi y}{2b} \exp (\pm \gamma z + j\omega t)\]  

Fig. 8.13

This solution differs mathematically from that of the TM case in having the guide walls coincide with the peaks of the cosine terms rather than their zeros. It would of course be equally possible to give a similar solution with the coordinates defined as in Fig. 8.10, with the result that (8-86) would contain sine rather than cosine terms. However, shifting the coordinate system results in a solution symmetric with (8-79), permitting direct comparison. The transverse field components for this wave, as in the previous case, are found by substituting the solution (8-86) into the general expressions for transverse fields (8-76) and (8-77). For the electric field, there is found

\[E_x = -H_{mn} \frac{j\omega \mu m\pi}{u^2} \cos \frac{m\pi x}{2a} \sin \frac{n\pi y}{2b} \exp (\pm \gamma z + j\omega t)\]  
\[E_y = H_{mn} \frac{j\omega \mu m\pi}{u^2} \sin \frac{m\pi x}{2a} \cos \frac{n\pi y}{2b} \exp (\pm \gamma z + j\omega t)\]

with similar expressions for the transverse magnetic field. It will be noted that the earlier expressions (8-81) for propagation constant \(\gamma\) and for cutoff frequency \(\omega_c\) are unaffected by the mode type and are equally valid for TE and TM modes.

One important difference between the two types of modes does exist, however. In either case, the numbers \(m\) and \(n\) must be real integers; but in the TE case, zero is a permissible value, while in the TM case it is not. This fact is readily evident if Equations (8-82) and (8-83) are compared with (8-87) and (8-88). If either \(m\) or \(n\) is set to zero, the transverse field can consist of only one component, for either the \(x\)- or \(y\)-component will contain \(\sin (0)\) and therefore vanish. In the TE case, this possibility is perfectly acceptable,
for a field crossing the guide in the manner indicated in Fig. 8.9 results. There will then exist a transverse displacement current, terminating at the guide walls and replaced there by a conduction current, as required by the continuity equation. On the other hand, a magnetic field consisting of only one component is not possible, for such a field would require magnetic flux lines terminating at the guide walls, thereby violating \( \text{div } \mathbf{B} = 0 \). Hence the lowest permissible value of \( m \) or \( n \) is unity in the TM case, but zero for TE waves. Only one of these integers may assume this value, of course, for if \( m = n = 0 \), all transverse field components vanish. For the first few TE modes, the electric field lines in a rectangular guide are sketched in Fig. 8.14. Except for the lowest-order modes, for which no magnetic equivalents exist, they are seen to resemble the magnetic field patterns of Fig. 8.12, but with different symmetry planes.

From a practical point of view, the mode with \( n = 0 \) is the most important one of all, for it possesses the lowest cutoff frequency. With \( n = 0, m = 1 \), Equation (8-81b) yields this cutoff frequency:

\[
\omega_c = \frac{\pi c}{2a}
\]

independently of the lateral dimension \( b \). Taking for illustrative purposes the same guide as above, with \( a = 2 \text{ cm}, b = 1 \text{ cm} \), there results \( \omega_c = 2.36 \times 10^9 \text{ radians per second} \), corresponding to a free space wavelength of twice the breadth of the guide. This limiting wavelength gives a simple indication of the size of guide needed for any particular application.

13. Numerical Solution of the Helmholtz Equation

Although the most widely used waveguide cross section is undoubtedly rectangular, many other shapes are employed in practice, both for reasons of mechanical convenience and because of the many peculiar desirable electri-
cal properties possessed by different types of guide. Unfortunately the analytic possibilities of the scalar Helmholtz equation are even more restricted than those of Laplace's equation. Very few shapes are amenable to solution by separation of variables, and it is in fact even possible to prove that no further ones can exist! However, solution by relaxation or iteration is readily possible. The method employed is a generalization of the technique used for solution of Laplace's equation in electrostatics and magnetostatics; it is equally applicable to the Helmholtz equation and other related equations. The fundamental problem is to solve

$$\nabla^2 \varphi + u^2 f(x, y) \varphi = 0 \quad (8-90)$$

in a specified closed region, subject to given boundary conditions. It will be noted that \( f(x, y) = 1 \) results in the Helmholtz equation, while other, more general, functions \( f \) may describe other problems, as shown in Section 15 below. The problem will be tackled on the basis of the usual rectangular grid and solution started by first assuming an arbitrary set of values of \( \varphi \) to exist throughout the region. The arbitrary values of \( \varphi \) do not satisfy Equation (8-90), of course; instead of yielding zero as the right-hand member, some other numbers, the residuals, are obtained. Using the finite-difference approximation to the Laplacian developed in Equations (2-35) to (2-38), the general equation (8-90) is replaced by

$$\varphi_x + \varphi_y + \varphi_s + \varphi_w - 4 \varphi_o + u^2 f_o h^2 \varphi_o = R_o \quad (8-91)$$

where \( f_o \) denotes the value of \( f(x, y) \) at the point \( O \). As for Laplace's equation, some iterative process for relaxing the residuals toward zero will yield an approximate solution to the problem. It must be emphasized once again that regardless of the nature of this iterative process, what is being solved by it is the set of finite-difference equations (8-91), not the differential equation (8-90); even though the residuals may be depressed exactly to zero, the result will give a reasonable approximation to the differential equation problem only if a sufficiently fine mesh is employed. In other words, there exist two principal, and independent, sources of error: the replacement of a differential equation by finite-difference approximations, and the solution of the latter.

Depression of residuals toward zero is best accomplished, as in previous cases, by Liebmann iteration with overrelaxation. In this process, each field value \( \varphi_o \) is replaced in turn by the old value plus a correction dependent on the residual:

$$\varphi_o^{\text{new}} = \varphi_o^{\text{old}} + \beta R_o^{\text{old}} \quad (8-92)$$

where \( \beta \) is a number yet to be determined. If simple relaxation is used, so that the replacement process (8-92) results in exactly zero residual,

$$\varphi_x + \varphi_y + \varphi_s + \varphi_w - (4 - f_o u^2 h^2) \varphi_o^{\text{new}} = 0 \quad (8-93)$$

On substitution of (8-92) into (8-93), there results

$$\sum_{N\text{FSW}} \varphi - (4 - f_o u^2 h^2) \varphi_o^{\text{old}} - \beta (4 - f_o u^2 h^2) R_o^{\text{old}} = 0 \quad (8-94)$$
whose two leftmost terms may be recognized from (8-91) to be simply the residual at point \( O \):

\[
R_{ij}^{\text{old}} - \beta (4 - f_0 u^2 h^2) R_0^{\text{old}} = 0
\]  

(8-95)

so that the unknown number \( \beta \) may be seen to be

\[
\beta = \frac{1}{4 - f_0 u^2 h^2}
\]  

(8-96)

The Liebmann process here, just as in the static field problem, consists of (a) using partly old, partly new field values to calculate each residual, and (b) overcorrecting each field value so that the new residual is not zero but some number of smaller magnitude than and opposite polarity to the old residual. The latter process is conveniently followed by using some overrelaxation factor \( \alpha \), which must be greater than unity to force the residual to alter sign, but must not exceed two, lest the new residual be larger than the old. The solution process thus becomes a sequential replacement quite analogous to Equation (2-39):

\[
\varphi_0^{\text{new}} = \varphi_0^{\text{old}} + \frac{\alpha}{4 - f_0 u^2 h^2} R_0^{\text{old}}
\]  

(8-97)

In the special case of the scalar Helmholtz equations (8-68) and (8-70), \( f_0 = 1 \), rendering the arithmetic required for the replacement (8-97) very simple, but not otherwise altering the argument at all. Similarly, the treatment of boundary conditions is identical to that in the static case. The scalar Helmholtz problem for waveguides is therewith reduced to the sequence of finite-difference problems

\[
(\sum_{\text{NWSW}} \varphi - 4\varphi_0) + u^2 h^2 \varphi_0 = 0
\]  

(8-98)

corresponding to the Helmholtz differential equation

\[
(\nabla^2 + u^2) \varphi = 0
\]  

(8-99)

and subject to appropriate boundary conditions.

One significant point of difference arises, however, between the finite-difference Helmholtz equation and the corresponding approximations to Laplace's or Poisson's equations. In order to solve the sequence of finite-difference equations (8-98), it is necessary to know the value of \( u^2 h^2 \); yet that value cannot be known unless the problem has already been solved. This dilemma may be resolved by the use of a doubly iterative calculation scheme. In addition to the initially guessed point values of \( \varphi \), a numerical value of \( u^2 h^2 \) is also assumed. Several relaxation passes are then made with this assuredly wrong value, so as to improve the point potential-function values \( \varphi \) to a reasonable degree; if the overrelaxation factor is not too far from its optimum value, only a few iterations will be needed. An improved estimate of \( u^2 h^2 \) is then calculated from the approximate field values and the whole cycle repeated. Iterations will be halted when convergence of all values to the desired degree has been achieved.
The improved estimate of $u^2 h^2$ may be obtained as follows. Let the Helmholtz equation (8-90) be multiplied by $\varphi$ and integrated over the cross-sectional area of the guide. There is obtained

$$u^2 = -\frac{\int \varphi \nabla^2 \varphi \, dS}{\int \varphi^2 f(x, y) \, dS} \quad (8-100)$$

In the theory of differential equations, this ratio is called the Rayleigh coefficient. It can be shown that it always furnishes a more accurate value of $u^2$ than the Helmholtz equation itself; this is also intuitively reasonable since overall averages rather than point values are involved, and errors may be expected to cancel at least partly. Using the same approximation to $\nabla^2 \varphi$ as previously, (8-100) may be replaced by the finite-difference equivalent

$$u^2 h^2 = -\frac{\sum \varphi_{jk}(\varphi_N + \varphi_E + \varphi_S + \varphi_W - 4\varphi_{jk})}{\sum \varphi_{jk} f_{jk}} \quad (8-101)$$

where the summations cover all points $(j, k)$ of the field, and the values $\varphi_N$ to $\varphi_W$ refer to the points immediately surrounding $(j, k)$. The calculations will then be carried out as shown in the flow chart of Fig. 8.15. It might be added that many refinements are possible. For example, there exist methods for

---

Fig. 8.15
estimating the optimum overrelaxation factor from the rate of solution convergence; in large-scale calculations, it is profitable to encase the actual field calculations in yet another iterative loop that establishes a new overrelaxation factor every few cycles.

14. Principal Modes of a Ridged Waveguide

In order to provide an illustration of the doubly iterative technique, the principal modes of a waveguide shaped as indicated in Fig. 8.16, essentially a rectangular guide with a raised ridge, will be computed. This is a common shape in practical applications and results in quite simple computer programming. As in the case of rectangular guides, the TM case will be considered first because it results from a particularly simple type of boundary condition and is easy to calculate. Only the principal mode will be sought, corresponding to the case \( m = n = 1 \) of Equation (8-79); higher modes may also be found by this technique but are more difficult to obtain, and other methods are normally employed if more than one or two are desired. It is readily perceived that the problem possesses symmetry about the center line of the guide and that only half need be calculated explicitly. In consequence, the boundary conditions for the TM mode become \( \phi = 0 \) along all wall surfaces, as usual, but \( \partial \phi / \partial n = 0 \) along the center line. Conversely, the TE mode will be found by imposing \( \partial \phi / \partial n = 0 \) along all walls, but \( \phi = 0 \) along the center line. Because these differences do exist, separate programmes are given here to illustrate the calculation method.

To solve for the TM mode, the flow chart of Fig. 8.15 may be followed directly. The only potentially troublesome point to be observed is that a trivial solution, \( E_z = 0 \), may be all too easily encountered. Unlike in static field problems, the guide walls are not only equipotential surfaces but all at the same potential \( E_z = 0 \). Thus the trivial solution satisfies (8-99) and will be obtained if the initial guess for all field values is also zero; at least one of the field points must be set at a different value initially, in order to permit at least one residual to start at a nonzero value. All values of \( E_z \) are set to a nonzero value to avoid convergence to the trivial solution in the computer programme shown in Appendix II.

Figure 8.17 shows numerical values of the longitudinal electric field and the magnetic field lines in a ridged guide, as obtained by the computer programme. It is evident that the principal TM mode resembles, in a qualitative sense, the mode \( m = n = 1 \) of a rectangular guide; but no quantitative
DOMINANT TM MODE OF RIDGE WAVEGUIDE

CUTOFF FREQUENCY = 308.8 MHZ DIVIDED BY GUIDE HEIGHT IN METERS

CONVERGENCE AFTER 54 ITERATIONS

Fig. 8.17
correspondence can be readily found. In particular, the regions to either side of the ridge are seen to contain a relatively low magnetic flux density and hence a comparatively small Poynting vector; the major proportion of energy is clearly transported in the narrow region above the ridge.

The computer programme shown is relatively straightforward, no particular attempt having been made to hold the number of iterations down by looking for an optimum overrelaxation factor or by attempting to hasten convergence to the final value of \( u^2 h^2 \). Examination of the programme will show that the residuals are not checked at all, the sole convergence criterion being the rate of change of \( u^2 h^2 \) with continuing iteration. If reasonably accurate field patterns are desired, prudence requires at least printing out the largest residual encountered on the final iterative pass so as to give some indication of the level to which residuals have decreased. On the other hand, if merely the cutoff frequency of the guide is desired, this convergence criterion is quite sufficient. In view of the fact that the final value of \( u^2 h^2 \) is only known to within an order of magnitude, it is best to specify its convergence in terms of a percentage change, unlike the residuals which are often required to converge to within an absolute error in static problems. It needs to be added also that in this Helmholtz problem, there exists no fixed nonzero field value anywhere and that the solution magnitudes will eventually settle down to some entirely unpredictable values dictated by the magnitudes of the initial guesses and the overrelaxation factor.

By finding the characteristic value \( u^2 h^2 \) of the Helmholtz problem, the cutoff frequency of the guide is implicitly determined. By definition,

\[
u^2 = k^2 + \gamma^2
\]

so that

\[
u^2 h^2 = (k^2 + \gamma^2) h^2
\]

Inserting the wave number of lossless media, \( k^2 = \omega^2 \mu e \),

\[
\frac{\nu^2 h^2}{h^2} = \frac{\omega^2}{c^2} + \gamma^2
\]  

(8-103)

As previously discussed, the cutoff frequency corresponds to \( \gamma = 0 \), where the propagation constant passes from pure imaginary to pure real values. Consequently the cutoff frequency is found from

\[
\omega_c = \frac{\sqrt{\nu^2 h^2 c}}{h}
\]  

(8-104)

a number inversely proportional to waveguide size, as might have been expected from the known results for the rectangular guide. It should be noted, however, that this result is valid for any shape whatever. Equation (8-104) is readily encodable for computer evaluation and is incorporated also in the sample programmes in Appendix II, in the modified form
\[ \omega_c = \frac{\sqrt{u^2 h^2 c}}{H} n \]  

(8-105)

where \( H = nh \) is some convenient dimension of the guide, here taken as the height, and \( n \) is the number of increments \( h \) into which this distance is subdivided in the calculation. This form has the evident advantage of removing the quantity \( h \), which is after all purely internal to the numerical calculation, from the end result.

