Electromagnetic Waves

Books From

Bell Telephone Laboratories

THEORY OF VIBRATING SYSTEMS AND SOUND. *By* Irving B. Crandall, *Late Member of the Technical Staff, Bell Telephone Laboratories.*

CONTEMPORARY PHYSICS. *By* Karl K. Darrow, *Member of the Technical Staff, Bell Telephone Laboratories.* Second Edition.

SPEECH AND HEARING. *By* Harvey Fletcher, *Acoustical Research Director, Bell Telephone Laboratories.* With an Introduction by H. D. Arnold, *Director of Research, Bell Telephone Laboratories.*

PROBABILITY AND ITS ENGINEERING USES. *By* Thornton C. Fry, *Member of the Technical Staff, Bell Telephone Laboratories.*

ELEMENTARY DIFFERENTIAL EQUATIONS. *By* Thornton C. Fry. Second Edition.

TRANSMISSIONS CIRCUITS FOR TELEPHONIC COMMUNICATION. METHODS OF ANALYSIS AND DESIGN. *By* K. S. Johnson, *Member of the Technical Staff, Bell Telephone Laboratories.*

A FUGUE IN CYCLES AND BELS. *By* John Mills, *Director of Publication, Bell Telephone Laboratories.*

TRANSMISSION NETWORKS AND WAVE FILTERS. *By* T. E. Shea, *Special Products Engineer, Bell Telephone Laboratories.*

ECONOMIC CONTROL OF QUALITY OF MANUFACTURED PRODUCT. *By* W. A. Shewhart, *Member of Technical Staff, Bell Telephone Laboratories.*

THE APPLICATION OF ELECTROMECHANICAL IMPEDANCE ELEMENTS IN TRANSDUCERS AND WAVE FILTERS. *By* Warren P. Mason, *Member of the Technical Staff, Bell Telephone Laboratories, Inc.*

RHOMBIC ANTENNA DESIGN. *By* A. E. Harper, *Bell Telephone Laboratories.*

POISSON'S EXPONENTIAL BINOMIAL LIMIT. *By* E. C. Molina, *Switching Theory Engineer, Bell Telephone Laboratories.*

ELECTROMAGNETIC WAVES. *By* S. A. Schelkunoff, *Member of the Technical Staff, Bell Telephone Laboratories.*

FUNDAMENTAL THEORY OF SERVOMECHANISMS. *By* LeRoy A. MacColl, *Member of the Technical Staff, Bell Telephone Laboratories, Inc.*

QUARTZ CRYSTALS FOR ELECTRICAL CIRCUITS. *By* Raymond A. Heising, *Radio Research Engineer, Bell Telephone Laboratories, Inc.*

Published by

D. Van Nostrand Company, Inc.

ELECTROMAGNETIC WAVES

By

S. A. SCHELKUNOFF

Member of the Technical Staff

Bell Telephone Laboratories, Inc.

FOURTH PRINTING

NEW YORK

D. VAN NOSTRAND COMPANY, Inc.

250 Fourth Avenue

First Published, April 1943
Reprinted July 1943, May 1944, October 1945

Produced by
TECHNICAL COMPOSITION CO.
BOSTON, MASS.

PRINTED IN THE UNITED STATES OF AMERICA

PREFACE

In the summer of 1942 it was my pleasure to give a course on Electromagnetic Waves at Brown University in connection with its Program of Advanced Instruction and Research in Mechanics. There I not only enjoyed the opportunity to test this book in manuscript but, through generous arrangements made by the University, I was enabled to put it in final shape for publication. To the Officers of Brown University, and particularly to R. G. D. Richardson, Dean of the Graduate School, I am grateful for their interest in the book and for the facilities which they put at my disposal.

As a whole, the book is an outgrowth of my research and consulting activities in Bell Telephone Laboratories. Its first draft was prepared in connection with courses of lectures in the Laboratories' "Out-of-Hour" program. Courses were given in 1933–34 and 1934–35, for which the lectures were mimeographed under the title "Electromagnetic Theory and Its Applications." A third course was given in 1941–42, when the notes were revised under the present title "Electromagnetic Waves."

If this book proves to be a "practical theory" of electromagnetic waves it will be largely due to my close association with experimentalists in the Bell Laboratories. Some credit for its final issuance is due to Dr. H. T. Friis who for years urged me to publish my notes. To Dr. M. J. Kelly and Dr. Thornton C. Fry I am grateful for arranging a leave of absence for my work at Brown University.

I am particularly indebted to Miss Marion C. Gray for her invaluable assistance throughout the entire preparation of this book.

S. A. S.

NEW YORK, N. Y.
January, 1943

TO THE READER

Since 1929 the opportunities for practical applications of electromagnetic theory have increased so spectacularly that a new approach has become almost a necessity. The old practice of working out each boundary value problem as if it were a new problem is being abandoned as repetitious and uneconomical because it fails to coordinate the various results. In the interest of unity, simplicity, compactness and physical interpretation, the conceptions of one-dimensional wave theory are being extended to waves in three dimensions and field theory is no longer considered as something apart from circuit and transmission line theories.

All physical fields are three dimensional; but in some circumstances either two or all three dimensions are unimportant; then they may be "integrated out" and thus "concealed"; in the first case the problem belongs to "transmission line theory" and in the second to "network theory." This suppression of some or all physical dimensions is analogous to the method of "ignoration of coordinates" in mechanics; and it may or may not involve approximations. It is a mistake to say that the circuit and line theories are approximate while only the field theory is exact. In fact in many important cases a three-dimensional problem is rigorously reducible to a set of one-dimensional problems. Once the one-dimensional problem has been solved in sufficiently general terms, the results can be used repeatedly in the solution of more general problems.

This point of view leads to a better understanding of wave phenomena; it saves time and labor; and it benefits the mathematician by suggesting to him more direct methods of attacking new problems. Once these ideas are more generally disseminated, large sections of electromagnetic theory can be explained in terms intelligible to persons with elementary engineering education.

The classical physicist, being concerned largely with isolated transmission systems, has emphasized only one wave concept, that of the velocity of propagation or more generally of the propagation constant. But the communication engineer who is interested in "chains" of such systems from the very start is forced to adopt a more general attitude and introduce the second important wave concept, that of the impedance. The physicist concentrates his attention on one particular wave: a wave of force, or a wave of velocity or a wave of displacement. His original differential equations may be of the first order and may involve both force and velocity; but by tradition he eliminates one of these variables, obtains a second order differential equation in the other and calls it the "wave equation." Thus he loses sight of the interdependence of force and velocity waves and he does not stress the difference which may exist between waves in different media even though the velocity of wave propagation is the same. The engineer, on the other hand, thinks in terms of the original "pair of wave equations" and keeps constantly in mind this interdependence between force and velocity waves. In this book I have injected the communication engineer's attitude into an orderly development of "field theory."

If the modern theory of electromagnetism were to be presented in four ideal volumes, then the first volume would treat the subject broadly rather than thoroughly, with emphasis on more elementary topics. The second volume would be devoted to electromagnetic waves in passive media free from space charge; in this volume electric generators would appear merely as given data, either as electric intensities tangential to the boundaries of the "generator regions" or as given currents inside these regions. Another volume, on "electromechanical transducers," would deal with interaction between mechanical and electrical forces and the final volume on "space charge waves" would be devoted to phenomena in vacuum tubes. The present book is confined to the material which would properly belong to the second of these volumes.

It is intended as a textbook and for reference. In it a practicing engineer will find basic theoretical information on radiation, wave propagation, wave guides and resonators. Those engaged in theoretical research will find a stock of equations which may serve as a starting point for further investigations.

Chapters 1 and 3, dealing with vector analysis and special functions, such as Bessel functions and Legendre functions, are intended for ready reference. These chapters are brief because it is only necessary for the reader to be familiar with the language of vectors and, in most cases, only elementary properties of the special functions are needed. Chapter 2 deals with applications of complex variables to the theory of oscillations and waves and Chapter 4 reviews the fundamental conceptions and equations. Elements of circuit theory are presented in Chapter 5; there the three-dimensional character of electromagnetic fields is suppressed and the discussion is conducted in terms of resistance, inductance and capacitance. Chapter 6 is concerned with some general aspects of waves in free space, on wires, and in wave guides. Its last few sections cover electrostatics and magnetostatics to the extent needed in wave theory. The one-dimensional wave theory is presented in great detail in Chapter 7. The following chapter treats the simplest types of waves in free space and in wave guides. Chapter 10 contains a more general, systematic treatment of such waves. Chapter 9 is devoted to radiation from known current distributions and to the directive properties of antennas, antenna arrays and electric horns. Chapter 11 presents a recent antenna theory and, finally, Chapter 12 deals with certain impedance discontinuities in wave guides.

There is enough material for an intensive six-hour course; the particular order adopted is best suited to students of communication engineering and microwave transmission. In the case of radio engineers, the first four sections of Chapter 8 may be followed by Chapter 9; and in the case of students of physics or applied mathematics these four sections may be followed directly by Chapter 10. For a shorter course the instructor will find it easy to select the material best suited to the needs of his students.

THE AUTHOR

CONTENTS

In the manufacture of this book, the publishers have observed the recommendations of the War Production Board and any variation from previous printings of the same book is the result of this effort to conserve paper and other critical materials as an aid to the war effort.

CHAPTER I

VECTORS AND COORDINATE SYSTEMS

1.1. *Vectors*

VECTOR is a generic name for such quantities as velocities, forces, electric intensities, etc. A vector can be represented graphically by a directed

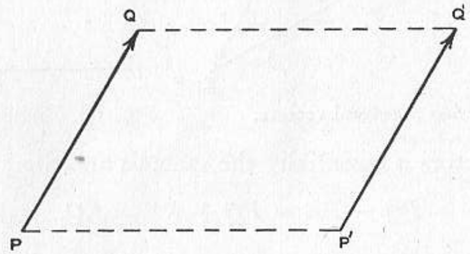


FIG. 1.1. Equal vectors.

segment PQ (Fig. 1.1) whose length is proportional to the magnitude of the vector. Two parallel vectors PQ and $P'Q'$ having the same magnitude and direction are considered equal.

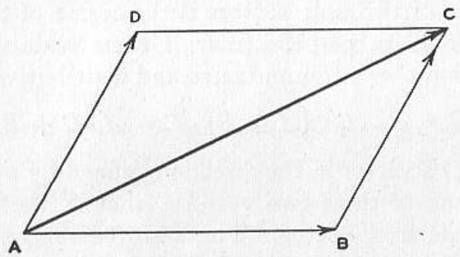


FIG. 1.2. Addition of two vectors.

The method of adding vectors is what distinguishes them from other quantities. This method consists in obtaining the diagonal of the parallelogram constructed on two vectors as adjacent sides (Fig. 1.2); thus*

$$AB + AD = AC.$$

*We shall use no special marks to designate vectors if the meaning is clear from the context; otherwise we shall use a bar over the letters.

Any number of vectors may be added by using the end of one vector as the origin of the next; the vector drawn from the origin of the first to the end of the last is the sum (Fig. 1.3).

Since by definition

$$AB + BA = 0, \quad \text{or} \quad BA = -AB,$$

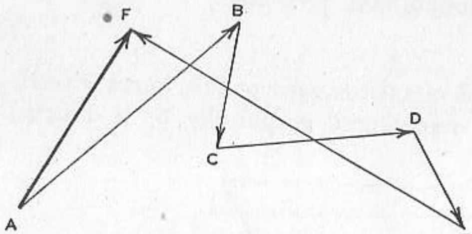


FIG. 1.3. Addition of several vectors.

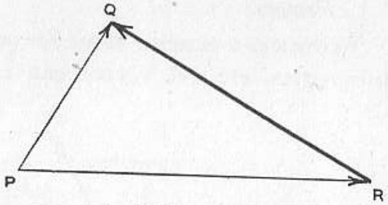


FIG. 1.4. Subtraction of vectors.

subtraction of vectors is essentially the same as addition; thus (Fig. 1.4)

$$PQ - PR = PQ + RP = RQ.$$

Hence the difference of two vectors drawn from the same origin is the vector connecting the end of the second to the end of the first.

The *scalar product* of two vectors is defined as the product of their magnitudes and the cosine of the angle between them; thus

$$A \cdot B = (A,B) = ab \cos \psi.$$

The scalar product of two unit vectors is the cosine of the angle between them. Two vectors are perpendicular if their scalar product is zero. Scalar multiplication obeys commutative and distributive laws

$$A \cdot B = B \cdot A, \quad (A + B) \cdot C = A \cdot C + B \cdot C.$$

The *component* of a vector in the direction defined by a given unit vector is the scalar product of these two vectors; that is, the projection of the given vector on the unit vector. The *direction components* of a vector drawn from point $P(x_1,y_1,z_1)$ to $Q(x_2,y_2,z_2)$, taken in the positive directions of the coordinate axes, are the differences $x_2 - x_1$, $y_2 - y_1$, $z_2 - z_1$. If l is the length of the vector and α, β, γ are the angles the vector makes with the coordinate axes, then

$$PQ_x = x_2 - x_1 = l \cos \alpha,$$

$$PQ_y = y_2 - y_1 = l \cos \beta,$$

$$PQ_z = z_2 - z_1 = l \cos \gamma.$$

The scalar product of any two vectors may be expressed as the sum of the products of their direction components

$$A' \cdot A'' = A'_x A''_x + A'_y A''_y + A'_z A''_z.$$

Hence the cosine of the angle ψ between the vectors is the sum of the products of their direction cosines

$$\cos \psi = \cos \alpha'_x \cos \alpha''_x + \cos \alpha'_y \cos \alpha''_y + \cos \alpha'_z \cos \alpha''_z.$$

The *vector product* $A \times B$ or $[A,B]$ of A and B is a vector perpendicular to both, pointing in the direction in which a right-handed screw would advance if turned from vector A to vector B through the smaller angle (Fig. 1.5); the magnitude of the vector product is the product of the magnitudes of A and B and of the sine of the angle between them, that is, the area of the parallelogram constructed on A and B as adjacent sides. For vector products we have

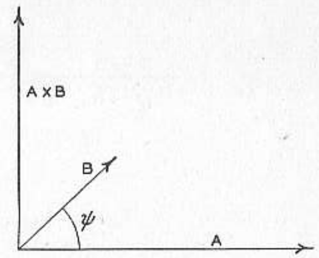


FIG. 1.5. The vector product.

$$A \times B = -B \times A, \quad (A + B) \times C = A \times C + B \times C.$$

The components of a vector product are expressed in terms of the direction components of the constituent vectors as follows

$$(A' \times A'')_x = A'_y A''_z - A'_z A''_y,$$

$$(A' \times A'')_y = A'_z A''_x - A'_x A''_z,$$

$$(A' \times A'')_z = A'_x A''_y - A'_y A''_x.$$

1.2. *Functions of Position*

A *function of position* or a *point function* is a function $f(x,y,z)$ depending only on the position of points. Loci of equal values of a point function are called *level surfaces* or *contour surfaces*; in the two dimensional case we have *level lines* or *contour lines*. Some level surfaces bear special names such as equipotential or isothermal or isobaric surfaces. Figure 1.6 illustrates how a two-dimensional point function may be represented graphically by drawing contour lines. The solid lines are the contour lines for $u = \log \rho_1/\rho_2$, where ρ_1 and ρ_2 are the distances from two fixed points; and the dotted lines are the contour lines for the angle ϑ made by BP with PA as shown in Fig. 1.6(a).

The rate of change of a point function depends not only on the position of a point but also on the particular direction of travel. If ΔV is the change in the value of a point function $V(x,y)$ as we pass from $A(x,y)$ to $B(x + \Delta x, y + \Delta y)$ and if Δs is the distance AB (Fig. 1.7), the ratio $\Delta V/\Delta s$

is the *average rate of change* of $V(x,y)$ in the direction AB. The limit of this ratio as B approaches A while remaining on the same straight line is the *directional derivative* of $V(x,y)$ in the direction AB. This derivative is denoted by $\partial V/\partial s$. Partial derivatives $\partial V/\partial x$ and $\partial V/\partial y$ are simply the directional derivatives taken along coordinate axes.

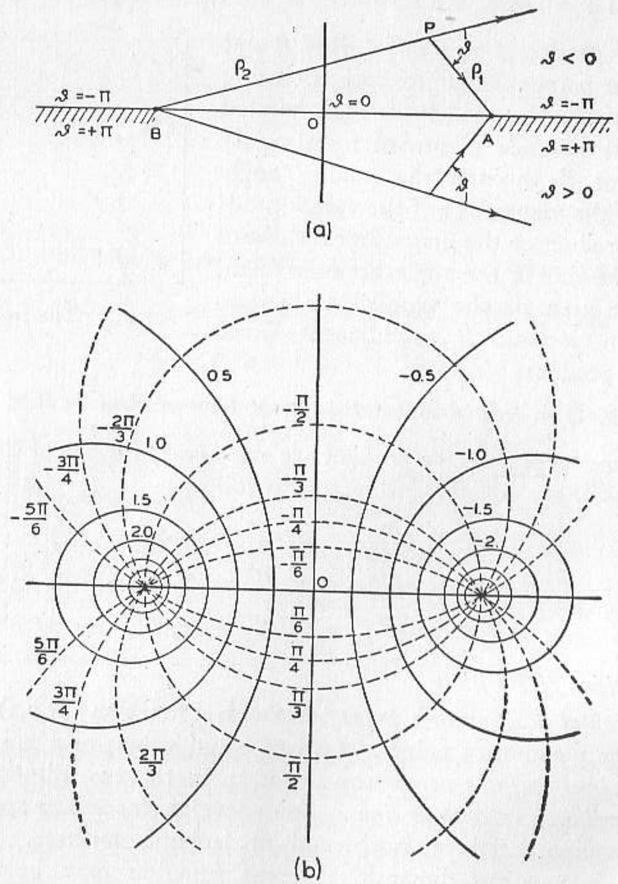


FIG. 1.6. Two families of contour lines.

The maximum rate of change is along the normal to the level line through A (Fig. 1.8). The *gradient* of V is defined as a vector along this normal

$$\text{grad } V = \frac{\partial V}{\partial n}\,\overline{n},$$

where $\overline{n}$ is the unit vector orthogonal to the level line. For an infinitesimal curvilinear triangle, we have

$$\Delta n = (\Delta s)\cos\psi,$$

and therefore

$$\frac{\partial V}{\partial s} = \frac{\partial V}{\partial n}\cos\psi. \tag{2-1}$$

Hence *the directional derivative of a point function is the component of its gradient in that particular direction.*

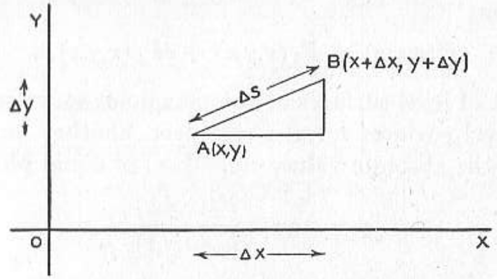


FIG. 1.7. Illustrating directional increments.

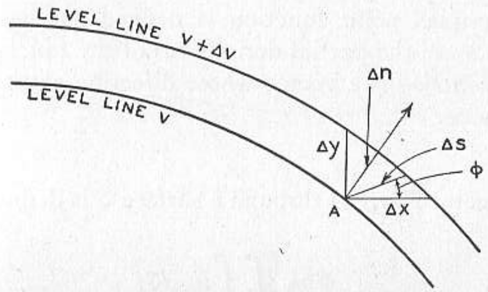


FIG. 1.8. Illustrating the gradient.

These equations are of course equally true for three-dimensional point functions. If α, β, γ are the angles made with the coordinate axes by the normal to the level surface at A, we have by (1)

$$\frac{\partial V}{\partial x} = \frac{\partial V}{\partial n}\cos\alpha, \quad \frac{\partial V}{\partial y} = \frac{\partial V}{\partial n}\cos\beta, \quad \frac{\partial V}{\partial z} = \frac{\partial V}{\partial n}\cos\gamma. \tag{2-2}$$

Thus the partial derivatives are the direction components of the gradient, and we have

$$\frac{\partial V}{\partial n} = \sqrt{\left(\frac{\partial V}{\partial x}\right)^2 + \left(\frac{\partial V}{\partial y}\right)^2 + \left(\frac{\partial V}{\partial z}\right)^2}.$$

Another expression for the normal derivative can be obtained if the equations in the set (2) are multiplied by cos α, cos β, cos γ respectively and then added

$$\frac{\partial V}{\partial n} = \frac{\partial V}{\partial x}\cos\alpha + \frac{\partial V}{\partial y}\cos\beta + \frac{\partial V}{\partial z}\cos\gamma.$$

This relation can be written down directly if we consider that the gradient is the sum of the projections of its components upon itself.

A *complex point function* is a function whose real and imaginary parts are point functions

$$V(x,y,z) = V_1(x,y,z) + iV_2(x,y,z).$$

We cannot speak of level surfaces of complex point functions since there is one family of level surfaces for the real part, another for the imaginary part, a third for the absolute value, etc. Loci of equal phase

$$\psi = \tan^{-1}\frac{V_2(x,y,z)}{V_1(x,y,z)}$$

of a complex point function are called *equiphase surfaces*; they are used in the classification into plane, cylindrical, spherical, etc., waves. The gradient of a complex point function is defined as the complex vector whose components are the partial derivatives of the function.

A *vector point function* is a vector whose direction components are ordinary point functions.

1.3. *Divergence*

The flux of a vector $F(x,y,z)$ through a surface S is defined as the surface integral

$$\Phi = \iint F_n\,dS,$$

where F_n is the component normal to the surface of integration. The outward flux of F through a simply connected closed surface S divided by the volume v enclosed by S is called the *average divergence* of F. The limit of the average divergence as S contracts to a point is the *divergence of* F *at that point*; thus

$$\operatorname{div} F = \lim \frac{\iint F_n\,dS}{v}, \text{ as } S \to 0.$$

Dividing the total volume v surrounded by the surface S into elementary cells, we observe that the total flux of F across the surface is the sum of the fluxes through the boundaries of the elementary cells, the fluxes through the common partitions between the cells contributing nothing to the whole. Since the flux through the boundary of a typical cell is div $F\,dv$, we have

$$\iint F_n\,dS = \iiint \operatorname{div} F\,dv. \tag{3–1}$$

The *surface divergence* is defined similarly; thus

$$\operatorname{div}' F = \lim \frac{\int F_n\,ds}{S}, \text{ as } s \to 0,$$

where s is the boundary of the elementary area S. The *linear divergence* is merely the ordinary derivative.

1.4. *Line Integral, Circulation, Curl*

The *line integral* of a vector F along a path AB (Fig. 1.9) is defined as the integral $\int_{(AB)} F_s\,ds$ of the tangential component of the vector. If F is a force, this integral represents the work done by F on a particle moving

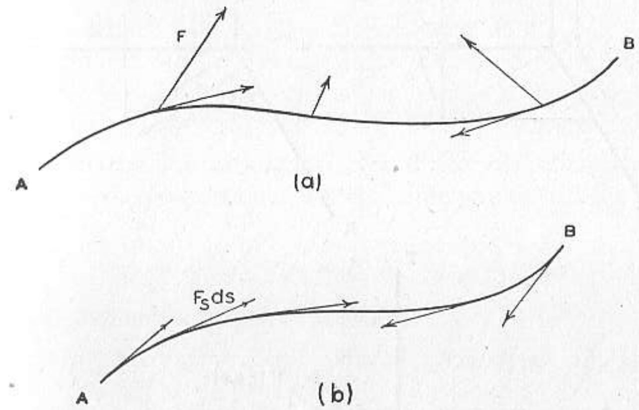


FIG. 1.9. Illustrating the line integral.

along AB. If the curve is closed, the line integral is called the *circulation*. The circulation per unit area of an infinitely small loop so oriented that the circulation is maximum is denoted by curl F; it is a vector perpendicular to the plane of the loop. The positive directions of this vector and circulation are related as shown in Fig. 1.10.

Consider a surface S bounded by a simple closed curve. Dividing S into elements, we observe that the circulation of F along the boundary of S is the sum of the circulations round the boundaries of the elements, since

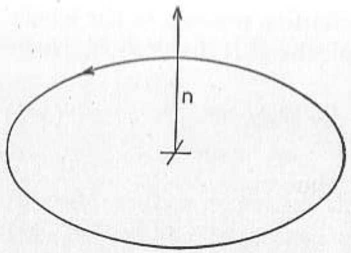


FIG. 1.10. The relationship between the positive directions of the circulation of F and the curl F.

the contributions due to the boundaries common to adjacent elements cancel out. Since the circulation round the boundary of each element is $\text{curl}_n\, F\, dS$ we have

$$\int F_s\, ds = \iint \text{curl}_n\, F\, dS. \quad (4\text{–}1)$$

1.5. *Coordinate Systems*

In practical applications the most frequently used coordinates are *rectangular*, *cylindrical*, and *spherical*; in these systems a typical point P is denoted by (x,y,z), (ρ,φ,z), (r,θ,φ) respectively. The meaning of these

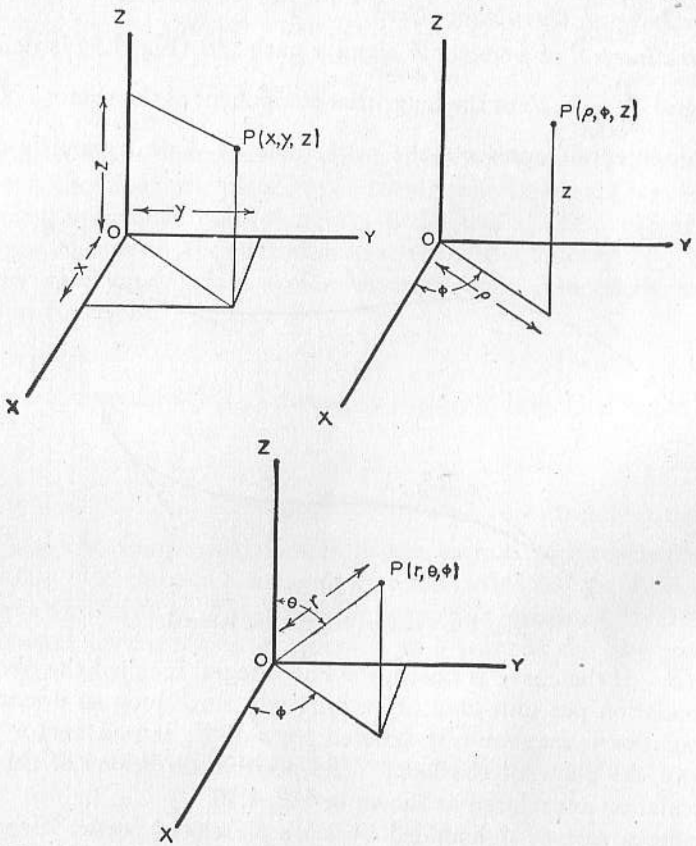


FIG. 1.11. Cartesian, cylindrical, and spherical systems of coordinates.

coordinates is explained in Fig. 1.11; x, y, z are the distances from three mutually perpendicular planes; ρ is the distance from the z-axis; r is the

distance from the origin; the "polar angle" θ is the angle between the radius r and the z-axis; the "longitude" φ is the angle between the xz-plane and the plane determined by the z-axis and the point P.

In a general system of coordinates, a point $P(u,v,w)$ is specified as a point of intersection of three surfaces

$$f_1(x,y,z) = u, \quad f_2(x,y,z) = v, \quad f_3(x,y,z) = w.$$

The lines of intersection of these *coordinate surfaces* are *coordinate lines*; thus u-lines are intersections of v- and w-surfaces.

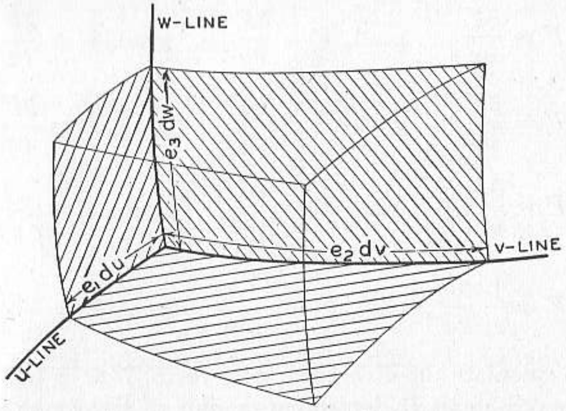


FIG. 1.12. An elementary coordinate cell.

If the coordinates are orthogonal, the differential distances along coordinate lines are proportional to the differentials of the coordinates (Fig. 1.12); thus

$$ds_u = e_1\, du, \quad ds_v = e_2\, dv, \quad ds_w = e_3\, dw.$$

For a general element of length ds we have

$$ds^2 = (ds_u)^2 + (ds_v)^2 + (ds_w)^2 = e_1^2\, du^2 + e_2^2\, dv^2 + e_3^2\, dw^2.$$

In rectangular, cylindrical, and spherical coordinates we have

$$ds_x = dx, \quad ds_y = dy, \quad ds_z = dz; \qquad ds^2 = dx^2 + dy^2 + dz^2;$$
$$ds_\rho = d\rho, \quad ds_\varphi = \rho\, d\varphi, \quad ds_z = dz; \qquad ds^2 = d\rho^2 + \rho^2\, d\varphi^2 + dz^2;$$
$$ds_r = dr, \quad ds_\theta = r\, d\theta, \quad ds_\varphi = r \sin\theta\, d\varphi; \quad ds^2 = dr^2 + r^2\, d\theta^2 + r^2 \sin^2\theta\, d\varphi^2.$$

For elementary areas in the coordinate surfaces we have

$$dS_u = ds_v\, ds_w, \qquad dS_v = ds_w\, ds_u, \qquad dS_w = ds_u\, ds_v;$$
$$dS_x = dy\, dz, \qquad dS_y = dz\, dx, \qquad dS_z = dx\, dy;$$
$$dS_\rho = \rho\, d\varphi\, dz, \qquad dS_\varphi = d\rho\, dz, \qquad dS_z = \rho\, d\rho\, d\varphi;$$
$$dS_r = r^2 \sin\theta\, d\theta\, d\varphi, \qquad dS_\theta = r \sin\theta\, dr\, d\varphi, \qquad dS_\varphi = r\, dr\, d\theta.$$

And, finally, the volume of an elementary coordinate cell is

$$d\tau = ds_u\,ds_v\,ds_w = dx\,dy\,dz = \rho\,d\rho\,d\varphi\,dz = r^2 \sin\theta\,dr\,d\theta\,d\varphi.$$

1.6. *Differential Expressions for Gradient, Divergence, Curl*

The components of the gradient of a point function V in the directions of the coordinate lines are the directional derivatives in the u-, v-, w-directions, that is, dV/ds_u, dV/ds_v, dV/ds_w. Substituting the values of the differential distances in the three coordinate systems as defined above, we have

$$\text{grad}_x V = \frac{\partial V}{\partial x}, \quad \text{grad}_y V = \frac{\partial V}{\partial y}, \quad \text{grad}_z V = \frac{\partial V}{\partial z};$$

$$\text{grad}_\rho V = \frac{\partial V}{\partial \rho}, \quad \text{grad}_\varphi V = \frac{\partial V}{\rho\,\partial \varphi}, \quad \text{grad}_z V = \frac{\partial V}{\partial z};$$

$$\text{grad}_r V = \frac{\partial V}{\partial r}, \quad \text{grad}_\theta V = \frac{\partial V}{r\,\partial \theta}, \quad \text{grad}_\varphi V = \frac{1}{r \sin\theta}\frac{\partial V}{\partial \varphi};$$

$$\text{grad}_u V = \frac{1}{e_1}\frac{\partial V}{\partial u}, \quad \text{grad}_v V = \frac{1}{e_2}\frac{\partial V}{\partial v}, \quad \text{grad}_w V = \frac{1}{e_3}\frac{\partial V}{\partial w}.$$

In order to calculate the divergence of a vector F at a point P we take an elementary cell about P, determine the flux of F through the surface of this cell, and divide it by its volume. The area of a u-surface through P, intercepted by the cell, is dS_u and the flux of F across this area is $F_u\,dS_u$. Since the rate of change of this flux in the u-direction is $D_u(F_u\,dS_u)$, the residual flux across the u-faces of the cell will be $D_u(F_u\,ds_u)\,du$. Similarly, we calculate the residual fluxes through the v-faces and the w-faces and obtain the value $D_u(F_u\,dS_u)\,du + D_v(F_v\,dS_v)\,dv + D_w(F_w\,dS_w)\,dw$ for the total outward flux of F from the cell. Substituting for the elementary areas their values in the various coordinate systems and dividing by the corresponding elementary volumes, we obtain

$$\text{div}\,F = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z},$$

$$\text{div}\,F = \frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho F_\rho) + \frac{1}{\rho}\frac{\partial F_\varphi}{\partial \varphi} + \frac{\partial F_z}{\partial z},$$

$$\text{div}\,F = \frac{1}{r^2 \sin\theta}\left[\sin\theta \frac{\partial}{\partial r}(r^2 F_r) + r\frac{\partial}{\partial \theta}(\sin\theta\,F_\theta) + r\frac{\partial F_\varphi}{\partial \varphi}\right],$$

$$\text{div}\,F = \frac{1}{e_1 e_2 e_3}\left[\frac{\partial}{\partial u}(e_2 e_3 F_u) + \frac{\partial}{\partial v}(e_1 e_3 F_v) + \frac{\partial}{\partial w}(e_1 e_2 F_w)\right].$$

By definition, the component $\text{curl}_u F$ of curl F in the u-direction is the circulation of F per unit area in the u-surface passing through P (Fig. 1.13). If we picture F as a mechanical force, $\text{curl}_u F$ is the work done by F per unit area in the u-surface. Consider an elementary area about P bounded by v- and w-lines. The work done by F along the w-line through P is $F_w\,ds_w$, its rate of change in the v-direction $D_v(F_w\,ds_w)$, and the total work

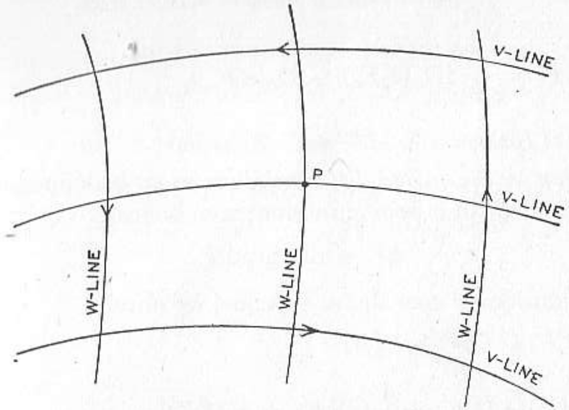


FIG. 1.13. Illustrating the derivation of the curl of a vector.

in the counter-clockwise direction along the w-paths of the loop bounding the elementary area is $D_v(F_w\,ds_w)\,dv$. Similarly, the work done along the remaining two sides of the loop is $D_w(F_v\,ds_v)\,dw$ in the clockwise direction. When the total work round the loop $D_v(F_w\,ds_w)\,dv - D_w(F_v\,ds_v)\,dw$ is divided by dS_u, the area enclosed by the loop, we have $\text{curl}_u F$. The remaining components are obtained by the cyclic permutation of u, v, w. Substituting the corresponding expressions for the differential elements and differential areas in the various coordinate systems, we obtain

$$\text{curl}_x F = \frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}, \quad \text{curl}_\rho F = \frac{1}{\rho}\frac{\partial F_z}{\partial \varphi} - \frac{\partial F_\varphi}{\partial z},$$

$$\text{curl}_y F = \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}, \quad \text{curl}_\varphi F = \frac{\partial F_\rho}{\partial z} - \frac{\partial F_z}{\partial \rho},$$

$$\text{curl}_z F = \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}; \quad \text{curl}_z F = \frac{1}{\rho}\left[\frac{\partial}{\partial \rho}(\rho F_\varphi) - \frac{\partial F_\rho}{\partial \varphi}\right];$$

$$\text{curl}_r F = \frac{1}{r \sin\theta}[D_\theta(\sin\theta\,F_\varphi) - D_\varphi(F_\theta)],$$

$$\text{curl}_\theta F = \frac{1}{r \sin\theta}[D_\varphi(F_r) - \sin\theta\,D_r(rF_\varphi)],$$

$$\text{curl}_\varphi F = \frac{1}{r}[D_r(rF_\theta) - D_\theta(F_r)];$$

$$\text{curl}_u F = \frac{1}{e_2e_3}[D_v(e_3F_w) - D_w(e_2F_v)],$$

$$\text{curl}_v F = \frac{1}{e_3e_1}[D_w(e_1F_u) - D_u(e_3F_w)],$$

$$\text{curl}_w F = \frac{1}{e_1e_2}[D_u(e_2F_v) - D_v(e_1F_u)].$$

1.7. *Differential Invariants and Green's Theorems*

The *Laplacian* or the *second differential invariant* is defined as the divergence of the gradient of a point function; symbolically

$$\Delta V = \text{div grad } V.$$

In the above considered coordinate systems, we obtain

$$\Delta V = D_x^2V + D_y^2V + D_z^2V,$$

$$\Delta V = \frac{1}{\rho}[D_\rho(\rho\, D_\rho V) + \frac{1}{\rho}(D_\varphi^2 V) + \rho(D_z^2 V)],$$

$$\Delta V = \frac{1}{r^2 \sin\theta}[\sin\theta\, D_r(r^2\, D_r V) + D_\theta(\sin\theta\, D_\theta V) + \frac{1}{\sin\theta}D_\varphi^2 V],$$

$$\Delta V = \frac{1}{e_1e_2e_3}\left[D_u\left(\frac{e_2e_3}{e_1}D_uV\right) + D_v\left(\frac{e_3e_1}{e_2}D_vV\right) + D_w\left(\frac{e_1e_2}{e_3}D_wV\right)\right]$$

The Laplacian of a vector F is the vector whose cartesian components are the Laplacians of the cartesian components of F.

The *first differential invariant* is defined as the scalar product of the gradient of two point functions; symbolically

$$\Delta(U,V) = (\text{grad } U, \text{grad } V) = D_xU\, D_xV + D_yU\, D_yV + D_zU\, D_zV.$$

Green's theorems are, then, expressed by the following equations:

$$\iiint \Delta(U,V)\, dv = \iint U\frac{\partial V}{\partial n}\, dS - \iiint U\,\Delta V\, dv, \quad (7\text{–}1)$$

and

$$\iiint (U\,\Delta V - V\,\Delta U)\, dv = \iint\left(U\frac{\partial V}{\partial n} - V\frac{\partial U}{\partial n}\right) dS, \quad (7\text{–}2)$$

where the surface integration is extended over the boundary of the volume and the normal derivatives are taken along the *outward* normals. The

second theorem is the consequence of the first: if U and V are interchanged in (1) and the result is subtracted from the original equation, (2) will follow. Equation (1) is proved by integrating the left hand side by parts.

1.8. *Miscellaneous Equations*

$$\text{div curl } F = 0, \quad \text{curl grad } V = 0, \quad (8\text{–}1)$$

$$\text{curl curl } F = \text{grad div } F - \Delta F, \quad (8\text{–}2)$$

$$\text{div } VF = V \text{ div } F + F\cdot\text{grad } V, \quad (8\text{–}3)$$

$$\text{curl } VF = V \text{ curl } F - F \times \text{grad } V, \quad (8\text{–}4)$$

$$\text{div } [F \times G] = G\cdot\text{curl } F - F\cdot\text{curl } G, \quad (8\text{–}5)$$

$$F = -\frac{1}{4\pi}\text{grad}\iiint\frac{\text{div } F}{R}\, dv + \frac{1}{4\pi}\text{curl}\iiint\frac{\text{curl } F}{R}\, dv. \quad (8\text{–}6)$$

CHAPTER II

Mathematics of Oscillations and Waves

2.1. *Complex Variables*

A *complex number*

$$z = x + iy$$

is a combination of real numbers x and y and an " *imaginary* " unit i subject to the following condition:

$$i^2 = -1.$$

The quantities x and y are called the *real* and *imaginary* parts of z; thus we write

$$x = \text{re}(z), \quad y = \text{im}(z).$$

Complex numbers are represented graphically by points in a plane (Fig. 2.1) or by vectors drawn from the origin to these points. Complex

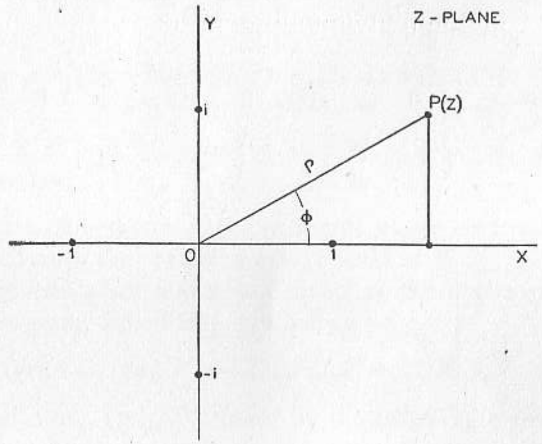


Fig. 2.1. Representation of a complex variable by points in a plane.

numbers obey the same arithmetical laws as real numbers. In the complex plane the addition and subtraction of complex numbers correspond to the addition and subtraction of vectors (Figs. 2.2 and 2.3).

In polar coordinates we have (Fig. 2.1)

$$z = \rho(\cos \varphi + i \sin \varphi);$$

ρ is called the *absolute value* or the *modulus* or the *amplitude* of z and φ is the *phase* of z. Thus we write

$$\rho = |z| = \text{am}(z), \quad \varphi = \text{ph}(z).$$

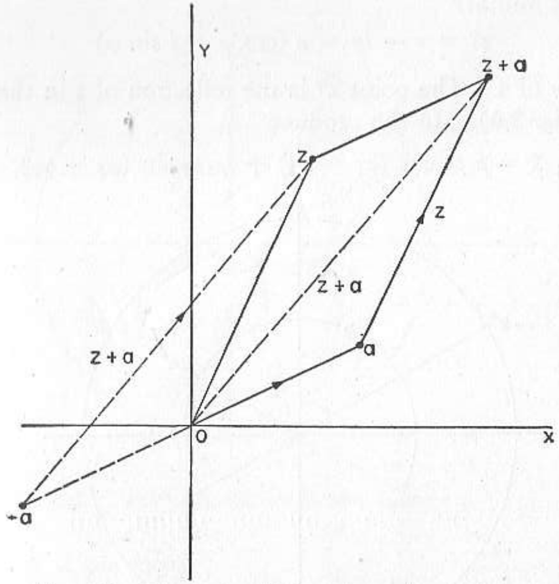


Fig. 2.2. Geometric addition of complex numbers.

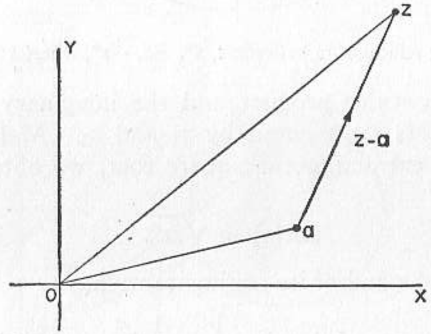


Fig. 2.3. Geometric representation of the difference between two complex numbers.

In polar form multiplication and division are particularly simple; thus

$$z_1 z_2 = \rho_1 \rho_2 [\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)],$$

$$\frac{z_1}{z_2} = \frac{\rho_1}{\rho_2} [\cos(\varphi_1 - \varphi_2) + i \sin(\varphi_1 - \varphi_2)].$$

The graphic interpretation of these equations is simple. When z_1 is multiplied by z_2, the vector represented by z_1 is stretched in the ratio $\rho_2 : 1$ and turned counterclockwise through the angle φ_2; and when z_1 is divided by z_2, it is shortened in the ratio $\rho_2 : 1$ and turned clockwise through the angle φ_2.

The complex number

$$z^* = x - iy = \rho\,(\cos\varphi - i\sin\varphi)$$

is the *conjugate* of z. The point z^* is the reflection of z in the real axis as in a mirror (Fig. 2.4). In the product

$$z_1 z_2^* = \rho_1\rho_2 \cos(\varphi_1 - \varphi_2) + i\rho_1\rho_2 \sin(\varphi_1 - \varphi_2),$$

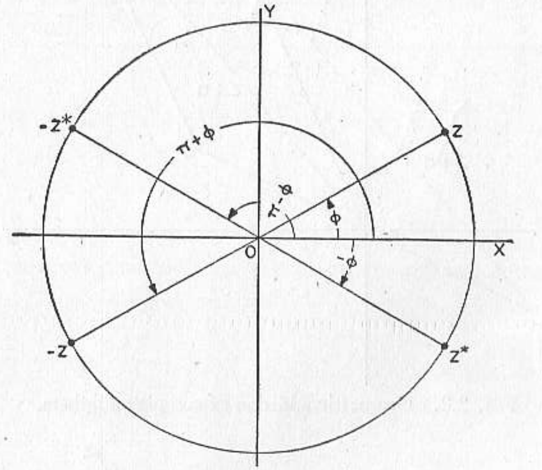


FIG. 2.4. Illustrating the relationship between z, z^*, $-z$, $-z^*$, where z^* is the conjugate of z.

the real part is the scalar product and the imaginary part the vector product of the vectors represented by z_1 and z_2. Making the complex numbers equal and extracting the square root, we obtain the following useful formula

$$\text{am}(z) = \sqrt{zz^*}.$$

For the reciprocals of z and of its conjugate we have

$$\frac{1}{z} = \frac{1}{\rho}(\cos\varphi - i\sin\varphi), \quad \frac{1}{z^*} = \frac{1}{\rho}(\cos\varphi + i\sin\varphi).$$

Points z and $1/z^*$ are said to be *inverse* with respect to the unit circle (Fig. 2.5). Points on the unit circle represent *unit complex numbers*.

Equations of curves can be written in terms of complex variables. Thus for a circle of radius r with its center at point a (Fig. 2.6) we have

$$(z - a)(z^* - a^*) = r^2.$$

This equation is of the following form:

$$zz^* + Az + A^*z^* + BB^* = 0, \tag{1-1}$$

where

$$a = -A^*, \quad r = \sqrt{AA^* - BB^*}.$$

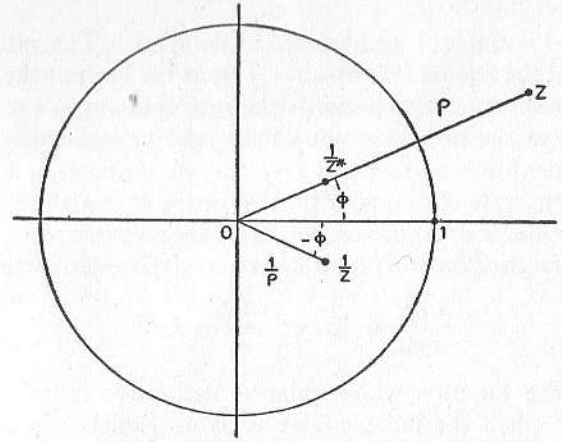


FIG. 2.5. Illustrating the relationship between z, $\frac{1}{z}$, $\frac{1}{z^*}$.

Linear equations of the form

$$Az + A^*z^* + BB^* = 0$$

represent straight lines.

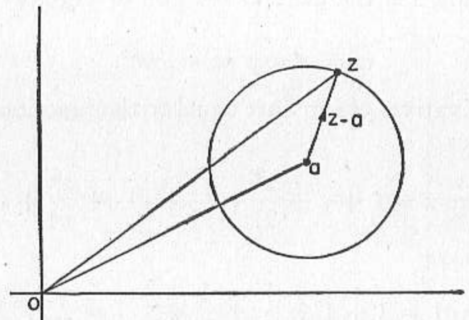


FIG. 2.6. In connection with the equation of a circle.

A variable w is a bilinear function of z and vice versa if

$$w = \frac{az + b}{cz + d}, \quad z = \frac{-dw + b}{cw - a}.$$

In transmission theory we shall use the following theorem: *If z describes a circle in the z-plane (including straight lines as special cases), then* w *also describes a circle in the* w-*plane.* This theorem is proved by substituting for z in (1) and showing that the resulting equation in w is of the same form.

2.2. *Exponential Functions*

Consider two variables z and w, real or complex. The ratios $\Delta z/z$ and $\Delta w/w$ are called the *relative increments*. Just as the limit of the ratio $\Delta w/\Delta z$ of two absolute increments represents the rate of change of w with respect to z, the limit of the ratio $\Delta w/w\Delta z$ of the relative increment in w to the absolute increment in z represents the *relative rate of change* of w with respect to z. The former rate of change is the derivative of w with respect to z and the latter the *relative derivative* or the *logarithmic derivative.*

An *exponential function* is a function whose relative derivative is constant

$$\frac{1}{w}\frac{dw}{dz} = k, \quad \text{or} \quad \frac{dw}{dz} = kw. \tag{2–1}$$

In particular the function whose relative derivative is unity and which becomes unity when the independent variable vanishes is designated as follows:*

$$w = \exp z = e^z;$$

thus by definition

$$\frac{d}{dz}\exp z = \exp z, \quad \exp 0 = 1.$$

In terms of this function the general solution of (1) may be expressed as follows:

$$w = A \exp kz = Ae^{kz}.$$

Since all the derivatives of exp z are equal to the function itself, we obtain from Taylor's series

$$e^z = \exp z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots + \frac{z^n}{n!} + \cdots.$$

In particular we have

$$e = \exp 1 = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots = 2.71828\cdots.$$

Either from (1) or by multiplying the power series the following addition theorem may be obtained:

$$\exp(z_1 + z_2) = (\exp z_1)(\exp z_2).$$

* The second notation anticipates some properties of exponential functions.

It is evident from the definition that all derivatives of an exponential function are proportional to the function itself. Hence a general linear differential equation with constant coefficients

$$a_n \frac{d^n w}{dz^n} + a_{n-1}\frac{d^{n-1}w}{dz^{n-1}} + \cdots + a_0 = ae^{kz} \tag{2–2}$$

will possess a particular solution of the following form:

$$w = be^{kz}. \tag{2–3}$$

The constant b is obtained by substituting in (2); thus

$$b = \frac{a}{Z(k)}, \quad Z(k) = a_n k^n + a_{n-1}k^{n-1} + \cdots + a_0. \tag{2–4}$$

A solution of this form exists for any value of k which is not a zero of $Z(k)$. On the other hand if

$$Z(\hat{k}_m) = 0, \tag{2–5}$$

then (2) possesses the following solution

$$w = b_m e^{\hat{k}_m z} \tag{2–6}$$

when $a = 0$. The most general solution of (2) when $a \neq 0$ is then

$$w = \frac{a}{Z(k)}e^{kz} + \sum b_m e^{\hat{k}_m z}. \tag{2–7}$$

2.3. *Exponential and Harmonic Oscillations*

Let the position of a point $P(z)$ in the complex plane be an exponential function of time; thus

$$z = Ae^{pt}, \tag{3–1}$$

where A and p are complex constants

$$A = ae^{i\varphi_0}, \quad p = \xi + i\omega; \tag{3–2}$$

then

$$z = ae^{\xi t}e^{i(\omega t + \varphi_0)}. \tag{3–3}$$

Thus the vector OP, drawn from the origin O to the point z, revolves about O with a constant angular velocity ω and its length varies exponentially with time

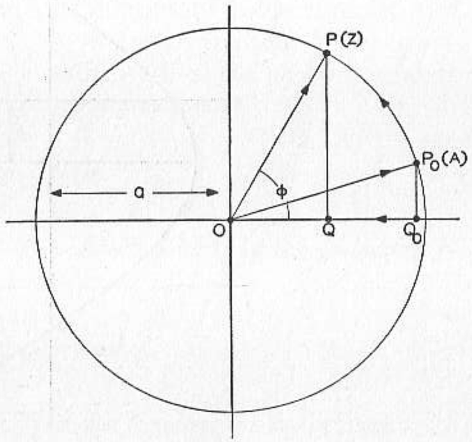


FIG. 2.7. Uniform rotation and harmonic oscillations.

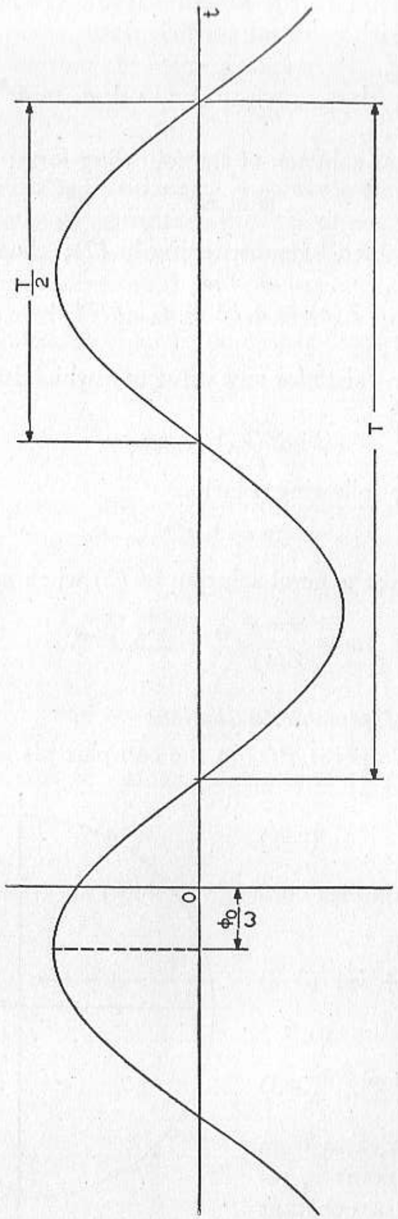


FIG. 2.8. The period T of harmonic oscillations.

If $\xi = 0$, the length of the vector OP remains constant and the point P moves along the circle of radius a with a uniform speed (Fig. 2.7). When $\omega > 0$, the movement is counterclockwise. The projection Q of the point P on the real axis is said to oscillate *harmonically* about O. The distance of Q from the center of the oscillation is a sinusoidal function of time (Fig. 2.8)

$$x = \text{re}(z) = a \cos(\omega t + \varphi_0).$$

The quantity

$$\varphi = \omega t + \varphi_0$$

is called the *phase of the oscillation* and a is the *amplitude of the oscillation*. The constant φ_0 is the *initial phase*.

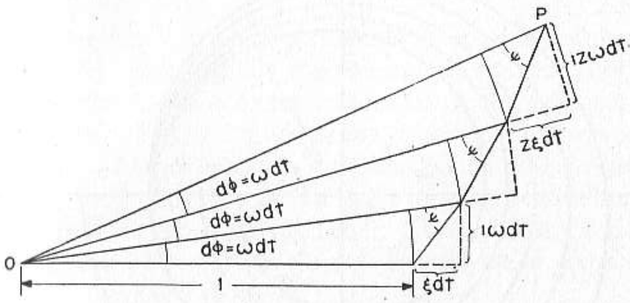


FIG. 2.9. Constant relative increments in the complex plane.

Two phases which differ by an integral multiple of 2π are regarded as the same because the points P and Q occupy the same positions. The interval T between two successive co-phase instants is called the *period of revolution* of P and the *period of oscillation* of Q; this period is

$$\omega T = 2\pi, \quad T = \frac{2\pi}{\omega}. \tag{3–4}$$

The number of revolutions of P or oscillations of Q is the *frequency* f in cycles per second; thus

$$f = \frac{1}{T}, \quad \omega = 2\pi f. \tag{3–5}$$

The constant ω is the angular velocity or the *frequency in radians per second*.

If $\xi \neq 0$, we have from (3)

$$d\rho = \rho\xi\, dt, \quad d\varphi = \omega\, dt,$$

and, therefore,

$$d\rho = \frac{\xi}{\omega} \rho \, d\varphi.$$

Hence the angle ψ between the radius and the trajectory (Fig. 2.9) is obtained from

$$\tan \psi = \frac{\omega}{\xi}, \quad \text{or} \quad \cot \psi = \frac{\xi}{\omega}.$$

The trajectories of point $P(z)$ are, thus, equiangular or logarithmic spirals; Fig. 2.10 shows several such spirals for different values of ξ/ω. In this case

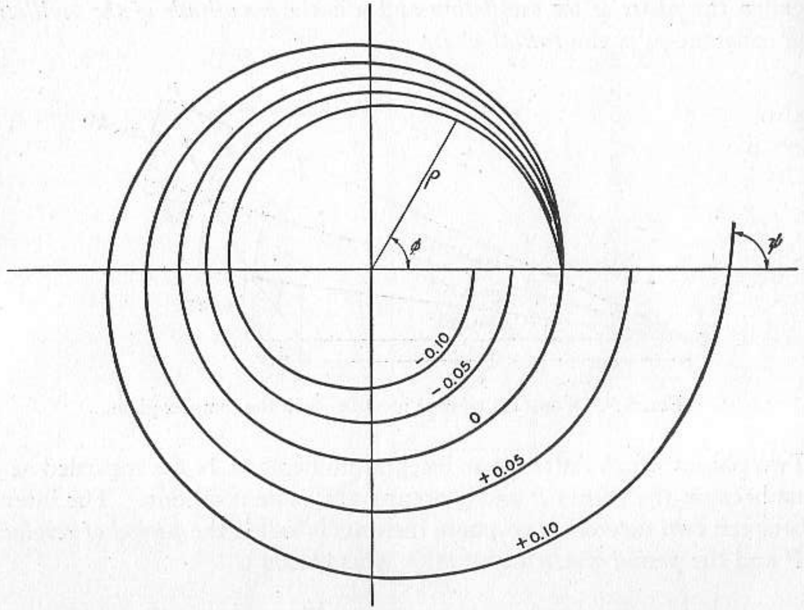


FIG. 2.10. Logarithmic spirals illustrating exponential functions of time; $z = e^{(\xi+i\omega)t} = e^{(\xi/\omega)\varphi+i\varphi}$; $\rho = e^{(\xi/\omega)\varphi}$. The so-called "$Q$" associated with oscillations may be defined as the magnitude of $\omega/2\xi$.

the distance $OQ = x$ is a sinusoidal function with an exponentially varying amplitude

$$x = \text{re}(z) = ae^{\xi t} \cos(\omega t + \varphi_0). \tag{3-6}$$

The constant ξ may be called the *growth constant*. The growth constant per cycle ξ/f is the *logarithmic increment* (or decrement).

Since there is one-to-one correspondence between points $P(z)$ moving in accordance with equation (1) and their projections on the real axis, exponential oscillations defined by equation (6) may be represented symbolically by complex exponentials of the form (1). The complex constant p is called the *oscillation constant*. The constant A gives simultaneous information about the initial amplitude and phase.

2.4. *Waves*

A *wave function* is a function of coordinates and time. For instance the function

$$V(x,t) = Ae^{-\Gamma x+pt}, \tag{4-1}$$

where A, Γ, p are complex constants, is a wave function. Let

$$A = ae^{i\varphi_0}, \quad p = \xi + i\omega, \quad \Gamma = \alpha + i\beta; \tag{4-2}$$

the real part of V, given by

$$\text{re}(V) = ae^{-\alpha x+\xi t} \cos(\omega t - \beta x + \varphi_0), \tag{4-3}$$

is also a wave function. At any point x this function is a sinusoidal function* of time, with exponentially varying amplitude; at any instant t, $\text{re}(V)$ is a sinusoidal function of the coordinate x, also with exponentially varying amplitude. Physical phenomena expressed by wave functions are called *waves*. As there is one-to-one correspondence between exponential functions of the form (1) and sinusoidal functions of the form (3), we may use the former to represent the latter. The constant Γ is called the *propagation constant*; its real part α is the *attenuation constant* and its imaginary part β the *phase constant*.

The quantity

$$\Phi = \omega t - \beta x + \varphi_0 \tag{4-4}$$

is called the *phase of the wave*. The distance λ between two successive equiphase points (at the same instant) is called the *wavelength*; thus

$$\beta\lambda = 2\pi, \quad \lambda = \frac{2\pi}{\beta}, \quad \beta = \frac{2\pi}{\lambda}. \tag{4-5}$$

The wavelength is analogous to the period.

The velocity v with which an observer should move in order to observe no change in phase is called the *velocity of the wave* or the *wave velocity* or the *phase velocity*. This velocity is obtained from the following condition

$$d\Phi = \omega \, dt - \beta \, dx = 0;$$

thus

$$v = \frac{dx}{dt} = \frac{\omega}{\beta}, \quad \beta = \frac{\omega}{v}. \tag{4-6}$$

* Except, of course, when $\omega = 0$.

From (5), (6), and (3–5) we obtain

$$f\lambda = v. \tag{4–7}$$

Consider now a general three-dimensional harmonic wave function

$$V = A(x,y,z)e^{-i\Phi(x,y,z)+i\omega t}$$

in which A and Φ are two real functions. The surfaces of equal phase (at the same instant), given by

$$\Phi(x,y,z) = \text{constant},$$

are called *equiphase surfaces.** The waves represented by V are called *plane*, *cylindrical*, *spherical*, etc., if the equiphase surfaces are plane, cylindrical, spherical, etc.

For any pair of infinitely close points in an equiphase surface, we have

$$\frac{\partial\Phi}{\partial x}dx + \frac{\partial\Phi}{\partial y}dy + \frac{\partial\Phi}{\partial z}dz = 0.$$

Replacing the differentials dx, dy, dz in this equation by $(X - x)$, $(Y - y)$, $(Z - z)$ where (x,y,z) is a point in the equiphase surface and (X,Y,Z) is a typical point in space, we obtain the equation of the tangent plane

$$\frac{\partial\Phi}{\partial x}(X - x) + \frac{\partial\Phi}{\partial y}(Y - y) + \frac{\partial\Phi}{\partial z}(Z - z) = 0.$$

The family of curves everywhere normal to the equiphase surfaces is given by the simultaneous differential equations

$$\frac{dx}{\dfrac{\partial\Phi}{\partial x}} = \frac{dy}{\dfrac{\partial\Phi}{\partial y}} = \frac{dz}{\dfrac{\partial\Phi}{\partial z}}. \tag{4–8}$$

These curves are called *wave normals.* Equation (8) implies that any wave normal is tangential to grad Φ. Consequently, *wave normals are curves along which the phase changes most rapidly.*

At two points infinitely close in space-time the phases are the same if

$$\omega\, dt - \frac{\partial\Phi}{\partial x}dx - \frac{\partial\Phi}{\partial y}dy - \frac{\partial\Phi}{\partial z}dz = 0.$$

If $dy = dz = 0$, the instantaneous rate with which the phase changes along an x-line is $\partial\Phi/\partial x$. The " phase constants " along coordinate lines, which

* For purposes of this definition phases differing by an integral multiple of π are regarded as equal.

generally are not constants at all, are

$$\beta_x = \frac{\partial\Phi}{\partial x}, \quad \beta_y = \frac{\partial\Phi}{\partial y}, \quad \beta_z = \frac{\partial\Phi}{\partial z}.$$

The phase constant along a wave normal is

$$\beta(x,y,z) = |\,\text{grad}\,\Phi\,|.$$

In three dimensions the phase constant may be regarded as a vector whose direction components are $\partial\Phi/\partial x$, $\partial\Phi/\partial y$, $\partial\Phi/\partial z$.

The phase velocity is usually defined by

$$v(x,y,z) = \frac{\omega}{\beta} = \frac{\omega}{|\,\text{grad}\,\Phi\,|}.$$

This is the instantaneous velocity along a wave normal. We may also speak of the phase velocities along the coordinate axes

$$v_x = \frac{\omega}{\beta_x} = \frac{\omega}{\dfrac{\partial\Phi}{\partial x}}, \quad v_y = \frac{\omega}{\beta_y} = \frac{\omega}{\dfrac{\partial\Phi}{\partial y}}, \quad v_z = \frac{\omega}{\beta_z} = \frac{\omega}{\dfrac{\partial\Phi}{\partial z}};$$

It is evident, however, that v_x, v_y, v_z are not the cartesian components of the phase velocity along the wave normal. On the other hand, the reciprocals of these phase velocities, being proportional to the phase constants, behave as vector components should. Thus, we define *phase slowness* S

$$S = \frac{1}{\omega}\,\text{grad}\,\Phi.$$

2.5. *Nepers, Bels, Decibels*

The logarithmic measure has come into use because in certain measurements the logarithm of a ratio of two quantities is more significant than the ratio itself. When the ratio of two quantities of the same kind is expressed in nepers, the number of nepers is computed from

$$N = \log\frac{A_1}{A_2}\ \text{nepers}. \tag{5–1}$$

From this we have

$$A_1 = A_2e^N, \quad A_2 = A_1e^{-N}.$$

Originally the logarithmic unit was introduced for the evaluation of power ratios and present laboratory units are the *bel* and the *decibel.* The number of bels and decibels (abbreviated db), expressing a power ratio W_1/W_2, is computed from

$$N = \log_{10}\frac{W_1}{W_2}\ \text{bels} = 10\log_{10}\frac{W_1}{W_2}\ \text{decibels}.$$

More recently the use of logarithmic units has been extended to "intensity ratios," that is, to voltage and current ratios; then the number of bels and decibels has been defined as

$$N = 2 \log_{10} \frac{E_1}{E_2} \text{ bels} = 20 \log_{10} \frac{E_1}{E_2} \text{ decibels.}$$

It is unfortunate that the size of the "bel" or "decibel" is not uniform but depends upon the nature of the measured quantity. We shall keep the neper as a fixed unit defined by (1) regardless of the nature of A. Thus in translating from nepers into decibels, at least when dealing with electrical quantities, we must multiply by different conversion factors. For power ratios we have

$$1 \text{ neper} = 10 \log_{10} e \simeq 4.343 \text{ db};$$

and for voltage ratios, current ratios, field intensity ratios, we have

$$1 \text{ neper} = 20 \log_{10} e \simeq 8.686 \text{ db}.$$

In accordance with the above definitions the attenuation constant is measured in nepers per unit length or in decibels per unit length. The phase constant is measured in radians per unit length. Similarly the time growth constant is measured in nepers or decibels per unit time and the frequency ω in radians per unit time.

2.6. *Stationary Waves*

The waves discussed in section 4 are called *progressive waves*. The sum (or the difference) of two unattenuated progressive waves, of equal amplitude, moving in opposite directions is called a *stationary wave* because different points oscillate always either in phase or 180° out of phase. For instance,

$$ae^{i(\omega t-\beta x)} + ae^{i(\omega t+\beta x)} = 2a \cos \beta x \, e^{i\omega t}.$$

2.7. *Impedance Concept*

The relationship between two quantities oscillating with the same frequency is given by the ratio Z of their complex representations; am(Z) gives the amplitude ratio and ph(Z) the phase difference. Consider, for example, a box containing a linear passive* electric network or an *impedor* and a pair of accessible terminals A, B (Fig. 2.11). Let the instantaneous values V_i, I_i of the voltage applied across the terminals and the current flowing in response to this voltage be

$$V_i = \text{re}(Ve^{i\omega t}) = V_a \cos(\omega t + \varphi_V),$$
$$I_i = \text{re}(Ie^{i\omega t}) = I_a \cos(\omega t + \varphi_I), \qquad (7\text{–}1)$$

* No source of power.

where V_a, I_a are the voltage and current amplitudes and φ_V, φ_I are the initial phases; then the ratio

$$Z = \frac{Ve^{i\omega t}}{Ie^{i\omega t}} = \frac{V}{I} = \frac{V_a}{I_a} e^{i(\varphi_V - \varphi_I)}$$

expresses completely the relationship between the voltage and current. This ratio is called the *impedance* across the terminals.

If V and I represented a mechanical force and the corresponding velocity response, then Z would be called the mechanical impedance. The ratio of a mechanical force to an electric current or of a voltage to a velocity is an electromechanical impedance. Occasionally a name is desired for the voltage/charge and force/displacement ratios; such ratios may be called electrical and mechanical *impediments* respectively.

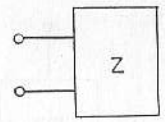


FIG. 2.11. A diagram representing an impedor.

The impedance is generally a complex number (Fig. 2.12); its real part is called the *resistance* and its imaginary part the *reactance*; thus

$$Z = R + iX.$$

For passive impedors Z is in the right half of the complex plane or on the imaginary axis. The reciprocal of the impedance is called the *admittance*

$$Y = \frac{1}{Z} = G + iB;$$

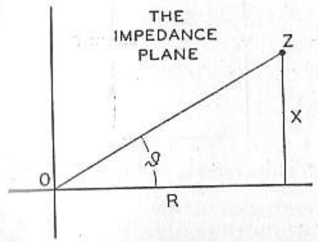


FIG. 2.12. Representation of impedances by points in a complex plane; R is the resistance component, X the reactance component, and ϑ the phase angle.

its real part is the *conductance* and its imaginary part the *susceptance*.

In diagrams it is often convenient to distinguish between numbers representing impedances and admittances; this can be accomplished by attaching the terminals to the sides of the rectangles representing corresponding impedors in the first case and to the vertices in the second (Fig. 2.13).

An impedor is a *resistor*, an *inductor*, or a *capacitor* according as the instantaneous voltage is proportional to the current, to its time derivative, or to its time integral

$$V_i = RI_i, \quad V_i = L\frac{dI_i}{dt}, \quad I_i = C\frac{dV_i}{dt};$$

the above coefficients of proportionality are called respectively the *resistance*, the *inductance*, the *capacitance*. In diagrams resistors, inductors, and capacitors are represented as shown in Fig. 2.14. For exponential and harmonic voltages and currents we have

$$V = RI, \quad V = pLI, \quad V = \frac{I}{pC},$$

$$V = RI, \quad V = i\omega LI, \quad V = \frac{I}{i\omega C}.$$

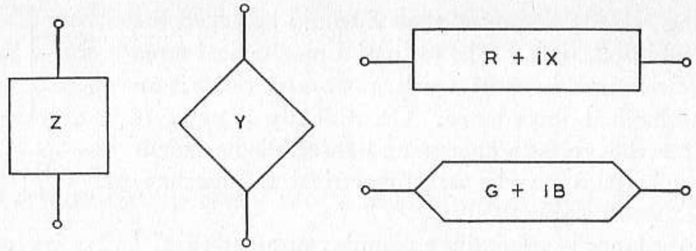


FIG. 2.13. Diagrammatic representation of impedances and admittances.

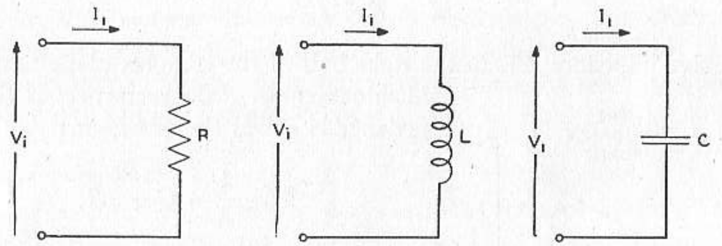


FIG. 2.14. Diagrammatic representation of resistances, inductances, and capacitances.

The impedance is a function of the frequency or more generally a function of the oscillation constant. To any exponential voltage Ve^{pt} there corresponds a finite response in the impedor (Fig. 2.11), given by

$$Ie^{pt} = \frac{Ve^{pt}}{Z(p)}, \tag{7-2}$$

provided p does not satisfy the following equation

$$Z(p) = 0. \tag{7-3}$$

Similarly, since

$$Ve^{pt} = Z(p)Ie^{pt}, \tag{7-4}$$

a definite exponential voltage exists across the terminals except when p is an infinity of the impedance function or a zero of the admittance function

$$Z(p) = \infty, \quad Y(p) = 0. \tag{7-5}$$

When p is a zero of the impedance function, then a finite exponential current may exist when $V = 0$, that is, when the terminals of the impedor are short-circuited; the zeros of $Z(p)$ are the *natural* oscillation constants of the impedor, with the terminals short-circuited. Similarly, if p is a pole of $Z(p)$, a finite voltage may exist if the terminals of the impedor are open; these poles determine the natural oscillation constants of the impedor, with the terminals open. The corresponding frequencies are the natural frequencies of the impedor. As we shall see, *natural oscillations* may be excited by an impulsive voltage. Natural oscillations may also occur when the applied voltage is exponential if this voltage is suddenly applied. These oscillations are also called *free oscillations* or *transient oscillations*. On the other hand, oscillations of the form (2) are often called *forced oscillations* or *steady-state oscillations*.

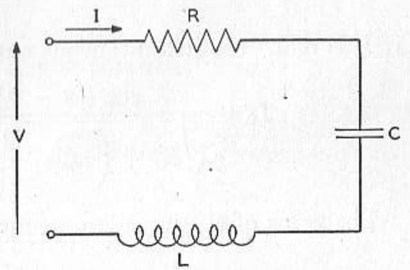


FIG. 2.15. A diagram showing an electric circuit comprised of a resistor, a capacitor, and an inductor connected in series.

Let us consider a special example of an impedor (or "electric circuit") consisting of a resistor, an inductor, and a capacitor in series (Fig. 2.15). The differential equation for this circuit is

$$L\frac{dI_i}{dt} + RI_i + \frac{\int I_i\,dt}{C} = V_i,$$

where V_i is the instantaneous value of the applied voltage and I_i is the corresponding current. When V_i and I_i are exponential functions of time and are represented by complex exponentials, the latter will also satisfy the above equation since the real part of the time derivative of a complex variable is the derivative of the real part. Thus we have

$$L\frac{d}{dt}(Ie^{pt}) + R(Ie^{pt}) + \frac{\int Ie^{pt}\,dt}{C} = Ve^{pt}.$$

After differentiation and integration the time factor is canceled and the solution becomes

$$I = \frac{V}{Z(p)}, \quad Z(p) = R + pL + \frac{1}{pC}. \tag{7-6}$$

When the impressed voltage is harmonic of frequency ω, then $p = i\omega$ and

$$I = \frac{V}{R + i\left(\omega L - \dfrac{1}{\omega C}\right)}.$$

If V is real,* the instantaneous current is

$$I_i = \frac{V \cos(\omega t - \vartheta)}{\sqrt{R^2 + \left(\omega L - \dfrac{1}{\omega C}\right)^2}}, \quad \vartheta = \tan^{-1}\frac{\omega^2 LC - 1}{\omega RC}.$$

The zeros of the impedance function (6) are

$$\hat{p}_{1,2} = -\frac{R}{2L} \pm \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}}.$$

These values are either real or complex depending upon whether R is greater than or less than $2\sqrt{L/C}$. In the latter case we have

$$\hat{p}_{1,2} = \hat{\xi} \pm i\hat{\omega}, \quad \hat{\xi} = -\frac{R}{2L}, \quad \hat{\omega} = \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}},$$

where $\hat{\omega}$ is the natural frequency.

In any physical problem the coefficients of the differential equation from which the impedance function is obtained are real; therefore $Z(p)$ has real coefficients and

$$Z(p^*) = Z^*(p). \tag{7–7}$$

Thus if p is a zero or a pole of the impedance function, p^* is also a zero or a pole. If

$$Z(i\omega) = R(\omega) + iX(\omega), \tag{7–8}$$

then by (7)

$$Z(-i\omega) = R(\omega) - iX(\omega).$$

On the other hand, replacing ω by $-\omega$ in (8), we have

$$Z(-i\omega) = R(-\omega) + iX(-\omega).$$

Comparing the last two equations, we find that the resistance function is an even function of the frequency and the reactance function is an odd function; thus

$$R(-\omega) = R(\omega), \quad X(-\omega) = -X(\omega).$$

* We can make it real by properly choosing the origin of time.

The impedance of a finite combination of resistors, inductors, and capacitors is a rational fraction; the difference between the degrees of the numerator and the denominator is either unity or zero or negative unity. The zeros and poles of a nondissipative impedor are simple; they lie on the imaginary axis and they separate each other. The zeros and poles of a passive dissipative impedor are always in the left half of the oscillation constant plane; they are not necessarily simple but usually so.

2.8. *Average Power and Complex Power*

The work performed by an applied voltage driving an electric current through an impedor is

$$\begin{aligned}\mathcal{E} &= \int_0^t V_i I_i\,dt = V_a I_a \int_0^t \cos(\omega t + \varphi_V)\cos(\omega t + \varphi_I)\,dt \\ &= \tfrac{1}{2}V_a I_a \int_0^t [\cos(\varphi_V - \varphi_I) + \cos(2\omega t + \varphi_V + \varphi_I)]\,dt \\ &= \tfrac{1}{2}V_a I_a t \cos(\varphi_V - \varphi_I) + \frac{1}{4\omega} V_a I_a \sin(2\omega t + \varphi_V + \varphi_I)\Big|_0^t.\end{aligned}$$

The second term is an oscillating function of time; hence the average power W contributed by the generator to the impedor over a long interval of time is

$$W = \tfrac{1}{2}V_a I_a \cos(\varphi_V - \varphi_I) = \text{re}\,\Psi, \tag{8–1}$$

where Ψ is the *complex power* defined by

$$\Psi = \tfrac{1}{2}VI^*. \tag{8–2}$$

Introducing the impedance and admittance we have

$$\Psi = \tfrac{1}{2}ZII^* = \frac{VV^*}{2Z^*} = \tfrac{1}{2}Y^*VV^* = \frac{II^*}{2Y}; \tag{8–3}$$

hence,

$$Z = \frac{2\Psi}{II^*} = \frac{VV^*}{2\Psi^*}, \quad Y = \frac{2\Psi^*}{VV^*} = \frac{II^*}{2\Psi}. \tag{8–4}$$

2.9. *Step and Impulse Functions**

Three functions are particularly important in wave theory: the *sinusoidal* function, or more generally the exponential function; the *step function* (Fig. 2.16) and the *impulse function* (Fig. 2.17). In the last case it is frequently assumed that the extent τ of the step is infinitely small and the step itself is infinitely large, while the *strength of the impulse*, represented by the area under the step is finite. The independent variable is usually either time or distance. A step function is called a *unit*

* Most of the contents of this and following sections are needed only in Chapter 10 and thus may be omitted on the first reading.

step if the sudden rise is from zero to unity. An impulse function is a *unit impulse* (or a *unit source*) if its strength is unity.

These functions are important in their own right; besides, by superposing either a finite or an infinite number of them one can obtain any function that may be met in practice. It is easy to see that this is so with impulse and step functions (Fig. 2.18);

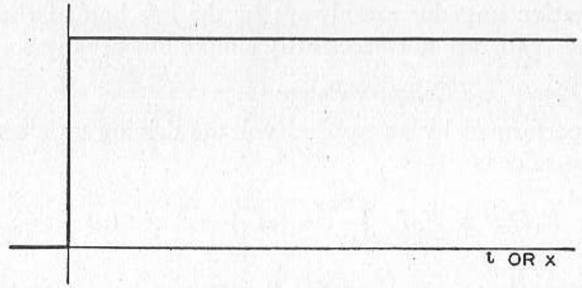


FIG. 2.16. A step function.

it requires some analysis to show that a function can be expanded either in a " Fourier series " or a " Fourier integral," representing addition of sinusoidal functions.