Programming the solution for TE modes is somewhat more complicated. First of all, the normal-gradient boundary conditions are lengthier to programme than the fixed values of the TM case. The programme shown in Appendix II handles them by reserving a memory allocation larger than the actual field, i.e., bordering the region of solution with an additional row of points, as indicated in Fig. 8.18. Iterations are only carried out in the interior region which actually corresponds to the guide cross section, the extra points merely serving, for the duration of one iterative pass, as fixed value boundaries. At the conclusion of each pass through the field, the newly calculated field values at the interior points are transferred to the artificial boundary, thereby enforcing the normal-gradient boundary condition for the next iteration. It should be clear that corner points of the type of point \( C \) (Fig. 8.18) must possess two field values simultaneously since they form external points for two distinct boundaries. This difficulty is dealt with by assigning to point \( C \) first one value, performing residual alterations until that value has been used, and changing the field value at \( C \) before continuing the iteration.

![Fig. 8.18](image_url)
A second, more fundamental difficulty is encountered in solution for the TE mode. Both the normal-gradient boundary condition and the Helmholtz equation are satisfied by the solution \( \phi = \text{constant} \), corresponding to \( u^2 h^2 = 0 \), and a solution once started will converge toward this undesired value. Physically, this solution corresponds to the impossible mode \( m = n = 0 \) of a rectangular guide. In order to keep the solution from converging toward this impossible mode, the additional requirement \( \text{div} \mathbf{B} = 0 \) needs to be imposed, i.e., that the total longitudinal magnetic flux of the guide must be zero. For the symmetric problem considered, this is equivalent to insisting that the longitudinal magnetic field \( H_z \) must be zero at the symmetry plane, but its normal derivative at the symmetry plane must not be zero, \( \partial H_z / \partial n \neq 0 \). A nonzero normal derivative is readily enforced by requiring that the average value of \( H_z \) at all points \( h \) removed from the symmetry plane have some nonzero value, chosen arbitrarily as 0.5 in the programme shown. This requirement is encoded as a scaling upward or downward of all the field values, after every iterative pass, so as to make the average value next to the symmetry line 0.5.

Finally, a complication arises in evaluating the Rayleigh coefficient according to Equations (8-100) and (8-101). In both denominator and numerator, the surface integrals are replaced by summations of the type

\[
\int F(x, y) \, dS = \sum_{j,k} F(x_j, y_k) \, \delta S_{jk}
\]

(8-106)

where \( \delta S_{jk} \) is the area element surrounding the \((j, k)\)th field point. Since the guide boundaries coincide with mesh points, these areas are not all equal; those in the interior of the guide are obviously simply \( h^2 \), but along the edges, values of \( 0.25h^2 \), \( 0.5h^2 \), and \( 0.75h^2 \) are encountered. In order to avoid cluttering the programme with area calculations, a separate function subroutine is provided which finds the area associated with any given field point when required. Only the symbol \( \text{AREA}(l, i) \), constituting a call for the function subroutine, appears in the main TE mode programme itself.

The TE mode programme, along with the area subroutine, appears in Appendix II, and results of the calculation are shown in Fig. 8.19. Just as in the analytic solution for the rectangular waveguide, the formulation for cutoff frequency is identical for TE and TM modes; and just as for the rectangular guide, the lowest TE mode is seen to have a substantially lower cutoff frequency than the lowest TM mode. Evidently the dominant TE mode is the most desirable one for practical work, for if the guide is chosen so as to have the operating frequency fall between these two cutoff frequencies, there is no likelihood that any but the principal mode (the lowest TE mode) will participate in energy transmission. Since a knowledge of the Rayleigh coefficient is a prerequisite for calculation of the field values, it should not be surprising that the latter should converge more slowly than the former. It is
### Dominant TE Mode of Ridge Waveguide

<table>
<thead>
<tr>
<th>Height (m)</th>
<th>Cutoff Frequency (MHz)</th>
<th>Domain (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.53</td>
<td>3.50</td>
<td>3.42</td>
</tr>
<tr>
<td>3.55</td>
<td>3.52</td>
<td>3.43</td>
</tr>
<tr>
<td>3.59</td>
<td>3.56</td>
<td>3.44</td>
</tr>
<tr>
<td>3.66</td>
<td>3.63</td>
<td>3.41</td>
</tr>
<tr>
<td>3.76</td>
<td>3.73</td>
<td>3.52</td>
</tr>
<tr>
<td>3.88</td>
<td>3.86</td>
<td>3.65</td>
</tr>
<tr>
<td>4.02</td>
<td>4.00</td>
<td>3.82</td>
</tr>
<tr>
<td>4.18</td>
<td>4.16</td>
<td>4.00</td>
</tr>
<tr>
<td>4.34</td>
<td>4.32</td>
<td>4.20</td>
</tr>
<tr>
<td>4.50</td>
<td>4.48</td>
<td>4.45</td>
</tr>
<tr>
<td>4.65</td>
<td>4.61</td>
<td>4.52</td>
</tr>
<tr>
<td>4.78</td>
<td>4.75</td>
<td>4.72</td>
</tr>
<tr>
<td>4.89</td>
<td>4.87</td>
<td>4.85</td>
</tr>
<tr>
<td>4.97</td>
<td>4.94</td>
<td>4.92</td>
</tr>
<tr>
<td>5.02</td>
<td>5.01</td>
<td>4.99</td>
</tr>
<tr>
<td>5.03</td>
<td>5.02</td>
<td>5.00</td>
</tr>
</tbody>
</table>

Cutoff Frequency = 62.6 MHz divided by guide height in meters
Convergence after 75 iterations

Fig. 8.19
in fact quite surprising how fast the Rayleigh coefficient does converge; for the TE mode calculation, the error values shown in the following table are obtained. This convergence rate, like any other, is naturally affected by the overrelaxation factor employed, as well as by the number of iterations between successive recalculations of the Rayleigh coefficient. Since an erroneous value of $u^2h^2$ cannot lead to the correct solution, there is little point in many intervening iterations tending to converge to an impossible result; on the other hand, continuous recomputation of the Rayleigh coefficient is not very useful either.

15. The Helmholtz Equation under Conformal Transformation

Many practical waveguide problems are not solvable by the method of variable separation because the Helmholtz equation is only separable in a very few coordinate systems. Equally, the iterative technique shown above may not be attractive because the problem boundaries may be so complicated as to lead to difficult programming or because higher-order modes may be desired, toward which the solution may converge poorly. In such cases, conformal mapping may be used to reduce the problem to one involving more convenient, e.g., rectangular or circular, boundaries. The solution is subsequently transformed back to the original problem boundaries.

Any conformal transformation from the Cartesian coordinate system $(x, y)$ to some other set of orthogonal coordinates $(p, q)$ may be given by the differential equation

$$\frac{dZ}{dW} = f(x, y)$$

(8-107)

where $W$ denotes the complex variable $W = p + jq$ and $Z$ denotes the complex variable $Z = x + jq$ (not the Cartesian coordinate $z$!). In order to trans-
from the Helmholtz equation, it is necessary to find out how the differential expression \( \nabla^2 \varphi \) will be transformed by (8-107). From differential calculus, it is known that

\[
\frac{\partial \varphi}{\partial p} = \frac{\partial x}{\partial p} \cdot \frac{\partial \varphi}{\partial x}
\]  

(8-108)

so that, on repeated application,

\[
\frac{\partial^2 \varphi}{\partial p^2} = \frac{\partial x}{\partial p} \cdot \frac{\partial}{\partial x} \left( \frac{\partial x}{\partial p} \cdot \frac{\partial \varphi}{\partial x} \right)
\]

\[
= \left( \frac{\partial x}{\partial p} \right)^2 \frac{\partial^2 \varphi}{\partial x^2}
\]  

(8-109)

A similar expression may also be developed to relate the second derivatives in \( q \) and \( y \). The basic assumption that the transformation of (8-107) must be conformal implies that the Cauchy–Riemann conditions must be satisfied. Hence

\[
\left( \frac{\partial x}{\partial p} \right)^2 = \left( \frac{\partial y}{\partial q} \right)^2 = \frac{\partial x}{\partial p} \frac{\partial y}{\partial q}
\]  

(8-110)

so that the Laplacian expressions \( \nabla^2 \varphi \) in the two coordinate systems are related by

\[
\frac{\partial^2 \varphi}{\partial p^2} + \frac{\partial^2 \varphi}{\partial q^2} = \frac{\partial x}{\partial p} \frac{\partial y}{\partial q} \left( \frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} \right)
\]  

(8-111)

Geometrically, the derivative \( \frac{\partial x}{\partial p} \) measures the transformation ratio between length increments in the \( x \)-direction and length increments in the \( p \)-direction. The Cauchy–Riemann conditions may, in this sense, be interpreted as saying that a conformal transformation is one in which length elements in the \( x \)-direction are magnified the same amount as elements in the \( y \)-direction. Hence all the ratios in Equation (8-110) are merely the magnification ratio by which a square area element \( dp \cdot dq \) is transformed between the \( Z \)- and \( W \)-planes:

\[
dx \cdot dy = \frac{\partial x}{\partial p} \frac{\partial y}{\partial q} \cdot dp \cdot dq
\]  

(8-112)

The transformation equation (8-107), it should be noted, was interpreted in a similar manner: the complex function \( f(x + jy) \) was thought to magnify the line element \( dW \) by its magnitude \( |f| \) and to rotate it by its angle \( \arg(f) \) in the process of transformation into the line element \( dZ \). For purposes of area transformation, the rotation is clearly irrelevant. Hence Equation (8-110) may be extended by adding yet another equivalent expression:

\[
\frac{\partial x}{\partial p} \frac{\partial y}{\partial q} = |f|^2
\]
so that the Helmholtz equation may finally be written
\[
\left( \nabla^2 + \frac{\mu^2}{|\mathbf{f}|^2} \right) \varphi = 0
\] (8-113)
where the differentiations are understood to be with respect to the variables \((x, y)\).

It may well be asked why this exercise in conformal mapping should have any value—has not a relatively simple differential equation been replaced by one substantially more complicated? Indeed, (8-113) is more likely to give rise to difficulties in solution than the simple Helmholtz equation (8-99). The greatest difficulty in solving partial differential equation boundary value problems, however, is not in finding solution functions of the equations themselves; it arises from the practical requirement of mathematically inconvenient boundary shapes. This is not only true for attempts at direct analytic (separation of variables) solutions but also for iterative methods. It is easy to write a programme to calculate residuals for complicated equations, but considerably harder to incorporate involved boundary shapes into the programme. Conformal transformation, on the other hand, overcomes much of this difficulty, for while Equation (8-113) is more complicated than the Helmholtz equation for which it was substituted, it is always associated with a very simple (rectangular) boundary shape. It follows that a single computer programme able to solve (e.g., by iteration) the general equation (8-113) in a rectangle will suffice to solve a large variety of problems, provided that mapping functions can be found and hence \(|\mathbf{f}|\) determined.

If solution of (8-113) is attempted by iterative methods, the usual necessity of calculating the Rayleigh coefficient by summations is encountered. As indicated in Equation (8-101), the mapping function enters this calculation also, specifically in the form
\[
u^2 h^2 = \frac{\sum_{jk} \varphi_{jk} \left( \sum_{\text{WAVE}} \varphi - 4 \varphi_{jk} \right)}{\sum_{jk} \varphi_{jk}^2 |\mathbf{f}|^2}
\] (8-114)
Evidently any sort of shape for which a mapping function can be found may be treated by this technique.

16. A High-Power Waveguide

As an illustrative example of the conformal transformation procedure, a particular shape of waveguide will now be analyzed and its principal mode (i.e., the lowest-order TE mode) field pattern, as well as its cutoff frequency, found.

From a brief examination of Fig. 8.14 and Equation (8-88), it is clear that the greatest value of electric field in a rectangular guide is found in the center when the guide is operated in its principal mode (i.e., lowest TE mode).
This factor forms one of the basic limitations for the power handling capability of a guide, for if the power transferred is to be great, the electric field may need to be sufficiently high to cause arcing in the center of the guide. In order to avoid arcing, one possibility is to construct the waveguide with a bulge in the center, as in Fig. 8.20a; the field distribution, it may be argued, will not likely be very much different from that in a rectangular guide, but the increased separation between top and bottom surfaces will perhaps serve to decrease the peak value of electric field.

![Fig. 8.20](image)

There is clearly very little hope of finding an analytic expression for the mapping function that transforms this waveguide shape into a rectangle in some other \( x-y \) plane, so that the necessary coordinate transformation is best obtained experimentally or graphically. To proceed experimentally, it is convenient to work with an electrolytic tank or, lacking one, with conductive paper. If a tank or paper sheet is taken of the same cross-sectional shape as the waveguide and electrodes placed to form equipotential surfaces as in Fig. 8.20b, a set of lines will be found, which may be imagined to denote one set of coordinate lines, say the lines \( x = \text{constant} \); their proper spacing is readily determined by simply placing them at equal potential increments. Subsequent exchange of equipotential and open-circuited boundaries, as in Fig. 8.20c, will then produce exactly the complementary set of lines. Lines of equal potential in one case will correspond to lines of equal stream function in the former case, and vice versa. The two families of lines, as explained in Chapters 2 and 3, must be orthogonal everywhere, and it only remains to decide on their relative spacing.

In Chapter 2 it was shown that the per unit length capacitance of any long structure could be evaluated from a field map drawn so as to form square figures:

\[
C = \frac{N_d e}{N_o} \quad (2-77)
\]

Similarly, the direct analogy between Laplace’s equation written for dielectrics or conductors permits the resistance of the sheet, measured as in Fig. 8.20b, to be written
Sec. 16 TRAVELING ELECTROMAGNETIC WAVES

\[ R_x = \frac{N_x}{N_y} R_0 \]  
(8-115)

where \( N_x \) is the number of squares in the \( x \)-direction and \( N_y \) the corresponding number in the \( y \)-direction of true squares laid out in the waveguide cross section. In exactly the same way, the measurement of resistance as indicated by Fig. 8.20c will give

\[ R_y = \frac{N_y}{N_x} R_0 \]  
(8-116)

In both cases, \( R_0 \) denotes the resistance of a single square and is a constant. The numbers \( N_y \) and \( N_x \) refer to the same field map—the map is not altered in going from Fig. 8.20b to 8.20c, only the potential and stream functions are interchanged—and are the same in (8-115) and (8-116). These two equations may be combined to yield, by dividing,

\[ \frac{N_y}{N_x} = \sqrt{\frac{R_y}{R_x}} \]  
(8-117)

In other words, the number of divisions in one direction must be related to the number of divisions in the other direction by a fixed number, which may be determined by making two resistance measurements. Either number may of course be chosen arbitrarily; it is conveniently picked so that both numbers will be very nearly integers. For the high-power waveguide cross section of Fig. 8.20a, a conductive-paper analogue yields the coordinate plot shown in Fig. 8.21. This set of curved lines gives the traces in the \( p-q \) plane of the lines \( x = \) constant and \( y = \) constant. It is therefore exactly what is wanted: a graphical representation of the conformal transformation function. (Figure 8.22 shows the TE mode field pattern, discussed below.)

To find the area transformation coefficients \(|f|^2\) is now easy. It is only necessary, at each grid point in the field defined by Fig. 8.21, to find the ratio of length measured in the \( p-q \) plane to the corresponding length in the \( x-y \) plane. This is most easily performed by simply measuring the distances between grid points, using any convenient scale, and subsequently normalizing the results so as to have the rectangle in the \( x-y \) plane possess an area equal to the actual waveguide cross section. The normalization is most easily performed by the same computer programme as is used to process the resulting information. Such a programme appears in Appendix II; a flow chart for it may be seen in Fig. 8.23.

Most of the programme details for subsequent calculation are similar to the foregoing iterative solutions. The normal-gradient boundary condition is dealt with, as in the last TE mode programme, by bordering the problem boundaries with an artificial exterior boundary and setting the field values at that boundary equal to the interior ones after each iterative pass. Naturally, only one-quarter of the problem need be solved, the rest is taken care of by
Define the field.

Read input data, giving \( I/\lambda \). Compute the area transformation and normalize for equal areas in both planes.

Set initial values for first iteration.

Start iteration count.

Set field values at artificial exterior boundaries.

Four iterations?

Yes

No

Calculate Rayleigh coefficient.

Perform one iteration.

Scale field values and augment iteration count.

Check convergence?

Yes

No

Calculate cutoff frequency.

Print results.

Fig. 8.23
symmetry. Convergence to the principal mode is assured, as before, by scaling the field values after each iterative pass so as to fix the normal derivative at the symmetry line. Both the Rayleigh coefficient and the maximum residual value are used as convergence indicators and, as is to be expected, substantially faster convergence obtains for the former than the latter. The actual data derived from Fig. 8.21 are appended to the computer programme in the appendix. It is only practical in this case to arrange the computer printout so as to produce a direct field plot in the $x$-$y$ plane; but its transfer into the $p$-$q$ plane is a simple matter of graphical plotting. The transverse electric field lines of the principal mode of this waveguide are shown in Fig. 8.22.

As in previous waveguide problems, it is found convenient to refer the results to a waveguide of some standard dimension, here taken as the height. The cutoff frequency is of course not solely a function of guide height—for a rectangular guide, it may be recalled, it did not even depend on the height at all—but it does vary, in general, with the guide shape in some complex way. Specifying any one dimension for a guide of fixed shape, however, specifies all other dimensions automatically. For the case at hand, a guide height of 1 m, with a corresponding width of 1.13 m, results in a calculated cutoff frequency of 117 MHz. Were the guide rectangular, the corresponding cutoff frequency would be 133 MHz; the shape alteration evidently has a significant effect.

There exist other methods for solving the transformed equation (8-113). A particularly promising and simple one is based on a two-dimensional Fourier expansion of the mapping function. That, as well as other fascinating possibilities, will be found by the interested reader in the periodical literature.

**READINGS**

There are a great many books to choose from on the subject of wave equation solutions; the real problem lies in finding those that give sufficient detail without excessive complexity. References suitable for first readings are Reitz and Milford (1), Slater and Frank (2), or the chapter in Javid and Brown already cited. Tralli (3) might well be an alternative source; Schelkunoff (14) is clear and easy to read. The little book by Coulson (4) deals with electromagnetic waves as one kind among many, but it furnishes an exceptionally pleasant introduction to some of the mathematics involved. A very thorough discussion of unbounded waves as well as waves guided by conductors will be found in Ramo, Whinnery, and Van Duzer (5). This book is recommended as a source for information on waveguides also. Weeks (6) is perhaps not quite so easy to read but does cover the subject of free-space waves well.
TEM waves propagating along conductors are of great importance in many branches of electrical engineering; in the microwave area, furthermore, many complicated shapes of line cross sections are encountered. Solutions of the various possible configurations have been attempted by any and all means available for solving Laplace's equation. Black and Higgins (7) solve a particular case, principally using conformal mapping; Wheeler (8) gives an approximate solution based on conformal transformation for a rather common type of line. Izatt (9) uses somewhat more difficult transformations to solve for the characteristic impedance of two special shapes. Numerical methods are lucidly described, and applied to particular problems, by Green (10) and Schneider (11).

The notion of intrinsic impedance is also highly useful in problems not directly connected with waveguides. Its relationship to network theory is clarified by Cooper (12) and to electric machines by Cullen and Barton (13).

For introductory reading on waveguides, Schelkunoff (14) may be recommended as both easy to read and informative. Jackson (15), in a chapter rather easier to read than might be expected from so comprehensive a book, deals with the same material at a more advanced level. Alternative readings on waveguides are to be found in the majority of the books cited above.

Numerical solution of the Helmholtz equation is not a subject as yet extensively treated in books. A quite comprehensive computer programme is described, and the method explained, by Davies and Muilwyk (16), while solutions by conformal mapping are discussed in a paper by Meinke, Lange, and Ruger (17).


**PROBLEMS**

8.1 Find the intrinsic impedance of free space.

8.2 A plane wave in free space has a peak amplitude of $E = 1$ volt per meter. Find the maximum energy density in this wave, and indicate how much of this energy is stored in the magnetic field. Assuming the frequency of this wave to be 10 GHz, find the peak amplitude of the magnetic field, and calculate the energy contained in a cubic meter of space.

8.3 Repeat the development of Equations (8–8) to (8–10) for a wave traveling in the opposite direction. Show that $E$ undergoes an additional sign change.

8.4 A right-handed elliptically polarized wave is reflected from a perfect conductor at normal incidence. Is the reflected wave right- or left-handed? Do the semiaxes of the ellipse undergo alteration?

8.5 A simple plane electromagnetic wave is reflected from a perfectly conducting surface at normal incidence. Find the instantaneous and time average values of the Poynting vector.

8.6 A plane electromagnetic wave traveling to the right along the z-axis is
reflected from a perfectly conductive surface moving toward the left at velocity $v$. Find the frequency difference between reflected and incident waves. Evaluate numerically for a frequency of 10 GHz and a velocity $v = 100$ km/hr.

8.7 A plane electromagnetic wave of 1 watt per square meter power density strikes a dielectric surface which has unit relative permeability and four times the permittivity of free space. Find the magnitude of the $\mathbf{H}$ vector in the dielectric.

8.8 Find the characteristic impedance of the structure of Problem 2.9. Does mesh size affect the result?

8.9 Show that the characteristic impedance of any structure may be found from an electrolytic tank or resistance-paper analogue by direct measurement. Use this fact to check the value given in the text for the structure of Fig. 2.16.

8.10 Prove that the Poynting vector of a standing plane wave alternates in direction. Express it in terms of the intrinsic impedance and the stored energy.

8.11 Show that a reflection from the short-circuit termination of a coaxial cable gives rise to standing waves of $V$, $i$ in the cable.

8.12 Calculate the characteristic impedance of a transmission line consisting of a deep three-sided trough 2 cm wide and a rectangular conductor 2 mm by 10 mm cross section placed symmetrically so as to clear the bottom of the trough by 10 mm (see Fig. 8.24). Assume the duct to extend infinitely high.

![Fig. 8.24](image)

8.13 Calculate the lowest 10 cutoff frequencies in a rectangular waveguide with 1 cm by 3 cm cross section.

8.14 Find the average power flow associated with a rectangular guide in the principal mode.

8.15 Prove that any two distinct modes of a rectangular waveguide are orthogonal, i.e., that the electric field of one and the magnetic field of the other always combine to give zero total energy.

8.16 Draw a plan view of a rectangular waveguide, so that $E$ of the principal mode points into the paper. On this plan view, sketch the lines of the Poynting vector associated with the principal mode.
8.17 What is the wavelength of the principal mode in a rectangular guide above
the cutoff frequency? at the cutoff frequency?

8.18 Two waves of equal amplitude, both of the principal mode, propagate along
a rectangular waveguide in opposite directions. Find their resultant, and
illustrate this situation with appropriate sketches.

8.19 Any multiconductor structure can support TE and TM as well as TEM waves.
Find the principal propagating mode (other than TEM) for the structure
of Problem 2.9.

8.20. Write an iteration programme to find the low-order TM mode of a right
triangular section waveguide. Use a mesh such as to relax residuals at roughly
100 points.

8.21 Write a similar programme for the dominant TE mode of a right triangular
waveguide.

8.22 Find the cutoff frequency and field pattern of a rectangular guide in its
principal mode, using the programme given for the high-power waveguide
analysis. How do your results compare with analytically known answers?

8.23 Find a suitable coordinate transformation graphically and then solve for the
principal mode of a circular waveguide. Compare with analytically known
answers.

8.24 Find the dominant TM mode of the high-power waveguide.
APPENDIX I

A Note on Units

The majority of electrical engineering and scientific work today employs the rationalized meter-kilogram-second-coulomb system of units, which forms a subset of the S.I. (scientific international) unit system. It is common for modern textbooks to state that "all equations in this book will hold in any rationalized system of units," or something to that effect; in fact, the only such system in general use is the RMKSC system. However, a number of authors still use the so-called Gaussian set of units, and some knowledge of at least the reasons for their existence is desirable. Furthermore, a number of American industrial organizations persists in employing a curious mixed system, in which the S.I. electrical units are intermixed with some Gaussian magnetic units and American units for mechanical quantities, producing such oddities as "kilolines per square inch" and "ampere-turns per inch." Sometimes this is referred to as the practical system of units, although the latter name is also attached to his pet system by nearly every author.

There are two basic questions to be settled in constructing a system of units: whether or not to rationalize the units, and whether or not to introduce
a fourth, dimensionally distinctive, entity in addition to mass, length, and time. It must be observed at the outset that the former problem concerns primarily a matter of convenience and unit size, for its practical effect is the introduction (or exclusion) of numerous factors $4\pi$ into the equations of electromagnetic theory. The second question, however, has a theoretical basis of somewhat greater importance and has been the source of considerable argument in the past.

1. Rationalized versus Unrationalized Units

In Chapter 1 the notions of electric charge and flux were introduced jointly, and the continuity postulate taken to be the basis of all electric phenomena. To put it another way, the development was begun from Gauss's law, and Coulomb's law was subsequently deduced as the only form possible that yielded a conservative electric field. This sequence makes it reasonable to define unit flux and unit charge as equivalent to each other, so that

$$q = \psi$$

(1-2)

from which it follows in due course that

$$F = \frac{q_1 q_2}{4\pi r^2 \varepsilon_0} 1_r$$

(1-29)

It is equally possible, however, to begin with Coulomb's law as the fundamental postulate and hence deduce the existence and nature of electric flux. If this line of argument is adopted, it is natural to start by defining unit charge so that the force between charges is simply proportional to the charges and inversely proportional to the square of distance:

$$F = K \frac{q_1 q_2}{r^2} 1_r$$

(1-1)

Of course it is possible to set $K = 1/(4\pi \varepsilon_0)$ so as to regain Equation (1-29). However, if the argument is started with the Coulomb force equation, there is no good reason to introduce anything so artificial as $4\pi$, and it is natural to let $K = 1/\varepsilon_0$ instead. In that case, the continuity postulate becomes, instead of (1-2),

$$\psi = 4\pi q$$

(1-2)

and, it is interesting to note,

$$D = \frac{q}{r^2} 1_r$$

(1-3)

In other words, the fundamental quantity now is not flux, but flux density, and unit flux density is defined as that obtained at unit distance from a unit charge. It should be noted that the unit sizes in which the various quantities
are to be measured are not all the same as before; this system is termed *unrationalized*, while that obtained by starting from Gauss's law is called *rationalized*. These names were coined by Heaviside, and merely refer to his opinion that the most rational set of units is that which results in the fewest factors $4\pi$ in the field equations.

2. Electrostatic and Electromagnetic Units

The constant $K$ introduced into Equation (I-1) purely arbitrarily may well be argued to be just as unnecessary as the $4\pi$ already eliminated. Since Equation (I-1) is in fact the defining equation for the unit of electric charge, why not define it so as to make $K = 1$ and be rid of another pointless constant? This manner of development eliminates the necessity for defining any such thing as a coulomb at all, for the dimensions of charge are thereby automatically $(\text{mass})^{1/2} (\text{length})^{3/2} (\text{time})^{-1}$. If the cgs units of force and distance (dynes and centimeters) are used, the resulting unit of charge is called a *franklin*, or merely 1 *esu* (electrostatic unit) of charge. Clearly, omission of the permittivity in Coulomb's law will also have repercussions in other equations. Chief among these is the lack of distinction between $\mathbf{D}$ and $\mathbf{E}$ in free space. Since the permittivity of free space is taken to be unity, and dimensionless, $\mathbf{D} = \mathbf{E}$.

Electromagnetic units are defined similarly but starting from a magnetic version of Coulomb's law. Unit magnetic field strength is defined as that obtained at unit distance from a unit pole, and a unit pole is defined in turn as one which produces unit force at unit distance on another pole of like strength. This is clearly not a very useful definition from an experimental point of view, and for actual measurements, these definitions have to be related to forces between dipoles or current-carrying loops. Much as in the electrostatic system of units, the permeability of free space is now a dimensionless quantity of unit value; hence $\mathbf{B} = \mathbf{H}$ in vacuo. Of course, it is only possible to define either the permeability or the permittivity of free space as unity; it is physically necessary that $\mu_0\varepsilon_0c^2 = 1$, so that defining one as unity requires the other to assume the value $1/c^2$.

The Gaussian system of units, used in engineering almost exclusively until the nineteen-fifties, is a combination of esu and emu, with electrical quantities expressed in esu and magnetic quantities in emu.

An interesting method of developing the electromagnetic equations is that followed by H. Arzeliès in his *Electricité* (Paris, Gauthier–Villars, 1963). In addition to the two constants $\mu_0$ and $\varepsilon_0$, a third is introduced. This procedure makes it possible to state all the equations of electromagnetic theory in such a way that the equations in any recognized system of units may be derived instantly by inserting the correct combination of values for
the three constants. Details of this technique are given in his Appendix II (p. 685).

3. The RMKSC System

The equations developed in the text have not generally been stated in terms of units, partly to avoid unnecessary complication, partly to emphasize that they represent relationships between physical quantities, not numbers. For convenient reference, the units natural to the RMKSC system, i.e., those that arise from the physical equations without addition of arbitrary conversion constants, are given in Table I below. The symbols given are those approved by the Institute of Electrical and Electronics Engineers for publication in IEEE journals (see IEEE spectrum, August 1965, pp. 111-115). Where the basic quantities are of inconvenient magnitude, prefixes are employed to scale unit sizes up or down; Table II lists the recommended prefixes and

<table>
<thead>
<tr>
<th>Table I</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The RMKSC Units</strong></td>
</tr>
<tr>
<td><strong>Unit</strong></td>
</tr>
<tr>
<td>ampere</td>
</tr>
<tr>
<td>coulomb</td>
</tr>
<tr>
<td>farad</td>
</tr>
<tr>
<td>henry</td>
</tr>
<tr>
<td>hertz</td>
</tr>
<tr>
<td>joule</td>
</tr>
<tr>
<td>kilogram</td>
</tr>
<tr>
<td>meter</td>
</tr>
<tr>
<td>newton</td>
</tr>
<tr>
<td>ohm</td>
</tr>
<tr>
<td>radian</td>
</tr>
<tr>
<td>second</td>
</tr>
<tr>
<td>siemens†</td>
</tr>
<tr>
<td>tesla</td>
</tr>
<tr>
<td>volt</td>
</tr>
<tr>
<td>watt</td>
</tr>
<tr>
<td>weber</td>
</tr>
</tbody>
</table>

*Numerical conversion tables for the different sets of units may be found in E. Weber, Electromagnetic theory: Static fields and their mapping. New York: Dover, 1965 (Appendices I and II); or in the handbook, Reference data for radio engineers. 4th ed. New York: International Telephone and Telegraph Corp., 1956 (pp. 36-37). Of course, many other sources of conversion tables exist.