In order to obtain the response of a linear system to an almost arbitrary force we need only find its response to any one of the three above mentioned standard functions

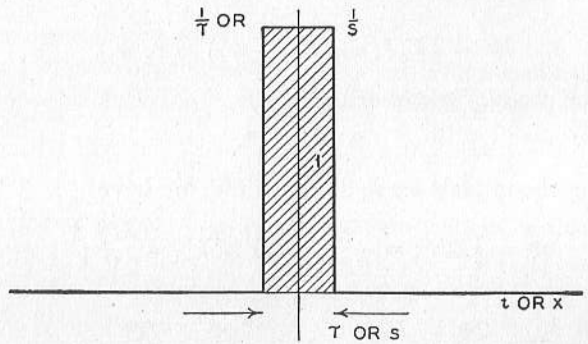


FIG. 2.17. An impulse function.

and integrate the result. Naturally, the principal requirement is that the integrals be convergent. For example, to find the electric current $I(t)$ through the terminals of a given impedor due to an electromotive force $V(t)$ applied across the terminals, we write

$$I(t) = \int_{-\infty}^{\infty} V(\hat{t})F(t,\hat{t})\,d\hat{t}, \tag{9–1}$$

where $F(t,\hat{t})$ is the current at time t in response to a unit impulse of electromotive force at time $\hat{t}$. In the present case we know *a priori* some properties of $F(t,\hat{t})$. This function is identically zero for $t < \hat{t}$ since the electromotive force is not retroactive; for

$t > \hat{t}$ its value depends only on $(t - \hat{t})$. Let $F(t,0) = F(t)$; then $F(t,\hat{t}) = F(t - \hat{t})$ and

$$I(t) = \int_{-\infty}^{\infty} V(\hat{t})F(t - \hat{t})\,d\hat{t} = \int_{-\infty}^{t} V(\hat{t})F(t - \hat{t})\,dt. \tag{9–2}$$

In potential theory, the function corresponding to $F(t,\hat{t})$ is called *Green's function*; we may apply this name to all responses to unit impulses.

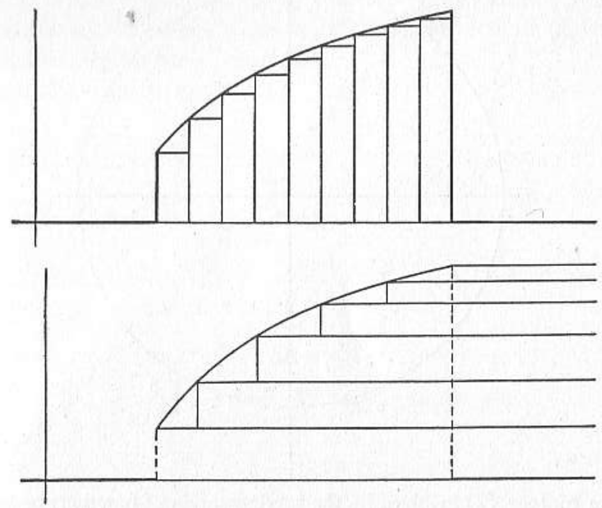


FIG. 2.18. Representation of arbitrary functions by superposition of impulse functions and step functions.

If $A(t)$ is the current in the network in response to a unit voltage step at $t = t_1$, then

$$I(t) = V(t_1)A(t - t_1) + \int_{t_1}^{t'} V'(\hat{t})A(t - \hat{t})\,d\hat{t}. \tag{9–3}$$

Here we have assumed that prior to $t = t_1$ the voltage is zero and that subsequently the voltage is a continuous function of time so that it changes in infinitely small steps, $V'(\hat{t})\,d\hat{t}$, where V' is the derivative. If there are finite discontinuities in $V(t)$, these must be taken care of in the same manner as the sudden change from zero to $V(t_1)$. For instance, if the electromotive force ceases to act at $t = t_2$, then

$$I(t) = V(t_1)A(t - t_1) + \int_{t_1}^{t_2} V'(\hat{t})A(t - \hat{t})\,d\hat{t} - V(t_2)A(t - t_2). \tag{9–4}$$

The upper limit of the integral can be t_2 just as well as t since $A(t)$ vanishes for negative values of t. More generally, we can replace (3) and (4) by

$$I(t) = \int_{-\infty}^{\infty} A(t - \hat{t})\,dV(\hat{t}), \tag{9–5}$$

provided we interpret this integral in the Stieltjes sense instead of the Riemann sense. This interpretation consists in taking the Riemann integral and adding to it $[V(\tau+0) - V(\tau-0)]A(t-\tau)$ at each point $t = \tau$ where $V(t)$ is discontinuous. In this sense the differential dV is permitted to be finite as well as infinitesimal. In electric circuit theory $A(t)$ is called the *indicial admittance* of the network.

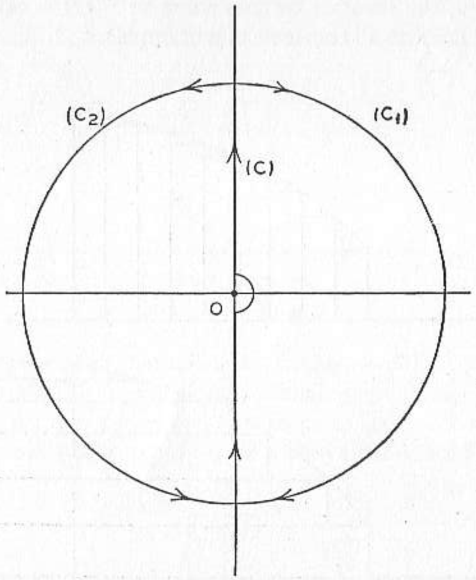


FIG. 2.19. The contour (C) involved in the representation of functions by contour integrals which can be interpreted as " sums of sinusoidal functions of infinitely small amplitudes."

We shall now express the unit step and the unit impulse functions as contour integrals. Consider the following integral

$$\frac{1}{2\pi i}\int_{(C)} \frac{e^{p(t-\hat{t})}}{p}\,dp \tag{9-6}$$

in the complex p-plane (Fig. 2.19). The contour (C) is along the imaginary axis indented to the right at the origin. If $t < \hat{t}$, we can add to this contour a semicircle (C_1) of infinite radius without changing the value of the integral. On this semicircle the real part of p is positive and $e^{p(t-\hat{t})}/p$ vanishes exponentially when $-\pi/2 < \mathrm{ph}(p) < \pi/2$; hence the integral over (C_1) is zero except, perhaps, over the portions corresponding to the phase angles infinitely close to $\pi/2$ or $-\pi/2$. A closer study of the integral in these regions would show that their contributions vanish as the radius of the semicircle becomes infinite. The integrand is single-valued and has no poles within $(C + C_1)$; hence the integral (6) is zero for $t < \hat{t}$.

If $t > \hat{t}$, then we can add to (C) an infinite semicircle (C_2) in the left half of the plane, without changing the value of (6). Within this contour there is a simple pole at $p = 0$; since the residue is unity, the value of the integral is unity. Thus the integral (6) represents a unit step at $t = \hat{t}$.

By superposing two step functions, of magnitude $1/\tau$ and $-1/\tau$, the first beginning at $t = -\tau/2$ and the second at $t = \tau/2$, we find that the unit impulse function, centered at $t = 0$, is

$$\frac{1}{2\pi i}\int_{(C)} \frac{\sinh\frac{p\tau}{2}}{\frac{p\tau}{2}}\, e^{pt}\,dp, \quad \text{as} \quad \tau \to 0. \tag{9-7}$$

In this expression it is not permissible to let $\tau = 0$ since the resulting integral is divergent in the usual sense. At times, however, the integrals derived from (7) converge for $\tau = 0$ and the substitution is permissible.

If the unit impulse is spatial rather than temporal, we write (7) in the form

$$\frac{1}{2\pi i}\int_{(C)} \frac{\sinh\frac{\gamma s}{2}}{\frac{\gamma s}{2}}\, e^{\gamma x}\,d\gamma, \quad \text{as} \quad s \to 0. \tag{9-8}$$

If in (7) and (8) we replace t by $t - \hat{t}$ and x by $x - \hat{x}$, we shift the center of the impulse from the origin to a typical point

$$\frac{1}{2\pi i}\int_{(C)} \frac{\sinh\frac{p\tau}{2}}{\frac{p\tau}{2}}\, e^{p(t-\hat{t})}\,dp, \quad \text{as} \quad \tau \to 0;$$

$$\frac{1}{2\pi i}\int_{(C)} \frac{\sinh\frac{\gamma s}{2}}{\frac{\gamma s}{2}}\, e^{\gamma(x-\hat{x})}\,d\gamma, \quad \text{as} \quad s \to 0. \tag{9-9}$$

In the preceding equations p is pure imaginary on (C) except in the immediate vicinity of $p = 0$. Thus the unit step and the unit impulse have been represented by superimposing sinusoids of infinitely small amplitudes with frequencies ranging from $-\infty$ to $+\infty$. The above contour integrals can be turned into more conventional forms of " Fourier Integrals " depending only on positive frequencies; but the present form is, on the whole, more useful. The values of the integrals will not be changed if (C) is deformed into any other contour provided this deformation takes place in the finite part of the plane and no poles are crossed in the process.

We shall now represent an arbitrary function $f(t)$ which is equal to zero for $t < 0$, in the form

$$f(t) = \int_{(C)} S(p)e^{pt}\,dp, \qquad S(p) = \frac{1}{2\pi i}\int_0^\infty f(t)e^{-pt}\,dt. \tag{9-10}$$

The function $S(p)$ is called the *complex spectrum* or simply the *spectrum** of $f(t)$. The

* In mathematical books $S(p)$ is also called the Laplace transform of $f(t)$.

function $f(t)$ can be regarded as the limit of the sum of impulse functions of strength $f(\hat{t})\,\Delta\hat{t}$, as $\Delta\hat{t}$ approaches zero; thus

$$f(t) = \lim \sum_{\hat{t}=0}^{\infty} \frac{f(\hat{t})\,\Delta\hat{t}}{2\pi i} \int_{(C)} \frac{\sinh \frac{p\,\Delta\hat{t}}{2}}{\frac{p\,\Delta\hat{t}}{2}} e^{p(t-\hat{t})}\,dp, \quad \text{as} \quad \Delta\hat{t} \to 0,$$

$$= \lim \frac{1}{2\pi i} \int_{(C)} e^{pt}\,dp \sum_{\hat{t}=0}^{\infty} \frac{\sinh \frac{p\,\Delta\hat{t}}{2}}{\frac{p\,\Delta\hat{t}}{2}} f(\hat{t}) e^{-p\hat{t}}\,\Delta\hat{t}$$

$$= \frac{1}{2\pi i} \int_{(C)} e^{pt}\,dp \int_0^{\infty} f(\hat{t}) e^{-p\hat{t}}\,d\hat{t},$$

and we have (10).

It should be noted that if $f(t)$ is given by (10), then

$$f(t - t_1) = \int_{(C)} S(p) e^{p(t-t_1)}\,dp.$$

The spectrum of this function is

$$S_1(p) = S(p) e^{-pt_1}.$$

If now the electromotive force $V(t)$ impressed on an impedor is expressed as an exponential contour integral, the current through the terminals may be obtained by superposition of the responses to each elementary exponential electromotive force. Thus

$$V(t) = \int_{(C)} S(p) e^{pt}\,dp, \quad \text{then} \quad I(t) = \int_{(C)} \frac{S(p)}{Z(p)} e^{pt}\,dp. \tag{9–11}$$

In particular the responses to a unit voltage step beginning at $t = 0$ and to a unit impulse centered at $t = 0$, are respectively

$$A(t) = \frac{1}{2\pi i} \int_{(C)} \frac{e^{pt}}{pZ(p)}\,dp, \quad B(t) = \frac{1}{2\pi i} \int_{(C)} \frac{\sinh \frac{p\tau}{2}}{\frac{p\tau}{2} Z(p)} e^{pt}\,dp. \tag{9–12}$$

For circuits consisting of physical elements, the second integral converges as $\tau \to 0$ and consequently

$$B(t) = \frac{1}{2\pi i} \int_{(C)} \frac{e^{pt}}{Z(p)}\,dp. \tag{9–13}$$

However, in some applications it is necessary to deal with impulses of finite width rather than with idealized impulses of zero width.

Substituting for $f(t)$ in (10) the functions $A(t)$ and $B(t)$ from (12), we have

$$\frac{1}{pZ(p)} = \int_0^{\infty} A(t) e^{-pt}\,dt, \qquad \frac{1}{Z(p)} = \int_0^{\infty} B(t) e^{-pt}\,dt.$$

Since the response $I(t)$ is zero prior to the application of the electromotive force, the contour (C) should be to the right of the poles of the integrand. If some poles are on the imaginary axis, the contour is indented as shown in Fig. 2.20.

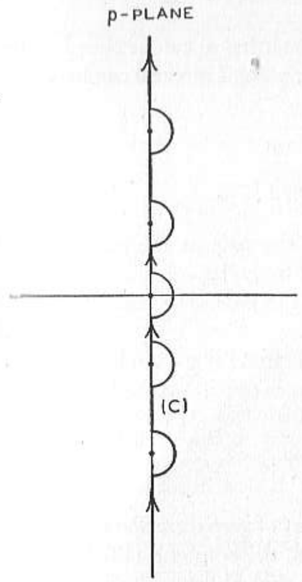


FIG. 2.20. Infinitely small indentations in the contour (C) when the poles of the integrand (the infinities of the admittance function) pass from the left half of the plane to the imaginary axis. This condition exists in non-dissipative (purely reactive) networks.

Consider now a special example. Take a resistor and an inductor in series (Fig. 2.21); the impedance function is $Z(p) = R + pL$. Hence the current $B(t)$ flowing in response to an infinitely short unit voltage impulse is

$$B(t) = \frac{1}{2\pi i} \int_{(C)} \frac{e^{pt}}{pL + R}\,dp.$$

When $t < 0$, this integral is automatically zero because of our choice of (C). When $t > 0$, we may

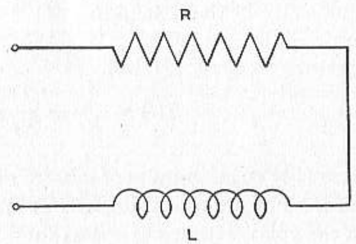


FIG. 2.21. A resistor and an inductor in series.

close (C) with an infinite semicircle in the left half of the plane. There is only one pole $p = -R/L$ within this contour and therefore

$$B(t) = \frac{1}{L} e^{-(R/L)t}. \tag{9–14}$$

If the zeros of $Z(p)$ are simple, we can obtain the current $B(t)$ in response to a unit voltage impulse in a simple form. In the vicinity of a typical zero we have $Z(p) = (p - p_m)Z'(p_m)$; hence the residue of the admittance is $Z'(p_m)$ and (13) becomes

$$B(t) = \sum \frac{e^{p_m t}}{Z'(p_m)}. \tag{9–15}$$

The summation is extended over all the zeros of $Z(p)$.

Similarly, for an impulse of finite duration, we obtain from (12)

$$B(t) = \sum \frac{2 \sinh \frac{p_m \tau}{2} e^{p_m t}}{p_m \tau Z'(p_m)} \tag{9–16}$$

for $t > \tau/2$, that is, in the interval after the impulse has ceased. During the operation of the impulse, we have $-\tau/2 < t < \tau/2$ and (C) in (12) cannot be closed with infinite semicircles since their contributions become infinite instead of infinitesimal. Substituting $\sinh p\tau/2 = \frac{1}{2}(e^{p\tau/2} - e^{-p\tau/2})$ in (12), we obtain

$$B(t) = \frac{1}{2\pi i\tau}\left[\int_{(C)} \frac{e^{p[t+(\tau/2)]}}{pZ(p)}\,dp - \int_{(C)} \frac{e^{p[t-(\tau/2)]}}{pZ(p)}\,dp\right].$$

Within our interval $t + \tau/2$ is positive and (C) in the first integral can be closed on the left; at the same time $t - \tau/2$ is negative and (C) in the second integral can be closed on the right. The second integral vanishes and we have

$$B(t) = \frac{1}{\tau Z(0)} + \frac{1}{\tau}\sum \frac{e^{p_m[t+(\tau/2)]}}{p_m Z'(p_m)}, \qquad (9\text{–}17)$$

assuming that $Z(0) \neq 0$. The first term arises from the pole at the origin. Evidently (17) represents the response to a voltage step of magnitude $1/\tau$ at $t = -\tau/2$. This is not surprising, since during the operation of the impulse the circuit does not know that the voltage will cease to operate.

The significance of the first term in (17) will be understood if we apply this equation to the circuit shown in Fig. 2.21. While the impulse is operating, we have

$$B(t) = \frac{1}{R\tau} - \frac{1}{R\tau}e^{-(R/L)[t+(\tau/2)]}. \qquad (9\text{–}18)$$

The response (14) to an impulse of infinite magnitude but of zero duration starts with a finite value. For a physical impulse of finite duration, the response (18) is zero at the instant the voltage begins to operate and builds up to the value

$$B\left(\frac{\tau}{2}\right) = \frac{1 - e^{-R\tau/L}}{R\tau},$$

at the instant $t = \tau/2$, when the voltage is off. Subsequently the response decreases exponentially. If the interval τ is so short that $R\tau/L$ is much less than unity, $B(\tau/2)$ becomes approximately $1/L$; this agrees with (14).

2.10. *Natural and Forced Waves*

Generally speaking, several wave functions are associated with a physical wave. When a wave is traveling along a string under tension, a typical point is not only displaced from the neutral position but it is also moving with some velocity and it is acted upon by some force. Thus we have a wave of displacement, a wave of velocity, and a wave of force. In electromagnetic waves we are confronted with interdependent waves of electric and magnetic intensities. The ratios of certain space-time wave functions play just as important a part in wave theory as the ratios of wave functions depending only on time play in the theory of oscillations. These generalized ratios are called *wave impedances* and are differentiated among themselves by qualifying adjectives and phrases.

For the present we shall confine ourselves to the simplest type of wave motion: waves in a transmission line. The equations governing these waves are

$$\frac{\partial V_i}{\partial x} = -\left(RI_i + L\frac{\partial I_i}{\partial t}\right) + E_i(x,t), \quad \frac{\partial I_i}{\partial x} = -\left(GV_i + C\frac{\partial V_i}{\partial t}\right), \qquad (10\text{–}1)$$

where: V_i and I_i are respectively the instantaneous transverse voltage across the line and the longitudinal electric current in it; E_i is the voltage per unit length, impressed along the line in " series " with it; R, L, G, C are constants representing the series

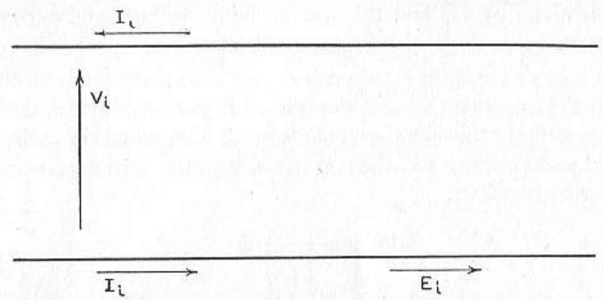


FIG. 2.22. A diagram explaining the convention regarding the positive directions of the transverse electromotive force V and the longitudinal electric current I in the lower wire (that is, the transverse magnetomotive force around the wire) of a transmission line consisting of two parallel wires.

resistance, the series inductance, the shunt conductance, the shunt capacitance—all per unit length of the line. A pair of parallel wires (Fig. 2.22) is a concrete example of a transmission line; the arrows in the figure explain the convention with regard to the positive directions of the variables V_i, I_i, E_i.

Let E_i, I_i, and V_i be harmonic functions; then

$$E_i(x,t) = \text{re}(\hat{E}e^{i\omega t}), \quad V_i = \text{re}(\hat{V}e^{i\omega t}), \quad I_i = \text{re}(\hat{I}e^{i\omega t}). \qquad (10\text{–}2)$$

As we have seen these complex exponentials will also satisfy equations (1). Substituting them in these equations and canceling the time factor $e^{i\omega t}$, we obtain a set of ordinary differential equations

$$\frac{d\hat{V}}{dx} = -(R + i\omega L)\hat{I} + \hat{E}(x), \quad \frac{d\hat{I}}{dx} = -(G + i\omega C)\hat{V}. \qquad (10\text{–}3)$$

The expressions $Z = R + i\omega L$, $Y = G + i\omega C$, are known as the series impedance and the shunt admittance per unit length.

Let us now suppose that the applied force $\hat{E}$ is distributed exponentially along the line. To solve the equations we assume that the response $\hat{V}$, $\hat{I}$ is also exponential and we write tentatively

$$\hat{E}(x) = Ee^{\gamma x}, \quad \hat{V}(x) = Ve^{\gamma x}, \quad \hat{I}(x) = Ie^{\gamma x}. \qquad (10\text{–}4)$$

Substituting in (3) and solving, we obtain

$$\gamma V + ZI = E, \quad YV + \gamma I = 0, \tag{10-5}$$

$$V = \frac{\gamma}{\gamma^2 - ZY} E, \quad I = \frac{Y}{ZY - \gamma^2} E.$$

Thus we find that a response of type (4) is possible except when the propagation constant γ satisfies the following equation

$$\gamma^2 - ZY = 0. \tag{10-6}$$

This response is given by (5) and the corresponding voltage and current waves are called *forced waves*, by analogy with forced oscillations.

When γ is a root of (6), that is, when $\gamma = \pm\sqrt{ZY}$, there is no finite response of type (4), unless the impressed force E vanishes. Exponential waves may exist in the transmission line without the operation of an applied force along the entire line. These waves are called *natural waves* and the corresponding propagation constants are called *natural propagation constants.*

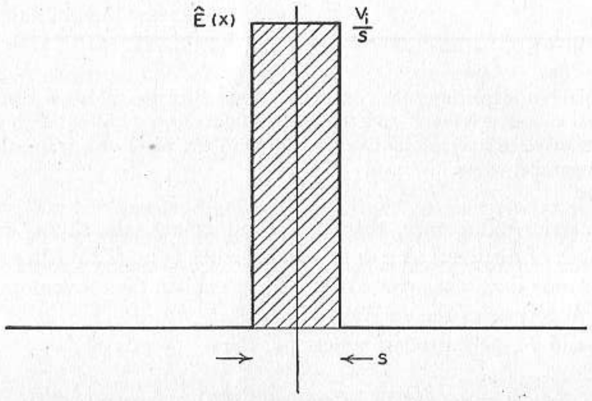


FIG. 2.23. A spatial impulse function representing a highly localized impressed electromotive force.

Just as natural oscillations can be produced by a temporal impulse of force, natural waves can be produced by a spatial impulse of force (Fig. 2.23). The magnitude of this impulse is the applied or impressed voltage

$$V^i = \int_{-s/2}^{s/2} \hat{E}(x)\, dx, \tag{10-7}$$

and it is represented by the area of the impulse. By (9-8) we have

$$\hat{E}(x) = \frac{V^i}{2\pi i} \int_{(C)} \frac{\sinh \frac{\gamma s}{2}}{\frac{\gamma s}{2}} e^{\gamma x}\, d\gamma,$$

where (C) is a contour in the γ-plane, shown in Fig. 2.24. Consequently from (4) and (5) we obtain

$$V(x) = \frac{V^i}{i\pi s} \int_{(C)} \frac{\sinh \frac{\gamma s}{2}}{\gamma^2 - ZY} e^{\gamma x}\, d\gamma, \quad I(x) = \frac{YV^i}{i\pi s} \int_{(C)} \frac{\sinh \frac{\gamma s}{2}}{\gamma(ZY - \gamma^2)} e^{\gamma x}\, d\gamma. \tag{10-8}$$

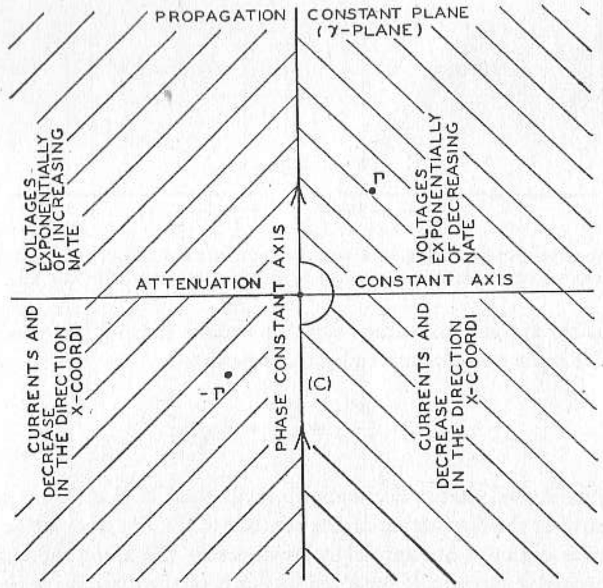


FIG. 2.24. The contour (C) in the propagation constant plane.

These integrals are convergent as $s \to 0$; thus for a point generator we have

$$V(x) = \frac{V^i}{2\pi i} \int_{(C)} \frac{\gamma e^{\gamma x}}{\gamma^2 - ZY} d\gamma, \quad I(x) = \frac{YV^i}{2\pi i} \int_{(C)} \frac{e^{\gamma x}}{ZY - \gamma^2} d\gamma. \tag{10-9}$$

Each integrand in (9) has two poles (Fig. 2.24) corresponding to the natural propagation constants; in the present case the origin is not a pole and thus contributes nothing to the value of the integral. Let $\Gamma = \sqrt{ZY}$ be that value of the square root which is located in the first quadrant of the γ-plane or on its boundaries.* If $x > 0$, (C) can be closed with an infinite semicircle in the left half of the γ-plane. Evaluating (9) we have

$$V(x) = \tfrac{1}{2} V^i e^{-\Gamma x}, \quad I(x) = \frac{YV^i}{2\Gamma} e^{-\Gamma x}; \quad x > 0. \tag{10-10}$$

If $x < 0$, (C) can be closed in the right half of the γ-plane; then we obtain from (9)

$$V(x) = -\tfrac{1}{2} V^i e^{\Gamma x}, \quad I(x) = \frac{YV^i}{2\Gamma} e^{\Gamma x}; \quad x < 0. \tag{10-11}$$

* Since R, L, G, C are positive, $\sqrt{ZY}$ is either in the first quadrant or in the third.

Thus if a harmonic voltage $V^i e^{i\omega t}$ is inserted at some point $x = 0$ of an infinitely long transmission line (Fig. 2.25) in series with it, two waves are originated, one traveling to the right and the other to the left. The current through the generator is

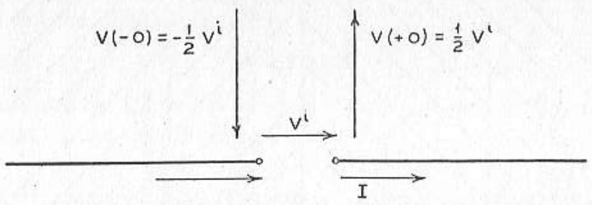


FIG. 2.25. The conditions existing in a transmission line extending to infinity in both directions when an electric generator is inserted in series with the line.

continuous but the transverse voltage suffers a sudden jump V^i in passing across the generator. The *input impedance* seen by the generator is

$$Z_i = \frac{V^i}{I(0)} = \frac{2\Gamma}{Y} = 2\sqrt{\frac{Z}{Y}}.$$

It will be remembered that in dealing with oscillations in electric circuits there was a question regarding the disposition of the contour (C). For dissipative circuits, the correct result was obtained automatically by choosing (C) along the imaginary axis; and for nondissipative circuits, if regarded as limits of the dissipative circuits, it was natural to indent (C) as in Fig. 2.20. Nevertheless the only valid reason for making (C) pass on the right of the poles (natural oscillation constants) is to satisfy the physical condition, not included in the differential equation, to the effect that there should be no response to a force before it begins to operate. A similar situation exists in the present case; this time, however, we expect the waves to travel in both directions from the point source and hence we want (C) to separate the poles (the natural propagation constants). But these poles might be separated as shown in Fig. 2.26, in which case we should obtain

$$V(x) = \tfrac{1}{2}V^i e^{\Gamma x}, \quad I(x) = -\frac{YV^i}{2\Gamma}e^{\Gamma x}; \quad x > 0;$$
$$\qquad (10\text{–}12)$$
$$V(x) = -\tfrac{1}{2}V^i e^{-\Gamma x}, \quad I(x) = -\frac{YV^i}{2\Gamma}e^{-\Gamma x}; \quad x < 0;$$

instead of (10) and (11). We can object to this result on two counts. In the first place, for dissipative lines the real part of Γ is positive and (12) states that the voltage and the current increase exponentially with the distance from the generator. This is contrary to our experience and would imply that infinite power could be dissipated in the line when finite power is supplied by the generator. In the second place, equations

(12) imply that power is not supplied by the generator to the line but that the line contributes power to the generator. Thus, by (12), the input impedances of dissipative and nondissipative lines are respectively

$$Z_i = \frac{V^i}{I(0)} = -\frac{2\Gamma}{Y}, \quad Z_i = -2\sqrt{\frac{L}{C}};$$

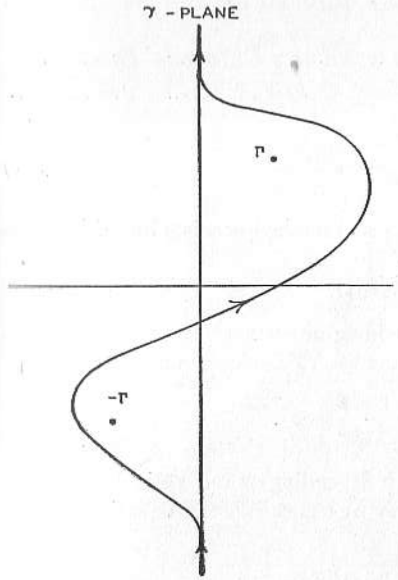


FIG. 2.26. A "forbidden" form of the contour of integration.

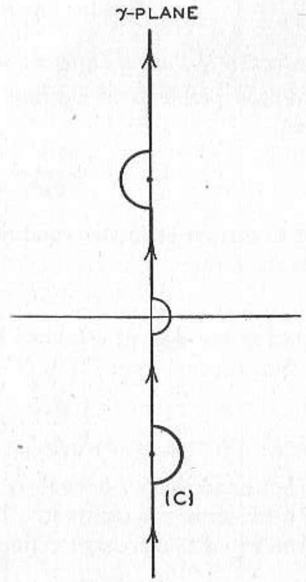


FIG. 2.27. Indentations in the contour of integration when the natural propagation constants move to the imaginary axis as the dissipation in the transmission line approaches zero.

hence the current flows in opposition to the impressed force. Therefore, we come to the conclusion that (C) must separate the natural propagation constants as shown in Fig. 2.24. For nondissipative lines this contour assumes the form shown in Fig. 2.27.

We have dismissed at the start the possibility that the wave could be started on one side of the generator and not on the other; for, when an alternating voltage is impressed at some point of the line, there is nothing to indicate on which side the wave should be if it is to be on one side only. In any case this possibility is contrary to experience.

The method just explained is particularly useful in more advanced chapters on wave theory. In the present case there is a simpler way of looking at this type of problem (see Chapter 7).

CHAPTER III

Bessel and Legendre Functions

3.1. *Reduction of Partial Differential Equations to Ordinary Differential Equations*

Numerous problems of electromagnetic theory depend on solving the following equation

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = \sigma^2 V, \tag{1-1}$$

subject to certain boundary conditions. The usual method consists in seeking solutions of the form

$$V = X(x)Y(y)Z(z), \tag{1-2}$$

and forming the desired solutions by either adding or integrating functions of this type. Substituting from (2) in (1) and dividing by XYZ, we obtain

$$\frac{1}{X}\frac{d^2X}{dx^2} + \frac{1}{Y}\frac{d^2Y}{dy^2} + \frac{1}{Z}\frac{d^2Z}{dz^2} = \sigma^2.$$

On the left-hand side we have three terms, each depending on one variable only; the sum of these terms is a constant. The only way we can satisfy this equation is to set each term equal to a constant; thus

$$\frac{d^2X}{dx^2} = \sigma_x^2 X, \quad \frac{d^2Y}{dy^2} = \sigma_y^2 Y, \quad \frac{d^2Z}{dz^2} = \sigma_z^2 Z. \tag{1-3}$$

The constants σ_x, σ_y, σ_z must satisfy

$$\sigma_x^2 + \sigma_y^2 + \sigma_z^2 = \sigma^2; \tag{1-4}$$

no other restrictions are imposed by the wave equation itself.

Functions of type (2) are particularly suitable for finding solutions satisfying prescribed boundary conditions at plane boundaries. If the boundaries are cylindrical or spherical, it is more practical to transcribe equation (1) in cylindrical or spherical coordinates and then separate the variables. Thus in cylindrical coordinates (1) becomes

$$\frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho \frac{\partial V}{\partial \rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 V}{\partial \varphi^2} + \frac{\partial^2 V}{\partial z^2} = \sigma^2 V. \tag{1-5}$$

Assuming

$$V = R(\rho)\Phi(\varphi)Z(z), \tag{1-6}$$

substituting in (5), and dividing by $R\Phi Z$, we have

$$\frac{1}{\rho R}\frac{d}{d\rho}\left(\rho \frac{dR}{d\rho}\right) + \frac{1}{\rho^2 \Phi}\frac{d^2\Phi}{d\varphi^2} + \frac{1}{Z}\frac{d^2Z}{dz^2} = \sigma^2. \tag{1-7}$$

Since all the terms except the third are independent of z, the third term must also be independent of z; thus

$$\frac{1}{Z}\frac{d^2Z}{dz^2} = \sigma_z^2, \quad \text{or} \quad \frac{d^2Z}{dz^2} = \sigma_z^2 Z.$$

Substituting in (7) and multiplying by ρ^2, we have

$$\frac{\rho}{R}\frac{d}{d\rho}\left(\rho \frac{dR}{d\rho}\right) + \frac{1}{\Phi}\frac{d^2\Phi}{d\varphi^2} = (\sigma^2 - \sigma_z^2)\rho^2. \tag{1-8}$$

The first and third terms are now independent of φ; hence the second term must also be independent of φ

$$\frac{1}{\Phi}\frac{d^2\Phi}{d\varphi^2} = \sigma_\varphi^2, \quad \text{or} \quad \frac{d^2\Phi}{d\varphi^2} = \sigma_\varphi^2 \Phi.$$

Substituting in (8), we have

$$\rho \frac{d}{d\rho}\left(\rho \frac{dR}{d\rho}\right) + [\sigma_\varphi^2 - (\sigma^2 - \sigma_z^2)\rho^2]R = 0,$$

or

$$\rho^2 \frac{d^2R}{d\rho^2} + \rho \frac{dR}{d\rho} + [\sigma_\varphi^2 - (\sigma^2 - \sigma_z^2)\rho^2]R = 0. \tag{1-9}$$

This is Bessel's equation. Equation (5) places no restrictions on the values of σ_z and σ_φ.

In spherical coordinates (1) becomes

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial V}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial V}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 V}{\partial \varphi^2} = \sigma^2 V. \tag{1-10}$$

Again letting $V = R(r)T(\theta,\varphi)$, substituting in (10), and using the above method of separation, we arrive at two equations

$$\sin\theta \frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial T}{\partial \theta}\right) + \frac{\partial^2 T}{\partial \varphi^2} + k^2 \sin^2\theta\, T = 0, \tag{1-11}$$

$$\frac{d}{dr}\left(r^2 \frac{dR}{dr}\right) = (k^2 + \sigma^2 r^2)R. \tag{1-12}$$

The first of these equations is the equation of Spherical Harmonics and the second is reducible to the Bessel equation.

Setting further, $T(\theta,\varphi) = \Theta(\theta)\Phi(\varphi)$ and substituting in (11), we have

$$\frac{d^2\Phi}{d\varphi^2} = -m^2\Phi, \tag{1-13}$$

$$\sin\theta \frac{d}{d\theta}\left(\sin\theta \frac{d\Theta}{d\theta}\right) + (k^2 \sin^2\theta - m^2)\Theta = 0. \tag{1-14}$$

This is the Associated Legendre Equation.

The constant m in (13) may be complex; but in many problems it is real and frequently an integer. The constant k may also be complex but often it is real and of the form $k^2 = n(n+1)$, where n is an integer. When m is an integer, Φ is a periodic function with period 2π. When n is an integer, some of the solutions of (14) are finite for all values of θ; for all other values of n the solutions of (14) become infinite either for $\theta = 0$ or for $\theta = \pi$.

There is an equation related to (12), which is important in subsequent work. Rewriting (12) in the form

$$r^2 \frac{d^2R}{dr^2} + 2r\frac{dR}{dr} = (k^2 + \sigma^2 r^2)R, \tag{1–15}$$

and substituting $\hat{R} = rR$, $R = \hat{R}/r$, we have

$$\frac{dR}{dr} = \frac{1}{r}\frac{d\hat{R}}{dr} - \frac{\hat{R}}{r^2}, \quad \frac{d^2R}{dr^2} = \frac{1}{r}\frac{d^2\hat{R}}{dr^2} - \frac{2}{r^2}\frac{d\hat{R}}{dr} + \frac{2\hat{R}}{r^3}.$$

Inserting in (15), we have

$$\frac{d^2\hat{R}}{dr^2} = \left(\sigma^2 + \frac{k^2}{r^2}\right)\hat{R}. \tag{1–16}$$

For $k = 0$ the solution of (16) is

$$\hat{R} = Ae^{-\sigma r} + Be^{\sigma r}, \tag{1–17}$$

and the corresponding solution of (17) becomes

$$R = \frac{Ae^{-\sigma r}}{r} + \frac{Be^{\sigma r}}{r}. \tag{1–18}$$

For sufficiently large values of r the term k^2/r^2 is small compared with σ^2; hence solutions of (16) approach asymptotically to (17) and the corresponding solutions of (15) approach (18).

3.2. *Boundary Conditions*

We have seen that in the case of solutions of type (1–2), two of the three constants σ_x, σ_y, σ_z are arbitrary. Various supplementary physical conditions will restrict these constants in one way or another. For instance, let us suppose that physical conditions require that V should vanish for $x = 0$ and $x = a$, that it should be independent of y, and that it should vanish at $z = \infty$. As a function of x, V will vanish as required if $X(0)$ and $X(a)$ vanish. Writing the general solution for X in the form

$$X(x) = A\cosh\sigma_x x + B\sinh\sigma_x x, \tag{2–1}$$

we find that $X(0)$ vanishes only if $A = 0$; $X(a)$ vanishes if

$$\sinh\sigma_x a = 0, \quad \sigma_x a = in\pi, \quad \sigma_x = \frac{in\pi}{a}, \tag{2–2}$$

where n is an integer. The constant B remains arbitrary.

V will be independent of y if Y is a constant; then $\sigma_y = 0$. Hence, from (2) and (1–4) we have

$$\sigma_z = \sqrt{\sigma^2 + \frac{n^2\pi^2}{a^2}}, \quad n = 1, 2, 3, \cdots. \tag{2–3}$$

If we take the particular value of the square root which lies in the first quadrant, the general expression for $Z(z)$ may be written as follows:

$$Z(z) = C_n e^{-\sigma_{z,n} z} + D_n e^{\sigma_{z,n} z}. \tag{2–4}$$

To ensure that V vanishes at $z = \infty$, we must take $D_n = 0$; C_n remains arbitrary.

Combining the above results, we find that

$$V = \sum_{n=1}^{\infty} C_n \sin\frac{n\pi x}{a}\, e^{-\sigma_{z,n} z} \tag{2–5}$$

satisfies the specified supplementary conditions. The coefficients C_n remain arbitrary and it is possible to find V so that it reduces to a given function $f(x)$ when z is given; we need only expand $f(x)$ in a sine series and compare the coefficients.

There may be other kinds of supplementary conditions imposed by physical considerations. For solutions of the type (1–6) Φ may be required to be periodic; for other types of solutions some factor may have to be finite in a certain range of the variable; all such conditions are best considered as they arise.

3.3. *Bessel Functions*

Bessel functions of order ν are defined as solutions of Bessel's equation

$$z^2\frac{d^2w}{dz^2} + z\frac{dw}{dz} + (z^2 - \nu^2)w = 0. \tag{3–1}$$

Being a second order differential equation, it possesses two linearly independent solutions. For nonintegral values of ν these solutions can be expressed by the following power series

$$J_\nu(z) = \sum_{m=0}^{\infty} \frac{(-)^m z^{\nu+2m}}{m!(\nu+m)!\,2^{\nu+2m}}, \quad J_{-\nu}(z) = \sum_{m=0}^{\infty} \frac{(-)^m z^{-\nu+2m}}{m!(-\nu+m)!\,2^{-\nu+2m}}, \tag{3–2}$$

where the generalized factorial is defined in terms of the Gamma function

$$p! = \Gamma(p+1). \tag{3–3}$$

The point $z = 0$ is seen to be a branch point. The most general solution of (1) is a linear combination of these two functions.

If ν is a positive integer n, then

$$J_{-n}(z) = (-)^n J_n(z), \tag{3–4}$$

and we are left with only one solution, regular at $z = 0$. A second solution is defined as follows. For any nonintegral order ν the function

$$N_\nu(z) = \frac{J_\nu(z)\cos\nu\pi - J_{-\nu}(z)}{\sin\nu\pi} \tag{3–5}$$

is a solution of the Bessel equation. Its limit as ν approaches an integer n

$$N_n(z) = \lim N_\nu(z) \quad \text{as} \quad \nu \to n \tag{3-6}$$

continues to be a solution of the Bessel equation. A series expansion for this function is

$$N_n(z) = -\frac{1}{\pi}\sum_{m=0}^{n-1} \frac{(n-m-1)!2^{n-2m}}{m!\,z^{n-2m}} + \frac{2}{\pi}(\log z + C - \log 2)J_n(z)$$
$$-\frac{1}{\pi}\sum_{m=0}^{\infty} \frac{(-)^m z^{n+2m}}{2^{n+2m} m!(m+n)!}[\varphi(m) + \varphi(m+n)],$$

$$N_0(z) = \frac{2}{\pi}(\log z + C - \log 2)J_0(z) + \frac{2}{\pi}\sum_{m=1}^{\infty}\frac{(-)^{m+1}z^{2m}}{2^{2m}(m!)^2}\varphi(m), \tag{3-7}$$

where the auxiliary function $\varphi(m)$ is defined by

$$\varphi(m) = 1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \frac{1}{m}, \tag{3-8}$$

and C is Euler's constant

$$C = 0.5772\cdots. \tag{3-9}$$

From (5) and (6) we have

$$N_n(z) = \frac{1}{\pi}\left[\frac{\partial J_\nu}{\partial \nu} + (-)^{n+1}\frac{\partial J_{-\nu}}{\partial \nu}\right]_{\nu=n}. \tag{3-10}$$

The series (7) is obtained by differentiating (2) with respect to ν. In differentiating power series one frequently encounters the logarithmic derivative of the generalized factorial function; this function $\psi(z)$ is

$$\psi(z) = \frac{d}{dz}\log(z!) = \frac{1}{z!}\frac{d(z!)}{dz} = -C + \sum_{n=1}^{\infty}\left(\frac{1}{n} - \frac{1}{n+z}\right). \tag{3-11}$$

For integral values of the independent variable we have (m is a positive integer)

$$\psi(0) = -C, \quad \psi(m) = -C + \varphi(m), \quad \psi(-m) = \infty. \tag{3-12}$$

Also

$$\psi(-\tfrac{1}{2}) = -C - 2\log 2 = -1.96351\cdots,$$
$$\psi(-\tfrac{1}{2} \pm m) = \psi(-\tfrac{1}{2}) + 2\left(1 + \tfrac{1}{3} + \tfrac{1}{5} + \cdots + \frac{1}{2m-1}\right). \tag{3-13}$$

The following reduction formulae are useful

$$\psi(z) = \psi(z-1) + \frac{1}{z},$$
$$\psi(-z) = \psi(z-1) + \pi\cot\pi z, \tag{3-14}$$
$$\psi(z-\tfrac{1}{2}) = \psi(-z-\tfrac{1}{2}) + \pi\tan\pi z.$$

For other formulae concerning $\psi(z)$ and its derivatives the reader is referred to Jahnke and Emde, Tables of Functions, p. 92.

The N-function as here defined is identical with the N-function of Jahnke and

Emde and is the same as the tabulated "Y-Function" of G. N. Watson. The N-notation is favored because it seems to be less confusing; there are several different "Y-functions."

Two linear combinations of J and N functions are known as Hankel's functions,

$$H_\nu^{(1)}(z) = J_\nu(z) + iN_\nu(z),$$
$$H_\nu^{(2)}(z) = J_\nu(z) - iN_\nu(z). \tag{3-15}$$

For small values of z we have approximately

$$J_\nu(z) = \frac{z^\nu}{2^\nu \nu!}, \quad N_\nu(z) = -\frac{2^\nu(\nu-1)!}{\pi z^\nu},$$
$$J_0(z) = 1, \quad N_0(z) = \frac{2}{\pi}(\log z + C - \log 2), \tag{3-16}$$

provided $\text{re}(\nu) > 0$ in the expression for N_ν. For large values of z we have asymptotic expressions

$$J_\nu(z) = \sqrt{\frac{2}{\pi z}}\cos\left(z - \frac{\nu\pi}{2} - \frac{\pi}{4}\right), \quad N_\nu(z) = \sqrt{\frac{2}{\pi z}}\sin\left(z - \frac{\nu\pi}{2} - \frac{\pi}{4}\right),$$
$$H_\nu^{(1)}(z) = \sqrt{\frac{2}{\pi z}}\exp i\left(z - \frac{\nu\pi}{2} - \frac{\pi}{4}\right), \quad H_\nu^{(2)}(z) = \sqrt{\frac{2}{\pi z}}\exp i\left(-z + \frac{\nu\pi}{2} + \frac{\pi}{4}\right). \tag{3-17}$$

These asymptotic expansions are valid only within certain phase limits: in the expressions for J_ν and N_ν, the phase of z must lie in the interval $(-\pi,\pi)$, in the expression for $H_\nu^{(1)}$ the phase is in the interval $(-\pi,2\pi)$, and in the expression for $H_\nu^{(2)}$ the phase is in the interval $(-2\pi,\pi)$. These restrictions are needed because Bessel functions are multiple-valued functions.

Complete asymptotic expansions of the various Bessel functions are

$$H_\nu^{(1)}(z) \sim \left(\frac{2}{\pi z}\right)^{1/2}\exp i(z - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi)\sum_{m=0}^{\infty}\frac{(-)^m(\nu,m)}{(2iz)^m},$$
$$H_\nu^{(2)}(z) \sim \left(\frac{2}{\pi z}\right)^{1/2}\exp i(-z + \tfrac{1}{2}\nu\pi + \tfrac{1}{4}\pi)\sum_{m=0}^{\infty}\frac{(\nu,m)}{(2iz)^m},$$
$$J_\nu(z) \sim \left(\frac{2}{\pi z}\right)^{1/2}\left[\cos(z - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi)\sum_{m=0}^{\infty}\frac{(-)^m(\nu,2m)}{(2z)^{2m}}\right.$$
$$\left. - \sin(z - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi)\sum_{m=0}^{\infty}\frac{(-)^m(\nu,2m+1)}{(2z)^{2m+1}}\right], \tag{3-18}$$
$$N_\nu(z) \sim \left(\frac{2}{\pi z}\right)^{1/2}\left[\sin(z - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi)\sum_{m=0}^{\infty}\frac{(-)^m(\nu,2m)}{(2z)^{2m}}\right.$$
$$\left. + \cos(z - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi)\sum_{m=0}^{\infty}\frac{(-)^m(\nu,2m+1)}{(2z)^{2m+1}}\right].$$

The auxiliary function (ν,m) is defined by

$$(\nu,m) = \frac{(\nu+m-\frac{1}{2})!}{m!(\nu-m-\frac{1}{2})!} = \frac{(4\nu^2-1^2)(4\nu^2-3^2)\cdots[4\nu^2-(2m-1)^2]}{2^{2m}m!}. \quad (3\text{–}19)$$

When $\nu = n+\frac{1}{2}$, then

$$(n+\tfrac{1}{2},m) = \frac{(n+m)!}{m!(n-m)!}. \quad (3\text{–}20)$$

The expansions (18) are divergent; but if a finite number of terms is retained, then as z increases the sum of these terms will represent the corresponding function with increasing accuracy.

3.4. *Modified Bessel Functions*

The modified Bessel equation is

$$z^2\frac{d^2w}{dz^2} + z\frac{dw}{dz} - (z^2+\nu^2)w = 0 \quad (4\text{–}1)$$

and its solutions are called modified Bessel functions. For nonintegral values of ν a set of two linearly independent solutions is

$$I_\nu(z) = \sum_{m=0}^{\infty}\frac{z^{\nu+2m}}{m!(\nu+m)!\,2^{\nu+2m}}, \quad I_{-\nu}(z) = \sum_{m=0}^{\infty}\frac{z^{-\nu+2m}}{m!(-\nu+m)!\,2^{-\nu+2m}}. \quad (4\text{–}2)$$

Another important solution is a linear combination of these two functions

$$K_\nu(z) = \frac{\pi}{2\sin\nu\pi}[I_{-\nu}(z) - I_\nu(z)]. \quad (4\text{–}3)$$

If ν is a positive integer n, then

$$I_{-n}(z) = I_n(z), \quad (4\text{–}4)$$

and the equations (2) represent only one solution of the modified Bessel equation. A second solution is obtained from (3) by allowing ν to approach n and passing to the limit

$$K_n(z) = \lim K_\nu(z) \quad \text{as} \quad \nu\to n,$$
$$= \frac{2}{\cos n\pi}\left(\frac{\partial I_{-n}}{\partial n} - \frac{\partial I_n}{\partial n}\right). \quad (4\text{–}5)$$

From this definition the following series are obtained

$$K_n(z) = \sum_{m=0}^{n-1}\frac{(-)^m(n-m-1)!\,2^{n-2m-1}}{m!\,z^{n-2m}} + (-)^{n+1}(\log z + C - \log 2)I_n(z)$$
$$+ (-)^n\sum_{m=0}^{\infty}\frac{z^{n+2m}}{2^{n+2m+1}m!(n+m)!}[\varphi(m)+\varphi(n+m)], \quad (4\text{–}6)$$

$$K_0(z) = -(\log z + C - \log 2)I_0(z) + \sum_{m=1}^{\infty}\frac{z^{2m}}{2^{2m}(m!)^2}\varphi(m).$$

For small values of z we have approximately

$$I_\nu(z) = \frac{z^\nu}{\nu!\,2^\nu}, \quad K_\nu(z) = \frac{(\nu-1)!\,2^{\nu-1}}{z^\nu},$$
$$I_0(z) = 1, \qquad K_0(z) = -(\log z + C - \log 2), \quad (4\text{–}7)$$

provided $\text{re}(\nu) > 0$ in the expression for K_ν.

If z is large, we have asymptotically

$$K_\nu(z) \sim \left(\frac{\pi}{2z}\right)^{1/2} e^{-z}\left[1 + \frac{4\nu^2-1^2}{1!8z} + \frac{(4\nu^2-1^2)(4\nu^2-3^2)}{2!(8z)^2} + \cdots\right] \quad (4\text{–}8)$$

provided $-3\pi/2 < ph(z) < 3\pi/2$.

Similarly for $I_\nu(z)$ we have an asymptotic expansion

$$I_\nu(z) \sim \frac{e^z}{\sqrt{2\pi z}}\sum_{m=0}^{\infty}\frac{(-)^m(\nu,m)}{(2z)^m} + \frac{e^{-z+[\nu+(1/2)]\pi i}}{\sqrt{2\pi z}}\sum_{m=0}^{\infty}\frac{(\nu,m)}{(2z)^m} \quad (4\text{–}9)$$

provided $-\pi/2 < ph(z) < 3\pi/2$.

The modified Bessel functions are related to Bessel functions; thus

$$I_\nu(iz) = e^{i\nu\pi/2}J_\nu(z),$$
$$K_\nu(iz) = \frac{\pi}{2}e^{-i(\nu+1)\pi/2}[J_\nu(z) - iN_\nu(z)]. \quad (4\text{–}10)$$

In particular, we have

$$I_0(iz) = J_0(z), \quad I_1(iz) = iJ_1(z),$$
$$K_0(iz) = \frac{\pi}{2i}[J_0(z) - iN_0(z)] = -\frac{\pi}{2}[N_0(z) + iJ_0(z)], \quad (4\text{–}11)$$
$$K_1(iz) = -\frac{\pi}{2}[J_1(z) - iN_1(z)].$$

In wave problems we find that for dissipative media the I and K functions are more suitable than the others, in view of the usual engineering convention with regard to propagation constants. If the arguments of the I and K functions are imaginary it is frequently desirable to separate real and imaginary parts; then we shall employ the J and N functions.

3.5. *Bessel Functions of Order $n+\frac{1}{2}$ and Related Functions*

If $\nu = n+\frac{1}{2}$, the asymptotic series representing Bessel functions and modified Bessel functions terminate. In this case Bessel functions reduce to elementary functions and the following related functions are better suited to wave problems than the original Bessel functions

$$\hat{K}_n(z) = \left(\frac{2z}{\pi}\right)^{1/2}K_{n+1/2}(z), \quad \hat{J}_n(z) = \left(\frac{\pi z}{2}\right)^{1/2}J_{n+1/2}(z), \quad (5\text{–}1)$$

$$\hat{I}_n(z) = \left(\frac{\pi z}{2}\right)^{1/2} I_{n+1/2}(z), \quad \hat{N}_n(z) = \left(\frac{\pi z}{2}\right)^{1/2} N_{n+1/2}(z). \tag{5–1}$$

The $\hat{K}$ and $\hat{I}$ functions are solutions of the following differential equation, related to (1–16),

$$\frac{d^2w}{dz^2} = \left[1 + \frac{n(n+1)}{z^2}\right] w. \tag{5–2}$$

The $\hat{J}$ and $\hat{N}$ functions satisfy

$$\frac{d^2w}{dz^2} = -w + \frac{n(n+1)}{z^2} w. \tag{5–3}$$

From (4–10) and (1) we have

$$\hat{I}_n(iz) = i^{n+1}\hat{J}_n(z), \quad \hat{I}'_n(iz) = i^n\hat{J}'_n(z),$$
$$\hat{K}_n(iz) = i^{-n-1}[\hat{J}_n(z) - i\hat{N}_n(z)], \quad \hat{K}'_n(iz) = i^{-n-2}[\hat{J}'_n(z) - i\hat{N}'_n(z)]. \tag{5–4}$$

The analytic expressions for the functions defined in (1) are

$$\hat{K}_n(z) = e^{-z} \sum_{m=0}^{n} \frac{(n+m)!}{m!(n-m)!(2z)^m},$$

$$\hat{I}_n(z) = \tfrac{1}{2}\left[e^z \sum_{m=0}^{n} \frac{(-)^m(n+m)!}{m!(n-m)!(2z)^m} + (-)^{n+1}e^{-z} \sum_{m=0}^{n} \frac{(n+m)!}{m!(n-m)!(2z)^m}\right],$$

$$\hat{J}_n(z) = \left(\cos\frac{n\pi}{2}\sin z - \sin\frac{n\pi}{2}\cos z\right) \sum_{m=0}^{2m\le n} \frac{(-)^m(n+2m)!}{2m!(n-2m)!(2z)^{2m}}$$
$$+ \left(\cos\frac{n\pi}{2}\cos z + \sin\frac{n\pi}{2}\sin z\right) \sum_{m=0}^{2m\le n-1} \frac{(-)^m(n+2m+1)!}{(2m+1)!(n-2m-1)!(2z)^{2m+1}}, \tag{5–5}$$

$$\hat{N}_n(z) = -\left(\cos\frac{n\pi}{2}\cos z + \sin\frac{n\pi}{2}\sin z\right) \sum_{m=0}^{2m\le n} \frac{(-)^m(n+2m)!}{2m!(n-2m)!(2z)^{2m}}$$
$$+ \left(\cos\frac{n\pi}{2}\sin z - \sin\frac{n\pi}{2}\cos z\right) \sum_{m=0}^{2m\le n-1} \frac{(-)^m(n+2m+1)!}{(2m+1)!(n-2m-1)!(2z)^{2m+1}}.$$

In particular we have

$$\hat{K}_0(z) = e^{-z}, \quad \hat{K}_1(z) = e^{-z}\left(1 + \frac{1}{z}\right), \quad \hat{K}_2(z) = e^{-z}\left(1 + \frac{3}{z} + \frac{3}{z^2}\right),$$

$$\hat{I}_0(z) = \sinh z, \quad \hat{I}_1(z) = \cosh z - \frac{\sinh z}{z}, \quad \hat{I}_2(z) = \left(1 + \frac{3}{z^2}\right)\sinh z - \frac{3}{z}\cosh z,$$

$$\hat{J}_0(z) = \sin z, \quad \hat{J}_1(z) = \frac{\sin z}{z} - \cos z, \quad \hat{J}_2(z) = \left(\frac{3}{z^2} - 1\right)\sin z - \frac{3}{z}\cos z, \tag{5–6}$$

$$\hat{N}_0(z) = -\cos z, \quad \hat{N}_1(z) = -\sin z - \frac{\cos z}{z}, \quad \hat{N}_2(z) = \left(1 - \frac{3}{z^2}\right)\cos z - \frac{3}{z}\sin z.$$

3.6. *Spherical Harmonics and Legendre Functions*

Solutions of (1–11) are called *spherical harmonics*. We have seen that it has solutions of the form

$$T(\theta,\varphi) = \Theta(\theta)(A\cos m\varphi + B\sin m\varphi), \tag{6–1}$$

where Θ is a solution of the associated Legendre equation

$$\sin\theta\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) + [n(n+1)\sin^2\theta - m^2]\Theta = 0. \tag{6–2}$$

For nonintegral values of n two independent solutions of (2) are $P_n^m(\cos\theta)$ and $P_n^m(-\cos\theta)$, where

$$P_n^m(\cos\theta) = 2^m\cos\frac{(n+m)\pi}{2}\,\frac{\left(\frac{n+m-1}{2}\right)!}{\left(\frac{n-m}{2}\right)!\left(-\frac{1}{2}\right)!}\sin^m\theta \times F\left(\frac{m+n+1}{2}, \frac{m-n}{2}; \frac{1}{2}; \cos^2\theta\right) +$$
$$+ 2^{m+1}\sin\frac{(n+m)\pi}{2}\,\frac{\left(\frac{n+m}{2}\right)!}{\left(\frac{n-m-1}{2}\right)!\left(-\frac{1}{2}\right)!}\sin^m\theta\cos\theta \times F\left(\frac{m+n+2}{2}, \frac{m-n+1}{2}; \frac{3}{2}; \cos^2\theta\right). \tag{6–3}$$

The hypergeometric function $F(\alpha,\beta;\gamma;x)$ appearing above is defined by the following power series

$$F(\alpha,\beta;\gamma;x) = 1 + \frac{\alpha\cdot\beta}{1\cdot\gamma}x + \frac{\alpha(\alpha+1)\beta(\beta+1)}{1\cdot 2\gamma(\gamma+1)}x^2 + \cdots. \tag{6–4}$$

This series is convergent for $|x| < 1$.

The series (3) represents Hobson's definition of P_n^m for *real values of the argument*, less than unity in absolute value. If we set

$$\mu = \cos\theta \tag{6–5}$$

and regard μ as a complex variable, the general function $P_n^m(\mu)$ has three branch points $\mu = 1$, $\mu = -1$ and $\mu = \infty$. In order to make this function single-valued a cut along the real axis is made from $\mu = -\infty$ to $\mu = +1$, along this cut P_n^m is discontinuous. On the cut itself, in the interval $-1 < \mu < 1$, the associated Legendre function is defined by

$$P_n^m(\mu) = e^{m\pi i/2}P_n^m(\mu + 0\cdot i) = e^{-m\pi i/2}P_n^m(\mu - 0\cdot i). \tag{6–6}$$

This definition makes P_n^m real in the interval $-1 < \mu < 1$.

For $m = 0$, we have

$$P_n(\cos\theta) = \sum_{s=0}^{\infty} \frac{(-)^s (n+s)!}{(n-s)!(s!)^2} \sin^{2s}\frac{\theta}{2}. \tag{6–7}$$

For integral values of m, we have

$$P_n^m(\cos\theta) = (-)^m \sin^m\theta \frac{d^m P_n(\cos\theta)}{d(\cos\theta)^m}. \tag{6–8}$$

In this case the function

$$T_n^m(\cos\theta) = (-)^m P_n^m(\cos\theta) = \sin^m\theta \frac{d^m P_n(\cos\theta)}{d(\cos\theta)^m} \tag{6–9}$$

is frequently used.

The function $P_n(\cos\theta)$ has a logarithmic singularity at $\theta = \pi$, and in this neighborhood the following series is more suitable for computing purposes

$$P_n(\cos\theta) = \frac{\sin n\pi}{\pi} \sum_{s=0}^{\infty} \frac{(-)^s(n+s)!}{(n-s)!} \left[2\log\cos\frac{\theta}{2} + \psi(n+s) \right.$$

$$\left. + \psi(n-s) - 2\psi(s) \right] \frac{\cos^{2s}\frac{\theta}{2}}{(s!)^2} + \cos n\pi \sum_{s=0}^{\infty} \frac{(-)^s(n+s)!}{(n-s)!(s!)^2} \cos^{2s}\frac{\theta}{2}. \tag{6–10}$$

For $\theta = 0$, we have

$$P_n(1) = 1. \tag{6–11}$$

Consequently the function $P_n(-\cos\theta)$ equals unity for $\theta = \pi$ and has a logarithmic singularity at $\theta = 0$.

When n is an integer $P_n(\cos\theta)$ becomes a polynomial in $\cos\theta$ and $P_n(-\cos\theta)$ differs from it only by a factor $(-)^n$. In this case the second solution of Legendre's equation is represented by another function $Q_n(\cos\theta)$, which is defined by

$$Q_n(\cos\theta) = P_n(\cos\theta)\log\cot\frac{\theta}{2}$$

$$-\left(P_{n-1}P_0 + \tfrac{1}{2}P_{n-2}P_1 + \tfrac{1}{3}P_{n-3}P_2 + \cdots + \frac{1}{n}P_0P_{n-1}\right), \tag{6–12}$$

$$Q_0(\cos\theta) = \log\cot\frac{\theta}{2}.$$

The second solution of the associated Legendre's equation is defined in terms of Q_n as follows

$$Q_n^m(\cos\theta) = (-)^m \sin^m\theta \frac{d^m Q_n(\cos\theta)}{d(\cos\theta)^m}. \tag{6–13}$$

This is the definition for real values of θ.

The Q-functions have two singular points $\theta = 0$ and $\theta = \pi$.

When $k = 0$, equation (1–11) for spherical harmonics becomes

$$\sin\theta \frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial T}{\partial\theta}\right) + \frac{\partial^2 T}{\partial\varphi^2} = 0. \tag{6–14}$$

The most general solution of this equation may be represented in the following form

$$T(\theta,\varphi) = \Phi\left(e^{i\varphi}\cot\frac{\theta}{2}\right) + \Psi\left(e^{-i\varphi}\cot\frac{\theta}{2}\right), \tag{6–15}$$

where Φ and Ψ are arbitrary functions. In particular we have solutions periodic in φ

$$T(\theta,\varphi) = {\cos \atop \sin}(m\varphi)\cot^m\frac{\theta}{2},\quad {\cos \atop \sin}(m\varphi)\tan^m\frac{\theta}{2}. \tag{6–16}$$

There are also solutions independent of φ

$$T(\theta) = \log\cot\frac{\theta}{2},\quad 1. \tag{6–17}$$

3.7. *Miscellaneous Formulae*

$$J_0'(x) = -J_1(x), \quad N_0'(x) = -N_1(x) \tag{7–1}$$

$$I_0'(x) = I_1(x), \quad K_0'(x) = -K_1(x) \tag{7–2}$$

$$e^{\pm\sigma\rho\sin\varphi} = I_0(\sigma\rho) + 2\sum_{n=1}^{\infty}(-)^n I_{2n}(\sigma\rho)\cos 2n\varphi \pm 2\sum_{n=0}^{\infty}(-)^n I_{2n+1}(\sigma\rho)\sin(2n+1)\varphi \tag{7–3}$$

$$e^{\pm\sigma\rho\cos\varphi} = I_0(\sigma\rho) + 2\sum_{n=1}^{\infty} I_n(\pm\sigma\rho)\cos n\varphi \tag{7–4}$$

$$e^{\pm i\beta\rho\sin\varphi} = J_0(\beta\rho) + 2\sum_{n=1}^{\infty} J_{2n}(\beta\rho)\cos 2n\varphi \pm 2i\sum_{n=0}^{\infty} J_{2n+1}(\beta\rho)\sin(2n+1)\varphi \tag{7–5}$$

$$e^{\pm i\beta\rho\cos\varphi} = J_0(\beta\rho) + 2\sum_{n=1}^{\infty}(-)^n J_{2n}(\beta\rho)\cos 2n\varphi \pm 2i\sum_{n=0}^{\infty}(-)^n J_{2n+1}(\beta\rho)\cos(2n+1)\varphi \tag{7–6}$$

$$\frac{1}{s^2}\int_{-s/2}^{s/2} dx \int_{-s/2}^{s/2} \log|x - \hat{x}|\, d\hat{x} = \log s - 1.5 \tag{7–7}$$

If $\beta s \ll 1$, then

$$\frac{1}{s^2}\int_{-s/2}^{s/2} dx \int_{-s/2}^{s/2} K_0(i\beta|x - \hat{x}|)\, d\hat{x} = \log\frac{\lambda}{s} - 0.222 - \frac{i\pi}{2} \tag{7–8}$$

$$\int_0^x x J_0(x)\, dx = x J_1(x) \tag{7–9}$$

$$\int_0^x x[J_n(kx)]^2\, dx = \frac{x^2}{2}\left[[J_n'(kx)]^2 + \left(1 - \frac{n^2}{k^2x^2}\right)J_n^2(kx)\right]$$

$$= \frac{x^2}{2}[-J_{n-1}(kx)J_{n+1}(kx) + J_n^2(kx)] \tag{7–10}$$

$$J_1(2.40) = 0.519, \quad 2.40 J_1(2.40) = 1.24 \tag{7–11}$$

$$K_n I_n' - K_n' I_n = I_{n+1} K_n + K_{n+1} I_n = \frac{1}{x} \tag{7–12}$$

$$J_n N_n' - J_n' N_n = J_{n+1} N_n - J_n N_{n+1} = \frac{2}{\pi x} \tag{7–13}$$

$$\frac{K_1(ix)}{K_0(ix)} = \frac{2}{\pi x (J_0^2 + N_0^2)} - i \frac{J_0 J_1 + N_0 N_1}{J_0^2 + N_0^2} \tag{7–14}$$

$$\frac{K_0(ix)}{K_1(ix)} = \frac{2}{\pi x (J_1^2 + N_1^2)} + i \frac{J_0 J_1 + N_0 N_1}{J_1^2 + N_1^2} \tag{7–15}$$

$$J_0^2 + N_0^2 \sim \frac{2}{\pi x} \sum_{m=0}^{\infty} \frac{(-)^m [1 \cdot 3 \cdot 5 \cdots (2m-1)]^4}{(2m)!(2x)^{2m}} \tag{7–16}$$

$$J_\nu^2 + N_\nu^2 \sim \frac{2}{\pi x} \sum_{m=0}^{\infty} [1 \cdot 3 \cdot 5 \cdots (2m-1)] \frac{(\nu, m)}{2^m x^{2m}} \tag{7–17}$$

$$\mathrm{Si}\, x = \int_0^x \frac{\sin t}{t}\, dt, \quad \mathrm{Ci}\, x = \int_\infty^x \frac{\cos t}{t}\, dt, \quad x > 0 \tag{7–18}$$

$$\int_0^x \frac{1 - \cos t}{t}\, dt = C + \log x - \mathrm{Ci}\, x, \quad \mathrm{Si}\, \infty = \frac{\pi}{2} \tag{7–19}$$

$$\mathrm{Si}\, x = x - \frac{x^3}{3!\,3} + \frac{x^5}{5!\,5} - \cdots \tag{7–20}$$

$$\mathrm{Ci}\, x = \log x + C - \frac{x^2}{2!\,2} + \frac{x^4}{4!\,4} - \frac{x^6}{6!\,6} + \cdots, \quad C = 0.577 \cdots \tag{7–21}$$

$$\mathrm{Si}\, x \sim \frac{\pi}{2} - \frac{\cos x}{x} \sum_{n=0}^{\infty} \frac{(-)^n (2n)!}{x^{2n}} - \frac{\sin x}{x} \sum_{n=0}^{\infty} \frac{(-)^n (2n+1)!}{x^{2n+1}} \tag{7–22}$$

$$\mathrm{Ci}\, x \sim \frac{\sin x}{x} \sum_{n=0}^{\infty} \frac{(-)^n (2n)!}{x^{2n}} - \frac{\cos x}{x} \sum_{n=0}^{\infty} \frac{(-)^n (2n+1)!}{x^{2n+1}} \tag{7–23}$$

$$Ei(\pm x) = \mathrm{Ci}\, x \pm i\, \mathrm{Si}\, x \tag{7–24}$$

$$\mathrm{Si}\,(-x) = -\,\mathrm{Si}\, x \tag{7–25}$$

$$\int_0^{-x} \frac{1 - \cos t}{t}\, dt = \int_0^x \frac{1 - \cos t}{t}\, dt \tag{7–26}$$

$$\int_0^\infty J_n(x)\, dx = 1 \tag{7–27}$$

$$\int_0^\infty \frac{J_n(kx)}{x}\, dx = \frac{1}{n}, \quad n = 1, 2, 3, \cdots \tag{7–28}$$

$$\int_0^x \frac{J_1(x)}{x}\, dx \sim 1 - \sum_{n=1}^{\infty} \frac{1 \cdot 3 \cdot 5 \cdots (2n-3) J_{n-1}(x)}{x^n} \tag{7–29}^*$$

$$\int_x^\infty \frac{J_n(x)}{x^m}\, dx \sim -\frac{J_{n+1}(x)}{x^m} - \frac{(m+n+1) J_{n+2}(x)}{x^{m+1}} - \frac{(m+n+1)(m+n+3) J_{n+3}(x)}{x^{m+2}} - \cdots \tag{7–30}$$

$$\int_0^x \exp\left(\frac{i\pi t^2}{2}\right) dt = C(x) + iS(x) \tag{7–31}$$

$$\int_0^x \cos \frac{\pi t^2}{2}\, dt = C(x), \quad \int_0^x \sin \frac{\pi t^2}{2}\, dt = S(x) \tag{7–32}$$

$$C(0) = S(0) = 0, \quad C(\infty) = S(\infty) = 0.5 \tag{7–33}$$

$$C(x) = \sum_{n=0}^{\infty} \frac{(-)^n \pi^{2n} x^{4n+1}}{2^{2n}(2n)!(4n+1)}, \quad S(x) = \sum_{n=0}^{\infty} \frac{(-)^n \pi^{2n+1} x^{4n+3}}{2^{2n+1}(2n+1)!(4n+3)} \tag{7–34}$$

$$C(x) \sim 0.5 + P(x) \cos \frac{\pi x^2}{2} - Q(x) \sin \frac{\pi x^2}{2} \tag{7–35}$$

$$S(x) \sim 0.5 + P(x) \sin \frac{\pi x^2}{2} + Q(x) \cos \frac{\pi x^2}{2} \tag{7–36}$$

$$P(x) = \frac{1}{\pi x} \sum_{n=0}^{\infty} \frac{(-)^{n+1} 1 \cdot 3 \cdot 5 \cdots (4n+1)}{(\pi x^2)^{2n+1}},$$
$$Q(x) = \frac{1}{\pi x} \sum_{n=0}^{\infty} \frac{(-)^{n+1} 1 \cdot 3 \cdot 5 \cdots (4n-1)}{(\pi x^2)^{2n}} \tag{7–37}^\dagger$$

$$C(x) = \sum_{n=0}^{\infty} J_{2n+(1/2)}(\tfrac{1}{2}\pi x^2), \quad S(x) = \sum_{n=0}^{\infty} J_{2n+(3/2)}(\tfrac{1}{2}\pi x^2) \tag{7–38}$$

$$\int_{z_1}^{z_2} \frac{\cos \beta(r+z)}{r}\, dz = \mathrm{Ci}\, \beta(r_2 + z_2) - \mathrm{Ci}\, \beta(r_1 + z_1)$$
$$r = \sqrt{\rho^2 + z^2}, \quad r_1 = \sqrt{\rho^2 + z_1^2}, \quad r_2 = \sqrt{\rho^2 + z_2^2} \tag{7–39}$$

$$\int_{z_1}^{z_2} \frac{\sin \beta(r+z)}{r}\, dz = \mathrm{Si}\, \beta(r_2 + z_2) - \mathrm{Si}\, \beta(r_1 + z_1) \tag{7–40}$$

$$\int_{z_1}^{z_2} \frac{\cos \beta(r-z)}{r}\, dz = \mathrm{Ci}\, \beta(r_1 - z_1) - \mathrm{Ci}\, \beta(r_2 - z_2) \tag{7–41}$$

$$\int_{z_1}^{z_2} \frac{\sin \beta(r-z)}{r}\, dz = \mathrm{Si}\, \beta(r_1 - z_1) - \mathrm{Si}\, \beta(r_2 - z_2) \tag{7–42}$$

* For $n = 1$ the numerical coefficient is unity.
† The first term of $Q(x)$ is $-1/\pi x$.

$$P_{n+\Delta}(\cos\theta) = P_n(\cos\theta) + 2\Delta\left[P_n(\cos\theta)\log\cos\frac{\theta}{2} + S_n'\right] \quad (7\text{–}43)^*$$

$$P_{n+\Delta}(-\cos\theta) = (-)^n P_n(\cos\theta) + 2\Delta\left[(-)^n P_n(\cos\theta)\log\sin\frac{\theta}{2} + S_n''\right] \quad (7\text{–}44)^*$$

$$S_n' = \sum_{\alpha=1}^{n}\frac{(-)^\alpha(n+\alpha)!}{\alpha!\alpha!(n-\alpha)!}\left(\frac{1}{n+\alpha}+\frac{1}{n+\alpha-1}+\cdots\frac{1}{n+1}\right)\sin^{2\alpha}\frac{\theta}{2} \quad (7\text{–}45)^*$$

$$S_n'' = \sum_{\alpha=1}^{n}\frac{(-)^\alpha(n+\alpha)!}{\alpha!\alpha!(n-\alpha)!}\left(\frac{1}{n+\alpha}+\frac{1}{n+\alpha-1}+\cdots\frac{1}{n+1}\right)\cos^{2\alpha}\frac{\theta}{2} \quad (7\text{–}46)^*$$

$$\sum_{n=1}^{\infty}\frac{\sin nx}{n} = \frac{\pi - x}{2}, \quad 0 < x < 2\pi \quad (7\text{–}47)$$

$$\sum_{n=1}^{\infty}\frac{\cos nx}{n} = -\tfrac{1}{2}\log 2(1-\cos x) = -\log\left(2\sin\frac{x}{2}\right), \quad 0 < x < 2\pi \quad (7\text{–}48)$$

$$\sum_{n=1}^{\infty}\frac{\cos nx}{n^2} = \frac{\pi^2}{6} - \frac{\pi x}{2} + \frac{x^2}{4}, \quad 0 \le x < 2\pi \quad (7\text{–}49)$$

$$\sum_{n=1}^{\infty}\frac{1-\cos nx}{n^2} = \frac{\pi x}{2} - \frac{x^2}{4}, \quad 0 \le x < 2\pi \quad (7\text{–}50)$$

$$\sum_{n=1}^{\infty}\frac{\sin nx}{n^3} = \frac{\pi^2 x}{6} - \frac{\pi x^2}{4} + \frac{x^3}{12}, \quad 0 \le x < 2\pi \quad (7\text{–}51)$$

$$\sum_{n=1}^{\infty}\frac{\sin nx}{n^2} = -x\log x + x + \frac{1}{9}\left(\frac{x}{2}\right)^3 + \frac{1}{450}\left(\frac{x}{2}\right)^5 + \cdots, \quad 0 \le x < 2\pi \quad (7\text{–}52)$$

$$\sum_{n=1}^{\infty}\frac{1-\cos nx}{n^3} = -\tfrac{1}{2}x^2\log x + \tfrac{3}{4}x^2 + \frac{x^4}{288} + \cdots, \quad 0 \le x < 2\pi \quad (7\text{–}53)$$

$$F(\theta,\varphi) = \sum_{n=0}^{\infty} a_n P_n(\mu) + \sum_{n=0}^{\infty}\sum_{m=1}^{n}(a_{n,m}\cos m\varphi + b_{n,m}\sin m\varphi)P_n^m(\mu), \quad (7\text{–}54)$$

$$a_n = \frac{2n+1}{4\pi}\int_0^{2\pi}\int_{-1}^{1} P_n(\hat\mu)F(\hat\theta,\hat\varphi)\,d\hat\mu\,d\hat\varphi, \quad \mu = \cos\theta, \quad \hat\mu = \cos\hat\theta,$$

$$a_{n,m} = \frac{2n+1}{2\pi}\frac{(n-m)!}{(n+m)!}\int_0^{2\pi}\int_{-1}^{1} P_n^m(\hat\mu)\cos m\hat\varphi F(\hat\theta,\hat\varphi)\,d\hat\mu\,d\hat\varphi, \quad (7\text{–}55)$$

$$b_{n,m} = \frac{2n+1}{2\pi}\frac{(n-m)!}{(n+m)!}\int_0^{2\pi}\int_{-1}^{1} P_n^m(\hat\mu)\sin m\hat\varphi F(\hat\theta,\hat\varphi)\,d\hat\mu\,d\hat\varphi$$

$$e^{-\Gamma r\cos\theta} = \frac{1}{\Gamma r}\sum_{n=0}^{\infty}(2n+1)\hat I_n(\Gamma r)P_n(\cos\theta) \quad (7\text{–}56)$$

* In equations (43),(44), (45) and (46) n is an integer.

$$e^{-i\beta r\cos\theta} = -\frac{1}{\beta r}\sum_{n=0}^{\infty}(2n+1)(-i)^n\hat J_n(\beta r)P_n(\cos\theta) \quad (7\text{–}57)$$

$$e^{i\beta r\cos\theta} = \frac{1}{\beta r}\sum_{n=0}^{\infty}(2n+1)i^n\hat J_n(\beta r)P_n(\cos\theta) \quad (7\text{–}58)$$

$$\frac{\exp\left[-\Gamma\sqrt{r^2+a^2-2ar\cos\theta}\right]}{\sqrt{r^2+a^2-2ar\cos\theta}} = \frac{1}{\Gamma ar}\sum_{n=0}^{\infty}(2n+1)\hat K_n(\Gamma a)\hat I_n(\Gamma r)P_n(\cos\theta), \quad r < a \quad (7\text{–}59)$$

$$\frac{\exp\left[-i\beta\sqrt{r^2+a^2-2ar\cos\theta}\right]}{\sqrt{r^2+a^2-2ar\cos\theta}} = \frac{1}{i\beta ar}\sum_{n=0}^{\infty}(2n+1)[\hat J_n(\beta a) - i\hat N_n(\beta a)]\hat J_n(\beta r)P_n(\cos\theta), \quad r < a \quad (7\text{–}60)$$

CHAPTER IV

Fundamental Electromagnetic Equations

4.1. *Fundamental Equations in the MKS System of Units*

It is assumed that the reader is familiar with fundamental electromagnetic concepts. An excellent description of a set of experiments underlying these concepts and the laws of electromagnetic induction may be found in the first four chapters of "Physical Principles of Electricity and Magnetism" by R. W. Pohl.* We shall use exclusively the meter-kilogram-second-coulomb system of units, commonly known as the Giorgi or the MKS system. In this system electromagnetic equations are particularly simple and correspond closely to physical ideas and measurements, common in engineering laboratories. Table I gives a list of quantities, symbols, names of units, and some dimensional equivalents of these units.