†An accepted alternative name is mho (ohm spelled backward).
TABLE II
RECOMMENDED MAGNITUDE PREFIXES

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Exponent</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>tera-</td>
<td>+12</td>
<td>T</td>
</tr>
<tr>
<td>giga-</td>
<td>+9</td>
<td>G</td>
</tr>
<tr>
<td>mega-</td>
<td>+6</td>
<td>M</td>
</tr>
<tr>
<td>kilo-</td>
<td>+3</td>
<td>k</td>
</tr>
<tr>
<td>hecto-</td>
<td>+2</td>
<td>h</td>
</tr>
<tr>
<td>deka-</td>
<td>+1</td>
<td>da</td>
</tr>
<tr>
<td>deci-</td>
<td>+0</td>
<td>d</td>
</tr>
<tr>
<td>centi-</td>
<td>+2</td>
<td>c</td>
</tr>
<tr>
<td>milli-</td>
<td>+3</td>
<td>m</td>
</tr>
<tr>
<td>micro-</td>
<td>+6</td>
<td>μ</td>
</tr>
<tr>
<td>nano-</td>
<td>-9</td>
<td>n</td>
</tr>
<tr>
<td>pico-</td>
<td>-12</td>
<td>p</td>
</tr>
<tr>
<td>femto-</td>
<td>-15</td>
<td>f</td>
</tr>
<tr>
<td>atto-</td>
<td>-18</td>
<td>a</td>
</tr>
</tbody>
</table>

their symbols. Prefixes exist for every factor of 1000, so that it should rarely be necessary to write more than one or two nonsignificant zeros.

The numbers in Table II denote powers of 10, e.g., 1 kilometer = 10^3 meters. Any prefix symbol may be combined with any unit symbol to form new combinations, e.g., millimeter mm, teratesla TT, picofarad pF, etc. Note that giga- is pronounced as it it were spelled “jiga-.” If numbers considerably beyond the table are required, prefixes may be combined, e.g., 1 megaterahertz = 10^18 hertz.

It is important to note that the RMKSC system is not identical to the metric system, although these two do share many common units and terms. For example, the liter, kilowatt-hour, and ton are metric but not S.I. units. More fundamentally, the metric system is based on the kilogram as a unit of force (or weight), while the S.I. system is based on the kilogram mass.
APPENDIX II

Examples of Computer Programmes

This appendix contains the actual working programmes of many of the examples worked out in the text, chosen primarily to illustrate the various numerical techniques and provide examples for the beginning fields analyst. Descriptions of these programmes, and flow charts of most of them, will be found in the appropriate locations in the text itself. The programmes are reproduced in the order of their appearance in the text:

1. Relaxation problem: strip in a square duct.
2. Liebmann iteration: strip in a square duct.
4. Line integrals as sums: Neumann's formula for triangular contours.
5. Surface integrals as sums: magnetic field of a direct-current conductor.
6. Iteration of Poisson's equation: leakage field of a transformer.
9. Helmholtz equation with equipotential boundaries: dominant TM mode of a ridged waveguide.
APPENDIX II

SAMPLE RELAXATION PROGRAMME

C RELAXATION PROBLEM - STRIP IN SQUARE DUCT
C
DIMENSION V(14,14), R(14,14)
DO 10 K = 1,14
DO 10 J = 1,K
10 V(K,J) = 2.5
DO 11 K = 1,14
11 V(K,K) = 0.
DO 12 K = 1,7
12 V(14,K) = 5.
KOUNT = 0
C
C CALCULATION OF RESIDUALS
C
1 DO 2 K = 2,13
   MAX = K - 1
   R(K,1) = 2. * V(K,2) + V(K-1,1) + V(K+1,1) - 4. * V(K,1)
   DO 2 J = 2,MAG
   R(K,J) = V(K-1,J) + V(K+1,J) + V(K,J-1) + V(K,J+1) - 4. * V(K,J)
CONTINUE
2 DO 3 J = 8,13
3 R(14,J) = 2. * V(13,J) + V(14,J-1) + V(14,J+1) - 4. * V(14,J)
C
C ADJUSTMENT OF POTENTIAL VALUES
C
INDIC = 0
DO 4 K = 2,14
   MAX = K - 1
   MIN = 1
   IF(K.EQ.14) MIN = 8
   DO 4 J = MIN, MAX
   DV = R(K,J) / 4.
   IF(ABS(DV).GE.5.E-4) INDIC = 1
   4 V(K,J) = V(K,J) + DV
   KOUNT = KOUNT + 1
   IF(KOUNT.GE.200) GO TO 5
   IF(INDIC.GE.1) GO TO 1
C
C OUTPUT STATEMENTS AND EXIT
C
5 PRINT 101
   PRINT 100, ((V(J,J), K = 1,14), J = 1,14)
   PRINT 102, KOUNT
100 FORMAT(1H-, 10X, 14F5.2)
101 FORMAT(1HI)
102 FORMAT(1H-,//24X, I4, 12H ITERATIONS)
STOP
END
ITERATION PROGRAMME - STRIP IN SQUARE DUCT

ALPHA = 1.5
RESMAX = 2.E-3
DIMENSION V(14,14)
DO 10 I = 1,14
10 V(I,J) = 2.5
DO 11 I = 0
11 V(I,J) = 0.
DO 12 I = 1,7
12 V(14,J) = 5.
KOUNT = 0
K = ALPHA / 4.

ITERATION CYCLE

1 INDIC = 0
RES = V(3,1) - 4.* V(2,1)
IF(ABS(RES).GE.RESMAX) INDIC = 1
V(2,1) = V(2,1) + R * RES
DO 2 I = 3,13
MAX = I-1
RES = 2.* V(I,2) + V(I-1,1) + V(I+1,1) - 4.* V(I,1)
IF(ABS(RES).GE.RESMAX) INDIC = 1
V(I,1) = V(I,1) + R * RES
DO 2 J = 2,MAX
RES = V(I,J+1) + V(I,J-1) + V(I-1,J) + V(I+1,J) - 4.* V(I,J)
IF(ABS(RES).GE.RESMAX) INDIC = 1
2 V(I,J) = V(I,J) + R * RES
DU 3 I = 8,13
RES = 2.* V(13,1) + V(14,1-1) + V(14,1+1) - 4.* V(14,1)
IF(ABS(RES).GE.RESMAX) INDIC = 1
3 V(14,1) = V(14,1) + R * RES
KOUNT = KOUNT + 1
IF(KOUNT.GE.200) GO TO 4
IF(INDIC.EQ.1) GO TO 1
EXIT ROUTINE

4 PRINT 100, ((V(J,K), K = 1,14), J = 1,14)
PRINT 101, KOUNT, ALPHA
100 FORMAT(1H1//1H+, 10X, 1H4F5.2)
101 FORMAT(1H-///11X, 13, 40H ITERATIONS WITH OVERRELAXATION FACTOR , 1 F6.3)
STOP
END
APENDIX II

303

CHARGE DISTRIBUTION ON ROD

This programme determines an approximation to the charge distribution on a thin rod whose radius and fineness of modelling are determined by the first two arithmetic assignment statements.

Externally supplied inverse hyperbolic sine function and matrix inversion subroutine are required.

Dimension P(30,30), Q(30)

NSEGMT = 15
R = 0.2
DX = 0.5 * (30. / FLOAT(NSEMT))
7 DO 1 J = 1, NSEMT
  XJ = (FLOAT(J) - 0.5) * 2. * DX
  DO 1 K = 1, J
    XK = (FLOAT(K) - 0.5) * 2. * DX
    P(J, K) = ARSH((ABS(XK - XJ) + DX) / R) - ARSH((ABS(XK + XJ) - DX) / R) + ARSH((ABS(XK + XJ) + DX) / R) - ARSH((ABS(XK - XJ) + DX) / R)
    P(K, J) = P(J, K)
  1 PRINT 100, ((P(L, M), L = 1, NSEMT), M = 1, NSEMT)

CALL INVERS(P, NSEMT, DET, PIV)
PRINT 102, ((P(L, M), L = 1, NSEMT), M = 1, NSEMT)

DU 4 J = 1, NSEMT
Q(J) = 0.
DU 4 K = 1, NSEMT
4 Q(J) = Q(J) + P(J, K)
PRINT 101, R, (Q(L), L = 1, NSEMT)
QT = 0.
DU 5 J = 1, NSEMT
5 QT = QT + Q(J)
QT = QT / FLOAT(NSEMT)
DU 6 J = 1, NSEMT
6 Q(J) = Q(J) / QT
PRINT 101, R, (Q(L), L = 1, NSEMT)
STOP

100 FORMAT(1H1/20X, 30H POTENTIAL COEFFICIENT MATRIX P//(6X, 15F8.4))
101 FORMAT(1H1/20X, 27H CHARGE DENSITY DISTRIBUTION//20X, 3HR =, F7.4//1 (6X, 15F8.4))
102 FORMAT(1H1/20X, 30H INFLUENCE COEFFICIENT MATRIX C//(6X, 15F8.4))
EXAMPLE OF NEUMANN'S FORMULA IN
SUMMATION FORM

TWO NONINTERSECTING TRIANGULAR CONTOURS ARE MODELLED BY UP TO 74
SEGMENTS EACH. ARRAYS PTA, PTB STORE THE (X,Y,Z) COORDINATES OF
SEGMENT ENDPOINTS ALONG TRIANGLES A AND B RESPECTIVELY. KURNER
STORES THE ENDPOINT NUMBERS CORRESPONDING TO TRIANGLE VERTICES.

NEUMANN'S FORMULA FOR TWO TRIANGLES

DIMENSION KORNER(2,4), DPT(3)
DIMENSION PTA(3,75), PTB(3,75)

READ IN CORNER LOCATIONS

11 DO 1 I = 1,4
    READ 100, J, (PTA(K,J), K = 1,3)
    KORNER(1,I) = J
100 FORMAT(12,8X,3F10.5)

NA = KORNER(1,4) - 1
NB = KORNER(2,4) - 1

CALCULATION OF SEGMENT ENDPOINTS

DO 6 I = 1,3
    MIN = KORNER(I,I) + 1
    MAX = KORNER(I,I+1) - 1
    SEGMTS = FLOAT(MAX - MIN + 2)
6 DO 2 K = 1,3
    DO 3 J = 1,3
        DPT(K) = (PTA(K,MAX+1) - PTA(K,MIN-1)) / SEGMTS
    3 DO 5 K = MIN,MAX
        PTA(J,K) = PTA(J,K-1) + DPT(J)
5 CONTINUE

DO 4 K = 1,3
    DO 5 J = 1,3
        DPT(K) = (PTB(K,MAX+1) - PTB(K,MIN-1)) / SEGMTS
    5 DO 7 K = MIN,MAX
        PTB(J,K) = PTB(J,K-1) + DPT(J)
7 CONTINUE
THE ENDPOINTS ARE NOW KNOWN. SUMMATION FOLLOWS.

TOTAL = 0.
DO 10 I = 1, NA
SUM = 0.
DO 9 J = 1, NB

EVALUATE DOT-PRODUCT

DOT = 0.
DO 7 K = 1, 3
7 DOT = DOT + (PTA(K, I+1) - PTA(K, I)) * (PTB(K, J+1) - PTB(K, J))

EVALUATE MIDPOINT DISTANCE

DISTSQ = 0.
DO 8 K = 1, 3
8 DISTSQ = DISTSQ + (PTA(K, I+1) - PTA(K, I))**2

9 SUM = SUM + DOT / SQRT(DISTSQ)
10 TOTAL = TOTAL + 2. * SUM
TOTAL = 0.1 * TOTAL

PRINT 101, TOTAL
101 FORMAT(1H1/1H/-/25X, 2*XHEUMANN'S FORMULA ANALYSIS //25X, 2*HOF TWO
1 TRIANGULAR CONTOURS //20X, 19H/MUTUAL INDUCTANCE =, F12.7, 14H M
2ICROHENRIES)

GO TO 11
END

ONE RUN OF TEST DATA TO VERIFY THE ABOVE PROGRAMME --

1  0.  0.  0.  A
1  0.  0.  1.  B
25  2.  0.  0.  A
25  2.  0.  1.  B
49  1.  1.7321 0.  A
49  1.  1.7321 1.  B
73  0.  0.  0.  A
73  0.  0.  1.  B
D-C BUS-BAR

THIS PROGRAMME EVALUATES MAGNETIC FIELD BY APPROXIMATING AN INTEGRAL.

DIMENSION A(14,18), XN(20), YN(20)
ET025 = EXP(25.)

C READ IN CONDUCTOR DESCRIPTION
READ 100, N, CURRNT, (XN(I), YN(I), I = 1,N)

DO 2 I = 1,14
   X = I - 1
   DO 2 J = 1,18
      Y = 18 - J
   C EVALUATE PRODUCT
   PRODUCT = 1.
   POWER = 0.
   DO 1 K = 1,N
      FACTR = (X - XN(K))**2 + (Y - YN(K))**2
   C EXAMINE FACTOR FOR ZERO VALUE
   IF(FACTR.LT.1.E-3) FACTR = 0.44705 **2
   PRODUCT = PRODUCT * FACTR * ((X + XN(K))**2 + (Y + YN(K))**2)
   1   PRODUCT = PRODUCT / ET025
      POWER = POWER + 25.
   CONTINUE
   2 A(I,J) = (CURRNT * 1.E-7 / FLOAT(4*N)) * (ALOG(PRODUCT) + POWER)

C SUBTRACT REFERENCE LEVEL AND SCALE THE RESULTS

SCALE = 0.
CENTRE = A(1,18)
DO 3 I = 1,14
   DO 3 J = 1,18
      A(I,J) = A(I,J) - CENTRE
3   IF(ABS(A(I,J)).GT.ABSISCALE) SCALE = A(I,J)
     SCALE = SCALE / 5.
   DO 4 I = 1,14
      DO 4 J = 1,18
4    A(I,J) = A(I,J) / SCALE

C PRINT 101, A, SCALE
PUNCH 101, A, SCALE
STOP

100 FORMAT(15, F10.4(2F10.5))
101 FORMAT(1H1 / 25X, 34HVECTOR POTENTIAL IN FIRST QUADRANT / 18(1H-, 1 7X, 14F5.2/)1H-, 8X, 14HSCALE FACTOR =, E13.5)
END
ITERATIVE SOLUTION OF POISSON'S EQUATION

THIS PROGRAMME ILLUSTRATES A POISSON'S-EQUATION PROBLEM BY CALCULATING THE MAGNETIC VECTOR POTENTIAL $A$ IN THE WINDOW SPACE OF A TRANSFORMER. BOTH THE EXISTENCE OF NORMAL-GRADIENT BOUNDARIES ON THREE SIDES, AND THE DISTRIBUTED SOURCES OF THE PROBLEM, CONTRIBUTE TO A HIGH OVERRELAXATION FACTOR AS WELL AS TO RELATIVELY SLOW CONVERGENCE. AS A SAFETY FEATURE, PROGRAMME EXECUTION IS HALTED AFTER 'MAXITR' ITERATIONS REGARDLESS OF WHETHER CONVERGENCE HAS BEEN ACHIEVED OR NOT.

TRANSFORMER LEAKAGE REACTANCE PROGRAMME

DIMENSION A(13,16), CURRNT(13,16)
UNITCR = 0.2
ALFA = 1.8
RESMAX = 0.0005
MAXITR = 400

DO 1 I = 1,13
DO 1 J = 1,16
A(I,J) = 0.
1 CURRNT(I,J) = 0.
DO 2 I = 3,11
CURRNT(I,16) = 0.5 * UNITCR
DO 2 J = 3,15
2 CURRNT(I,J) = UNITCR
ALFA4 = ALFA / 4.

ITERATION START

ITERAT = 0
3 ITERAT = ITERAT + 1
BIGEST = 0.

DO POINTS ALONG TOP EDGE OF FIELD
DO 4 I = 2,12
IF(ABS(RESDL).GT.BIGEST) BIGEST = ABS(RESDL)
4 A(I,J) = A(I,J) + ALFA4 * RESDL

DO POINT (13,1)
IF(ABS(RESDL).GT.BIGEST) BIGEST = ABS(RESDL)
A(13,J) = A(13,J) + ALFA4 * RESDL

DO INTERIOR POINTS
DO 7 J = 2,15

DO POINTS IN MID-FIELD
DO 8 I = 2,12
IFI(ABS(RESDL).GT.BIGEST) BIGEST = ABS(RESDL)
8 A(I,J) = A(I,J) + ALFA4 * RESDL
RIGHT EDGE POINTS
1 + CURRNT(J,13)
IFABS(RESOL).GT.BIGEST) BIGEST = ABS(RESOL)
A(J,13) = A(J,13) + ALFA4 * RESOL
7 CONTINUE

DO POINTS ALONG BOTTOM EDGE NEXT
DU 9 I = 2,12
9 IFABS(RESOL).GT.BIGEST) BIGEST = ABS(RESOL)
A(J,16) = A(J,16) + ALFA4 * RESOL

LOWER RIGHT CORNER POINT (13,16)
RESOL = 2. * A(12,16) + 2. * A(13,15) - 4. * A(13,16)
IFABS(RESOL).GT.BIGEST) BIGEST = ABS(RESOL)
A(J,16) = A(J,16) + ALFA4 * RESOL

EXAMINE SOLUTION FOR CONVERGENCE AND EXIT
IF(ITERAT.GT.MAXITR) GO TO 5
IF(BIGEST.GT.RESMAX) GO TO 3
5 PRINT 100, A

CALCULATION OF LEAKAGE INDUCTANCE
ENERGY = 0.
DU 6 I = 1,13
DU 6 J = 1,16
6 ENERGY = ENERGY + A(I,J) * CURRNT(I,J)
PRINT 102, HENRY
PRINT 101, BIGEST, ITERAT
STOP

100 FORMAT(1H1, 1H /17X, 47HLEAKAGE FLUX PATTERN OF TWO-WINDING TRANSFORMER/ (1H- , 8X, 13F5.2))
101 FORMAT(1H0/ 8X, 18HLARGEST RESIDUAL = , F8.5, 2X, SHAFTER, I4, 11H 1 ITERATIONS)
102 FORMAT(1H0/ 8X, 36HLEAKAGE INDUCTANCE PER METER DEPTH = , F8.5, 13H MICROHENRIES /38X, 24H* (NO. OF TURNS Squared))
END
APPENDIX II

POT-CORE INDUCTOR

THIS PROGRAMME ILLUSTRATES THE METHOD OF LIEBMANN ITERATION CARRIED OUT IN CYLINDRICAL COORDINATES. IT MUST BE OBSERVED THAT THE MESH USED IS COARSE, AND ACCURACY CONSEQUENTLY NOT VERY HIGH.

DIMENSION A(14,16), RADIUS(13), RINV SQ(13)
ALPHA = 1.4
KOUNT = 0
RESMAX = 0.00025
MAXITR = 200
ALFA4 = ALPHA / 4.

C
CALCULATE RADIAL MULTIPLIERS

C
do 1 i = 2,13
rinvsw(i) = 1. / float((i-1)**2)
1 r radius(i) = 1. / float(2 * (i-1))
C
SET BOUNDARY VALUES OF A

C
do 2 i = 1,14
do 2 j = 1,16
a(i,j) = 0.
2
do 3 i = 5,10
do 3 j = 5,16
a(i,j) = 4.
3
C
SET NORMAL-GRADIENT BOUNDARY

C
4 do 5 i = 1,14
5 a(i,16) = a(i,14)
C
DU ONE ITERATIVE PASS

C
bigest = 0.
dj 6 j = 2,15
imin = 2
imax = 13
if(j.lt.5) go to 6
imax = 4
6 do 7 i = imin,imax
residl = a(i,j-1) + a(i,j+1) + (1. - radius(i)) * a(i-1,j) + (1. - radius(i)) * a(i+1,j) - (4. + rinv sq(i)) * a(i,j)
if(abs(residl).gt.bigest) bigest = abs(residl)
7 a(i,j) = a(i,j) + alfa4 * residl
if(imax.eq.13) go to 8
imin = 11
imax = 13
go to 6
8 continue
kount = kount + 1
EXAMINE SOLUTION FOR CONVERGENCE

IF(KOUNT.GE.MAXITR) GO TO 9
IF(BIGEST.GT.RESMAX) GO TO 4
9 PRINT 100, (A(I,J), I = 1,14), J = 1,5), (A(I,J), I = 1,5),
1 (A(I,J), I = 10,14), J = 6,15); KOUNT, BIGEST
C
C LINE INTEGRATION
C
SUM = 0.
C
UP THE LINE R = 11.5 AND DOWN R = 2.5
DO 10 J = 3,15
1 * A(3,J)) / 2.5
IF(J.EQ.15) TERM = TERM / 2.
10 SUM = SUM + TERM
C
ACROSS ALONG THE LINE Z = 12.5
DO 11 I = 4,12
TERM = A(I,2) - A(I,4)
11 SUM = SUM + TERM
C
C CALCULATE COIL AMPERE- Turns AND FLUX
C
C
C CALCULATE INDUCTANCE
C
HENRY = FLUX / CURRENT
PRINT 101, HENRY
STOP
C
100 FORMAT(1HI//25X, 33HVECTO R EQUI POTENTIALS IN POT CORE //
  1 5(1H-, 6X, 14F5.2/), 10(1H-, 6X, 5F5.2, 20X, 5F5.2/), 1H-, 7X,
  2 22HLARGEST RESIDUAL AFTER, 14, 13H ITERATIONS =, F8.5)
101 FORMAT(1HO, 7X, 12HINDUCTANCE =, 1PE10.3, 33H HENRIES X (NO. OF TU
  1RNS SQUARED))
END
SKIN EFFECT IN A BUS-BAR

This programme calculates alternating current distribution in a rectangular conductor 100 mm x 5 mm, by solving the matrix approximation to the inhomogeneous integral equation of current density distribution. Complex arithmetic is avoided by employing double-sized matrices and working with pure real numbers only.

```
DIMENSION Z(30,30), X(10,10), CURNT(20), VOLTS(20), SIZE(10),
1 ANGL(10)
G = 1. E+8/2.83
FREQ = 377.
H = 5. E-3
PI = 3. 1415927
PERM = 4. * PI * 1. E-7
EO = 1.

C CONSTRUCT X-MATRIX
C
CONST = (FREQ * PERM * G * H**2) / (2. * PI)
DO 1 I = 1,10
XI = (FLOAT(I) - 0.5) * H
X(I,I) = CONST * ALOG(0.44705 * H * XI * 2.)
DO 1 J = 1,10
IF(J.EQ.I) GO TO 1
X(J,I) = CONST * ALOGABS(XI-XJ) * ABS(XI+XJ))
X(J,J) = X(I,J)
1 CONTINUE

C CONSTRUCT Z-MATRIX
C
DO 2 I = 1,20
DO 2 J = 1,20
Z(I,J) = 0.
2 Z(I,I) = 1.
DO 3 I = 1,10
DO 3 J = 1,10
Z(I+10,J) = -X(I,J)
3 Z(I,J+10) = X(I,J)

C SET UP VOLTAGES AND COMPUTE CURRENTS
C
V = G * EO
DO 4 I = 1,10
VOLTS(I) = V
4 VOLTS(I+10) = 0.
CALL INVER(Z, 20, DETR, PIVT)
DO 5 I = 1,20
CURNT(I) = 0.
5 DO 5 J = 1,20
CURNT(I) = CURNT(I) + Z(J,I) * VOLTS(J)
```
C

PRINT MAGNITUDES AND ARGUMENTS OF CURRENT

C

RADIUS = 180./PI
DO I = 1,10
   SIZE(I) = SQRT(CURRENT(I)**2 + CURRENT(I+10)**2)
   ANGL(I) = ATAN2(CURRENT(I+10), CURRENT(I)) * RADIUS
   FREU = FREU / (12.*PI)
   PRINT 100, FREU, (I, SIZE(I), ANGL(I), I = 1,10)
END

100 FORMAT(1H1//20X, 23HCURRENT DISTRIBUTION AT , F5.1, 3H Hz//20X, 2H
   INU, 3X, 9HMAGNITUDE, 8X, 5HANGLE//(20X, 12, 1P2E15.4))
END
ITERATIVE SOLUTION OF HELMHOLTZ EQUATION

THE FOLLOWING PROGRAMME SOLVES THE SCALAR HELMHOLTZ EQUATION TO FIND RELATIVE VALUES OF AXIAL ELECTRIC FIELD IN A RIDGE WAVEGUIDE. ONLY ONE-HALF THE ACTUAL SYMMETRICAL GUIDE IS MODELLED, THE OTHER HALF BEING SUPPLIED BY REFLECTION AT THE PLANE OF SYMMETRY. IT SHOULD BE NOTED THAT THIS PROBLEM HAS ZERO BOUNDARY VALUES AND IS SATISFIED BY THE TRIVIAL SOLUTION E = 0 EVERYWHERE. A NONZERO STARTING GUESS FOR THE ITERATIONS IS THEREFORE MANDATORY.

AS A SAFETY FEATURE, THE PROGRAMME EXECUTION IS TERMINATED AFTER 'MAXITR' ITERATIONS WHETHER CONVERGENCE HAS BEEN ACHIEVED OR NOT.

DUMINANT TM MODE OF RIDGE WAVEGUIDE

DIMENSION EZ(I3,16)
MAXITR = 400
ALPHA = 1.3
USQHSQ = 0.25
DELTAS = 0.00005

DO 1 I = 1,13
  DO 1 J = 1,16
  EZ(I,J) = 0.
DO 1 I = 1,12
  EZ(I,J) = 4.

ALFA4 = ALPHA / 4.
KUNVRG = 0
ITERAT = 0

FIVE ITERATIONS OF POINT VALUES OF EZ

DO 4 K = 1,5
  DU 3 J = 2,7
  RESDL = 2. * EZ(2,J) + EZ(1,J-1) + EZ(1,J+1) + (USQHSQ - 4.) * EZ(1,J)
  EZ(1,J) = EZ(1,J) + ALFA4 * RESDL
  DU 3 I = 2,12
  RESDL = EZ(I-1,J) + EZ(I+1,J) + EZ(I,J+1) + EZ(I,J-1) + (USQHSQ - 4.) * EZ(I,J)
  EZ(I,J) = EZ(I,J) + ALFA4 * RESDL

4 EZ(I,J) = EZ(I,J) + ALFA4 * RESDL
ITERAT = ITERAT + K
IF(KUNVRG.EQ.1) GO TO 7
CALCULATION OF RAYLEIGH COEFFICIENT

RCN = 0.
RCO = 0.
DO 5 J = 2, 7
   RCN = RCN + (2. * EZ(2,J) + EZ(1,J-1) + EZ(1,J+1) - 4. * EZ(1,J))
   * EZ(1,J)
   RCO = RCO + EZ(1,J) ** 2
   DO 5 I = 2, 12
      RCN = RCN + EZ(I,J) * (EZ(I+1,J) + EZ(I-1,J) + EZ(I,J-1)
      + EZ(I,J+1) - 4. * EZ(I,J))
   5 RCN = RCN + EZ(I,J) ** 2
   DO 6 J = 8, 15
      DO 6 I = 7, 12
         RCN = RCN + EZ(I,J) * (EZ(I+1,J) + EZ(I-1,J) + EZ(I,J-1)
         + EZ(I,J+1) - 4. * EZ(I,J))
      6 RCN = RCN + EZ(I,J) ** 2
   UHNEW = - RCN/RCO

EXAMINE FOR CONVERGENCE AND CALCULATE CUTOFF FREQUENCY

IF(ABS((UHNEW - USQHSG)/UHNEW).LT.DELTAU) KONVRG = 1
IF(ITERAT.GE.MAXITR) GO TO 7
   USQHSG = UHNEW
   ALFA4 = ALPHA / (4. - USQHSG)
   GO TO 2

7 FREQCY = SQRT(USQHSG) * 15. * 300. / (2. * 3.1415927)
   PRINT 100, (EZ(I,J), I = 1, 13), J = 1, 8
   PRINT 101, (EZ(I,J), I = 6, 13), J = 9, 16
   PRINT 102, FREQCY, ITERAT
   STOP

100 FORMAT(1HI/1HO/23X, 35HDOMINANT TM MODE OF RIDGE WAVEGUIDE //1H-1, 6X, 13F5.2)
101 FORMAT(1H-, 31X, 8F5.2)
102 FORMAT(1HO/ 8X, 18HCUTOFF FREQUENCY = , F6.1, 38H MHZ DIVIDED BY G I D E HEIGHT IN METERS //8X, 17HCORVERGENCE AFTER, 14, 1X, 10HITERA 2TIONS)
END
APPENDIX II

315

SOLUTION OF HELMHOLTZ EQUATION WITH NORMAL-GRADIENT BOUNDARIES

C
C DOMINANT TE MODE OF RIDGE WAVEGUIDE
C
C DIMENSION HZ(14,18)
ALPHA = 1.75
MAXITR = 250
RESMAX = 0.0001
DELTAU = 0.000005
C
C ENTER INITIAL GUESS FOR FIELD VALUES
UHSQ = 0.1
DO 2 J = 1,16
HZ(14,J) = 0.
DO 2 I = 1,13
2 HZ(I,J) = 0.5