The definitions of common electromagnetic terms may be summarized as follows. The basic idea of electric charge has to be gained from experience. When two bodies are electrified certain forces between them are attributed to the "electric charges" in them and the force E per unit "positive" charge is called the *electric intensity*. Electricity appears to be atomic in structure and the smallest particle of negative charge is called the *electron*. In some substances, called *conductors*, there are many "free" electrons, easily detachable from atoms; in such substances the *electric current density* J, defined as the time rate of flow of electric charge per unit area normal to the lines of flow, is proportional to the electric intensity (Ohm's law); thus

$$J = gE,$$

where the coefficient of proportionality is the *conductivity* of the substance. By definition, the direction of the electric current coincides with the direction of moving positive charge and is opposite to the direction of moving negative charge.

The *electromotive force* V (or the "voltage") acting along a path joining points P and Q is defined as the line integral of the electric intensity

$$V = \int_{(PQ)} E_s \, ds. \tag{1-1}$$

* Blackie & Son Limited (1930).

Table I

Name of Quantity	Symbol	Name of Unit	Dimensional Equivalent
Length	—	meter	—
Mass	—	kilogram	—
Time	t	second	—
Energy	—	joule	volt-coulomb, newton-meter
Power	—	watt	joule per second
Force	—	newton	joule per meter
Electric charge, electric displacement	q	coulomb	ampere-second
Displacement density	D	coulomb per meter^2	
Electric current	I	ampere	coulomb per second
Current density	J	ampere per meter^2	
Electromotive force	V	volt	joule per coulomb
Electric intensity	E	volt per meter	newton per coulomb
Impedance (electric)	Z,K	ohm	volt per ampere
Admittance (electric)	Y,M	mho	ampere per volt
Inductance	L	henry	ohm-second
Permeability	μ	henry per meter	
Capacitance	C	farad	
Dielectric constant	ϵ	farad per meter	
Conductivity	g	mho per meter	
Magnetomotive force	U	ampere	
Magnetic intensity	H	ampere per meter	
Magnetic flux	Φ	weber	volt-second
Magnetic charge	m	weber	volt-second
Magnetic flux density	B	weber per meter^2	
Magnetic current	K[1]	volt	
Magnetic current density	M[2]	volt per meter^2	

[1] It will be clear from the context whether K stands for magnetic current or impedance.

[2] It will be clear from the context whether M designates magnetic current density or admittance.

Thus the electromotive force represents the work done by E on a unit positive charge moving along PQ and the work done on a charge q is Vq. The electromotive force is not a true force in the usual mechanical sense.* Consider now a homogeneous conducting rod of length l and let its cross-

* The electromotive force may be regarded as the *generalized force* and the electric charge as the *generalized coordinate* in Lagrange's sense.

section be S; then the current I in the rod is SJ, the electromotive force is lE, and therefore

$$I = GV, \quad G = \frac{gS}{l}; \quad V = RI, \quad R = \frac{l}{gS}.$$

The coefficients of proportionality R and G are called respectively the *resistance* and the *conductance* of the rod. The work done by V per second is Vq/t or $VI = GV^2$ watts; this work appears as heat. The electric power dissipated in heat per unit volume is evidently $JE = gE^2$.

The magnitudes of the volt and the ampere have been chosen to suit the most common experience. Thus the voltages supplied to the homes for lighting and cooking purposes are between 110 and 120 volts; the current through a 100 watt electric bulb is about 0.9 ampere and the resistance of the bulb is about 122 ohms. For the mechanical units, similarly, the weight of 1 kilogram is about 9.8 newtons; the potential energy of a kilogram-mass 1 meter above ground is 9.8 joules; and the kinetic energy of 1 kilogram moving with a velocity equal to 1 meter per second is half a joule.

A *perfect dielectric* is a medium possessing no detachable electric particles; in such a medium $g = 0$. Vacuum is the only physical example of a perfect dielectric; in the first approximation, however, many other media may be treated as perfect dielectrics. Consider now the field produced by electric charges placed in a perfect dielectric. If a neutral conductor is introduced in the field, the free electrons will move under the influence of E until inside the conductor the electric intensity of the separated or " induced " charges becomes equal to $-E$. At the surface of the conductor the total tangential component of the electric intensity must be zero or the electrons would still be moving; on the other hand, the normal component will be different from zero. In order to explore this " electrostatic induction " quantitatively we may take a conductor formed by two separable thin flat discs of small area and insert it at some point in the field. The discs are then separated, removed from the field, and the charges on them measured. It will be found that these charges are equal and opposite and that their magnitudes are proportional to the product of the electric intensity, the area of the discs, and the cosine of the angle formed by the normal to the discs with some fixed direction. The maximum charge per unit area is called the *displacement density* D at the point in question. Generally the components of D are linear functions of the components of E and the directions of D and E are different; but in *isotropic* media D and E are in the same direction and

$$D = \epsilon E.$$

The coefficient of proportionality ϵ is called the *dielectric constant*. In this book we are concerned only with isotropic media.

For example, consider two parallel conducting plates with equal and opposite charges. If the distance between the plates is small compared with the dimensions of the plates, the electric field between the plates is nearly uniform except near the edges and electric lines (lines tangential

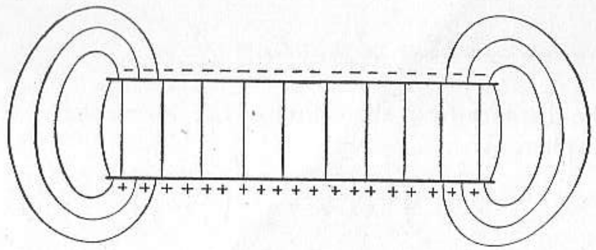


FIG. 4.1. Electric lines of force between the plates of a capacitor.

to E) look like those in Fig. 4.1. Thus experience indicates that between two infinite uniformly charged plates the field would be uniform and the charges on the two conducting plates introduced in the field would separate as shown in Fig. 4.2. The displacement density is found to be equal to the electric charge per unit area on the positively charged plate.

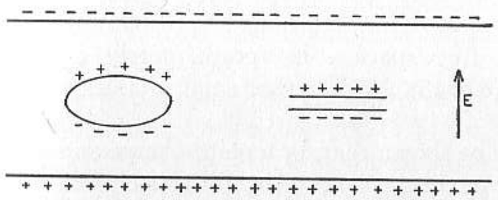


FIG. 4.2. Electrostatic induction in the field between two equally and oppositely charged planes.

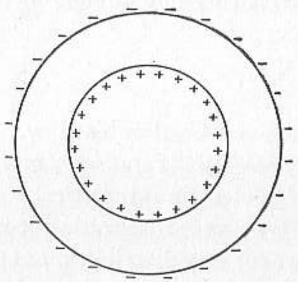


FIG. 4.3. Concentric spheres.

In his original experiments on electrostatic induction Faraday employed two metal spheres (Fig. 4.3). He placed a fixed charge q on one sphere and then enclosed it within the other. After connecting the outer sphere to ground, he measured the charge remaining on the sphere. He found that this charge is equal and opposite to the charge on the inner sphere regardless of the dimensions of the spheres and the medium between them. In the case of concentric spheres electric lines of force are radial and therefore the total *displacement* through any sphere concentric with the sphere containing q is

$$\iint D_n \, dS = q. \qquad (1\text{–}2)$$

The voltage is found to be proportional to q; hence D is proportional to E. In the region between concentric spheres we have then

$$D_r = \frac{q}{4\pi r^2}, \quad E_r = \frac{q}{4\pi \epsilon r^2}, \qquad (1\text{–}3)$$

where r is the distance from the center of the inner sphere. The voltage between the spheres is

$$V = \int_a^b E_r \, dr = \frac{q}{4\pi\epsilon}\left(\frac{1}{a} - \frac{1}{b}\right).$$

The ratio q/V is called the *capacitance*; hence the capacitance of two concentric spheres is

$$C = \frac{4\pi\epsilon ab}{b - a}.$$

When the outer sphere is removed to infinity, then $C = 4\pi\epsilon a$.

Let the outer sphere be infinite and the inner vanishingly small. Then any other "point charge" q' will not affect the field of q and the force exerted by one charge on the other is

$$F = \frac{qq'}{4\pi\epsilon r^2}.$$

This is Coulomb's law. In free space ϵ is approximately equal to $(1/36\pi)10^{-9}$ farads per meter (8.854×10^{-12}); thus small charges produce large electric intensities.

By direct integration it may be shown that for a point charge and therefore for any distribution of charge the displacement through a closed surface is equal to the enclosed charge so that (2) is true regardless of the shape of (S).

An electric current exerts a force on each end of a compass needle and thus is surrounded by a *magnetic field*. In the case of two coaxial cylindrical sheets or two plane sheets, carrying equal and oppositely directed steady currents (Figs. 4.4*a* and 4.4*b*), the field is confined to the region between the cylinders or the planes. In the first case magnetic lines (that is, lines tangential to the forces on the ends of the compass needle) are circles coaxial with the cylinders and in the second case they are straight lines parallel to the planes and perpendicular to the direction of current flow. The

arrows in Fig. 4.4 indicate the direction of the force on the north-seeking or the "positive" end of the needle; $+I$ is supposed to be flowing toward the reader. Between two plane current sheets the *magnetic intensity* H is uniform; it depends on the linear current density in the sheets but not on the distance between them. This linear current density is taken as the

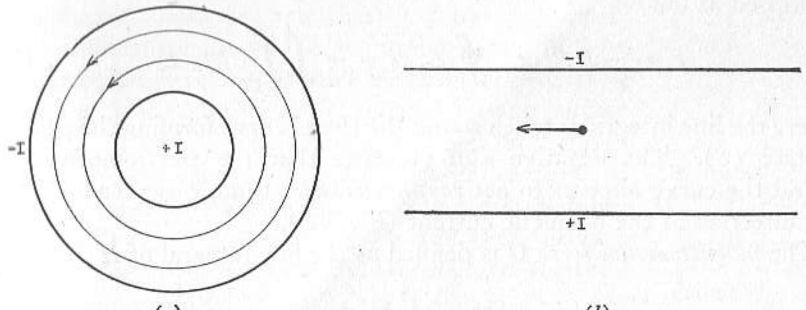


FIG. 4.4. The magnetic intensity between coaxial cylinders and parallel planes.

measure of H; hence the unit of H is the ampere per meter. Similarly the magnetic field is uniform inside a closely wound cylindrical coil; the magnetic intensity is independent of the shape and size of the cross-section of the cylinder, is parallel to the generators of the cylinder, and is equal to the circulating current per unit length of the coil. In this case H is also equal to the number of "ampere-turns," that is, to the product of the current in the wire and the number of turns per unit length of the coil.

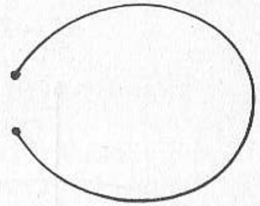


FIG. 4.5. A wire loop

It is an experimental fact, discovered by Faraday, that a voltage exists across the terminals of a loop (Fig. 4.5) in a varying magnetic field. For a small loop this voltage is proportional to the cosine of the angle between H and the axis of the loop; the maximum voltage is proportional to the product of the area S of the loop and the time derivative of H; the coefficient of proportionality μ depends on the medium and is called the *permeability*. The time integral of the voltage is called the *magnetic flux* or the *magnetic displacement* Φ through the loop; thus for a small locp

$$\Phi = SB, \quad B = \mu H, \qquad (1\text{–}4)$$

where B is called the *magnetic flux density*. The magnetic flux through any surface is defined as the surface integral of the normal component of B

$$\Phi = \iint_{(S)} B_n \, dS. \qquad (1\text{–}5)$$

The following time derivatives

$$K = \frac{\partial \Phi}{\partial t}, \quad M = \frac{\partial B}{\partial t}, \tag{1–6}$$

are called respectively the *magnetic current* and the *magnetic current density*. The first law of electromagnetic induction (Faraday's law) may then be expressed as follows

$$V = -K, \quad \text{or} \quad \oint E_s\, ds = -\iint_{(S)} M_n\, dS, \tag{1–7}$$

where the line integral is taken round the closed curve forming the edge of a surface (S). The negative sign indicates that the electromotive force round the curve appears to act *counterclockwise* to an observer looking in the direction of the magnetic current (Fig. 4.6).

The *magnetomotive force* U is defined as the line integral of H

$$U = \int H_s\, ds. \tag{1–8}$$

The second law of electromagnetic induction (Ampère's law, amended by Maxwell) may then be expressed as follows

$$U = I, \quad \text{or} \quad \oint H_s\, ds = \iint_{(S)} J_n\, dS; \tag{1–9}$$

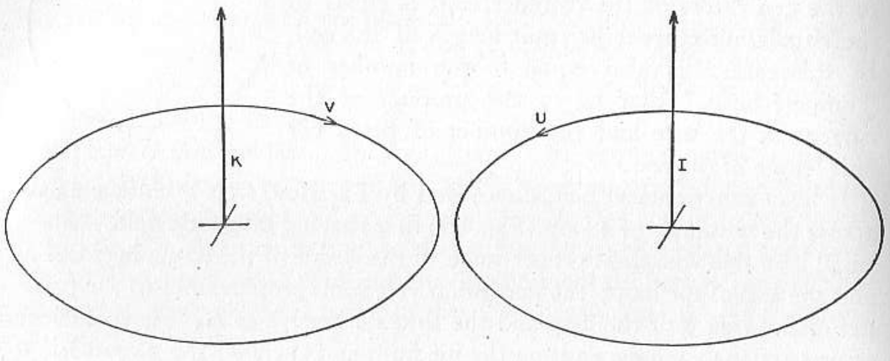


FIG. 4.6. Relative directions between magnetic and electric currents and electromotive and magnetomotive forces induced by them.

that is, the magnetomotive force round a closed curve equals the electric current passing through any surface bounded by the curve. The magnetomotive force appears clockwise to an observer looking in the direction of the current (Fig. 4.6).

These two fundamental laws of induction form the basis of electromagnetic wave theory. They are abstractions from experiments performed

under restricted conditions and as such they are postulates which may be considered valid as long as theoretical conclusions derived from them agree with all available experimental evidence.

Inside a conductor the electric current density is gE; the magnetomotive forces may be calculated from this current density only if the surface (S) in equation (9) lies completely inside the conductor. Suppose now that we connect two electrically charged spheres with a conducting wire (Fig. 4.7). The magnetomotive force round a small loop encircling the wire is I; if equation (9) is to hold for any surface such as (S) in Fig. 4.7 through which no conduction current is flowing we should include in J another term. If the positive direction of I is chosen to be from the lower sphere to the upper, then $I = \partial q/\partial t$ and there exists a time rate of change of electric displacement through (S) which, when regarded as an "electric current," is just sufficient to give the right value of the magnetomotive force. Thus we define the *displacement current density*

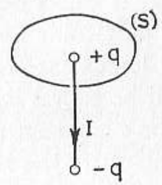


FIG. 4.7. The time derivative of the inward displacement through (S) is equal to the outward conduction current.

$$J_d = \frac{\partial D}{\partial t} = \epsilon \frac{\partial E}{\partial t}. \tag{1–10}$$

Displacement currents must be included if (9) is to be true irrespective of the choice of (S); but, of course, there is no *à priori* physical reason requiring that (9) should be so general and only experiment can decide if it is. If (9) is really general, we should expect magnetomotive forces to exist where there are no other currents but displacement currents; this happens to be the case. Consider for example a circular loop round which we measure the magnetomotive force U and an electric charge q moving on the axis of this loop (Fig. 4.8); the displacement through the plane of the loop increases until the charge reaches this plane. If q comes from a great distance, the time integral $\int U\, dt$ is $\frac{1}{2}q$; as q passes on to a great distance the time integral becomes* q.

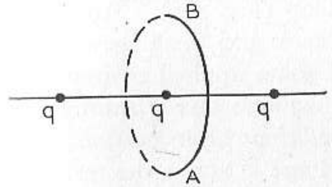


FIG. 4.8. An electric charge moving on the axis of a circular loop.

For complete generality a third term should be included in the current density in (9). Imagine a cloud of electric charge moving with a velocity v;

* The absolute value of D_n diminishes but its sign is negative.

if the volume density of electric charge is q_r the current density is

$$J^i = q_r v. \tag{1–11}$$

This is the *convection current* and it differs from the conduction current in that it exists outside conductors and its density is not proportional to E. In particular we shall be interested in *forced* convection currents produced by forces external to the field (chemical, mechanical, etc.); such currents will be called *impressed currents*. In our equations such currents provide a link between an electromagnetic field and its external cause. Thus we shall write the total electric current density in the following form

$$J = gE + \epsilon \frac{\partial E}{\partial t} + J^i. \tag{1–12}$$

The first term alone exists in conductors;* in perfect dielectrics we ordinarily have only the second term; in ordinary good dielectrics this term predominates but a small conduction current term will usually exist; and finally in "electric generators" we have J^i. In electrolytic cells and vacuum tubes we have convection currents; in a dynamo the current is of conduction type but its density is not equal to gE, where E is the electric intensity of the electromagnetic field. In a dynamo the current density is $g(E + E^i)$, where E^i is the "motional electromotive force" which is an example of an *impressed electromotive force*. If the wires of the dynamo were "perfect conductors," g would be infinite and $E + E^i$ would equal zero; J^i, however, would be finite.

In wave theory we use J^i for closing circuits and introducing hypothetical generators of simplified type. Suppose we wish to find what happens when a variable current is flowing through a wire loop (Fig. 4.5). We assume that the electric charge is being transferred back and forth between the terminals of the loop under the influence of some applied or impressed force. The current in the loop produces a magnetic flux Φ through the loop; this flux is proportional to I and the coefficient of proportionality L is called the inductance of the loop. The voltage V across the terminals of the loop is $-\partial\Phi/\partial t$ and the impressed force V^i needed for the transfer of charge between the terminals against this "electromotive force of self-induction" is

$$V^i = -V = L\frac{\partial I}{\partial t}. \tag{1–13}$$

This is not the total impressed force needed for the transfer of charge; the wire has a resistance and an "internal" inductance which will be considered in a later chapter.

Physically there is only one type of magnetic current, namely the displacement current defined as the time rate of change of magnetic displacement or magnetic flux. Compass needles exert forces on each other consistent with an assumption that there exist "magnetic charges" near the ends of the needles and that the forces between these charges are similar to those between electric charges; but efforts to separate magnetic charges have consistently failed and all present experimental evidence indicates that there are no physical magnetic charges and that consequently there can be no physical magnetic conduction and convection currents. On some occasions, nevertheless, considerable mathematical simplifications may be secured if we write Maxwell's equations in a more symmetrical form by including hypothetical magnetic charges and corresponding convection currents. These should be defined to be analogous to electric charges and currents in order to preserve whatever symmetry there is in Maxwell's equations. The field of a magnetic "pole" is assumed to be radial and the magnitude m of the magnetic charge is assumed to be equal to the magnetic flux emerging from the pole. If then in Fig. 4.8 the moving electric particle is replaced by a magnetic charge m, the time integral of the electromotive force round the loop AB as m passes from $-\infty$ to $+\infty$ will be m. The magnetic convection current density M^i is defined as equal to $m_r v$ where m_r is the volume density and v the velocity of the charge. The total magnetic current density will be expressed as

$$M = \frac{\partial B}{\partial t} + M^i = \mu\frac{\partial H}{\partial t} + M^i. \tag{1–14}$$

In general therefore we shall write the fundamental electromagnetic equations in the following form

$$\oint E_s\, ds = -\iint \mu\frac{\partial H_n}{\partial t}\, dS - \iint M_n\, dS,$$
$$\oint H_s\, ds = \iint \left(gE_n + \epsilon\frac{\partial E_n}{\partial t}\right) dS + \iint J_n\, dS, \tag{1–15}$$

where the conduction and displacement current densities are included explicitly and the superscripts designating the impressed current densities have been dropped.

Since we are concerned primarily with fields varying harmonically with time, we replace the instantaneous field intensities and current densities by the corresponding complex variables and write Maxwell's equations as follows

$$\int E_s\, ds = -\iint i\omega\mu H_n\, dS - \iint M_n\, dS,$$
$$\int H_s\, ds = \iint (g + i\omega\epsilon)E_n\, dS + \iint J_n\, dS. \tag{1–16}$$

* Except at optical frequencies when the second is appreciable.

One more remark about magnetic charges. The force on a magnetic charge in a magnetic field is assumed to be Hm. The dimensions are correct but a numerical factor might have been included. The torque on a *magnetic doublet* (two charges m, $-m$ separated by distance l) of *moment* ml placed normally to the lines of force is then Hml. In Chapter 6 we shall find that with the above definitions in mind the field of a magnetic doublet is the same as that of an elementary electric current loop provided $ml = \mu SI$, where I is the current and S is the area of the loop. The torque on the loop whose axis is normal to the magnetic lines would seem to be $\mu HSI = BSI$; this happens to be the case and we have a machinery for replacing in calculations circulating electric currents by equivalent magnetic doublets. Coulomb's law for the force between magnetic charges as above defined is evidently $m_1m_2/4\pi\mu r^2$.

Maxwell's equations in the form in which we have expressed them possess considerable symmetry; E and H correspond to each other, the first being measured in volts per meter and the second in amperes per meter; D and B correspond to each other, the first being measured in ampere-seconds per square meter and the second in volt-seconds per square meter; electric and magnetic currents correspond to each other, the first being measured in amperes and the second in volts. In literature one finds arguments to the effect that " physically " E and B (and D and H) are similar and that B is more " basic " than H. All such arguments seem sterile since electric and magnetic quantities are physically different; whatever similarity there is comes from the equations.

4.2. *Impressed Forces*

In equations (1–15) and (1–16) electric generators are represented by the current densities J and M whose values are supposed to be given. Of course, in order to obtain given values of the generator currents we must have properly distributed *impressed intensities* which sustain these currents against the forces of the electromagnetic field. These impressed intensities are not included in E and H in equations (1–15) and (1–16); they are equal and opposite to the field intensities against which the charges in the generator have to be moved.

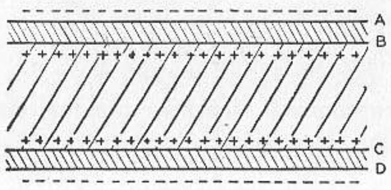


FIG. 4.9. A diagram of an electrolytic cell.

For example, in an electrolytic cell there exist contact forces tending to separate positive and negative charges. In Fig. 4.9 the lightly shaded region represents an electrolyte and the heavily shaded regions are greatly

magnified regions of contact between the electrolyte and the metal plates A and D. The contact forces move the negative charge to the metal plates until the electromotive forces of the separated charges just balance the contact electromotive forces. Contact forces are different for different metals and when *two* plates A and D are submerged there may exist a net impressed electromotive force between A and D (Fig. 4.10) which is equal

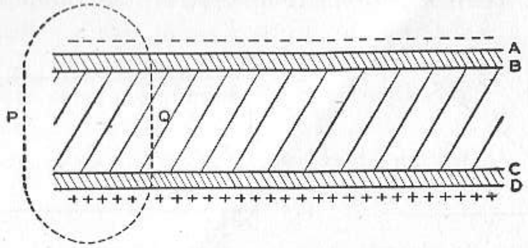


FIG. 4.10. The electromotive force between the metal plates of the cell due to the electric charges separated within it is the same for any path either inside or outside the cell. Inside the cell there are also contact forces acting on the charges.

and opposite to the voltage V_{AQD} produced by the separated charges. The voltage $V_{APD} = -V_{AQD}$ and the total electromotive force of the field round the closed path is zero, which is consistent with (1–15) since there is no magnetic current through the circuit. If A and D are connected by a conducting wire, the electric intensity of the field of separated charges will move the electrons in the wire and a current will be produced.

As another example let us take a fixed pair of wires AB and CD and a sliding wire MN of length l equal to the separation between AB and CD (Fig. 4.11). Let there be a uniform magnetic field H perpendicular to the plane of the paper and directed toward the reader. Let the velocity of MN be v. Consider a circuit $MNPM$ of which MN is the only moving part. In time dt the magnetic flux through this circuit changes by $Blv\,dt$ and the electromotive force in the circuit due to the motion of MN is

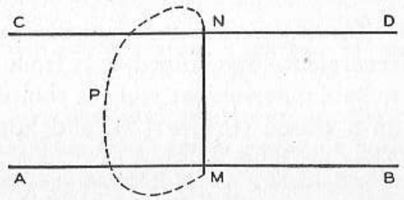


FIG. 4.11. A conducting wire sliding along a pair of parallel wires in a magnetic field.

$$V^i = -Blv.$$

This force is independent of the fixed part of the circuit and hence resides in the moving wire MN; for this reason it is called the *motional electromotive force*. The relative directions of H, v and the motional E are shown in Fig. 4.12. More generally the force on a charge moving with velocity v

in a magnetic field is

$$F = qv \times B.$$

In subsequent work we shall assume as given either impressed or generator currents or impressed electromotive forces, according to circumstances. In a pair of parallel wires (Fig. 4.13) for instance, we may start with a given current I^i flowing from B to A, determine the charge and

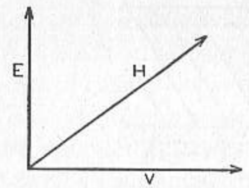


FIG. 4.12. The relative directions of the motional electric intensity, the magnetic intensity, and the velocity.

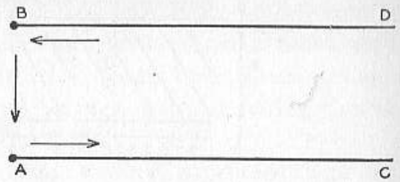


FIG. 4.13. A diagram illustrating the simplified conception of an electric generator in wave theory.

current distribution in the wires, the field due to these charges and currents, and hence the electromotive force V of this field acting from A to B. The impressed electromotive force needed to sustain I^i against V is $V^i = -V$. Or we may start with a given V^i and determine the corresponding I^i.

4.3. *Currents across a Closed Surface*

The total electric and magnetic currents across a closed surface vanish. This theorem follows immediately from (1–15) when these equations are assumed to hold independently of the choice of (S). We simply draw a closed curve on a closed surface (S), and apply (1–15) to each part of the surface. Thus we have

$$\iint \left(gE_n + \epsilon \frac{\partial E_n}{\partial t}\right) dS = -\iint J_n \, dS = -I,$$

$$\iint \mu \frac{\partial H_n}{\partial t} \, dS = -\iint M_n \, dS = -K, \qquad (3\text{–}1)$$

where I and K are the impressed currents flowing out of the volume bounded by (S).

In perfect dielectrics $g = 0$. Substituting in (1) and integrating with respect to t, we have

$$\iint \epsilon E_n \, dS = \iint D_n \, dS = -\int_{-\infty}^{t} I \, dt = q, \qquad (3\text{–}2)$$

$$\iint \mu H_n \, dS = \iint B_n \, dS = -\int_{-\infty}^{t} K \, dt = m, \qquad (3\text{–}2)$$

where q and m are the electric and magnetic charges inside (S) at time t.

4.4. *Differential Equations of Electromagnetic Induction and Boundary Conditions*

Applying the integral equations (1–15) to an infinitely small loop and using the definition of the curl of a vector, we obtain

$$\text{curl } E = -\mu \frac{\partial H}{\partial t} - M, \quad \text{curl } H = gE + \epsilon \frac{\partial E}{\partial t} + J. \qquad (4\text{–}1)$$

Similarly, equations (1–16) for harmonic fields become

$$\text{curl } E = -i\omega\mu H - M, \quad \text{curl } H = (g + i\omega\epsilon)E + J. \qquad (4\text{–}2)$$

At a boundary (S) between two media the above equations are not necessarily satisfied because E and H may be discontinuous. A connection between the fields on opposite sides of (S) is obtained from the integral

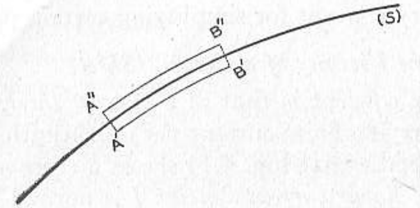


FIG. 4.14. A cross-section of a boundary between two media and a rectangle having two sides parallel to this boundary and the other two sides vanishingly small.

equations. Thus, assuming that all variables and constants in these equations are finite, and applying the equations to a typical rectangle with two sides, one in each medium, close to and parallel to (S), such as the rectangle $A'B'B''A''$ in Fig. 4.14, we have

$$E'_t = E''_t, \quad H'_t = H''_t. \qquad (4\text{–}3)$$

Hence the *tangential components of* E *and* H *are continuous at the interface of two media.*

Since the circulation of the tangential component of H per unit area is the normal component of J, the latter is continuous across S. The normal component of M is also continuous and we have

$$J'_n = J''_n, \quad M'_n = M''_n. \qquad (4\text{–}4)$$

For harmonic fields in source-free regions, these equations become

$$(g' + i\omega\epsilon')E'_n = (g'' + i\omega\epsilon'')E''_n, \quad \mu' H'_n = \mu'' H''_n. \qquad (4\text{–}5)$$

For static fields in perfect dielectrics, the conditions are

$$\epsilon' E'_n = \epsilon'' E''_n, \quad \mu' H'_n = \mu'' H''_n. \tag{4–6}$$

In *perfect conductors* ($g = \infty$) the electric intensity is zero for finite currents and the condition at the boundary is

$$E_t = 0, \quad \text{or} \quad H_n = 0. \tag{4–7}$$

The conception of perfect conductors is valuable chiefly because it helps to simplify mathematical calculations and to provide approximations to solutions of problems involving good conductors. In the future we shall assume all perfect conductors to be infinitely thin sheets. For reasons that will become evident later we may describe perfectly conducting sheets as *sheets of zero impedance*.

A *sheet of infinite impedance* is defined by the boundary conditions complementary to (7), that is by

$$H_t = 0, \quad \text{or} \quad E_n = 0. \tag{4–8}$$

Such a sheet can be pictured as having an infinite permeability and it is useful as an auxiliary concept for simplifying certain problems.

4.5. *Conditions in the Vicinity of a Current Sheet*

Another auxiliary concept is that of a *current sheet*, defined as an infinitely thin sheet carrying finite current per unit length normal to the lines of flow. Let us suppose that Fig. 4.14 shows a cross-section of an *electric* current sheet whose *linear current density* J is normal to the plane of the figure and is directed to the reader. Applying (1–15) to the rectangle $A'B'B''A''$, we obtain

$$H'_t - H''_t = J, \quad E'_t = E''_t. \tag{5–1}$$

The positive directions of the current density, the tangential component of H and the normal to the sheet are assumed to form a right-handed triad. Similarly for a *magnetic* current sheet of density M, we have

$$E'_t - E''_t = -M, \quad H'_t = H''_t. \tag{5–2}$$

The discontinuities in the tangential components of the field intensities imply discontinuities in the normal components of the field current densities. Imagining a pill box with its broad faces infinitely close and parallel to the electric current sheet on its opposite sides and then calculating the current into the pill box and out of it, we have

$$J''_n - J'_n = -\text{div}'\, J, \quad M'_n = M''_n. \tag{5–3}$$

Similarly for the magnetic current sheet, we obtain

$$M''_n - M'_n = -\text{div}'\, M, \quad J'_n = J''_n. \tag{5–4}$$

4.6. *Conditions in the Vicinity of Linear Current Filaments*

These conditions are obtained directly from (1–7) and (1–9). Thus in the immediate vicinity of an infinitely thin electric current filament I, and magnetic current filament K, we have

$$H_\varphi = \frac{I}{2\pi\rho}, \quad E_\varphi = -\frac{K}{2\pi\rho}, \tag{6–1}$$

assuming that the filaments coincide with the z-axis.

4.7. *Moving Surface Discontinuities*

We shall now consider the case in which the time derivatives of E and H are infinite as, for example, at a *wavefront* defined as the boundary between a finite moving field and a field-free space. Without loss of generality we may ignore the impressed currents. For reasons of simplification our discussion is restricted to homogeneous perfect dielectrics. Since there is no surface charge on the wavefront (S) (Fig. 4.15), the normal components of the electric and magnetic displacement densities are continuous; in homogeneous media this means that the normal components of the electric and magnetic intensities are also continuous. Since the field is identically zero on one side of the wavefront, the normal components vanish and E *and* H *are tangential to the wavefront.*

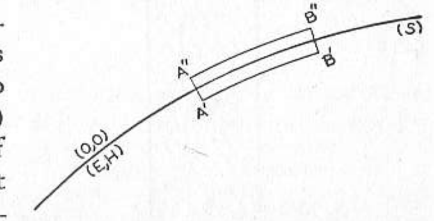


FIG. 4.15. A cross-section of the wavefront, that is, the boundary separating a field from field free space.

Let us assume that the positive directions of E, H and the velocity v of the field (normal to S) form a right-handed triplet and consider a rectangle $A'B'B''A''$ (Fig. 4.15) in which $A'B'$ is normal to H. The magnetic displacement through the rectangle increases at a rate μHvl, where l is the length of $A'B'$. This must be equal to the electromotive force El around the rectangle, where E is the electric intensity along $A'B'$; in view of our convention regarding the positive directions of E, H, v, we have $E = \mu vH$. Similarly, if we choose $A'B'$ in the direction of H, we obtain $H = \epsilon vE$. These equations connect H and the component of E normal to it. If there were a residual component of E, then, starting with this component and proceeding as above, we should have to acknowledge the existence of H normal to it which would be inconsistent with the original assumption that we have started with the total H. Multiplying and dividing the above equations we have

$$\mu\epsilon v^2 = 1, \quad v = \pm\frac{1}{\sqrt{\mu\epsilon}}, \quad E = \pm\eta H, \quad \eta = \sqrt{\frac{\mu}{\epsilon}}.$$

The velocity of the wavefront and the ratio E/H on it are thus fixed by the properties of the medium; this velocity will be called the *characteristic velocity* and the ratio $\eta = E/H$ the *intrinsic impedance.* In free space we have approximately* $v_0 = 3 \times 10^8$ m/sec, $\eta_0 = 120\pi \simeq 377$ ohms; in pure water $v = \frac{1}{3} \times 10^8$ m/sec, $\eta \simeq 42$ ohms.

Since E is positive when the product of H and v is positive, the field is moving in the direction in which a right-handed screw would advance when turned from E to H through 90°. Depending upon the relative directions of E and H, the field is either moving into the field-free space or receding from it. Imagine for instance a uniform " field slice " (Fig. 4.16) in which E and H have constant values between two parallel planes. As we have shown,

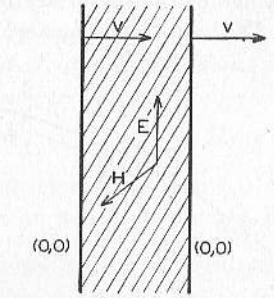


FIG. 4.16. The cross-section of a " field slice. "

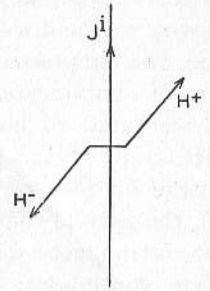


FIG. 4.17. An electric current sheet generating a field slice.

such a slice cannot remain stationary but must move with the characteristic velocity. That such a moving field could conceivably be generated may be seen as follows. Imagine an infinite plane sheet containing uniformly distributed equal and opposite charges and let a constant impressed intensity E^i set these charges in motion at the instant $t = 0$. At this moment magnetic intensities must appear on each side of the current sheet (Fig. 4.17) and $H^+ = -H^- = \frac{1}{2}J^i$ where J^i is the current density in the sheet. For the electric intensities on the two sides we have $E^+ = E^- = -E^i$. Considering the relative directions of E and H, we find that on both sides the field will be propagated away from the sheet. Between the two wavefronts the field remains uniform until E^i ceases to operate, at the instant $t = T$, let us say. Thereafter we shall have two field slices of thickness $l = vT$ moving in opposite directions. The work per unit area performed by E^i in sustaining J^i during the interval $(0,T)$ is E^iJ^iT and the energy contributed to the field is carried away by the field slices.

Similar but spherical field shells expanding outwards are created whenever an electric particle is accelerated or decelerated.

* The subscript zero is used to indicate specifically that the constants refer to free space.

4.8. *Energy Theorems*

Starting with the fundamental equations of electromagnetic induction (4–1), let us take the scalar product of the first equation and H and subtract from it the scalar product of the second equation and E

$$H \cdot \text{curl } E - E \cdot \text{curl } H = -M \cdot H - E \cdot J - gE^2 - \mu H \cdot \frac{\partial H}{\partial t} - \epsilon E \cdot \frac{\partial E}{\partial t}.$$

Integrating over a volume (τ) bounded by a closed surface (S), using equations (1.8–5) and then (1.3–1), and rearranging the terms we obtain

$$-\iiint_{(\tau)} E \cdot J \, d\tau - \iiint_{(\tau)} M \cdot H \, d\tau = \iiint_{(\tau)} gE^2 \, d\tau$$
$$+ \frac{\partial}{\partial t} \iiint_{(\tau)} \tfrac{1}{2}\epsilon E^2 \, d\tau + \frac{\partial}{\partial t} \iiint_{(\tau)} \tfrac{1}{2}\mu H^2 \, d\tau + \iint_{(S)} (E \times H)_n \, dS. \quad (8\text{–}1)$$

As usual the positive normal n to (S) is directed outwards. Integrating (1) with respect to t in the interval $(-\infty, t)$ and assuming that originally the space was field-free, we obtain

$$-\int_{-\infty}^{t} dt \iiint_{(\tau)} (E \cdot J + M \cdot H) \, d\tau = \int_{-\infty}^{t} dt \iiint_{(\tau)} gE^2 \, d\tau$$
$$+ \iiint_{(\tau)} (\tfrac{1}{2}\epsilon E^2 + \tfrac{1}{2}\mu H^2) \, d\tau + \int_{-\infty}^{t} dt \iint_{(S)} (E \times H)_n \, dS. \quad (8\text{–}2)$$

The left side in (1) is the rate at which work is done by the impressed forces against the forces of the field in sustaining the impressed currents and the left side in (2) is the total work performed by the impressed forces up to the instant t. In accordance with the principle of conservation of energy we say that this work appears as electromagnetic energy and we explain the various terms as follows. The first term on the right of (1) is the rate at which electric energy is converted into heat and the first term in (2) is the total energy so converted.* The second term in (1) may be interpreted as the time rate of change of the electric energy within (S) and the third term as the time rate of change of magnetic energy; the corresponding terms in (2) represent the electric and magnetic energies within (S) at the instant t. The last term in (1), being a surface integral, is interpreted as the time rate of energy flow across (S); similarly the last

* Since gE is the conduction current density $gE\,dt$ is the electric charge moving in response to E; consequently $gE^2\,dt$ is the work done by the field and must appear as some other form of energy. This form is heat and gE^2 is the power conversion per unit volume.

term in (2) is interpreted as the total flow of energy across (S) up to the instant t.

In this interpretation we assume that the electromagnetic energies are distributed throughout the field just as the energy dissipated in heat is known to be distributed. In conformity with (2) the volume densities of electric and magnetic energies are assumed to be

$$\mathcal{E}_e = \tfrac{1}{2}\epsilon E^2, \quad \mathcal{E}_m = \tfrac{1}{2}\mu H^2. \tag{8–3}$$

It is consistent with (1) and (2) to interpret the *Poynting* vector

$$P = E \times H \tag{8–4}$$

as a vector representing the time rate of energy flow per unit area. Certainly the surface integral of this vector over a closed surface represents

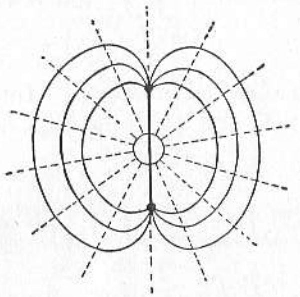


FIG. 4.18 The field of a magnetic doublet and an electric charge at the center of the doublet.

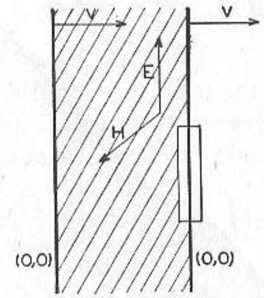


FIG. 4.19. The cross-section of a moving field slice and a "pill-box" with the flat faces parallel to the wavefront at vanishingly small distances from it.

the difference between the energy contributed to the field inside (S) and the energy accounted for within (S). On the other hand it is also true that the value of this integral remains the same if the curl of an arbitrary function is added to P. Furthermore, in the case of a magnet and an electric charge at its center (Fig. 4.18), P does not vanish; and yet in this case we are averse to assuming an actual flow of energy even though such an assumption is permissible.

In another instance, however, the interpretation of P as power flow per unit area is attractive. Consider a uniform field slice, moving in a perfect dielectric, and a pill box with its broad surfaces parallel to and on opposite sides of the wavefront (Fig. 4.19). In this case $J = M = 0$, $g = 0$, and (1) becomes

$$EHS = (\tfrac{1}{2}\epsilon E^2 + \tfrac{1}{2}\mu H^2)vS, \tag{8–5}$$

where S is the area of each of the broad surfaces of the pill box. The vector P is in the direction of the advancing wave and it seems as if the energy associated with the wave were actually traveling with velocity v, which is reasonable since the wave itself is advancing with this velocity.

If in the volume bounded by (S) there are no sources of energy, the energy dissipated in heat should enter the volume across (S). Consider for instance a direct current I in a cylindrical wire of radius a. If E and H are the components tangential to the wire, then the energy flowing into a section of length l in time t is

$$(2\pi a H)(lE)t = VIt, \tag{8–6}$$

where V is the voltage along the surface of the wire. Since It is the charge which has passed through the wire in time t, VIt is indeed the work done by the forces of the field.

We shall now derive another energy theorem, particularly suitable to harmonic fields. Multiplying scalarly the first equation of the set (4–2) by H^*, the conjugate of the second by E, and subtracting, we obtain

$$H^* \cdot \operatorname{curl} E - E \cdot \operatorname{curl} H^* = -M \cdot H^* - E \cdot J^* - gE \cdot E^* + i\omega\epsilon E \cdot E^* - i\omega\mu H \cdot H^*.$$

Integrating this over a volume (τ) bounded by a closed surface (S), using (1.8–5) and (1.3–1), rearranging the terms and dividing by two, we have

$$-\tfrac{1}{2}\iiint_{(\tau)} (E \cdot J^* + M \cdot H^*)\,d\tau = \tfrac{1}{2}\iiint_{(\tau)} gE \cdot E^*\,d\tau + \tfrac{1}{2}i\omega \iiint_{(\tau)} \mu H \cdot H^*\,d\tau - \tfrac{1}{2}i\omega \iiint_{(\tau)} \epsilon E \cdot E^*\,d\tau + \tfrac{1}{2}\iint_{(S)} (E \times H^*)_n\,dS. \tag{8–7}$$

The real part of the expression on the left of this equation is the average power spent by the impressed forces in sustaining the field. Some of this power is transformed into heat and the precise amount is given by the first term on the right; the rest flows out of the volume across (S) and the amount is represented by the real part of the last term. The second and third terms on the right are equal to the product of 2ω and the difference between the average magnetic and electric power stored inside (S).

The last term in (7) is called the *complex power flow* across (S) and is designated by Ψ

$$\Psi = \tfrac{1}{2}\iint (E \times H^*)_n\,dS. \tag{8–8}$$

The vector $P = \frac{1}{2}E \times H^*$ is the *complex Poynting vector*; its real part is the average power flow per unit area.

If (S) is a perfect conductor, the tangential component of E vanishes; hence there is no flow of energy across (S) and P is parallel to the surface. A closed perfectly conducting sheet separates space into two electromagnetically independent regions. Similar complete separation is afforded by a closed surface of infinite impedance ($H_t = 0$). In the physical world metals are in some respects good approximations to perfect conductors; but there are no good approximations to infinite impedance sheets except at zero frequency (or nearly zero) when substances with extremely high permeability act as high impedance media.

Only the tangential components of E and H contribute to Ψ. If u and v are orthogonal coordinates on (S) and if u, v, n form a right-handed triplet of directions, then

$$\Psi = \tfrac{1}{2}\int\int (E_u H_v^* - E_v H_u^*)\, dS. \tag{8-9}$$

Introducing the ratios

$$Z_{uv} = \frac{E_u}{H_v}, \quad Z_{vu} = -\frac{E_v}{H_u}, \tag{8-10}$$

we obtain

$$\Psi = \tfrac{1}{2}\int\int (Z_{vu} H_u H_u^* + Z_{uv} H_v H_v^*)\, dS. \tag{8-11}$$

If now

$$Z_{uv} = Z_{vu} = Z_n, \tag{8-12}$$

then

$$\Psi = \tfrac{1}{2}\int\int Z_n (H_u H_u^* + H_v H_v^*)\, dS. \tag{8-13}$$

In this case Z_n is called the *impedance normal to the surface* (S).

Consider now a conducting surface of thickness t. The linear current density $\mathcal{J}$ is equal to Jt, where J is the surface current density. If g is the conductivity, then $J = gE$ and consequently $\mathcal{J} = gtE$. If t approaches zero and g increases so that the product $G = gt$ remains constant, we have

$$\mathcal{J} = GE, \quad E = R\mathcal{J}, \tag{8-14}$$

where G and R are called respectively the *surface conductance* and the *surface resistance* of the sheet. More generally we define the *surface admittance* Y_s and the *surface impedance* Z_s by equations similar to the above

$$\mathcal{J} = Y_s E, \quad E = Z_s \mathcal{J}, \tag{8-15}$$

where the constants of proportionality are complex.

Imagine now two infinitely close sheets, one with infinite surface impedance and the other with finite impedance Z_s. Since the component of H tangential to an infinite impedance sheet is zero, the component of H tangential to that side of the finite impedance sheet which is not adjacent to the infinite impedance sheet is equal to the linear current density $\mathcal{J}$; thus

$$H_u = \mathcal{J}_v, \quad H_v = -\mathcal{J}_u, \tag{8-16}$$

and the impedance normal to the combination of the two sheets* is equal to the surface impedance. Equation (13) then becomes

$$\Psi = \tfrac{1}{2}\int\int Z_s (\mathcal{J}_u \mathcal{J}_u^* + \mathcal{J}_v \mathcal{J}_v^*)\, dS. \tag{8-17}$$

4.9. *Secondary Electromagnetic Constants*

The conductivity g, the dielectric constant ϵ, and the permeability μ are the primary electromagnetic constants of the medium in the sense that they appear directly in the formulation of the electromagnetic equations. In equations (4–2) the terms on the left are three dimensional derivatives corresponding to ordinary derivatives in one-dimensional problems. The transmission line equations (2.10–3) represent a special case of Maxwell's equations and the terminology of the former may be extended to the latter. Thus we may call $i\omega\mu$ the distributed series impedance of the medium and $(g + i\omega\epsilon)$ the distributed shunt admittance. The constants μ, g, ϵ are respectively the distributed series inductance, shunt conductance, shunt capacitance. In wave theory the important constants are not the primary constants. Thus in transmission line theory two secondary constants are introduced: the *propagation constant* Γ and the *characteristic impedance* K, the first being defined as the square root of the product of the series impedance and the shunt admittance and the second as the square root of their ratio. Likewise, in three dimensional theory the important constants are the *intrinsic propagation constant* σ and the *intrinsic impedance* η defined by

$$\sigma = \sqrt{i\omega\mu(g + i\omega\epsilon)}, \quad \eta = \sqrt{\frac{i\omega\mu}{g + i\omega\epsilon}}. \tag{9-1}$$

These constants are independent of the geometry of the wave; hence the adjective "intrinsic," as synonymous with "characteristic of the medium." The characteristic impedances of various types of waves will contain η as a factor.

The primary constants are non-negative except when the frequencies

* As seen from the side of the finite impedance sheet.

are very high;* hence the square roots in the above equations are either in the first quadrant or in the third. The definitions (1) are made unambiguous if it is agreed that σ and η lie in the first quadrant or on its boundaries. In fact, σ is never below the bisector of the first quadrant and η is never above it. For non-dissipative media (perfect dielectrics) σ is on the positive imaginary axis and η on the positive real; for good conductors $\omega\epsilon$ is negligible compared with g (except at optical frequencies) and both σ and η are on the bisector.

In general σ and η are complex quantities

$$\sigma = \alpha + i\beta, \quad \eta = \mathcal{R} + i\mathcal{X}. \tag{9-2}$$

The quantities α and β are respectively the *attenuation constant* and the *phase constant* of the medium; $\mathcal{R}$ and $\mathcal{X}$ are the intrinsic resistance and reactance.

Bearing in mind the wave terminology introduced in section 2.4 we have the following expressions for perfect dielectrics

$$\sigma = i\beta, \quad \beta = \omega\sqrt{\mu\epsilon} = \frac{\omega}{v} = \frac{2\pi}{\lambda}, \quad \eta = \sqrt{\frac{\mu}{\epsilon}}, \quad f\lambda = v,$$
$$v = \frac{1}{\sqrt{\mu\epsilon}}, \quad \lambda = \frac{2\pi}{\beta}, \quad \mu = \frac{\eta}{v}, \quad \epsilon = \frac{1}{\eta v}. \tag{9-3}$$

The phase velocity v as defined by these equations is called the *characteristic velocity* of the medium. For some electromagnetic waves the phase velocity is equal to this characteristic velocity; but in general the characteristic velocity is only one factor in determining the actual phase velocity of a wave. Similarly the wavelength as defined above is called the *characteristic wavelength*; it is one of the factors determining the actual wavelength.

For free space we have the following numerical values of various constants

$$\eta_0 = \sqrt{\frac{\mu_0}{\epsilon_0}} = 376.7 \simeq 377 \simeq 120\pi \text{ ohms},$$
$$v_0 = \frac{1}{\sqrt{\mu_0\epsilon_0}} = 2.998 \times 10^8 \simeq 3 \times 10^8 \text{ meters/second},$$
$$\eta_0^{-1} = \sqrt{\frac{\epsilon_0}{\mu_0}} = 2.654 \times 10^{-3} \simeq \frac{1}{120\pi} \text{ mhos}, \tag{9-4}$$

* At optical frequencies ϵ may be negative.

$$\mu_0 = 4\pi \times 10^{-7} = 1.257 \times 10^{-6} \text{ henries/meter},$$
$$\epsilon_0 = 8.854 \times 10^{-12} \simeq \frac{1}{36\pi} \times 10^{-9} \text{ farads/meter}.$$

The value 120π for the impedance of free space corresponds to the value 3×10^8 for the characteristic velocity; for most engineering purposes these values are sufficiently accurate and we shall use them in all subsequent numerical calculations.

The permeability and dielectric constant of any medium relative to free space are called the *relative permeability* and the *relative dielectric constant* of the medium. It is the relative permeabilities and dielectric constants that are usually given in tables of physical constants. These constants are dimensionless; we shall designate them by μ_r and ϵ_r. The square root of ϵ_r is called the *index of refraction.*

For conductors we have

$$\sigma = \sqrt{i\omega\mu g}, \quad \eta = \sqrt{\frac{i\omega\mu}{g}},$$
$$\alpha = \beta = \sqrt{\pi\mu f g} = \frac{2\pi\sqrt{30 g\mu_r}}{\sqrt{\lambda_0}} = 34.4\sqrt{\frac{g\mu_r}{\lambda_0}}, \tag{9-5}$$
$$\mathcal{R} = \mathcal{X} = \sqrt{\frac{\pi\mu f}{g}} = \frac{2\pi\sqrt{30\mu_r}}{\sqrt{g\lambda_0}} = \frac{34.4\sqrt{\mu_r}}{\sqrt{g\lambda_0}},$$

where λ_0 is the characteristic wavelength in free space. The simple relation

$$\alpha = \mathcal{R}g, \quad \mathcal{R} = \frac{\alpha}{g}, \tag{9-6}$$

between $\mathcal{R}$ and α should be noted.

For pure copper we assume $g = 5.8005 \times 10^7 \simeq 5.8 \times 10^7$; then we have

$$\alpha = 15.1\sqrt{f} = 47.8\sqrt{\frac{f}{10}} = \frac{26.2 \times 10^4}{\sqrt{\lambda_0}} = \frac{82.9 \times 10^4}{\sqrt{10\lambda_0}}, \tag{9-7}$$
$$\mathcal{R} = 2.61 \times 10^{-7}\sqrt{f} = 8.25 \times 10^{-7}\sqrt{\frac{f}{10}} = \frac{4.52 \times 10^{-3}}{\sqrt{\lambda_0}} = \frac{0.0143}{\sqrt{10\lambda_0}}.$$

The " Q " of a medium is defined as the ratio of the displacement current density to the conduction current density

$$Q = \frac{\omega\epsilon}{g}. \tag{9-8}$$

If $Q \gg 1$, the displacement currents are much stronger than the conduction currents and the medium is called a *quasi-dielectric*; if $Q \ll 1$, then the medium is a *quasi-conductor*. Some media belong to one class or the other over the entire working frequency range; others may change their character within this range. The frequency $\hat{f}$ or the corresponding free space characteristic wavelength $\hat{\lambda}_0$ for which

$$Q = 1 \tag{9–9}$$

may conveniently be used to indicate whether in a given frequency range the medium is a quasi-conductor or a quasi-dielectric. From the above equations we have

$$\hat{f} = \frac{g}{2\pi\epsilon}, \quad \hat{\lambda}_0 = \frac{v_0}{\hat{f}}. \tag{9–10}$$

Introducing Q into the definitions (1), we have

$$\sigma = \sqrt{i\omega\mu g(1 + iQ)} = i\omega\sqrt{\mu\epsilon\left(1 - \frac{i}{Q}\right)},$$
$$\eta = \sqrt{\frac{i\omega\mu}{g(1 + iQ)}} = \sqrt{\frac{\mu}{\epsilon}}\left(1 - \frac{i}{Q}\right)^{-1}. \tag{9–11}$$

Separating real and imaginary parts, we obtain

$$\alpha = \sqrt{\tfrac{1}{2}\omega\mu(\sqrt{g^2 + \omega^2\epsilon^2} - \omega\epsilon)} = \tfrac{1}{2}g\sqrt{\frac{\mu}{\epsilon}}\sqrt{\frac{2}{1 + \sqrt{1 + \frac{1}{Q^2}}}},$$

$$\beta = \sqrt{\tfrac{1}{2}\omega\mu(\sqrt{g^2 + \omega^2\epsilon^2} + \omega\epsilon)} = \omega\sqrt{\mu\epsilon}\sqrt{\frac{1}{2}\left(1 + \sqrt{1 + \frac{1}{Q^2}}\right)},$$

$$\mathcal{R} = \frac{\beta}{\sqrt{g^2 + \omega^2\epsilon^2}} = \sqrt{\frac{\mu}{\epsilon}}\sqrt{\frac{1 + \sqrt{1 + \frac{1}{Q^2}}}{2\left(1 + \frac{1}{Q^2}\right)}}, \tag{9–12}$$

$$\mathcal{X} = \frac{\alpha}{\sqrt{g^2 + \omega^2\epsilon^2}}, \quad |\eta| = \sqrt{\omega\mu}(g^2 + \omega^2\epsilon^2)^{-1/4} = \sqrt{\frac{\mu}{\epsilon}}\left(1 + \frac{1}{Q^2}\right)^{-1/4};$$

$$\eta = |\eta|\,e^{i\vartheta}, \quad \vartheta = \frac{\pi}{4} - \frac{1}{2}\tan^{-1} Q,$$

$$\alpha = \sqrt{\pi\mu f g}\,(1 - \tfrac{1}{2}Q + \tfrac{1}{8}Q^2 + \tfrac{1}{16}Q^3 - \tfrac{5}{128}Q^4 - \tfrac{7}{256}Q^5 + \cdots),$$

$$\beta = \sqrt{\pi\mu f g}\,(1 + \tfrac{1}{2}Q + \tfrac{1}{8}Q^2 - \tfrac{1}{16}Q^3 - \tfrac{5}{128}Q^4 + \tfrac{7}{256}Q^5 + \cdots),$$

$$\alpha = \tfrac{1}{2}g\sqrt{\frac{\mu}{\epsilon}}\left(1 - \frac{1}{8Q^2} + \frac{7}{128Q^4} - \cdots\right),$$

$$\beta = \omega\sqrt{\mu\epsilon}\left(1 + \frac{1}{8Q^2} - \frac{5}{128Q^4} + \cdots\right), \tag{9–12 (cont'd.)}$$

$$\mathcal{R} = \sqrt{\frac{\pi\mu f}{g}}\,(1 + \tfrac{1}{2}Q - \tfrac{3}{8}Q^2 - \tfrac{5}{16}Q^3 + \tfrac{35}{128}Q^4 + \tfrac{63}{256}Q^5 - \cdots),$$

$$\mathcal{X} = \sqrt{\frac{\pi\mu f}{g}}\,(1 - \tfrac{1}{2}Q - \tfrac{3}{8}Q^2 + \tfrac{5}{16}Q^3 + \tfrac{35}{128}Q^4 - \tfrac{63}{256}Q^5 - \cdots),$$

$$\mathcal{R} = \sqrt{\frac{\mu}{\epsilon}}\left(1 - \frac{3}{8Q^2} + \frac{35}{128Q^4} - \cdots\right),$$

$$\mathcal{X} = \frac{1}{2Q}\sqrt{\frac{\mu}{\epsilon}}\left(1 - \frac{5}{8Q^2} + \frac{63}{128Q^4} - \cdots\right).$$

The last two series are appropriate for $Q > 1$ and the preceding pair for $Q < 1$. While there exist rapidly converging series expansions in the neighborhood of $Q = 1$ it is more practicable in this range to compute directly from the first four equations.

The first terms of the last two series are first approximations for quasi-dielectrics ($Q \gg 1$) and the first two terms in the preceding pair are first approximations for quasi-conductors ($Q \ll 1$). The frequency and free-space wavelength for which $Q = 1$ are determined from (10) or still more conveniently from

$$\hat{f} = \frac{1.8 \times 10^{10} g}{\epsilon_r}, \quad \hat{\lambda}_0 = \frac{2\pi\epsilon_r}{\eta_0 g} = \frac{\epsilon_r}{60g}. \tag{9–13}$$

The following table illustrates the extent to which media may differ from each other electromagnetically,

Mica,	$g = 1.1 \times 10^{-14}$,	$5.7 < \epsilon_r < 7$,	$2.8 \times 10^{-5} < \hat{f} < 3.5 \times 10^{-5}$,
Quartz,	$g = 8.3 \times 10^{-13}$,	$\epsilon_r = 4.5$,	$\hat{f} = 0.00332$,
Dry Soil,	$g = 0.015$,	$\epsilon_r = 10$,	$\hat{\lambda}_0 = 11$,
Wet Soil,	$g = 0.015$,	$\epsilon_r = 30$,	$\hat{\lambda}_0 = 33$,
Sand,	$g = 0.002$,	$\epsilon_r = 10$,	$\hat{\lambda}_0 = 83$,
Sea Water,	$g = 5$,	$\epsilon_r = 78$,	$\hat{\lambda}_0 = 0.26$.

(9–14)

The figures for soil and water are subject to considerable variation from place to place and are given here in order to give some idea of their magnitudes.

If g is independent of the frequency the Q may be conveniently expressed in the following form

$$Q = \frac{f}{\hat{f}} = \frac{\hat{\lambda}}{\lambda}. \tag{9–15}$$

It should be noted, however, that for many dielectrics g is nearly proportional to f and hence Q is nearly constant.

The propagation constant of a good conductor is always much larger than that of a dielectric. Thus for a conductor and free space we have

$$\frac{|\sigma|}{\beta_0} = \sqrt{60g\lambda_0\mu_r}. \tag{9–16}$$

Since g is of the order 10^7, this ratio is large even for the highest "engineering" frequencies.

4.10. *Waves in Dielectrics and Conductors*

If the impressed current densities J and M are differentiable, then either E or H can be eliminated from Maxwell's equations (4–1) and (4–2). It is then found that E and H and more generally $(gE + \epsilon(\partial E/\partial t))$ and H satisfy non-homogeneous "wave equations." In practice, however, J and M are usually discontinuous and hence non-differentiable; in such cases E and H cannot be eliminated without introducing auxiliary functions, called potential functions.

In source-free regions, on the other hand, we can always eliminate either E or H. Thus in homogeneous media from (4–2) we have

$$\Delta E = \sigma^2 E, \quad \Delta H = \sigma^2 H.$$

More generally from (4–1) we have

$$\Delta E = \mu\epsilon\frac{\partial^2 E}{\partial t^2} + \mu g\frac{\partial E}{\partial t} = \frac{1}{v^2}\frac{\partial^2 E}{\partial t^2} + \frac{2\bar{\alpha}}{v}\frac{\partial E}{\partial t}, \quad v = \frac{1}{\sqrt{\mu\epsilon}}, \quad \bar{\alpha} = \tfrac{1}{2}g\sqrt{\frac{\mu}{\epsilon}},$$

and a similar equation for H.

We shall now consider exponential waves. Designating by V a typical cartesian component of either E or H, we have

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = \sigma^2 V. \tag{10–1}$$

Evidently this equation is satisfied by the following wave function

$$V = e^{-\Gamma_x x - \Gamma_y y - \Gamma_z z}, \tag{10–2}$$

where the propagation constants Γ_x, Γ_y, Γ_z in the directions of the coordinate axes are subject to the following condition

$$\Gamma_x^2 + \Gamma_y^2 + \Gamma_z^2 = \sigma^2 = i\omega\mu g - \omega^2\mu\epsilon. \tag{10–3}$$

In nondissipative media we have

$$\Gamma_x^2 + \Gamma_y^2 + \Gamma_z^2 = -\beta^2 = -\omega^2\mu\epsilon = -\frac{4\pi^2}{\lambda^2}. \tag{10–4}$$

Thus the laws of induction impose only one condition on the three propagation constants. Two of these constants are controlled by the distribution of sources producing the field. If this distribution is uniform in planes parallel to the xy-plane, for example, we should expect the field to be similarly uniform and $\Gamma_x = \Gamma_y = 0$; then the propagation constant in the z-direction is equal to the intrinsic propagation constant.

Consider a typical plane through the origin

$$x\cos A + y\cos B + z\cos C = 0, \tag{10–5}$$

where $\cos A$, $\cos B$ and $\cos C$ are the direction cosines of a normal to the plane. The distance s from this plane along the normal passing through the origin may be expressed as

$$s = x\cos A + y\cos B + z\cos C.$$

Hence if the field is uniform in planes parallel to (5), we have

$$V = e^{-\sigma s} = e^{-\sigma(x\cos A + y\cos B + z\cos C)}. \tag{10–6}$$

The propagation constants along the coordinate axes are

$$\Gamma_x = \sigma\cos A, \quad \Gamma_y = \sigma\cos B, \quad \Gamma_z = \sigma\cos C, \tag{10–7}$$

and for uniform plane waves in nondissipative media, we have

$$\beta_x = \beta\cos A, \quad \beta_y = \beta\cos B, \quad \beta_z = \beta\cos C;$$

$$\lambda_x = \lambda\sec A, \quad \lambda_y = \lambda\sec B, \quad \lambda_z = \lambda\sec C;$$

$$v_x = v\sec A, \quad v_y = v\sec B, \quad v_z = v\sec C; \tag{10–8}$$

$$\frac{1}{\lambda_x^2} + \frac{1}{\lambda_y^2} + \frac{1}{\lambda_z^2} = \frac{1}{\lambda^2}; \quad \frac{1}{v_x^2} + \frac{1}{v_y^2} + \frac{1}{v_z^2} = \frac{1}{v^2}.$$

Thus in the case of uniform plane waves the phase velocities in various directions are never smaller than the characteristic phase velocity and the phase constants β_x, β_y, β_z never greater than β.

If the propagation constants Γ_x, Γ_y, Γ_z are all imaginary, then

$$\beta_x^2 + \beta_y^2 + \beta_z^2 = \beta^2. \tag{10–9}$$

None of the phase constants is greater than the characteristic phase constant and we can identify planes in which the field is uniform as those normal to the straight line whose direction cosines are

$$\cos A = \frac{\beta_x}{\beta}, \quad \cos B = \frac{\beta_y}{\beta}, \quad \cos C = \frac{\beta_z}{\beta}. \tag{10–10}$$

But if the phase constant in some direction is greater than β, then there must be a real propagation constant in some perpendicular direction. Thus let $\beta_z > \beta$; then

$\Gamma_x^2 + \Gamma_y^2 = \beta_z^2 - \beta^2 > 0$. For instance if $\beta_z = \beta\sqrt{2}$ and therefore $v_z = v/\sqrt{2}$, and if $\Gamma_y = 0$, then $\Gamma_x = \beta = 2\pi/\lambda$, $\Gamma_x\lambda = 2\pi$. The attenuation in the x-direction per characteristic wavelength is about 6.28 nepers or 54.6 decibels; the field intensity is reduced to 0.00187 of its value if x is increased by λ.

In the general case of complex propagation constants we have

$$V = e^{-(\alpha_x x+\alpha_y y+\alpha_z z)}\, e^{-i(\beta_x x+\beta_y y+\beta_z z)}.$$

From (4) we obtain (for nondissipative media)

$$\beta_x^2 + \beta_y^2 + \beta_z^2 - \alpha_x^2 - \alpha_y^2 - \alpha_z^2 = \beta^2, \quad \alpha_x\beta_x + \alpha_y\beta_y + \alpha_z\beta_z = 0. \quad (10\text{–}11)$$

The second equation shows that the equiamplitude planes

$$\alpha_x x + \alpha_y y + \alpha_z z = \text{constant}$$

are perpendicular to the equiphase planes

$$\beta_x x + \beta_y y + \beta_z z = \text{constant}.$$

Thus in nondissipative media equiamplitude planes either coincide with or are normal to equiphase planes. In the former case the waves are uniform on equiphase planes in the sense that E and H each have constant values at all points of a given equiphase plane at a given instant; in the other case the amplitude varies exponentially, the fastest variation being in the direction given by the direction components $(\alpha_x,\alpha_y,\alpha_z)$.

In dissipative media the second equation of the set (11) becomes

$$\alpha_x\beta_x + \alpha_y\beta_y + \alpha_z\beta_z = \tfrac{1}{2}\omega\mu g,$$

and equiamplitude planes are no longer perpendicular to equiphase planes.

The foregoing general conclusions concerning waves of exponential type (2) have a broader significance than appears at first sight. The constant Γ_x represents the relative rate of change of V in the x-direction and we have

$$\Gamma_x = -\frac{1}{V}\frac{\partial V}{\partial x}, \quad \Gamma_x^2 = \frac{1}{V}\frac{\partial^2 V}{\partial x^2}.$$

The second equation is also satisfied by $-\Gamma_x$ and in it V may be a sum of two exponential terms, one proportional to $e^{-\Gamma_x x}$ and the other to $e^{\Gamma_x x}$. If the wave function is not exponential we may still define Γ_x by the second of the above equations; Γ_y and Γ_z may be defined similarly. If these quantities vary slowly from point to point, the solutions of the wave equation will be approximately exponential, and the above properties of exponential waves will be applicable in sufficiently small regions.

Some broad conclusions may be drawn with regard to waves at an interface between two homogeneous media whose intrinsic propagation constants are of different orders of magnitude, as is the case for conductors and dielectrics. Consider a plane interface (the xy-plane) between air (substantially free space) above the plane and some conductor below the plane (Fig. 4.20). For an exponential wave of type (2) the propagation constants Γ_x, Γ_y in directions parallel to the boundary must be the same in both media in order that the boundary conditions may be satisfied at all points of the air-conductor interface. This is evidently so if V represents a component of either E or H parallel to the boundary; in this instance V is continuous across the boundary and the continuity cannot be satisfied at all points unless Γ_x, Γ_y are the same on both sides of the boundary. The same argument applies to the normal com-

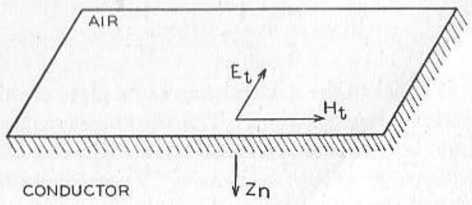


FIG. 4.20. A plane boundary between two semi-infinite media.

ponents of the current densities $(g + i\omega\epsilon)E_z$ and $i\omega\mu H_z$. Thus we have

$$\Gamma_x^2 + \Gamma_y^2 + \Gamma_{z,0}^2 = -\beta_0^2, \quad \Gamma_x^2 + \Gamma_y^2 + \Gamma_z^2 = \sigma^2.$$

Subtracting the first from the second, we have

$$\Gamma_z^2 = \sigma^2 + \beta_0^2 + \Gamma_{z,0}^2.$$

We have seen that the propagation constant for a conductor is always much larger than that for free space. $\Gamma_{z,0}$ is the propagation constant in free space in the direction normal to the interface; it may be comparable to β_0 if the direction of the wave is nearly normal to the conductor, or much smaller than β_0 if the wave direction is nearly parallel to the conductor. Hence in the conductor the propagation constant normal to the interface is substantially equal to the intrinsic propagation constant at all engineering frequencies.

Since the current density normal to the interface is continuous, we have

$$gE_n'' = i\omega\epsilon_0 E_n';$$

hence the normal component of E in the conductor is negligibly small compared with the normal component of E in the air.

Even at moderately high frequencies the attenuation constant in the conductor is large and the field becomes quite small at rather small distances from the interface. For frequencies of 10^3, 10^6, and 10^9 cycles per second the attenuation constant for copper is respectively 0.478, 15.1, 478 nepers per millimeter; or 4.15, 131, 4150 decibels per millimeter. Each 20 decibels represents a 10 to 1 intensity ratio; thus at a million cycles the field intensity one millimeter from the surface of the conductor is less than one millionth of the intensity at the surface. Except at low frequencies the fields are confined largely to thin skins of conductors.

The current density at the surface of the conductor is gE_t and elsewhere it is $gE_t e^{-\sigma z}$, where z is the normal distance from the surface. Then the total current per unit length normal to the lines of flow is

$$J = \int_0^\infty gE_t e^{-\sigma z}\, dz = \frac{g}{\sigma}E_t = \frac{1}{\eta}E_t. \quad (10\text{–}12)$$

On the other hand, $H_t = J$ and therefore the impedance normal to the interface is equal to the intrinsic impedance of the conductor. The conductor may be replaced by a sheet whose surface impedance is η, adjacent to a sheet of infinite impedance which would effectively exclude the space previously occupied by the conductor.

The resistance normal to a sufficiently thick metal plate may be expressed as

$$\mathcal{R} = \frac{\alpha}{g} = \frac{1}{gt}, \quad t = \frac{1}{\alpha}; \tag{10–13}$$

hence this resistance is equal to the d–c resistance of a plate of thickness t, defined by the reciprocal of the attenuation constant. This thickness is called the "skin depth"; but the term should not be interpreted as meaning that the rest of the conductor could be removed without changing its a–c resistance. The attenuation through the skin depth is only 1 neper and the field is reduced to only 0.368 of its value on the surface. If the entire current were compelled to flow in the "skin" of thickness t, given by the above equation, the a–c resistance would be 8.6 per cent higher than the actual resistance. The field is reduced to one tenth of its value when the distance from the surface is 2.3 times the skin depth.

It will follow from the equations of section 8.1 that the surface impedance of a conducting plate whose thickness t is small compared with the radius of curvature is $\eta \coth \sigma t$.

4.11. *Polarization*

In electromagnetic wave theory the differences between various media are expressed by three primary constants g, ϵ, μ. Inasmuch as material media are regions of free space in which are imbedded various material and electric particles, one can expect that in the final analysis there is only one medium, this being free space, with $g = 0$, $\epsilon = \epsilon_0$, $\mu = \mu_0$. An electromagnetic field acts on the electric particles of the medium; their spatial distribution and velocities are changed; and they act as secondary sources of the field. The macroscopic effect of these secondary sources is described by g, $\epsilon - \epsilon_0$, $\mu - \mu_0$. In wave theory we are not interested in physical explanations of the electromagnetic differences between various media, the constants g, ϵ, μ are supposed to be known, and we are not concerned whether their values have been obtained experimentally or somehow computed; but the principle of replacing one medium by another with compensating secondary sources is sometimes useful and can profitably be examined.

Before passing to generalities let us consider a few examples. Take a pair of concentric conducting spheres (Fig. 4.21). If the electric charge on the inner sphere is q, that on the outer (after being grounded) is $-q$, regardless of the dielectric between the spheres. By (1–3) the electric intensity is

$$E_r = \frac{q}{4\pi\epsilon r^2}, \tag{11–1}$$

while in free space it would be

$$E_r' = \frac{q}{4\pi\epsilon_0 r^2}. \tag{11–2}$$

The difference is

$$E_r - E_r' = \frac{q}{4\pi r^2}\left(\frac{1}{\epsilon} - \frac{1}{\epsilon_0}\right) = \frac{-q(\epsilon - \epsilon_0)}{4\pi\epsilon_0\epsilon r^2}. \tag{11–3}$$

This difference in intensities could be produced in free space by an electric charge $-q[1 - (\epsilon_0/\epsilon)]$ on the inner sphere and $q[1 - (\epsilon_0/\epsilon)]$ on the outer. If we postulate these charges on the surfaces of the dielectric adjacent to the spheres, we can account for the actual electric intensity on the assumption that the dielectric constant between the spheres is ϵ_0 instead of ϵ. In order to explain the postulated charges we may assume a reservoir of equal and opposite quantities of electricity in the dielectric so distributed as to render it neutral in the absence of an electric force. After the inner sphere has received an electric charge $+q$ and the outer $-q$, an electric field (2) is established. Under the influence of this field the electrified particles in the dielectric are displaced, negative particles toward the inner sphere and positive toward the outer. The total effect is to give rise to surface charges on the boundaries of the dielectric. These charges produce a field acting against the field (2), thus reducing the intensity to the value given by (1).

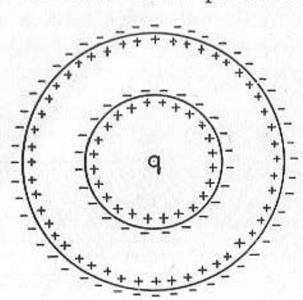


FIG. 4.21. Polarization of dielectrics.

The displacement density between the spheres is

$$D_r = \epsilon E_r = \epsilon_0 E_r + (\epsilon - \epsilon_0)E_r. \tag{11–4}$$

It differs from the displacement density that would have been produced by the same intensity in free space. The difference is called the *polarization* of the dielectric

$$P = (\epsilon - \epsilon_0)E. \tag{11–5}$$

As another example let us take a pair of conducting planes with a stratified dielectric between them (Fig. 4.22). If on the lower plane we have charge q and on the upper plane $-q$ then

$$D = q = \epsilon_1 E_1 = \epsilon_2 E_2. \tag{11–6}$$

We can now say that the dielectric constant is ϵ_1 everywhere between the conducting planes but that in the shaded region (Fig. 4.22) the medium has become polarized relative to the surrounding medium. The relative polarization P is defined by

$$P = (\epsilon_2 - \epsilon_1)E. \tag{11–7}$$

In the polarized region we now have

$$D = \epsilon_1 E + P. \tag{11–8}$$

On the boundaries between the polarized and the unpolarized regions E and $\epsilon_1 E$ are discontinuous and since this discontinuity is no longer ascribed to a difference in dielectric constants, it must be explained by surface charges. At the upper boundary the discontinuity in $\epsilon_1 E$ is $\epsilon_1(E_1 - E_2)$; by (6) this is equal to $(\epsilon_2 - \epsilon_1)E_2$ and therefore to the polarization P. Thus the density of the surface charge on the upper boundary is P; similarly on the lower boundary the density is $-P$ If $\epsilon_2 < \epsilon_1$, P is negative and the surface charges on the boundaries are of the same sign as those on the conductors nearest to them.

If instead of a stratified dielectric between the parallel plates we have a stratified conductor through which an electric current of density J is flowing, we shall have exactly the same results as above with conductivities g_1 and g_2 in place of dielectric constants ϵ_1 and ϵ_2 and with J in place of q.

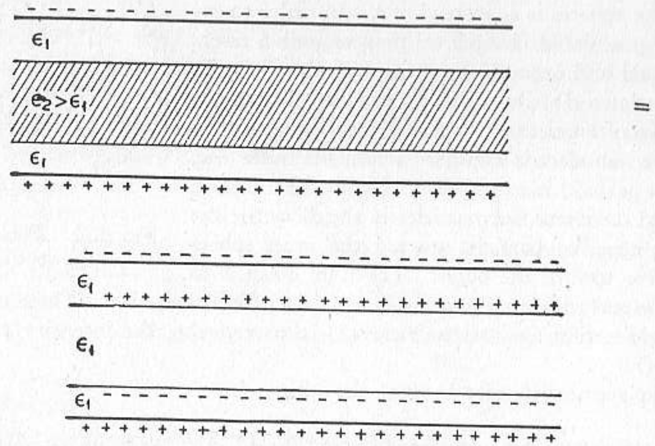


FIG. 4.22. The medium may be regarded as electrically homogeneous if we assume the existence of compensating charges.

More generally let us consider a medium which is homogeneous except for an "island" (Fig. 4.23) and suppose that the island is source free. Within the island we have

$$\operatorname{curl} E = -i\omega\mu''H, \quad \operatorname{curl} H = (g'' + i\omega\epsilon'')E, \tag{11-9}$$

and at the boundary

$$E'_t = E''_t, \quad H'_t = H''_t, \quad \mu'H'_n = \mu''H''_n,$$
$$(g' + i\omega\epsilon')E'_n = (g'' + i\omega\epsilon'')E''_n. \tag{11-10}$$

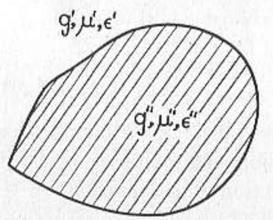


FIG. 4.23. The cross-section of an "island" in an otherwise homogeneous medium.