C
C FIVE ITERATIVE CYCLES FOLLOW
KONVRG = 0
ITERAT = 0
3 ALFA = ALPHA / (4. - UHSQ)
DO 15 K = 1,5
C
C ITERATIVE PASS THROUGH FIELD
BIGEST = 0.
IMAX = 13
4 J = J + 1
DO 5 I = 2,IMAX
RESDL = HZ(I-1,J) + HZ(I+1,J) + HZ(I,J-1) + HZ(I,J+1) + (UHSQ - 4.)*HZ(I,J)
1 IF(ABS(RESDL).GT.BIGEST) BIGEST = ABS(RESDL)
5 HZ(I,J) = HZ(I,J) + ALFA * RESDL
IF(J-9) 4, 6, 7
6 IMAX = 9
HZ(10,10) = HZ(8,10)
7 IF(J.LT.17) GO TO 4
C
C SET HZ VALUES AT ARTIFICIAL FIELD POINTS
DO 8 I = 2,13
HZ(I,1) = HZ(I,3)
8 HZ(I,18) = HZ(I,16)
DO 9 I = 10,13
9 HZ(I,10) = HZ(I,8)
DO 10 J = 1,18
10 HZ(1,J) = HZ(3,J)
DO 11 J = 11,18
11 HZ(10,J) = HZ(8,J)

C
C FIND AVERAGE HZ NEAR SYMMETRY EDGE
TOTAL = 0.5 * (HZ(13,2) + HZ(13,9))
DO 12 J = 3,8
12 TOTAL = TOTAL + HZ(13,J)
AVG = 0.5 * TOTAL / TOTAL

C
C SCALE FIELD VALUES
DO 13 I = 1,14
DO 13 J = 1,18
13 HZ(I,J) = AVG * HZ(I,J)
15 ITERAT = ITERAT + 1
C
C END OF SINGLE ITERATION
CALCULATION OF RAYLEIGH COEFFICIENT

\[
\begin{align*}
RCN &= 0, \\
RCD &= 0, \\
DU 16 J &= 2,17 \\
DU 16 I &= 2,13 \\
FCTR &= AREA(I,J) \\
RCN &= RCN + FCTR \times (HZ(I,J) + HZ(I-1,J) + HZ(I+1,J) + HZ(I,J-1) + HZ(I,J+1) - 4 \times HZ(I,J)) \\
RCD &= RCD + FCTR \times (HZ(I,J) ^ 2) \\
RCN &= RCN \times 0.5 \times (HZ(9,10) \times (HZ(8,10) - HZ(10,8))) \\
\text{RAYLGH} &= \frac{-RCN}{RCD} \\
\text{CHECK FOR CONVERGENCE} \\
\text{IF(ABS((RAYLGH - UHSQ)/RAYLGH)} &< \text{D} \text{E} \text{T} \text{A} \text{U} \\
\text{KONVRG} &= 1 \\
\text{IF(ITEKAT.GT.MAXITR) GO TO 19} \\
UHSW &= RAYLGH \\
\text{IF(KONVRG.EQ.O) GO TO 3} \\
\text{IF(BIGEST.GT.RESMAX) GO TO 3} \\
FREOCY &= \text{SQRT(UHSW)} \times 15. \times 300. / (2. \times 3.1415927) \\
\text{PRINT 100, ((HZ(I,J), I = 2,14), J = 2,9), ((HZ(I,J), I = 2,9), J = 10,17), FREOCY, ITERAT} \\
\text{STOP} \\
\end{align*}
\]

FUNCTION AREA(I,J)

\[
\begin{align*}
\text{AREA} &= 1. \\
\text{IF(I-2) 30,1,2} \\
1 \text{ AREA} &= \text{AREA} / 2. \\
2 \text{ IF(J-9) 12,8,3} \\
3 \text{ IF(I-9) 5, 4,30} \\
4 \text{ AREA} &= \text{AREA} / 2. \\
5 \text{ IF(J-17) 7, 6,30} \\
6 \text{ AREA} &= \text{AREA} / 2. \\
7 \text{ RETURN} \\
8 \text{ IF(I-9) 7, 9,10} \\
9 \text{ AREA} &= 0.75 \\
10 \text{ RETURN} \\
11 \text{ IF(I-14) 7, 6,30} \\
12 \text{ IF(J-2) 30,10,11} \\
30 \text{ AREA} &= 0. \\
\text{RETURN} \\
\text{END} \\
\end{align*}
\]
APPENDIX II

GENERAL HELMHOLTZ EQUATION WITH NORMAL-GRADIENT BOUNDARIES

THIS PROGRAMME FINDS THE DOMINANT TE MODE OF A WAVEGUIDE OF ARBITRARY SHAPE, GIVEN THE AREA TRANSFORMATION FUNCTION FOR SUITABLE CURVILINEAR COORDINATES. THE CUTOFF FREQUENCY IS ALSO CALCULATED.

DIMENSION H(16,12), F(16,12)
ALPHA = 1.6
DELTAU = 0.0001
REMAX = 0.001
MAXITR = 300
KONVRG = 0
ITERAT = 0

C C COMPUTE AREA TRANSFORMATION FROM DATA
READ 100, ((F(I,J), I = 1,15), J = 2,11)
TOTAL = 0.
DO 1 I = 1,15
DO 1 J = 2,11
TERM = F(I,J)**2
IF(I.EQ. 1) TERM = 0.5 * TERM
IF(J.EQ.15) TERM = 0.5 * TERM
IF(I.EQ. 2) TERM = 0.5 * TERM
IF(J.EQ.11) TERM = 0.5 * TERM
1 TOTAL = TOTAL + TERM
SCALE = TOTAL / 126.
DO 2 I = 1,15
DO 2 J = 2,11
2 F(I,J) = (F(I,J)**2)/SCALE

C C SET STARTING VALUES
DO 3 J = 1,11
H(1,J) = 0.
DO 3 I = 2,15
3 H(I,J) = 0.5
USQMSQ = 0.5

C C SET ARTIFICIAL BOUNDARY VALUES
5 DO 6 I = 2,15
H(I,1) = H(I,3)
6 H(I,12) = H(I,10)
DO 7 J = 2,11
7 H(16,J) = H(14,J)
IF(ITER.EQ.4) GO TO 11
BIGEST = 0.

C C SINGLE ITERATION STARTS
DO 8 I = 2,15
DO 8 J = 2,11
FACTOR = 4. - F(I,J) * USQMSQ
RESDL = H(I-1,J) + H(I+1,J) + H(I,J-1) + H(I,J+1) - FACTOR* H(I,J)
IF(ABS(RESDL).GT.BIGEST) BIGEST = ABS(RESDL)
8 H(I,J) = H(I,J) + RESDL * ALPHA / FACTOR
ITER = ITER + 1
ITERAT = IERAT + 1
APPENDIX II

SCALE FIELD VALUES
AVRG = 0.5 * (H(2,1) + H(2,11))
DO 9 J = 3,10
9 AVRG = AVRG * H(2,J)
SCALE = 0.5 * AVRG
DO 10 I = 2,15
DO 10 J = 2,11
10 H(I,J) = SCALE * H(I,J)
GO TO 5

FIND RAYLEIGH COEFFICIENT
11 RAYD = 0.
RAYN = 0.
DO 15 I = 2,15
DO 15 J = 2,11
TERM = F(I,J) * H(I,J)**2
IF(I.NE.15) GO TO 12
TERMN = 0.5 * TERMN
TERM = 0.5 * TERM
12 IF(J.EQ.2) GO TO 13
IF(J.NE.11) GO TO 14
13 TERMN = 0.5 * TERMN
TERM = 0.5 * TERM
14 RAYD = RAYD + TERM
15 RAYN = RAYN + TERM
RAYLGH = (RAYD + RAYN) / 2

CHECK FOR CONVERGENCE AND EXIT
IF(ABS(RAYLGH - USQHSQ)/USQHSQ.LT.DELTAU) KONVRY = 1
USQHSQ = RAYLGH
IF(ITERAT.GE.MAXITR) GO TO 16
IF(BIGEST.GT.RESMAX) GO TO 4
16 FREQ = SORT(USQHSQ) * 3.E+2 * 9. / 6.28318
PRINT 101, ((H(I,J), I = 1,15), J = 2,11), FREQ, BIGEST, ITERAT
STOP

THE FOLLOWING DATA CARDS DESCRIBE THE HIGH-POWER WAVEGUIDE COORDINATES

0.95 1.02 1.08 1.00 0.88 0.75 0.65 0.57 0.51 0.46 0.41 0.35 0.30 0.24 0.19
0.89 0.92 0.95 0.90 0.81 0.72 0.64 0.57 0.51 0.46 0.41 0.35 0.32 0.30 0.26
0.83 0.85 0.85 0.82 0.71 0.69 0.62 0.57 0.52 0.47 0.42 0.37 0.35 0.36 0.31
0.79 0.80 0.80 0.75 0.69 0.66 0.61 0.57 0.52 0.47 0.43 0.40 0.39 0.38 0.36
0.76 0.77 0.75 0.72 0.67 0.62 0.60 0.57 0.52 0.48 0.45 0.43 0.41 0.40 0.39
0.73 0.74 0.73 0.69 0.65 0.62 0.60 0.57 0.52 0.48 0.46 0.44 0.43 0.43 0.40
0.71 0.72 0.72 0.68 0.64 0.61 0.58 0.57 0.52 0.48 0.47 0.44 0.44 0.43 0.42
0.70 0.70 0.71 0.68 0.63 0.60 0.57 0.57 0.52 0.48 0.47 0.45 0.45 0.45 0.42
0.70 0.69 0.70 0.68 0.63 0.60 0.57 0.57 0.52 0.48 0.47 0.45 0.45 0.45 0.44
0.70 0.70 0.70 0.68 0.63 0.60 0.57 0.57 0.52 0.49 0.47 0.45 0.45 0.45 0.45
Solution of meaningful field problems very often involves the use of some of the special functions of mathematical physics or manipulations with large matrices. This appendix gives subroutines for the calculation of the most important functions occurring in the problems at chapter ends and brief tabulations of those functional values not found in the most common sets of tables. It may be noted that many more significant figures are given than warranted by the interval of tabulation. This is quite intentional; these values are intended primarily for checking the proper working of subroutines, rather than for use as tables. The subroutines included here are

1. Complete elliptic integrals of the first and second kinds.
2. The Kelvin functions $\text{ber}(x)$, $\text{bei}(x)$, $\text{ber}'(x)$, $\text{bei}'(x)$.
3. The inverse hyperbolic functions $\sinh^{-1} x$, $\cosh^{-1} x$.
4. Inverse circular and hyperbolic sines and cosines of complex arguments.
5. A matrix inversion subroutine for $30 \times 30$ real matrices.
COMPLETE ELLIPTIC INTEGRAL SUBROUTINES

These subroutines produce the complete elliptic integrals of the first and second kinds, using a polynomial approximation. The error lies below two units in the eighth significant figure.

C C THIS SUBROUTINE RETURNS VALUES OF THE COMPLETE ELLIPTIC INTEGRAL
C OF THE FIRST KIND TO THE MODULUS Z
C
FUNCTION ELINK(Z)
  P = 1. - Z**2
  IF(P.EQ.0.) GO TO 1
  ELINK = 1.38629436 + P * (0.096663443 + P * (0.035900924 +
  1. P * (0.037425637 + 0.014511962 * P)) - ALOG(P) * (0.5 + P *
  2. (0.12498594 + P * (0.068802486 + P * (0.033283553 +
  3. 0.0044178701 * P))))
  RETURN
1. ELINK = EXP(88.)
  RETURN
END

C C THIS SUBROUTINE RETURNS VALUES OF THE COMPLETE ELLIPTIC INTEGRAL
C OF THE SECOND KIND TO THE MODULUS Z
C
FUNCTION ELINE(Z)
  P = 1. - Z**2
  IF(P.EQ.0.) GO TO 1
  ELINE = 0.44325141 + P * (0.062606012 + P * (0.047573836 +
  1. P * (0.017365065 * P)) - ALOG(P) * (0.24998368 +
  2. P * (0.0920018 + P * (0.040696975 + P * 0.0052644964 * P)))
  RETURN
1. ELINE = 1.
  RETURN
END
## COMPLETE ELLIPTIC INTEGRALS