Since equations (9) can be written in the form,

$$\operatorname{curl} E = -i\omega\mu'H - M,$$
$$\operatorname{curl} H = (g' + i\omega\epsilon')E + J, \tag{11-11}$$

where

$$J = (g'' - g')E + i\omega(\epsilon'' - \epsilon')E, \quad M = i\omega(\mu'' - \mu')H, \tag{11-12}$$

it is theoretically possible to assume that the electromagnetic constants of the island are the same as those of the surrounding medium and that the field external to the island has induced in the latter the electric and magnetic currents given by (12). The island is said to be *polarized relative to the external medium* and J, M are called *polarization currents*. The latter act as impressed currents in addition to those producing the field.

Furthermore there is nothing to prevent us from ascribing an arbitrary set of electromagnetic constants to the entire medium provided we compensate for the new values by introducing polarization currents equal to the difference between the actual currents and the currents that would flow in the hypothetical medium in response to the same field intensities. Polarization relative to free space may be called *absolute polarization*.

If the island is homogeneous then div E and div H vanish. Likewise we have

$$\operatorname{div} J = 0, \quad \operatorname{div} M = 0. \tag{11-13}$$

J and M appear to originate on the boundary of the island; from (10) we obtain the densities of polarization currents appearing to flow out of the boundary

$$J_n = (g' + i\omega\epsilon')(E'_n - E''_n) = (g'' - g')E''_n + i\omega(\epsilon'' - \epsilon')E''_n,$$
$$M_n = i\omega\mu'(H'_n - H''_n) = i\omega(\mu'' - \mu')H''_n. \tag{11-14}$$

If $\omega = 0$, (14) become

$$J_n = (g'' - g')E''_n, \quad M_n = 0. \tag{11-15}$$

If $g' = g'' = 0$, the surface charge densities on the boundary of the island are

$$q_s = (\epsilon'' - \epsilon')E''_n, \quad m_s = (\mu'' - \mu')H''_n. \tag{11-16}$$

Within the island we have an electric polarization P^e and a magnetic polarization P^m

$$P^e = (\epsilon'' - \epsilon')E, \quad P^m = (\mu'' - \mu')H. \tag{11-17}$$

The surface charges (16) are seen to be equal to the normal components of polarization

$$q_s = P^e_n, \quad m_s = P^m_n. \tag{11-18}$$

If the island is not homogeneous then div J and div M represent the sources within the island. In nonconducting media we have volume distributions of charge given by

$$q_v = -\operatorname{div} P^e, \quad m_v = -\operatorname{div} P^m, \tag{11-19}$$

as well as surface distributions (18).

The reader must have been impressed by the artificial character of the foregoing transformations and it should be admitted that from the point of view of wave theory alone the concept of polarization is highly artificial and for the most part useless. J and M, as given by (12), depend on the *à priori* unknown field intensities and generally cannot be computed prior to the solution of the problem itself; after a solution has been found, J and M are only of academic interest. On the other hand, if the electromagnetic constants of the island are very different from those of the surrounding medium, it may be possible to obtain an approximate expression for the internal field without solving the complete problem and the action of this field on the surrounding medium may then be evaluated. An outstanding example of this is the antenna problem where the properties of the conducting wire serving as an antenna are very different from the properties of the surrounding space. Likewise, if the constants of the island and the external medium are nearly equal, the difference may be ignored in the first approximation, the problem solved, and then the polarization currents computed and used as secondary sources to obtain the first correction.

Polarization currents are not true applied currents and contribute no energy to the field.

4.12. *Special Forms of Maxwell's Equations in Source-Free Regions*

The most useful sets of Maxwell's equations appropriate to source-free regions are those expressed in cartesian, cylindrical or spherical coordinates. The general equations and some of the more important special forms will be listed for convenient reference. In cartesian coordinates we have

$$\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = -i\omega\mu H_x, \quad \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = (g + i\omega\epsilon)E_x,$$
$$\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -i\omega\mu H_y, \quad \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = (g + i\omega\epsilon)E_y, \qquad (12\text{–}1)$$
$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -i\omega\mu H_z, \quad \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = (g + i\omega\epsilon)E_z.$$

In cylindrical coordinates we have

$$\frac{\partial E_z}{\partial \varphi} - \rho\frac{\partial E_\varphi}{\partial z} = -i\omega\mu\rho H_\rho, \quad \frac{\partial H_z}{\partial \varphi} - \rho\frac{\partial H_\varphi}{\partial z} = (g + i\omega\epsilon)\rho E_\rho,$$
$$\frac{\partial E_\rho}{\partial z} - \frac{\partial E_z}{\partial \rho} = -i\omega\mu H_\varphi, \quad \frac{\partial H_\rho}{\partial z} - \frac{\partial H_z}{\partial \rho} = (g + i\omega\epsilon)E_\varphi, \qquad (12\text{–}2)$$
$$\frac{\partial}{\partial \rho}(\rho E_\varphi) - \frac{\partial E_\rho}{\partial \varphi} = -i\omega\mu\rho H_z, \quad \frac{\partial}{\partial \rho}(\rho H_\varphi) - \frac{\partial H_\rho}{\partial \varphi} = (g + i\omega\epsilon)\rho E_z.$$

In spherical coordinates we have

$$\frac{\partial}{\partial \theta}(\sin\theta\, E_\varphi) - \frac{\partial E_\theta}{\partial \varphi} = -i\omega\mu r \sin\theta\, H_r, \quad \frac{\partial}{\partial r}(rE_\theta) - \frac{\partial E_r}{\partial \theta} = -i\omega\mu r H_\varphi,$$
$$\frac{\partial}{\partial \theta}(\sin\theta\, H_\varphi) - \frac{\partial H_\theta}{\partial \varphi} = (g + i\omega\epsilon) r \sin\theta\, E_r, \quad \frac{\partial}{\partial r}(rH_\theta) - \frac{\partial H_r}{\partial \theta} = (g + i\omega\epsilon) r E_\varphi,$$
$$\frac{\partial E_r}{\partial \varphi} - \sin\theta \frac{\partial}{\partial r}(rE_\varphi) = -i\omega\mu r \sin\theta\, H_\theta, \qquad (12\text{–}3)$$
$$\frac{\partial H_r}{\partial \varphi} - \sin\theta \frac{\partial}{\partial r}(rH_\varphi) = (g + i\omega\epsilon) r \sin\theta\, E_\theta.$$

If the field is uniform in the direction of the z-axis, that is if the field is independent of the z-coordinate ($\partial/\partial z = 0$), then the foregoing equations separate into two independent sets. Thus in cartesian coordinates we have the following sets connecting E_z, H_x, H_y and the remaining components H_z, E_x, E_y.

$$\frac{\partial E_z}{\partial y} = -i\omega\mu H_x, \quad \frac{\partial E_z}{\partial x} = i\omega\mu H_y, \quad \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = (g + i\omega\epsilon)E_z; \qquad (12\text{–}4)$$

$$\frac{\partial H_z}{\partial y} = (g + i\omega\epsilon)E_x, \quad \frac{\partial H_z}{\partial x} = -(g + i\omega\epsilon)E_y, \quad \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} = i\omega\mu H_z. \qquad (12\text{–}5)$$

In cylindrical coordinates we have

$$\frac{\partial E_z}{\partial \varphi} = -i\omega\mu\rho H_\rho, \quad \frac{\partial E_z}{\partial \rho} = i\omega\mu H_\varphi, \quad \frac{\partial}{\partial \rho}(\rho H_\varphi) - \frac{\partial H_\rho}{\partial \varphi} = (g + i\omega\epsilon)\rho E_z; \qquad (12\text{–}6)$$
$$\frac{\partial H_z}{\partial \varphi} = (g + i\omega\epsilon)\rho E_\rho, \quad \frac{\partial H_z}{\partial \rho} = -(g + i\omega\epsilon)E_\varphi, \quad \frac{\partial}{\partial \rho}(\rho E_\varphi) - \frac{\partial E_\rho}{\partial \varphi} = -i\omega\mu\rho H_z. \qquad (12\text{–}7)$$

If the field is circularly symmetric, that is, if it is independent of the φ-coordinate, then the general set of six equations again breaks into two independent sets. In cylindrical coordinates we have

$$\frac{\partial E_\varphi}{\partial z} = i\omega\mu H_\rho, \quad \frac{\partial}{\partial \rho}(\rho E_\varphi) = -i\omega\mu\rho H_z, \quad \frac{\partial H_\rho}{\partial z} - \frac{\partial H_z}{\partial \rho} = (g + i\omega\epsilon)E_\varphi; \qquad (12\text{–}8)$$
$$\frac{\partial H_\varphi}{\partial z} = -(g + i\omega\epsilon)E_\rho, \quad \frac{\partial}{\partial \rho}(\rho H_\varphi) = (g + i\omega\epsilon)\rho E_z, \quad \frac{\partial E_z}{\partial \rho} - \frac{\partial E_\rho}{\partial z} = i\omega\mu H_\varphi. \qquad (12\text{–}9)$$

In spherical coordinates we have

$$\frac{\partial}{\partial \theta}(\sin\theta\, E_\varphi) = -i\omega\mu r \sin\theta\, H_r, \quad \frac{\partial}{\partial r}(rE_\varphi) = i\omega\mu r H_\theta,$$
$$\frac{\partial}{\partial r}(rH_\theta) - \frac{\partial H_r}{\partial \theta} = (g + i\omega\epsilon) r E_\varphi; \qquad (12\text{–}10)$$
$$\frac{\partial}{\partial \theta}(\sin\theta\, H_\varphi) = (g + i\omega\epsilon) r \sin\theta\, E_r, \quad \frac{\partial}{\partial r}(rH_\varphi) = -(g + i\omega\epsilon) r E_\theta,$$
$$\frac{\partial}{\partial r}(rE_\theta) - \frac{\partial E_r}{\partial \theta} = -i\omega\mu r H_\varphi. \qquad (12\text{–}11)$$

If the field is independent of two coordinates, x and y let us say, then equations (1) become

$$\frac{dE_x}{dz} = -i\omega\mu H_y, \quad \frac{dH_y}{dz} = -(g + i\omega\epsilon)E_x; \qquad (12\text{–}12)$$
$$\frac{dE_y}{dz} = i\omega\mu H_x, \quad \frac{dH_x}{dz} = (g + i\omega\epsilon)E_y; \qquad (12\text{–}13)$$
$$E_z = 0, \quad H_z = 0. \qquad (12\text{–}14)$$

We shall also have occasion to use the following sets, substantially the same as (4) and (5). If the field is independent of the x-coordinate, then

$$\frac{\partial E_x}{\partial z} = -i\omega\mu H_y, \quad \frac{\partial E_x}{\partial y} = i\omega\mu H_z, \quad \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = (g + i\omega\epsilon)E_x; \qquad (12\text{–}15)$$

$$\frac{\partial H_z}{\partial z} = (g + i\omega\epsilon)E_y, \quad \frac{\partial H_z}{\partial y} = -(g + i\omega\epsilon)E_z, \quad \frac{\partial E_y}{\partial z} - \frac{\partial E_z}{\partial y} = i\omega\mu H_x. \qquad (12\text{–}16)$$

If the field is independent of the y-coordinate, then

$$\frac{\partial E_y}{\partial z} = i\omega\mu H_x, \quad \frac{\partial E_y}{\partial x} = -i\omega\mu H_z, \quad \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = (g + i\omega\epsilon)E_y; \qquad (12\text{–}17)$$

$$\frac{\partial H_y}{\partial z} = -(g + i\omega\epsilon)E_x, \quad \frac{\partial H_y}{\partial x} = (g + i\omega\epsilon)E_z, \quad \frac{\partial E_z}{\partial x} - \frac{\partial E_x}{\partial z} = i\omega\mu H_y. \qquad (12\text{–}18)$$

CHAPTER V

IMPEDORS, TRANSDUCERS, NETWORKS

5.1. *Impedors and Networks*

An *impedor* is any combination of conductors and dielectrics, with two accessible terminals (Fig. 2.11). It may be as simple in structure as a laboratory resistor or as complicated as an antenna. In the latter case the impedor includes the wires of the antenna proper and the surrounding medium, including the earth, the Heaviside layer, etc. The impedor is linear if in the steady state the harmonic electromotive force between the terminals is proportional to the current

$$V = ZI, \qquad (1\text{–}1)$$

where the coefficient Z, called the impedance of the impedor, is a function of the frequency and in general of the oscillation constant.

Strictly speaking, we should specify the path between the terminals A and B, along which we compute or measure the electromotive force. For any two paths the difference of the electromotive forces is equal to the magnetic current through a surface bounded by these two paths. If H is the average magnetic intensity over the surface whose area is S, then the difference between the two voltages is*

$$V_1 - V_2 = i\omega\mu_0 HS = i\frac{240\pi^2 SH_{av}}{\lambda_0}. \qquad (1\text{–}2)$$

In the immediate vicinities of the wires the magnetic intensity is equal to the current divided by the length l of their circumference; hence H_{av} is certainly less than I/l, in fact considerably less since the integral of H along the radius will vary as log l. But even with l in the denominator of (2), the voltage difference is small if the distance between the terminals is a small fraction of the wavelength. Thus at low frequencies it becomes unnecessary to specify the path, except in precision calculations. At high frequencies we shall assume that the path is a straight line connecting the terminals, unless otherwise specified.

In the unrestricted frequency range the impedance is a complicated

* Assuming that the paths are in free space.

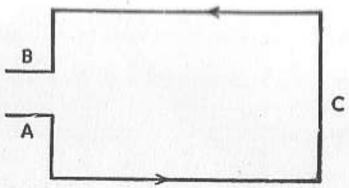


FIG. 5.1. A wire loop as an electric circuit.

function of the frequency; but its expansion in the vicinity of $\omega = 0$ generally is*

$$Z(i\omega) = \frac{1}{i\omega C} + R + i\omega L + \cdots. \tag{1-3}$$

If $C = \infty$ and $R \neq 0$, then for sufficiently low frequencies the impedor is a resistor; if $C = \infty$ and $R = 0$, then the impedor is an inductor; and if $R = L = 0$ but $C \neq \infty$, it is a capacitor. In practice R and L may be very small but they never vanish.

Consider for example a conducting loop (Fig. 5.1). From Faraday's law (4.1-7) we have

$$\int_{(ACB)} E_s\,ds + \int_{(BA)} E_s\,ds = -i\omega\Phi, \tag{1-4}$$

the first integral being taken along the conducting wire and the second along the straight line joining the terminals. The impressed electromotive force V, needed to transfer the charge from B to A against the field produced by the charge and the current in the wire must be equal and opposite to the second integral

$$V = -\int_{(BA)} E_s\,ds. \tag{1-5}$$

Substituting in (4), rearranging the terms, and dividing by the current flowing out of A, we have

$$Z = \frac{V}{I} = \frac{\int_{(ACB)} E_s\,ds}{J} + \frac{i\omega\Phi}{I}. \tag{1-6}$$

The first term, representing the ratio of the electromotive force along the surface of the wire to the input current, is called the *internal impedance* or the *surface impedance* of the wire; the second term is the *external impedance*.

At $\omega = 0$ for a homogeneous wire of length l and of uniform cross-section S, we have

$$J = gE_s, \quad J = \frac{I}{S}, \quad \int E_s\,ds = \frac{lI}{gS} = RI.$$

* Always, for actual physical structures; but for idealized physical structures the origin may sometimes be a branch point.

The impedance of the loop is just its resistance. The magnetic flux Φ is proportional to I. If L is the coefficient of proportionality at $\omega = 0$, then at low frequencies the external impedance is approximately proportional to the frequency. In the next chapter we shall obtain the next higher term in the expansion for the impedance of the loop.

From (2) we find that except at rather high frequencies the impedance of a loop of practical dimensions is small. This impedance may be increased by winding the wire into a coil (Fig. 5.2). Thus inside a long coil the magnetic intensity is approximately equal to the number of ampere-

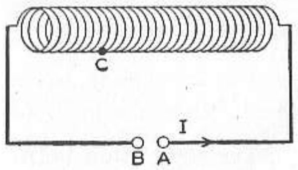


FIG. 5.2. A solenoid as an inductor.

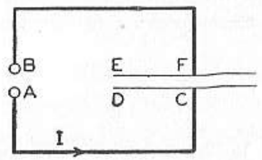


FIG. 5.3. An electric circuit containing a capacitor.

turns per unit length. The magnetic flux through the coil is then μSnI, where n is the number of turns per unit length. The electromotive force round each turn is $-i\omega\mu SnI$, and per unit length it is $-i\omega\mu Sn^2I$. The impressed electromotive force needed to drive the current through the coil against this electromotive force of " self-induction " is equal and opposite. Hence by increasing the number of turns, the impedance of the coil may be raised.

Consider now another structure consisting of a conducting wire and a pair of closely spaced conducting plates (Fig. 5.3). Applying the first induction law to the circuit $ACDEFBA$, we have

$$\int_{(BA)} E_s\,ds + \int E_s\,ds + \int_{(DE)} E_s\,ds = -i\omega\Phi, \tag{1-7}$$

where the second integral is taken along ACD and EFB. The first term is equal and opposite to the impressed electromotive force V. The time derivative of the charge q on the lower plate is the current I_c flowing into the capacitor and

$$i\omega q = I_c, \quad q = \frac{I_c}{i\omega}. \tag{1-8}$$

When $\omega = 0$, the charge is proportional to the voltage across the capacitor and therefore

$$\int_{(DE)} E_s\,ds = \frac{I_c}{i\omega C}, \tag{1-9}$$

except for the terms vanishing with ω. Hence from (7) we obtain

$$Z = \frac{V}{I} = Z_i + \frac{i\omega\Phi}{I} + \frac{1}{i\omega C}\frac{I_c}{I}, \qquad (1\text{–}10)$$

where I is again the input current and Z_i is the internal impedance of the wire and the outer surfaces of the capacitor. The first two terms are small and the impedance is nearly inversely proportional to the frequency. The ratio I_c/I is substantially unity unless the wires are "long." The meaning of "small" and "large," "short" and "long," will be discussed in more detail in the next chapter.

From elementary considerations we find that approximately

$$C = \frac{\epsilon S}{s}, \qquad (1\text{–}11)$$

where S is the area of each plate and s is the separation between them. This is a small quantity even for the smallest practicable values of s. In

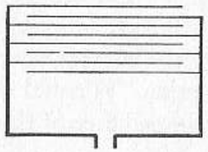


Fig. 5.4. A method for securing larger capacitances.

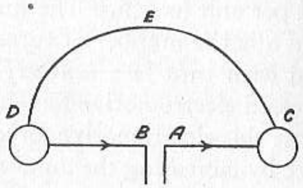


Fig. 5.5. A short doublet surmounted by two spheres.

order to increase C and make the impedance smaller, the area S is made larger in some such way as shown in Fig. 5.4.

Take another structure consisting of two short wires surmounted by spheres (Fig. 5.5). Applying the first law to the circuit $ACEDBA$ and dividing by the input current, we obtain equation (10) in which C is approximately the capacitance between the spheres at $\omega = 0$.

One of the fundamental problems of electromagnetic theory is to calculate the impedances of certain basic structures. The province of network theory is to study the impedance functions of various combinations or *networks* of such structures in order to design networks possessing specified desirable properties. While it is outside our province to be concerned extensively with network theory, we should know its elements in order to present our results in usable form. Calculation of impedances usually involves solution of auxiliary problems of wave propagation and the ultimate object (from the point of view of applications) might easily be forgotten.

Consider a number of impedors connected *in series* (Fig. 5.6). Assuming that their impedances are not too small, we may neglect the impedance of the connecting wires and write

$$V_{BC} + V_{DE} + V_{FG} + V_{HA} = 0, \qquad (1\text{–}12)$$

where the separate terms are the electromotive forces of the field which act between the various terminals in the order indicated. The last term is $-V$, where V is the impressed electromotive force between H and A. The above equation expresses the first *Kirchhoff law*. From (12) we then have

$$V = (Z_1 + Z_2 + Z_3)I, \quad Z = Z_1 + Z_2 + Z_3, \qquad (1\text{–}13)$$

where Z is the impedance of the entire circuit. The internal impedances of the connecting wires and the external impedance of the connecting loop could be added to (13). In the above equations we have assumed that the current through each impedor is equal to the input current of the entire

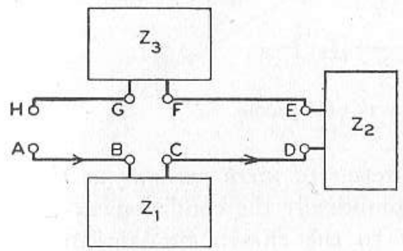


Fig. 5.6. A series connection of impedors.

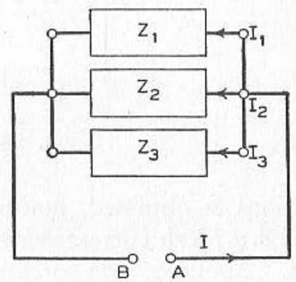


Fig. 5.7. A parallel connection of impedors.

circuit and thus disregarded the displacement currents between the connecting wires. Since the dielectric constant is small, it may be anticipated that these currents are negligible even at comparatively high frequencies. In order to obtain more precise information about their magnitudes and effect on the input impedance of the circuit we shall have to consider wave propagation on wires.

Consider now a number of impedors *in parallel* (Fig. 5.7). At a branch point the total current flowing in or out is zero; this is the second Kirchhoff law and it follows from Ampère's law of induction if we neglect the displacement current flowing from the branch point. Thus

$$I = I_1 + I_2 + I_3. \qquad (1\text{–}14)$$

Applying Faraday's law to various circuits in Fig. 5.7 and to a circuit in which the parallel combination is replaced by an "equivalent" impedor Z, we have

$$V = Z_1 I_1 = Z_2 I_2 = Z_3 I_3 = ZI.$$

Consequently

$$\frac{1}{Z} = \frac{1}{Z_1} + \frac{1}{Z_2} + \frac{1}{Z_3}, \quad \text{or} \quad Y = Y_1 + Y_2 + Y_3. \tag{1–15}$$

Let us now consider a more general network (Fig. 5.8). We could apply Kirchhoff's laws to different circuits or *meshes* of the network and write a number of equations connecting the impressed voltages with the currents through various impedors or *branches of the network*. A simpler set of

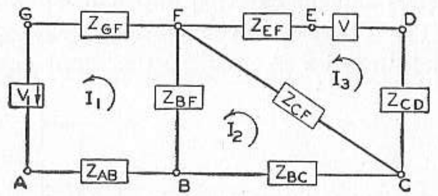


FIG. 5.8. A network of impedors.

equations is obtained, however, in terms of *mesh currents* as shown in Fig. 5.8. Mesh currents satisfy automatically the conditions at a branch point. Applying the circuital law to the chosen fundamental set of meshes, and substituting the mesh currents for the branch currents, we obtain the following typical set of equations for an n-mesh electric network

$$\begin{aligned} Z_{11}I_1 + Z_{12}I_2 + Z_{13}I_3 + \cdots + Z_{1n}I_n &= V_1, \\ Z_{21}I_1 + Z_{22}I_2 + Z_{23}I_3 + \cdots + Z_{2n}I_n &= V_2, \\ Z_{31}I_1 + Z_{32}I_2 + Z_{33}I_3 + \cdots + Z_{3n}I_n &= V_3, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ Z_{n1}I_1 + Z_{n2}I_2 + Z_{n3}I_3 + \cdots + Z_{nn}I_n &= V_n, \end{aligned} \tag{1–16}$$

where the V's are the total applied voltages in the corresponding meshes. The coefficient Z_{mm} is called the *impedance of the* m*th mesh* and the coefficients Z_{mk} are the *mutual impedances* between meshes m and k. If the electric current in the kth mesh is 1 ampere and the currents in the remaining meshes are equal to zero, then the voltage in the mth mesh is Z_{mk} volts. In Fig. 5.8 we have: $Z_{12} = -Z_{BF}$, $Z_{13} = 0$, $Z_{23} = -Z_{CF}$, etc.

If the matrix of the coefficients in (16) is nonsingular, we can solve the

equations and obtain

$$\begin{aligned} I_1 &= Y_{11}V_1 + Y_{12}V_2 + Y_{13}V_3 + \cdots + Y_{1n}V_n, \\ I_2 &= Y_{21}V_1 + Y_{22}V_2 + Y_{23}V_3 + \cdots + Y_{2n}V_n, \\ I_3 &= Y_{31}V_1 + Y_{32}V_2 + Y_{33}V_3 + \cdots + Y_{3n}V_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ I_n &= Y_{n1}V_1 + Y_{n2}V_2 + Y_{n3}V_3 + \cdots + Y_{nn}V_n, \end{aligned} \tag{1–17}$$

where the Y's are the *admittance coefficients* of the network. If the electromotive force in the kth mesh is 1 volt and the voltages in the remaining meshes are equal to zero, then the electric current in the mth mesh is Y_{mk} amperes.

Solving (16) for the I's and comparing with (17), we have

$$Y_{mk} = \frac{D_{km}}{D}, \tag{1–18}$$

where D is the determinant of the coefficients in (16) and D_{km} is the cofactor of the element Z_{km} in D.

Similarly solving (17) for the V's and comparing with (16), we obtain

$$Z_{mk} = \frac{\Delta_{km}}{\Delta}, \tag{1–19}$$

where Δ is the determinant of the coefficients in (17) and Δ_{km} is the cofactor of the element Y_{km} in Δ.

The impedance seen by the generator in the mth mesh is the ratio V_m/I_m when all the V's, except V_m, are equal to zero; hence this impedance is

$$Z_m = \frac{1}{Y_{mm}} = \frac{D}{D_{mm}}. \tag{1–20}$$

The impedance and admittance matrices are symmetric

$$Z_{mk} = Z_{km}, \quad Y_{mk} = Y_{km}. \tag{1–21}$$

That is, *the electromotive force in the* m*th mesh due to a unit current in the* k*th mesh is the same as the electromotive force in the* k*th mesh due to a unit current in the* m*th mesh, and also the current in the* m*th mesh due to a unit electromotive force in the* k*th mesh is the same as the current in the* k*th mesh due to a unit electromotive force in the* m*th mesh.* This is the Reciprocity Theorem. Since it is possible to choose the fundamental meshes in such a way that any two given branches belong to two different meshes and to no others, the reciprocity theorem implies that an interchange of the positions of a generator and an ammeter does not change the ammeter reading.

To prove the theorem we shall first establish the following lemma: let $V'_1, V'_2, \cdots V'_n$ be the electromotive forces in the various meshes of the electric network and $I'_1, I'_2, \cdots I'_n$ the corresponding currents, and let $I''_1, I''_2, \cdots I''_n$ be the currents in response to another set of electromotive forces $V''_1, V''_2, \cdots V''_n$; then

$$\sum_{\alpha=0}^{n} V'_\alpha I''_\alpha = \sum_{\beta=0}^{n} V''_\beta I'_\beta. \qquad (1\text{–}22)$$

On the left side of this equation we replace a typical voltage V'_α by the sum of the branch voltages in that mesh and group the terms having common branch voltages. If any particular branch PQ is common to several meshes, the voltage V'_{PQ} multiplied by the respective mesh currents will occur in several terms, the sum of which will be $V'_{PQ}I''_{PQ}$; therefore

$$\sum_{\alpha=0}^{n} V'_\alpha I''_\alpha = \sum_{(PQ)} V'_{PQ}I''_{PQ} = \sum_{(PQ)} Z_{PQ}I'_{PQ}I''_{PQ},$$

where the last two summations are taken over all the branches. The last expression is symmetric in the primed and double primed I's; hence (22) is true and our lemma is proved.

Since $I'_1, I'_2, \cdots I'_n$ and $I''_1, I''_2, \cdots I''_n$ are two independent sets of quantities, we may set

$$I'_\alpha = 1, \text{ if } \alpha = m; \quad I''_\beta = 1, \text{ if } \beta = k;$$
$$= 0, \text{ if } \alpha \neq m; \quad = 0, \text{ if } \beta \neq k.$$

Substituting in (22) we obtain

$$V'_k = V''_m, \quad \text{and} \quad Z_{km} = Z_{mk}.$$

Similarly choosing

$$V'_\alpha = 1, \text{ if } \alpha = m; \quad V''_\beta = 1, \text{ if } \beta = k;$$
$$= 0, \text{ if } \alpha \neq m; \quad = 0, \text{ if } \beta \neq k;$$

we obtain

$$I''_m = I'_k, \quad \text{and} \quad Y_{mk} = Y_{km}.$$

Thus the reciprocity theorem has been proved.

5.2. *Transducers*

A *four-terminal transducer* or simply a transducer is any combination of conductors and dielectrics with *two pairs* of accessible terminals (Fig. 5.9). The pairs of terminals may be those of a transformer, or of a telephone transmission line between two cities, or of two antennas. In the last case the transducer includes the space between the antennas, the ground, etc.

If the transducers are linear we have *à priori* equations

$$V_1 = Z_{11}I_1 + Z_{12}I_2,$$
$$V_2 = Z_{21}I_1 + Z_{22}I_2; \qquad (2\text{–}1)$$

similarly the currents are linear functions of the voltages

$$I_1 = Y_{11}V_1 + Y_{12}V_2,$$
$$I_2 = Y_{21}V_1 + Y_{22}V_2. \qquad (2\text{–}2)$$

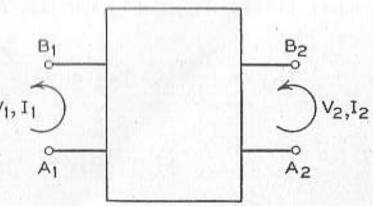


FIG. 5.9. A diagram for a four terminal transducer.

The Z's and Y's are functions of the electrical properties of the transducer and of the frequency but not of the V's and I's. The coefficient Z_{11} is called the impedance seen from the first pair of terminals and Z_{22} the impedance seen from the second pair; Z_{12} and Z_{21} are the *mutual impedances* or the *transfer impedances*. Similarly Y_{11} is called the admittance seen from the first pair of terminals and Y_{22} the admittance seen from the second pair; Y_{12} and Y_{21} are the *mutual admittances* or the *transfer admittances*.

If we leave the second pair of terminals " open " so that $I_2 = 0$, (1) becomes

$$V_1 = Z_{11}I_1, \quad V_2 = Z_{21}I_1. \qquad (2\text{–}3)$$

Thus if one ampere is passing through the first pair of terminals, Z_{11} is the voltage across this pair and Z_{21} is the voltage across the second pair. Similarly if the second pair of terminals is " short-circuited " so that $V_2 = 0$, then (2) becomes

$$I_1 = Y_{11}V_1, \quad I_2 = Y_{21}V_1. \qquad (2\text{–}4)$$

Hence if a unit voltage is impressed on the first pair of terminals, then Y_{11} is the current through this pair and Y_{21} is the current through the second pair.

Consider an n-mesh passive network of impedors with two pairs of accessible terminals, one pair in the mth mesh and the other in the kth. Then in (1–16) all the V's are zero except V_m and V_k. Eliminating all the I's except I_m and I_k, we obtain equations of the form (2) and hence the impedance coefficients of the transducer. By the theory of determinants it may be shown that the transfer impedances of a transducer consisting of a network with two pairs of accessible terminals are equal. For each type of transducer the corresponding reciprocity theorem should be proved separately as there is no *à priori* reason why the matrix of the impedance coefficients should be symmetric. In connection with cylindrical waves we shall encounter generalized transducers whose impedance matrices are not symmetric.

The admittances can be expressed in terms of the impedances and vice versa; thus solving (1) for the I's and comparing with (2), we have

$$Y_{11} = \frac{Z_{22}}{D}, \quad Y_{12} = -\frac{Z_{12}}{D}, \quad Y_{21} = -\frac{Z_{21}}{D}, \quad Y_{22} = \frac{Z_{11}}{D},$$
$$D = Z_{11}Z_{22} - Z_{12}Z_{21} = Z_{11}Z_{22} - Z_{12}^2. \tag{2–5}$$

Similarly we obtain

$$Z_{11} = \frac{Y_{22}}{\Delta}, \quad Z_{12} = -\frac{Y_{12}}{\Delta}, \quad Z_{21} = -\frac{Y_{21}}{\Delta}, \quad Z_{22} = \frac{Y_{11}}{\Delta},$$
$$\Delta = Y_{11}Y_{22} - Y_{12}Y_{21} = Y_{11}Y_{22} - Y_{12}^2. \tag{2–6}$$

These equations show that if $Z_{21} = Z_{12}$, then $Y_{21} = Y_{12}$ and vice versa. If the expression for any of the Z's from (6) is substituted in the expression for the corresponding Y in (5), we obtain $D\Delta = 1$.

Multiplying the first equation of set (1) by I_1^*, the second by I_2^*, and taking half the sum, we obtain an expression for the complex power

$$\Psi = \tfrac{1}{2}(V_1I_1^* + V_2I_2^*) = \tfrac{1}{2}[Z_{11}I_1I_1^* + Z_{12}(I_1I_2^* + I_1^*I_2) + Z_{22}I_2I_2^*]. \tag{2–7}$$

If $\hat{I}_1$ and $\hat{I}_2$ are the amplitudes of I_1 and I_2 and if ϑ is the phase angle between them, then (7) becomes

$$\Psi = \tfrac{1}{2}(Z_{11}\hat{I}_1^2 + 2Z_{12}\hat{I}_1\hat{I}_2 \cos\vartheta + Z_{22}\hat{I}_2^2). \tag{2–8}$$

The real part of Ψ is the average power contributed by the impressed forces to the transducer.

Multiplying the first equation of the set (2) by V_1^*, the second by V_2^* and taking half the sum, we have the corresponding expressions for the conjugate complex power

$$\begin{aligned}\Psi^* &= \tfrac{1}{2}[Y_{11}V_1V_1^* + Y_{12}(V_1V_2^* + V_1^*V_2) + Y_{22}V_2V_2^*] \\ &= \tfrac{1}{2}(Y_{11}\hat{V}_1^2 + 2Y_{12}\hat{V}_1\hat{V}_2 \cos\psi + Y_{22}\hat{V}_2^2),\end{aligned} \tag{2–9}$$

where ψ is the phase angle between V_1 and V_2.

It is convenient to express Ψ as the sum of three terms

$$\Psi = \Psi_{11} + 2\Psi_{12} + \Psi_{22}, \tag{2–10}$$

and to call the term $2\Psi_{12}$, depending on both I_1 and I_2, the *mutual power*.

The mutual impedance and admittance can be expressed in terms of Ψ_{12}; thus

$$Z_{12} = \frac{2\Psi_{12}}{I_1I_2^* + I_1^*I_2} = \frac{\Psi_{12}}{\hat{I}_1\hat{I}_2 \cos\vartheta},$$
$$Y_{12} = \frac{2\Psi_{12}^*}{V_1V_2^* + V_1^*V_2} = \frac{\Psi_{12}^*}{\hat{V}_1\hat{V}_2 \cos\psi}. \tag{2–11}$$

If I_1 and I_2 are impressed in phase we can assume their initial phases to be zero, then

$$Z_{12} = \frac{\Psi_{12}}{I_1I_2}. \tag{2–12}$$

In (1) the electromotive forces V_1 and V_2 are the forces necessary to sustain I_1 and I_2 against the forces of reaction of the transducer; they are the total electromotive forces developed by the generators only if the internal impedances of these generators are equal to zero. If the generators possess internal impedances Z_1 and Z_2, then the total impressed forces $\bar{V}_1$ and $\bar{V}_2$ are

$$\bar{V}_1 = V_1 + Z_1I_1, \quad \bar{V}_2 = V_2 + Z_2I_2. \tag{2–13}$$

Thus it is easy to extend (1) to include the generator impedances; we have only to add Z_1 to Z_{11} and Z_2 to Z_{22}

$$\bar{V}_1 = (Z_1 + Z_{11})I_1 + Z_{12}I_2,$$
$$\bar{V}_2 = Z_{21}I_1 + (Z_2 + Z_{22})I_2. \tag{2–14}$$

If in (1) the second pair of terminals is short-circuited, $V_2 = 0$ and we have

$$I_2 = -\frac{Z_{21}}{Z_{22}}I_1, \quad V_1 = \left(Z_{11} - \frac{Z_{12}^2}{Z_{22}}\right)I_1. \tag{2–15}$$

More generally if, instead of a generator, an "output impedance" Z_0 is inserted across the second pair of terminals, or "output terminals," the "input impedance" across the first pair is

$$Z_i = \frac{V_1}{I_1} = Z_{11} - \frac{Z_{12}^2}{Z_{22} + Z_0}. \tag{2–16}$$

This is the ratio of the voltage across the input terminals to the current in the input circuit and it is obtained from (13) and (14) by setting $\bar{V}_2 = 0$ and $Z_2 = Z_0$. The total impedance seen by the generator is

$$\frac{\bar{V}_1}{I_1} = Z_g + Z_i, \tag{2–17}$$

where Z_g is the internal impedance of the generator.

5.3. *Iterated Structures*

Consider a semi-infinite chain of identical transducers (Fig. 5.10). The input impedance K of this chain may be obtained very readily since I_0 and I_1 will not be altered if we replace the chain to the right of the terminals A_1, B_1 by an impedance equal to K. Thus we have the following equations

$$Z_{11}I_0 + Z_{12}I_1 = V_0, \quad Z_{21}I_0 + (Z_{22} + K)I_1 = 0, \tag{3–1}$$

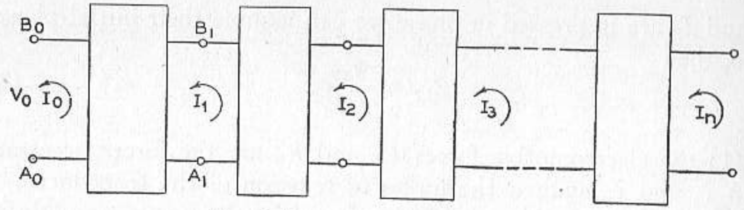


FIG. 5.10. A chain of transducers.

where V_0 is the counter-electromotive force of the chain which acts from A_0 to B_0 and hence is equal to the impressed electromotive force acting from B_0 to A_0. Eliminating I_1, we obtain

$$V_0 = \left(Z_{11} - \frac{Z_{12}^2}{Z_{22} + K}\right) I_0. \tag{3–2}$$

Since by definition the ratio V_0/I_0 is equal to K, we have

$$K = Z_{11} - \frac{Z_{12}^2}{Z_{22} + K}. \tag{3–3}$$

Solving, we obtain

$$K^2 + (Z_{22} - Z_{11})K - (Z_{11}Z_{22} - Z_{12}^2) = 0, \tag{3–4}$$

$$K_{1,2} = \tfrac{1}{2}(Z_{11} - Z_{22}) \pm \sqrt{\tfrac{1}{4}(Z_{11} + Z_{22})^2 - Z_{12}^2}. \tag{3–5}$$

The current transfer ratio I_1/I_0 may be found from (1); thus

$$\frac{I_1}{I_0} = -\frac{Z_{12}}{Z_{22} + K} = -\frac{2Z_{12}}{Z_{11} + Z_{22} \pm \sqrt{(Z_{11} + Z_{22})^2 - 4Z_{12}^2}}$$

$$= \frac{-(Z_{11} + Z_{22}) \pm \sqrt{(Z_{11} + Z_{22})^2 - 4Z_{12}^2}}{2Z_{12}}, \tag{3–6}$$

where the signs in front of the square root correspond to those in (5). This is the common ratio of the currents in two successive meshes of the chain; hence the transfer ratio across n transducers is

$$\frac{I_n}{I_0} = \frac{I_1}{I_0}\cdot\frac{I_2}{I_1}\cdot\frac{I_3}{I_2}\cdots\frac{I_n}{I_{n-1}} = \frac{(I_1)^n}{(I_0)^n} = \frac{(I_0)^{-n}}{(I_1)^{-n}}.$$

It is evident on physical grounds that the input impedance of the chain and the transfer ratio across each transducer are unique and we are faced with the problem of choosing the proper sign for the square roots in (5) and (6). The product of the two values χ_1, χ_2 for the current transfer ratio is unity. Thus if the absolute value of χ_1 is less than unity, the absolute value of χ_2 is greater than unity. In a dissipative chain the amplitude of the current must necessarily decrease and we must choose that sign of the square root in (6) which makes the absolute value of the current transfer ratio less than unity. This choice determines uniquely the sign in the expression for K. It is apparent from (6) that for small values of Z_{12} the proper sign is positive.

We may represent the current transfer ratio as an exponential function

$$\frac{I_1}{I_0} = e^{-\Gamma};$$

the constant Γ is the *propagation constant* or the *transfer constant* of the chain. The real part of Γ is positive and is called the *attenuation constant*; the imaginary part of Γ is called the *phase constant.*

Since one of the values of the current ratio in (6) is $e^{-\Gamma}$ and the other is its reciprocal e^{Γ}, we have

$$\cosh\Gamma = \frac{e^{\Gamma} + e^{-\Gamma}}{2} = -\frac{Z_{11} + Z_{22}}{2Z_{12}}. \tag{3–7}$$

The current and voltage across the terminals A_n, B_n may now be expressed in the form

$$I_n^+ = I_0^+ e^{-\Gamma n}, \quad V_n^+ = V_0^+ e^{-\Gamma n}, \quad V_0^+ = K^+ I_0^+,$$

where the superscript "plus" is used specifically to indicate a wave traveling from the source toward the right in an infinite chain (Fig. 5.11) of which the semi-infinite chain forms a part. For a wave traveling to the left in an infinite chain, we should have

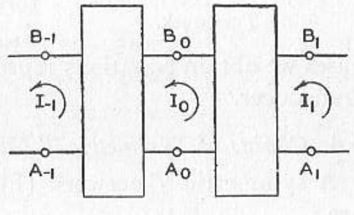


FIG. 5.11. Two sections of a chain extending to infinity in both directions.

$$I_n^- = I_0^- e^{\Gamma n},$$

since in this case the amplitude should decrease as n decreases. Here the current ratio is represented by the second value in (6). For the voltage we have

$$V_n^- = V_0^- e^{\Gamma n}, \quad V_0^- = -K^- I_0^-,$$

where $-K^-$ is the value of K in (5) other than the one designated by K^+. The impedance K^- is the impedance of the semi-infinite chain extending to the left, as seen from any pair of terminals. In fact if K^+ happens to correspond to the upper sign in (5), so that

$$K^+ = \tfrac{1}{2}(Z_{11} - Z_{22}) + \sqrt{\tfrac{1}{4}(Z_{11} + Z_{22})^2 - Z_{12}^2}; \qquad (3\text{–}8)$$

then, in accordance with the above definition,

$$K^- = \tfrac{1}{2}(Z_{22} - Z_{11}) + \sqrt{\tfrac{1}{4}(Z_{11} + Z_{22})^2 - Z_{12}^2}. \qquad (3\text{–}9)$$

If the elements of the chain are symmetric, then $Z_{11} = Z_{22}$ and $K^+ = K^-$. Since K^+ and K^- are impedances of passive networks, their real parts cannot be negative. These two impedances are called the *characteristic impedances* of the chain of transducers.

The expressions for the current and voltage in a chain consisting of a finite number of transducers may now be written in the following form

$$I_n = Ae^{-\Gamma n} + Be^{\Gamma n}, \quad V_n = K^+Ae^{-\Gamma n} - K^-Be^{\Gamma n},$$

where A and B are constants obtainable in terms of the terminal conditions. For instance, let the total number of transducers in the chain be m and let an impedance Z be inserted across the mth pair of terminals; then we have

$$Z = \frac{V_m}{I_m} = \frac{K^+Ae^{-\Gamma m} - K^-Be^{\Gamma m}}{Ae^{-\Gamma m} + Be^{\Gamma m}}.$$

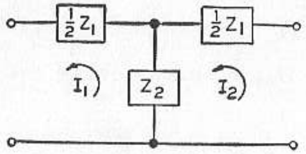


FIG. 5.12. A symmetric T-network.

From this equation we can express B in terms of A. The constant A can then be found in terms of the input voltage V_0 or the input current I_0.

Similarly A and B can be expressed in terms of V_0 and I_0, or in terms of V_0 and V_m, or in terms of I_0 and I_m. In the last two cases we obtain equations representing the chain of transducers as a single transducer.

5.4. *Chains of Symmetric T-Networks*

A symmetric T-network (Fig. 5.12) is a transducer whose impedances are

$$Z_{11} = Z_{22} = \tfrac{1}{2}Z_1 + Z_2, \quad Z_{12} = -Z_2. \qquad (4\text{–}1)$$

Substituting in (3–7) and (3–8), we have

$$\Gamma = \cosh^{-1}\frac{Z_1 + 2Z_2}{2Z_2} = \cosh^{-1}\left(\frac{Z_1}{2Z_2} + 1\right),$$

$$K^+ = K^- = \tfrac{1}{2}\sqrt{Z_1(Z_1 + 4Z_2)}. \qquad (4\text{–}2)$$

These are the constants for the iterated network shown in Fig. 5.13.

Evidently any symmetric transducer may be represented by a symmetric T-network; thus from (1) we have

$$Z_1 = 2(Z_{11} + Z_{12}), \quad Z_2 = -Z_{12}. \qquad (4\text{–}3)$$

Hence expressions (2) may be used for any chain of symmetric transducers.

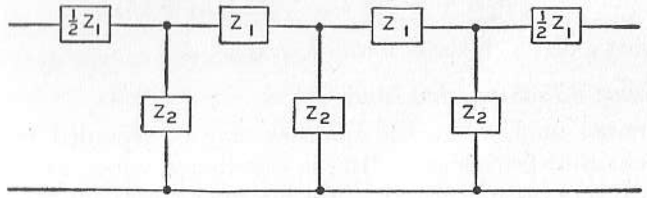


FIG. 5.13. A chain of symmetric T-networks.

5.5. *Chains of Symmetric Π-Networks*

Consider now a symmetric Π-network (Fig. 5.14). Starting with the following mesh equations

$$2Z_2I_1 - 2Z_2I + 0 = V_1, \quad -2Z_2I_1 + (Z_1 + 4Z_2)I - 2Z_2I_2 = 0,$$
$$0 - 2Z_2I + 2Z_2I_2 = V_2,$$

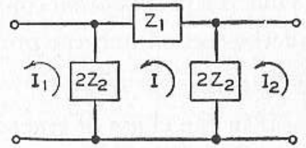


FIG. 5.14. A symmetric Π-network.

and eliminating the current I in the intermediate mesh, we obtain

$$\frac{2Z_2(Z_1 + 2Z_2)}{Z_1 + 4Z_2}I_1 - \frac{4Z_2^2}{Z_1 + 4Z_2}I_2 = V_1,$$
$$-\frac{4Z_2^2}{Z_1 + 4Z_2}I_1 + \frac{2Z_2(Z_1 + 2Z_2)}{Z_1 + 4Z_2}I_2 = V_2.$$

Hence we have

$$Z_{11} = Z_{22} = \frac{2Z_2(Z_1 + 2Z_2)}{Z_1 + 4Z_2}, \quad Z_{12} = -\frac{4Z_2^2}{Z_1 + 4Z_2}.$$

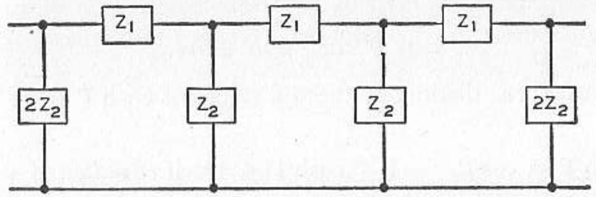


FIG. 5.15. A chain of symmetric Π-networks.

Therefore

$$K = \sqrt{Z_{11}^2 - Z_{12}^2} = \sqrt{(Z_{11} + Z_{12})(Z_{11} - Z_{12})} = 2Z_2\sqrt{\frac{Z_1}{Z_1 + 4Z_2}} \qquad (5\text{–}1)$$

is the characteristic impedance of a chain of Π-networks (Fig. 5.15).

The propagation constant is evidently the same as in the case of the chain of T-networks; K has a different value only because of terminal differences between the two chains. In fact, we can obtain (1) from the characteristic impedance in (4–2). Designating the latter by K and the former by K', we have

$$K' = \frac{(K + \frac{1}{2}Z_1)2Z_2}{K + \frac{1}{2}Z_1 + 2Z_2} = \frac{4Z_2K + 2Z_1Z_2}{2K + Z_1 + 4Z_2}.$$

It may be shown that K' is identical with K in (1).

5.6. *Continuous Transmission Lines*

A continuous uniform transmission line may be regarded as a limiting case of a chain of transducers. If the distributed series impedance and shunt admittance per unit length of the line are respectively Z and Y, then

$$Z_1 = Z\,dx, \quad Z_2 = \frac{1}{Y\,dx},$$

where dx is an element of length. By (4–2) we have

$$K = \sqrt{\frac{Z}{Y}}, \quad \cosh\Gamma = 1 + \tfrac{1}{2}\Gamma^2 + \cdots = 1 + \tfrac{1}{2}ZY\,dx^2, \quad \Gamma = \sqrt{ZY}\,dx.$$

Thus the propagation constant per "section" is proportional to the length of the section and the propagation constant per unit length is $\sqrt{ZY}$.

5.7. *Filters*

If in the chain of transducers Z_1 and Z_2 are pure reactances, their ratio is real and therefore $\cosh\Gamma$ is also real. Let α and β be the attenuation constant and the phase constant, then

$$\cosh\Gamma = \cosh(\alpha + i\beta) = \cosh\alpha\cos\beta + i\sinh\alpha\sin\beta.$$

This expression is real if

$$\sinh\alpha\sin\beta = 0;$$

that is, if

$$\alpha = 0 \quad \text{or} \quad \beta = \pm n\pi.$$

Thus there are three distinct ranges of values of $\cosh\Gamma$ to be considered, namely

$$\begin{aligned} \cosh\Gamma &= \cos\beta, \quad -1 \le \cosh\Gamma \le 1, \quad \text{if } \alpha = 0,\\ \cosh\Gamma &= \cosh\alpha, \quad 1 \le \cosh\Gamma < \infty, \quad \text{if } \beta = 0,\\ \cosh\Gamma &= -\cosh\alpha, \quad -\infty < \cosh\Gamma \le -1, \quad \text{if } \beta = \pi. \end{aligned} \tag{7–1}$$

The value of $\cosh\Gamma$ depends on the frequency; at some frequencies there will be no attenuation while at others the attenuation may be very high. Hence our chain of transducers will act as a *filter*. The frequency

interval of zero attenuation is called the *pass-band* of the filter; the frequency interval of nonzero attenuation is called the *stop-band*. By (4–2) and (1) the pass-band is determined by the following inequality

$$-4 \le \frac{Z_1}{Z_2} \le 0, \quad \text{or} \quad 0 \le -\frac{Z_1}{Z_2} \le 4. \tag{7–2}$$

The end points of the pass-band are called the *cut-off* frequencies. The pass-band may also be obtained from

$$-2 \le \frac{Z_{11} + Z_{22}}{Z_{12}} \le 2. \tag{7–3}$$

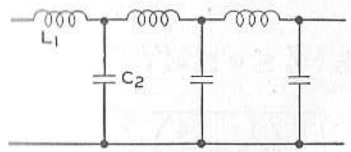


FIG. 5.16. A low-pass filter.

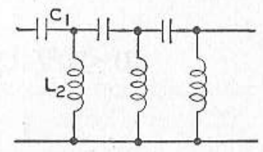


FIG. 5.17. A high-pass filter.

The following are a few simple examples of filter structures. In the chain of series inductances and shunt capacitances shown in Fig. 5.16, we have

$$Z_1 = i\omega L_1, \quad Z_2 = \frac{1}{i\omega C_2}, \quad \frac{Z_1}{Z_2} = -\omega^2 L_1 C_2;$$

in this case the pass-band is specified by

$$0 \le \omega \le \omega_c, \quad \omega_c = \frac{2}{\sqrt{L_1C_2}}. \tag{7–4}$$

The frequencies below the cut-off are passed and the filter is a *low-pass filter*. As L_1 and C_2 become smaller, ω_c becomes greater; in the limiting case of uniformly distributed constants all frequencies are passed.

For the chain shown in Fig. 5.17, we have

$$Z_1 = \frac{1}{i\omega C_1}, \quad Z_2 = i\omega L_2, \quad \frac{Z_1}{Z_2} = -\frac{1}{\omega^2 L_2 C_1};$$

thus the pass-band is determined by

$$\omega > \omega_c, \quad \omega_c = \frac{1}{2\sqrt{L_2C_1}}, \tag{7–5}$$

and the chain is a *high-pass* filter. In the limiting case of uniformly distributed series capacity and shunt inductance, no frequency is passed.

Our next example is the chain shown in Fig. 5.18. Here we have

$$Z_1 = i\omega L_1 + \frac{1}{i\omega C_1}, \quad Z_2 = \frac{1}{i\omega C_2}, \quad \frac{Z_1}{Z_2} = -\omega^2 L_1 C_2 + \frac{C_2}{C_1}, \qquad (7\text{–}6)$$

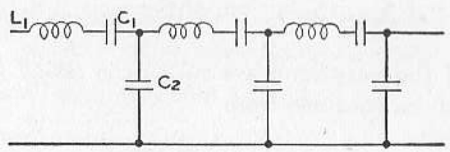


FIG. 5.18. A band-pass filter.

and

$$0 \le \omega^2 L_1 C_2 - \frac{C_2}{C_1} \le 4, \quad \omega_c' \le \omega \le \omega_c'',$$

$$\omega_c' = \frac{1}{\sqrt{L_1 C_1}}, \quad \omega_c'' = \sqrt{\frac{1}{L_1}\left(\frac{1}{C_1} + \frac{4}{C_2}\right)}.$$

This is a band-pass filter. If however L_1, C_1, C_2 are continuously distributed, we have

$$L_1 = \bar{L}_1\, dx, \quad C_1 = \frac{\bar{C}_1}{dx}, \quad C_2 = \bar{C}_2\, dx, \qquad (7\text{–}7)$$

where $\bar{L}_1$, $\bar{C}_1$ and $\bar{C}_2$ refer to unit length. The upper cut-off frequency becomes infinite and the transmission line has the characteristic of a high-pass filter with a cut-off given by

$$\omega_c = \frac{1}{\sqrt{\bar{L}_1 \bar{C}_1}}. \qquad (7\text{–}8)$$

We shall see that propagation of transverse magnetic waves is governed by equations of this type.

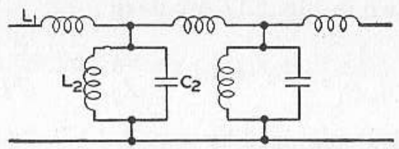


FIG. 5.19. A band-pass filter.

For the structure shown in Fig. 5.19 we have

$$Z_1 = i\omega L_1, \quad \frac{1}{Z_2} = i\omega C_2 + \frac{1}{i\omega L_2}, \quad \frac{Z_1}{Z_2} = -\omega^2 L_1 C_2 + \frac{L_1}{L_2}. \qquad (7\text{–}9)$$

This structure is also a band-pass filter in which the lower and the upper cut-offs are specified by

$$\omega_c' = \frac{1}{\sqrt{L_2 C_2}}, \quad \omega_c'' = \sqrt{\frac{1}{C_2}\left(\frac{1}{L_2} + \frac{4}{L_1}\right)}. \qquad (7\text{–}10)$$

When such a structure becomes a continuous line, then

$$L_1 = \bar{L}_1\, dx, \quad C_2 = \bar{C}_2\, dx, \quad L_2 = \frac{\bar{L}_2}{dx}, \quad \omega_c' = \frac{1}{\sqrt{\bar{L}_2 \bar{C}_2}}, \qquad (7\text{–}11)$$

and the upper cut-off has receded to infinity. Propagation of transverse electric waves is governed by equations of this type.

5.8. *Forced Oscillations in a Simple Series Circuit*

One of the simplest electric networks is a circuit consisting of a series combination of a resistor, an inductor, and a capacitor (Fig. 5.20). The impedance of such a circuit is

$$Z = R + i\omega L + \frac{1}{i\omega C} = R + i\left(\omega L - \frac{1}{\omega C}\right). \qquad (8\text{–}1)$$

The reactance component vanishes when

$$\omega = \bar{\omega} = \frac{1}{\sqrt{LC}}; \qquad (8\text{–}2)$$

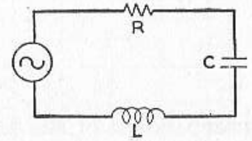


FIG. 5.20. A simple series circuit.

the frequency so defined is called the *resonant frequency* of the circuit. At the resonant frequency the reactances of the inductor and capacitor are equal except for sign; thus

$$\bar{\omega} L = \frac{1}{\bar{\omega} C} = \sqrt{\frac{L}{C}} = K. \qquad (8\text{–}3)$$

The quantity K is called the characteristic impedance of the circuit. The impedance of the circuit at any frequency can now be expressed in the form

$$Z = R + iK\left(\frac{\omega}{\bar{\omega}} - \frac{\bar{\omega}}{\omega}\right). \qquad (8\text{–}4)$$

At resonance the absolute value of the impedance is minimum and the current is maximum. The sharpness of the resonance curve (current vs. frequency) is seen to depend on the ratio of the characteristic impedance of the circuit to the resistance, that is, the "Q" of the circuit

$$Q = \frac{K}{R} = \frac{\bar{\omega} L}{R} = \frac{1}{\bar{\omega} RC}. \qquad (8\text{–}5)$$

This quantity can be defined in terms of the total energy $\mathcal{E}$ stored in the circuit at resonance and the average power W dissipated in R. Starting with the definitions of the resistor, inductor, and capacitor (section 2.7) and obtaining the work done by the applied electromotive forces, we find that at any particular instant the power dissipated in the resistor, the energy stored in the inductor, and the energy stored in the capacitor are respectively RI_i^2, $\frac{1}{2}LI_i^2$, $\frac{1}{2}CV_{C,i}^2$, where I_i is the instantaneous current in the resistor or the inductor and $V_{C,i}$ is the instantaneous voltage across the capacitor. Since at resonance the reactance seen by the generator is zero, the energy $\mathcal{E}$ stored in the circuit is constant and is equal to the energy stored in the inductor when the current is maximum and hence the voltage across the capacitor is zero, or to the energy stored in the capacitor when the voltage across it is maximum and the current is zero; the average dissipated power W is one half of the power dissipated when the current is maximum; that is,

$$\mathcal{E} = \tfrac{1}{2}L\hat{I}^2, \quad W = \tfrac{1}{2}R\hat{I}^2,$$

where $\hat{I}$ is the amplitude of I_i. Substituting L and R from these equations in (5), we have an alternative definition of Q

$$Q = \frac{\bar{\omega}\mathcal{E}}{W}, \tag{8–6}$$

independent of the concepts of resistance and inductance.

The current is proportional to the admittance Y of the circuit. The ratio of this admittance to the admittance at resonance depends only on Q and the ratio $\omega/\bar{\omega}$; thus

$$RY = \frac{1}{1 + iQ\left(\frac{\omega}{\bar{\omega}} - \frac{\bar{\omega}}{\omega}\right)}. \tag{8–7}$$

Introducing δ defined by

$$\delta = \frac{\omega - \bar{\omega}}{\bar{\omega}} = \frac{f - \bar{f}}{f}, \tag{8–8}$$

we obtain

$$RY = \frac{1}{1 + i2\delta Q(1 + \frac{1}{2}\delta)(1 + \delta)^{-1}}.$$

For small values of δ this is approximately

$$RY = \frac{1}{1 + i2\delta Q}. \tag{8–9}$$

Thus for "high Q circuits" the input resistance and reactance are equal when

$$2\delta Q = \pm 1, \quad \delta = \pm\frac{1}{2Q}. \tag{8–10}$$

At these frequencies the current amplitude is $50\sqrt{2} \simeq 71$ per cent of its maximum value and the power absorbed is one-half the maximum power that can be absorbed by the circuit. The quantity

$$|\,2\delta\,| = \frac{2\,|f - \bar{f}|}{\bar{f}} = \frac{1}{Q} \tag{8–11}$$

is called the *relative width* of the resonance curve.

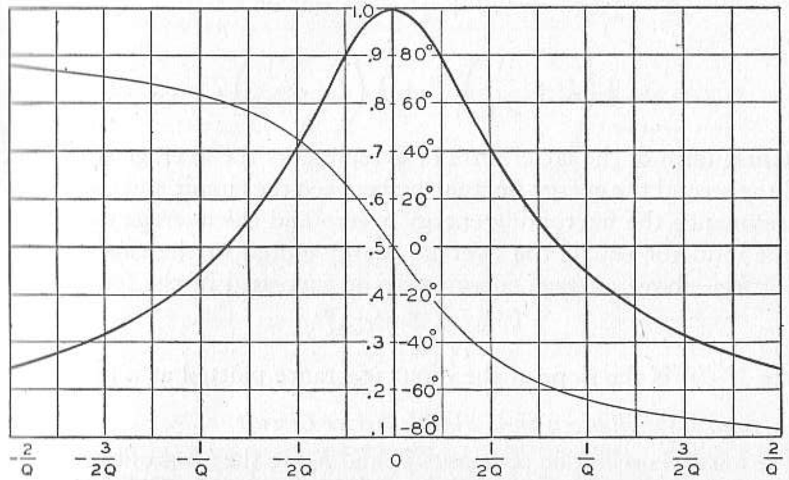


FIG. 5.21. Universal resonance curves. The heavy curve represents the amplitude and the light curve the phase.

Plotting the amplitude and phase of the ratio RY as a function of δ we obtain the *universal* resonance curves (Fig. 5.21). The heavy curve represents the absolute value of RY and is proportional to the amplitude of the current in the circuit when the impressed voltage is constant. The light curve is the phase of RY.

Other useful forms for the input admittance of the series circuit are

$$Y = \frac{\omega}{R\omega + i(\omega^2 - \bar{\omega}^2)L} = \frac{i\omega}{\left(\bar{\omega}^2 - \omega^2 + \frac{i\bar{\omega}\omega}{Q}\right)L}. \tag{8–12}$$

In the nondissipative case we have simply

$$Y = \frac{i\omega}{L(\bar{\omega}^2 - \omega^2)}. \qquad (8\text{–}13)$$

Let us now consider the instantaneous energy fluctuations. Choosing the origin of time so that the initial phase of I is zero, we have the following expressions for the instantaneous current in the circuit and the instantaneous voltage across the capacitor*

$$I_i = \hat{I}\cos\omega t, \quad V_{C,i} = \frac{\hat{I}}{\omega C}\cos\left(\omega t - \frac{\pi}{2}\right) = \frac{\hat{I}}{\omega C}\sin\omega t.$$

Therefore the energy stored in the inductor and the capacitor is

$$\mathfrak{S} = \tfrac{1}{2}L\hat{I}^2\cos^2\omega t + \tfrac{1}{2}C\frac{\hat{I}^2}{\omega^2C^2}\sin^2\omega t$$

$$= \tfrac{1}{4}\left(L + \frac{1}{\omega^2 C}\right)\hat{I}^2 + \tfrac{1}{4}\left(L - \frac{1}{\omega^2 C}\right)\hat{I}^2\cos 2\omega t.$$

The first term of the latter form of $\mathfrak{S}$ represents the average stored energy and the second the energy fluctuating between the circuit and the generator. At resonance the fluctuating energy is zero and the average energy stored in the inductor equals the average energy stored in the capacitor. Evidently the above average energy may be expressed in the form

$$\mathfrak{S}_a = \tfrac{1}{4}X'(\omega)\hat{I}^2, \qquad (8\text{–}14)$$

where $X'(\omega)$ is the slope of the input reactance plotted as a function of ω.

5.9. *Natural Oscillations in a Simple Series Circuit*

The natural oscillation constants $\hat{p}_1$ and $\hat{p}_2$ are the zeros of the impedance function (2.7–6). The impedance and admittance functions can be factored and thus expressed in terms of these zeros:

$$Z(p) = L\frac{(p - \hat{p}_1)(p - \hat{p}_2)}{p}, \quad Y(p) = \frac{p}{L(p - \hat{p}_1)(p - \hat{p}_2)}. \qquad (9\text{–}1)$$

Depending on the relative values of the circuit constants the natural oscillation constants may be either real or complex. Thus if

$$\frac{R}{2L} \geq \frac{1}{\sqrt{LC}}, \quad \text{or} \quad R \geq 2K,$$

the constants are real and the "oscillations" degenerate into an exponential decay. On the other hand if $R < 2K$, the oscillation constants are conjugate complex

* The impedance of the capacitor is $-i/\omega C$ and the voltage is lagging behind the current by 90°.

$$\hat{p}_1 = \hat{p} = \hat{\xi} + i\hat{\omega}, \quad \hat{p}_2 = \hat{p}^* = \hat{\xi} - i\hat{\omega}, \qquad (9\text{–}2)$$

where the amplitude constant $\hat{\xi}$ and the natural frequency $\hat{\omega}$ are given by

$$\hat{\xi} = -\frac{R}{2L} = -\frac{W}{2\mathfrak{S}} = -\frac{\bar{\omega}}{2Q},$$

$$\hat{\omega} = \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}} = \sqrt{\bar{\omega}^2 - \hat{\xi}^2} = \bar{\omega}\sqrt{1 - \frac{1}{4Q^2}}. \qquad (9\text{–}3)$$

For high Q circuits the natural frequency is nearly equal to the resonant frequency $\hat{\omega} = \bar{\omega}$.

5.10. *Forced Oscillations in a Simple Parallel Circuit*

The theory of a parallel combination of an inductor, a capacitor, and a resistor (Fig. 5.22) is very similar to the theory of the series circuit. The input admittance of the circuit is

$$Y = G + i\omega C + \frac{1}{i\omega L}. \qquad (10\text{–}1)$$

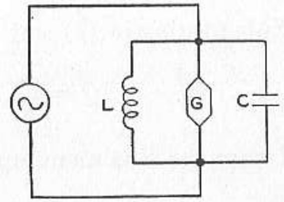


FIG. 5.22. A simple parallel circuit.

Comparing this with the input impedance (8–1) of the series circuit, we observe that the equations of section (8) can be adapted to parallel circuits if we interchange L and C, Z and Y, and replace R by G. We should also replace the characteristic impedance K by the characteristic admittance M; but subsequently it may be more convenient to reintroduce K. Thus we have the following expression for the input impedance of the circuit

$$GZ = \frac{1}{1 + iQ\left(\frac{\omega}{\bar{\omega}} - \frac{\bar{\omega}}{\omega}\right)}, \qquad (10\text{–}2)$$

where Q is defined by

$$Q = \frac{\omega\mathfrak{S}}{W} = \frac{M}{G} = \frac{1}{\bar{\omega}GL} = \frac{1}{KG}. \qquad (10\text{–}3)$$

If the input current is fixed, the voltage across the circuit varies with the frequency exactly as does the current in the case of the series circuit (Fig. 5.21).

Another type of parallel circuit is that shown in Fig. 5.23. In general the frequency characteristics of this circuit are different from those of the circuit in Fig. 5.22. If, however, both circuits have high Q values, then their behavior in the neighborhood of resonance is approximately the same.

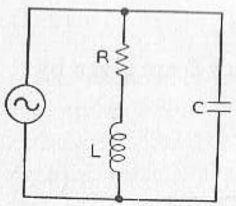


FIG. 5.23. Another type of simple parallel circuit.

To prove this, consider a parallel combination of two impedances Z_1 and Z_2 such that $|Z_2| \gg |Z_1|$. For the input impedance Z of this combination we have approximately

$$Z = \frac{Z_1 Z_2}{Z_1 + Z_2} \simeq Z_1\left(1 - \frac{Z_1}{Z_2}\right) = Z_1 - \frac{Z_1^2}{Z_2}.$$

If Z_1 is a pure reactance and Z_2 a pure resistance, then the last term is positive real and a large resistance Z_2 in parallel with Z_1 may be replaced by a small resistance in series with Z_1 or vice versa. Hence the circuits in Figs. 5.22 and 5.23 are approximately equivalent in the neighborhood of resonance if

$$R = \bar{\omega}^2 L^2 G \quad \text{or} \quad G = \frac{R}{\bar{\omega}^2 L^2} = \frac{R}{K^2}. \tag{10–4}$$

Substituting in (2) and (3), we have

$$Q = \frac{K}{R}, \quad Z = \frac{K^2}{R}\left[1 + iQ\left(\frac{\omega}{\bar{\omega}} - \frac{\bar{\omega}}{\omega}\right)\right]^{-1}. \tag{10–5}$$

Hence the maximum input impedance is

$$Z_{\max} = \frac{K^2}{R} = KQ. \tag{10–6}$$

If a generator is connected as shown in Fig. 5.24, then the maximum input impedance is

$$Z_{1,\max} = \frac{K_1^2}{R} = \frac{\bar{\omega}^2 L_1^2}{R}. \tag{10–7}$$

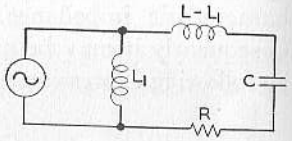


FIG. 5.24. A tapped parallel circuit.

Assuming that K_1 is still large compared with R, the effect of shifting the terminals of the circuit is to reduce the inductance to L_1 and to increase the effective capacity. The resonant frequency is evidently unchanged; but the maximum input impedance is reduced in the following ratio

$$\frac{Z_{1,\max}}{Z_{\max}} = \frac{L_1^2}{L^2}. \tag{10–8}$$

Of course, if L_1 is so small that $\bar{\omega}L_1$ is no longer large compared with R, the above formulae must be modified.

For the circuit shown in Fig. 5.22 and approximately for the one shown

in Fig. 5.23 we can obtain an equation similar to (8–14); thus

$$\mathcal{E}_a = \tfrac{1}{4}B'(\omega)\hat{V}^2, \tag{10–9}$$

where $\hat{V}$ is the maximum voltage amplitude across the circuit.

5.11. *Expansion of the Input Impedance Function*

The input impedance (1–20) as seen from the terminals of a generator in a typical mesh of an electric network is a rational fraction when considered as a function of the oscillation constant p. The numerator and the denominator are factorable and the impedance may be represented as a ratio of two products

$$Z(p) = A\frac{(p - \hat{p}_1)(p - \hat{p}_2)\cdots}{(p - p_1)(p - p_2)\cdots}, \tag{11–1}$$

where A is a constant. The zeros $\hat{p}_1, \hat{p}_2, \cdots$ of $Z(p)$ are infinities of $Y(p)$; they represent the natural oscillation constants of the network when the voltage across the input terminals is zero and hence when the terminals are short-circuited. The infinities $p_1, p_2, \cdots$ of $Z(p)$ are the zeros of $Y(p)$; they represent the natural oscillation constants of the network when the current through the input terminals is zero and hence when the network is open at these terminals.

A rational fraction can be expanded in partial fractions. If all the zeros of the admittance function are simple we may write $Z(p)$ in the following form

$$Z(p) = \frac{a_1}{p - p_1} + \frac{a_2}{p - p_2} + \cdots + f(p), \tag{11–2}$$

where $f(p)$ is a polynomial in p. Multiplying by $(p - p_1)$ and letting p approach p_1, we have

$$a_1 = \lim\,(p - p_1)Z(p) = \lim \frac{p - p_1}{Y(p)} \quad \text{as} \quad p \to p_1.$$

Therefore

$$a_1 = \frac{1}{Y'(p_1)}; \tag{11–3}$$

consequently we have

$$Z(p) = \sum \frac{1}{(p - p_m)Y'(p_m)} + f(p), \tag{11–4}$$

where the summation is extended over all the zeros of $Y(p)$.

Similarly the admittance function may be represented as follows

$$Y(p) = \sum \frac{1}{(p - \hat{p}_m)Z'(\hat{p}_m)} + g(p). \tag{11–5}$$

In network theory it is shown that $f(p)$ and $g(p)$ are polynomials of degree less than 2.

We have seen that the complex zeros and poles of $Z(p)$ occur in conjugate pairs;

thus for typical pairs of zeros and infinities we have

$$\hat{p}_m = \hat{\xi}_m + i\hat{\omega}_m, \quad \hat{p}_m^* = \hat{\xi}_m - i\hat{\omega}_m,$$
$$p_m = \xi_m + i\omega_m, \quad p_m^* = \xi_m - i\omega_m. \tag{11-6}$$

The real parts ξ_m and $\hat{\xi}_m$ are never positive since positive values would mean that the amplitudes of the currents and voltages in the network were steadily increasing. Then infinite power would be dissipated in the resistors and infinite energy stored in the inductors and capacitors without a continuous operation of an impressed force, that is, without a continuous supply of energy to the network.

Let us now consider the values of $Z(p)$ and $Y(p)$ on the imaginary axis

$$Z(i\omega) = R(\omega) + iX(\omega),$$
$$Y(i\omega) = G(\omega) + iB(\omega). \tag{11-7}$$

$R(\omega)$ and $G(\omega)$ are never negative; if they were negative for some value of ω, then at this frequency power would be contributed to the generator by the supposedly passive network. If the network is only slightly dissipative, $R(\omega)$ and $G(\omega)$ are small. In this case the zeros of $Z(p)$ and $Y(p)$ are given approximately by

$$X(\hat{\omega}_m) = 0, \quad B(\omega_m) = 0. \tag{11-8}$$

In order to obtain the second approximation we note that

$$Z(i\hat{\omega}_m + \hat{\delta}_m) = Z(i\hat{\omega}_m) + \hat{\delta}_m Z'(i\hat{\omega}_m) + \cdots = 0,$$
$$Y(i\omega_m + \delta_m) = Y(i\omega_m) + \delta_m Y'(i\omega_m) + \cdots = 0. \tag{11-9}$$

Solving, we obtain

$$\hat{\delta}_m = -\frac{Z(i\hat{\omega}_m)}{Z'(i\hat{\omega}_m)}, \quad \delta_m = -\frac{Y(i\omega_m)}{Y'(i\omega_m)}. \tag{11-10}$$

Differentiating (7)

$$iZ'(i\hat{\omega}_m) = R'(\hat{\omega}_m) + iX'(\hat{\omega}_m),$$
$$iY'(i\omega_m) = G'(\omega_m) + iB'(\omega_m), \tag{11-11}$$

and substituting in (10), we obtain

$$\hat{\delta}_m = -\frac{R(\hat{\omega}_m)}{X'(\hat{\omega}_m) - iR'(\hat{\omega}_m)} = -\frac{R(\hat{\omega}_m)}{X'(\hat{\omega}_m)} - i\frac{R(\hat{\omega}_m)R'(\hat{\omega}_m)}{[X'(\hat{\omega}_m)]^2},$$
$$\delta_m = -\frac{G(\omega_m)}{B'(\omega_m) - iG'(\omega_m)} = -\frac{G(\omega_m)}{B'(\omega_m)} - i\frac{G(\omega_m)G'(\omega_m)}{[B'(\omega_m)]^2}. \tag{11-12}$$

Thus the approximate expressions for the real parts of the natural oscillation constants are

$$\hat{\xi}_m = -\frac{R(\hat{\omega}_m)}{X'(\hat{\omega}_m)}, \quad \xi_m = -\frac{G(\omega_m)}{B'(\omega_m)}. \tag{11-13}$$

Since R and G are positive and the ξ's are negative, we have the following inequalities

$$X'(\hat{\omega}_m) > 0, \quad B'(\omega_m) > 0. \tag{11-14}$$

Substituting the above approximations in the general equations (4) and (5), we have*

$$Z(i\omega) = \sum \frac{2i\omega}{(\omega_m^2 - \omega^2 - 2i\omega\xi_m)B'(\omega_m)} + f(i\omega),$$
$$Y(i\omega) = \sum \frac{2i\omega}{(\hat{\omega}_m^2 - \omega^2 - 2i\omega\hat{\xi}_m)X'(\hat{\omega}_m)} + g(i\omega). \tag{11-15}$$

Comparing the second of the above equations with (8–12), we find that a slightly dissipative network behaves like a parallel combination of simple series circuits whose inductances and Q's are given by

$$L_m = \tfrac{1}{2}X'(\hat{\omega}_m), \quad \hat{Q}_m = -\frac{\hat{\omega}_m}{2\hat{\xi}_m} = \frac{\hat{\omega}_m X'(\hat{\omega}_m)}{2R(\hat{\omega}_m)}. \tag{11-16}$$

Likewise we can regard the network as approximately equivalent to a series combination of parallel resonant circuits whose capacitances and Q's are

$$C_m = \tfrac{1}{2}B'(\omega_m), \quad Q_m = \frac{\omega_m B'(\omega_m)}{2G(\omega_m)}. \tag{11-17}$$

In view of (8–14) and (10–9), equations (15) can be expressed in terms of the energies stored in the circuit at the various resonant frequencies. Thus we have

$$Z(i\omega) = \sum \frac{i\omega}{2\mathcal{E}_m\left(\omega_m^2 - \omega^2 + \dfrac{i\omega\omega_m}{Q_m}\right)} + f(i\omega), \tag{11-18}$$

where $\mathcal{E}_m$ is the energy stored in the circuit at the mth resonant frequency when the input terminals are open and when the voltage amplitude at these terminals is unity. Similarly we have

$$Y(i\omega) = \sum \frac{i\omega}{2\hat{\mathcal{E}}_m\left(\hat{\omega}_m^2 - \omega^2 + \dfrac{i\omega\hat{\omega}_m}{\hat{Q}_m}\right)} + g(i\omega), \tag{11-19}$$

where $\hat{\mathcal{E}}_m$ is the energy stored in the circuit at the mth resonant frequency when the input terminals are short-circuited and when the current amplitude at these terminals is unity.

So far we have tacitly assumed that none of the natural oscillation constants falls on the real axis. Let us now suppose that $p = p_0$ is a real simple zero of $Y(p)$.

* It should be recalled that $X(\omega)$ and $B(\omega)$ are odd functions and therefore $X'(\omega)$ and $B'(\omega)$ are even. We retain only the principal terms in the final approximations.

Then the corresponding term in (4) is not paired with any other; it is

$$Z_0(p) = \frac{1}{(p - p_0)Y'(p_0)}. \tag{11–20}$$

If p_0 is small, we have approximately

$$p_0 = \delta_0 = -\frac{G(0)}{B'(0)}, \quad Y'(p_0) = B'(0);$$

so that, for $p = i\omega$, equation (20) becomes

$$Z_0(i\omega) = \frac{1}{G(0) + i\omega B'(0)}. \tag{11–21}$$

$G(0)$ is the direct current conductance of the network and $B'(0)$ the direct current capacitance; thus

$$G(0) = G = W_0, \quad B'(0) = C = 2\mathcal{E}_0, \tag{11–22}$$

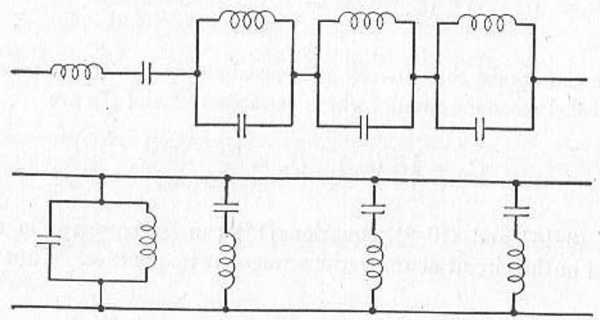


FIG 5.25. Two equivalent representations of a general reactive network.

where W_0 is the power lost in the conductance and $\mathcal{E}_0$ is the energy stored in the capacitance when a unit voltage is applied to the network. The term (21), which is to be added to (18), now becomes

$$Z_0(i\omega) = \frac{1}{G + i\omega C} = \frac{1}{W_0 + 2i\omega\mathcal{E}_0}. \tag{11–23}$$

Similarly if $Z(p)$ has a simple zero on the real axis, we should add to (19) the following term

$$Y_0(i\omega) = \frac{1}{R(0) + i\omega X'(0)} = \frac{1}{R + i\omega L} = \frac{1}{\hat{W}_0 + 2i\omega\hat{\mathcal{E}}_0}, \tag{11–24}$$

where R and L are respectively the d-c resistance and inductance of the network, $\hat{W}_0$ is the power lost in the resistance and $\hat{\mathcal{E}}_0$ is the energy stored in the inductance when a unit current is passing through the input terminals.

The foregoing results may be summarized graphically as shown in Fig. 5.25. A purely reactive finite network may be represented either as a series combination of a series resonant circuit (or in a special case nonresonant) and a succession of parallel resonant circuits or as a parallel combination of a parallel resonant circuit and a succession of series resonant circuits. In the first case the series circuit is obtained when one of the parallel resonant circuits degenerates into an inductance and another into a capacitance; the parallel circuit in the second case is obtained similarly. The degenerate circuits correspond to the zeros and poles of $Z(p)$ at the origin and at infinity. If the network is slightly dissipative, then in the first approximation the equivalent networks will differ from those in Fig. 5.25 only in that each series and parallel branch will contain a resistance element.

CHAPTER VI

About Waves in General

6.0. *Introduction*

If the impressed currents are known throughout an infinite homogeneous medium, the field can be calculated fairly easily; we need only obtain the field of a current element and then use the principle of superposition. The solution of this problem is useful even though most practical problems are concerned with fields in media composed of homogeneous parts and not in completely homogeneous media. Thus if the medium is homogeneous except for isolated islands, it is sometimes possible to obtain approximate polarization currents which can be used as virtual sources in an otherwise homogeneous medium (section 2.11). The first few sections of this chapter are devoted to this problem.

The boundaries between media with different electromagnetic properties or the " discontinuities " may have a profound effect on wave propagation. In a homogeneous medium, for example, the energy from a given source will travel in all directions; but in the presence of parallel conducting wires at least a fraction of this energy will flow in the direction of the wires. The effects of such discontinuities will be studied in detail in subsequent chapters; but some general considerations are introduced in this chapter.

A brief discussion of electrostatics and magnetostatics is also included in this chapter. These topics are of interest in wave theory for the following two reasons: (1) they furnish approximations to slowly varying fields, (2) they furnish exact solutions of certain two-dimensional wave problems.

6.1. *The Field Produced by a Given Distribution of Currents in an Infinite Homogeneous Medium*

Our problem is to solve the electromagnetic equations (4.4–2) for harmonic fields. This is the most important case in practice; besides, the solution of the most general case can then be expressed in the form of a contour integral in the oscillation constant plane. The usual procedure for solving a simultaneous system of equations is to eliminate all dependent variables except one. In the present case this procedure would be unnecessarily restrictive since we should have to differentiate J and M and hence assume that they are continuous and differentiable functions. In practical problems J and M are localized and for all practical purposes

the regions occupied by them have sharp boundaries. Thus it is convenient to introduce a set of auxiliary functions, generally called potential functions.

To begin with let us write

$$E = E' + E'', \quad H = H' + H'', \tag{1-1}$$

where (E',H') and (E'',H'') are solutions of

$$\begin{aligned} \text{curl } E' &= -i\omega\mu H', \qquad & \text{curl } H' &= J + (g + i\omega\epsilon)E', \\ \text{curl } E'' &= -M - i\omega\mu H'', \quad & \text{curl } H'' &= (g + i\omega\epsilon)E''. \end{aligned} \tag{1-2}$$

The field (E',H') is produced by electric currents and (E'',H'') by magnetic currents. Their sum satisfies (4.4–2).

Taking the divergence of each equation in the set (2), we have

$$\begin{aligned} \text{div } H' &= 0, \qquad & \text{div } E' &= -\frac{\text{div } J}{g + i\omega\epsilon}, \\ \text{div } H'' &= -\frac{\text{div } M}{i\omega\mu}, \quad & \text{div } E'' &= 0. \end{aligned} \tag{1-3}$$

The second and third of these equations require J and M to be continuous and differentiable; but one form of the solution of our problem is obtained without using these equations. In the other form of the solution which depends on them we may assume J and M differentiable to begin with and then extend the results to include discontinuous distributions. The first and last equations show that H' and E'' can be represented as the curls of certain vector point functions

$$H' = \text{curl } A, \quad E'' = -\text{curl } F. \tag{1-4}$$

Substituting from (4) into (2), we obtain

$$E' = -i\omega\mu A - \text{grad } V, \quad H'' = -(g + i\omega\epsilon)F - \text{grad } U, \tag{1-5}$$

where V and U are two new point functions which are introduced because the equality of the curls of two vectors does not imply that the vectors are identical.

From (4) and (5) and the two remaining equations in (2), we obtain

$$\begin{aligned} \text{curl curl } A &= J - \sigma^2 A - (g + i\omega\epsilon) \text{ grad } V, \\ \text{curl curl } F &= M - \sigma^2 F - i\omega\mu \text{ grad } U. \end{aligned} \tag{1-6}$$

Using (1.8–2) we have

$$\begin{aligned} \Delta A - \text{grad div } A &= -J + \sigma^2 A + (g + i\omega\epsilon) \text{ grad } V, \\ \Delta F - \text{grad div } F &= -M + \sigma^2 F + i\omega\mu \text{ grad } U. \end{aligned} \tag{1-7}$$

Thus we have expressed E and H in terms of two vectors A and F and two scalars V and U, the new functions being connected by two vector equations. So far the vectors are somewhat arbitrary since equations (4) are unchanged if we add to A and F the gradients of arbitrary functions. The functions V and U are completely arbitrary. Hence we have an opportunity to impose further conditions on these functions to suit our convenience. For instance we may set

$$V = -\frac{\operatorname{div} A}{g + i\omega\epsilon}, \quad U = -\frac{\operatorname{div} F}{i\omega\mu}, \tag{1–8}$$

so that equations (7) become

$$\Delta A = \sigma^2 A - J, \quad \Delta F = \sigma^2 F - M. \tag{1–9}$$

When specified in the above manner, the functions A, F, V and U are called *wave potentials*, the first two being vector potentials and the last two scalar potentials. More specifically A is called the magnetic vector potential, F the electric vector potential, V the electric scalar potential and U the magnetic scalar potential. Lorentz was the first to introduce these wave potentials in dealing with nondissipative media and he called them *retarded potentials* for reasons that will soon become obvious. In general the wave potentials are not only "retarded" but also "attenuated," and the more general designation "wave potentials" is more appropriate.

Thus we have the following expressions for the field produced by a given distribution of impressed currents

$$E = -i\omega\mu A - \operatorname{grad} V - \operatorname{curl} F,$$
$$H = \operatorname{curl} A - \operatorname{grad} U - (g + i\omega\epsilon)F, \tag{1–10}$$

where V and U are defined by (8) and A and F are the solutions of (9).

If J and M are differentiable functions, V and U satisfy equations similar to (9). Thus taking the divergence of (9) and substituting from (8), we have

$$\Delta V = \sigma^2 V + \frac{\operatorname{div} J}{g + i\omega\epsilon}, \quad \Delta U = \sigma^2 U + \frac{\operatorname{div} M}{i\omega\mu}. \tag{1–11}$$

In nondissipative media div J is equal to the negative time derivative of the volume density of electric charge; thus

$$\operatorname{div} J = -i\omega q_v, \qquad \operatorname{div} M = -i\omega m_v. \tag{1–12}$$

Consequently

$$\Delta V = -\beta^2 V - \frac{1}{\epsilon} q_v, \quad \Delta U = -\beta^2 U - \frac{1}{\mu} m_v. \tag{1–13}$$

From the physical point of view the vector potentials can be obtained much more satisfactorily by the method explained in the next section than by solving equations (9) formally.

6.2. *The Field of an Electric Current Element*

Consider a short current filament (Fig. 6.1) and assume that the current I is uniform and steady between the terminals A and B so that the entire current is forced to flow out of B into the external medium and then back into A. If the medium is a perfect dielectric this would mean a concentration of electric charge at B at the rate I amperes per second and an ever increasing electric field around the filament. The product Il of the current and the length of the filament is called the *moment* of the electric current element.

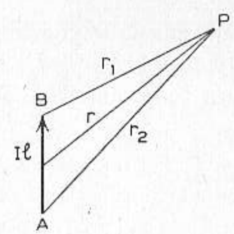


FIG. 6.1. An electric current element.

Let us suppose that the current element is centered at the origin along the z-axis. From a point source the current would flow outwards uniformly in all directions; the density would then be

$$J_r = \frac{I}{4\pi r^2} = \frac{\partial}{\partial r}\left(-\frac{I}{4\pi r}\right).$$

Hence for two point sources separated by distance l the current density, at distances large compared with l, is the gradient of the following function

$$P = -l\frac{\partial}{\partial z}\left(-\frac{I}{4\pi r}\right) = -\frac{Il\cos\theta}{4\pi r^2}.$$

Consequently

$$J_r = \frac{Il\cos\theta}{2\pi r^3}, \quad J_\theta = \frac{Il\sin\theta}{4\pi r^3}. \tag{2–1}$$

The dotted lines in Fig. 6.2 are the flow lines.

The magnetic lines of force are circles coaxial with the element and in order to obtain the magnetic intensity we need only calculate the magnetomotive force round the circumference of a circle PP' coaxial with the element (Fig. 6.2). This magnetomotive force is equal to the electric current $I(\theta)$ passing through any surface bounded by PP'. Choosing the surface as a sphere concentric with the origin, we have

$$I(\theta) = \int_0^{2\pi}\int_0^{\theta} J_r r^2 \sin\theta \, d\theta \, d\varphi = \frac{Il\sin^2\theta}{2r}; \tag{2–2}$$

hence

$$H_\varphi = \frac{I(\theta)}{2\pi r \sin\theta} = \frac{Il \sin\theta}{4\pi r^2}. \tag{2–3}$$

Equation (1–10) shows that H is expressible as the curl of a vector A. From (1–9) we conclude that each cartesian component of A depends only

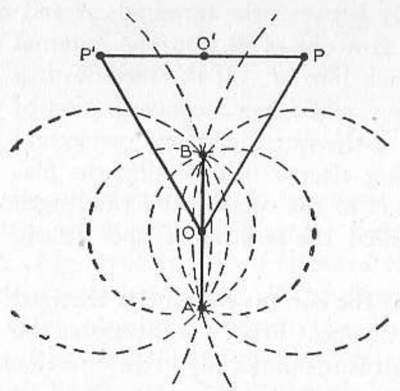


FIG. 6.2. Electric lines of force in the vicinity of an electric current element.

on the corresponding component of the impressed current density. Thus in the present case A is parallel to the z-axis and we should have

$$H_\varphi = -\frac{\partial A_z}{\partial \rho} = -\frac{\partial A_z}{\partial r}\frac{\partial r}{\partial \rho} = -\frac{\partial A_z}{\partial r}\sin\theta. \tag{2–4}$$

Comparing with (3), we have

$$A_z = \frac{Il}{4\pi r}. \tag{2–5}$$

Let us now suppose that the current is a harmonic function of time. As the frequency approaches zero the field must approach that given by the above equations where I, A, H are now complex amplitudes of the corresponding quantities and the time factor $e^{i\omega t}$ is omitted. From (1) we obtain the electric intensities

$$E_r = \frac{Il\cos\theta}{2\pi(g + i\omega\epsilon)r^3}, \quad E_\theta = \frac{Il\sin\theta}{4\pi(g+i\omega\epsilon)r^3}. \tag{2–6}$$

Next we seek that solution of Maxwell's equations which approaches (6) as $\omega \to 0$. At all points external to the current element the magnetic vector potential A must satisfy (1–9) with $J = 0$ and by (5) it must be independent of θ and φ; hence

$$\frac{\partial}{\partial r}\left(r^2\frac{\partial A_z}{\partial r}\right) = \sigma^2 r^2 A_z.$$

This is equation (3.1–15) with $k = 0$ and its general solution is

$$A_z = \frac{Pe^{-\sigma r}}{r} + \frac{Qe^{\sigma r}}{r}.$$

In dissipative media the second term increases exponentially with the distance from the element and hence cannot represent the field produced by the element. The first term approaches (5) as ω (and therefore σ) approaches zero if $P = Il/4\pi$. Thus we have

$$A_z = \frac{Ile^{-\sigma r}}{4\pi r}. \tag{2–7}$$

Nondissipative media may be regarded as limiting cases of dissipative media and then

$$A_z = \frac{Ile^{-i\beta r}}{4\pi r}. \tag{2–8}$$

By (1–8) we have

$$V = -\frac{1}{g + i\omega\epsilon}\frac{\partial A_z}{\partial z} = \frac{\eta Il}{4\pi r}\left(1 + \frac{1}{\sigma r}\right)e^{-\sigma r}\cos\theta. \tag{2–9}$$

The field intensities are now obtained from (1–10); thus

$$E_r = \frac{\eta Il}{2\pi r^2}\left(1+\frac{1}{\sigma r}\right)e^{-\sigma r}\cos\theta, \quad H_\varphi = \frac{\sigma Il}{4\pi r}\left(1+\frac{1}{\sigma r}\right)e^{-\sigma r}\sin\theta,$$
$$E_\theta = \frac{i\omega\mu Il}{4\pi r}\left(1 + \frac{1}{\sigma r} + \frac{1}{\sigma^2 r^2}\right)e^{-\sigma r}\sin\theta. \tag{2–10}$$

In nondissipative media we have

$$E_r = \frac{\eta Il}{2\pi r^2}\left(1+\frac{1}{i\beta r}\right)e^{-i\beta r}\cos\theta, \quad H_\varphi = \frac{iIl}{2\lambda r}\left(1+\frac{1}{i\beta r}\right)e^{-i\beta r}\sin\theta,$$
$$E_\theta = \frac{i\eta Il}{2\lambda r}\left(1 + \frac{1}{i\beta r} - \frac{1}{\beta^2 r^2}\right)e^{-i\beta r}\sin\theta. \tag{2–11}$$

If $\omega \to 0$, these expressions approach (3) and (6).

Similarly the field of a magnetic current element of moment Kl is obtained from the following electric vector potential

$$F_z = \frac{Kle^{-\sigma r}}{4\pi r}. \tag{2–12}$$

Any given distribution of applied currents may be subdivided into elements, and the field can be obtained by superposition of the fields of individual elements. Take an infinitesimal volume bounded by the lines of flow and two surfaces normal to them. The current in this element is $I = J\,dS$, where dS is the cross-section of the tube of flow; hence, the moment Il is equal to $J\,dv$, where dv is the volume of the element. Thus we shall have

$$A = \iiint \frac{Je^{-\sigma r}}{4\pi r}\,dv, \quad F = \iiint \frac{Me^{-\sigma r}}{4\pi r}\,dv, \tag{2–13}$$

where r is the distance between a typical element and a typical point in space. The scalar potentials are then computed from (1–8).

If, however, J and M are differentiable functions, then V and U can also be computed from equations similar to the above. We note the similarity between equations (1–9), (1–11), and (1–12) and write

$$V = -\iiint \frac{\operatorname{div} J\,e^{-\sigma r}}{4\pi(g + i\omega\epsilon)r}\,dv, \quad U = -\iiint \frac{\operatorname{div} M\,e^{-\sigma r}}{4\pi i\omega\mu r}\,dv,$$
$$V = \iiint \frac{q_v e^{-i\beta r}}{4\pi\epsilon r}\,dv, \qquad U = \iiint \frac{m_v e^{-i\beta r}}{4\pi\mu r}\,dv. \tag{2–14}$$

These equations can be extended to include the case of nondifferentiable J and M by adding appropriate surface integrals, and, more generally, by adding line integrals and discrete terms representing the potentials of point charges. Thus in the case of an electric current element in a nondissipative medium (Fig. 6.1) we have two point charges at the terminals

$$q_A = -\frac{I}{i\omega}, \quad q_B = \frac{I}{i\omega}, \tag{2–15}$$

and the scalar potential is then

$$V = \frac{I}{4\pi i\omega\epsilon}\left(\frac{e^{-i\beta r_1}}{r_1} - \frac{e^{-i\beta r_2}}{r_2}\right), \tag{2–16}$$

which leads to (9) when l is very small compared with r.

For surface and line distributions of impressed currents, the expressions for A and F are similar to (13), the surface and line integrals appearing in place of the volume integrals. For any filament carrying current $I(s)$ we have

$$A = \int \frac{I(s)e^{-\sigma r}}{4\pi r}\,\overline{ds}, \tag{2–17}$$

where $\overline{ds}$ is a directed element of length.

We shall now define the terms "large distance" and "small distance" as used in wave theory. A given distance r is large if $|\sigma r| \gg 1$ and small if $|\sigma r| \ll 1$. In perfect dielectrics this means that r is large if $\beta r = 2\pi r/\lambda$ is large compared with unity; r is small if βr is small compared with unity. For example, $r = 2\lambda$ is fairly large since $2\beta\lambda = 4\pi = 12.57$; on the other hand $r = \lambda/80$ is small to about the same degree since $\beta\lambda/80 = \pi/40 = 1/12.7$. The length $r = \lambda/2\pi = 0.16\lambda$ may be taken as the reference length.

At large distances from the element the field is particularly simple. Thus in a nondissipative medium we have approximately

$$H_\varphi = \frac{iIl}{2\lambda r}\,e^{-i\beta r}\sin\theta, \quad E_\theta = \eta H_\varphi, \quad E_r = 0. \tag{2–18}$$

6.3. *Radiation from an Electric Current Element*

The flow of power across an infinitely large sphere concentric with the element is*

$$W = \tfrac{1}{2}\int_0^{2\pi}\int_0^{\pi} E_\theta H_\varphi^* r^2 \sin\theta\,d\theta\,d\varphi = \frac{\eta I^2 l^2}{8\lambda^2}\int_0^{2\pi}\int_0^{\pi}\sin^3\theta\,d\theta\,d\varphi.$$

Integrating, we have

$$W = \tfrac{1}{3}\pi\eta\left(\frac{Il}{\lambda}\right)^2 = \tfrac{1}{3}\pi\eta\frac{p^2}{\lambda^2} = 40\pi^2\left(\frac{Il}{\lambda_0}\right)^2 = \frac{40\pi^2 p^2}{\lambda_0^2}, \tag{3–1}$$

where p is the moment of the element; the last two expressions represent the power radiated in free space.

This radiated power must come from the source; it can also be calculated from the work done by the electromotive force impressed on the current element. From (2–11) we have the electric intensity on the axis of the element

$$E_z = \frac{\eta Il}{2\pi}\left(\frac{1}{i\beta r^3} - \frac{i\beta}{2r} - \tfrac{1}{3}\beta^2 + \cdots\right).$$

The first two terms are in quadrature with I and on the average do no work; but the third term is 180° out of phase with I and work is done against the field by the impressed electromotive force. The in-phase component of this force is then

$$\operatorname{re}(V^i) = -l\operatorname{re}(E_z) = \frac{\eta}{6\pi}\beta^2 l^2 I = \frac{2\pi\eta l^2 I}{3\lambda^2}. \tag{3–2}$$

* By properly choosing the origin of time we can make the initial phase of I zero; hence I will be real and $II^* = I^2$.

The work done by this force per second is seen to be equal to W in (1). The ratio

$$R = \frac{\mathrm{re}(V^i)}{I} = \frac{2\pi\eta l^2}{3\lambda^2} = 80\pi^2\left(\frac{l}{\lambda_0}\right)^2 \tag{3–3}$$

is called the radiation resistance of the element.

The reactive forces in the vicinity of the element are very large. Assuming a finite radius for the element, we can compute these forces at the element itself; since, however, they depend on r, we should subdivide the element into smaller elements and then integrate the effects. In order to sustain a uniform current the impressed forces must be distributed along the entire element. It will be shown (section 6.8) that if the element is energized at the center (Fig. 6.3), the current distribution is approximately linear. Then the *moment of the current distribution* is

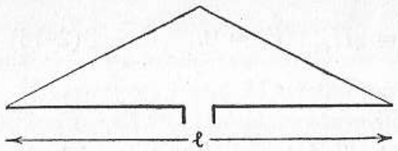


FIG. 6.3. A short wire energized at the center.

$$p = \int I(s)\, ds = \tfrac{1}{2}Il, \tag{3–4}$$

where I is now the input current at the center. If the element is short the distant field and hence the radiated power will be determined by the moment. Thus from (1) we obtain

$$W = \frac{10\pi^2 I^2 l^2}{\lambda_0^2}, \quad R = 20\pi^2\frac{l^2}{\lambda_0^2} = 197\frac{l^2}{\lambda_0^2}. \tag{3–5}$$

It is sometimes convenient to express the field of the current element in terms of the radiated power. From (1) we obtain

$$p = \frac{\sqrt{12\pi W}}{\beta\sqrt{\eta}} = \lambda\sqrt{\frac{3W}{\pi\eta}} = \frac{\lambda_0\sqrt{W}}{2\pi\sqrt{10}}; \tag{3–6}$$

hence for the distant field in free space we have

$$|H_\varphi| = \frac{\sqrt{W}\sin\theta}{4\pi\sqrt{10}\,r}, \quad |E_\theta| = \frac{3}{r}\sqrt{10W}\sin\theta. \tag{3–7}$$

6.4. *The Mutual Impedance between Two Current Elements and the Mutual Radiated Power*

Consider two infinitesimal current elements of moments $I_1\,ds_1$ and $I_2\,ds_2$, and let ψ be the angle between the positive direction of one element and the E-vector of the other (Fig. 6.4). Let E_1 be the electric intensity due to the first element and E_2 the intensity due to the second. The electromotive force of the field of the first element acting along the second element is $E_{1,s}(s_2)\,ds_2$, where $E_{1,s}(s_2)$ is the component of E_1 in the direction of ds_2; hence the electromotive force which should be impressed on the second element in order to counteract the force of this field is

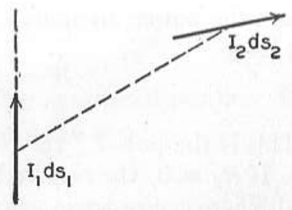


FIG. 6.4. Two current elements.

$$-E_{1,s}(s_2)\,ds_2. \tag{4–1}$$

The ratio of this impressed force to the current in the first element is the mutual impedance between two current elements

$$Z_{12} = -\frac{E_{1,s}(s_2)\,ds_2}{I_1} = -\frac{E_{2,s}(s_1)\,ds_1}{I_2}$$
$$= -\frac{E_1(s_2)\cos\psi\,ds_2}{I_1} = -\frac{E_2(s_1)\cos\psi\,ds_1}{I_2}. \tag{4–2}$$

The reciprocity implied by this equation follows immediately on writing explicit expressions for the forces involved. For example, the mutual impedance of two distant parallel elements, perpendicular to the line joining their centers (Fig. 6.5) is

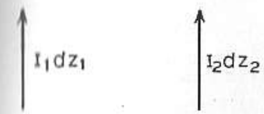


FIG. 6.5. Two parallel current elements.

$$Z_{12} = \frac{i\eta e^{-i\beta r}}{2\lambda r}\,dz_1\,dz_2. \tag{4–3}$$

If the reactive components of the self-impedances Z_{11} and Z_{22} are tuned out, then

$$V_1 = R_{11}I_1 + Z_{12}I_2, \qquad V_2 = Z_{12}I_1 + R_{22}I_2, \tag{4–4}$$

where V_1 and V_2 are the applied electromotive forces. If the elements are of equal length, then

$$R_{11} = R + R_1, \quad R_{22} = R + R_2, \quad R = 80\pi^2\left(\frac{l}{\lambda}\right)^2, \tag{4–5}$$

where R_1 and R_2 are the internal resistances of the elements and R is the radiation resistance.

If $V_2 = 0$, then

$$I_2 = -\frac{Z_{12}I_1}{R + R_2}, \tag{4–6}$$

and the power dissipated in R_2 is

$$W = \tfrac{1}{2}R_2 I_2 I_2^* = \frac{R_2 Z_{12} Z_{12}^*}{2(R + R_2)^2} I_1 I_1^*. \tag{4–7}$$

This is the power "received" by the "load" resistance R_2.

If $R_2 = 0$, the received power is zero; if $R_2 = \infty$, the received power is also zero. For some value of R_2 the received power must be a maximum; this maximum value is obtained from

$$\frac{\partial W}{\partial R_2} = 0. \tag{4–8}$$

Thus we find that for maximum reception the load resistance must "match" the radiation resistance

$$R_2 = R = 80\pi^2 \left(\frac{l}{\lambda}\right)^2. \tag{4–9}$$

The received power is then

$$W_{\max} = \frac{Z_{12} Z_{12}^*}{8R} | I_1 |^2. \tag{4–10}$$

Substituting from (3) and (9), we obtain*

$$W_{\max} = \frac{3\eta \, | I_1 l |^2}{64\pi r^2} = \frac{45}{8r^2} | I_1 l |^2. \tag{4–11}$$

In terms of the power radiated by the first element, we have

$$| I_1 |^2 = \frac{2W_t}{R};$$

hence the received power is

$$W_r = \frac{45 l^2 W_t}{4Rr^2} = \left(\frac{3\lambda}{8\pi r}\right)^2 W_t = 0.0142 \left(\frac{\lambda}{r}\right)^2 W_t. \tag{4–12}$$

Equation (12) gives the power received by the load resistance R_2; the total power $\hat{W}$ received by the second element may be taken as

$$\hat{W} = \tfrac{1}{2}(R + R_2) I_2 I_2^* = \frac{Z_{12} Z_{12}^*}{2(R + R_2)} I_1 I_1^*, \tag{4–13}$$

of which the following amount

$$\overline{W} = \tfrac{1}{2} R I_2 I_2^* = \frac{R Z_{12} Z_{12}^*}{2(R + R_2)^2} I_1 I_1^* \tag{4–14}$$

* The first expression holds for any dielectric and the second is for free space.

is "reradiated." When $R_2 = R$, we have

$$W_r = \overline{W} = \tfrac{1}{2}\hat{W}, \tag{4–15}$$

and the power absorbed by the load is equal to the reradiated power. If $R_2 = 0$, the "received" power is completely reradiated and

$$\overline{W} = \hat{W} = \frac{Z_{12} Z_{12}^*}{2R} I_1 I_1^*. \tag{4–16}$$

The power radiated by the two elements is the real part of (5.2–8)

$$W = \tfrac{1}{2}(R_{11}\hat{I}_1^2 + 2R_{12}\hat{I}_1\hat{I}_2 \cos\vartheta + R_{22}\hat{I}_2^2), \tag{4–17}$$

where $\hat{I}_1$ and $\hat{I}_2$ are the amplitudes of I_1 and I_2 and ϑ is the phase difference. In this equation R_{11} and R_{22} are, of course, the radiation resistances

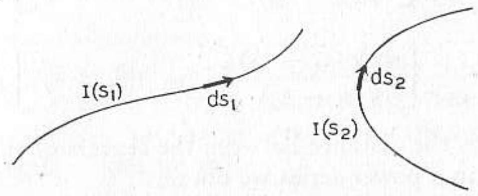


FIG. 6.6. Two current filaments.

of the elements and do not include the internal resistances of the generators driving the currents. We have seen that the R's are proportional to the products of the lengths of the elements

$$R_{12} = k_{12}\, ds_1\, ds_2, \quad R_{11} = k_{11}\, ds_1^2, \quad R_{22} = k_{22}\, ds_2^2. \tag{4–18}$$

It is also evident that

$$k_{11} = k_{22}. \tag{4–19}$$

The total power radiated by any two current filaments of arbitrary shape and length (Fig. 6.6) can be expressed in the form

$$W = W_{11} + 2W_{12} + W_{22}, \tag{4–20}$$

where W_{11} is the work done by the impressed forces in sustaining the current in the first filament against the forces produced by this current, with W_{22} defined similarly for the second filament; W_{12} is the work done in sustaining the current in the first filament against the forces produced by the current in the second filament. While W_{11} and W_{22} are inherently positive, W_{12} may be either positive or negative. For the mutual power radiated by two arbitrary filaments we have

$$2W_{12} = \int\int k_{12}(s_1, s_2)\hat{I}(s_1)\hat{I}(s_2) \cos\vartheta \, ds_1\, ds_2. \tag{4–21}$$

For *coincident* filaments this becomes

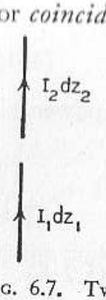


Fig. 6.7. Two coaxial current elements.

$$W_{11} = W_{12}$$

$$= \tfrac{1}{2}\int\int k_{12}(s_1,s_2)\hat{I}(s_1)\hat{I}(s_2)\cos\vartheta\, ds_1\, ds_2. \tag{4–22}$$

If the elements are collinear (Fig. 6.7), then

$$R_{12} = -\frac{\text{re}(E_r\, dz_2)}{I_1}, \tag{4–23}$$

where E_r is due to the lower element. Using (2–11) we obtain

$$R_{12} = \frac{\eta}{2\pi(z_1 - z_2)^2}\left[\frac{\sin\beta(z_1 - z_2)}{\beta(z_1 - z_2)} - \cos\beta(z_1 - z_2)\right]dz_1\, dz_2,$$

$$k_{12} = \frac{\eta}{2\pi(z_1 - z_2)^2}\left[\frac{\sin\beta(z_1 - z_2)}{\beta(z_1 - z_2)} - \cos\beta(z_1 - z_2)\right], \tag{4–24}$$

where $|z_1 - z_2|$ is the distance between the centers of the elements.

Expanding k_{12} in a power series we obtain

$$k_{12} = \frac{2\pi\eta}{3\lambda^2}\left[1 - \frac{2\pi^2(z_1 - z_2)^2}{5\lambda^2} + \frac{2\pi^4(z_1 - z_2)^4}{35\lambda^4} - \cdots\right]. \tag{4–25}$$

6.5. *Impressed Currents Varying Arbitrarily with Time*

In nondissipative media the vector potential of a given electric current distribution varying harmonically with time is

$$A = \int\int\int \frac{Je^{-i\beta r}}{4\pi r}\, dv. \tag{5–1}$$

If the phase of J is ωt, the phase of the corresponding component of the vector potential is $\omega t - \beta r = \omega[t - (r/v)]$, where v is the characteristic velocity of the medium. The time delay r/v is independent of the frequency; hence all frequency components of a general function $J(x,y,z;t)$ are shifted equally on the time scale and A will depend on $J[x,y,z;t - (r/v)]$. Thus we have

$$A(x,y,z;t) = \int\int\int \frac{J\left(x,y,z;t - \frac{r}{v}\right)}{4\pi r}\, dv. \tag{5–2}$$

A similar equation is obtained for the scalar potential. Then the field is obtained from

$$E = -\mu\frac{\partial A}{\partial t} - \text{grad}\, V, \quad H = \text{curl}\, A. \tag{5–3}$$

Representing J by a contour integral of the form (2.9–10), the proof can be made more formal. In the dissipative case no simple formula analogous to (2) exists.

Let us now consider an electric current filament of length l along the z-axis at the origin and suppose that the current starts from zero at $t = 0$ and is an arbitrary continuous function of time thereafter; thus

$$I(t) = 0, \quad t \leq 0; \quad \frac{dI}{dt} \text{ is finite.} \tag{5–4}$$

The charge $q(t)$ at the upper end is zero when $t < 0$ and

$$q(t) = \int_0^t I(t)\, dt \quad \text{when } t > 0. \tag{5–5}$$

At the lower end the charge is $-q(t)$.

Computing the field we find that it is composed of three parts. One of these parts (E',H') depends only on the time derivative of the current; another (E'',H'') depends on the current alone; the remainder E''' depends on the charges. Thus we write

$$E = E' + E'' + E''', \; H = H' + H'' + H''', \; H''' = 0,$$

$$E'_\theta = \eta H'_\varphi, \qquad E'_r = 0, \qquad H'_\varphi = \frac{lI'\left(t - \frac{r}{v}\right)}{4\pi v r}\sin\theta;$$

$$E''_\theta = \eta H''_\varphi, \qquad E''_r = 2E''_\theta\cot\theta, \qquad H''_\varphi = \frac{lI\left(t - \frac{r}{v}\right)}{4\pi r^2}\sin\theta; \tag{5–6}$$

$$E'''_\theta = \frac{lq\left(t - \frac{r}{v}\right)}{4\pi\epsilon r^3}\sin\theta, \qquad E'''_r = \frac{lq\left(t - \frac{r}{v}\right)}{2\pi\epsilon r^3}\cos\theta.$$

To an observer moving radially with velocity v the first part (E',H') of the total field would appear varying inversely as the distance from the element, the second part (E'',H'') inversely as the square of the distance, and the third part inversely as the cube of the distance. At sufficiently great distances only (E',H') will be sensibly different from zero although this particular fraction of the field is very small unless the electric current is changing very rapidly. The entire field is zero outside the spherical surface of radius vt with its center at the element. This sphere is the wavefront of the wave emitted by the element and on it (E'',H'') and (E''',H''') vanish. At the wavefront E and H are perpendicular to the radius.

6.6. *Potential Distribution on Perfectly Conducting Straight Wires*

Let the current $I(z)$ on a perfectly conducting straight wire of radius "a" (Fig. 6.8) be longitudinal and be distributed uniformly around the wire. This is substantially the case under any conditions if the wire is thin; if the "wire" is a cylindrical shell of large radius, circulating currents

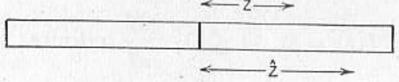


FIG. 6.8. A cylindrical wire.

will exist on it unless the electric intensity is impressed uniformly around the shell. Under the postulated conditions the vector potential is parallel to the axis of the wire. Let its value on the surface of the wire be $\Pi(z)$; then the corresponding value of the scalar electric potential V is

$$V = -\frac{1}{i\omega\epsilon}\frac{\partial \Pi}{\partial z}. \tag{6–1}$$

Since the electric intensity tangential to the surface of the wire is zero except in the region of impressed forces, we have

$$E_z(a) = -i\omega\mu\Pi - \frac{\partial V}{\partial z} = 0. \tag{6–2}$$

Thus we have obtained two equations connecting the values of V and Π on the surface of the wire

$$\frac{dV}{dz} = -i\omega\mu\Pi, \quad \frac{d\Pi}{dz} = -i\omega\epsilon V. \tag{6–3}$$

Eliminating either Π or V, we find

$$\frac{d^2V}{dz^2} = -\beta^2 V, \quad \frac{d^2\Pi}{dz^2} = -\beta^2\Pi; \tag{6–4}$$

hence V and Π are sinusoidal functions of the distance along the wire and the velocity of propagation is equal to the characteristic velocity of the surrounding medium. The equations, however, do not show where the nodes and antinodes of V and Π are located with respect to the ends of the wire. Since the radial component E_ρ is determined solely by the gradient of V, V can be defined as the electromotive force acting along a radius from the surface of the wire to infinity; but V is not a quantity which can readily be measured.

Consider now two parallel wires (Fig. 6.9) energized in "push-pull." Of the total impressed electromotive force V^i, one half is in series with the lower wire and the other half is in series with the upper wire and acts in the opposite direction. Under these conditions the currents in the two

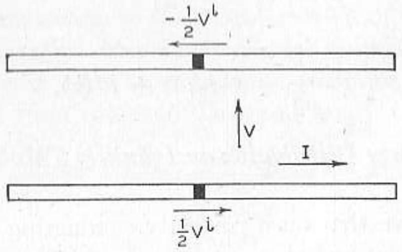


FIG. 6.9. Two parallel wires energized in push-pull.

wires are equal and opposite. Let V_1, Π_1 be the values of the potential functions on the surface of the lower wire and V_2, Π_2 the corresponding values on the upper wire; then we have

$$\frac{dV_1}{dz} = -i\omega\mu\Pi_1, \quad \frac{d\Pi_1}{dz} = -i\omega\epsilon V_1;$$
$$\frac{dV_2}{dz} = -i\omega\mu\Pi_2, \quad \frac{d\Pi_2}{dz} = -i\omega\epsilon V_2 \tag{6–5}$$

everywhere on the wires except where the impressed forces are acting. Subtracting we obtain

$$\frac{dV}{dz} = -i\omega\mu\Pi, \quad \frac{d\Pi}{dz} = -i\omega\epsilon V, \tag{6–6}$$

where

$$V = V_1 - V_2 = 2V_1, \quad \Pi = \Pi_1 - \Pi_2 = 2\Pi_1. \tag{6–7}$$

The potential difference V is now the transverse electromotive force acting from the lower wire to the upper along any path between the wires, lying completely in the plane normal to them. When the distance between the wires is small, V is a measurable quantity.

If the impressed forces acting on the wires are equal and in phase, the currents will also be equal and, then, V and Π in (4) refer to either wire.

If the generator is in series with one wire, we can replace it by two pairs of generators, one pair acting in push-pull and the other in phase

$$\tfrac{1}{2}V^i, \quad -\tfrac{1}{2}V^i;$$
$$\tfrac{1}{2}V^i, \quad \tfrac{1}{2}V^i. \tag{6–8}$$

Equations (3) apply only to those parts of the wire which are free from impressed forces. If $E^i(z)$ is the impressed intensity, then

$$E^i(z) = -E_z(a)$$

and

$$\frac{dV}{dz} = -i\omega\mu\Pi + E^i(z). \qquad (6\text{–}9)$$

6.7. *Current and Charge Distribution on Infinitely Thin Perfectly Conducting Wires*

We shall now prove that on a perfectly conducting wire of vanishingly small radius the current and charge are sinusoidal functions of the distance along the wire except in the immediate vicinity of the nodal points of the current and charge and in the vicinity of sudden bends. If $I(z)$ and $q(z)$ are the current and charge in the wire per unit length, then

$$\Pi(z) = \int \frac{I(\hat{z})e^{-i\beta r}}{4\pi r}\,d\hat{z}, \quad V(z) = \int \frac{q(\hat{z})e^{-i\beta r}}{4\pi\epsilon r}\,d\hat{z}, \qquad (7\text{–}1)$$

where the integration is extended over the length of the wire,

$$r = \sqrt{a^2 + (\hat{z} - z)^2}, \qquad (7\text{–}2)$$

and a is the radius of the wire. As a approaches zero the major contribution to Π and V is made by the current and charge in the vicinity of the point $\hat{z} = z$. Thus we shall have approximately

$$\Pi(z) = I(z)\int_{z-l}^{z+l} \frac{d\hat{z}}{4\pi r} = k\,I(z), \quad V(z) = q(z)\int_{z-l}^{z+l} \frac{d\hat{z}}{4\pi\epsilon r} = \frac{k}{\epsilon}q(z), \quad (7\text{–}3)$$

where l is small and

$$k = \int_{z-l}^{z+l} \frac{d\hat{z}}{4\pi r}. \qquad (7\text{–}4)$$

Calculating k we have

$$k = \frac{1}{4\pi}\int_{z-l}^{z+l} \frac{d\hat{z}}{\sqrt{a^2 + (\hat{z} - z)^2}} = \frac{1}{4\pi}\int_{-l}^{l} \frac{dx}{\sqrt{a^2 + x^2}}$$
$$= \frac{1}{4\pi}\log\left(x + \sqrt{a^2 + x^2}\right)\Big|_{-l}^{l} = \frac{1}{4\pi}\log\frac{\sqrt{a^2 + l^2} + l}{\sqrt{a^2 + l^2} - l}. \qquad (7\text{–}5)$$

Assuming that a is small compared with l, we obtain

$$\sqrt{l^2 + a^2} \simeq l\left(1 + \frac{a^2}{2l^2}\right) = l + \frac{a^2}{2l}, \quad k = \frac{1}{2\pi}\log\frac{2l}{a}. \qquad (7\text{–}6)$$

The quantity k increases indefinitely as the radius of the wire approaches zero and l is kept constant. The contributions to Π and V from the rest of the wire remain finite; hence equations (3) represent first approximations to Π and V everywhere except in the neighborhoods of the nodes of I and q, where the principal terms become small and contributions from more distant parts of the wire must be included in these first approximations.

Substituting for Π from (3) in (6–3), we have

$$\frac{dV}{dz} = -i\omega\mu k I, \quad \frac{dI}{dz} = -\frac{i\omega\epsilon}{k}V. \qquad (7\text{–}7)$$

Substituting for V from (3), we have also

$$\frac{dq}{dz} = -i\omega\mu\epsilon I, \quad \frac{dI}{dz} = -i\omega q. \qquad (7\text{–}8)$$

The second equation in this set is really exact; it may be obtained directly from the principle of conservation of charge.

Thus we have proved the theorem stated at the beginning of this section for the case of straight wires. If the wire is curved (Fig. 6.10), our arguments are still valid except in the immediate vicinity of "angular points" where the wire suddenly changes its direction. In this case the exact equation connecting the scalar potential with the vector potential becomes

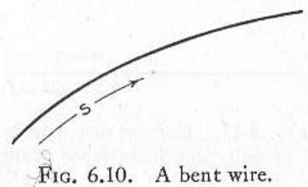


FIG. 6.10. A bent wire.

$$\frac{\partial V}{\partial s} = -i\omega\mu A_s.$$

The same approximations can be made as for straight wires except in the vicinity of angular points. Our conclusions are still valid if the radius of the wire is varying so long as the rate of change is finite.

Let us now return to the case of two parallel wires energized in push-pull (Fig. 6.9). Here Π_1, as defined in the preceding section, consists of two parts

$$\Pi_1 = \Pi_1' + \Pi_1'',$$

each due to the current in one of the wires. If the distance d between the axes of the wires is small, the approximate expressions (3) are particularly good since the contributions from distant portions of one wire are nearly canceled by contributions from similar portions of the other wire. For k we have

$$k = \frac{1}{4\pi}\int_{-l}^{l}\left[\frac{1}{\sqrt{a^2 + x^2}} - \frac{1}{\sqrt{d^2 + x^2}}\right]dx = \frac{1}{2\pi}\log\frac{d}{a},$$

as long as d and a are both small compared with l. Thus k has become independent of the indefinite length l. By (6–7) and (3) we now have

$$\frac{dV}{dz} = -i\omega LI, \quad \frac{dI}{dz} = -i\omega CV, \tag{7–9}$$

where

$$L = \frac{\mu}{\pi}\log\frac{d}{a}, \quad C = \frac{\pi\epsilon}{\log\frac{d}{a}}. \tag{7–10}$$

6.8. *Radiation from a Wire Energized at the Center*

We are now in a position to calculate the power radiated by an infinitely thin perfectly conducting wire energized at the center.* We have proved that the current distribution is sinusoidal; the ends of an infinitely thin wire must be current nodes; and the current $I(z)$ must be an even function of the distance z from the center (Fig. 6.11). Therefore

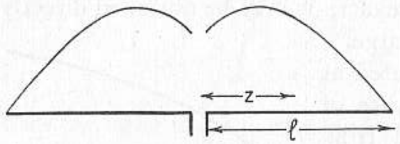


FIG. 6.11. Current distribution on a wire of finite length energized at the center.

$$I(z) = I\sin\beta(l - z), \quad z > 0;$$
$$= I\sin\beta(l + z), \quad z < 0; \tag{8–1}$$

where I is the maximum amplitude of the current.

If the length $2l$ of the wire is equal to a half wavelength, (1) becomes

$$I(z) = I\cos\beta z, \tag{8–2}$$

where the maximum amplitude is now at the center. The radiated power can be calculated by (4–22). Using only the first three terms of the power series for k_{12}, we obtain (for free space)

$$W = \tfrac{1}{2}RI^2, \quad R = 73.2 \text{ ohms.} \tag{8–3}$$

More accurate calculation gives $R = 73.129$. In Chapter 11 we shall prove that the exact value of the input resistance depends on the radius of the wire, particularly for lengths greater than a half wavelength; there we shall obtain expressions for R as well as for the reactive component of the input impedance as functions of the radius of the antenna.

6.9. *The Mutual Impedance between Two Current Loops; the Impedance of a Loop*

Let us now consider two current loops carrying uniform currents I_1 and I_2 in phase with each other (Fig. 6.12). The component of the vector

* Or at any other point for that matter. The general formulae will be obtained in Chapter 9.

potential due to I_2, along the element ds_1, is

$$A_{s,1} = \frac{I_2}{4\pi}\int \frac{e^{-i\beta r_{12}}\cos\psi}{r_{12}}\,ds_2, \tag{9–1}$$

where ψ is the angle between the elements ds_1 and ds_2. The electromotive force round the first circuit due to the field of the second is

$$\int E_{s,1}ds_1 = -i\omega\mu\int A_{s,1}ds_1 = -\frac{i\omega\mu I_2}{4\pi}\int\int\frac{e^{-i\beta r_{12}}\cos\psi}{r_{12}}\,ds_1\,ds_2, \tag{9–2}$$

since the integral of $\partial V/\partial s$ round the circuit vanishes. The electromotive force which should be impressed on the first circuit in order to sustain I_1

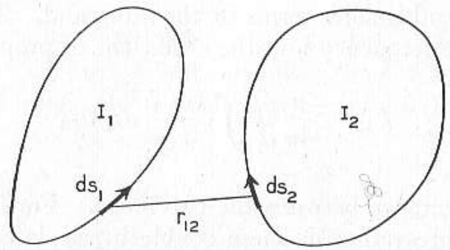


FIG. 6.12. Two current loops.

against the electromotive force induced by I_2 is the negative of (2). The ratio of this impressed electromotive force to I_2 is the mutual impedance between the two loops

$$Z_{12} = \frac{i\omega\mu}{4\pi}\int\int\frac{e^{-i\beta r_{12}}\cos\psi}{r_{12}}\,ds_1\,ds_2. \tag{9–3}$$

The real and imaginary components of this impedance are

$$R_{12} = \frac{\omega\mu}{4\pi}\int\int\frac{\sin\beta r_{12}}{r_{12}}\cos\psi\,ds_1\,ds_2,$$
$$X_{12} = \frac{\omega\mu}{4\pi}\int\int\frac{\cos\beta r_{12}}{r_{12}}\cos\psi\,ds_1\,ds_2. \tag{9–4}$$

R_{12} represents the mutual radiation resistance.

The above expressions are exact if I_1 and I_2 are uniform as we have assumed; but uniform currents can be sustained only by properly *distributed* impressed forces. The usual method of energizing a loop is to impress an electromotive force between a pair of terminals (Fig. 5.1.) In section 7 it has been shown that the current distribution on an infinitely thin wire is represented by a sinusoidal function of the distance along the loop. Furthermore the current entering the loop from the generator at A equals the current leaving the loop at B; hence the current is an even function of the distance s from the midpoint C of the loop. The even sinusoidal function of s is $\cos \beta s$, and this is nearly constant for small values of s. Thus the above equations should apply approximately to small loops energized by concentrated impressed forces.

Furthermore for loops which are not too far apart we have approximately

$$X_{12} = \frac{\omega\mu}{4\pi} \int\int \frac{\cos\psi}{r_{12}}\, ds_1\, ds_2, \qquad (9\text{–}5)$$

neglecting $\frac{1}{2}\beta^2 r_{12}$ and smaller terms in the integrand. This is seen to be proportional to the frequency and the coefficient of proportionality

$$L_{12} = \frac{\mu}{4\pi} \int\int \frac{\cos\psi}{r_{12}}\, ds_1\, ds_2 \qquad (9\text{–}6)$$

is the mutual inductance between the two loops. For loops of small but finite radius the integration in these double integrals is performed along the axes of the wires, although for more accurate computation the wires must be divided into elementary filaments.

If the loops are coincident the mutual impedance becomes the self impedance of the loop. Except at rather low frequencies the current is distributed near the surface of the wire. The vector potential of such a current distribution, at points external to the wire, can be computed by assuming that the current is distributed along the axis; but the second integration should be performed where the current actually happens to be and the corresponding curve of integration in the above double integrals must be taken on the surface of the wire. In particular these remarks should be kept in mind in computing the self-reactance or self-inductance of a loop; when calculating its radiation resistance, no great error is made if both integrations are taken along the axis of the loop. The reason for this is that the error involved in shifting the second path of integration from the surface to the axis is greatest for small values of r_{12}, and for these values the integrand in R_{12} is nearly independent of r_{12}. On the other hand the greatest contribution to X_{12} comes from small values of r_{12}.

Even if the current is distributed throughout the cross-section of the wire, the "external" inductance of the loop is obtained from (6) by integrating once along the axis of the wire and once along a parallel curve on its surface. The "internal" inductance is computed separately.

The mutual impedance between two closed uniform current filaments can also be expressed as follows:

$$Z_{12} = \frac{i\omega\Phi_{12}}{I_2}, \qquad (9\text{–}7)$$

where Φ_{12} is the magnetic flux through the first loop due to the current in the second. Equation (6) shows that $\Phi_{21} = \Phi_{12}$. The radiation resistance appears through the component of Φ_{12} in quadrature with I_2.

For the power radiated by two loops carrying currents differing in phase by ϑ, the mutual power term contains the factor $\cos\vartheta$; this factor appears in the expression for Ψ for any transducer as shown by equation (5.2–8). If two current elements or two filaments are in quadrature they radiate independently of each other.

6.10. *Radiation from a Small Plane Loop Carrying Uniform Current*

If the loop is small, we may obtain an approximate value for the radiation resistance by retaining only the first two terms of the power series for $\sin \beta r_{12}$ in (9–4); thus we have

$$\begin{aligned} R &= \frac{\omega\mu}{4\pi} \int\int (\beta - \tfrac{1}{6}\beta^3 r_{12}^2) \cos\psi\, ds_1\, ds_2 \\ &= \frac{\eta\beta^2}{4\pi} \int\int \cos\psi\, ds_1\, ds_2 - \frac{\eta\beta^4}{24\pi} \int\int r_{12}^2 \cos\psi\, ds_1\, ds_2. \end{aligned} \qquad (10\text{–}1)$$

Let ϑ_1 and ϑ_2 be the inclinations of the elements ds_1 and ds_2 (Fig. 6.13); then

$$\cos\vartheta_1 = \frac{dx_1}{ds_1}, \quad \sin\vartheta_1 = \frac{dy_1}{ds_1},$$

$$\cos\vartheta_2 = \frac{dx_2}{ds_2}, \quad \sin\vartheta_2 = \frac{dy_2}{ds_2},$$

$$\cos\psi = \cos(\vartheta_1 - \vartheta_2) = \frac{dx_1\, dx_2 + dy_1\, dy_2}{ds_1\, ds_2}.$$

Remembering that dx_1 and dy_1 are independent of dx_2 and dy_2, we obtain

$$\begin{aligned} \int\int \cos\psi\, ds_1\, ds_2 &= \int\int (dx_1\, dx_2 + dy_1\, dy_2) \\ &= \int dx_1 \int dx_2 + \int dy_1 \int dy_2 = 0, \end{aligned}$$

since the total change of either the abscissa or the ordinate around a closed curve is zero.

Similarly we have

$$\iint r_{12}^2 \cos\psi \, ds_1 \, ds_2 = \iint [(x_2 - x_1)^2 + (y_2 - y_1)^2](dx_1 \, dx_2 + dy_1 \, dy_2)$$
$$= -2\left(\int x \, dy\right)^2 - 2\left(\int y \, dx\right)^2 = -4S^2,$$

where S is the area bounded by the loop. All other terms in this integral vanish because the total variation of either a coordinate or a function of the coordinate vanishes for the complete cycle.

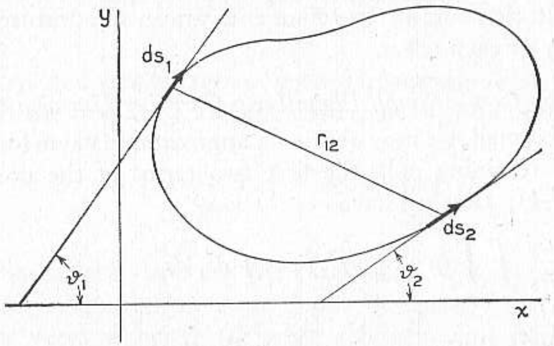


FIG. 6.13. A loop carrying uniform current.

Substituting these results in (1), we have

$$R = \frac{\eta}{6\pi}(\beta^2 S)^2 = \tfrac{8}{3}\pi^3\eta\frac{S^2}{\lambda^4}. \tag{10–2}$$

In free space this becomes

$$R = 20(\beta^2 S)^2 = 320\pi^4\frac{S^2}{\lambda^4} = 31{,}000\frac{S^2}{\lambda^4}. \tag{10–3}$$

6.11. *Transmission Lines and Wave Guides*

We have seen that even in nondissipative homogeneous media the amplitude of a wave decreases with the distance from the source. On the other hand, the equations of section 7 show that the current in perfectly conducting wires is approximately sinusoidal and that consequently the amplitude of the waves is independent of the distance along the wire or wires. The wires act as *transmission lines* or *wave guides*. We shall use the term " transmission line " in a restricted sense for wave guides whose transverse dimensions are small.

Of course, if the wires and the medium are dissipative, the amplitude of guided waves will also decrease with the distance because of power absorption. More general transmission equations for a pair of parallel wires (Fig. 6.14) may be derived directly from the fundamental electromagnetic equations. Applying Faraday's law to the rectangle $ABCDA$ in which $AB = 1$, we have

$$V_{AB} + V_{BC} + V_{CD} + V_{DA} = -i\omega\Phi. \tag{11–1}$$

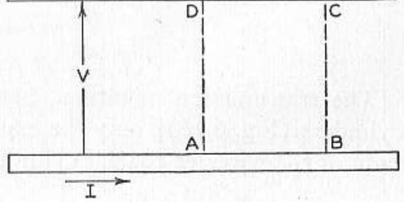


FIG. 6.14. Two parallel wires.

If the internal or surface impedances of the wires per unit length are Z_1 and Z_2, then for the particular transmission mode in which the currents in the wires are equal and oppositely directed we have

$$V_{AB} = Z_1 I, \quad V_{CD} = Z_2 I. \tag{11–2}$$

If V is the transverse voltage between the wires, then

$$V_{BC} + V_{DA} = V_{BC} - V_{AD} = \frac{dV}{dz}. \tag{11–3}$$

Finally, Φ is proportional to I; consequently equation (1) becomes

$$\frac{dV}{dz} = -(Z_1 + Z_2 + i\omega L)I. \tag{11–4}$$

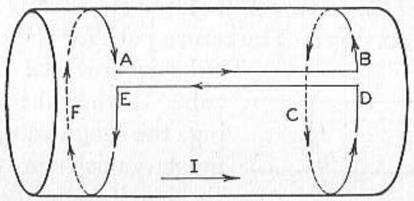


FIG. 6.15. A closed path on the surface of a wire.

Applying Ampère's law to the circuit $ABCDEFA$ on the surface of the lower wire (Fig. 6.15) we have

$$U_{AB} + U_{BCD} + U_{DE} + U_{EFA} = I_t,$$

or

$$U_{BCD} + U_{EFA} = I_t.$$

The left side is the difference between the currents flowing in the wire at A and B; if $AB = 1$, then

$$U_{BCD} + U_{EFA} = -\frac{dI}{dz} = I_t. \tag{11–5}$$

On the other hand the transverse current is proportional to the voltage

$$I_t = (G + i\omega C)V, \tag{11-6}$$

and consequently

$$\frac{dI}{dz} = -(G + i\omega C)V. \tag{11-7}$$

The transmission equations (4) and (7) apply equally well to coaxial cylinders (Fig. 6.16); only the expressions for L, G, C are different. If the radii of the wires or coaxial cylinders vary slowly with the distance z along

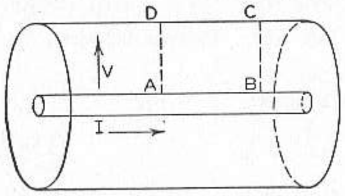


Fig. 6.16. Two coaxial cylinders.

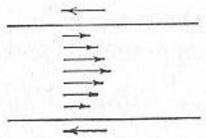


Fig. 6.17. Illustrating a possible distribution of the longitudinal displacement current inside a metal tube.

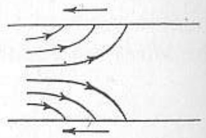


Fig. 6.18. Electric lines inside a metal tube.

the wires, then L, G, C are functions of z. Generally these transmission equations are approximate; but in Chapter 8 we shall find that under certain conditions they may be exact.

Let us now remove the inner conductor of the coaxial pair and see if wave transmission is still possible. The return path for the current in the metal tube is now the dielectric inside the tube. If the tube is perfectly conducting, the longitudinal electric intensity must vanish on the boundary and the longitudinal displacement current might be distributed as shown in Fig. 6.17. The lines of displacement current flow would then look like those in Fig. 6.18. If the field is symmetric, magnetic lines are circles coaxial with the tube. Let V be the transverse voltage from the axis of the tube to its periphery and I the total longitudinal displacement current (Fig. 6.19). Applying Faraday's law to a rectangle $ABCDA$ we obtain equation (1). The voltage V_{CD} is given by (2); similarly, we have equation (3); but V_{AB} is now equal to the longitudinal electric intensity on the axis

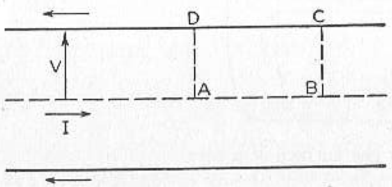


Fig. 6.19. A metal tube and a rectangular path formed by two radii and the lines joining their ends.

$$V_{AB} = E_0. \tag{11-8}$$

In the case of a coaxial pair this voltage is small, equal to zero, in fact, if the inner conductor is perfect; but with no inner conductor there is every reason to suppose that it will prove to be significant.

Assuming that the longitudinal electric intensity is maximum on the axis, we have

$$E_z = E_0 f(\rho), \quad f(0) = 1. \tag{11-9}$$

Since the total current is

$$I = i\omega\epsilon E_0 \int\int f(\rho)\rho \, d\rho \, d\varphi,$$

we have

$$E_0 = \frac{I}{i\omega \hat{C}}, \quad \hat{C} = \epsilon\left(\frac{1}{S}\int\int f(\rho)\rho \, d\rho \, d\varphi\right) S, \tag{11-10}$$

where S is the area of the cross-section of the tube. The quantity in parentheses is the average value of $f(\rho)$ over the cross-section of the tube. Φ is proportional to I but the coefficient of proportionality L is naturally different from that for coaxial pairs. Equation (1) now assumes the following form

$$\frac{dV}{dz} = -\left(Z_2 + i\omega L + \frac{1}{i\omega \hat{C}}\right) I. \tag{11-11}$$

This equation differs from the equations for coaxial pairs and parallel pairs in that it contains a term representing distributed series capacity $\hat{C}$. This, of course, was to be expected.

The second transmission equation is simply the equation of conservation of electric charge

$$\frac{dI}{dz} = i\omega q,$$

where q is the charge on the surface of the tube per unit length. This charge is proportional to the radial component E_ρ of the electric intensity and therefore to the transverse voltage V; taking into consideration our convention with regard to the positive direction of V, we have

$$\frac{dI}{dz} = -i\omega CV. \tag{11-12}$$

Comparing (11) and (12) with (5.7-6) and (5.7-7) we find that the metal tube behaves as a high pass filter. The cutoff frequency is determined by

$$\omega_c = \frac{1}{\sqrt{L\hat{C}}}. \tag{11-13}$$

It is easy to make a rough estimate of this frequency. From the physical picture underlying the present transmission mode it is clear that the error will not be excessive if we assume a uniformly distributed longitudinal displacement current. Then $f(\rho) = 1$ and $\hat{C} = \epsilon S = \epsilon\pi a^2$. In this case the inductance per unit length is $L = \mu/4\pi$. Substituting in (13), we have

$$\omega_c = \frac{2}{a\sqrt{\mu\epsilon}}, \quad \frac{2\pi a}{\lambda_c} = 2. \tag{11–14}$$

Thus the cutoff wavelength is equal roughly to the circumference of the wave guide divided by 2.* Longer waves are not transmitted. The exact cutoff is determined if we use 2.40 instead of 2. More powerful methods for obtaining the cutoff frequencies will be described in later chapters. The above estimate has been made in order to show that, starting from a physical picture of a given field and applying the electromagnetic laws in their integral form, it is possible to obtain qualitative and even fairly satisfactory quantitative results.

(a) (b)

FIG. 6.20. Cross-sections of metal tubes of rectangular and semi-circular cross-section and magnetic lines.

In tubes of noncircular cross-section magnetic lines will be deformed (Fig. 6.20), the numerical values of L, C, $\hat{C}$ will be altered, but the essential picture will remain the same. We can even make an estimate of the cutoff wavelength by expressing λ_c for the circular guide in terms of the area of the cross-section instead of the radius.

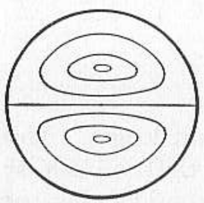


FIG. 6.21. Illustrating a transmission mode with two sets of closed magnetic lines.

The direction of the conduction current in the tube is related to the direction of the magnetic lines of force. If the magnetic lines are counterclockwise, the current in the tube flows away from the observer. Consider now a circular tube with an infinitely thin perfectly conducting axial partition (Fig. 6.21) and assume that waves of equal intensity of the type shown in Fig. 6.20 have been set up in such a way that at all times one set of magnetic lines is counterclockwise and the other clockwise. The total current in the axial partition is zero and the partition has no effect

* On the energy basis $L = \mu/8\pi$ and $\lambda_c = 2\pi a/2.8$. Much better results are obtained by taking $f(\rho) = 1-\rho^2/a^2$ or $\cos(\pi\rho/2a)$ so that $f(a) = 0$.

on the field inside the circular tube. This partition can, therefore, be removed and we are left with a new mode of transmission in a circular tube. In this mode the conduction current flows in opposite directions in opposite halves of the tube; longitudinal displacement currents also flow in opposite directions. The cutoff frequency for this mode is higher than that for the first mode; the ratio of these frequencies is equal to the ratio of the cutoff frequencies for the first mode in the circular tube and in a semicircular tube of half the area. Hence, the approximate ratio of the two cutoff frequencies is $\sqrt{2}$ or 1.4; the exact ratio is nearly 1.6. The magnetic lines " avoid " the corners in the semicircular tube and this tendency makes the effective area of the tube smaller than the actual area.

This synthetic method of construction of field configurations can be extended. The circular tube can be divided by radial planes into an even number of sectors. Assuming a wave in each sector, traveling in the first

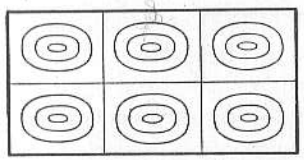


FIG. 6.22. A transmission mode with six sets of closed magnetic lines.

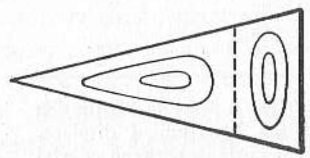


FIG. 6.23. A possible mode of transmission in a tube of triangular cross-section.

mode, and assuming relative directions of magnetic lines so as to make radial planes current free and hence removable, we obtain a sectorial wave in the circular tube. Any rectangular tube can be divided into equal rectangular tubes of smaller cross-section (Fig. 6.22); assuming the first mode in each in such a way that the adjacent lines of magnetic force point in the same direction, we obtain a higher transmission mode in the original tube. All these field configurations can be constructed on the basis of symmetry. They furnish us with qualitative ideas when symmetry is no longer a guide. We feel certain, for example, that in the tube whose cross-section is shown in Fig. 6.23, there exists a mode with two sets of magnetic lines of force as shown in the figure; but without more complete analysis we do not know just how the available space is divided between these sets of lines. All we can say is that the area of the left sector will be larger than that of the right sector because the magnetic lines avoid corners, particularly sharp corners. The magnetic lines surround the longitudinal displacement current which is proportional to the longitudinal electric intensity; but the latter must vanish on the boundaries of the tube and it will approach zero more rapidly when two boundaries are close together. In coaxial pairs similar waves can exist; thus in the circularly symmetric case the longitudinal displacement current may be distributed as shown in Fig. 6.24.

All waves of the above type are called *transverse magnetic waves* or *TM-waves* because the magnetic vector is perpendicular to the direction of wave propagation. If the electric vector is perpendicular to the direction of wave propagation, then the waves are called *transverse electric waves* or *TE-waves*. Finally, if both vectors are perpendicular to the direction of wave

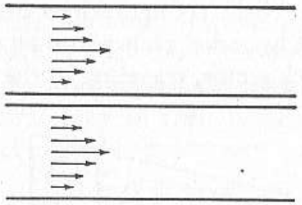


FIG. 6.24. A possible distribution of the longitudinal displacement current between coaxial cylinders.

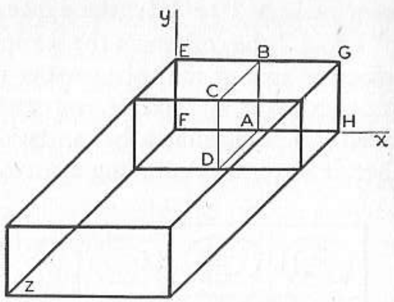


FIG. 6.25. A rectangular metal tube.

propagation, then the waves are *transverse electromagnetic* (TEM-waves). In general both field intensities have longitudinal and transverse components; such waves are called *hybrid waves*. There are no electromagnetic waves in which either the electric intensity or the magnetic intensity is totally longitudinal.

A general idea of transverse electric waves may be obtained as follows. Consider two parallel metal strips, whose width is large compared with the distance between them. Such strips form a transmission line similar to a pair of parallel wires. Between the strips the electric field is almost uniform, except near the edges; the magnetic lines surround each strip and between the strips they are nearly parallel to them. The wave is transverse electromagnetic. The longitudinal currents in the strips flow in opposite directions and the circuit is made complete with the aid of transverse displacement currents. Let us now connect the edges of the strips metallically and form a rectangular tube (Fig. 6.25). The electric intensity, which we assume to be parallel to the y-axis, must vanish at the boundaries to which it is parallel. Let us assume that E is maximum in the middle plane $ABCD$ and that it is distributed as shown in Fig. 6.26. Magnetic lines cannot cross the conducting boundaries and must form loops (Fig. 6.27) surrounding the transverse displacement current. The longitudinal magnetic intensity is associated with the transverse conduction current in the tube.

Now let I be the total longitudinal current in the lower face of the tube and $-I$ the corresponding current in the upper face. Let V be the voltage from the lower face to the upper along a typical " central " line AB. This voltage is equal to the total longitudinal magnetic current flowing through

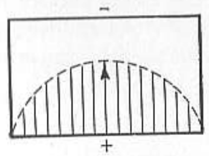


FIG. 6.26. A possible distribution of the transverse displacement current in a rectangular tube.

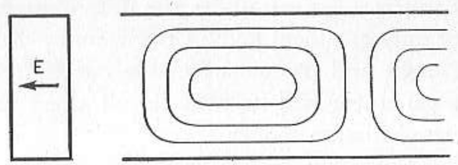


FIG. 6.27. Magnetic lines in planes normal to the E-vector.

the rectangle $ABEF$ in the negative z-direction, or the total current through $ABGH$ in the positive direction. Assuming that the tube is perfectly conducting and applying the first law of induction to the rectangle $ABCD$, we obtain the first transmission equation. Expressing the variation of I with z in terms of the shunt displacement and conduction currents we obtain the second transmission equation. Thus we have

$$\frac{dV}{dz} = -i\omega LI, \quad \frac{dI}{dz} = -\left(i\omega C + \frac{1}{i\omega L}\right)V. \qquad (11\text{–}15)$$

Thus in the case of transverse electric waves the tube also behaves as a high pass filter (Fig. 5.19). The constants can be calculated if more specific assumptions are made with regard to the distribution of the transverse displacement current. However, in Chapter 8 we shall obtain the complete and exact solution of this problem. The purpose of the present discussion is to stimulate the development of physical ideas as we proceed with mathematical analysis.

In Chapter 10 we shall solve rigorously the problem of cylindrical wave guides and confirm the existence of an infinite number of *transmission modes*. Each mode is characterized by a definite field pattern in a typical plane normal to the guide. This field pattern determines completely the constants in the transmission equations (11) and (12) or (15), depending upon whether the wave is of transverse magnetic or transverse electric type. The propagation constant and the velocity in the guide depend upon the frequency and the particular transmission mode. The cutoff frequencies for various transmission modes may be arranged in ascending order of magnitude. The lowest of these frequencies is called the *absolute cutoff frequency* for the guide and the corresponding mode is the principal or the *dominant* transmission mode. If the guide is energized at some frequency lower than the absolute cutoff frequency, the propagation constants for all modes are real (if there is no dissipation) and the field intensity approaches zero with increasing distance from the generator. If, however, the frequency is above the absolute cutoff but below the next higher, then at a sufficient distance from the generator the wave will be traveling along the guide substantially in the dominant mode. All the other modes represent only the local field in the vicinity of the generator. As the frequency increases and passes successive cutoff frequencies, the energy supplied by the generator will be transferred along the guide in an increasing number of transmission modes.

Transmission modes are analogous to oscillation modes in electric networks. A simple series circuit has only one natural frequency and one oscillation mode; an n-mesh network has n oscillation modes; and a section of a transmission line has an infinite number of oscillation modes. Actual physical circuits are always multiple circuits, possessing in fact an infinite number of oscillation modes. The lowest natural frequency of some circuits, however, is so much lower than all the others that in a limited frequency range they may be approximated by *simple* circuits possessing only one natural frequency. Similarly all physical wave guides admit of an infinite number of transmission modes; but some wave guides, such as coaxial pairs, admit of one mode for which the cutoff frequency is zero and of other modes with very high cutoff frequencies. In a restricted frequency range such wave guides may be treated as *simple transmission lines*, possessing only one transmission mode. The absolute cutoff of metal tubes is high and the cutoff frequencies of higher modes are close to it (on the ratio basis); in such cases the existence of other transmission modes cannot be forgotten even if the operating frequency is such that only the dominant mode takes part in energy transmission. However, at frequencies between the first and second cutoffs, the higher transmission modes represent only local fields in the vicinity of discontinuities such as generators, receivers, sudden bends or changes in the transverse dimensions of the guide. Under these conditions the wave guide acts as a simple transmission line in which the local fields associated with the discontinuities are represented by reactors either in series or in shunt with the line.

6.12. *Reflection*

In sections 1 and 2 we have calculated the field produced by a given distribution of sources in an infinite homogeneous medium. Let us now suppose that the medium consists of two homogeneous regions separated by a surface (S). Without loss of generality we may assume that one of these regions is source-free. If the sources are distributed throughout both

regions, we may regard the total field as due to the superposition of two fields, each produced by sources located in one region only.

Thus let the sources be in region (1) as shown in Fig. 6.28. If the two regions had the same electromagnetic properties, the field of these sources would be found from the equations of sections 1 and 2. But when the electromagnetic properties are different the field (E^i,H^i) thus obtained is not the actual field. In region (1) it represents the primary field of the sources and is called the *impressed field*. The field (E^r,H^r) which must be added to give the actual field in region (1) is called the *reflected field*. We may think of the reflected field as produced by polarization currents in region (2); in so far as these virtual* sources are concerned region (1) is source-free and the reflected field should satisfy the homogeneous form of Maxwell's equations

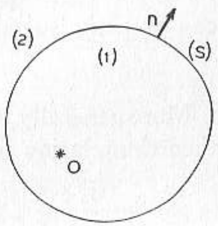


FIG. 6.28. A surface enclosing the source of the field.

$$\text{curl } E^r = -i\omega\mu_1 H^r, \quad \text{curl } H^r = (g_1 + i\omega\epsilon_1)E^r. \qquad (12\text{–}1)$$

Let the actual field in region (2) be (E^t,H^t); this field is called the *transmitted* (or "refracted") field and it also satisfies the homogeneous equations

$$\text{curl } E^t = -i\omega\mu_2 H^t, \quad \text{curl } H^t = (g_2 + i\omega\epsilon_2)E^t. \qquad (12\text{–}2)$$

At the interface (S) of the two media the tangential components of E and H are continuous

$$E_t^i + E_t^r = E_t^t, \quad H_t^i + H_t^r = H_t^t. \qquad (12\text{–}3)$$

This set of equations constitutes one formulation of the problem of determining the field of a given system of sources when the medium consists of two homogeneous regions. The method can be extended to any number of homogeneous regions.

If the boundary (S) is a perfectly conducting sheet, then the tangential component of E should vanish on (S)

$$E_t^i + E_t^r = 0, \quad \text{or} \quad E_t^r = -E_t^i. \qquad (12\text{–}4)$$

A perfectly conducting sheet can support finite electric current and the tangential component of H is no longer continuous across (S). In fact, energy cannot flow across a perfect conductor and the field in region (2) due to the sources in region (1) is equal to zero. The component of H tangential to (S) in region (1) represents the current density J on (S). By the second law of induction J is normal to the tangential component of H; hence if n is a unit normal to (S), regarded as positive when pointing

* As distinct from true sources.

into the source-free region (2), then

$$J = (H_t^i + H_t^r) \times n. \tag{12–5}$$

Since the vector product of n and the normal component of H is zero, we can drop the subscript " tangential " and write

$$J = (H^i + H^r) \times n. \tag{12–6}$$

More generally if the impedance Z_n normal to the boundary is prescribed, then the relation between the tangential components in region (1) is

$$E_t^t = Z_n H_t^t \times n, \quad \text{or} \quad E_t^i + E_t^r = Z_n(H_t^i + H_t^r) \times n. \tag{12–7}$$

Wave propagation in wave guides may be regarded as a case of reflection. We start with a certain system of sources inside a metal tube, for example; then the total field inside the tube is the sum of the impressed and reflected fields in the sense defined in this section.

6.13. *The Induction Theorem*

Let us rewrite (12–3) as follows

$$E_t^t - E_t^r = E_t^i, \quad H_t^t - H_t^r = H_t^i, \tag{13–1}$$

and concentrate our attention on the " induced " field $(\hat{E},\hat{H})$ consisting of the reflected field (E^r,H^r) in region (1) and the transmitted field (E^t,H^t) in region (2). This field satisfies the homogeneous equations (12–1) and (12–2) everywhere except on (S) and it may be obtained from a distribution of sources on (S) as well as from the original sources.

It has been shown in section 4.5 that the discontinuities in E and H across (S) could be produced by current sheets on (S) of densities

$$M = (E_t^t - E_t^r) \times n = E_t^i \times n, \qquad J = n \times (H_t^t - H_t^r) = n \times H_t^i. \tag{13–2}$$

Since the vector product of n and a normal component of the field is zero, we have

$$M = E^i \times n, \quad J = n \times H^i. \tag{13–3}$$

Thus if we wish to determine the field whose only sources are the currents on (S) given by (3), we have to solve exactly the same equations as those used in the preceding section to obtain the induced field. In other words the induced field $(\hat{E},\hat{H})$ could be produced by electric and magnetic current sheets of densities given by (3); this is the Induction Theorem.

6.14. *The Equivalence Theorem*

Let us now suppose that (S) is a surface in a *homogeneous* medium, separating a source-free region (2) from the rest. In this case the " reflected " field is evidently zero and the transmitted field is the actual field in the source-free region. Thus we obtain the following Equivalence Theorem: the field in a source-free region bounded by a surface (S) could be produced by a distribution of electric and magnetic currents on this surface and in this sense the actual source distribution can be replaced by an "equivalent" distribution (13–3).

6.15. *Stationary Fields*

Stationary fields are fields independent of time and may be regarded as special cases of variable fields. For example from (1–10) and (2–14) we obtain the following expression for the electrostatic field produced by a given distribution of electric charge in a perfect dielectric

$$E = -\text{grad } V, \quad V = \int \frac{dq}{4\pi\epsilon r}, \tag{15–1}$$

where dq is a typical element of charge. The function V is now called the electrostatic potential. A similar expression may be obtained for the magnetostatic field of a given distribution of magnetic charge

$$H = -\text{grad } U, \quad U = \int \frac{dm}{4\pi\mu r}, \tag{15–2}$$

where dm is a typical element of magnetic charge. In this case the function U is called the magnetostatic potential. These expressions are also the limits of the harmonic field when ω approaches zero.

In an infinite homogeneous conductor we have from (2–13) and (1–8)

$$A = \int \frac{dp}{4\pi r}, \quad V = -\frac{1}{g} \text{div } A, \tag{15–3}$$

where dp is the moment of a typical impressed current element. From (1–10) we have

$$E = -\text{grad } V, \quad H = \text{curl } A. \tag{15–4}$$

The field due to currents in conductors surrounded by a homogeneous dielectric medium can be obtained from

$$\text{curl } H = gE + J^i \tag{15–5}$$

if we use (1.8–6) and recall that in the absence of magnetic charges $\text{div } H = 0$. Thus we have

$$H = \text{curl} \iiint \frac{gE + J^i}{4\pi r} \, dv. \tag{15–6}$$

This formula can also be used, of course, for homogeneous media but it is more complicated than (3) and (4) which give H in terms of the impressed currents alone. Thus the latter equations give immediately the field of an impressed current element (equations 2–3, 2–5, 2–6) while this could only be obtained from (6) by integrating over the entire infinite medium.

6.16. *Conditions in the Vicinities of Simple and Double Layers of Charge*

Consider a stationary distribution of electric charge on some surface (S) or a simple layer (Fig. 6.29). The normal component of E is discontinuous across the layer; thus by (4.3–2) we have

$$E_{n,2} - E_{n,1} = \frac{q_S}{\epsilon}, \qquad (16\text{–}1)$$

where q_S is the charge on the layer per unit area and ϵ is the dielectric constant of the surrounding medium. The tangential component of E is continuous. In terms of the electrostatic potential V these boundary conditions become

$$\frac{\partial V_1}{\partial s} = \frac{\partial V_2}{\partial s}, \quad \frac{\partial V_1}{\partial n} - \frac{\partial V_2}{\partial n} = \frac{q_S}{\epsilon}. \qquad (16\text{–}2)$$

From the expression for the potential in terms of the charge distribution it is evident that the potential is continuous across the simple layer; this condition implies the continuity of the tangential components of V.