<table>
<thead>
<tr>
<th>M</th>
<th>K(M)</th>
<th>E(M)</th>
<th>H(M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>1.570796</td>
<td>1.570796</td>
<td>0.000000</td>
</tr>
<tr>
<td>0.02</td>
<td>1.570953</td>
<td>1.570639</td>
<td>0.000004</td>
</tr>
<tr>
<td>0.04</td>
<td>1.571425</td>
<td>1.570168</td>
<td>0.000012</td>
</tr>
<tr>
<td>0.06</td>
<td>1.572213</td>
<td>1.569382</td>
<td>0.000042</td>
</tr>
<tr>
<td>0.08</td>
<td>1.573319</td>
<td>1.568280</td>
<td>0.000101</td>
</tr>
<tr>
<td>0.10</td>
<td>1.574746</td>
<td>1.566862</td>
<td>0.000198</td>
</tr>
<tr>
<td>0.12</td>
<td>1.576497</td>
<td>1.565126</td>
<td>0.000343</td>
</tr>
<tr>
<td>0.14</td>
<td>1.578579</td>
<td>1.563071</td>
<td>0.000547</td>
</tr>
<tr>
<td>0.16</td>
<td>1.580997</td>
<td>1.560694</td>
<td>0.000820</td>
</tr>
<tr>
<td>0.18</td>
<td>1.583757</td>
<td>1.557995</td>
<td>0.001174</td>
</tr>
<tr>
<td>0.20</td>
<td>1.586868</td>
<td>1.554969</td>
<td>0.001619</td>
</tr>
<tr>
<td>0.22</td>
<td>1.590338</td>
<td>1.551614</td>
<td>0.002170</td>
</tr>
<tr>
<td>0.24</td>
<td>1.594179</td>
<td>1.547927</td>
<td>0.002837</td>
</tr>
<tr>
<td>0.26</td>
<td>1.598402</td>
<td>1.543904</td>
<td>0.003636</td>
</tr>
<tr>
<td>0.28</td>
<td>1.603020</td>
<td>1.539541</td>
<td>0.004580</td>
</tr>
<tr>
<td>0.30</td>
<td>1.608049</td>
<td>1.534833</td>
<td>0.005686</td>
</tr>
<tr>
<td>0.32</td>
<td>1.613504</td>
<td>1.529777</td>
<td>0.006971</td>
</tr>
<tr>
<td>0.34</td>
<td>1.619404</td>
<td>1.524366</td>
<td>0.008453</td>
</tr>
<tr>
<td>0.36</td>
<td>1.625771</td>
<td>1.518594</td>
<td>0.010152</td>
</tr>
<tr>
<td>0.38</td>
<td>1.632628</td>
<td>1.512455</td>
<td>0.012090</td>
</tr>
<tr>
<td>0.40</td>
<td>1.640000</td>
<td>1.505942</td>
<td>0.014291</td>
</tr>
<tr>
<td>0.42</td>
<td>1.647917</td>
<td>1.499046</td>
<td>0.016782</td>
</tr>
<tr>
<td>0.44</td>
<td>1.656411</td>
<td>1.491760</td>
<td>0.019591</td>
</tr>
<tr>
<td>0.46</td>
<td>1.665520</td>
<td>1.484075</td>
<td>0.022752</td>
</tr>
<tr>
<td>0.48</td>
<td>1.675284</td>
<td>1.475979</td>
<td>0.026300</td>
</tr>
<tr>
<td>0.50</td>
<td>1.685750</td>
<td>1.467462</td>
<td>0.030277</td>
</tr>
<tr>
<td>0.52</td>
<td>1.696972</td>
<td>1.458512</td>
<td>0.034729</td>
</tr>
<tr>
<td>0.54</td>
<td>1.709009</td>
<td>1.449114</td>
<td>0.039708</td>
</tr>
<tr>
<td>0.56</td>
<td>1.721930</td>
<td>1.439255</td>
<td>0.045273</td>
</tr>
<tr>
<td>0.58</td>
<td>1.735815</td>
<td>1.428918</td>
<td>0.051494</td>
</tr>
<tr>
<td>0.60</td>
<td>1.750754</td>
<td>1.418083</td>
<td>0.058449</td>
</tr>
<tr>
<td>0.62</td>
<td>1.766853</td>
<td>1.406732</td>
<td>0.066231</td>
</tr>
<tr>
<td>0.64</td>
<td>1.784236</td>
<td>1.394842</td>
<td>0.074947</td>
</tr>
<tr>
<td>0.66</td>
<td>1.803050</td>
<td>1.382386</td>
<td>0.084725</td>
</tr>
<tr>
<td>0.68</td>
<td>1.823466</td>
<td>1.369337</td>
<td>0.095717</td>
</tr>
<tr>
<td>0.70</td>
<td>1.845694</td>
<td>1.355661</td>
<td>0.108108</td>
</tr>
<tr>
<td>0.72</td>
<td>1.869985</td>
<td>1.341322</td>
<td>0.122121</td>
</tr>
<tr>
<td>0.74</td>
<td>1.896650</td>
<td>1.326276</td>
<td>0.138032</td>
</tr>
<tr>
<td>0.76</td>
<td>1.926075</td>
<td>1.310473</td>
<td>0.156189</td>
</tr>
<tr>
<td>0.78</td>
<td>1.958748</td>
<td>1.293854</td>
<td>0.177033</td>
</tr>
<tr>
<td>0.80</td>
<td>1.995303</td>
<td>1.276350</td>
<td>0.201140</td>
</tr>
<tr>
<td>0.82</td>
<td>2.036575</td>
<td>1.257875</td>
<td>0.229275</td>
</tr>
<tr>
<td>0.84</td>
<td>2.083701</td>
<td>1.238326</td>
<td>0.262489</td>
</tr>
<tr>
<td>0.86</td>
<td>2.138283</td>
<td>1.217570</td>
<td>0.302270</td>
</tr>
<tr>
<td>0.88</td>
<td>2.202677</td>
<td>1.195437</td>
<td>0.350826</td>
</tr>
<tr>
<td>0.90</td>
<td>2.280549</td>
<td>1.171697</td>
<td>0.416210</td>
</tr>
<tr>
<td>0.92</td>
<td>2.378071</td>
<td>1.146025</td>
<td>0.490536</td>
</tr>
<tr>
<td>0.94</td>
<td>2.506864</td>
<td>1.117924</td>
<td>0.598740</td>
</tr>
<tr>
<td>0.96</td>
<td>2.693143</td>
<td>1.086546</td>
<td>0.761659</td>
</tr>
<tr>
<td>0.98</td>
<td>3.020980</td>
<td>1.050092</td>
<td>1.061660</td>
</tr>
<tr>
<td>1.00</td>
<td>INFINITE</td>
<td>1.000000</td>
<td>INFINITE</td>
</tr>
</tbody>
</table>
KELVIN FUNCTION SUBROUTINES

The following subroutines return values of the four Kelvin functions \( \text{BER}(z), \text{BEI}(z), \text{BER}^*(z), \text{BEI}^*(z) \). The error does not exceed two units in the eighth significant figure, for values of \( z \) smaller than 8. For larger values of \( z \), these routines should not be used.

**FUNCTION BER(Z)**

\[
C = (z/8.1)^4
\]

\[
\text{BER} = 1. + X * (-64. * X*(113.777778 + X*(-32.3634565 + X*(12.6449114 + X*(-0.08344609 + X*(0.00122552 + 0.00000901 * X))))))
\]

RETURN END

**FUNCTION BEI(Z)**

\[
C = (z/8.1)^2
\]

\[
X = Y**2
\]

\[
\text{BEI} = 16. + X*(-113.777778 + X*(72.817777 + X*(-10.5676578 + X*(0.52185615 + X*(-0.01103667 + 0.00011346 * X)))))
\]

RETURN END

**FUNCTION BEIP(Z)**

\[
C = (z/8.1)^4
\]

\[
\text{BEIP} = 0.5 + Y*(-10.6666667 + Y*(11.3777777 + Y*(-2.31167514 + Y*(0.14677204 + Y*(-0.00379386 + 0.00004609 * Y)))))
\]

RETURN END

**FUNCTION BERP(Z)**

\[
C = (z/8.1)^4
\]

\[
\text{BERP} = -4. + Y*(14.222222 + Y*(-6.0681481 + Y*(0.66047849 + Y*(-0.06092953 + Y*(0.0000394 * Y)))))
\]

RETURN END
### Kelvin Function Table

<table>
<thead>
<tr>
<th>( x )</th>
<th>( \text{BER}(x) )</th>
<th>( \text{BEI}(x) )</th>
<th>( \text{BER}^*(x) )</th>
<th>( \text{BEI}^*(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>1.000000</td>
<td>0.000000</td>
<td>-0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>0.2</td>
<td>0.999975</td>
<td>0.010000</td>
<td>-0.000000</td>
<td>0.099999</td>
</tr>
<tr>
<td>0.4</td>
<td>0.999960</td>
<td>0.039998</td>
<td>-0.000010</td>
<td>0.199973</td>
</tr>
<tr>
<td>0.6</td>
<td>0.999797</td>
<td>0.089980</td>
<td>-0.000076</td>
<td>0.299798</td>
</tr>
<tr>
<td>0.8</td>
<td>0.999360</td>
<td>0.159886</td>
<td>-0.000320</td>
<td>0.399147</td>
</tr>
<tr>
<td>1.0</td>
<td>0.984382</td>
<td>0.249566</td>
<td>-0.000976</td>
<td>0.497397</td>
</tr>
<tr>
<td>1.2</td>
<td>0.967629</td>
<td>0.358704</td>
<td>-0.002426</td>
<td>0.593523</td>
</tr>
<tr>
<td>1.4</td>
<td>0.940075</td>
<td>0.486734</td>
<td>-0.005235</td>
<td>0.686008</td>
</tr>
<tr>
<td>1.6</td>
<td>0.897891</td>
<td>0.632726</td>
<td>-0.009182</td>
<td>0.77240</td>
</tr>
<tr>
<td>1.8</td>
<td>0.836722</td>
<td>0.795262</td>
<td>-0.018285</td>
<td>0.850927</td>
</tr>
<tr>
<td>2.0</td>
<td>0.751734</td>
<td>0.972292</td>
<td>-0.030817</td>
<td>0.917014</td>
</tr>
<tr>
<td>2.2</td>
<td>0.637690</td>
<td>1.160970</td>
<td>-0.049308</td>
<td>0.966609</td>
</tr>
<tr>
<td>2.4</td>
<td>0.489048</td>
<td>1.357485</td>
<td>-0.075528</td>
<td>0.994429</td>
</tr>
<tr>
<td>2.6</td>
<td>0.300092</td>
<td>1.556878</td>
<td>-0.111448</td>
<td>0.994263</td>
</tr>
<tr>
<td>2.8</td>
<td>0.065112</td>
<td>1.752851</td>
<td>-0.159160</td>
<td>0.958965</td>
</tr>
<tr>
<td>3.0</td>
<td>-0.221380</td>
<td>1.937587</td>
<td>-0.220760</td>
<td>0.880482</td>
</tr>
<tr>
<td>3.2</td>
<td>-0.564376</td>
<td>2.101573</td>
<td>-0.298179</td>
<td>0.749924</td>
</tr>
<tr>
<td>3.4</td>
<td>-0.968039</td>
<td>2.234346</td>
<td>-0.392499</td>
<td>0.557690</td>
</tr>
<tr>
<td>3.6</td>
<td>-1.435305</td>
<td>2.319864</td>
<td>-0.505896</td>
<td>0.293662</td>
</tr>
<tr>
<td>3.8</td>
<td>-1.967423</td>
<td>2.454333</td>
<td>-0.636751</td>
<td>0.052527</td>
</tr>
<tr>
<td>4.0</td>
<td>-2.563416</td>
<td>2.629260</td>
<td>-0.783663</td>
<td>-0.491137</td>
</tr>
<tr>
<td>4.2</td>
<td>-3.219480</td>
<td>2.814268</td>
<td>-0.942624</td>
<td>-1.031862</td>
</tr>
<tr>
<td>4.4</td>
<td>-3.928306</td>
<td>1.872564</td>
<td>-1.106776</td>
<td>-1.683251</td>
</tr>
<tr>
<td>4.6</td>
<td>-4.678357</td>
<td>1.461037</td>
<td>-1.265363</td>
<td>-2.452012</td>
</tr>
<tr>
<td>4.8</td>
<td>-5.453076</td>
<td>0.883657</td>
<td>-1.404216</td>
<td>-3.342181</td>
</tr>
<tr>
<td>5.0</td>
<td>-6.230082</td>
<td>0.116035</td>
<td>-1.502086</td>
<td>-4.354140</td>
</tr>
<tr>
<td>5.2</td>
<td>-6.980346</td>
<td>-0.865839</td>
<td>-1.532394</td>
<td>-5.43504</td>
</tr>
<tr>
<td>5.4</td>
<td>-7.667394</td>
<td>-2.084516</td>
<td>-1.460896</td>
<td>-6.719859</td>
</tr>
<tr>
<td>5.6</td>
<td>-8.246576</td>
<td>-3.559746</td>
<td>-1.245070</td>
<td>-8.045364</td>
</tr>
<tr>
<td>5.8</td>
<td>-8.664445</td>
<td>-5.306844</td>
<td>-0.833385</td>
<td>-9.433251</td>
</tr>
<tr>
<td>6.0</td>
<td>-8.858316</td>
<td>-7.334745</td>
<td>-0.164858</td>
<td>-10.846223</td>
</tr>
<tr>
<td>6.2</td>
<td>-8.756063</td>
<td>-9.643738</td>
<td>0.830977</td>
<td>-12.234814</td>
</tr>
<tr>
<td>6.4</td>
<td>-8.276250</td>
<td>-12.22862</td>
<td>2.235040</td>
<td>-13.535755</td>
</tr>
<tr>
<td>6.6</td>
<td>-7.328688</td>
<td>-15.046992</td>
<td>4.129665</td>
<td>-14.670413</td>
</tr>
<tr>
<td>6.8</td>
<td>-5.815516</td>
<td>-18.073623</td>
<td>6.611577</td>
<td>-15.543406</td>
</tr>
<tr>
<td>7.2</td>
<td>-0.673698</td>
<td>-24.456478</td>
<td>13.703240</td>
<td>-16.032857</td>
</tr>
<tr>
<td>7.4</td>
<td>3.169454</td>
<td>-27.608769</td>
<td>18.481599</td>
<td>-15.367002</td>
</tr>
<tr>
<td>7.6</td>
<td>7.999380</td>
<td>-30.548262</td>
<td>24.166297</td>
<td>-13.875334</td>
</tr>
<tr>
<td>7.8</td>
<td>13.908909</td>
<td>-33.091539</td>
<td>30.783303</td>
<td>-11.372740</td>
</tr>
<tr>
<td>8.0</td>
<td>20.973953</td>
<td>-35.016725</td>
<td>38.311320</td>
<td>-7.660322</td>
</tr>
</tbody>
</table>
INVERSE HYPERBOLIC FUNCTIONS OF A REAL ARGUMENT

FUNCTION ARSH(X)

THIS SUBROUTINE RETURNS VALUES OF THE INVERSE HYPERBOLIC SINE.

Z = ABS(X)
ARSH = ALOG(Z * SQRT(Z**2 + 1.))
IF(Z.NE.X) ARSH = - ARSH
RETURN
END

FUNCTION ARCOSH(X)

THIS FUNCTION SUBROUTINE RETURNS VALUES OF THE INVERSE HYPERBOLIC COSINE. ARGUMENTS BETWEEN -1. AND +1. ARE INVALID AND CAUSE AN ERROR MESSAGE TO BE PRINTED. FOR NEGATIVE ARGUMENTS, NEGATIVE FUNCTIONAL VALUES ARE RETURNED.

Z = ABS(X)
IF(Z.GE.1.) GO TO 2
PRINT 1, X
1 FORMAT(4X, 24HATTEMPT TO FIND ARCCOSH, E10.3, 9H) INVALID)
RETURN
2 ARCOSH = ALOG(Z + SQRT(Z**2 - 1.))
IF(Z.NE.X) ARCOSH = - ARCOSH
RETURN
END
COMPLEX INVERSE FUNCTIONS

THE SUBROUTINE BELOW CALCULATES INVERSE FUNCTIONS OF COMPLEX ARGUMENTS \((x + jy)\). IT USES NORMAL LIBRARY FUNCTIONS ARCSIN AND ARCCOS, AS WELL AS ARCCOSH, WHICH MUST BE SUPPLIED BY THE USER.

SUBROUTINE CINFCT(U, V, NFCT, X, Y)
C THIS SUBROUTINE FINDS VALUES OF THE INVERSE TRIGONOMETRIC AND HYPERBOLIC SINES AND COSINES OF COMPLEX ARGUMENTS.
C
FOR NFCT = 1, U + JV = ARC SIN \((x + jy)\)
C NFCT = 2, U + JV = ARC COS \((x + jy)\)
C NFCT = 3, U + JV = ARC SINH \((x + jy)\)
C NFCT = 4, U + JV = ARC COSH \((x + jy)\)
C
IF(NFCT.NE.3) GO TO 1
Z = X
X = Y
Y = Z
1 Z = 1.
IF(Y.LT.0.) Z = -Z
P = (1. + X)**2 + Y**2
Q = P - 4.*X
P = SQRT(P)
Q = SQRT(Q)
R = 2. * X / (P + Q)
S = (P + Q) / 2.
IF(R.GT.1.) R = 1.
IF(S.LT.1.) S = 1.
GO TO (2, 3, 2, 3), NFCT
2 U = ARSIN(R)
V = Z * ARCOSH(S)
IF(NFCT.EQ.1) RETURN
Z = U
U = V
V = Z
Z = X
X = Y
Y = Z
RETURN
3 U = ARCOS(R)
V = -Z * ARCOSH(S)
IF(NFCT.EQ.2) RETURN
R = -Z * V
V = Z * U
U = R
RETURN
END
SUBROUTINE INVERS(A, N, DELTA, EPS)

C THIS SUBROUTINE INVERTS THE MATRIX A(N,N) AND RETURNS ITS DETERMI-
C NANT DELTA. EPS IS THE SMALLEST PIVOT USED IN THE GAUSSIAN INVER-
C SION. THE MAIN PROGRAMME MUST DIMENSION A(30,30).

C DIMENSION A(30,30), C(30), D(30), IZ(30)
DELTA = 1.
DO 1 J = 1, N
1 IZ(J) = J
DO 8 I = 1, N
   K = I
   Y = A(I,I)
   IP = I + 1
   IF(I.GE.N) GO TO 3
   DO 2 J = IP, N
      W = A(I,J)
      IF(ABS(W).LE.ABS(Y)) GO TO 2
      K = J
      Y = W
   CONTINUE
2 CONTINUE
3 DELTA = DELTA * Y
   IF(I.EQ.1) EPS = ABS(Y)
   EPS = AMIN1(EPS, ABS(Y))
   IF(EPS.EQ.0.) RETURN
   Y = 1./Y
   DO 4 J = 1, N
      C(J) = A(I,J)
   4 A(I,J) = C(J) + Y
   DO 5 J = 1, N
      A(I,J) = A(I,J) * Y
   5 CONTINUE
   DO 9 I = 1, N
      IF(I.EQ.1) GO TO 9
      DO 8 J = 1, N
         IF(J.EQ.I) A(I,J) = A(I,J) - D(J) * C(I)
      8 CONTINUE
   9 CONTINUE
   DO 13 I = 1, N
      IF(I.EQ.1) GO TO 13
      DO 12 J = 1, N
         W = A(I,J)
         A(I,J) = A(I,J) - D(J) * C(I)
      12 CONTINUE
   13 CONTINUE
RETURN
END
Index

A

Action at a distance, 14
Ampère's law, 127
Analogues:
  conductive paper, 91
  electrolytic tank, 90, 286
  resistance mesh, 91
Analytic functions, 94
Anomalous loss, 227
Anomaly factor, 227
Area transformation, 284
  by electrolytic tank, 287
Attenuation constant, 257

B

ber, bei functions, 229
Bessel functions, 228
Bessel's equation, of order zero, 228
Bilinear transformation, 195
Boundary conditions:
  at material surfaces, 133
  electrostatic, 36
  for TEM waves, 259
  in waveguides, 267
  normal-gradient, 184
Bundle conductors, 49
  potential distribution, 50

C

Capitance, 73
Capacitance coefficients, 69
Capacitor, 73
Cassini curves, 50
Cauchy-Riemann conditions, 93, 188, 284
Characteristic impedance, 259
calculation of, 260
Charge, 15, 18
  independence of velocity, 117
  surface, 24
Charge density, surface, 35
Charge distribution:
  energy of, 30
  on metal rod, 71
Circular polarization, 251
Coaxial cable:
  capacitance of, 74
  electrostatics of, 42
  TEM waves in, 259
Complex potentials, 95
Conductive paper analogues, 91
Conductivity, 86
Conformal mapping, 98
  by electrolytic tank, 287
determination of constants, 111
  edge of thin plate, 109
  Helmholtz equation, 283
  magnetic fields, 189
  parallel-plate capacitor, 109
  Schwarz-Christoffel, 106
  semicircle, 104
  vertices at infinity, 108
Conjugate functions, 96
Conservative field, 26
Constitutive relations, 128, 213
Continuity equation, 84, 129
Continuity principle, 15
Convergence rate of iteration, 63, 283
Corona, 48
Coulomb convention, 144
Coulomb's law, 20
  for moving charges, 122
Current density, 85
  surface, 134
Current distribution in flat bar, 235
Current flow around a hole, 104
Curvilinear square, 55
Cutoff frequency, 270
  rectangular waveguide, 273
  Cutoff frequency (Cont.)
    ridged waveguide, 279

D
Depth of penetration, 232
Diffusion equation, 215
  in pulse transformer, 218
Dipole energy, 150
Dipole sheet, 150
Displacement current, 130
Double Fourier series, 222
Doubly iterative solutions, 275

E
Earnshaw's theorem, 33
Eddy currents:
  sinusoidal, 226
  transient, 220, 223
Electric charge, 18
Electric field, 22
  relativistic transformation, 118
Electric flux, 15
Electric potential, 27
Electric scalar potential, 233
Electrodynamic field, 120
Electrodynamic potentials, 233
Electrodynamics, 116
Electrolytic tank, 90
double-layer, 94
Electromagnetic units, 297
Electromechanical devices, 76
Electromotive force, 135
Electrostatic field and TEM waves, 258
Electrostatic generator, 77
Electrostatic units, 297
Electrostatic voltmeter, 76
Elliptic integral, complete:
  first kind, 30, 155
  second kind, 155
Elliptic polarization, 251
Energy, electrostatic, 30, 69
Energy density:
  electric field, 31
  magnetic field, 153
INDEX

Equipotentials:
- electric, 33
- magnetic scalar, 188
Ether, 3, 14

F
Faraday's law, 136
Felici machine, 77
Field:
- distribution, rectangular waveguide, 272
- energy, magnetic, 151
maps:
- characteristic impedance calculation, 261
- inductance calculation from, 206
of moving charge, 116
TE or TM in waveguide, 258
time-varying, 212
Filamentary source in a slot, 194
Finite differences, error estimate, 114
Flat strip conductor, field of, 101
Flux:
- electric, 15
- magnetic, 124
Flux density:
- electric, 16
- transformation rule, 117
Flux distribution in transformer core, 192
Force:
- between moving charges, 119
- between point particles, 22
- density, 128
- magnetic, 122
Fourier series:
- one-dimensional, 219
- two-dimensional, 222
Fundamental postulates, 18

G
Gaussian units, 297
Gauss's law, 18
Geometric mean distance, 176, 237
Graphical relaxation, 54

H
Helmholtz equation:
- conformal mapping, 283
- numerical solutions, 273
- plane waves, 256
- scalar, for waveguides, 266
- vector form, 216
- vector, periodic case, 217
High-power waveguide, 285
Horizontal polarization, 250
Hysteresis, 227

I
Images:
- electric, 52
- magnetic, 179
- partial magnetic, 179
Incident wave, 254
Inductance, 157
- calculation from field maps, 206
- direct calculation, 163
- from flux and m.m.f., 200
- general expression, 183
- leakage, 183
- mutual:
  - triangular loops, 161
  - two circular rings, 157
- pot-core, approximate, 203
- related to reluctance, 201
self:
  - circular ring, 158
  - parallel-line, 164
- square loop, 159
- solenoidal coils, 165
- toroidal coil, approximate, 201
Induction coefficients, 66, 69
Inertial frames, 2
Integral equation, 236
- solved by inversion, 237
Integral-value problems, 230
Interference, of plane waves, 251
Interval, 12
Intrinsic impedance, 248, 255
Ionization, 48
Irrotational field, 26
Iteration, Liebmann’s method, 62

K
Kelvin functions, 229
Kinematics, relativistic, 8
Kirchhoff’s circuit laws, 137

L
Laplace’s equation, 32
cylindrical coordinates, 42
in conductors, 87
and TEM waves, 258
Laplacian, finite-difference, 274
cylindrical coordinates, 204
Leakage flux, 182
Leakage inductance, 187
Length contraction, relativistic, 7
Liebmann iteration, 62
cylindrical coordinates, 203
Helmholtz equation, 274
illustrative programme, 63
Poisson’s equation, 183
Light, velocity of, 3
Line integral, finite-difference approximation, 206
Linear polarization, 250
Lorentz condition, 234
Lorentz contraction, 7
Lorentz factor, 6
addition of, 9
Lorentz force, 125
low-velocity approximation, 126
Lorentz transformation, 5
Lossless wave equation, 245

M
Magnetic dipole, 149
Magnetic field, 122
conformal mapping, 189
current-carrying conductor, 172
current-carrying ring, 154
energy, 151
filamentary current in slot, 197
Magnetic field (Cont.)
finite-difference solution, 174
flat conductor, 173
hollow rectangular conductor, 177
solenoidal property, 124
two-dimensional, integral form, 172
Magnetic flux, 124
Magnetic flux density, 124
Magnetic images, 179
Magnetic moment, 147
Magnetic Ohm’s law, 200
Magnetic scalar potential, 187
Magnetic susceptibility, 131
Magnetic torque, 147
Magnetic vector potential, 143, 233
explicit calculation, 153
physical view, 147
Magnetization vector, 131
Magnetizing:
current, 225
impedance, 224
Magnetomotive force, 199
Magnetostatic field and TEM waves, 258
Material boundaries, fields at, 35, 132
Material media:
-electric properties, 23
-magnetic properties, 131
Matrix inversion, complex, 237
Matter-energy, 19
Maxwell’s displacement current, 130
Maxwell’s equations:
differential form, 130
integral form, 131
Modes of propagation in waveguides, 263
Monte Carlo method, 65
Moving media, 216

N
Neumann’s formula, 156
computer programme, 161
in summation form, 160
Neutrality principle, 20
INDEX

Nodes, of a network, 137
Normal modes, 221
Normal-gradient boundaries, in waveguides, 280
Normal-gradient boundary condition, 262
Numerical relaxation, 56

Oblique polarization, 250
Ohm's law, 86
Ohm's law, magnetic, 200
Orthogonal modes, 220
Overrelaxation, 62
Overrelaxation factor, 275
" optimum, 277

Parallel-plate capacitor:
 field at edges, 109
 infinite, 75
Parallel-wire line:
 electrostatics of, 44, 96
 magnetic field, 164
Parallel-wire screen, potential distribution, 99
Partial images, 179
Permeability, 122, 132
Permittivity, 25
Phase, surface of constant, 249
Phase constant, 257
Phase velocity, 257, 269
Plane waves, 247, 249
general, 255
Poisson's equation, 32
 finite-difference, 57
 iterative solution of, 183
 vector, 144
 formal solution, 154
Polarization, 24
 of plane waves, 249
Pot-core inductor:
 iterative solution, 205
 reluctance of, 202
Potential:
 coefficient matrix, 67
 coefficients, 66
 complex, 95
electric, 27
electrodynamic, 233
gradient, 27
magnetic:
 scalar, 187
 vector, 143
Power, 88
density, 89
flow density, 240
Poynting's theorem, 238
Poynting's vector, 239
plane wave, 249
Principal mode of waveguide, 281
Propagation constant, 257
 rectangular waveguide, 269
 unbounded wave, 258
Pulse transformer, fields in core, 217

Random walk, 65
Rationalized units, 296
Rayleigh coefficient, 276
computation, 281
general case, 285
Recording head, field near, 190
Rectangular core transient problem, 220
Reference frames, inertial, 2
Reflected wave, 253
Reflection coefficient, 255
Reflection of plane waves, 253
Relativity, special, 1
Relaxation (see also Iteration), 53
graphical, 54
numerical, 56
Reluctance, 200
Residual, 57
Resistance, 90
Ridged waveguide, principal modes, 277
RMKSC units, 298
INDEX

S
Scalar potential:
  electric, 27
  integral solution, 27
  magnetic, 187
Schwarz-Christoffel:
  differential equation, 108
  transformation, 106, 190, 192
Separation of variables, 215, 221
  plane wave problem, 256
Skin depth, 232
Skin effect:
  flat bar, 235
  high frequencies, 230
  round wire at low frequency, 227
Space-time, 14
Standing waves, 252
Stream function, 95
Subareas, method of, 70
Superposition:
  electric potentials, 44
  of partial images, 179
Surface charge:
  accompanying current flow, 88
  density, 35
  on flat conductor, 103
Surface current:
  at reflective boundary, 254
  density, 134, 232
Surface of constant phase, 249
Susceptibility:
  electric, 24
  magnetic, 131

T
TE waves, 265
  rectangular waveguide, 271
TEM waves, 255
  transmission line, 258
Time dilation, 7
Time-varying fields, 212
TM waves, 264
  in rectangular waveguide, 269
Torque, on current-carrying loop, 149
Transformer:
  flux distribution in, 192
  lamination, field in, 225
  leakage field of, 182
  magnetizing impedance of, 224
Transmission coefficient, 255
Transmitted wave, 254
Traveling waves, 245
Two-wire line (see Parallel-wire line)

U
Unit magnetic dipole, 149
Units, 16, 23, 295
  electromagnetic, 297
  electrostatic, 297
  Gaussian system, 297
  rationalized, 296
  rmksc system, 298
  unrationalized, 296

V
Variables, separation of, 215
Vector Helmholtz equation, 216
Vector Poisson equation, 144
Vector potential, magnetic, 143
Velocity addition, relativistic, 9
Velocity of light, 3
  wave propagation at, 258
Velocity of propagation, unbounded wave, 258
Vertical polarization, 250
Volume force density, 128

W
Wave equation:
  field-vector, 212
  free-space, 214
  general homogeneous, 213
  general inhomogeneous, 214
  lossless, 214, 245
Wave number, 216
  complex, 224
Work in conservative fields, 25
the current periodical literature and illustrate practical applications of the methods introduced in the text.

- Illustrative examples are drawn from practical problems associated with simple devices already familiar to the reader.

- Complete computer programs are provided for nearly all numerical examples to assist problem solving. Programs for nearly all assigned numerical problems appear in an accompanying solutions manual.

P. Silvester, Ph.D., is presently Associate Professor of Electrical Engineering, McGill University, and is a regular contributor to professional journals.
FIELD-EFFECT TRANSISTORS: Physics, Technology and Applications
edited by J. TORKEL WALLMARK, Chalmers University, Gothenburg, Sweden, and HARWICK JOHNSON, Electronic Research Laboratory, RCA Laboratories

This volume covers all the important aspects of the field-effect transistor, including surface physics, device fabrication, and electronic circuit theory. A great deal of the material is based on very recent developments and has never before been published in book form. The author of each chapter is an expert in the particular phase of the subject with which he deals.

In addition to the theoretical discussion, the authors provide a great deal of practical information, such as: formulas for the time and temperature of oxide growth; equations for device performance; experimental information on radiation tolerance; transistor processing sequence and technology; thin-film transistor technology; typical admittances as a function of frequency; and circuit design formulas.

Published 1966 376 pages

FOURIER TRANSFORMS AND THE THEORY OF DISTRIBUTIONS
by J. ARSAC, Numerical Analysis Services, Meudon (Paris) Observatory
The Translators: ALLEN NUSSBAUM, University of Minnesota and GRETCHEN C. HEIM, Northwestern Bell Telephone Company

In writing this book, the author has merged the viewpoints of the mathematician and the physicist and has run the middle course between extreme rigor and none at all. An example is the widely divergent treatment of the Dirac delta function, ranging from the intuitive handling accorded it by the physicist to the fundamental exposition in the Theory of Distributions by L. Schwartz. By reducing the theory of distributions to its fundamentals and avoiding the fine points of function theory, the author has been able to treat existence and convergence conditions without too much difficulty, using conventional means. Thus, the book fills a definite gap and will be recognized as the work of an expert who is making his specialized knowledge available to a wider audience.

The book is divided into three sections. The first is purely mathematical, involving basic concepts that are to be used later. Part Two is concerned with applications. One topic considered is the study of linear filters. The final section covers numerical calculations and emphasizes general theory rather than specific problems.

Published 1966 416 pages

PRENTICE-HALL, Inc.
Englewood Cliffs, New Jersey

59311