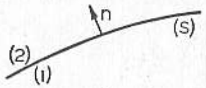


FIG. 6.29. A surface layer of charge.

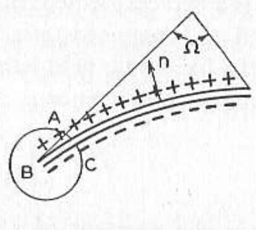


FIG. 6.30. A double layer of charge.

Two close layers of equal and opposite charges constitute a *double layer* (Fig. 6.30). Such a layer may be subdivided into elementary doublets. If q_S is the charge per unit area and l is the separation between the simple layers, then $dp = q_S l\, dS$ is the moment of an elementary doublet. The moment χ per unit area is called the *strength* of the layer. For an ideal double layer l is vanishingly small and q_S is infinitely large, while their product χ is finite. We have seen that the potential of a doublet is

$$dV = \frac{\chi \cos\psi}{4\pi\epsilon r^2}\, dS, \qquad (16\text{–}3)$$

where ψ is the angle made with the axis of the doublet by the line joining

the doublet with a typical point P (Fig. 6.31). The potential of the entire layer is therefore

$$V(P) = \frac{1}{4\pi\epsilon}\iint_{(S)} \frac{\chi\cos\psi}{r^2}\, dS. \qquad (16\text{–}4)$$

By the definition of the double layer and by (4.3–2) the normal component of the electric intensity is continuous; hence this is also true of the normal derivative of the potential of a double layer in a homogeneous medium. On the other hand, the potential itself is discontinuous. Inside the layer the electric intensity is $-q_S/\epsilon$ and the potential rise across the layer in the direction of the normal as indicated in Fig. 6.30 is

$$V_0 = \frac{q_S l}{\epsilon} = \frac{\chi}{\epsilon}. \qquad (16\text{–}5)$$

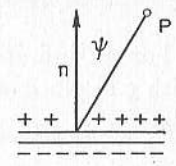


FIG. 6.31. An element of a double layer.

In terms of this potential discontinuity across the layer, the potential outside the layer given by (4) becomes

$$V(P) = \frac{1}{4\pi}\iint_{(S)} \frac{V_0\cos\psi}{r^2}\, dS. \qquad (16\text{–}6)$$

If a conical surface is generated by sliding the radius from a fixed point along a closed curve, the space enclosed is called a *solid angle*. The measure Ω of the solid angle is the area intercepted by the angle on a unit sphere with its center at the apex of the solid angle. The solid angle at P subtended by an element of the double layer of area dS (Fig. 6.31) is

$$d\Omega = \frac{\cos\psi}{r^2}\, dS, \qquad (16\text{–}7)$$

if $\cos\psi$ is positive. We shall regard this equation as defining the solid angle subtended by a "directed element of area" by permitting $\cos\psi$ to assume negative values as well as positive. This will make the solid angle subtended by a closed surface zero for an external point and $\pm 4\pi$ for an internal point. Substituting from (7) in (4) and (6), we have

$$V(P) = \frac{1}{4\pi\epsilon}\int \chi\, d\Omega = \frac{1}{4\pi}\int V_0\, d\Omega. \qquad (16\text{–}8)$$

If the layer is uniform, then

$$V(P) = \frac{\chi\Omega}{4\pi\epsilon} = \frac{1}{4\pi} V_0 \Omega, \qquad (16\text{–}9)$$

where Ω is the solid angle subtended at P by the layer (Fig. 6.30).

Similarly the potential of a magnetostatic double layer in an infinite homogeneous medium is

$$U(P) = \frac{1}{4\pi\mu}\int \chi \, d\Omega = \frac{1}{4\pi}\int U_0 \, d\Omega, \tag{16–10}$$

where χ is the strength of the magnetic layer, defined as the magnetic moment per unit area, and U_0 is the sudden rise of the magnetic potential in passing across the layer. If the layer is uniform, then

$$U(P) = \frac{\chi\Omega}{4\pi\mu} = \frac{1}{4\pi} U_0\Omega. \tag{16–11}$$

For an infinite homogeneous conductor the potential is given by (8) with g in place of ϵ.

6.17. *Equivalence of an Electric Current Loop and a Magnetic Double Layer*

Consider a uniform magnetic double layer (Fig. 6.30) of strength $\chi = \mu U_0$. The magnetomotive force along a path ABC leading from a point A on the positive side of the layer to an opposite point C on the negative side is U_0, since the total magnetomotive force round $ABCA$ is zero. Imagine now an electric current loop along the edge of the double layer and let the current I in the loop be regarded as positive when it appears counterclockwise to an observer on the positive side of the layer. The magnetomotive force of the field produced by this current, round any contour such as ABC in Fig. 6.30 is I. Thus in so far as points external to the layer are concerned, the layer and the loop are equivalent if

$$I = U_0. \tag{17–1}$$

Inside the layer the two fields are, of course, very different. Substituting from (1) in (16–11) we have the magnetic potential of an electric current loop

$$U(P) = \frac{I\Omega}{4\pi}, \tag{17–2}$$

at all points outside some surface (S) bounded by the loop.

Let us now consider an infinitely small plane current loop of area S and the corresponding magnetic double layer. The total moment of this magnetic doublet is μIS. Assuming that I is variable, the magnetic doublet becomes a magnetic current element of moment

$$p = Kl = \mu S \frac{\partial I}{\partial t}, \tag{17–3}$$

where K is the magnetic current and l is the length of the element.*

* In dealing with magnetic current elements we are concerned only with the moment Kl, and apart from this product neither K nor l need have definite values.

For harmonic currents we have

$$p = Kl = i\omega\mu SI. \tag{17–4}$$

This relationship between elementary current loops and magnetic current elements makes it very easy to obtain the field of the loop. We have already calculated the field of an electric current element of moment Il. We have also seen that the fields of magnetic currents are obtainable from an electric vector potential F which differs from the magnetic vector potential A only in that magnetic currents appear in the place of electric currents. Thus for an electric current loop in a nondissipative medium we have

$$F = \frac{Kle^{-\sigma r}}{4\pi r} = \frac{i\omega\mu SIe^{-\sigma r}}{4\pi r} = \frac{i\eta\beta SIe^{-i\beta r}}{4\pi r},$$

$$\begin{aligned} E_\varphi &= \frac{\eta\beta^2 SI}{4\pi r}\left(1 + \frac{1}{i\beta r}\right) e^{-i\beta r} \sin\theta, \\ H_\theta &= -\frac{\beta^2 SI}{4\pi r}\left(1 + \frac{1}{i\beta r} - \frac{1}{\beta^2 r^2}\right) e^{-i\beta r} \sin\theta, \\ H_r &= \frac{i\beta SI}{2\pi r^2}\left(1 + \frac{1}{i\beta r}\right) e^{-i\beta r} \cos\theta. \end{aligned} \tag{17–5}$$

At great distances from the loop the field is

$$H_\theta = -\frac{\beta^2 SIe^{-i\beta r}\sin\theta}{4\pi r} = -\frac{\pi SIe^{-i\beta r}\sin\theta}{\lambda^2 r}, \qquad E_\varphi = -\eta H_\theta, \qquad H_r = 0, \tag{17–6}$$

while near the loop it is

$$H_\theta = \frac{SI\sin\theta}{4\pi r^3}, \quad H_r = \frac{SI\cos\theta}{2\pi r^3}, \quad E_\varphi = -\frac{i\omega\mu SI\sin\theta}{4\pi r^2}. \tag{17–7}$$

A large loop carrying current I, uniform over the loop but varying with time, is also equivalent to a uniform double layer over a surface (S) bounded by the loop. In order to show this we need only imagine that (S) is divided into a large number of elementary loops filling the entire surface, each carrying current I in the same direction (Fig. 6.32). The electromagnetic effects of the currents in adjacent sections of the elementary loops cancel out, and the system of loops is equivalent to the large loop. But each elementary loop is equivalent to a magnetic doublet or to an element of the double

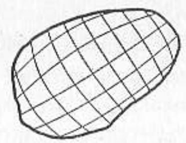


FIG. 6.32. Representation of a large loop carrying uniform current by subdivision into small loops, each carrying the same current.

layer, and hence the loop as a whole will be equivalent to a uniform double layer over (S).

6.18. *Induction and Equivalence Theorems for Stationary Fields*

Next in simplicity to a homogeneous infinite medium is a medium which is homogeneous in each of two regions (1) and (2) separated by a closed surface (S) (Fig. 6.28). Let us suppose that we have a distribution of electric charge in region (1) while region (2) is source-free. Let V^i be the potential of this distribution in an infinite medium with a dielectric constant ϵ_1, equal to the dielectric constant of region (1); we shall call this potential the impressed potential and the corresponding field the impressed field. Let the difference between the actual potential in region (1) and the impressed potential be V^r; we shall call this the reflected potential and the corresponding field the reflected field. Finally let the actual field in region (2) be represented by the transmitted potential V^t. The reflected and the transmitted potentials satisfy Laplace's equation

$$\Delta V^r = 0, \quad \Delta V^t = 0. \tag{18–1}$$

Assuming that there are no sources on the interface (S) between the two regions, we have the following conditions to be satisfied over (S):

$$V^i + V^r = V^t, \quad \epsilon_1 \frac{\partial V^i}{\partial n} + \epsilon_1 \frac{\partial V^r}{\partial n} = \epsilon_2 \frac{\partial V^t}{\partial n}. \tag{18–2}$$

The first of these conditions states that there is no double layer of charge over (S) and the second that there is no simple layer. The above equations together with supplementary requirements of finiteness, continuity and proper behavior at infinity suffice for the calculation of V^r and V^t.

Equations (2) may be rewritten as follows

$$V^t - V^r = V^i, \quad \epsilon_2 \frac{\partial V^t}{\partial n} - \epsilon_1 \frac{\partial V^r}{\partial n} = \epsilon_1 \frac{\partial V^i}{\partial n}. \tag{18–3}$$

Suppose now we have a double layer in region (1) at the boundary (S), with potential discontinuity V^i, and a simple layer of density

$$q_S = \epsilon_1 \frac{\partial V^i}{\partial n}. \tag{18–4}$$

Furthermore let the rest of space be source-free. In order to obtain the field of these surface sources we have to satisfy equations (1) and (3) and supplementary requirements of finiteness, continuity, and proper behavior at infinity which are the same as in the previous problem. In other words the two problems are indistinguishable and the field consisting of the reflected field in region (1) and the transmitted field in region (2) may be produced by the postulated simple and double layers of electric charge. This is the electrostatic version of the Induction Theorem.

If the dielectric constant of region (2) is equal to that of region (1), then there is no reflected field and the transmitted field is identical with the impressed field given by V^i. The induction theorem becomes now an equivalence theorem which states that the simple and double layers defined by (3) produce a field which is equal to zero in region (1) and to the actual field in region (2); that is, the field in the source-free region (2) produced by a system of sources distributed throughout region (1) may also be produced by a proper system of sources over the boundary (S) separating the two regions.

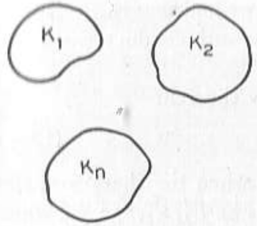


FIG. 6.33. A system of conductors.

In the situation contemplated in the above equivalence theorem the entire space is homogeneous and the expression for the potential of the simple and double layers over (S) can be written at once. Dropping the superscript i, we have

$$V(P) = \frac{1}{4\pi} \iint_{(S)} \left(\frac{1}{r} \frac{\partial V}{\partial n} + \frac{V \cos \psi}{r^2} \right) dS = \frac{1}{4\pi} \iint_{(S)} \left[\frac{1}{r} \frac{\partial V}{\partial n} - V \frac{\partial}{\partial n} \left(\frac{1}{r} \right) \right] dS. \tag{18–5}$$

The magnetostatic field can be treated similarly.

6.19. *Potential and Capacitance Coefficients of a System of Conductors*

Consider a system of n conductors $K_1, K_2, \cdots, K_n$ (Fig. 6.33) with total charges respectively equal to $q_1, q_2, \cdots, q_n$. The potential V is a linear function of the total charges on the conductors

$$V(P) = f_1 q_1 + f_2 q_2 + \cdots + f_n q_n, \tag{19–1}$$

where the coefficients $f_1, f_2, \cdots, f_n$ are functions of position. The function f_m represents the potential due to a unit charge on the mth conductor when the remaining conductors have zero charges. On a conductor electricity moves freely so that, when a steady state has been reached, the tangential component of the electric intensity vanishes and the surface of the conductor becomes an equipotential surface. Designating by $V_1, V_2, \cdots, V_n$ the potentials of the conductors, we have

$$\begin{aligned} V_1 &= p_{11}q_1 + p_{12}q_2 + p_{13}q_3 + \cdots + p_{1n}q_n, \\ V_2 &= p_{21}q_1 + p_{22}q_2 + p_{23}q_3 + \cdots + p_{2n}q_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ V_n &= p_{n1}q_1 + p_{n2}q_2 + p_{n3}q_3 + \cdots + p_{nn}q_n, \end{aligned} \tag{19–2}$$

where the p's are the corresponding values of the f's. These coefficients are called the *potential coefficients*. The set $p_{1m}, p_{2m}, \cdots, p_{nm}$ represents the potentials of the conductors when a unit charge is placed on the mth conductor while the other conductors remain uncharged.

Solving (2) for the q's we have

$$\begin{aligned} q_1 &= c_{11}V_1 + c_{12}V_2 + c_{13}V_3 + \cdots + c_{1n}V_n, \\ q_2 &= c_{21}V_1 + c_{22}V_2 + c_{23}V_3 + \cdots + c_{2n}V_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ q_n &= c_{n1}V_1 + c_{n2}V_2 + c_{n3}V_3 + \cdots + c_{nn}V_n. \end{aligned} \tag{19–3}$$

The capacitance of a conductor is defined as the ratio of its charge to its potential when all other conductors are kept at zero potential. Accordingly, $c_{11}, c_{22}, \cdots, c_{nn}$ are the capacitances of the respective conductors. The remaining c's are known as the coefficients of mutual electrostatic induction. From (3) we find that $c_{1m}, c_{2m}, \cdots, c_{nm}$ are the charges on the conductors if the potential of the mth conductor is unity while all the other conductors are kept at zero potential.

We shall now prove the following electrostatic reciprocity theorem

$$p_{mk} = p_{km}, \quad c_{mk} = c_{km}. \tag{19–4}$$

Let the potentials of $K_1, K_2, \cdots, K_n$ be $V_1, V_2, \cdots, V_n$ when the charges on the conductors are respectively $q_1, q_2, \cdots, q_n$; let the potentials be $V'_1, V'_2, \cdots, V'_n$ when the charges are $q'_1, q'_2, \cdots, q'_n$. Forming the following expressions

$$\sum_m q_m V'_m = \sum_m \iint_{(K_m)} q_v \, dS \iint_{(K)} \frac{q'_v \, dS}{4\pi\epsilon r'} = \iint_{(K)} \iint_{(K)} \frac{q_v q'_v \, dS_1 \, dS_2}{4\pi\epsilon r'}, \tag{19–5}$$

$$\sum_m q'_m V_m = \sum_m \iint_{(K_m)} q'_v \, dS \iint_{(K)} \frac{q_v \, dS}{4\pi\epsilon r'} = \iint_{(K)} \iint_{(K)} \frac{q'_v q_v \, dS_1 \, dS_2}{4\pi\epsilon r'},$$

and comparing, we find

$$\sum_m q_m V'_m = \sum_m q'_m V_m. \tag{19–6}$$

From the way in which this equation has been derived, it is evident that either all the charges or all the potentials are arbitrary. Choosing

$$\begin{aligned} q_m &= 0, \text{ if } m \neq \alpha, & q'_m &= 0, \text{ if } m \neq \beta, \\ &= 1, \text{ if } m = \alpha, & &= 1, \text{ if } m = \beta, \end{aligned} \tag{19–7}$$

and substituting in (6), we have

$$V'_\alpha = V_\beta. \tag{19–8}$$

Thus, *the potential of* K_α *due to a unit charge on* K_β *is the same as the potential of* K_β *due to a unit charge on* K_α.

Similarly if we let

$$\begin{aligned} V_m &= 0, \text{ if } m \neq \alpha, & V'_m &= 0, \text{ if } m \neq \beta, \\ &= 1, \text{ if } m = \alpha, & &= 1, \text{ if } m = \beta, \end{aligned} \tag{19–9}$$

and substitute again in (6), we obtain

$$q_\beta = q'_\alpha; \tag{19–10}$$

that is, *the charge on* K_β *when* K_α *is kept at a unit potential and the remaining conductors at zero potential, is the same as the charge on* K_α *when* K_β *is kept at a unit potential while the remaining conductors are at zero potential.* Equations (8) and (10) are equivalent to (4).

6.20. *Representation of a System of Conductors by an Equivalent Network of Capacitors*

In practical applications it is convenient to represent systems of conductors as networks of capacitors (Fig. 6.34). In this diagram each conductor appears as a terminal.

Since $c_{mk} = c_{km}$, the number of independent c's in a system of n conductors is $n(n+1)/2$. This is equal to the number of capacitors in the corresponding network provided we regard the ground (or infinity, in the case of a system in free space) as the $(n+1)$-th terminal.

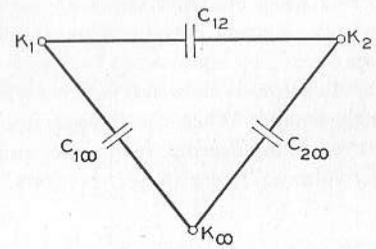


FIG. 6.34. Representation of a system of conductors by a network of capacitors.

On the two plates of a capacitor the electric charges are equal but of opposite sign and are proportional to the voltage between the plates. Our object is to express the relations between the actual potentials of the conductors and the charges on them in terms of the capacitances of the network. The charge q_1 on conductor K_1 is represented in the network as the sum of the charges on all plates connected to the terminal (1). The charge on each plate is the product of the capacitance of the capacitor and the electromotive force from this plate to the opposite plate. This electromotive force is equal to the difference between the potential of the first plate and that of the second. Thus we have the following set of equations

$$\begin{aligned} q_1 &= C_{1\infty}V_1 + C_{12}(V_1 - V_2) + C_{13}(V_1 - V_3) + \cdots + C_{1n}(V_1 - V_n), \\ q_2 &= C_{21}(V_2 - V_1) + C_{2\infty}V_2 + C_{23}(V_2 - V_3) + \cdots + C_{2n}(V_2 - V_n), \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ q_n &= C_{n1}(V_n - V_1) + C_{n2}(V_n - V_2) + C_{n3}(V_n - V_3) + \cdots + C_{n\infty}V_n, \end{aligned} \tag{20–1}$$

where $C_{m\infty}$ is the direct capacity of the mth conductor to infinity (or to ground) which is taken at zero potential. Collecting terms according to the V's and comparing the result with (19–3), we obtain

$$\begin{aligned} c_{11} &= C_{1\infty} + C_{12} + C_{13} + \cdots + C_{1n}, \\ c_{22} &= C_{2\infty} + C_{21} + C_{23} + \cdots + C_{2n}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ c_{nn} &= C_{n\infty} + C_{n1} + C_{n2} + \cdots + C_{n,n-1}, \\ c_{mk} &= -C_{mk} \text{ if } m \neq k, \quad C_{mk} = C_{km}. \end{aligned} \tag{20–2}$$

Solving for the C's, we have

$$\begin{aligned} C_{mk} &= -c_{mk}, \text{ if } m \neq k, \quad m, k = 1, 2, 3, \cdots n, \\ C_{m\infty} &= c_{m1} + c_{m2} + c_{m3} + \cdots + c_{mm} + \cdots + c_{mn}. \end{aligned} \tag{20–3}$$

Since the capacitance of a capacitor is essentially positive, the coefficients of electrostatic induction are essentially negative. The quantity C_{12} is called the *direct capacitance* between conductors K_1 and K_2; $C_{1\infty}$ is the direct capacitance between K_1 and the ground or infinity. The capacitance of a given conductor is seen to be equal to the sum of the direct capacitances between this conductor and the remaining conductors, including the ground or infinity.

6.21. *Energy Theorems for Stationary Fields*

The energy theorems for stationary fields are special cases of the general theorems derived in section 4.8. Consider, for example, an electrostatic field. During the period when electric charges are separated and moved to their final positions or to infinity, electric currents flow and impressed forces perform work. This work is represented by equation (4.8–2) with $M = 0$. The surface (S) surrounding the separated charges is chosen so far from them that the electrostatic field outside it is vanishingly small. When the charges assume their final positions and thus the impressed currents stop flowing, H vanishes and E becomes stationary within a rapidly expanding volume; hence (4.8–2) becomes

$$W = \tfrac{1}{2}\iiint_{(\tau)} \epsilon E^2 \, d\tau + W_0, \qquad (21\text{–}1)$$

where W is the total work performed by the impressed forces and W_0 is a constant representing the first and third terms on the right-hand side of (4.8–2). This constant W_0 represents the energy lost in heat and by radiation and it can be made arbitrarily small by moving the charges sufficiently slowly. We account for the work W by saying that the system of charges has acquired potential energy. The idea that this energy is distributed throughout the field is consistent with equation (1); but naturally this is not a necessary conclusion from the equation.

By Green's theorem (1.7–1) we can transform (1) into

$$W = \tfrac{1}{2}\epsilon\iiint E^2 \, dv = \tfrac{1}{2}\epsilon\iiint (\text{grad } V, \text{grad } V)\, dv$$
$$= -\tfrac{1}{2}\epsilon\iint_{(K)} V\frac{\partial V}{\partial n}\,dS - \tfrac{1}{2}\epsilon\iiint V\,\Delta V\,dv = -\tfrac{1}{2}\epsilon\iint_{(K)} V\frac{\partial V}{\partial n}\,dS, \qquad (21\text{–}2)$$

where n is the normal pointing into the volume occupied by the field and (K) represents the surface consisting of the surfaces of all the conductors together with a surface at infinity. The latter, however, contributes nothing to W. Over the surface of each conductor V is constant and

$$-\epsilon\iint_{(K_m)} \frac{\partial V}{\partial n}\,dS = \iint_{(K_m)} D_n\,dS = q_m. \qquad (21\text{–}3)$$

Substituting in (2), we have

$$W = \tfrac{1}{2}\sum_{m=0}^{n} V_m q_m. \qquad (21\text{–}4)$$

In view of (19–2) and (19–3) we also have

$$W = \tfrac{1}{2}\sum_{m,\,k}\sum p_{mk} q_m q_k = \tfrac{1}{2}\sum_{m,\,k}\sum c_{mk} V_m V_k. \qquad (21\text{–}5)$$

An equation similar to (1) may be obtained for a magnetostatic field

$$W = \tfrac{1}{2}\iiint_{(\tau)} \mu H^2\,d\tau + W_0. \qquad (21\text{–}6)$$

This equation holds also for the stationary magnetic field due to steady electric currents except that W_0 will not be constant. It will be remembered that one of the terms in W_0 represents the energy transformed into heat by currents flowing in conductors and in a stationary state this term is proportional to time.

6.22. *The Method of Images*

The potential of a system of isolated electric point charges in an infinite homogeneous medium is

$$V = \frac{q_1}{4\pi\epsilon r_1} + \frac{q_2}{4\pi\epsilon r_2} + \cdots + \frac{q_n}{4\pi\epsilon r_n}, \qquad (22\text{–}1)$$

where the r's are the distances from a point in space to the corresponding point charges. Imagine now that an infinitely thin perfectly conducting sheet is introduced over an equipotential surface. Since the component of E tangential to an equipotential surface is zero, the boundary condition at the conducting sheet is satisfied by the existing field which is therefore unaltered. The charges on one side of the sheet are said to be the *images* of the charges on the other side. If all the charges are on the same side of the sheet, they are the images of a point charge at infinity.

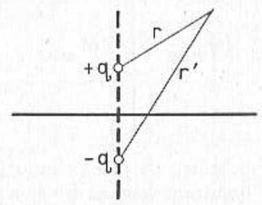


FIG. 6.35. A point charge above a conducting plane and its image.

For example, in the case of two equal and opposite charges (Fig. 6.35) the plane perpendicular to the bisector of the line joining the charges is an equipotential. Here the two charges are "mirror images" of each other. The potential

$$V = \frac{q}{4\pi\epsilon}\left(\frac{1}{r} - \frac{1}{r'}\right) \qquad (22\text{–}2)$$

determines the field of the two charges both before and after the introduction of the conducting sheet. Let us now remove the charge $-q$ to infinity below the plane. This movement will not affect the field above the plane and will reduce the field below the plane to zero. Above the plane the effect of the image charge is replaced by a distribution of charge on the plane. The density of this distribution is

$$D_n = \epsilon E_n = -\epsilon\frac{\partial V}{\partial z} = -\frac{q}{4\pi}\frac{\partial}{\partial z}\left(\frac{1}{r} - \frac{1}{r'}\right), \qquad (22\text{–}3)$$

where the differentiation is performed at the plane. If h is the distance from the point charge to the plane, then (3) becomes

$$D_n = -\frac{qh}{2\pi r^3}. \qquad (22\text{–}4)$$

Thus the density of the charge induced on the plane by the given point charge varies as the cube of the distance from the point charge. By integration we can demonstrate that the total induced charge is $-q$.

Similarly we find the potential of a point charge inside a quadrant formed by two perpendicular conducting planes (Fig. 6.36) in the form

$$V = \frac{q}{4\pi\epsilon}\left(\frac{1}{r_1} - \frac{1}{r_2} + \frac{1}{r_3} - \frac{1}{r_4}\right). \tag{22-5}$$

In the case of two parallel planes the number of images is infinite. Let the two planes (Fig. 6.37) be $z = 0$ and $z = c$ and the point charge be at $z = h$. By placing an image charge $-q$ at $z = -h$, we make $z = 0$ an equipotential plane. In order to

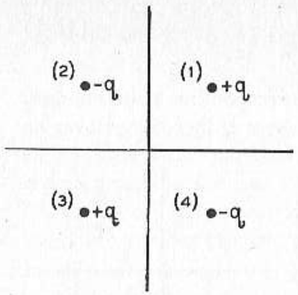


FIG. 6.36. A point charge in a quadrant formed by two conducting planes and its images.

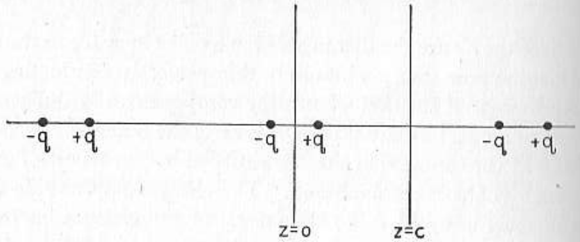


FIG. 6.37. A point charge between two parallel planes and its images.

make $z = c$ an equipotential, we place a charge $-q$ at $z = 2c - h$ and another charge q at $z = 2c + h$. These charges change the status of $z = 0$ and in order to make this plane once more an equipotential we place q at $z = -(2c - h)$ and $-q$ at $z = -(2c + h)$. This process has to be continued indefinitely; thus we have an infinite system of charges

$$\begin{aligned} &q : h, \quad -2c + h, \quad 2c + h, \quad -4c + h, \quad 4c + h, \cdots \\ &-q : -h, \quad 2c - h, \quad -2c - h, \quad 4c - h, \quad -4c - h, \cdots \end{aligned} \tag{22-6}$$

and the potential of the given point charge between the two planes may be expressed as an infinite series

$$V = \frac{q}{4\pi\epsilon}\sum_{-\infty}^{\infty}\frac{1}{\sqrt{\rho^2 + (z - h + 2nc)^2}} - \frac{q}{4\pi\epsilon}\sum_{-\infty}^{\infty}\frac{1}{\sqrt{\rho^2 + (z + h + 2nc)^2}}, \tag{22-7}$$

where ρ is the distance from the line of charges.

The method of images can be used to satisfy other boundary conditions. Thus if the image source in Fig. 6.35 is of the same sign as the given source, the normal derivative of the potential will vanish at the plane. This is the boundary condition at a perfect " magnetic conductor " in the case of electrostatic fields, at a perfect electric conductor in the case of magnetostatic fields, and at a perfect insulator in the case of steady electric current flow. In each case either the normal component of displacement or the normal component of current density is required to vanish. Thus if instead of a point charge q in Fig. 6.35, we have a point source of electric current I,

then the potential in the semi-infinite homogeneous conductor bounded by a perfectly insulating plane is

$$V = \frac{I}{4\pi g}\left(\frac{1}{r} + \frac{1}{r'}\right), \tag{22-8}$$

where g is the conductivity of the medium.

To summarize: the image of a simple source in an infinite plane is equal in strength to the source but of opposite sign if the potential or the tangential component of the field intensity is required to vanish at the plane; the image is equal in strength to the source and has the same sign if the normal derivative of the potential or the normal component of the field intensity has to vanish at the plane.

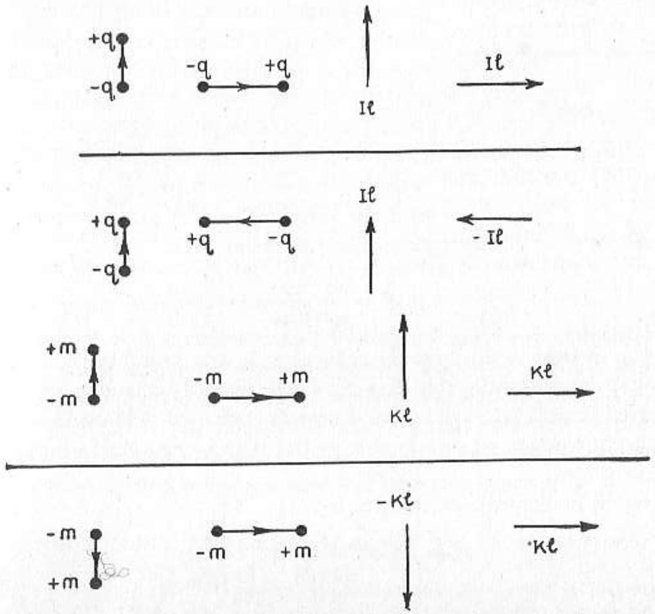


FIG. 6.38. The images of various doublets and current elements in a perfectly conducting plane.

This rule can be broadened to include doublets and current elements. Consider for example a perfectly conducting plane and a variety of doublets and current elements (Fig. 6.38). The images of the following sources have the same sign as the sources: the electric doublet and the electric current element normal to the plane; the magnetic doublet and the magnetic current element tangential to the plane. The remaining sources have images of opposite sign: the electric doublet and the electric current element parallel to the plane; the magnetic doublet and the magnetic current element normal to the plane.

It is easy to verify that the above rule applies to electric and magnetic current elements with variable moments. The rule for the images of electric loops is identical, of course, with the rule for magnetic current elements.

We shall now extend the method of images in another direction. Let the plane be an interface between two homogeneous dielectrics (Fig. 6.39). Consider a point charge q at point A in the upper medium. As has been explained in section 18 we may regard the total field in this medium as the sum of the impressed field, defined as the field of the point charge on the assumption that $\epsilon_2 = \epsilon_1$, and the reflected field. We already know that in at least two cases the reflected field is equal to the one produced by an image charge at point B, which is the geometric image of point A. Thus if $\epsilon_2 = 0$, then the displacement density in the lower medium is identically zero and therefore the normal component of the displacement density in the upper medium should vanish at the boundary; in this case the image charge "producing" the reflected field is $q^r = q$. If $\epsilon_2 = \infty$, then the electric intensity in the lower medium must vanish, or else the displacement density would be infinite; in this case the tangential component of the electric intensity in the upper medium should vanish at the boundary and the reflected field could be produced by $q^r = -q$. Furthermore if $\epsilon_2 = \epsilon_1$, the reflected field should vanish. With this information in mind, we assume tentatively that the reflected field in general could be produced by an image charge at B, having the following value

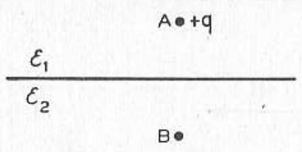


FIG. 6.39. Illustrating the image theory for two dielectric media.

$$q^r = \frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2} q. \tag{22-9}$$

The ratio q^r/q defined in this manner reduces to 1, −1, and 0 in the three cases considered above; but naturally this does not mean that (9) is true in general. In fact, we do not even know that in the general case the reflected field could be produced by an isolated point charge; we merely start with (9) as a hypothesis which can be either proved or disproved.

The potential of the pair of charges is

$$V + V^r = \frac{q}{4\pi\epsilon_1 r_1} + \frac{q^r}{4\pi\epsilon_1 r_2}, \tag{22-10}$$

where r_1 and r_2 are respectively the distances from A and B. The potential along the boundary is

$$\bar{V} + \bar{V}^r = \frac{q + q^r}{4\pi\epsilon_1 r_1} = \frac{q}{2\pi(\epsilon_1 + \epsilon_2) r_1}. \tag{22-11}$$

The normal component of the displacement density (also at the boundary) is

$$\bar{D} + \bar{D}^r = \epsilon_1 \frac{\partial(\bar{V} + \bar{V}^r)}{\partial z} = \frac{(q - q^r)h}{4\pi r_1^3}. \tag{22-12}$$

This displacement density is such as could be produced by a charge q^t at point A, where

$$q^t = q - q^r = \frac{2\epsilon_2}{\epsilon_1 + \epsilon_2} q. \tag{22-13}$$

Naturally the field of q^t is source-free in the lower region where it satisfies Laplace's equation. Assume tentatively that the transmitted field in the lower region is given by

$$V^t = \frac{q^t}{4\pi\epsilon_2 r_1} = \frac{q}{2\pi(\epsilon_1 + \epsilon_2) r_1}. \tag{22-14}$$

At the boundary this potential is equal to that given by (11). Thus both boundary conditions are satisfied and we may finally say that the field of a point charge q, located at a point A in a semi-infinite homogeneous medium separated by a plane from another semi-infinite homogeneous medium, may be represented as follows: (1) on the same side of the boundary as point A, the field is the sum of the fields which would be produced in an infinite medium by the original charge at A and by an image charge q^r at B, assuming that the dielectric constant of the medium is ϵ_1; (2) on the other side of the boundary the field is the same as that which would be produced by a charge q^t, placed at A, in an infinite medium with the dielectric constant ϵ_2.

For magnetic fields we have a similar theorem. In the above formulae electric charges are replaced by magnetic charges and ϵ's are replaced by μ's. The rules for doublets can be formulated very readily since doublets are pairs of point charges.

These theorems do not apply in general to variable fields. That this is the case is obvious when the intrinsic propagation constants of the two media are different; the fields of simple point sources cannot possibly be matched along the entire plane boundary. But it is conceivable that such fields could be matched when the propagation constants are the same; and this is actually found to be the case.

6.23. *Two-Dimensional Stationary Fields*

A two-dimensional field is defined as a field depending on two coordinates and, in particular, as a field depending on two cartesian coordinates. While such fields are special cases of three-dimensional fields, the simplifications resulting from the decrease in the number of effective coordinates are so great that two-dimensional fields are usually studied separately. Assuming that the field is independent of the z-coordinate, we obtain the following equations for source-free homogeneous regions under different conditions.

For an electrostatic field the potential satisfies the two-dimensional Laplace's equation

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0, \tag{23-1}$$

which in polar coordinates becomes

$$\rho \frac{\partial}{\partial \rho}\left(\rho \frac{\partial V}{\partial \rho}\right) + \frac{\partial^2 V}{\partial \varphi^2} = 0. \tag{23-2}$$

The electric intensity is equal to $-\text{grad}\, V$; thus

$$E_x = -\frac{\partial V}{\partial x}, \quad E_y = -\frac{\partial V}{\partial y}; \tag{23-3}$$

$$E_\rho = -\frac{\partial V}{\partial \rho}, \quad E_\varphi = -\frac{\partial V}{\rho \partial \varphi}. \tag{23-4}$$

Exactly the same set of equations describes the steady current flow. For magnetostatic fields the equations are similar, with the magnetic potential U taking the place of the electric potential V and H appearing in place of E.

Magnetic fields produced by electric currents are derivable from the vector potential A, which, when the current is parallel to the z-axis, has only one component A_z. Thus such fields depend essentially on one scalar function $A_z = \Psi$, usually called the *stream function*. In terms of this function we have

$$H_x = \frac{\partial \Psi}{\partial y}, \quad H_y = -\frac{\partial \Psi}{\partial x}; \tag{23–5}$$

$$H_\rho = \frac{\partial \Psi}{\rho \partial \varphi}, \quad H_\varphi = -\frac{\partial \Psi}{\partial \rho}. \tag{23–6}$$

In source-free regions the stream function satisfies equations (1) and (2).

A two-dimensional electrostatic field is produced by a system of uniform filaments of electric charge parallel to the z-axis. These filaments may form either a discrete or a continuous set. A uniform *line source* is an elementary source of such a field in the same sense as a point source is an elementary source of a three-dimensional field. The potential of a line source is independent of the φ-coordinate. Hence from (2) we have

$$\rho \frac{\partial V}{\partial \rho} = P, \quad V = P \log \rho + \text{constant}. \tag{23–7}$$

Therefore,

$$E_\rho = -\frac{P}{\rho}. \tag{23–8}$$

If q is the electric charge per unit length of the filament, then by taking the radial displacement over the surface of a cylinder concentric with the filament we obtain

$$2\pi\rho\epsilon E_\rho = q, \quad \text{and} \quad P = -\frac{q}{2\pi\epsilon}. \tag{23–9}$$

Substituting in (7), we have

$$V = -\frac{q}{2\pi\epsilon} \log \frac{\rho}{a}, \tag{23–10}$$

where a is a constant length which remains arbitrary. For the electric intensity itself we have

$$E_\rho = \frac{q}{2\pi\epsilon\rho}, \quad E_\varphi = 0. \tag{23–11}$$

The last two equations can be derived without using Laplace's equation. Thus (11) follows directly from symmetry considerations and from the divergence equation (4.3–2). Furthermore it is evident that E can be expressed as the gradient of a function depending only on ρ and that this function is given by (10). Laplace's equation becomes of real value, however, when only a part of the complete distribution of electric charge is known and the information regarding the remaining charges is replaced by boundary conditions. Consider for example a conducting cylindrical tube and a known line charge parallel to the axis of this tube. Instead of being given the distribution of electric charge on the cylinder we are required to find it, using the boundary condition that the component of E tangential to the tube vanishes. This time our

problem is to find a reflected field satisfying (2) and having a tangential component equal and opposite to the tangential component of the field which would be produced by the line charge in an infinite medium.

In the case of magnetic fields produced by a distribution of parallel currents the elementary line source is a uniform infinitely thin current filament carrying current I. By (4.6–1) we have

$$H_\varphi = \frac{I}{2\pi\rho}, \quad H_\rho = 0. \tag{23–12}$$

When the field is steady, there is no displacement current parallel to the filament and (12) is valid at any distance from the element, not only in its immediate vicinity. Comparing (12) with (6), we find that H may be obtained from the following stream function

$$\Psi = -\frac{I}{2\pi} \log \frac{\rho}{a}, \tag{23–13}$$

where a is an arbitrary constant.

The value of the stream function becomes evident when we attempt to find the field of several current filaments. Such a field may be calculated directly from (12) by adding vectorially the magnetic intensities of the individual current filaments. On the other hand the stream function is a scalar and the addition of stream functions is much simpler. For example in the case of two filaments (Fig. 6.40) one passing through point $(l/2,0)$ carrying current I and the other through $(-l/2,0)$ with current $-I$, we have

$$\Psi = \frac{I}{2\pi} \log \frac{\rho_2}{\rho_1}. \tag{23–14}$$

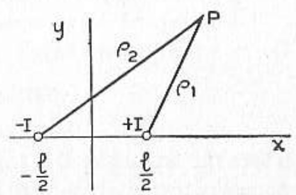


FIG. 6.40. The cross-section of two infinitely long parallel wires carrying equal and opposite currents.

When ρ_1 and ρ_2 are large compared with $l/2$, we have

$$\rho_2 = \rho + \frac{l}{2} \cos \varphi, \quad \rho_1 = \rho - \frac{l}{2} \cos \varphi, \tag{23–15}$$

and (14) becomes

$$\Psi = \frac{Il \cos \varphi}{2\pi\rho}. \tag{23–16}$$

By (6) we now have

$$H_\rho = -\frac{Il \sin \varphi}{2\pi\rho^2}, \quad H_\varphi = \frac{Il \cos \varphi}{2\pi\rho^2}. \tag{23–17}$$

It should be noted that in general V and Ψ become logarithmically infinite at infinity. But when the total charge is zero, while the total amount of charge of either sign is finite, then the potential vanishes at infinity; similarly when the total current is zero, while the currents flowing in either direction are finite, Ψ vanishes at infinity. This is certainly the case for a pair of oppositely directed current filaments. Any other distribution, for which the total current is zero, may be subdivided into pairs of oppositely directed current filaments. Thus we have the general theorem.

The method of images can evidently be applied to two-dimensional fields. The rules for the magnitudes and the signs of the images of line sources are the same as for corresponding point sources.

6.24. *The Inductance of a System of Parallel Currents*

Generally, in the case of steady parallel currents, the stream function satisfies the two-dimensional Poisson's equation

$$\Delta\Psi = \frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2} = -J, \qquad (24\text{–}1)$$

where J is the total current density. The energy of the magnetic field produced by this current distribution may be expressed in terms of Ψ; thus the energy per unit length in the z-direction is

$$W = \tfrac{1}{2}\mu\iint (H_x^2 + H_y^2)\, dS = \tfrac{1}{2}\mu\iint\left[\left(\frac{\partial\Psi}{\partial x}\right)^2 + \left(\frac{\partial\Psi}{\partial y}\right)^2\right] dS, \qquad (24\text{–}2)$$

where the integration is extended over any plane normal to the z-axis.

To begin with let us consider the above integral extended over a finite area. By Green's theorem we have

$$W = \tfrac{1}{2}\mu\int \Psi\frac{\partial\Psi}{\partial n}\, ds - \tfrac{1}{2}\mu\iint \Psi\Delta\Psi\, dS, \qquad (24\text{–}3)$$

where the line integral is taken over the periphery of the chosen area. Let this area increase indefinitely in both linear dimensions and assume that J is distributed over a finite area. If the total current in the z-direction is different from zero W will also increase indefinitely. On the other hand if the total current is zero, as is the case in practice, then Ψ varies ultimately as* $1/\rho$ and $\partial\Psi/\partial n$ as $1/\rho^2$, consequently the line integral in (3) approaches zero. Substituting from (1) in (3), we thus obtain

$$W = \tfrac{1}{2}\mu\iint J\Psi\, dS. \qquad (24\text{–}4)$$

Effectively, this integration is extended only over the areas occupied by the current.

Let us now consider a system of parallel wires and let the currents in these wires be uniformly distributed throughout their cross-sections. In this case J is constant for each wire and (4) becomes

$$W = \tfrac{1}{2}\mu\sum J_m\iint_{(S_m)} \Psi\, dS, \qquad (24\text{–}5)$$

where $S_1, S_2, \cdots$ are the cross-sections of the various wires. Introducing the average values of the stream function Ψ over each cross-section

$$\Psi_m = \frac{1}{S_m}\iint_{(S_m)} \Psi\, dS, \qquad (24\text{–}6)$$

* Ψ may contain a constant which does not affect the field and hence may be taken as zero.

and noting that the total current I_m in the mth wire is $I_m = J_mS_m$, we transform (5) into

$$W = \tfrac{1}{2}\mu\sum\Psi_mI_m. \qquad (24\text{–}7)$$

If there are only two wires, carrying equal and opposite currents, then

$$I_1 = I, \quad I_2 = -I,$$

and the energy of the field per unit length along the wires is

$$W = \tfrac{1}{2}\mu(\Psi_1 - \Psi_2)I. \qquad (24\text{–}8)$$

The values of Ψ_1 and Ψ_2 are proportional to I so that

$$W = \tfrac{1}{2}LI^2,$$

where the coefficient L is seen to be the inductance per unit length of the wires and its value is

$$L = \mu\frac{\Psi_1 - \Psi_2}{I}. \qquad (24\text{–}9)$$

For two pairs of wires let

$$I_3 = -I_1, \quad I_4 = -I_2;$$

then equation (7) becomes

$$W = \tfrac{1}{2}\mu(\Psi_1 - \Psi_3)I_1 + \tfrac{1}{2}\mu(\Psi_2 - \Psi_4)I_2. \qquad (24\text{–}10)$$

The average values of the stream functions are linear functions of I_1 and I_2: thus

$$\mu(\Psi_1 - \Psi_3) = L_{11}I_1 + L_{12}I_2,$$
$$\mu(\Psi_2 - \Psi_4) = L_{21}I_1 + L_{22}I_2.$$

Substituting in (10), we have

$$W = \tfrac{1}{2}L_{11}I_1^2 + \tfrac{1}{2}(L_{12} + L_{21})I_1I_2 + \tfrac{1}{2}L_{22}I_2^2.$$

These formulae may be extended to n pairs of wires.

Thus in order to compute the inductance coefficients of a system of parallel currents we have to compute first the stream function Ψ and then its average values over the

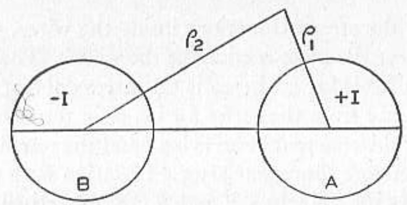


FIG. 6.41. The cross-section of two parallel cylinders.

cross-sections of the different wires. Two examples will illustrate the procedure. First let us take a pair of wires of circular cross-section (Fig. 6.41). We have already assumed that the current distribution is uniform throughout the cross-section of each wire. From symmetry considerations we conclude that the stream function for each wire, in the region external to it, is equal to the stream function that would be obtained if the entire current were concentrated along the axis of the wire. Thus for points external to both wires the stream function is given by (23–14). Inside the wire A the magnetic intensity due to the current in the wire itself is

$$H'_{\varphi_1} = \frac{\rho_1 I}{2\pi a^2}, \tag{24–11}$$

where (ρ_1, φ_1) are cylindrical coordinates referred to the axis of A and a is the radius of the wire. We now obtain that part Ψ' of the total Ψ which is due to the current in A by integrating (23–6) using (11); thus

$$\Psi' = -\frac{\rho_1^2 I}{4\pi a^2} + P, \tag{24–12}$$

where P is a constant of integration. This constant is determined so as to make Ψ' continuous at the surface of the wire; thus

$$-\frac{I}{2\pi}\log a = -\frac{I}{4\pi} + P, \quad P = \frac{I}{4\pi} - \frac{I}{2\pi}\log a,$$

and equation (12) becomes

$$\Psi' = -\frac{\rho_1^2 I}{4\pi a^2} + \frac{I}{4\pi} - \frac{I}{2\pi}\log a.$$

This is the stream function due to the current in A; adding to it the stream function due to the current in the wire B we have the total stream function in the region $\rho_1 \leq a$ in the following form

$$\Psi = \frac{I}{2\pi}\log \rho_2 + \Psi' = \frac{I}{2\pi}\log\frac{\rho_2}{a} + \frac{I}{4\pi}\left(1 - \frac{\rho_1^2}{a^2}\right).$$

Similarly for the region $\rho_2 \leq b$, where b is the radius of the wire B, we have

$$\Psi = -\frac{I}{2\pi}\log\frac{\rho_1}{b} - \frac{I}{4\pi}\left(1 - \frac{\rho_2^2}{b^2}\right).$$

Having determined the stream functions inside the wires, we proceed to compute their average values over the cross-sections of the wires. The average value of $\log \rho_2$ over the cross-section of A is $\log l$, where l is the interaxial distance between the wires. This result follows at once from the series for $\log \rho_2$ in terms of the polar coordinates referred to A, in which the constant term is $\log l$ and the remaining terms are periodic functions of φ_1 and therefore disappear after integration with respect to φ_1 from 0 to 2π. The total averages for the wires A and B are respectively

$$\Psi_1 = \frac{I}{2\pi}\log l - \frac{I}{2\pi}\log a + \frac{I}{8\pi},$$

$$\Psi_2 = -\frac{I}{2\pi}\log l + \frac{I}{2\pi}\log b - \frac{I}{8\pi}.$$

Substituting in (9), we find the inductance of the pair of wires per unit length

$$L = \frac{\mu}{2\pi}\log\frac{l^2}{ab} + \frac{\mu}{4\pi} = \frac{\mu}{\pi}\log\frac{l}{\sqrt{ab}} + \frac{\mu}{4\pi}.$$

The simplicity of this method for calculating inductances is remarkable and no complications arise in the case of several pairs. Since the self-inductances of each pair are already known, we need to calculate only the mutual inductances. For the two pairs of wires in Fig. 6.42 we have to find the averages of the logarithms of the distances from A and B over the cross-sections of the other two wires and vice versa; these averages are the logarithms of the interaxial distances. Paying proper attention to the algebraic signs, we obtain

$$L_{12} = L_{21} = \frac{\mu}{2\pi}\log\frac{l_{AD}l_{BC}}{l_{AC}l_{BD}}.$$

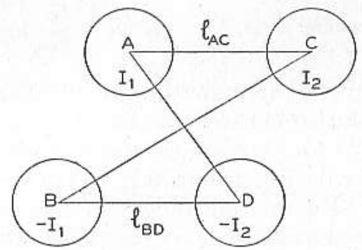


FIG. 6.42. The cross-section of two pairs of cylinders.

Incidentally we have proved the reciprocity theorem for the mutual inductance coefficients.

6.25. *Functions of Complex Variables and Stationary Fields*

We have seen that in source-free regions static potentials and stream functions satisfy the two-dimensional Laplace's equation. This equation is also satisfied by the real and imaginary parts of an arbitrary monogenic function of the complex variable $z = x + iy$. In the theory of partial differential equations it is shown in fact that the most general solution of

$$\frac{\partial^2 W}{\partial x^2} + \frac{\partial^2 W}{\partial y^2} = 0 \tag{25–1}$$

may be represented in the following form

$$W = F(z) + G(z^*), \tag{25–2}$$

where F and G are arbitrary functions and

$$z = x + iy, \quad z^* = x - iy. \tag{25–3}$$

If W is to be real, then F and G should be chosen as identical functions of their respective variables; hence the most general potential in a source-free region may be regarded as the real part of a "complex potential" W. The same real potential is the imaginary part of iW, which is also a function of the complex variable z. Likewise the stream function can be represented either as a real part or as an imaginary part of a "complex stream function" W.

If the real and imaginary parts of a complex potential are separated

$$W = V + i\Psi, \tag{25–4}$$

then the imaginary part Ψ is the negative* of the stream function for the field represented by the potential V. This follows from the definitions of the potential and stream function, inherent in (23–3) and (23–5), and from the Cauchy-Riemann equations. Similarly if the stream function is represented by the real part of a complex stream function, then the imaginary part represents the potential of the same field.

Consider for example the following complex potential

$$W = -\frac{q}{2\pi\epsilon}\log z = -\frac{q}{2\pi\epsilon}\log\rho - i\frac{q}{2\pi\epsilon}\varphi; \tag{25–5}$$

its real part represents the potential of a line charge of density q. The same field can also be obtained from the imaginary part with the aid of an electric analog of equations (23–6). If q/ϵ is replaced by I, then the real part will represent the stream function for the field produced by an electric current filament.

The field intensities may also be represented as the real and imaginary parts of appropriate complex functions. Thus for an electrostatic field we have

$$E_x - iE_y = -\frac{\partial V}{\partial x} + i\frac{\partial V}{\partial y} = -\frac{\partial V}{\partial x} - i\frac{\partial \Psi}{\partial x} = -\frac{dW}{dx}. \tag{25–6}$$

Hence the real and imaginary parts of the derivative of the complex potential give, except for an algebraic sign, the cartesian components of the electric intensity and the magnitude of the intensity is

$$\sqrt{E_x^2 + E_y^2} = \sqrt{\frac{dW}{dz}\frac{dW^*}{dz^*}}. \tag{25–7}$$

Starting from any complex potential, we can obtain a distribution of sources which s capable of producing the corresponding field. Take, for example,

$$W = i\frac{q}{2\pi\epsilon}\log z, \tag{25–8}$$

which differs from (5) only by a constant $-i$. The real potential is

$$V = -\frac{q}{2\pi\epsilon}\varphi. \tag{25–9}$$

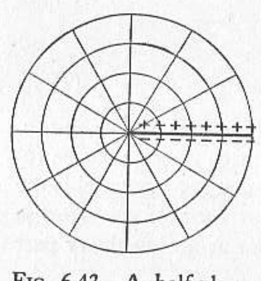


FIG. 6.43. A half-plane double layer.

This function happens to be a multiple-valued function while the electric potential in source-free regions is single-valued. If we make (9) single-valued by restricting φ to the interval $(0,2\pi)$, we assume in effect that the sources of the field are located on the positive xz-plane. Since the potential rise across this plane is q/ϵ, there must be a double layer on this plane of strength q (Fig. 6.43). The electric intensity is

$$E_\rho = 0, \quad E_\varphi = \frac{q}{2\pi\epsilon\rho}. \tag{25–10}$$

* The algebraic sign of potential and stream functions is a matter of convention.

Since E_φ is continuous, there are no simple layers of charge. The electric lines are circles and the equipotential lines are radii.

Since the radial planes are equipotential, we can assume any pair of them to be perfect conductors insulated from each other along the line passing through the origin (Fig. 6.44). The displacement density along one plane, passing through the x-axis, is

$$D_\varphi = \epsilon E_\varphi = \frac{q}{2\pi\rho}. \tag{25–11}$$

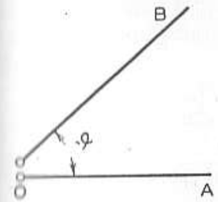


FIG. 6.44. A wedge formed by two half-planes.

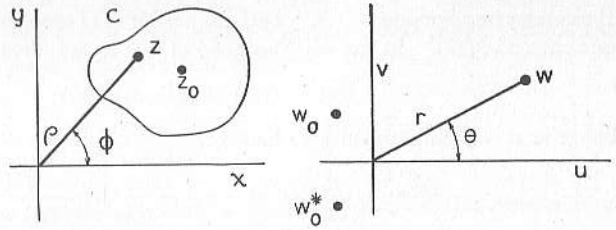


FIG. 6.45. Illustrating the conformal transformation of a region bounded by a closed curve into the upper half-plane.

The displacement density at the other plane is the negative of this. Since the potential difference between the planes is $q\vartheta/2\pi\epsilon$, where ϑ is the angle between them, the capacitance between the planes per unit area at distance ρ from their adjacent edges is

$$C = \frac{\epsilon}{\rho\vartheta}. \tag{25–12}$$

Let us now consider the general problem of a line charge in the presence of a perfectly conducting cylindrical boundary whose generators are parallel to the line charge. Let (C) be the contour of the conducting boundary (Fig. 6.45) and let the complex number z_0 designate the position of the line charge. Suppose that we have found a function

$$w = u + iv = f(z), \tag{25–13}$$

such that the curve (C) goes point by point into the u-axis and the interior of the region bounded by (C) goes into the upper half of the w-plane. Let w_0 be the point corresponding to z_0 and assume that a line charge of density q is passing through w_0. The complex potential of this charge is

$$W_1 = -\frac{q}{2\pi\epsilon}\log(w - w_0). \tag{25–14}$$

If a perfectly conducting plane is assumed to pass through the u-axis, then its effect on the field in the upper half-plane may be represented by the potential of the image charge of density $-q$, located at the image point w_0^*; this image potential is

$$W_2 = \frac{q}{2\pi\epsilon}\log(w - w_0^*). \tag{25–15}$$

The total potential is therefore

$$W = -\frac{q}{2\pi\epsilon}\log\frac{w - w_0}{w - w_0^*}. \tag{25–16}$$

If now we substitute from (13) into (16), we obtain

$$W = -\frac{q}{2\pi\epsilon}\log\frac{f(z) - f(z_0)}{f(z) - f(z_0^*)}, \tag{25–17}$$

which is the complex potential of the line charge passing through z_0 in the presence of the conducting boundary (C). The real part of (17) reduces to zero on (C) because of our choice of $f(z)$. In the neighborhood of $z = z_0$, we have

$$f(z) - f(z_0) = (z - z_0)f'(z_0); \tag{25–18}$$

hence in this neighborhood (17) becomes

$$W = -\frac{q}{2\pi\epsilon}\log(z - z_0) - \frac{q}{2\pi\epsilon}\log\frac{f'(z_0)}{f(z_0) - f(z_0^*)}. \tag{25–19}$$

The first term of this expression represents the potential of the line charge in the infinite medium and the second term its modification due to the boundary. The second term is constant and does not affect the charge density on the source; hence the real part of (17) satisfies all the requirements of our original problem.

More generally (17) may be represented in the form

$$W = -\frac{q}{2\pi\epsilon}\log(z - z_0) - \frac{q}{2\pi\epsilon}\log\frac{f(z) - f(z_0)}{(z - z_0)[f(z) - f(z_0^*)]}, \tag{25–20}$$

in which the effect of the boundary on the potential is given explicitly. The second term in (20) has no singularities in the region bounded by (C).

If instead of an infinitely thin filament, we assume a thin circular wire of radius a, the complex potential of the wire is given approximately by (19). In this approximation we ignore the redistribution of charge round the wire due to the fact that the reflected potential is not really constant but varies from point to point; actually the wire is in a transverse electric field which forces some positive charge from one side of the wire to the other. In fact, following this line of thought we may obtain a more accurate expression for the field of charge on a conducting wire of finite radius. However, if the radius is small compared with the shortest distance from the wire to the boundary, then (19) is a good approximation and in this case the capacitance of the wire per unit length is

$$C = \frac{q}{V} = \frac{2\pi\epsilon}{\log|f(z_0) - f(z_0^*)| - \log|f'(z_0)| - \log a}. \tag{25–21}$$

In the above problem the image source was the negative of the given source. If the boundary condition is such that the image source is of the same sign as the given source, then instead of (17) we have

$$W = -\frac{q}{2\pi\epsilon}\log[f(z) - f(z_0)][f(z) - f(z_0^*)]. \tag{25–22}$$

With a few appropriate modifications all the above formulae can be used for the magnetic field produced by an electric current filament in the presence of a perfectly conducting cylinder. The real part of the complex function

$$W = -\frac{I}{2\pi}\log(w - w_0) \tag{25–23}$$

is now taken to represent the stream function. This stream function satisfies the same boundary condition as the electric potential. Hence our final result will be (17) with q/ϵ replaced by I. The inductance per unit length of a thin wire of radius a is found in the same way as the capacitance; thus we have

$$L = \frac{\mu}{2\pi}[\log|f(z_0) - f(z_0^*)| - \log|f'(z_0)| - \log a]. \tag{25–24}$$

We shall now consider a few special problems. Figure 6–46 shows the cross-section of a wedge formed by two perfectly conducting planes. Let a line charge be at point z_0. The function

$$w = z^n = \rho^n e^{in\varphi} \tag{25–25}$$

is positive real for $\varphi = 0$ and negative real for $n\varphi = \pi$. Hence if we choose n so that

$$n\vartheta = \pi, \quad n = \frac{\pi}{\vartheta}, \tag{25–26}$$

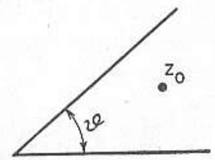


FIG. 6.46. The cross-section of a wedge and a charged filament.

then one boundary of the wedge will go into the positive real axis in the w-plane and the other into the negative axis. For the intermediate values of φ the imaginary part of w is positive and the points interior to the wedge are transformed into the upper half of the w-plane. Hence by (17) we have

$$W = -\frac{q}{2\pi\epsilon}\log\frac{z^n - z_0^n}{z^n - z_0^{*n}}. \tag{25–27}$$

The real part of this expression is

$$V = -\frac{q}{2\pi\epsilon}\log\frac{\rho^{2n} - 2\rho^n\rho_0^n\cos n(\varphi - \varphi_0) + \rho_0^{2n}}{\rho^{2n} - 2\rho^n\rho_0^n\cos n(\varphi + \varphi_0) + \rho_0^{2n}}. \tag{25–28}$$

If $\vartheta = \pi/2$, then $n = 2$ and (27) becomes

$$W = -\frac{q}{2\pi\epsilon}\log\frac{z^2 - z_0^2}{z^2 - z_0^{*2}} = -\frac{q}{2\pi\epsilon}\log\frac{(z - z_0)(z + z_0)}{(z - z_0^*)(z + z_0^*)}. \tag{25–29}$$

This shows that the effect of the wedge can be represented in this case by three image sources at points $-z_0$, z_0^* and $-z_0^*$; the first of these image sources is of the same sign as that at z_0 and the remaining sources are of the opposite sign.

If n is an integer, then

$$z^n - z_0^n = (z - z_0)(z - z_0 e^{2i\vartheta})(z - z_0 e^{4i\vartheta}) \cdots (z - z_0 e^{2(n-1)i\vartheta}), \qquad (25\text{–}30)$$

where ϑ is the wedge angle. Similarly we can factorize the denominator of (27). This factorization leads to $(2n - 1)$ image sources and V can be expressed in terms of the logarithms of the distances from a typical point to the source and its images.

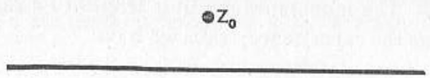


FIG. 6.47. A charged filament and a conducting half-plane.

When n is not an integer there is no system of image sources which could represent the effect of the wedge. For example, a half plane (Fig. 6.47) can be regarded as a wedge with $\vartheta = 2\pi$ and $n = \frac{1}{2}$. In this case (27) and (28) become

$$W = -\frac{q}{2\pi\epsilon} \log \frac{\sqrt{z} - \sqrt{z_0}}{\sqrt{z} - \sqrt{z_0^*}}$$

$$V = -\frac{q}{2\pi\epsilon} \log \frac{\rho - 2\sqrt{\rho\rho_0} \cos \frac{1}{2}(\varphi - \varphi_0) + \rho_0}{\rho - 2\sqrt{\rho\rho_0} \cos \frac{1}{2}(\varphi + \varphi_0) + \rho_0}, \qquad (25\text{–}31)$$

and no factorization is possible.

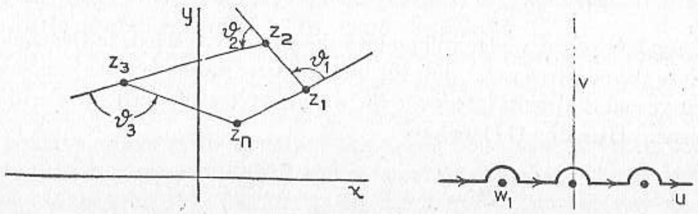


FIG. 6.48. Illustrating the transformation of a region enclosed by a polygon into the upper half-plane.

A general function transforming a polygon (Fig. 6.48) in the z-plane into the real axis of the w-plane was discovered by Schwarz. Let us set up the following integral

$$z = A \int^w (w - w_1)^{n_1}(w - w_2)^{n_2}(w - w_3)^{n_3} \cdots dw \qquad (25\text{–}32)$$

and examine the changes in z as we follow the real axis in the w-plane, indented at w_1, w_2, etc. The indentations may be taken as infinitely small semicircles in the upper half-plane and are needed to make the integrand an unambiguous function of w when the n's are fractions. As we follow the u-axis in the positive direction, the phase of the integrand remains unchanged so long as we are on the straight part of the path;

hence the increments of z are in phase and z must follow a straight line. Let us suppose that we are on the left of w_1. In the z-plane we are on some straight line $z_n z_1$ and are moving toward the point corresponding to w_1. As we go round the first infinitesimal indentation, all the factors of the integrand except $(w - w_1)^{n_1}$ are constant. The absolute value of $(w - w_1)^{n_1}\,dw$ is r^{n_1+1}, where r is the infinitesimal radius; hence if $n_1 + 1 > 0$, the magnitude of the increment in z is infinitely small during the motion round the semicircle. But while z remains unchanged, its new increments beyond w_1 will have a different phase. This is because the phase of $w - w_1$ has decreased by π and therefore the phase of $(w - w_1)^{n_1}$ has changed by $\vartheta_1 = -n_1\pi$. The new increments make an angle ϑ_1 with the old ones and we now follow a straight line $z_1 z_2$ making the angle ϑ_1 with $z_n z_1$. The second bending of the path takes place at z_2 corresponding to w_2, etc. The transformation (32) can now be expressed in the form

$$z = A \int^w (w - w_1)^{-\vartheta_1/\pi}(w - w_2)^{-\vartheta_2/\pi}(w - w_3)^{-\vartheta_3/\pi} \cdots dw, \qquad (25\text{–}33)$$

where ϑ_1, ϑ_2, etc. are the external angles of the polygon.

If from the point at infinity on the positive u-axis we follow round an infinite semicircle in the upper half of the w-plane, the path in the w-plane becomes closed. The total change in z around this path is zero since there are no singularities in the upper half-plane. Thus we shall return to the original value of z. The region enclosed by the polygon is transformed into the upper half-plane. The term "enclosed" is defined as follows: that region is enclosed by a curve which is on the left of an observer following the boundary counterclockwise.

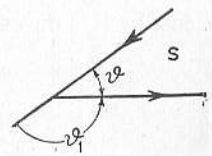


FIG. 6.49. A wedge of angle ϑ.

As our first example let us take a wedge (Fig. 6.49). In order to transform the region (S) into the upper half-plane we follow its boundary in the counterclockwise direction and imagine that the contour is completed with an arc of an infinitely large circle. Choosing the vertex of the wedge to correspond to the point $w = 0$ and seeing that the angle ϑ_1 is positive and equal to $\pi - \vartheta$, we have

$$z = A \int^w w^{-(\pi-\vartheta)/\pi}\,dw = \frac{A\pi}{\vartheta} w^{\vartheta/\pi} = w^{\vartheta/\pi}, \qquad (25\text{–}34)$$

where we have taken $A = \vartheta/\pi$. This agrees with (25).

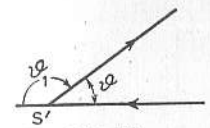


FIG. 6.50. The complementary wedge.

If we wish, we can transform the complementary region (S') into the upper half-plane; then we must follow the contour round this region in the counterclockwise direction (Fig. 6.50). In this case ϑ_1 is negative and its absolute value is $\pi - \vartheta$; hence

$$z = A \int w^{(\pi-\vartheta)/\pi}\,dw = \frac{A\pi}{2\pi - \vartheta} w^{(2\pi-\vartheta)/\pi} = w^{(2\pi-\vartheta)/\pi}. \qquad (25\text{–}35)$$

Since $2\pi - \vartheta$ is the angle of the wedge (S') in the same sense as ϑ is the angle of (S) the two results agree.

In our next example (Fig. 6.51) let the vertices A and B correspond to $w = -1$ and $w = +1$. Here $\vartheta_1 = \vartheta_2 = \pi/2$ and

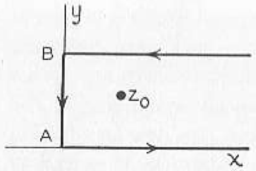


FIG. 6.51. The cross-section of a cylindrical surface with a charged filament inside.

$$z = A\int^w (w+1)^{-1/2}(w-1)^{-1/2}\,dw$$

$$= A\int^w \frac{dw}{\sqrt{w^2-1}} = A\cosh^{-1} w; \tag{25–36}$$

hence

$$w = \cosh\frac{z}{A}. \tag{25–37}$$

The constant A can be expressed in terms of the distance $AB = l$ between the planes. Separating the real and imaginary parts of (37) we have

$$w = \cosh\frac{x+iy}{A} = \cosh\frac{x}{A}\cos\frac{y}{A} + i\sinh\frac{x}{A}\sin\frac{y}{A}. \tag{25–38}$$

When $y = l$, w must be real; therefore $l/A = \pi$ and

$$w = \cosh\frac{\pi z}{l}. \tag{25–39}$$

Hence by (17) the complex potential of a line charge q at z_0 is

$$W = -\frac{q}{2\pi\epsilon}\log\frac{\cosh\dfrac{\pi z}{l} - \cosh\dfrac{\pi z_0}{l}}{\cosh\dfrac{\pi z}{l} - \cosh\dfrac{\pi z_0^*}{l}}. \tag{25–40}$$

This can be transformed into

$$W = -\frac{q}{2\pi\epsilon}\log\frac{\sinh\dfrac{\pi(z+z_0)}{2l}}{\sinh\dfrac{\pi(z+z_0^*)}{2l}} - \frac{q}{2\pi\epsilon}\log\frac{\sinh\dfrac{\pi(z-z_0)}{2l}}{\sinh\dfrac{\pi(z-z_0^*)}{2l}}. \tag{25–41}$$

As z_0 moves to the right away from the y-axis, the first term approaches zero and in the limit we have the potential of a line charge between two parallel planes

$$W = -\frac{q}{2\pi\epsilon}\left[\log\sinh\frac{\pi(z-z_0)}{2l} - \log\sinh\frac{\pi(z-z_0^*)}{2l}\right]. \tag{25–42}$$

In the arrangement shown in Fig. 6.52, we have

$$z = A\int^w (w+1)^{1/2}(w-1)^{1/2}\,dw = A\int^w \sqrt{w^2-1}\,dw. \tag{25–43}$$

Integrating, we obtain

$$z = \tfrac{1}{2}A[w\sqrt{w^2-1} - \cosh^{-1} w]. \tag{25–44}$$

Unfortunately in this case it is impossible to express w explicitly as a function of z.

As our last example, let us consider a line charge in a tube of rectangular cross-section (Fig. 6.53). We shall represent the corners of the rectangle by two pairs of points in the w-plane symmetric about the origin. The Schwarz integral becomes

$$z = A\int_0^w \frac{dw}{\sqrt{(1-w^2)(1-k^2w^2)}} = A\,\mathrm{sn}^{-1} w; \tag{25–45}$$

hence

$$w = \mathrm{sn}\frac{z}{A}, \tag{25–46}$$

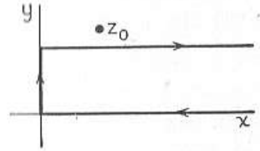


FIG. 6.52. The cross-section of a cylindrical surface with a charged filament outside.

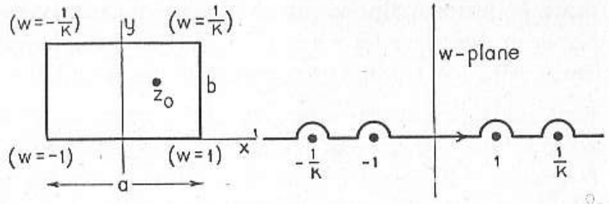


FIG. 6.53. A rectangular metal tube and a charged filament inside it.

where sn x is the elliptic sine of x. The constant A and the modulus k are determined by the dimensions of the rectangle. Thus if the complete elliptic integrals K and K' are defined by

$$K = \int_0^1 \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}},\quad K' = \int_0^1 \frac{dt}{\sqrt{(1-t^2)(1-k'^2t^2)}}, \tag{25–47}$$

$$k^2 + k'^2 = 1;$$

then

$$\frac{a}{2A} = K,\quad \frac{b}{A} = K',\quad \frac{K'}{K} = \frac{2b}{a}. \tag{25–48}$$

The last equation defines the modulus k; then A may be calculated from either of the first two equations. The transformation becomes

$$w = \mathrm{sn}\frac{2Kz}{a}; \tag{25–49}$$

hence the complex potential of the line charge is

$$W = -\frac{q}{2\pi\epsilon}\log\frac{\mathrm{sn}\dfrac{2Kz}{a} - \mathrm{sn}\dfrac{2Kz_0}{a}}{\mathrm{sn}\dfrac{2Kz}{a} - \mathrm{sn}\dfrac{2Kz_0^*}{a}}. \tag{25–50}$$

CHAPTER VII

Transmission Theory

7.0. *Introduction*

In section 6.11 we have shown that in source free regions the approximate equations connecting the harmonic transverse electromotive force V between two parallel wires (Fig. 7.1) and the longitudinal current I in the lower wire (or the magnetomotive force round the wire) are

$$\frac{dV}{dx} = -ZI, \quad \frac{dI}{dx} = -YV, \tag{0–1}$$

where the distributed series impedance Z and shunt admittance Y per unit length are complex constants

$$Z = R + i\omega L, \quad Y = G + i\omega C, \tag{0–2}$$

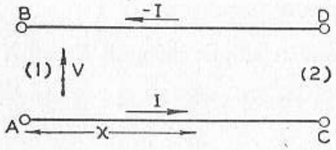


Fig. 7.1. A diagram representing a section of a transmission line.

and x is the distance along the line. The positive direction for V has been chosen from the lower wire to the upper and that for I in the direction of increasing x-coordinate. If one of these positive directions is reversed, the negative signs in equations (1) become positive. If the distance between the wires is variable, Z and Y are functions of x and the equations are general linear differential equations of the first order.

Equations (1) are not restricted to transmission lines alone but play an important role in the general theory of wave propagation. In the case of waves in three dimensions the field intensities E and H usually appear in place of V and I. This difference is superficial since E is the electromotive force and H the magnetomotive force per unit length and V and I are the integrated values of E and H. It happens that at low frequencies it is easier to measure V and I, while at very high frequencies E and H are more readily measured. In the case of waves in three dimensions the field intensities are generally functions of three coordinates; nevertheless under certain conditions wave propagation along, let us say, all x-lines is the same, and the remaining coordinates may be ignored in so far as wave transmission in the x-direction is concerned. Moreover the more general types of waves may frequently be decomposed into simpler types traveling in

accordance with equations (1). For expository convenience however we shall discuss these equations as applied to a pair of parallel wires.

7.1. *Impressed Forces and Currents*

Sources of energy may be of two types: (1) electric generators of zero impedance in series with the line, and (2) electric generators of infinite impedance in shunt with the line. The first type is represented by an impressed electromotive force $E(x)$ per unit length of the line and the second by an impressed transverse current $J(x)$, also per unit length of the line. The assumption that the internal impedances of the generators are respectively zero and infinite will not restrict the generality of our results since the actual internal impedances may be included in Z and Y. The transmission equations in regions with given source distributions may be obtained by the method used in section 6.11 for deriving (0–1). Let us assume that the impressed electromotive force $E(x)$ per unit length is acting in series with the lower wire* and let the positive directions of $E(x)$ and $J(x)$ be as shown in Fig. 7.2. By taking the electromotive force round a rectangle $ABCDA$ in which $AB = 1$, we obtain equation (6.11–1) in which the electromotive force of the field along AB is now given by $V_{AB} = Z_1I - E(x)$ and not by (6.11–2). Similarly the expression for the total transverse current per unit length is $I_t = (G + i\omega C)V + J(x)$ and not the one given by (6.11–6). Thus the transmission equations become

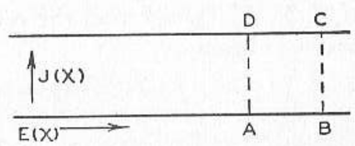


Fig. 7.2. The convention regarding the positive directions of the impressed series e.m.f. per unit length $E(x)$ and the shunt current per unit length $J(x)$.

$$\frac{dV}{dx} = -ZI + E(x), \quad \frac{dI}{dx} = -YV - J(x). \tag{1–1}$$

7.2. *Point Sources*

In practice the impressed sources are sometimes distributed over long sections of a transmission line and are sometimes highly concentrated in the vicinity of a point. An example of one type of distribution is furnished by a radio wave impinging on an open wire telephone line and an example of the other type is an ordinary generator connected to the line. In theory

* Strictly speaking the series impressed forces should be applied to both wires in a balanced "push-pull" manner, that is $\frac{1}{2}E(x)$ in series with the lower wire and $-\frac{1}{2}E(x)$ in series with the upper wire. Otherwise the longitudinal currents in the wires will not be equal and opposite (see section 6.6). Our assumption does not affect the results in so far as the balanced mode of propagation is concerned. The unbalanced mode will be considered in Chapter 8 in connection with waves on a single wire.

it is convenient to idealize concentrated distributions and regard them as point sources.

Let $E(x)$ be distributed in the interval $\xi - s/2 < x < \xi + s/2$ and let $J(x) = 0$. Integrating (1–1) in this interval, we have

$$V\left(\xi + \frac{s}{2}\right) - V\left(\xi - \frac{s}{2}\right) = -\int_{\xi-s/2}^{\xi+s/2} ZI\,dx + \int_{\xi-s/2}^{\xi+s/2} E(x)\,dx,$$
$$I\left(\xi + \frac{s}{2}\right) - I\left(\xi - \frac{s}{2}\right) = -\int_{\xi-s/2}^{\xi+s/2} YV\,dx. \tag{2–1}$$

Assume that as s approaches zero, the applied electromotive force approaches a finite limit

$$\lim \int_{\xi-s/2}^{\xi+s/2} E(x)\,dx = \hat{V}(\xi) \quad \text{as} \quad s \to 0. \tag{2–2}$$

If Z, Y, V, I are finite, the remaining integrals in (1) vanish in the limit and we obtain

$$V(\xi + 0) - V(\xi - 0) = \hat{V},$$
$$I(\xi + 0) - I(\xi - 0) = 0. \tag{2–3}$$

Thus the current is continuous at $x = \xi$ while the transverse voltage rises by an amount $\hat{V}$. Everywhere else V and I satisfy the homogeneous equations (0–1). These are the conditions for a point generator of zero impedance in series with the transmission line.

Similarly if $E(x) = 0$ and $J(x)$ is concentrated at $x = \xi$, we have

$$\int_{\xi-s/2}^{\xi+s/2} J(x)\,dx = \hat{I}(\xi) \tag{2–4}$$

and the conditions for a point generator of infinite impedance in shunt with the line at $x = \xi$ become

$$V(\xi + 0) - V(\xi - 0) = 0,$$
$$I(\xi + 0) - I(\xi - 0) = -\hat{I}. \tag{2–5}$$

In this case the voltage is continuous and the longitudinal current drops by an amount $\hat{I}$. Everywhere else V and I satisfy the homogeneous equations (0–1).

In the above equations $\hat{V}$ and $\hat{I}$ may be regarded as general discontinuities in the transverse voltage and longitudinal current and not merely as applied voltage and current. Thus if an impedance $\hat{Z}$ is inserted at $x = \xi$ in series with the line, the voltage drop across the impedance is $\hat{Z}I$ and the discontinuity $\hat{V}$ in the transverse voltage across the line is $\hat{V} = -\hat{Z}I$. Similarly if an admittance $\hat{Y}$ is inserted in shunt with the line, the transverse voltage is continuous and the discontinuity in the longitudinal current is $\hat{I} = \hat{Y}V$.

If the solutions of the transmission equations, subject to whatever supplementary conditions may be necessary, are known for point sources, then the general solutions of (1–1), subject to the same supplementary conditions, may be found by integration. We need only superpose the waves of elementary sources $E\,dx$ and $J\,dx$.

7.3. *The Energy Theorem*

The method of obtaining the energy equations for transmission lines is the same as in the general case of three dimensional electromagnetic fields. Starting with the fundamental equations (1–1), multiplying the first by I^* and the conjugate of the second by V, and adding, we obtain

$$I^* \frac{dV}{dx} + V \frac{dI^*}{dx} = -ZII^* - Y^*VV^* + EI^* - VJ^*.$$

The left side is the derivative of VI^*; hence multiplying each side by $\frac{1}{2}dx$, integrating from x_1 to x_2, and rearranging the terms, we have

$$\frac{1}{2}\int_{x_1}^{x_2} (EI^* - VJ^*)\,dx = \frac{1}{2}\int_{x_1}^{x_2} (ZII^* + Y^*VV^*)\,dx + \tfrac{1}{2}V(x_2)I^*(x_2) - \tfrac{1}{2}V(x_1)I^*(x_1). \tag{3–1}$$

The left-hand side represents the complex work done by the impressed forces and hence the power introduced into the transmission line; thus $\frac{1}{2}EI^*\,dx$ is the complex work done by an elementary applied electromotive force in driving the current I while $-\frac{1}{2}VJ^*\,dx$ is the work done by an elementary shunt generator which introduces the current $J\,dx$ against the transverse voltage V of the line. If we designate this total complex power by $\hat{\Psi}$ and replace Z and Y by their values from (0–2), we have

$$\hat{\Psi} = \frac{1}{2}\int_{x_1}^{x_2} RII^*\,dx + \frac{1}{2}\int_{x_1}^{x_2} GVV^*\,dx + i\omega\int_{x_1}^{x_2} (\tfrac{1}{2}LII^* - \tfrac{1}{2}CVV^*)\,dx + \tfrac{1}{2}V(x_2)I^*(x_2) - \tfrac{1}{2}V(x_1)I^*(x_1). \tag{3–2}$$

The real part of $\hat{\Psi}$ is the average power contributed to the line. The first two terms on the right represent the average power dissipated in the section (x_1,x_2). The difference between the power contributed to the section and that dissipated in it is represented by the real part of the last two terms. The power may be either entering or leaving the section at its ends; hence we may interpret the real part of $\frac{1}{2}V(x)I^*(x)$ as the average power passing across the point x of the transmission line in the direction of increasing x-coordinate. The imaginary term represents the fluctuating power.

If the sources of power are located at points $x < x_1$, then $\hat{\Psi} = 0$ and (2) becomes

$$\tfrac{1}{2}V(x_1)I^*(x_1) = \tfrac{1}{2}\int_{x_1}^{x_2} RII^*\,dx + \tfrac{1}{2}\int_{x_1}^{x_2} GVV^*\,dx + i\omega\int_{x_1}^{x_2} (\tfrac{1}{2}LII^* - \tfrac{1}{2}CVV^*)\,dx + \tfrac{1}{2}V(x_2)I^*(x_2). \tag{3-3}$$

In this case the complex power $\Psi = \frac{1}{2}V(x_1)I^*(x_1)$ is entering the line at $x = x_1$ and it is accounted for by the various terms on the right.

In the above interpretation of (1) we have implicitly assumed that the impressed intensity E acts on the total longitudinal current and that the impressed current J is acted upon by the total transverse voltage. This is necessarily true when the longitudinal current is localized as in conventional two conductor transmission lines. But if the longitudinal current is distributed, as in hollow metal tubes, for instance, then the left side of (1) will no longer represent the complex power contributed by the sources and the interpretation of other terms in the equation must be correspondingly modified. In these more general situations it is better to rely on the three-dimensional energy theorem (4.8–7) of which the present theorem is a special case, than to try to obtain appropriate modifications of the foregoing equations. The various energy terms will usually differ from those in this section by a constant factor.

7.4. *Fundamental Sets of Wave Functions for Uniform Lines*

Consider now a source-free section of a uniform transmission line. If either V or I is eliminated from (0–1), we obtain a second order linear differential equation with constant coefficients; thus

$$\frac{d^2I}{dx^2} = \Gamma^2 I, \quad \frac{d^2V}{dx^2} = \Gamma^2 V, \tag{4-1}$$

where the propagation constant Γ is defined by

$$\Gamma = \sqrt{ZY} = \sqrt{(R + i\omega L)(G + i\omega C)}. \tag{4-2}$$

The general solutions for V and I may therefore be expressed either as exponential or as hyperbolic functions.

Expressing I in terms of exponential functions, we have

$$I(x) = I^+e^{-\Gamma x} + I^-e^{\Gamma x}, \tag{4-3}$$

where I^+ and I^- are arbitrary constants. By (0–1) V is completely determined by I; thus

$$V(x) = -\frac{1}{Y}\frac{dI}{dx} = KI^+e^{-\Gamma x} - KI^-e^{\Gamma x}, \tag{4-4}$$

where the characteristic impedance K is defined by

$$K = \sqrt{\frac{Z}{Y}} = \sqrt{\frac{R + i\omega L}{G + i\omega C}} = \frac{\Gamma}{G + i\omega C} = \frac{R + i\omega L}{\Gamma}. \tag{4-5}$$

By definition Γ is a complex constant which lies in the first quadrant of the propagation constant plane or on its boundaries* and the equations

$$I^+(x) = I^+e^{-\Gamma x} = I^+e^{-\alpha x - i\beta x} = I^+e^{-\alpha x}e^{-i\beta x}, \quad V^+(x) = KI^+(x)$$

represent a progressive wave moving in the positive x-direction, with an amplitude which is attenuated at the rate of α nepers per meter. Likewise the equations

$$I^-(x) = I^-e^{\Gamma x} = I^-e^{\alpha x}e^{i\beta x}, \quad V^-(x) = -KI^-(x)$$

represent a progressive wave moving in the negative x-direction with an amplitude which is also attenuated at the rate of α nepers per meter. The constants I^+ and I^- are the amplitudes of these progressive waves at $x = 0$. They can be expressed in terms of the values of either the voltage or the current at the terminals (1) and (2) of the section of the line (Fig. 7.1). If the terminals C, D are at infinity and the generator is connected across A and B (or to the left of A, B), then $I^- = 0$ and only one progressive wave is present. In this case the impressed voltage in the direction from B to A is given by $V^+ = KI^+$. On the other hand if the terminals A, B are at infinity and the generator is connected across C and D, then $I^+ = 0$ and we have only the progressive wave moving leftward. For this case the voltage impressed from D to C is

$$V^-(l) = -KI^-(l) = -KI^-e^{\Gamma l}.$$

Expressing the wave functions in terms of the applied voltage we have

$$V^-(x) = V^-(l)e^{-\Gamma(l-x)}, \quad I^-(x) = -\frac{V^-(l)}{K}e^{-\Gamma(l-x)}.$$

More generally we can arbitrarily assign either the values of the voltages or of the currents at each pair of terminals, or the voltage at one pair of terminals and the current at the other pair, or the voltage (or the current) at one pair and the voltage/current ratio at the other. To start with let

* For non-negative values of R, G, L and C the two values of the complex constant $\pm\Gamma$ are respectively in the first and third quadrants. It is a matter of convention which value is designated as $+\Gamma$.

us consider some simple special cases and obtain a wave function for which the current vanishes at $x = 0$. This function follows from the general expression (3) if we let $I^- = -I^+$. For future convenience, however, let us first express the general solution in terms of hyperbolic functions

$$I(x) = A \cosh \Gamma x + B \sinh \Gamma x. \tag{4–6}$$

The corresponding expression for V is derived by differentiating as in (4); thus

$$V(x) = -KA \sinh \Gamma x - KB \cosh \Gamma x. \tag{4–7}$$

It is now immediately obvious that the wave for which I vanishes at $x = 0$ corresponds to $A = 0$. Expressing the remaining constant B in terms of the voltage at $x = 0$, we have

$$V(x) = V(0) \cosh \Gamma x, \quad I(x) = -\frac{V(0)}{K} \sinh \Gamma x. \tag{4–8}$$

This wave corresponds to the case in which the line is *electrically open* at $x = 0$. Similarly the wave for which the voltage vanishes at $x = 0$ is expressed by

$$I(x) = I(0) \cosh \Gamma x, \quad V(x) = -KI(0) \sinh \Gamma x. \tag{4–9}$$

This condition exists when the line is *short-circuited* at $x = 0$. The wave for which the voltage and the current assume given values at $x = 0$ is obtained by superposition of (8) and (9); thus

$$\begin{aligned} V(x) &= V(0) \cosh \Gamma x - KI(0) \sinh \Gamma x, \\ I(x) &= -\frac{V(0)}{K} \sinh \Gamma x + I(0) \cosh \Gamma x. \end{aligned} \tag{4–10}$$

The wave for which the current vanishes at $x = l$ is

$$V(x) = V(l) \cosh \Gamma(l - x), \quad I(x) = \frac{V(l)}{K} \sinh \Gamma(l - x). \tag{4–11}$$

These equations may be obtained from (8) by regarding $x = l$ as a new origin of the coordinate system and then substituting $x - l$ for the new x. Similarly the wave for which the voltage vanishes at $x = l$ is

$$I(x) = I(l) \cosh \Gamma(l - x), \quad V(x) = KI(l) \sinh \Gamma(l - x). \tag{4–12}$$

And finally the wave whose voltage and current assume given values at $x = l$ is obtained by superposition of (11) and (12); thus

$$\begin{aligned} V(x) &= V(l) \cosh \Gamma(l - x) + KI(l) \sinh \Gamma(l - x), \\ I(x) &= \frac{V(l)}{K} \sinh \Gamma(l - x) + I(l) \cosh \Gamma(l - x). \end{aligned} \tag{4–13}$$

For nondissipative lines we have

$$\Gamma = i\beta = i\omega\sqrt{LC}, \quad K = \sqrt{\frac{L}{C}}. \tag{4–14}$$

In this case the hyperbolic functions can be replaced by circular functions

$$\cosh \Gamma x = \cosh i\beta x = \cos \beta x,$$

$$\sinh \Gamma x = \sinh i\beta x = i \sin \beta x.$$

Waves represented by the corresponding nondissipative forms of (8), (9), (11) and (12) are standing or stationary waves. There is no change in phase along the line (except for a complete reversal each half wavelength); the voltage waves are in quadrature with the current waves; and consequently on the average there is no transfer of energy along the line. In the dissipative case there are no standing waves in the strict sense of the term; but for convenience we shall apply the term to all wave functions of the type (8), (9), (11) and (12).

7.5. *Characteristic Constants of Uniform Transmission Lines*

In most transmission problems the derived constants K and Γ are more important than the primary constants R, G, L and C. The latter can always be expressed in terms of the former, thus

$$R + i\omega L = K\Gamma, \quad G + i\omega C = \frac{\Gamma}{K}. \tag{5–1}$$

For nondissipative lines these equations become

$$\omega L = K\beta = \frac{2\pi K}{\lambda}, \quad \omega C = \frac{\beta}{K} = \frac{2\pi}{K\lambda}, \tag{5–2}$$

and the series reactance and the shunt susceptance per wavelength depend solely on the characteristic impedance

$$\omega L\lambda = 2\pi K, \quad \omega C\lambda = \frac{2\pi}{K}. \tag{5–3}$$

Since $\omega = 2\pi v/\lambda$, L and C can be expressed in terms of the characteristic impedance and velocity

$$L = \frac{K}{v}, \quad C = \frac{1}{Kv}. \tag{5–4}$$

In slightly dissipative lines we have approximately

$$\Gamma = i\omega\sqrt{LC}\left(1 - \frac{iR}{\omega L}\right)^{1/2}\left(1 - \frac{iG}{\omega C}\right)^{1/2}$$

$$= i\omega\sqrt{LC}\left[1 - \frac{i}{\omega}\left(\frac{R}{2L} + \frac{G}{2C}\right) + \frac{1}{8\omega^2}\left(\frac{R}{L} - \frac{G}{C}\right)^2\right].$$

Hence the attenuation and phase constants are given by

$$\alpha = \tfrac{1}{2}R\sqrt{\frac{C}{L}} + \tfrac{1}{2}G\sqrt{\frac{L}{C}}, \quad \beta = \omega\sqrt{LC}\left[1 + \frac{1}{8\omega^2}\left(\frac{R}{L} - \frac{G}{C}\right)^2\right]. \tag{5-5}$$

The first order effect of dissipation is to cause the amplitude of progressive waves to decrease with the distance from the source; the second order effect is to diminish the wave velocity

$$v = \frac{\omega}{\beta} = \frac{1}{\sqrt{LC}}\left[1 - \frac{1}{8\omega^2}\left(\frac{R}{L} - \frac{G}{C}\right)^2\right]. \tag{5-6}$$

These approximations are applicable when the resistance per unit length is small compared with the reactance and when the conductance per unit length is small compared with the susceptance.

The effect of dissipation on the characteristic impedance is to introduce a small reactance; thus we have

$$K = \sqrt{\frac{L}{C}} - i\sqrt{\frac{L}{C}}\left(\frac{R}{2\omega L} - \frac{G}{2\omega C}\right). \tag{5-7}$$

In transmission lines which are used for efficient transmission of energy and for communication purposes the reactive component of K is very small and may usually be neglected. Over long distances the attenuation may be large even though the attenuation constant is small. Even the small term of the second order in the expression for the wave velocity has to be retained when distortion of signals is being considered. In practical applications very small correction terms may be important under some conditions and much larger terms unimportant under other conditions.

Using for K the approximation given by the first term in (7), we have

$$\alpha = \frac{R}{2K} + \tfrac{1}{2}GK. \tag{5-8}$$

The attenuation constant may be expressed in terms of the power W carried by a progressive wave across a given cross section of the line and the power $\hat{W}$ per unit length which is dissipated in the line; thus

$$\alpha = \frac{RII^*}{2KII^*} + \frac{GK^2II^*}{2KII^*} = \frac{RII^* + GVV^*}{2KII^*}. \tag{5-9}$$

Since $\frac{1}{2}KII^*$ is the power carried by a progressive wave, $\frac{1}{2}RII^*$ the power dissipated in the series resistance and $\frac{1}{2}GVV^*$ the power dissipated in the shunt conductance, we have

$$\alpha = \frac{\hat{W}}{2W}. \tag{5-10}$$

This formula is convenient for the calculation of α when R, G, L and C are not explicitly given. For this reason we shall prove it from a more general viewpoint. If the power W carried by a wave is dissipated uniformly, then

$$\frac{dW}{W} = -k\,dx, \tag{5-11}$$

where k is a constant and x is the distance in the direction of the wave. If the change in W is due to dissipation of power in heat, then

$$-\frac{dW}{dx} = \hat{W} \tag{5-12}$$

is the power dissipated per unit length and consequently

$$k = \frac{\hat{W}}{W}. \tag{5-13}$$

This is the attenuation constant for power. From (11) we have

$$W = W_0 e^{-kx}. \tag{5-14}$$

If W is proportional to the square of some quantity such as voltage, or current, or one of the field intensities, the attenuation constant for the latter is $\frac{1}{2}k$ and we obtain equation (10).

7.6. *The Input Impedance*

Let us assume that a section of a transmission line of length l (Fig. 7.1) is terminated in an impedance $Z(l)$ at C, D. Since

$$Z(l) = \frac{V(l)}{I(l)}, \tag{6-1}$$

the "input impedance" at A, B is obtained from (4–13)

$$Z(0) = \frac{V(0)}{I(0)} = K\frac{Z(l)\cosh\Gamma l + K\sinh\Gamma l}{K\cosh\Gamma l + Z(l)\sinh\Gamma l}. \tag{6-2}$$

If the line is short-circuited at $x = l$, then

$$Z(0) = K\tanh\Gamma l; \tag{6-3}$$

and if the line is open, then

$$Z(0) = K \coth \Gamma l. \tag{6–4}$$

If the "output impedance" $Z(l)$ is equal to the characteristic impedance then

$$Z(0) = K \tag{6–5}$$

and the input impedance is equal to the characteristic impedance for all values of l. Thus we expect the voltage-current distribution in the section of the line to be the same as for an infinitely long line which is indeed the case.

At times it is more convenient to express various relations in terms of admittances. These forms are analogous to those involving impedances. Thus for the input admittance we have

$$Y(0) = M\frac{Y(l)\cosh \Gamma l + M \sinh \Gamma l}{M \cosh \Gamma l + Y(l)\sinh \Gamma l}, \tag{6–6}$$

where the *characteristic admittance* M is the reciprocal of K.

For nondissipative lines the input impedance is

$$Z(0) = K\frac{Z(l)\cos \beta l + iK \sin \beta l}{K \cos \beta l + iZ(l)\sin \beta l}. \tag{6–7}$$

When $l = n\lambda/2$, where n is an integer, the input impedance equals the output impedance

$$Z(0) = Z\left(\frac{n\lambda}{2}\right). \tag{6–8}$$

If $l = n\lambda \pm \lambda/4$, then

$$Z(0) = \frac{K^2}{Z\left(n\lambda \pm \frac{\lambda}{4}\right)}. \tag{6–9}$$

In this case if the input impedance is small compared with K then the output impedance is large and vice versa. Either impedance is said to be the *inverse* of the other with respect to K. The ratio $Z(0)/K$ depends only on the ratio $Z(l)/K$ and on βl.

There exists a very simple and useful geometric interpretation of the behavior of the input impedance of a nondissipative line when the output impedance is fixed and the length of the line is varied. If the numerator and the denominator of (7) are divided by $\cos \beta l$, we have

$$Z(0) = K\frac{Z(l) + iK \tan \beta l}{K + iZ(l)\tan \beta l}.$$

In a complex plane $Z(0)$ is a bilinear function of another complex variable w, the latter assuming only the real values $w = \tan \beta l$; hence (section 2.1) the input impedance $Z(0)$ must describe a circle. The center of this *impedance circle* is on the real axis. Thus if any point Z is on the impedance circle, its inverse K^2/Z with respect to the circle of radius K is also on the circle. This follows from equation (9) which relates the input impedance of a section of length l and the input impedance $Z(\lambda/4)$ of a section of length $l - \lambda/4$. Since K is real, the phase of the inverse impedance is equal to the phase of the original impedance and opposite in sign. This would be impossible if the center of the impedance circle were off the real axis as there would then be some points on one side of the axis whose phases would be larger than the phases of all points on the other side.

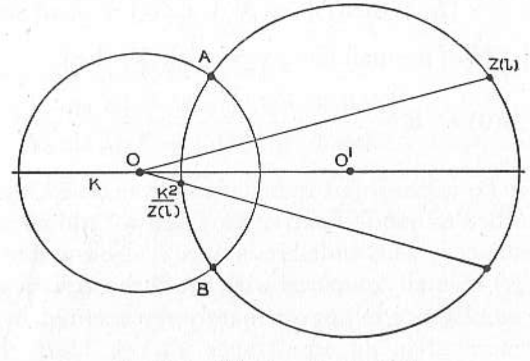


FIG. 7.3. The impedance circle.

Since the impedance circle passes through points $Z(l)$ and its inverse $K^2/Z(l)$, we can draw a perpendicular through the midpoint of the line joining these points and thus determine the center. The impedance circle (Fig. 7.3) is useful for visualizing certain properties of the input impedance as a function of the length of the line. Thus the impedance circle is unaltered if the output impedance is replaced by its inverse. The impedance circle becomes the imaginary axis if $Z(l)$ is a pure reactance. If $Z(l)$ is a pure resistance, then one of the two values $Z(l)$ and $K^2/Z(l)$ is the maximum input impedance and the other is the minimum. If $|Z(l)|$ is small compared with K, then $R(l)$ is nearly the minimum of $Z(0)$ regardless of $X(l)$ and $K^2/R(l)$ is nearly the maximum. The maximum reactance may be either positive or negative and its value is represented by the radius of the circle. For large impedance circles, the maximum resistance (and the maximum impedance) is nearly $2X_{max}$ and the minimum resistance (as well as the minimum impedance) is nearly $K^2/2X_{max}$.

The position of the center of the impedance circle can be obtained by the methods of complex variables or analytic geometry; thus the abscissa x of the center is

$$x_c = \frac{K^2 + |Z(l)|^2}{2R(l)}.$$

The radius r of the impedance circle is obtained at once from the triangle OAO'; thus $r = \sqrt{x_c^2 - K^2}$.

For dissipative lines the impedance circle becomes a spiral converging to the point K. If the attenuation constant or the length l or both are so small that αl is small compared with unity, then the input impedance can be expressed in the following approximate form

$$Z(0) = K\frac{[Z(l) + K\alpha l] \cos \beta l + i[K + \alpha l Z(l)] \sin \beta l}{[K + \alpha l Z(l)] \cos \beta l + i[Z(l) + K\alpha l] \sin \beta l}.$$

Furthermore, if $Z(l)\alpha l$ is small compared with K, then

$$Z(0) = K\frac{[Z(l) + K\alpha l] \cos \beta l + iK \sin \beta l}{K \cos \beta l + i[Z(l) + K\alpha l] \sin \beta l}.$$

In this case, in so far as the input impedance is concerned, the section of the line may be regarded as nondissipative provided we add a resistance $K\alpha l$ to the output impedance. This added resistance is also equal to $\frac{1}{2}Rl + \frac{1}{2}K^2Gl$. Similarly if $\alpha l Y(l)$ is small compared with the characteristic admittance M, then the input admittance is approximately represented by a nondissipative section terminated in an admittance $Y(l) + M\alpha l$; the added conductance is equal to $\frac{1}{2}Gl + \frac{1}{2}M^2Rl$.

The voltage and current distribution may be expressed in terms of the output impedance and either the input voltage or the input current; thus from (1) and (4–13) we have

$$V(x) = \frac{Z(l) \cosh \Gamma(l - x) + K \sinh \Gamma(l - x)}{Z(l) \cosh \Gamma l + K \sinh \Gamma l} V(0),$$

$$I(x) = \frac{K \cosh \Gamma(l - x) + Z(l) \sinh \Gamma(l - x)}{K \cosh \Gamma l + Z(l) \sinh \Gamma l} I(0). \tag{6–10}$$

If the line is short-circuited, then

$$V(x) = \frac{\sinh \Gamma(l - x)}{\sinh \Gamma l} V(0), \quad I(x) = \frac{\cosh \Gamma(l - x)}{\cosh \Gamma l} I(0); \tag{6–11}$$

and if the line is open

$$V(x) = \frac{\cosh \Gamma(l - x)}{\cosh \Gamma l} V(0), \quad I(x) = \frac{\sinh \Gamma(l - x)}{\sinh \Gamma l} I(0). \tag{6–12}$$

7.7. *Transmission Lines as Transducers*

A section of a uniform line is a symmetric transducer. By (5.2–3) its self-impedances are equal to the open-circuit impedance of the line and the transfer impedance may be found from (6–12); thus

$$Z_{11} = Z_{22} = K \coth \Gamma l, \quad Z_{12} = -K \operatorname{csch} \Gamma l.$$

If the transmission line is represented by a T-network (see Fig. 5.12), then by (5.4–1) the impedances of the shunt arm and of each series arm are respectively

$$Z_2 = K \operatorname{csch} \Gamma l, \quad \tfrac{1}{2}Z_1 = K \tanh \frac{\Gamma l}{2}.$$

Regardless of the length l of the section we have

$$Z_{11}Z_{22} - Z_{12}^2 = K^2.$$

7.8. *Waves Produced by Point Sources*

Let a section of the line of length l be terminated in impedances Z_1 and Z_2 (Fig. 7.4). Let $V_1(x,\xi)$ and $I_1(x,\xi)$ be the voltage and the current at point $x = x$ when a unit electromotive force is impressed at $x = \xi$ in series with the line.* To the left and to the right of the generator we have respectively

$$\begin{aligned} I_1(x,\xi) &= P_1 \cosh \Gamma x + Q_1 \sinh \Gamma x, \\ V_1(x,\xi) &= -K[P_1 \sinh \Gamma x + Q_1 \cosh \Gamma x], \quad x < \xi; \\ I_1(x,\xi) &= P_2 \cosh \Gamma(l - x) + Q_2 \sinh \Gamma(l - x), \\ V_1(x,\xi) &= K[P_2 \sinh \Gamma(l - x) + Q_2 \cosh \Gamma(l - x)], \quad x > \xi. \end{aligned}$$

At $x = 0$, the voltage-current ratio is $-Z_1$ and therefore $P_1 = PK$, $Q_1 = PZ_1$, where P is a disposable constant. Similarly at $x = l$ the voltage-current ratio is Z_2 and therefore $P_2 = QK$, $Q_2 = QZ_2$. At $x = \xi$ we have

$$I_1(\xi + 0,\xi) - I_1(\xi - 0,\xi) = 0,$$
$$V_1(\xi + 0,\xi) - V_1(\xi - 0,\xi) = 1. \tag{8–1}$$

Making the necessary substitutions we obtain

$$Q[K \cosh \Gamma(l - \xi) + Z_2 \sinh \Gamma(l - \xi)] = P[K \cosh \Gamma\xi + Z_1 \sinh \Gamma\xi],$$

$$Q[K \sinh \Gamma(l - \xi) + Z_2 \cosh \Gamma(l - \xi)] + P[K \sinh \Gamma\xi + Z_1 \cosh \Gamma\xi] = \frac{1}{K}.$$

The first of these equations is satisfied if we let

$$DQ = K \cosh \Gamma\xi + Z_1 \sinh \Gamma\xi, \quad DP = K \cosh \Gamma(l - \xi) + Z_2 \sinh \Gamma(l - \xi).$$

* The coordinate x is the distance from Z_1.

Substituting in the second, we have

$$D = K[K \cosh \Gamma\xi + Z_1 \sinh \Gamma\xi][K \sinh \Gamma(l-\xi) + Z_2 \cosh \Gamma(l-\xi)] + K[K \sinh \Gamma\xi + Z_1 \cosh \Gamma\xi][K \cosh \Gamma(l-\xi) + Z_2 \sinh \Gamma(l-\xi)].$$

Multiplying and collecting terms, we have

$$D = K[(K^2 + Z_1 Z_2) \sinh \Gamma l + K(Z_2 + Z_1) \cosh \Gamma l].$$

Thus all disposable constants are determined and we have

$$\begin{aligned} DI_1(x,\xi) &= [K \cosh \Gamma x + Z_1 \sinh \Gamma x] \times \\ &\quad [K \cosh \Gamma(l-\xi) + Z_2 \sinh \Gamma(l-\xi)], \quad x < \xi, \\ &= [K \cosh \Gamma\xi + Z_1 \sinh \Gamma\xi] \times \\ &\quad [K \cosh \Gamma(l-x) + Z_2 \sinh \Gamma(l-x)], \quad x > \xi, \end{aligned} \tag{8-2}$$

$$\begin{aligned} DV_1(x,\xi) &= -K[K \sinh \Gamma x + Z_1 \cosh \Gamma x] \times \\ &\quad [K \cosh \Gamma(l-\xi) + Z_2 \sinh \Gamma(l-\xi)], \quad x < \xi, \\ &= K[K \cosh \Gamma\xi + Z_1 \sinh \Gamma\xi] \times \\ &\quad [K \sinh \Gamma(l-x) + Z_2 \cosh \Gamma(l-x)], \quad x > \xi. \end{aligned}$$

It is easy to see that $I_1(x,\xi)$ is symmetric

$$I_1(x,\xi) = I_1(\xi,x).$$

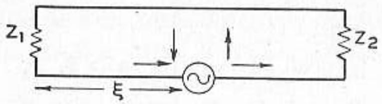


FIG. 7.4. A section of a line energized by a series generator of zero impedance.

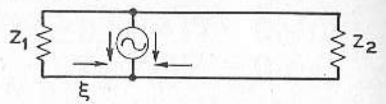


FIG. 7.5. A section of a line energized by a shunt generator of infinite impedance.

This proves the reciprocity theorem under the conditions stated at the beginning of the section. The arrows in Fig. 7.4 show that the current at the generator flows in the direction of the impressed electromotive force on both sides while the transverse voltages are in opposite directions on the two sides.

If a unit current is impressed at $x = \xi$ in shunt with the line (Fig. 7.5), then the voltage and current waves $V_2(x,\xi)$ and $I_2(x,\xi)$ satisfy the following conditions at $x = \xi$

$$\begin{aligned} V_2(\xi + 0,\xi) - V_2(\xi - 0,\xi) &= 0, \\ I_2(\xi + 0,\xi) - I_2(\xi - 0,\xi) &= -1. \end{aligned} \tag{8-3}$$

In this case we obtain the following solutions

$$\begin{aligned} DI_2(x,\xi) &= K[K \cosh \Gamma x + Z_1 \sinh \Gamma x] \times \\ &\quad [K \sinh \Gamma(l-\xi) + Z_2 \cosh \Gamma(l-\xi)], \quad x < \xi, \\ &= -K[K \sinh \Gamma\xi + Z_1 \cosh \Gamma\xi] \times \\ &\quad [K \cosh \Gamma(l-x) + Z_2 \sinh \Gamma(l-x)], \quad x > \xi, \end{aligned} \tag{8-4}$$

$$\begin{aligned} DV_2(x,\xi) &= -K^2[K \sinh \Gamma x + Z_1 \cosh \Gamma x] \times \\ &\quad [K \sinh \Gamma(l-\xi) + Z_2 \cosh \Gamma(l-\xi)], \quad x < \xi, \\ &= -K^2[K \sinh \Gamma\xi + Z_1 \cosh \Gamma\xi] \times \\ &\quad [K \sinh \Gamma(l-x) + Z_2 \cosh \Gamma(l-x)], \quad x > \xi. \end{aligned}$$

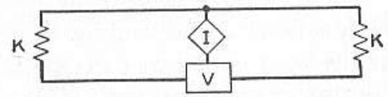


FIG. 7.6. A series voltage and a shunt current applied at the same point of the line.

The voltage wave function is now symmetric

$$V_2(x,\xi) = V_2(\xi,x).$$

The arrows in Fig. 7.5 show that at the generator the voltage is in the same direction on both sides while the currents are in opposite directions.

In the special case when $Z_1 = Z_2 = K$, we have

$$\begin{aligned} V_1(x,\xi) &= -\tfrac{1}{2}e^{-\Gamma(\xi-x)}, & I_1(x,\xi) &= \frac{1}{2K}e^{-\Gamma(\xi-x)}, \quad x < \xi; \\ &= \tfrac{1}{2}e^{-\Gamma(x-\xi)}, & &= \frac{1}{2K}e^{-\Gamma(x-\xi)}, \quad x > \xi \end{aligned} \tag{8-5}$$

for a unit electromotive force impressed in series with the line. Similarly for a unit current impressed in shunt with the line, we obtain

$$\begin{aligned} V_2(x,\xi) &= -\frac{K}{2}e^{-\Gamma(\xi-x)}, & I_2(x,\xi) &= \tfrac{1}{2}e^{-\Gamma(\xi-x)}, \quad x < \xi; \\ &= -\frac{K}{2}e^{-\Gamma(x-\xi)}, & &= -\tfrac{1}{2}e^{-\Gamma(x-\xi)}, \quad x > \xi. \end{aligned} \tag{8-6}$$

If a current $-I$ is impressed in shunt with the line and a voltage $V = KI$ in series at the same point $x = \xi$ (Fig. 7.6) then from (5) and (6) we find that the wave to the left of $x = \xi$ vanishes and that to the right becomes

$$V(x,\xi) = KIe^{-\Gamma(x-\xi)}, \quad I(x,\xi) = Ie^{-\Gamma(x-\xi)}.$$

This conclusion could easily be reached from considerations of symmetry and of the relative directions of the voltages and currents at the generators as illustrated by the arrows in Figs. 7.4 and 7.5. Once we have come to the conclusion that, for given values of applied current and voltage, there is no wave to the left of $x = \xi$, it becomes evident that the impedance K at the left end of the line could be replaced by any other impedance and the left section could be completely removed.

In practice these conditions can be realized only approximately since they demand generators of zero impedance (or " constant voltage generators ") and generators of infinite impedance (" constant current generators ").

7.9. *Waves Produced by Arbitrary Distributions of Sources*

Knowing the wave functions corresponding to point sources we can immediately construct the wave functions corresponding to any given distribution of sources by proper superposition. Thus if a series electromotive force $E(x)$ per unit length and a shunt current $J(x)$ per unit length are distributed in the interval (x_1,x_2) then we have the solutions of (1–1) in the following form

$$V(x) = \int_{x_1}^{x_2} E(\xi)V_1(x,\xi)\,d\xi + \int_{x_1}^{x_2} J(\xi)V_2(x,\xi)\,d\xi,$$

$$I(x) = \int_{x_1}^{x_2} E(\xi)I_1(x,\xi)\,d\xi + \int_{x_1}^{x_2} J(\xi)I_2(x,\xi)\,d\xi. \tag{9–1}$$

That these functions satisfy (1–1) can be proved by direct substitution. It should be recalled, however, that $V_1(x,\xi)$ and $I_2(x,\xi)$ are discontinuous and hence nondifferentiable functions of x at $x = \xi$. For this reason we break up the integrands as follows

$$V(x) = \int_{x_1}^{x-0} E(\xi)V_1(x,\xi)\,d\xi + \int_{x+0}^{x_2} E(\xi)V_1(x,\xi)\,d\xi + \int_{x_1}^{x_2} J(\xi)V_2(x,\xi)\,d\xi,$$

$$I(x) = \int_{x_1}^{x_2} E(\xi)I_1(x,\xi)\,d\xi + \int_{x_1}^{x-0} J(\xi)I_2(x,\xi)\,d\xi + \int_{x+0}^{x_2} J(\xi)I_2(x,\xi)\,d\xi. \tag{9–2}$$

Each integral is now a differentiable function. In taking the derivatives of $V(x)$ and $I(x)$ we use

$$\frac{d}{dx}\int_{x_1}^{x} f(x,\xi)\,d\xi = \int_{x_1}^{x} \frac{d}{dx}f(x,\xi)\,d\xi + f(x,x),$$

$$\frac{d}{dx}\int_{x}^{x_2} f(x,\xi)\,d\xi = \int_{x}^{x_2} \frac{d}{dx}f(x,\xi)\,d\xi - f(x,x).$$

Thus we have

$$\frac{dV}{dx} = \int_{x_1}^{x-0} E(\xi)\frac{d}{dx}V_1(x,\xi)\,d\xi + \int_{x+0}^{x_2} E(\xi)\frac{d}{dx}V_1(x,\xi)\,d\xi$$
$$+ E(x)V_1(x,x-0) - E(x)V_1(x,x+0) + \int_{x_1}^{x_2} J(\xi)\frac{d}{dx}V_2(x,\xi)\,d\xi.$$

Since V_1 and V_2 are solutions of (0–1) we may substitute $-ZI_1$ and $-ZI_2$ for the derivatives under the integral signs; by (2) the sum of the integrals is then equal to* $-ZI(x)$. The remaining two terms are equal to $E(x)$ in virtue of (8–1). Thus we have proved that the first equation in the set (1–1) is satisfied. Similarly we can show that the second equation is satisfied. Finally it can be verified that the boundary conditions are fulfilled.

7.10. *Nonuniform Transmission Lines*

Let us now assume that Z and Y are functions of x. Eliminating first I and then V, we find that in source-free regions both variables satisfy general homogeneous linear differential equations of the second order

$$\frac{d^2V}{dx^2} - \frac{Z'}{Z}\frac{dV}{dx} - YZV = 0, \quad \frac{d^2I}{dx^2} - \frac{Y'}{Y}\frac{dI}{dx} - YZI = 0. \tag{10–1}$$

A second order differential equation of this type possesses two linearly independent solutions and its general solution is a linear function of these solutions. Thus we have

$$I(x) = AI^+(x) + BI^-(x),$$

where A and B are two disposable constants. The corresponding solution for V is then

$$V(x) = AV^+(x) + BV^-(x),$$

where V^+ and V^- are obtained from

$$V^+(x) = -\frac{1}{Y(x)}\frac{dI^+(x)}{dx}, \quad V^-(x) = -\frac{1}{Y(x)}\frac{dI^-(x)}{dx}. \tag{10–2}$$

Alternatively a pair of fundamental voltage wave functions might be selected and the corresponding current wave functions then defined as follows

$$I^+(x) = -\frac{1}{Z(x)}\frac{dV^+(x)}{dx}, \quad I^-(x) = -\frac{1}{Z(x)}\frac{dV^-(x)}{dx}. \tag{10–3}$$

* This is true even if Z is a function of x.

A wave impedance may be associated with each fundamental pair of wave functions; thus

$$K^+(x) = \frac{V^+(x)}{I^+(x)} = -\frac{1}{YI^+}\frac{dI^+}{dx} = -\frac{1}{Y}\frac{d}{dx}\log I^+ = -\frac{ZV^+}{\frac{d}{dx}V^+} = -\frac{Z}{\frac{d}{dx}\log V^+}, \tag{10-4}$$

$$K^-(x) = -\frac{V^-(x)}{I^-(x)} = \frac{1}{YI^-}\frac{dI^-}{dx} = \frac{1}{Y}\frac{d}{dx}\log I^- = \frac{ZV^-}{\frac{d}{dx}V^-} = \frac{Z}{\frac{d}{dx}\log V^-}. \tag{10-5}$$

In the strict sense of the term there are no progressive waves in nonuniform transmission lines since any local nonuniformity in an otherwise uniform line will generate a reflected wave. However, in some instances wave functions may exist which bear considerable resemblance to the exponential wave functions and hence may be said to represent "progressive" waves in nonuniform lines. This is apt to happen when Z and Y are slowly varying functions of x. Even then it may be more convenient to select other sets of wave functions for the fundamental set. Thus in general we should look upon $K^+(x)$ and $K^-(x)$ as factors to be used in passing from a given current wave to the corresponding voltage wave and vice versa.

Other ratios are useful in the general theory. Thus the *voltage transfer ratios* are defined by

$$\chi_V^+(x_1,x_2) = \frac{V^+(x_2)}{V^+(x_1)}, \quad \chi_V^-(x_1,x_2) = \frac{V^-(x_2)}{V^-(x_1)}. \tag{10-6}$$

Similarly the *current transfer ratios* are defined by

$$\chi_I^+(x_1,x_2) = \frac{I^+(x_2)}{I^+(x_1)}, \quad \chi_I^-(x_1,x_2) = \frac{I^-(x_2)}{I^-(x_1)}. \tag{10-7}$$

Consider now a section of a nonuniform line extending from $x = x_1$ to $x = x_2$. If the output impedance at $x = x_2$ is $Z(x_2)$, then it is easy to show that the input impedance is

$$Z(x_1) = \frac{\begin{vmatrix} V^+(x_1) & I^+(x_2) \\ V^-(x_1) & I^-(x_2) \end{vmatrix} Z(x_2) - \begin{vmatrix} V^+(x_1) & V^+(x_2) \\ V^-(x_1) & V^-(x_2) \end{vmatrix}}{\begin{vmatrix} I^+(x_1) & I^+(x_2) \\ I^-(x_1) & I^-(x_2) \end{vmatrix} Z(x_2) - \begin{vmatrix} I^+(x_1) & V^+(x_2) \\ I^-(x_1) & V^-(x_2) \end{vmatrix}}. \tag{10-8}$$

In order to obtain the impedance of the section at $x = x_2$ when an impedance $Z(x_1)$ is across the line at $x = x_1$, we merely interchange x_1 and x_2 in

(8) and reverse the signs of $Z(x_1)$ and $Z(x_2)$. The reversal of the sign corresponds to the change in the direction of the impedance.

7.11. *Calculation of Nonuniform Wave Functions by Successive Approximations*

Consider a section of a nonuniform line of length l and let

$$Z = Z_0 + \hat{Z}, \quad Y = Y_0 + \hat{Y}, \tag{11-1}$$

where Z_0 and Y_0 are constants. These constants may be taken to represent the average values of Z and Y in the interval $(0,l)$

$$Z_0 = \frac{1}{l}\int_0^l Z(x)\,dx, \quad Y_0 = \frac{1}{l}\int_0^l Y(x)\,dx. \tag{11-2}$$

Assuming that there are no sources in the chosen interval, the transmission equations are

$$\frac{dV}{dx} = -Z_0I - \hat{Z}I, \quad \frac{dI}{dx} = -Y_0V - \hat{Y}V. \tag{11-3}$$

We now seek that solution of these equations for which the initial values of the voltage and current are given

$$V(0) = V_0, \quad I(0) = I_0. \tag{11-4}$$

For this purpose we first find the solution of

$$\frac{dV}{dx} = -Z_0I, \quad \frac{dI}{dx} = -Y_0V, \tag{11-5}$$

which has the following discontinuities in V and I at $x = \xi$

$$V(\xi + 0) = V(\xi - 0) - \hat{V}, \quad I(\xi + 0) = I(\xi - 0) - \hat{I}. \tag{11-6}$$

In the interval $(0,\xi)$ we evidently have (see equation 4–10)

$$V(x) = V_0(x) = V_0 \cosh \Gamma_0 x - K_0 I_0 \sinh \Gamma_0 x,$$
$$I(x) = I_0(x) = -\frac{V_0}{K_0}\sinh \Gamma_0 x + I_0 \cosh \Gamma_0 x, \tag{11-7}$$

where

$$\Gamma_0 = \sqrt{Z_0 Y_0}, \quad K_0 = \sqrt{\frac{Z_0}{Y_0}}. \tag{11-8}$$

At $x = \xi$, V and I are decreased by $\hat{V}$ and $\hat{I}$ and in order to obtain the solution in the interval (ξ,l) we need only add to equations (7) analogous expressions in which the initial values are $-\hat{V}$ and $-\hat{I}$; thus

$$V(x) = V_0(x) - \hat{V} \cosh \Gamma_0(x - \xi) + K_0\hat{I} \sinh \Gamma_0(x - \xi),$$
$$I(x) = I_0(x) + \frac{\hat{V}}{K_0} \sinh \Gamma_0(x - \xi) - \hat{I} \cosh \Gamma_0(x - \xi). \tag{11–9}$$

We now consider a " continuous distribution of discontinuities "

$$\hat{V} = \hat{Z}(\xi)I(\xi)\, d\xi, \quad \hat{I} = \hat{Y}(\xi)V(\xi)\, d\xi, \tag{11–10}$$

and construct the following solution of (3)

$$V(x) = V_0(x) - \int_0^x \hat{Z}(\xi)I(\xi) \cosh \Gamma_0(x-\xi)\, d\xi + K_0 \int_0^x \hat{Y}(\xi)V(\xi) \sinh \Gamma_0(x-\xi)\, d\xi,$$
$$I(x) = I_0(x) + \frac{1}{K_0}\int_0^x \hat{Z}(\xi)I(\xi) \sinh \Gamma_0(x-\xi)\, d\xi - \int_0^x \hat{Y}(\xi)V(\xi) \cosh \Gamma_0(x-\xi)\, d\xi. \tag{11–11}$$

We have not really solved the original differential equations since the unknown functions appear under the integral signs; but we have converted the differential equations into integral equations.

That these integral equations define functions satisfying the differential equations (3) and the initial conditions can be proved directly. At $x = 0$ the integrals vanish and $V(0)$, $I(0)$ evidently reduce to V_0, I_0. Differentiating $V(x)$ we have

$$\frac{dV(x)}{dx} = \frac{dV_0(x)}{dx} - \Gamma_0 \int_0^x \hat{Z}(\xi)I(\xi) \sinh \Gamma_0(x - \xi)\, d\xi - \hat{Z}(x)I(x)$$
$$+ Z_0 \int_0^x \hat{Y}(\xi)V(\xi) \cosh \Gamma_0(x - \xi)\, d\xi.$$

The right-hand side of this equation is identical with the right-hand side of (3) if we take into account the expression (11) for $I(x)$ and the following equation

$$\frac{dV_0(x)}{dx} = -Z_0I_0(x).$$

Similarly it can be shown that the second equation of the set (3) is satisfied.

From the integral equations (11) $V(x)$ and $I(x)$ can be calculated by successive approximations. Thus we set

$$V(x) = V_0(x) + V_1(x) + V_2(x) + \cdots$$
$$I(x) = I_0(x) + I_1(x) + I_2(x) + \cdots, \tag{11–12}$$

where

$$V_{n+1}(x) = -\int_0^x \hat{Z}(\xi)I_n(\xi) \cosh \Gamma_0(x - \xi)\, d\xi$$
$$+ K_0 \int_0^x \hat{Y}(\xi)V_n(\xi) \sinh \Gamma_0(x - \xi)\, d\xi, \tag{11–13}$$
$$I_{n+1}(x) = \frac{1}{K_0}\int_0^x \hat{Z}(\xi)I_n(\xi) \sinh \Gamma_0(x - \xi)\, d\xi$$
$$- \int_0^x \hat{Y}(\xi)V_n(\xi) \cosh \Gamma_0(x - \xi)\, d\xi.$$

Evidently the differential equations (3) can be transformed into other integral equations in which Z_0 and Y_0 are not constants provided we can obtain solutions of the corresponding equations (4) and (5). Just as equations (11) are most useful when the transmission line is only slightly nonuniform, other integral equations may be particularly useful when a given nonuniform line deviates slightly from another nonuniform line with known wave functions.

It should be noted that the solutions (12) are valid even if $\hat{Z}(x)$ and $\hat{Y}(x)$ are discontinuous functions.

7.12. *Slightly Nonuniform Transmission Lines*

When Z and Y are nearly equal to their average values Z_0 and Y_0, so that the relative deviations $\hat{Z}/Z_0$ and $\hat{Y}/Y_0$ are small, only the first corrections $V_1(x)$ and $I_1(x)$, or at most the second corrections V_2 and I_2, need be considered. When the deviations are large in a given section of the line, the section may be subdivided into smaller sections. Taking $n = 0$ in (11–13), substituting from (11–7), and rearranging the terms, we obtain the first corrections

$$V_1(x) = V_0[B(x) \cosh \Gamma_0x - A(x) \sinh \Gamma_0x + C(x) \sinh \Gamma_0x]$$
$$-K_0I_0[A(x) \cosh \Gamma_0x - B(x) \sinh \Gamma_0x + C(x) \cosh \Gamma_0x], \tag{12–1}$$
$$I_1(x) = -\frac{V_0}{K_0}[B(x) \sinh \Gamma_0x - A(x) \cosh \Gamma_0x + C(x) \cosh \Gamma_0x]$$
$$+ I_0[A(x) \sinh \Gamma_0x - B(x) \cosh \Gamma_0x + C(x) \sinh \Gamma_0x],$$

where

$$A(x) = \tfrac{1}{2}\int_0^x \left(\frac{\hat{Z}}{K_0} - K_0\hat{Y}\right) \cosh 2\Gamma_0\xi\, d\xi, \quad C(x) = \tfrac{1}{2}\int_0^x \left(\frac{\hat{Z}}{K_0} + K_0\hat{Y}\right) d\xi,$$
$$B(x) = \tfrac{1}{2}\int_0^x \left(\frac{\hat{Z}}{K_0} - K_0\hat{Y}\right) \sinh 2\Gamma_0\xi\, d\xi. \tag{12–2}$$

In some nonuniform transmission lines the product ZY is constant; then we choose

$$\Gamma_0 = \sqrt{ZY}, \quad Z_0 = Z_{av}, \quad Y_0 = \frac{\Gamma_0^2}{Z_0}. \tag{12-3}$$

We now have as the first approximation

$$\frac{\hat{Y}}{Y_0} = -\frac{\hat{Z}}{Z_0}, \quad -K_0\hat{Y} = \frac{\hat{Z}}{K_0}. \tag{12-4}$$

Consequently $C(x) = 0$ and

$$K_0A(x) = \int_0^x \hat{Z}(\xi) \cosh 2\Gamma_0\xi \, d\xi, \quad K_0B(x) = \int_0^x \hat{Z}(\xi) \sinh 2\Gamma_0\xi \, d\xi. \tag{12-5}$$

7.13. *Reflection in Uniform Lines*

Consider a semi-infinite uniform transmission line terminated in an impedance Z and let a progressive wave arrive from infinity. If $Z = K$, the voltage associated with this incident wave is exactly equal to the voltage across the terminal impedance if the entire incident current should flow through this impedance. The terminal impedance "absorbs" the incident wave completely and causes no disturbance in the line.

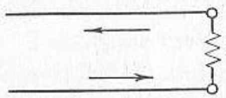


FIG. 7.7. Reflection in a transmission line at a point of discontinuity.

On the other hand if $Z \neq K$, the absorption cannot be complete and a reflected wave is initiated at the terminals. Let V^i, I^i be the incident voltage and current at the terminals of Z; let the reflected voltage and current be V^r, I^r; and the total or "transmitted" values be V^t, I^t. Since the voltage and current must be continuous at Z, we have

$$V^i + V^r = V^t, \quad I^i + I^r = I^t. \tag{13-1}$$

The reflected wave travels back to infinity and hence is also a progressive wave. Thus we have the following conditions

$$V^i = KI^i, \quad V^r = -KI^r, \quad V^t = ZI^t. \tag{13-2}$$

Substituting from (2) in (1) and solving we obtain the following expressions for the *reflection coefficients*

$$q_I = \frac{I^r}{I^i} = \frac{K - Z}{K + Z} = \frac{1 - k}{1 + k}, \quad k = \frac{Z}{K},$$
$$q_V = \frac{V^r}{V^i} = \frac{Z - K}{Z + K} = \frac{k - 1}{k + 1}, \quad q_V = -q_I, \tag{13-3}$$

and the corresponding expressions for the *transmission coefficients*

$$p_I = \frac{I^t}{I^i} = \frac{2K}{K + Z} = \frac{2}{1 + k} = 1 + q_I,$$
$$p_V = \frac{V^t}{V^i} = \frac{2Z}{K + Z} = \frac{2k}{1 + k} = 1 + q_V. \tag{13-4}$$

Thus the reflection and transmission coefficients depend on the ratio of the terminal impedance to the characteristic impedance. If this ratio is unity, the impedances are said to be "matched," and there is no reflection. If k equals either zero or infinity, the reflection is complete; in the former case the current at the terminals is doubled and the voltage is annihilated while in the latter the current is annihilated and the voltage doubled. For all impedance ratios the voltage reflection coefficient is the negative of the current reflection coefficient.

The voltage reflection and transmission coefficients have exactly the same form as the corresponding current coefficients if expressed in terms of admittances; thus

$$q_V = \frac{M - Y}{M + Y}, \quad p_V = \frac{2M}{M + Y}. \tag{13-5}$$

The expressions for the incident and reflected waves may then be written in the following form

$$V^i(x) = V^i e^{-\Gamma x}, \quad I^i(x) = I^i e^{-\Gamma x};$$
$$V^r(x) = q_V V^i e^{\Gamma x}, \quad I^r(x) = q_I I^i e^{\Gamma x} \tag{13-6}$$

assuming that Z is located at the origin. If Z is a semi-infinite transmission line whose characteristic impedance and propagation constant are K_1, Γ_1, then for the transmitted wave we have

$$V^t(x) = p_V V^i e^{-\Gamma_1 x}, \quad I^t(x) = p_I I^i e^{-\Gamma_1 x}. \tag{13-7}$$

The reflection coefficient depends on the ratio $k = K_1/K$ of the characteristic impedances and is independent of the propagation constants of the two lines.

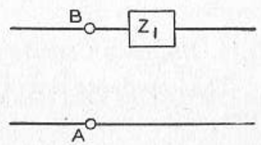


FIG. 7.8. An impedance in series with a line.

Let us consider a special case of reflection caused by an impedance Z_1 inserted in series with the line (Fig. 7.8). The impedance Z seen to the right from the terminals A, B is $Z = Z_1 + K$. Substituting in (3) and

(4), we have

$$q_V = -q_I = \frac{Z_1}{2K+Z_1}, \quad p_I = \frac{2K}{2K+Z_1}. \qquad (13\text{–}8)$$

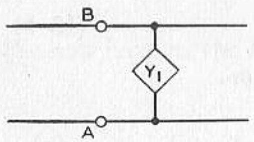


FIG. 7.9. An admittance in shunt with a line.

Taking the reciprocals, we obtain

$$\frac{2K}{Z_1} = \frac{1}{q_V} - 1 = -\frac{1}{q_I} - 1, \quad \frac{Z_1}{2K} = \frac{1}{p_I} - 1. \qquad (13\text{–}9)$$

Thus the ratio Z_1/K can be expressed quite simply in terms of the reflection and transmission coefficients which in certain circumstances can be measured more readily than the ratio itself.

If an admittance Y_1 is inserted in parallel with the line (Fig. 7.9) we have

$$Y = Y_1 + M, \quad q_I = -q_V = \frac{Y_1}{2M+Y_1}, \quad p_V = \frac{2M}{2M+Y_1}, \qquad (13\text{–}10)$$

$$\frac{2M}{Y_1} = \frac{1}{q_I} - 1 = -\frac{1}{q_V} - 1, \quad \frac{Y_1}{2M} = \frac{1}{p_V} - 1.$$

Let us now consider reflection and transmission of power. The transmitted power W^t is

$$W^t = \tfrac{1}{2}\mathrm{re}(V^t I^{t*}) = \tfrac{1}{2}\mathrm{re}(p_V p_I^* V^i I^{i*}).$$

For an incident progressive wave in a nondissipative line V^i is in phase with I^i; and therefore for the incident power we have

$$W^i = \tfrac{1}{2}\mathrm{re}(V^i I^{i*}) = \tfrac{1}{2} V^i I^{i*}.$$

The power transmission and reflection coefficients are then

$$p_W = \frac{W^t}{W^i} = \mathrm{re}(p_V p_I^*) = \frac{4\,\mathrm{re}(k)}{(1+k)(1+k^*)},$$

$$q_W = \frac{W^r}{W^i} = 1 - p_W = |q_I|^2 = |q_V|^2.$$

7.14. *Reflection Coefficients as Functions of the Impedance Ratio*

The impedance ratio k and the voltage reflection coefficient q_V are complex quantities

$$k = R + iX = Ae^{i\varphi}, \quad q_V = ae^{i\vartheta},$$

where the amplitudes A and a are essentially positive. The phase of k lies in the interval $(-\pi/2, \pi/2)$ while the phase of q_V is in $(-\pi, \pi)$. The reflection coefficient is the ratio of two complex quantities represented by the lines AP and BP (Fig. 7.10) joining points $A(1)$ and $B(-1)$ to $P(k)$. The phase ϑ of q_V and the phase ϑ_I of q_I are the angles formed by these lines as indicated in Fig. 7.10; the amplitude a is the ratio of the lengths AP and BP.

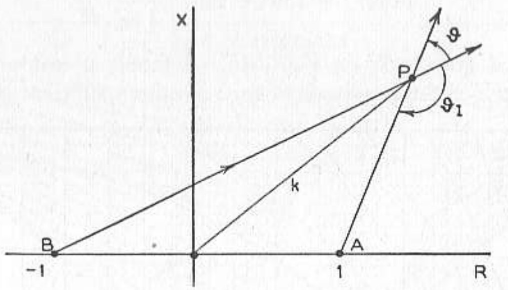


FIG. 7.10. The complex plane for representing the impedance ratio.

Taking the square of the amplitude of the reflection coefficient, we obtain

$$\frac{(R-1)^2 + X^2}{(R+1)^2 + X^2} = a^2, \quad \text{or} \quad \left(R - \frac{1+a^2}{1-a^2}\right)^2 + X^2 = \frac{4a^2}{(1-a^2)^2}. \qquad (14\text{–}1)$$

This equation represents a family of circles surrounding A and B as illustrated in Fig. 1.6. If $X = 0$, then

$$R_1 = \frac{1-a}{1+a}, \quad R_2 = \frac{1+a}{1-a}$$

are the real values of the impedance ratio for which the reflection coefficient is $\mp a$.

The unit circle represents points for which the absolute value of the impedance ratio is unity. On this circle the phase of the reflection coefficient is $\pm 90°$. The points of intersection of this circle and (1) are

$$R = \frac{1-a^2}{1+a^2}, \quad X = \pm\frac{2a}{1+a^2}.$$

The loci of points for which the phase of the reflection coefficient is constant are circles passing through A and B. The equation of this family of circles is

$$\tan^{-1}\frac{X}{R-1} - \tan^{-1}\frac{X}{R+1} = \vartheta, \quad \text{or} \quad R^2 + (X - \cot\vartheta)^2 = \csc^2\vartheta.$$

In making a chart (Fig. 7.11) showing the dependence of the reflection coefficient on the impedance ratio we could limit ourselves to one quadrant inside the unit circle. The absolute values of the reflection coefficient are equal for k, $1/k$, k^* and $1/k^*$. The phase of q for the impedance ratio $1/k$ is different from that for k by 180° and the phase for k^* is the negative of that for k.

The amplitude of the reflection coefficient as a function of A and φ is

$$\mathrm{am}(q) = \sqrt{\frac{1 - 2A\cos\varphi + A^2}{1 + 2A\cos\varphi + A^2}}.$$

Figure 7.12 shows the variation of am(q) with A and φ. The phase (in degrees) of the voltage reflection coefficient is

$$\mathrm{ph}(qv) = 180° - \tan^{-1}\frac{2A\sin\varphi}{1 - A^2}$$

and is represented graphically in Fig. 7.13. A family of contour lines connecting points (A,φ) for which the amplitude of the reflection coefficient is constant is shown

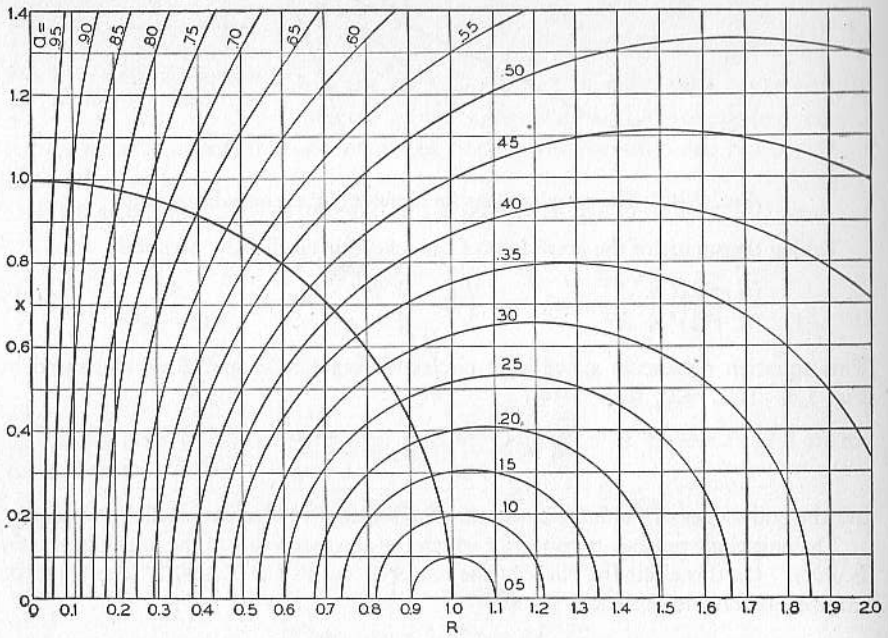


FIG. 7.11. A reflection chart. The characteristic impedance is unity, $R + iX$ is the terminal impedance, and a is the amplitude of the reflection coefficient.

in Fig. 7.14; the curves are symmetric about the axis $\varphi = 0$. Another family in Fig. 7.15 represents contour lines along which ϑ is constant. The equations of these two families are respectively

$$\cos\varphi = \tfrac{1}{2}\left(A + \frac{1}{A}\right)\frac{1 - a^2}{1 + a^2}, \quad \sin\varphi = \tfrac{1}{2}\left(A - \frac{1}{A}\right)\tan\vartheta.$$

If k is small, then we have approximately

$$qv = -1 + 2k - 2k^2 + \cdots;$$

and if k is large, then

$$qv = 1 - \frac{2}{k} + \frac{2}{k^2} - \cdots.$$

If the line is nondissipative, the amplitude of the incident progressive wave is independent of the position x along the line. When a reflected wave is superposed on it, the amplitude will have maxima and minima whose values depend on the amplitude a

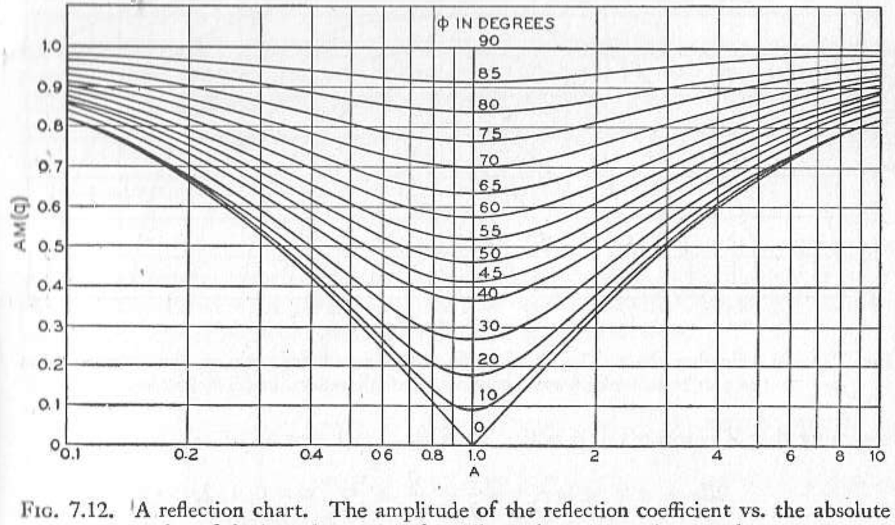


FIG. 7.12. A reflection chart. The amplitude of the reflection coefficient vs. the absolute value of the impedance ratio for different impedance phase angles.

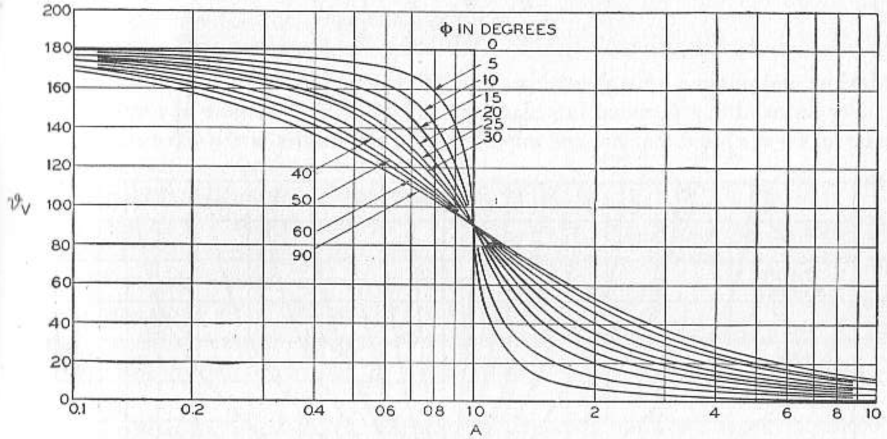


FIG. 7.13. A reflection chart. The phase of the voltage reflection coefficient vs. the absolute value of the impedance ratio for different impedance phase angles.

and whose positions are determined by the phase ϑ of the reflection coefficient. Thus the amplitude B of the combined wave is

$$B = |\, e^{-i\beta x} + ae^{i(\beta x + \vartheta)} \,| = \sqrt{1 + 2a\cos(2\beta x + \vartheta) + a^2},$$

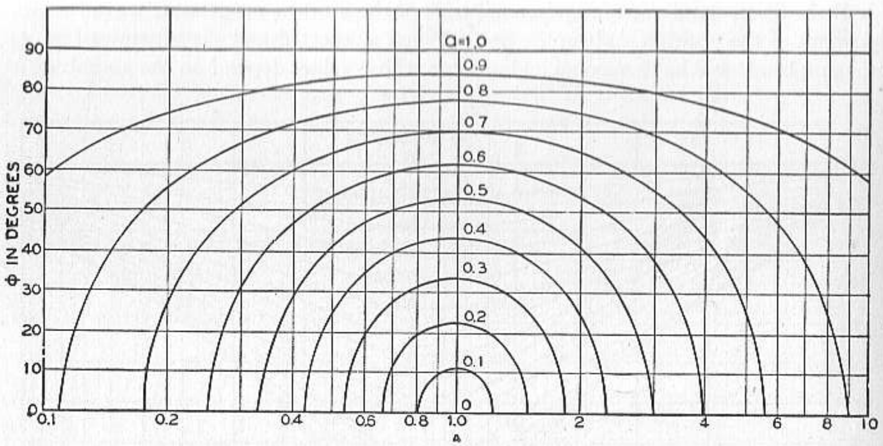


FIG. 7.14. A reflection chart. The phase angle of the impedance ratio vs. the absolute value of this ratio for different amplitudes of the reflection coefficient.

assuming that Z is located at $x = 0$. Then $B_{\max} = 1 + a$ at

$$2\beta x + \vartheta = \pm 2n\pi, \quad -\frac{x}{\lambda} = \frac{\vartheta}{4\pi} \pm \frac{n}{2}, \quad n = 0, 1, 2, \cdots.$$

Similarly, $B_{\min} = 1 - a$ at

$$-\frac{x}{\lambda} = \frac{\vartheta}{4\pi} \pm \frac{n}{2} \pm \frac{1}{4}.$$

Maxima and minima are a quarter wavelength apart.

For an arbitrary terminal impedance the phase ϑ may assume any value in the interval $(-\pi,\pi)$ and maxima and minima of the amplitudes of the wave can occur at

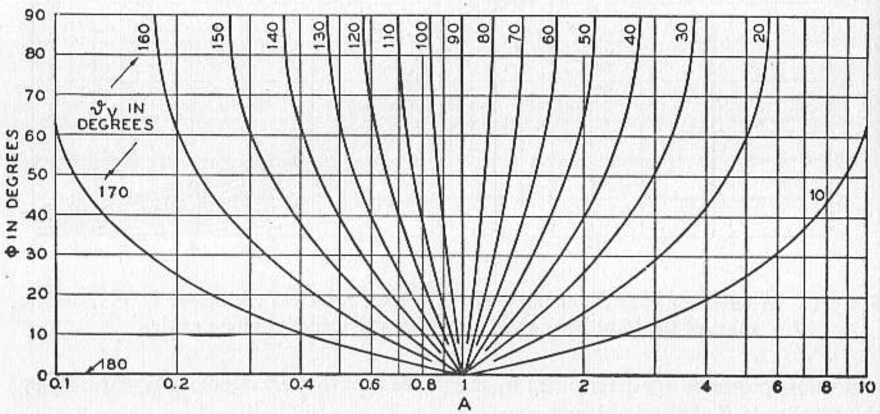


FIG. 7.15. A reflection chart The phase angle of the impedance ratio vs. the absolute value of this ratio for different phases of the voltage reflection coefficient

any distance from the discontinuity. But if an impedance is inserted in series with the line as shown in Fig. 7.8, then k is on some perpendicular which crosses the R-axis to the right of $R = 1$; hence, $-\pi/2 \leq \vartheta \leq \pi/2$. Positive and negative values of ϑ correspond respectively to positive and negative reactance components of the inserted impedance. Thus for positive reactances the maxima occur within the first eighth of a wavelength from the discontinuity and then at half-wavelength intervals; for negative reactances the first maximum will be between $\frac{3}{8}\lambda$ and $\frac{1}{2}\lambda$ from the discontinuity. The first minimum will occur in the remainder $(\frac{1}{8}\lambda,\frac{3}{8}\lambda)$ of the first half-wavelength. The current maxima and minima coincide respectively with voltage minima and maxima.

If an admittance is inserted in parallel with the line as shown in Fig. 7.9, then it is the first current maximum that will occur either in the interval $(0, \frac{1}{8}\lambda)$ or in $(\frac{3}{8}\lambda,\frac{1}{2}\lambda)$ and the first minimum in $(\frac{1}{8}\lambda,\frac{3}{8}\lambda)$. This result follows at once from the admittance diagram for reflection coefficients; or, if we use the impedance diagram, we may show that the impedance of a parallel combination of the characteristic impedance with another impedance must lie within the right half of the unit circle.

7.15. *Induction and Equivalence Theorems for Transmission Lines*

At an impedance discontinuity Z (Fig. 7.7) equations (13–1) may be written as follows

$$V^t - V^r = V^i, \quad I^t - I^r = I^i. \tag{15–1}$$

The impedance Z may, of course, be a section of a transmission line with a different characteristic impedance (Fig. 7.16). Comparing (1) to (2–3) and (2–5), we find that the wave comprised of the reflected wave to the left of the discontinuity A, B and of the transmitted wave to the right of it (or in Z) could be produced by a combination of a generator of zero impedance, acting at point A, in series with the line, and a generator of infinite impedance, acting in shunt with the line across A, B. The electromotive force impressed by the series generator should be $\hat{V} = V^i$ and the current impressed by the shunt generator $\hat{I} = -I^i$. The relations between the directions of the voltage and current in the incident wave (at A, B) and the voltage and current impressed by these generators are shown by the arrows in Fig. 7.16.

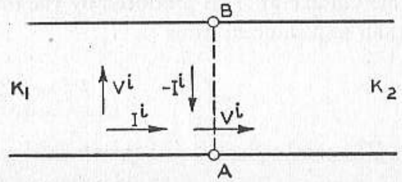


FIG. 7.16. Illustrating the induction and equivalence theorems for transmission lines.

If K_2 is zero, the impressed current is diverted entirely to K_2 and hence the reflected current in this case is produced solely by the series generator. The reflected wave is the wave which would be generated by a voltage equal to the incident voltage, acting in series with the line at its short-circuited end. If K_2 is infinite, the series generator has no effect on the reflected wave which in this case could be produced by a fixed current generator placed across the open end. These conditions are realized approximately if K_2 is either small or large compared with K_1.

If $K_2 = K_1 = K$, then $\frac{1}{2}I^i$ flows in K_2 and $\frac{1}{2}I^i$ in K_1. The impressed voltage

$V^i = KI^i$ acts on the impedance $2K$ and produces a current $\frac{1}{2}I^i$ through A. The total current on the left of A is canceled and that on the right becomes I^i. The two generators send a wave to the right of A, B and none to the left.

More generally the current produced at A by the series generator is the same on both sides

$$I_1^r = I_1^t = \frac{V^i}{K_1 + K_2}. \tag{15–2}$$

The corresponding line voltages are then

$$V_1^r = -K_1 I_1^r = -\frac{K_1}{K_1 + K_2} V^i, \quad V_1^t = K_2 I_1^t = \frac{K_2}{K_1 + K_2} V^i.$$

The impressed current due to the shunt generator is divided in the inverse ratio of the impedances K_1 and K_2; thus

$$I_2^t = \frac{K_1}{K_1 + K_2} I^i, \quad I_2^r = -\frac{K_2}{K_1 + K_2} I^i.$$

The corresponding voltages are

$$V_2^t = K_2 I_2^t = \frac{K_1 K_2}{K_1 + K_2} I^i = \frac{K_2}{K_1 + K_2} V^i,$$

$$V_2^r = -K_1 I_2^r = \frac{K_1 K_2}{K_1 + K_2} I^i = \frac{K_2}{K_1 + K_2} V^i.$$

The analogy between these equations and (2) becomes clear if, when calculating the line voltage at A, B produced by the impressed current I^i, we use admittances rather than impedances; thus

$$V_2^r = V_2^t = \frac{I^i}{M_1 + M_2}.$$

The total reflected and transmitted waves are then given in the following form

$$V^r = V_1^r + V_2^r, \quad I^r = I_1^r + I_2^r,$$
$$V^t = V_1^t + V_2^t, \quad I^t = I_1^t + I_2^t.$$

7.16. *Conditions for Maximum Delivery of Power to an Impedance*

Let a generator with an internal impedance Z_1 be connected to a load impedance Z_2 (Fig. 7.17). In general we may have $Z_1 = R_1 + iX_1$, $Z_2 = R_2 + iX_2$. The amplitude of the current in the circuit is

$$|I| = \frac{|V|}{|Z_1 + Z_2|} = \frac{|V|}{\sqrt{(R_1 + R_2)^2 + (X_1 + X_2)^2}}.$$

If W_1 is the power absorbed by Z_1 and W_2 that absorbed by Z_2, then

$$W_1 = \frac{R_1 VV^*}{2[(R_1+R_2)^2+(X_1+X_2)^2]}, \quad W_2 = \frac{R_2 VV^*}{2[(R_1+R_2)^2+(X_1+X_2)^2]}.$$

The condition for maximum absorption of power by Z_2 is $dW_2 = 0$. If Z_1 is fixed, this requires $R_2 = R_1$, $X_2 = -X_1$. Thus Z_2 is the conjugate of Z_1

$$Z_2 = Z_1^*. \tag{16–1}$$

If Z_1 and Z_2 are pure resistances, then the condition for maximum delivery of power from the generator becomes simply the equality or "matching" of the impedances.

When the conditions for maximum delivery of power are fulfilled, then

$$W_1 = W_2 = \frac{VV^*}{8R}, \tag{16–2}$$

where R is the load resistance.

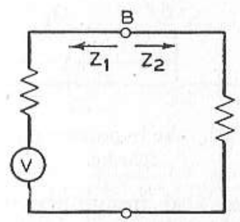


FIG. 7.17. Illustrating the transfer of power from a generator to a load impedance.

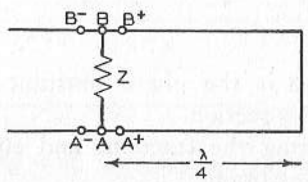


FIG. 7.18. A quarter wave section effectively opens the line at A,B.

7.17. *Transformation and Matching of Impedances*

At high frequencies it is usually impossible to open a transmission line electrically simply by discontinuing it mechanically; thus it may be impossible to terminate the line in its characteristic impedance by inserting a proper impedance at the physical end of the line. On the other hand it is usually possible to short-circuit the line. If now an impedance Z is inserted in shunt with the line a quarter wavelength from the short-circuited end, the line will see the impedance Z from the terminals A^-, B^- just to the left of the terminals A, B (Fig. 7.18). The impedance looking to the right of the terminals A^+, B^+ is infinite and thus the line is electrically open at these terminals. If $Z = K$, the line will be terminated in its characteristic impedance and there will be no reflection.

Let us now suppose that we wish to absorb completely the power carried by a line with an impedance K_1 in a resistance $R \neq K_1$. If a quarter wavelength section of a line with an impedance K_2 is inserted between R and the given line (Fig. 7.19), then the impedance looking to the right of

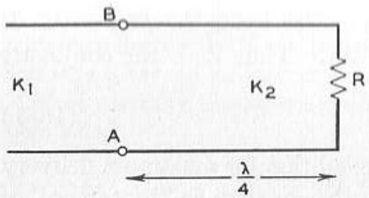


FIG. 7.19. A quarter wave section as an impedance transformer.

A, B is K_2^2/R and we shall have a match if

$$K_1 = \frac{K_2^2}{R}, \quad \text{or} \quad K_2 = \sqrt{K_1 R}. \quad (17\text{–}1)$$

Since the length of the matching line depends on the wavelength, a perfect match can be obtained only at one frequency.

Another method of matching the line consists of shunting the line with a proper susceptance iB at a proper distance from the given output conductance (Fig. 7.20). The condition for matching is

$$M_1 = iB + M_2 \frac{G \cos \beta l + iM_2 \sin \beta l}{M_2 \cos \beta l + iG \sin \beta l}, \quad (17\text{–}2)$$

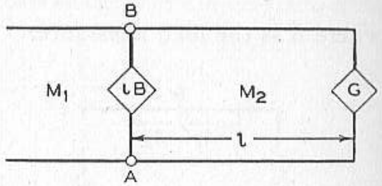


FIG. 7.20. An impedance matching scheme.

where β is the phase constant of the matching section.

Clearing the fractions and equating the real and imaginary parts to zero, we obtain

$$M_2(G - M_1) \cos \beta l = BG \sin \beta l, \quad (M_1 G - M_2^2) \sin \beta l = BM_2 \cos \beta l.$$

Multiplying and dividing these equations term by term, we have

$$(G - M_1)(GM_1 - M_2^2) = B^2 G, \quad \frac{M_2(G - M_1)}{GM_1 - M_2^2} = \frac{G}{M_2} \tan^2 \beta l.$$

From this we obtain

$$\frac{B}{M_2} = \pm \sqrt{\left(1 - \frac{M_1}{G}\right)\left(\frac{GM_1}{M_2^2} - 1\right)}, \quad \tan \beta l = \frac{M_2}{B}\left(1 - \frac{M_1}{G}\right). \quad (17\text{–}3)$$

Thus we have explicit expressions for B and l when the line is matched to G. In a special case M_2 may be equal to M_1.

If G is large compared with M_1 and M_2, then we have approximately

$$\frac{B}{M_2} = \pm \frac{\sqrt{GM_1}}{M_2}, \quad \tan \beta l = \frac{M_2}{B}, \quad \frac{l}{\lambda} = \frac{1}{2} \pm \frac{M_2}{2\pi B},$$

and the matching section is nearly equal to a half wavelength.

Another simple matching circuit is shown in Fig. 7.21; in this case the condition for matching is

$$K_1 = iX + K_2 \frac{R \cos \beta l + iK_2 \sin \beta l}{K_2 \cos \beta l + iR \sin \beta l}.$$

Comparing this with (2), we observe that one equation goes into the other if the admittances are replaced by the corresponding impedances. The solution corresponding to (3) is therefore

$$\frac{X}{K_2} = \pm \sqrt{\left(\frac{K_1}{R} - 1\right)\left(1 - \frac{RK_1}{K_2^2}\right)}, \quad \tan \beta l = \frac{K_2}{X}\left(1 - \frac{K_1}{R}\right).$$

If R is small, then we have approximately

$$\frac{X}{K_2} = \pm \sqrt{\frac{K_1}{R}}, \quad \tan \beta l = \mp \sqrt{\frac{K_1}{R}}, \quad \frac{l}{\lambda} = \frac{1}{4} \pm \frac{1}{2\pi} \sqrt{\frac{R}{K_1}}.$$

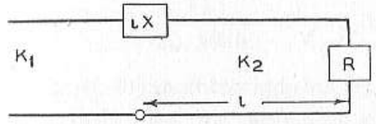


FIG. 7.21. An impedance matching scheme.

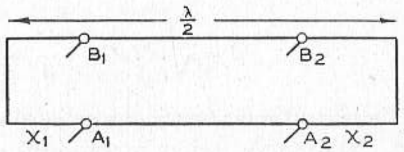


FIG. 7.22. An impedance matching scheme.

Finally let us consider a half wavelength section (Fig. 7.22). If the line is nondissipative, the voltage distribution at resonance is sinusoidal $V(x) = V_0 \sin \beta x$, where x is the distance from either end. The voltage nodes are at the ends of the section. Let us suppose that power is contributed to the line at A_1, B_1 and withdrawn from it at A_2, B_2. Let us also assume that the power withdrawn is a small fraction of the average energy stored in the matching section so that the current distribution is not affected appreciably. Then if R_1 is the resistance seen from A_1, B_1 and R_2 is the resistance across A_2, B_2, the power transfer is complete when

$$\frac{V_1^2}{R_1} = \frac{V_2^2}{R_2}, \quad \frac{R_1}{R_2} = \frac{V_1^2}{V_2^2} = \frac{\sin^2 \beta x_1}{\sin^2 \beta x_2}.$$

If x_1 and x_2 are small, then

$$\frac{R_1}{R_2} = \frac{x_1^2}{x_2^2}.$$

The above equations are approximate in that as soon as power is withdrawn across A_2, B_2, the impedance seen from A_1, B_1 acquires a small reactive component.

7.18. *Tapered Transmission Lines and Impedance Matching*

An obvious method for approximate impedance matching is to vary L and C gradually and thus change the characteristic impedance from K_1 to K_2. The match will not be perfect, of course; on the other hand, it will not be as sensitive to frequency variations as in the matching methods considered in the preceding section. Since the reflection coefficient depends on the ratio of the impedance discontinuity to the characteristic impedance of the line, it may be expected that a very satisfactory " taper " of the line will be exponential when the relative rates of change of L and C are constant.

If L and C vary exponentially with x, then voltage wave functions may be expressed in the following form

$$V^+(x) = V^+ \exp\left(-\frac{kx}{2} - i\beta x\right), \quad V^-(x) = V^- \exp\left(-\frac{kx}{2} + i\beta x\right). \qquad (18\text{–}1)$$

The constant k is the relative rate at which L decreases and C increases; hence negative values of k will mean an increasing L and decreasing C. The phase constant β is

$$\beta = \sqrt{\omega^2 L_0 C_0 - \frac{k^2}{4}} = \omega\sqrt{L_0 C_0}\sqrt{1 - \frac{k^2}{4\omega^2 L_0 C_0}}. \qquad (18\text{–}2)$$

The current wave functions associated with (1) are obtained from (10–3); thus

$$I^+(x) = \frac{\left(i\beta + \frac{k}{2}\right)V^+}{i\omega L_0} \exp\left(\frac{kx}{2} - i\beta x\right), \quad I^-(x) = -\frac{\left(i\beta - \frac{k}{2}\right)V^-}{i\omega L_0} \exp\left(\frac{kx}{2} + i\beta x\right).$$

The corresponding characteristic impedances are then

$$K^+(x) = \frac{i\omega L_0}{i\beta + \frac{k}{2}} e^{-kx}, \quad K^-(x) = \frac{i\omega L_0}{i\beta - \frac{k}{2}} e^{-kx}.$$

If the rate of taper k is small, then we have approximately

$$\beta = \omega\sqrt{L_0 C_0}, \quad K^+(x) = K^-(x) = e^{-kx}\sqrt{\frac{L_0}{C_0}}. \qquad (18\text{–}3)$$

Hence the impedance change is approximately independent of the frequency and the characteristic impedance at each point is nearly equal to the characteristic impedance of a uniform line. The rate k is small if

$$\frac{k^2}{4\omega^2 L_0 C_0} = \frac{k^2}{4\beta^2} = \frac{k^2\lambda^2}{16\pi^2} \ll 1;$$

that is if $k\lambda \ll 4\pi = 12.57\cdots$. Hence the relative rate of taper should be such that the change in impedance per wavelength is substantially less than 12.5 nepers. If the total change is 2 nepers per wavelength, then the approximate value of β in (3) differs from the exact value in (2) by less than 2%, this represents an impedance change in the ratio nearly equal to 7 4. The exact characteristic impedances of an exponentially tapered line have reactive components depending on $k\lambda$, which prevents an exact match to a uniform line even at a single frequency.*

It may be seen from equation (2) that below a certain critical frequency the propagation constant is real and the waves will be attenuated. Thus an exponential transmission line has the properties of a high pass filter. The cut-off frequency depends on the rate of taper; the larger the rate the higher the cut-off.

A practical example of this kind of impedance matching is an acoustic horn, an exponentially tapered horn in particular.

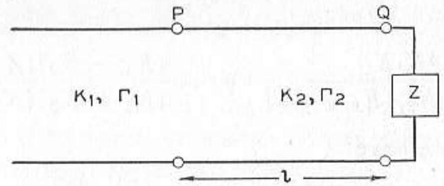


FIG. 7.23. Transmission across a " barrier."

7.19. *Transmission Across a Section of a Uniform Line*

So far we have considered reflection at a single discontinuity. Let us now suppose that we have two discontinuities at points P and Q. As shown in Fig. 7.23 we have a transmission line section of length l inserted between another line with different characteristic constants and an impedance Z.

Transmission coefficients across PQ are defined as follows

$$T_I = \frac{I_Q}{I_i} = \frac{I_Q I_P}{I_P I_i}, \quad T_V = \frac{V_Q}{V_i} = \frac{V_Q V_P}{V_P V_i}, \qquad (19\text{–}1)$$

where V_i, I_i are the incident voltage and current at P and the remaining quantities are the actual values, taking reflection into account. To find the effect produced by insertion of the section PQ into a line terminated in Z, we need only divide the above transmission coefficients by their values corresponding to $l = 0$. Likewise if we wish to know the effect produced when we replace a section of length l of a given line by another line of length l, then we merely take the ratio of the transmission coefficients as defined above for the two conditions under consideration.

Assuming that P is at $x = 0$ and Q at $x = l$, we obtain from (6–10) the following ratios

$$\frac{I_Q}{I_P} = \frac{K_2}{K_2 \cosh \Gamma_2 l + Z \sinh \Gamma_2 l}, \quad \frac{V_Q}{V_P} = \frac{Z}{Z \cosh \Gamma_2 l + K_2 \sinh \Gamma_2 l}.$$

* Without introducing a compensating reactance, of course.

Then we have from (13–4)

$$\frac{I_P}{I_i} = \frac{2K_1}{K_1 + Z_P}, \quad \frac{V_P}{V_i} = \frac{2Z_P}{K_1 + Z_P},$$

where Z_P is the impedance at P looking toward Q. Substituting from these equations in (1), we obtain

$$T_I = p_I(1 - q_I e^{-2\Gamma_2 l})^{-1} e^{-\Gamma_2 l}, \quad T_V = \frac{Z}{K_1} T_I, \tag{19–2}$$

where

$$p_I = \frac{4K_1K_2}{(K_1 + K_2)(K_2 + Z)}, \quad q_I = \frac{(K_1 - K_2)(Z - K_2)}{(K_1 + K_2)(K_2 + Z)}. \tag{19–3}$$

When $Z = K_1$, we have

$$p = \frac{4k}{(k+1)^2}, \quad q = \frac{(k-1)^2}{(k+1)^2}, \tag{19–4}$$

where k is the ratio of the characteristic impedances.

If PQ is so long that $e^{-2\alpha l}$ is small compared with unity, then we have approximately

$$T_I = p_I e^{-\Gamma_2 l}, \quad T_V = \frac{Z}{K_1} T_I. \tag{19–5}$$

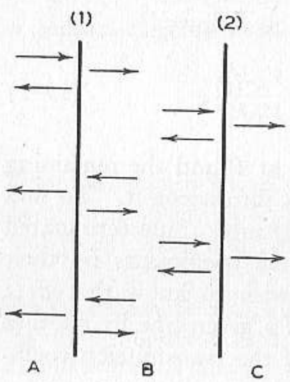


FIG. 7.24. Illustration of multiple reflections from two points of discontinuity.

The constants p_I and q_I are products of certain transmission and reflection coefficients at the two discontinuities considered separately. These relations may be better understood if we think of reflection from a pair of discontinuities as a series of reflections as represented diagrammatically in Fig. 7.24. A wave coming from region (A) suffers a reflection at the first discontinuity. Assuming that this wave is "unaware" of the existence of the second discontinuity, the transmitted part of this wave depends only on the impedances of regions (A) and (B). Let the transmission coefficient be p_1; then, if the intensity of the incident wave is unity, the intensity of the transmitted wave is p_1. The intensity of the transmitted wave becomes $p_1 e^{-\Gamma_2 l}$ when this wave reaches the second discontinuity. There the wave is partly reflected and partly transmitted; let the reflection and transmission coefficients be q_2 and p_2, on the assumption that the

reflected wave is "unaware" that it will again encounter the first discontinuity. The transmitted intensity is then $p_1p_2e^{-\Gamma_2 l}$. The reflected intensity is $p_1e^{-\Gamma_2 l}q_2$; this becomes $p_1e^{-\Gamma_2 l}q_2e^{-\Gamma_2 l}$ when the wave reaches the first discontinuity. Let the reflection coefficient be q_1; then the intensity of the part of the wave which is reflected back is $p_1e^{-\Gamma_2 l}q_2e^{-\Gamma_2 l}q_1$. This intensity is multiplied once more by the transmission ratio $e^{-\Gamma_2 l}$ when the wave reaches the second discontinuity and by p_2 when the wave passes across it. The intensity of this transmitted component is therefore

$$p_1e^{-\Gamma_2 l}q_2e^{-\Gamma_2 l}q_1e^{-\Gamma_2 l}p_2 = p_1p_2e^{-\Gamma_2 l}q_1q_2e^{-2\Gamma_2 l}.$$

The process of successive reflection is repeated indefinitely and the transmission coefficient T is obtained in the form of an infinite series in which the individual terms represent intensities of waves transmitted across the second discontinuity after successive reflections. If

$$p = p_1p_2, \quad q = q_1q_2, \tag{19–6}$$

then the series is

$$\begin{aligned} T &= pe^{-\Gamma_2 l} + pqe^{-3\Gamma_2 l} + pq^2e^{-5\Gamma_2 l} + pq^3e^{-7\Gamma_2 l} + \cdots \\ &= p(1 - qe^{-2\Gamma_2 l})^{-1}e^{-\Gamma_2 l}, \end{aligned} \tag{19–7}$$

which agrees with (2).

The values of the factors p and q depend on the type of wave we are considering. It is also necessary to remember that while p_1 is the transmission coefficient for a wave passing from region (A) to region (B), q_1 is the reflection coefficient for a wave traveling in (B) toward (A). Thus we have (see section 13)

$$\begin{aligned} p_{I,1} &= \frac{2K_1}{K_1 + K_2}, \quad p_{I,2} = \frac{2K_2}{K_2 + Z}; \quad p_{V,1} = \frac{2K_2}{K_1 + K_2}, \quad p_{V,2} = \frac{2Z}{K_2 + Z}; \\ q_{I,1} &= \frac{K_2 - K_1}{K_2 + K_1}, \quad q_{I,2} = \frac{K_2 - Z}{K_2 + Z}; \quad q_{V,1} = \frac{K_1 - K_2}{K_1 + K_2}, \quad q_{V,2} = \frac{Z - K_2}{Z + K_2} \end{aligned} \tag{19–8}$$

Hence for the voltage wave we have

$$p_V = \frac{4K_2Z}{(K_1 + K_2)(K_2 + Z)}, \quad q_V = q_I,$$

and, using this value of p in (7), we obtain T_V. The apparent lack of symmetry in equations (2) is caused by the use of p_I in both equations.

As a reminder of the directions for which the partial reflection and transmission coefficients should be employed we may write (6) in the following form

$$p = p_1^+p_2^+, \quad q = q_1^-q_2^+. \tag{19–9}$$

The above calculations can be extended to any number of sections between the given line and Z; thus for two sections, we have, instead of equation (1),

$$T_I = \frac{I_R}{I_i} = \frac{I_R I_Q I_P}{I_Q I_P I_i},$$

where R marks the end of the second section.

Noting how (2) and (3) have been formed from the equations which precede them, it is easy to write a general formula for any number of inserted sections. Let there be n sections; let the constants of a typical section be K_m, Γ_m, l_m; let the impedance looking to the right at the beginning of each section be Z_m; let K be the characteristic impedance of the line in which the sections have been inserted; then the transmission coefficient is

$$T_I = p_I[(1 - q_{I,1}e^{-2\Gamma_1 l_1})(1 - q_{I,2}e^{-2\Gamma_2 l_2}) \cdots (1 - q_{I,n}e^{-2\Gamma_n l_n})]^{-1} \times e^{-\Gamma_1 l_1 - \Gamma_2 l_2 - \cdots - \Gamma_n l_n},$$

$$p_I = \frac{2K \cdot 2K_1 \cdot 2K_2 \cdots 2K_n}{(K + K_1)(K_1 + K_2)(K_2 + K_3) \cdots (K_n + Z_{n+1})}, \quad (19\text{–}10)$$

$$q_{I,1} = \frac{(K_1 - K)(K_1 - Z_2)}{(K_1 + K)(K_1 + Z_2)}, \quad q_{I,m} = \frac{(K_m - K_{m-1})(K_m - Z_{m+1})}{(K_m + K_{m-1})(K_m + Z_{m+1})},$$

where Z_{n+1} is the terminal impedance. If the attenuation in each section is high, then Z_m is approximately equal to K_m; in this case the expression for T_I becomes simply

$$T_I = p_I e^{-\Gamma_1 l_1 - \Gamma_2 l_2 - \cdots - \Gamma_n l_n}. \quad (19\text{–}11)$$

The voltage transmission coefficient is obtained if we multiply T_I by Z_{n+1}/K.

7.20. *Reflection in Nonuniform Lines*

The equations at an impedance discontinuity are the same for uniform and nonuniform lines (13–1). Equations (13–2) are replaced, however, by the following more general equations

$$V^i = K^+ I^i, \quad V^r = -K^- I^r, \quad V^t = Z I^t. \quad (20\text{–}1)$$

The impedances K^+, K^- associated with the incident and reflected waves need no longer be equal. Solving (13–1), subject to the above conditions, we have

$$q_I = \frac{K^+ - Z}{K^- + Z}, \quad p_I = \frac{K^- + K^+}{K^- + Z},$$
$$q_V = \frac{M^+ - Y}{M^- + Y}, \quad p_V = \frac{M^- + M^+}{M^- + Y}, \quad (20\text{–}2)$$

where the admittances M^+, M^-, Y are the reciprocals of the corresponding impedances.

The transmission coefficient across a section (x_1,x_2) of another line inserted between the given line and an impedance Z may be obtained at once in the form of an infinite series analogous to (19–7). The principal difference consists in replacing the exponential factors $e^{-\Gamma_2 l}$ by proper transmission ratios as given by equations (10–6) and (10–7). Thus we set

$$p = p^+(x_1)p^+(x_2), \quad q = q^-(x_1)q^+(x_2), \quad \chi = \chi^+(x_1,x_2)\chi^-(x_2,x_1), \quad (20\text{–}3)$$

and replace (19–7) by

$$T = p(1 + q\chi + q^2\chi^2 + \cdots)\chi^+(x_1,x_2) = \frac{p}{1 - q\chi}\chi^+(x_1,x_2). \quad (20\text{–}4)$$

7.21. *Formation of Wave Functions with the Aid of Reflection Coefficients*

In section 10 we have seen that the most general wave functions are linear combinations of pairs of the "fundamental" wave functions. If we wish to form wave functions $V(x)$, $I(x)$, whose ratio is prescribed at $x = x_2$ by terminating the line in a given impedance, then we may proceed as follows. We choose one fundamental set $V^+(x)$, $I^+(x)$ as a wave which is "incident" on the given impedance and the second set as a reflected wave; then we write

$$V(x) = V^+(x) + V^+(x_2)q_V(x_2)\frac{V^-(x)}{V^-(x_2)} = V^+(x) + q_V(x_2)\frac{V^+(x_2)}{V^-(x_2)}V^-(x). \quad (21\text{–}1)$$

Similarly we obtain

$$I(x) = I^+(x) + q_I(x_2)\frac{I^+(x_2)}{I^-(x_2)}I^-(x). \quad (21\text{–}2)$$

The impedance at point $x = x_1$ may then be expressed as follows

$$\frac{V(x_1)}{I(x_1)} = K^+(x_1)\frac{1 + q_V(x_2)\chi_V^+(x_1,x_2)\chi_V^-(x_2,x_1)}{1 + q_I(x_2)\chi_I^+(x_1,x_2)\chi_I^-(x_2,x_1)}. \quad (21\text{–}3)$$

For example, in the case of uniform lines we may choose the following fundamental wave functions

$$I^+(x) = I_0e^{-\Gamma x}, \quad I^-(x) = I_1e^{\Gamma x}, \quad V_0 = KI_0,$$
$$V^+(x) = V_0e^{-\Gamma x}, \quad V^-(x) = V_1e^{\Gamma x}, \quad V_1 = -KI_1, \quad (21\text{–}4)$$

Then (1), (2) and (3) become (if $x_2 - x_1 = l$)

$$V(x) = V_0e^{-\Gamma x} + V_0q_Ve^{-2\Gamma l}e^{\Gamma x}, \quad I(x) = I_0e^{-\Gamma x} + I_0q_Ie^{-2\Gamma l}e^{\Gamma x},$$
$$Z(x_1) = K\frac{1 + q_Ve^{-2\Gamma l}}{1 + q_Ie^{-2\Gamma l}}. \quad (21\text{–}5)$$

It is important to observe that there are no restrictions on the choice of fundamental pairs of wave functions V^+, I^+ and V^-, I^- beyond the requirement of linear independence. In dealing with uniform lines the choice of progressive wave functions is, perhaps, more generally useful; in the case of nonuniform lines other choices are frequently preferable. Even in the case of uniform lines we may find it desirable in some problems to choose a different fundamental set.

For example, let us choose the following set as the fundamental set

$$V^+(x) = V_0 e^{-\Gamma x}, \quad I^+(x) = I_0 e^{-\Gamma x}, \quad V_0 = KI_0,$$
$$V^-(x) = V_1 \cosh \Gamma x, \quad I^-(x) = I_1 \sinh \Gamma x, \quad V_1 = -KI_1, \tag{21–6}$$

where V^-, I^- represent a stationary wave with the current equal to zero at $x = 0$. Substituting these expressions in (1) and (2), we have

$$V(x) = V_0 e^{-\Gamma x} + q_V(l) V_0 e^{-\Gamma x} \frac{\cosh \Gamma x}{\cosh \Gamma l},$$
$$I(x) = I_0 e^{-\Gamma x} + q_I(l) I_0 e^{-\Gamma x} \frac{\sinh \Gamma x}{\sinh \Gamma l}. \tag{21–7}$$

The above choice of V^-, I^- is such that "reflection" does not affect the original current at $x = 0$ and is particularly suited to problems in which the original wave is generated by a fixed current generator at $x = 0$. Under these circumstances the choice of a progressive wave to represent reflection would necessitate a consideration of multiple reflections since the wave reflected from an impedance Z at $x = l$ would be reflected again at the generator. The choice (6) takes care of these multiple reflections in one operation.

The input impedance is now expressed as follows

$$Z(0) = K + q_V(l) \frac{e^{-\Gamma l}}{\cosh \Gamma l} K. \tag{21–8}$$

The first term represents the input impedance of a wave which does not see the discontinuity at $x = l$; the second term may be called the *induced impedance* or the impedance *coupled* to the generator in consequence of reflection from the far end of the line. The induced impedance represents the effect of the environment as changed by the terminal impedance and it may be called the "mutual" impedance.

Even though the line is uniform, the reflection coefficients should now be calculated from the general formulae (20–2) since $K^-(l)$ is no longer equal to $K^+(l) = K$. Thus

$$K^-(l) = -\frac{V^-(l)}{I^-(l)} = K \coth \Gamma l, \quad q_V = \frac{(Z-K)\cosh \Gamma l}{K \cosh \Gamma l + Z \sinh \Gamma l}.$$

The impedance is then

$$Z(0) = K + \frac{(Z-K)e^{-\Gamma l}}{K \cosh \Gamma l + Z \sinh \Gamma l} K.$$

Naturally this expression gives the same total value for $Z(0)$ as (6–2).

Another choice of fundamental functions is particularly suitable in the case when the line is energized by a fixed voltage generator (zero internal impedance); here we have

$$V^-(x) = V_1 \sinh \Gamma x, \quad I^-(x) = I_1 \cosh \Gamma x, \quad V_1 = -KI_1,$$
$$V(x) = V_0 e^{-\Gamma x} + q_V(l) V_0 e^{-\Gamma l} \frac{\sinh \Gamma x}{\sinh \Gamma l},$$
$$I(x) = I_0 e^{-\Gamma x} + q_I(l) I_0 e^{-\Gamma l} \frac{\cosh \Gamma x}{\cosh \Gamma l}.$$

In this case the voltage at the generator remains unaffected and the effect of the impedance at the far end of the line is represented by the "reflected current." The input admittance is

$$Y(0) = M + q_I(l) \frac{e^{-\Gamma l}}{\cosh \Gamma l} M,$$

and the effect of reflection is represented by an *induced admittance*. In this case

$$M^-(l) = M \coth \Gamma l, \quad q_I(l) = \frac{(Y-M)\cosh \Gamma l}{M \cosh \Gamma l + Y \sinh \Gamma l},$$
$$Y(0) = M + \frac{(Y-M)e^{-\Gamma l}}{M \cosh \Gamma l + Y \sinh \Gamma l} M.$$

7.22. *Natural Oscillations in Uniform Transmission Lines*

In equations (0–2) for the distributed series impedance Z and shunt admittance Y we have explicitly assumed that the transverse voltage and the longitudinal current are harmonic; but throughout the greater part of the succeeding sections no use has been made of this assumption and the results obtained apply equally well to the general case in which the oscillation constant p is complex, that is,

$$Z = R + pL, \quad Y = G + pC. \tag{22–1}$$

In the circuit shown in Fig. 7.25, where Z_1 and Z_2 are the impedances seen from the terminals A, B to the left and to the right, the current is

$$I = \frac{V}{Z_1 + Z_2}. \tag{22–2}$$

An exceptional case arises when the oscillation constant is a root of the following equation

$$Z_1 + Z_2 = 0. \tag{22–3}$$

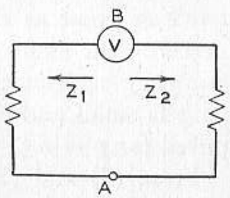


FIG. 7.25. Illustrating the condition for natural oscillations.

In Chapter 2 these roots have been named "natural oscillation constants." When p is a root of (3), an electric current Ie^{pt} may flow in the circuit without a continuous applied voltage Ve^{pt}. Such natural oscillations may be started by an impulsive force and they may be calculated by the methods outlined in section 2.9, when the roots of (3) are known.

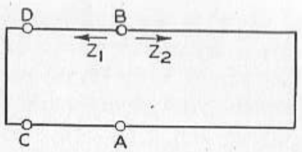


FIG. 7.26. A section of a line short-circuited at both ends.

Let us now consider a uniform transmission line of length l, short-circuited at both ends (Fig. 7.26). Here Z_1 and Z_2 are the impedances seen from any pair of terminals A, B and in particular, from terminals C, D at one of the short-circuited ends. In the latter case $Z_1 = 0$ and Z_2 is given by (6–3); thus

$$\tanh \Gamma l = 0, \quad \text{or} \quad \sinh \Gamma l = 0,$$

$$\Gamma_n l = n\pi i, \quad n = 0, \pm 1, \pm 2, \cdots, \tag{22–4}$$

where

$$\Gamma = \sqrt{(R + pL)(G + pC)}. \tag{22–5}$$

From (5) we obtain p

$$p = -\left(\frac{R}{2L} + \frac{G}{2C}\right) \pm \sqrt{\frac{\Gamma^2 - RG}{LC} + \left(\frac{R}{2L} + \frac{G}{2C}\right)^2}.$$

If R and G are small, we have approximately

$$p = -\left(\frac{R}{2L} + \frac{G}{2C}\right) \pm \frac{\Gamma}{\sqrt{LC}}. \tag{22–6}$$

Substituting from (4), we obtain

$$p_n = \xi \pm i\omega_n, \quad \xi = -\left(\frac{R}{2L} + \frac{G}{2C}\right), \quad \omega_n = \frac{n\pi}{l\sqrt{LC}} = \frac{n\pi v}{l}, \tag{22–7}$$

where v is the wave velocity in the line. The phase constant and wavelength (in the line) corresponding to the natural frequency ω_n are $\beta_n = n\pi/l$, $\lambda_n = 2l/n$. The lowest natural frequency corresponds to a wavelength twice as great as the length of the section. The above solution for p is based on the assumption that R and G are independent of p while in practice they are functions of p. However if R and G are small, their effect on p is small and equation (6) is approximately true if R and G are computed for $p = i\omega_n$.

From (7) and (5.11–16) we have an expression for Q

$$Q_n = -\frac{\omega_n}{2\xi} = \left(\frac{R}{\omega_n L} + \frac{G}{\omega_n C}\right)^{-1}.$$

Substituting for ω_n from (7), we express Q in terms of the characteristic impedance and the attenuation constant

$$Q_n = \frac{n\pi}{l}\left(\frac{R}{K} + GK\right)^{-1} = \frac{n\pi}{2\alpha l} = \frac{\beta_n}{2\alpha} = \frac{\pi}{\alpha\lambda_n}. \tag{22–8}$$

Exactly the same results are obtained for a line which is open at both ends. But if one end is open and the other short-circuited, then the equation for the natural oscillation constants is coth $\Gamma l = 0$, or cosh $\Gamma l = 0$, and its roots are $\Gamma_n l = i(n + \frac{1}{2})\pi$, $n = 0, \pm 1, \pm 2, \cdots$. In this case $\lambda_n = 4l/(2n + 1)$, and the lowest natural frequency corresponds to a wavelength (in the line) four times as great as the length of the section.

More generally if the line is terminated into an impedance Z_1 at one end and an impedance Z_2 at the other, the equation for natural oscillation constants is

$$\frac{Z_2 \cosh \Gamma l + K \sinh \Gamma l}{K \cosh \Gamma l + Z_2 \sinh \Gamma l} + \frac{Z_1}{K} = 0. \tag{22–9}$$

Frequently this equation can be solved by successive approximations. Suppose for example that Z_1 and Z_2 are pure resistances, small compared with K. Then as a first approximation we assume them equal to zero, and, if the line itself is nondissipative, we have $pl\sqrt{LC} = n\pi i$. Then we write $pl\sqrt{LC} = n\pi i + \Delta$, where Δ is a small quantity, substitute in (9), and retain only first order terms in Δ to obtain

$$\frac{R_2 + K\Delta}{K + R_2\Delta} + \frac{R_1}{K} = 0, \quad \Delta = -\frac{R_1 + R_2}{K}.$$

Thus the effect of small terminal resistances on the natural frequencies is at least of the second order; the first order effect consists of damping.

As another example let us take $Z_1 = 0$, $Z_2 = R_2 + pL_2$, and assume that the line itself is nondissipative. Equation (9) becomes

$$(R_2 + L_2 p) \cosh pl\sqrt{LC} + K \sinh pl\sqrt{LC} = 0. \tag{22–10}$$

Let

$$pl\sqrt{LC} = iu, \quad p = \frac{iu}{l\sqrt{LC}};$$

substituting in (10), we have

$$-iR_2 \cos u + \frac{KL_2}{lL} u \cos u + K \sin u = 0. \tag{22–11}$$

If R_2 is small, the first approximation is a root of

$$\frac{\tan \hat{u}}{\hat{u}} = -\frac{L_2}{lL}. \tag{22–12}$$

Letting $u = \hat{u} + \Delta$, substituting in (11), and neglecting powers of Δ above the first, we have

$$\Delta = \frac{iR_2 \cos^2 \hat{u}}{K\left(1 - \dfrac{\sin 2\hat{u}}{2\hat{u}}\right)}, \quad p = -\frac{R_2 \cos^2 \hat{u}}{Ll\left(1 - \dfrac{\sin 2\hat{u}}{2\hat{u}}\right)} + \frac{i\hat{u}}{l\sqrt{LC}}.$$

Roots of (12) may be obtained either graphically or numerically. If, however, the right-hand side is small compared with unity, then we have approximately $\hat{u} = n\pi + \delta$, where δ is a small quantity. Hence $\delta = -n\pi L_2/lL$, as long as n is not too large, and therefore $\hat{u} = n\pi[1 - (L_2/lL)]$.

7.23. *Conditions for Impedance Matching and Natural Oscillations in Terms of the Reflection Coefficient*

Let one of the impedances in Fig. 7.25 be the characteristic impedance of a transmission line (Fig. 7.27). We assume that an incident wave is originated at infinity and that the reflected wave goes back to infinity without further reflection. The reflection coefficient (for the current for example) is

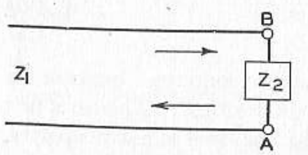


FIG. 7.27. Illustrating the conditions for impedance matching and for natural oscillations.

$$q = \frac{Z_1 - Z_2}{Z_1 + Z_2}. \tag{23–1}$$

If the impedances are equal, q vanishes

$$Z_1 - Z_2 = 0, \quad q = 0, \tag{23–2}$$

and there is no reflected wave. On the other hand if

$$Z_1 + Z_2 = 0, \quad q = \infty, \tag{23–3}$$

then a reflected wave may exist without an incident wave. This is the condition for natural oscillations.

In circuit theory the reflection coefficient is used as a measure of impedance mismatch for any pair of impedances (Fig. 7.25). If the impedances are equal the reflection coefficient is zero and the power is evenly distributed between them; if the sum of the impedances is zero, the reflection coefficient is infinite and oscillations may exist without a continuously applied electromotive force.

7.24. *Expansions in Partial Fractions*

The input impedance and admittance can be expanded in partial fractions in terms of the natural oscillation constants. If the line is slightly dissipative, such expansions can be obtained by computing the energies associated with different oscillation modes and using equations (5.11–18) and (5.11–19). In the present case this is unnecessary since the impedances have already been calculated and it is easy to obtain the expansions directly. For example, for short-circuited and open-circuited lines we use the well known expansions

$$\tanh x = \sum_{n=1}^{\infty} \frac{8x}{(2n-1)^2\pi^2 + 4x^2}, \quad \coth x = \frac{1}{x} + \sum_{n=1}^{\infty} \frac{2x}{n^2\pi^2 + x^2}. \tag{24–1}$$

Hence for a line of length l, short-circuited at the output end, we have

$$Z(0) = 8K \sum_{n=1}^{\infty} \frac{\Gamma l}{(2n-1)^2\pi^2 + 4\Gamma^2 l^2}, \quad Y(0) = \frac{1}{K\Gamma l} + \frac{2}{K} \sum_{n=1}^{\infty} \frac{\Gamma l}{n^2\pi^2 + \Gamma^2 l^2}. \tag{24–2}$$

Since $K\Gamma = Z$ and $\Gamma/K = Y$, we have

$$Z(0) = Zl \sum_{n=1}^{\infty} \frac{8}{(2n-1)^2\pi^2 + 4\Gamma^2 l^2}, \quad Y(0) = \frac{1}{Zl} + Yl \sum_{n=1}^{\infty} \frac{2}{n^2\pi^2 + \Gamma^2 l^2}. \tag{24–3}$$

These expressions show how the low frequency impedance Zl of the line is modified at high frequencies. The second equation, in particular, gives the effect as an impedance in parallel with the low frequency impedance.

Substituting $\Gamma = \alpha + i\beta$, neglecting α^2 in the denominator and α in the numerator we have

$$Z(0) = K \sum_{n=1}^{\infty} \frac{8i\beta l}{[(2n-1)^2\pi^2 - 4\beta^2 l^2] + 8i\alpha\beta l^2},$$
$$Y(0) = \frac{1}{K}\left[\frac{1}{\alpha l + i\beta l} + \sum_{n=1}^{\infty} \frac{2i\beta l}{(n^2\pi^2 - \beta^2 l^2) + 2i\alpha\beta l^2}\right]. \tag{24–4}$$

The approximate natural frequencies can be found from the above equations by inspection (since $\beta = \omega\sqrt{LC}$). Thus if the input terminals are short-circuited, the natural frequencies are

$$\hat{\omega}_n = \frac{n\pi}{l\sqrt{LC}}; \tag{24–5}$$

and if the input terminals are open, then

$$\omega_n = \frac{(2n-1)\pi}{2l\sqrt{LC}}. \tag{24–6}$$

In terms of the natural frequencies and corresponding Q's, we have

$$Z(0) = \frac{2}{Cl} \sum_{n=1}^{\infty} \frac{i\omega}{\omega_n^2 - \omega^2 + \dfrac{i\omega\omega_n}{Q_n}}, \quad Y(0) = \frac{1}{(R + i\omega L)l} + \frac{2}{Ll} \sum_{n=1}^{\infty} \frac{i\omega}{\hat{\omega}_n^2 - \omega^2 + \dfrac{i\omega\hat{\omega}_n}{\hat{Q}_n}}. \tag{24–7}$$

Similarly for a line of length l, open at the output end, we obtain

$$Z(0) = \frac{1}{(G + i\omega C)l} + \frac{2}{Cl} \sum_{n=1}^{\infty} \frac{i\omega}{\hat{\omega}_n^2 - \omega^2 + \dfrac{i\omega\hat{\omega}_n}{\hat{Q}_n}}, \quad Y(0) = \frac{2}{Ll} \sum_{n=1}^{\infty} \frac{i\omega}{\omega_n^2 - \omega^2 + \dfrac{i\omega\omega_n}{Q_n}}, \tag{24–8}$$

where $\hat{\omega}_n$ and ω_n are given by (5) and (6).

Let us now consider a section of length l, short-circuited at both ends, with the input terminals at distance $x = d$ from one of the ends (Fig. 7.28). It is particularly easy to obtain the expansion for the input admittance $Y(d)$ using the method explained in section 5.11. The infinities of $Y(d)$ are the natural frequencies when the input terminals are short-circuited and hence are given by (5). The current in the line (if the line is only slightly dissipative) is substantially proportional to $\cos \beta_n x$, where x is the distance from one of the ends; hence, for a unit amplitude at the input terminals we have

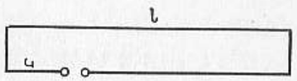


FIG. 7.28. A section of a line short-circuited at both ends.

$$I_n(x) = \frac{\cos \beta_n x}{\cos \beta_n d}. \tag{24–9}$$

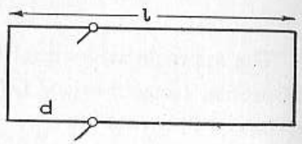
For the nth oscillation mode the energy stored in the line is

$$\hat{\mathcal{E}}_n = \tfrac{1}{2}L \int_0^l [I_n(x)]^2\, dx = \frac{Ll}{4 \cos^2 \beta_n d}, \quad \hat{\mathcal{E}}_0 = \tfrac{1}{2}Ll. \tag{24–10}$$

Substituting in (5.11–19) and including the term corresponding to (5.11–24) we have

$$Y(d) = \frac{1}{(R + i\omega L)l} + \frac{2}{Ll}\sum_{n=1}^{\infty} \frac{i\omega \cos^2 \beta_n d}{\hat{\omega}_n^2 - \omega^2 + \dfrac{i\omega\hat{\omega}_n}{\hat{Q}_n}}. \tag{24–11}$$

Similarly it is easy to find the expansion for the input impedance in the case shown in Fig. 7.29 because, if the terminals are open, the natural frequencies are the same as in the preceding case. The voltage in the line is proportional to $\sin \beta_n x$ and, adjusting it for a unit amplitude at $x = d$, we have



$$V_n(x) = \frac{\sin \beta_n x}{\sin \beta_n d}. \tag{24–12}$$

FIG. 7.29. A section of a line short-circuited at both ends.

The energy of a typical oscillation mode is then

$$\mathcal{E}_n = \tfrac{1}{2}C \int_0^l [V_n(x)]^2\, dx = \frac{Cl}{4 \sin^2 \beta_n d}. \tag{24–13}$$

Substituting in (5.11–18), we have

$$Z(d) = \frac{2}{Cl}\sum_{n=1}^{\infty} \frac{i\omega \sin^2 \beta_n d}{\omega_n^2 - \omega^2 + \dfrac{i\omega\omega_n}{Q_n}}. \tag{24–14}$$

In order to obtain the input impedance in the case shown in Fig. 7.28 we have to determine the natural frequencies when the terminals are open and then calculate the energies of the corresponding oscillations. In this case the two sections of the line are in series and the input impedance is the sum of two expansions similar to the expansion for $Z(0)$ in (7). In the case shown in Fig. 7.29 the two sections of the line are in parallel and the input admittance is the sum of two expansions similar to the expansion for $Y(0)$ in (7).

7.25. *Multiple Transmission Lines*

Equations (0–1) are said to describe a *simple* transmission line or a line admitting of only one transmission mode. In general transmission equations are of the following form

$$\frac{dV_m}{dx} = -\sum_{k=1}^{n} Z_{mk} I_k, \quad \frac{dI_m}{dx} = -\sum_{k=1}^{n} Y_{mk} V_k. \tag{25–1}$$

For example, if two transmission lines run parallel to each other waves in one may influence waves in the other, since alternating longitudinal currents in one line induce longitudinal voltages in the other, and alternating transverse voltages induce transverse currents. Thus for a pair of interacting transmission lines we have

$$\frac{dV_1}{dx} = -Z_{11}I_1 - Z_{12}I_2, \quad \frac{dI_1}{dx} = -Y_{11}V_1 - Y_{12}V_2,$$
$$\frac{dV_2}{dx} = -Z_{21}I_1 - Z_{22}I_2, \quad \frac{dI_2}{dx} = -Y_{21}V_1 - Y_{22}V_2. \tag{25–2}$$

Z_{12} is the distributed mutual impedance per unit length and Y_{12} is the distributed mutual admittance per unit length.

Equations (1) are linear differential equations with constant coefficients and hence possess solutions of exponential form

$$V_m(x) = V_m e^{-\Gamma x}, \quad I_m(x) = I_m e^{-\Gamma x}. \tag{25–3}$$

Substituting in (1), we obtain

$$\Gamma V_m = \sum Z_{mk} I_k, \quad \Gamma I_m = \sum Y_{mk} V_k. \tag{25–4}$$

Thus we have $2n$ linear homogeneous equations connecting $2n$ variables V_m, I_m. These equations will possess solutions, not vanishing identically, only if the determinant of their coefficients is zero. This determinant is an equation of the nth degree in Γ^2 and its solutions represent the natural propagation constants of the multiple transmission line. For each value of Γ we can determine the ratios of V_m, I_m to some one variable. Thus there are $2n$ arbitrary constants at our disposal and these may be determined to satisfy assigned boundary or initial conditions. The line is said to possess n transmission modes. Each transmission mode is characterized by its pair $\pm\Gamma_m$ of propagation constants and by a relative distribution of voltages and currents peculiar to it.

In practical applications the mutual coefficients are frequently very small. Then equations (1) may be solved by successive approximations, the first being the solution of n independent pairs of equations

$$\frac{d\hat{V}_m}{dx} = -Z_{mm}\hat{I}_m, \quad \frac{d\hat{I}_m}{dx} = -Y_{mm}\hat{V}_m, \tag{25–5}$$

in which the interaction between the individual transmission lines has been neglected. The values of the voltages and currents obtained from (5) are now substituted in all the small terms of (1); thus we obtain

$$\frac{dV_m}{dx} = -Z_{mm}I_m - \sum{}'Z_{mk}\hat{I}_k, \quad \frac{dI_m}{dx} = -Y_{mm}V_m - \sum{}'Y_{mk}\hat{V}_k, \qquad (25\text{–}6)$$

where the primes indicate the omission in the summation of terms for which $k = m$. These equations are the equations of simple transmission lines with given distributions of applied voltages and currents. Solving these equations we obtain corrections to the previous solutions. The process can be repeated as often as may be necessary; but usually the first corrections are sufficient.

In communication engineering the interference between neighboring lines is called *crosstalk*. The interference due to the mutual impedances Z_{mk} is called the *impedance crosstalk*; similarly the interference due to the mutual admittances Y_{mk} is called the *admittance crosstalk*.

7.26. *Iterative Structures*

If a pair of sections (Fig. 7.30) of uniform transmission lines is repeated an indefinite number of times, an iterative structure is obtained which may have properties radically different from the properties of the original lines. Thus if the original lines were capable of transmitting all frequencies, the iterative structure might suppress some frequency bands.

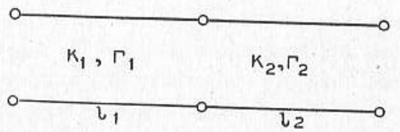


FIG. 7.30. A transducer formed by two line sections in tandem.

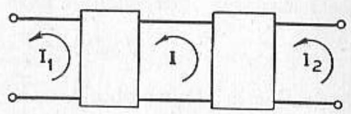


FIG. 7.31. A chain of transducers.

The equations for the present iterative structure may be obtained from the general equations of section 5.3 as soon as the constants of the transducer shown in Fig. 7.30 are calculated. This transducer consists of two transducers in tandem (Fig. 7.31). Using single primes for the constants of the first transducer and double primes for those of the second, we have

$$Z'_{11}I_1 + Z'_{12}I = V_1, \quad Z''_{21}I + Z''_{22}I_2 = V_2,$$
$$Z'_{21}I_1 + (Z'_{22} + Z''_{11})I + Z''_{12}I_2 = 0. \qquad (26\text{–}0)$$

Solving the last equation for I and substituting in the remaining equations, we obtain the constants of the combined transducer

$$Z_{11} = Z'_{11} - \frac{Z'_{12}Z'_{21}}{Z'_{22} + Z''_{11}}, \quad Z_{12} = -\frac{Z'_{12}Z''_{12}}{Z'_{22} + Z''_{11}},$$
$$Z_{22} = Z'_{22} - \frac{Z''_{12}Z''_{21}}{Z'_{22} + Z''_{11}}, \quad Z_{21} = -\frac{Z''_{21}Z'_{21}}{Z'_{22} + Z''_{11}}. \qquad (26\text{–}1)$$

In the case of uniform transmission lines the transfer impedances of the constituent transducers are equal and consequently $Z_{12} = Z_{21}$.

From (1) and from equations of section 7, we obtain the following expressions for the transducer shown in Fig. 7.30

$$Z_{11} = \frac{K_1^2 + K_1K_2 \coth \Gamma_1 l_1 \coth \Gamma_2 l_2}{K_1 \coth \Gamma_1 l_1 + K_2 \coth \Gamma_2 l_2},$$
$$Z_{22} = \frac{K_2^2 + K_1K_2 \coth \Gamma_1 l_1 \coth \Gamma_2 l_2}{K_1 \coth \Gamma_1 l_1 + K_2 \coth \Gamma_2 l_2}, \qquad (26\text{–}2)$$
$$Z_{12} = -\frac{K_1K_2 \operatorname{csch} \Gamma_1 l_1 \operatorname{csch} \Gamma_2 l_2}{K_1 \coth \Gamma_1 l_1 + K_2 \coth \Gamma_2 l_2}.$$

By (5.3–7) the propagation constant Γ per section of the iterative structure is

$$\cosh \Gamma = \frac{K_1^2 + K_2^2}{2K_1K_2} \sinh \Gamma_1 l_1 \sinh \Gamma_2 l_2 + \cosh \Gamma_1 l_1 \cosh \Gamma_2 l_2$$
$$= \frac{1}{p}[\cosh (\Gamma_1 l_1 + \Gamma_2 l_2) - q^2 \cosh (\Gamma_1 l_1 - \Gamma_2 l_2)],$$

where q is the reflection coefficient and p the product of the transmission coefficients

$$p = \frac{4K_1K_2}{(K_1 + K_2)^2}, \quad q^2 = \frac{(K_1 - K_2)^2}{(K_1 + K_2)^2}.$$

Let us suppose that the transmission lines are nondissipative; then $\Gamma_1 = i\omega/v_1$, $\Gamma_2 = i\omega/v_2$, and we have

$$\cosh \Gamma = \frac{1}{p}(\cos \omega T - q^2 \cos \omega\tau),$$
$$T = \frac{l_1}{v_1} + \frac{l_2}{v_2}, \quad \tau = \frac{l_1}{v_1} - \frac{l_2}{v_2}.$$

For some values of ω, $\cosh \Gamma$ will be in the interval $(-1,+1)$ and Γ will be imaginary; for other values Γ will be either real or complex. Hence the structure will transmit some frequencies and suppress the remaining. Pass and stop bands may be determined by plotting $\cosh \Gamma$ as a function of ωT or $\omega\tau$.

7.27. *Resonance in Slightly Nonuniform Transmission Lines*

Consider a section of a slightly nonuniform nondissipative transmission line of length l and let this section be open at $x = 0$ and short-circuited at $x = l$. In the first approximation the longest resonant wavelength is

$$\lambda = 4l, \quad l = \frac{\lambda}{4}. \qquad (27\text{–}1)$$

The first correction may be obtained from the equations of sections 11 and 12. The current I_0 at $x = 0$ and voltage $V(l) \simeq V_0(l) + V_1(l)$ at $x = l$

vanish; thus the equation for resonance is approximately

$$[1 + B(l)] \cos \beta l - iA(l) \sin \beta l + iC(l) \sin \beta l = 0. \tag{27–2}$$

Assuming that A, B and C are small compared with unity, we let

$$\lambda = 4l(1 - \delta), \quad l = \frac{\lambda}{4}(1 + \delta), \tag{27–3}$$

where δ is small compared with unity. Substituting in (2) and ignoring small quantities of the second order, we obtain

$$-\frac{\pi\delta}{2} - iA(l) + iC(l) = 0, \quad \delta = \frac{2i}{\pi}[C(l) - A(l)]. \tag{27–4}$$

The values of $A(l)$ and $C(l)$ are computed* from (12–2)

$$A(l) = \tfrac{1}{2}\int_0^l \left(\frac{\hat{Z}}{K_0} - K_0\hat{Y}\right) \cos \frac{\pi\xi}{l}\, d\xi, \quad C(l) = \tfrac{1}{2}\int_0^l \left(\frac{\hat{Z}}{K_0} + K_0\hat{Y}\right) d\xi. \tag{27–5}$$

If the line is short-circuited at $x = 0$ and open at $x = l$, so that $V_0 = 0$ and $I(l) = 0$ then the first approximation to the principal resonant wavelength is (1) and the correction δ is

$$\delta = \frac{2i}{\pi}[A(l) + C(l)]. \tag{27–6}$$

In many practical cases the product ZY is constant and (see 12–3)

$$\sqrt{ZY} = \sqrt{Z_0Y_0} = i\beta, \quad \frac{\hat{Y}}{Y_0} = -\frac{\hat{Z}}{Z_0} = -\frac{\hat{K}}{K_0}. \tag{27–7}$$

K_0 is the average characteristic impedance which may be defined by

$$K_0 = \frac{1}{l}\int_0^l K(x)\, dx, \quad K(x) = \sqrt{\frac{Z(x)}{Y(x)}}, \tag{27–8}$$

and $\hat{K}$ is the deviation of the "nominal" characteristic impedance $K(x)$ from this average impedance $\hat{K}(x) = K(x) - K_0$. In this case $C(l) = 0$ and

$$iA(l) = -\beta\int_0^l \frac{\hat{K}(x)}{K_0} \cos \frac{\pi x}{l}\, dx = \frac{\pi}{2l}\int_0^l \left[1 - \frac{K(x)}{K_0}\right] \cos \frac{\pi x}{l}\, dx. \tag{27–9}$$

Hence if the line is open at $x = 0$ and shorted at $x = l$, then

$$\delta = \frac{1}{l}\int_0^l \left[\frac{K(x)}{K_0} - 1\right] \cos \frac{\pi x}{l}\, dx = \frac{1}{lK_0}\int_0^l K(x) \cos \frac{\pi x}{l}\, dx. \tag{27–10}$$

* In the integrand of $A(l)$ we replace λ by its first approximation $4l$.

If the line is shorted at $x = 0$ and open at $x = l$, then

$$\bar{\delta} = \frac{1}{l}\int_0^l \left[1 - \frac{K(x)}{K_0}\right] \cos \frac{\pi x}{l}\, dx = -\frac{1}{lK_0}\int_0^l K(x) \cos \frac{\pi x}{l}\, dx. \tag{27–11}$$

The integrand in the last terms of the above equations is positive near $x = 0$ and negative near $x = l$. If the capacitance per unit length varies more or less uniformly and if it is larger near the open end than near the shorted end, then the principal resonant wavelength is somewhat shorter than $4l$. If the capacitance is larger near the shorted end, then the resonant wavelength is longer than $4l$.

If the line is shorted at both ends $x = 0$ and $x = l$, then the principal resonance occurs approximately when $\lambda = 2l$, $l = \lambda/2$. It is left to the reader to show that the more accurate expressions are

$$\lambda = 2l(1 + \chi), \quad l = \frac{\lambda}{2}(1 - \chi), \tag{27–12}$$

where $\chi = (1/i\pi)[A(l) + C(l)]$ and

$$A(l) = \tfrac{1}{2}\int_0^l \left(\frac{\hat{Z}}{K_0} - K_0\hat{Y}\right) \cos \frac{2\pi x}{l}\, dx, \quad C(l) = \tfrac{1}{2}\int_0^l \left(\frac{\hat{Z}}{K_0} + K_0\hat{Y}\right) dx. \tag{27–13}$$

If ZY is constant, then we have

$$\chi = \frac{1}{K_0 l}\int_0^l [K(x) - K_0] \cos \frac{2\pi x}{l}\, dx = \frac{1}{K_0 l}\int_0^l K(x) \cos \frac{2\pi x}{l}\, dx. \tag{27–14}$$

If both ends are open, then the correction term is $-\chi$.

The above approximate expressions give the two principal terms, one independent of and the other varying inversely as the average characteristic impedance. By continuing the process of successive approximations in the solution of nonuniform transmission lines more accurate expressions for the resonant lengths of such lines can be obtained; but usually the above formulae are satisfactory for practical purposes. The method applies also when the line is terminated in some reactance at either or both ends, as in this case we may start with the prescribed value of $V(0)/I(0)$ and plot the ratio $V(l)/I(l)$ as a function of l until we obtain the prescribed value of this ratio.

In the simple cases when the ends of the line are either open or shorted, another method of treatment yields the above results and some similar approximations. Under these terminal conditions the energy equation (3–3) for nondissipative lines becomes

$$\int_0^l CVV^*\, dx = \int_0^l LII^*\, dx. \tag{27–15}$$

This equation simply reiterates the fact that at resonance the maximum electric and magnetic energies are equal. In the important special case in which $LC = v^{-2}$ = constant, the transmission equations may be expressed in the following form

$$I = \frac{iv^2 C}{\omega}\frac{dV}{dx}, \quad V = \frac{iv^2 L}{\omega}\frac{dI}{dx}.$$

Substituting in (15), we obtain

$$\frac{\omega^2}{v^2} = \frac{4\pi^2}{\lambda^2} = \frac{\int_0^l L\left|\frac{dI}{dx}\right|^2 dx}{\int_0^l L\,|I|^2\,dx} = \frac{\int_0^l C\left|\frac{dV}{dx}\right|^2 dx}{\int_0^l C\,|V|^2\,dx}. \tag{27–16}$$

If now the line is open at $x = 0$ and short-circuited at $x = l$, we assume the current and voltage distributions that would exist in a uniform line

$$I(x) = I \sin\frac{\pi x}{2l}, \quad V(x) = V \cos\frac{\pi x}{2l},$$

and substitute in (16). Thus we obtain

$$\frac{16l^2}{\lambda^2} = \frac{1+\delta}{1-\delta} = \frac{1-\xi}{1+\xi}, \tag{27–17}$$

where

$$\delta = \frac{\int_0^l L\cos\frac{\pi x}{l}\,dx}{\int_0^l L\,dx}, \quad \xi = \frac{\int_0^l C\cos\frac{\pi x}{l}\,dx}{\int_0^l C\,dx}. \tag{27–18}$$

The above value of δ is the same as that given by equation (10).

If the line is shorted at $x = 0$ and open at $x = l$, then

$$\frac{16l^2}{\lambda^2} = \frac{1-\delta}{1+\delta} = \frac{1+\xi}{1-\xi}. \tag{27–19}$$

If the line is shorted at both ends, we assume

$$I(x) = I\cos\frac{\pi x}{l}, \quad V(x) = V\sin\frac{\pi x}{l},$$

and find

$$\frac{4l^2}{\lambda^2} = \frac{1-\chi}{1+\chi} = \frac{1+\hat{\chi}}{1-\hat{\chi}}, \tag{27–20}$$

where

$$\chi = \frac{\int_0^l L\cos\frac{2\pi x}{l}\,dx}{\int_0^l L\,dx}, \quad \hat{\chi} = \frac{\int_0^l C\cos\frac{2\pi x}{l}\,dx}{\int_0^l C\,dx}. \tag{27–21}$$

For a line open at both ends we have

$$\frac{4l^2}{\lambda^2} = \frac{1+\chi}{1-\chi} = \frac{1-\hat{\chi}}{1+\hat{\chi}}. \tag{27–22}$$

The above approximations in terms of L and in terms of C are actually different approximations except when the deviation of L and C from uniform distribution is small, which is the condition postulated in deriving the formulae. Nevertheless in cases of substantial nonuniformities when one formula may fail altogether the other may give surprisingly good results. We shall encounter, for example, lines in which C is proportional and L inversely proportional to x. The formulae depending on L fail, but those depending on C give very accurate results.

CHAPTER VIII

Waves, Wave Guides and Resonators — 1

8.0. *Introduction*

If we were to start with a physical situation met in practice, we should be faced with a complicated mathematical problem having a complicated solution. Frequently after studying the solution numerically we should find that it reduces to a simple approximate solution. In practical work it is more expedient to make approximations *ab initio*; this, however, requires experience and good physical sense that may develop from it. The mathematician's "laboratory experiments" consist in looking for simple special solutions and studying them; then he can look for physical conditions under which such solutions might be realized either exactly or approximately. This chapter is devoted to just such experiments, chosen either for their value in explaining electromagnetic waves or for their practical value in general and in communication engineering in particular. A more comprehensive treatment of Maxwell's equations is reserved for Chapter 10.

8.1. *Uniform Plane Waves*

A *plane wave* is a wave whose equiphase surfaces form a family of parallel planes. In a *uniform plane wave* the field intensities are independent of the coordinates in each equiphase plane. Choosing the *xy*-plane parallel to the equiphase planes and setting $\partial/\partial x = \partial/\partial y = 0$ in the general electromagnetic equations, we have $E_z = 0$, $H_z = 0$, and

$$\frac{dE_x}{dz} = -i\omega\mu H_y, \quad \frac{dH_y}{dz} = -(g + i\omega\epsilon)E_x; \tag{1–1}$$

$$\frac{dE_y}{dz} = i\omega\mu H_x, \quad \frac{dH_x}{dz} = (g + i\omega\epsilon)E_y. \tag{1–2}$$

Thus uniform plane waves are transverse electromagnetic.

The four transverse components in (1) and (2) are divided into two independent pairs E_x, H_y and E_y, H_x. In the wave associated with each pair E and H are mutually perpendicular. The difference in the algebraic signs in (1) and (2) is removed by reversing the positive direction of one component, H_x for example, and equations (1) become general if the positive directions of E, H, and the phase velocity v are related as shown in Fig. 8.1.

Equations (1) are seen to be the equations for a transmission line whose series inductance is μ, shunt conductance g and shunt capacitance ϵ, all taken per unit length. The electric intensity E_x plays the part of the transverse electromotive force and the magnetic intensity that of the transverse magnetomotive force or the longitudinal current. The longitudinal current as such is not in evidence because our equations refer to a typical "unit wave tube" rather than to the complete wave, the boundaries of which have been made inaccessible by the assumption of infinite dimensions. We can introduce a pair of parallel perfectly conducting planes perpendicular to E, without disturbing the electrodynamic equilibrium; we can assume that the field has been annihilated except in the region between the planes since perfect conductors provide a perfect barrier to the flow of energy; and finally we may cut out of the planes infinitely long narrow slits (Fig. 8.2) and separate a pair of parallel infinitely long strips from the rest of the plane conductors which, however, are retained as "guards" to keep the electric and magnetic lines straight. The longitudinal current I in the lower strip and the transverse voltage from the lower strip to the upper are* then

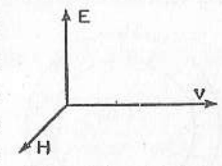


Fig. 8.1. Relative directions of E, H, v in a uniform plane wave.

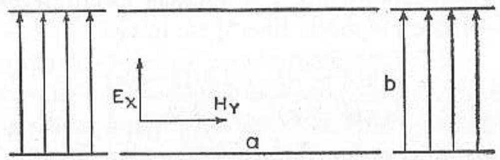


Fig. 8.2. A cross-section of two parallel strips with guards to eliminate the edge effect.

$$I = aH_y, \quad V = bE_x.$$

Substituting in (1), we have the following equations for our transmission line

$$\frac{dV}{dz} = -i\omega LI, \quad \frac{dI}{dz} = -(G + i\omega C)V,$$

$$L = \frac{\mu b}{a}, \quad G = \frac{ga}{b}, \quad C = \frac{\epsilon a}{b}.$$

The characteristic impedance and propagation constant are

$$K = \frac{\eta b}{a} = \frac{b}{a}\sqrt{\frac{i\omega\mu}{g + i\omega\epsilon}}, \quad \Gamma = \sigma = \sqrt{i\omega\mu(g + i\omega\epsilon)}.$$

* Assuming I positive in the positive z-direction which is away from the reader in Fig. 8.2.

If the guard plates are removed, the … near the edges of our parallel strips will bulge out as shown … nd the magnetic lines will bend round to enclose one conduct… r. Subsequent analysis will show that a wave of this modifi… exist at all frequencies and that the shape … ion of the electric and magnetic lines are in… f the frequency. For such a wave the edg… all if b is small compared to a since th… istributed largely between the plates.

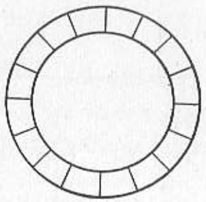


FIG. 8.3. Coaxial cylinders of nearly equal radii.

If b is small comp… he strips can be bent into cylinders to for… nductors with nearly equal radii (Fig. 8.3) … lines will run along radii and magnetic li… oaxial circles between the conductors. Th… a slight " curvature effect " instead of the edge effect. The… ect is comparatively small; thus if the radii of the conduct… b $(b > a)$, then by the parallel plane formula (using the a… nference for the approximate length of the magnetic lines),

$$K = \frac{\eta(b-a)}{\pi(b+a)} = \ldots$$

If $b = 2a$, this gives $K = 40$ ohms; the … 41.6.

Since the voltages along various parts … dius are added while the magnetomotive force is the same for … lines, the characteristic impedance of a coaxial pair is the sum … cteristic impedances of coaxial shells into which the space be… onductors might be subdivided. Thus if $b - a$ is divided int… rts, the exact value of K may be expressed in the following for…

$$K = \lim 120(b-a)\left[\frac{1}{(2n-1)a+b} + \frac{}{(2n} \ldots \frac{1}{a+(2n-1)b}\right]$$

as $n \to \infty$.

Taking again $b = 2a$ and choosing $n = 2$, … $= 41.1$; this value differs from the exact value by about 1 pe…

Let us now return to waves in an unlim… n. With transverse dimensions fading out of the picture, we … tention on the field intensities E and H, rather than on thei… values, and define the ratio E/H as the *wave impedance* in th… f wave propagation. A uniform plane wave can be generated by … ent sheet of uniform

* When a numerical value is ascribed to the intr… e, free space is usually assumed.

density. Consider such a sheet in the xy-plane and let its density be J_x. Since the electric intensity is continuous at the sheet while the magnetic intensity is discontinuous, we have

$$E_x(+0) = E_x(-0), \quad H_y(+0) - H_y(-0) = -J_x.$$

The current sheet acts as a shunt generator and sends out plane waves in both directions

$$E_x^+(z) = -\tfrac{1}{2}\eta J_x e^{-\sigma z}, \quad H_y^+(z) = -\tfrac{1}{2}J_x e^{-\sigma z}, \quad z > 0,$$
$$E_x^-(z) = -\tfrac{1}{2}\eta J_x e^{\sigma z}, \quad H_y^-(z) = \tfrac{1}{2}J_x e^{\sigma z}, \quad z < 0.$$

The complex power (per unit area) contributed to the field by the impressed forces is

$$\Psi = -\tfrac{1}{2}E_x(0)J_x^* = \tfrac{1}{4}\eta J_x J_x^*.$$

If the medium is nondissipative, then the power carried by each wave per unit area in an equiphase plane is

$$\Psi^+ = \tfrac{1}{2}E_x^+(z)[H_y^+(z)]^* = \tfrac{1}{8}\eta J_x J_x^*, \quad \Psi^- = -\tfrac{1}{2}E_x^-(z)[H_y^-(z)]^* = \tfrac{1}{8}\eta J_x J_x^*.$$

The sum is equal to the power contributed to the field.

The total power carried by a uniform plane wave in an unlimited medium is infinite and the wave cannot possibly be started by an ordinary generator. The principal reason for considering such waves at all is their simplicity, combined with the fact that at great distances from any antenna and in a sufficiently limited region the wave is nearly plane.

If the medium is nondissipative it is possible to send all the energy in one direction only. Consider two parallel equal current sheets (1) and (2), a quarter wavelength apart, and let the currents be in quadrature. If the current in the left-hand sheet (2) is 90 degrees ahead, then the right-hand wave generated by it will be in phase with the right-hand wave generated by the sheet (1); the two waves will reinforce each other. The left-hand wave from (1) will be 180 degrees out-of-phase with the left-hand wave from (2); the two waves will destroy each other to the left of the plane (2). The electric intensity of the wave produced by the sheet (1) will directly oppose the electric intensity of the second sheet and reduce the total intensity at that sheet to zero; hence the second sheet contributes no power and may be taken to be a perfect conductor. The electric intensities of the two waves reinforce each other at the sheet (1). Assuming that this sheet is in the plane $z = 0$, we have therefore

$$E_x^+(z) = -\eta J_x e^{-i\beta z}, \quad H_y^+(z) = -J_x e^{-i\beta z}, \quad z > 0.$$

The power emitted by the sheet is twice that which would be emitted by an isolated sheet.

Let us look at the situation in another way and assume at the start that the plane (2) is a perfect conductor. By (7.6–3) the impedance as seen from plane (1) leftward is $Z_z^- = \eta \tanh i\beta l = i\eta \tan \beta l$, where l is the distance between the planes. If $l = \lambda/4$, this impedance is infinite and no power will flow to the left of plane (1).

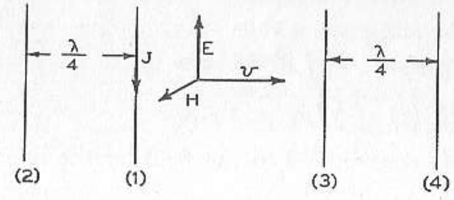


FIG. 8.4. Passing of waves through and reflection from resistance sheets.

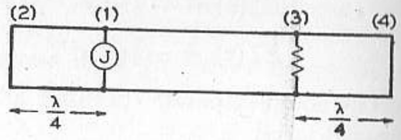


FIG. 8.5. Transmission diagram representing the case in Fig. 8.4.

The wave to the right of sheet (1) can be completely absorbed by a thin conducting sheet (3), with surface resistance equal to η, if the sheet has a perfectly conducting sheet (4) a quarter wavelength behind (Fig. 8.4). To facilitate the use of the transmission theory of the preceding chapter we construct a transmission line diagram (Fig. 8.5) in which the current sheet (1) is shown as a shunt generator, the resistance sheet (3) as a shunt resistance and the perfect conductors (2) and (4) as zero resistances at the ends of the line. Without the reflector (4) the impedance of the sheet (3) would be in parallel with the intrinsic impedance of the medium behind it, the impedance presented to the incoming wave would be only $\frac{1}{2}\eta$, and some of the wave would be reflected. It should be noted that the absorber (3) will function just as well even if the medium between the resistance sheet and the reflector is different from that between the resistance sheet and the generator.

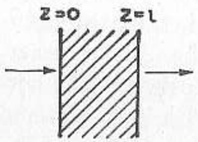


FIG. 8.6. A cross-section of an infinite metal plate.

The impedance normal to a plate of thickness l (Fig. 8.6) is in general

$$Z_z(0) = \eta \frac{Z_z(l) \cosh \sigma l + \eta \sinh \sigma l}{\eta \cosh \sigma l + Z_z(l) \sinh \sigma l}, \qquad (1\text{–}3)$$

where $Z_z(l)$ is the impedance looking to the right of the plane $z = l$. If the latter plane is a perfect conductor, then $Z_z(l) = 0$ and we have $Z_z(0) = \eta \tanh \sigma l$. If $z = l$ is a sheet of infinite impedance, then

$$Z_z(0) = \eta \coth \sigma l. \qquad (1\text{–}4)$$

As we have already pointed out, in practice an infinite impedance sheet at $z = l$ can be provided by placing a zero impedance sheet at $z = l + \lambda/4$.

If the plate is a good conductor its intrinsic impedance η is very small even at very high frequencies. If the medium to the right of $z = l$ is free space, then $Z_z(l) = 377$. This impedance is so large compared with η that equation (4) represents an excellent approximation to the impedance normal to a plate of high conductivity provided l is not too small. If η is much smaller than $Z_z(l)$, then, regardless of the thickness of the plate, we can ignore the second term in the numerator of (3) and obtain the following approximation

$$Y_z(0) = \frac{1}{\eta \coth \sigma l} + \frac{1}{Z_z(l)}.$$

Thus the input admittance is represented as equivalent to two admittances in parallel, the admittance of the plate on the assumption that $Z_z(l) = \infty$ and the admittance $Y_z(l)$ itself.

When l is very small the "open-circuit" impedance (4) for any quasi-conductor becomes approximately

$$Z_z(0) = \eta\left(\frac{1}{\sigma l} + \tfrac{1}{3}\sigma l + \cdots\right) = \frac{\eta}{\sigma l} = \frac{1}{gl}.$$

This impedance is equal to the free-space impedance if $l = 1/377g$. A sheet of this thickness with a reflector behind it to provide an open-circuit condition will completely absorb a plane wave incident normally to the plate. It should be noted however that for very thin films the value of g is different from that for the substance in bulk.

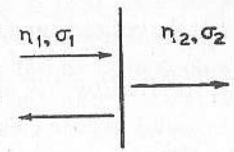


FIG. 8.7. Reflection at normal incidence.

The formulae for the reflection of uniform plane waves from a plane interface between two homogeneous media (Fig. 8.7), when the incidence is normal to the interface, follow immediately from (7.13–3) and (7.13–4); thus we have

$$q_E = \frac{\eta_2 - \eta_1}{\eta_2 + \eta_1}, \quad q_H = \frac{\eta_1 - \eta_2}{\eta_1 + \eta_2}; \quad p_E = \frac{2\eta_2}{\eta_1 + \eta_2}, \quad p_H = \frac{2\eta_1}{\eta_1 + \eta_2}. \qquad (1\text{–}5)$$

At a metal surface the reflection is almost complete, E practically vanishes and H is doubled; almost pure standing waves are formed with nodal planes for E parallel to the metal surface at distances $0, \lambda/2, \lambda, \cdots$ from it and nodal planes for H at distances $\lambda/4, 3\lambda/4, 5\lambda/4, \cdots$. The planes of maximum H coincide with the nodal planes for E and the planes for maximum E are the nodal planes for H.

The shielding effectiveness of metals is great; it can be judged by using (7.6–10) to obtain the ratio of magnetic intensities at the two surfaces of a plate in free space

$$\frac{H(l)}{H(0)} = \frac{\eta}{\eta \cosh \sigma l + 377 \sinh \sigma l}. \qquad (1\text{–}6)$$

Even for quite thin plates $\sinh \sigma l$ is approximately equal to $\frac{1}{2}e^{\sigma l}$ and

$$\frac{H(l)}{H(0)} = \frac{2\eta}{377} e^{-\sigma l}, \quad \left|\frac{H(l)}{H(0)}\right| = \frac{2\sqrt{2\mathcal{R}}}{377} e^{-\alpha l},$$

where $\mathcal{R}$ is the intrinsic resistance of the plate. As the frequency diminishes, η and σ approach zero and the limiting value of (6) is $1/377gl$.

If these formulae are to be applied to frequencies in the optical range one should bear in mind that the electromagnetic "constants" may vary with the frequency. While there is no evidence, for example, that the conductivity of good metals varies in the frequency range $(0, 10^{10})$, there is definite evidence that at optical frequencies it is a complicated function of the frequency. The conductivity of dielectrics is usually variable even at low frequencies.

If the incident waves come from free space and strike water, then $\eta_1 = 377$ and $\eta_2 = 42$; thus from (5) we obtain

$$q_E = -0.8, \quad q_H = 0.8; \quad p_E = 0.2, \quad p_H = 1.8.$$

For waves coming from water the coefficients for E and H are interchanged.

8.2. *Elliptically Polarized Plane Waves*

Waves whose electric and magnetic vectors have directions independent of time,* as for instance the waves given by (1–1) and (1–2), are called *linearly polarized*.†

Let us now consider two linearly polarized progressive waves

$$E_x = E_1 e^{-\sigma z + i\omega t}, \quad H_y = \frac{E_1}{\eta} e^{-\sigma z + i\omega t};$$

$$E_y = E_2 e^{-\sigma z + i(\omega t + \vartheta)}, \quad H_x = -\frac{E_2}{\eta} e^{-\sigma z + i(\omega t + \vartheta)}.$$

We choose the origin of time to make E_1 real; E_2 denotes the amplitude of the second wave and ϑ its initial phase in the plane $z = 0$. The instantaneous electric intensities are the real parts of the above complex quantities; thus in the plane $z = 0$

$$\hat{E}_x = E_1 \cos \omega t, \quad \hat{E}_y = E_2 \cos(\omega t + \vartheta).$$

* Although they change their sense twice during each cycle.

† We discard the term "plane polarized," frequently used in optics, in favor of the term used by radio engineers.

Unless $\vartheta = 0, \pi$ the vector representing the resultant intensity will rotate. Consider for instance the case $\vartheta = -\pi/2$ when the two component intensities are in quadrature; then

$$\hat{E}_x = E_1 \cos \omega t, \quad \hat{E}_y = E_2 \sin \omega t.$$

At $t = 0$, the electric vector is along the positive x-axis; at $t = T/4$, where $T = 1/f$ is the period of oscillations, the vector is along the y-axis; at $t = T/2$ the vector is along the negative x-axis. As the vector rotates, its magnitude changes. The locus of the end point of the vector is found by eliminating t from the above equations

$$\frac{\hat{E}_x^2}{E_1^2} + \frac{\hat{E}_y^2}{E_2^2} = 1.$$

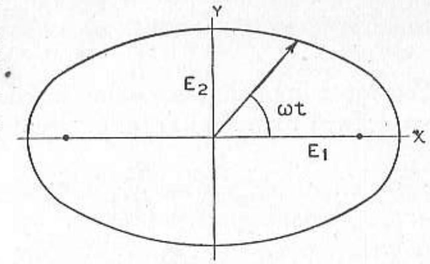


FIG. 8.8. Elliptic polarization.

This equation represents an ellipse (Fig. 8.8) whose semiaxes are E_1 and E_2. The wave is said to be *elliptically polarized*; it is *circularly polarized* if $E_1 = E_2$.

If $\vartheta = \pi/2$, the ellipse of polarization is exactly the same but the vector rotates clockwise instead of counterclockwise. This polarization is said to be *left-handed* as distinct from the *right-handed* polarization in the preceding example. If in the right-handed polarization the electric vector is represented by the handle of a corkscrew, then as the vector rotates the screw advances in the direction of wave propagation. For values of ϑ other than $\pm 90°$, the wave is still elliptically polarized but the axes of the ellipse do not coincide with the coordinate axes.

So far we have considered the electric vector in the plane $z = 0$. For $z > 0$ the amplitudes of both components of E are multiplied by $e^{-\alpha z}$ and the ellipse becomes smaller; the phases of both components are retarded by βz but this simultaneous retardation does not affect the orientation of the ellipse.

The magnetic vector describes another ellipse. In nondissipative media H is evidently perpendicular to E at all times; but in dissipative media this is not the case. When $\vartheta = 0$ or π, the E-ellipse and the H-ellipse degenerate into straight lines and the wave becomes linearly polarized.

8.3. *Wave Impedances at a Point*

In an orthogonal system of coordinates the components of the complex Poynting vector are

$$P_u = \tfrac{1}{2}(E_v H_w^* - E_w H_v^*), \quad P_v = \tfrac{1}{2}(E_w H_u^* - E_u H_w^*),$$
$$P_w = \tfrac{1}{2}(E_u H_v^* - E_v H_u^*).$$

The real part of each component represents the average power per unit area flowing parallel to the corresponding axis. The following ratios are defined as the *wave impedances at a typical point* looking in the directions of increasing coordinates

$$Z^+_{uv} = \frac{E_u}{H_v}, \quad Z^+_{vw} = \frac{E_v}{H_w}, \quad Z^+_{wu} = \frac{E_w}{H_u};$$

$$Z^+_{vu} = -\frac{E_v}{H_u}, \quad Z^+_{wv} = -\frac{E_w}{H_v}, \quad Z^+_{uw} = -\frac{E_u}{H_w}.$$

The wave impedances looking in the directions of decreasing coordinates are defined by a similar set of equations

$$Z^-_{uv} = -\frac{E_u}{H_v}, \quad Z^-_{vw} = -\frac{E_v}{H_w}, \quad Z^-_{wu} = -\frac{E_w}{H_u};$$

$$Z^-_{vu} = \frac{E_v}{H_u}, \quad Z^-_{wv} = \frac{E_w}{H_v}, \quad Z^-_{uw} = \frac{E_u}{H_w}.$$

The u-component of the Poynting vector becomes

$$P_u = \tfrac{1}{2}(Z^+_{wv}H_vH_v^* + Z^+_{vw}H_wH_w^*) = -\tfrac{1}{2}(Z^-_{wv}H_vH_v^* + Z^-_{vw}H_wH_w^*);$$

the remaining components are obtained by cyclic permutations of u, v, w. The algebraic signs in the definitions of the wave impedances have been so chosen that, if the real part of any given impedance is positive, the corresponding average power flow is in the direction of the impedance.

We have seen that the impedance concept plays an important part in transmission theory, but the general formulae of the preceding chapter have been obtained for simple transmission lines having only one impedance in a given direction. At a junction between two simple transmission lines two variables V and I must be continuous and the reflection coefficients depend on their ratio. Transmission theory of this type can be extended to a transmission line with two transmission modes, when there are four variables V_1, I_1 and V_2, I_2 which must satisfy continuity requirements at a junction. The resulting formulae are so complicated that it is doubtful if they would actually save labor in solving problems. At any rate, until a sufficiently large number of problems involving such two-mode transmission lines arises, it is preferable to treat each individual problem by itself, particularly since in many practical problems double mode lines can be approximated by two nearly independent single mode lines. In considering waves in three dimensions the situation is in general vastly more complex. For a general wave the wave impedances are point functions and no advantage is derived from their introduction; one might just as well

deal with the actual field intensities. But let us suppose that two impedances Z_{uv} and Z_{vu}, for example, associated with a given wave, are independent of the u and v coordinates; then in effect we have a double mode transmission line. If at any surface $w = w_0$ the properties of the medium are suddenly altered, the four tangential components have to satisfy continuity conditions at one point only — the continuity conditions elsewhere are automatically satisfied as soon as they are satisfied at this point. The amplitudes of the reflected and transmitted waves will depend on the associated wave impedances. If, furthermore, the two wave impedances in the same direction are equal

$$Z^+_u = \frac{E_v}{H_w} = -\frac{E_w}{H_v}, \quad Z^+_v = \frac{E_w}{H_u} = -\frac{E_u}{H_w}, \quad Z^+_w = \frac{E_u}{H_v} = -\frac{E_v}{H_u},$$

with the corresponding set for the impedances looking in the opposite directions, then the transmission theory of the preceding chapter can be applied in full. E_u, H_v, Z^+_w and E_v, $-H_u$, Z^+_w form right-handed triplets. It is not necessary that all the wave impedances should satisfy these equations. If we are concerned with reflection of waves at the surface $w = w_0$, only Z^+_w and Z^-_w need exist and be independent of the u and v coordinates. Likewise, we are concerned only with Z^+_u and Z^-_u when considering reflection at the surface $u = u_0$.

8.4. *Reflection of Uniform Plane Waves at Oblique Incidence*

In considering reflection of uniform plane waves falling at an arbitrary angle on a plane interface between two homogeneous media (Fig. 8.9) it becomes necessary to distinguish between two orientations of the field vectors: (1) the case in which H is parallel to the interface, (2) the case in which E is parallel to the interface. The impedances normal to the interface are different in the two cases.* If neither E nor H is parallel to the interface, the wave is resolved into two waves, one having the first of the above properties and the other the second. This resolution is always possible since the E-vector for example can be resolved into two components, one parallel to the interface and the other in the *plane of incidence*, that is in the plane determined by the wave normal and the normal to the interface. The second component is along the line of intersection of the equiphase plane and the plane of incidence. The component of H associ-

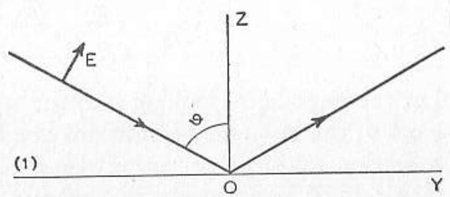


FIG. 8.9. Reflection of uniform plane waves incident obliquely at a plane boundary; H is parallel to the boundary.

* For a general orientation the impedance normal to the interface does not exist.

ated with this component of E is perpendicular to it and to the wave normal; therefore this H is parallel to the interface.

First we shall assume that H is perpendicular to the plane of incidence and hence parallel to the boundary between the two media; in Fig. 8.9 the positive direction of H is in the positive x-direction (toward the reader). The angle ϑ between the wave normal and the normal to the boundary is called the *angle of incidence.* In radio engineering its complement $\frac{1}{2}\pi - \vartheta$ is frequently used. The equations for the incident wave are

$$E = E_0 e^{-\sigma s}, \quad H = H_0 e^{-\sigma s}, \quad E_0 = \eta H_0,$$

where E_0, H_0 are the field intensities at O and s is the distance from O along the wave normal. In cartesian coordinates we have $s = y \sin \vartheta - z \cos \vartheta$, and

$$H_x = H_0 e^{-\sigma s}, \quad E_y = E_0 \cos \vartheta e^{-\sigma s}, \quad E_z = E_0 \sin \vartheta e^{-\sigma s};$$

thus the equations for the incident wave may be written in the following form

$$H_x = H_0 e^{-\Gamma_y y + \Gamma_z z}, \quad E_y = E_0 \cos \vartheta e^{-\Gamma_y y + \Gamma_z z},$$

$$E_z = E_0 \sin \vartheta e^{-\Gamma_y y + \Gamma_z z}; \quad \Gamma_y = \sigma \sin \vartheta, \quad \Gamma_z = \sigma \cos \vartheta. \qquad (4\text{–}1)$$

These equations may be interpreted as the equations of propagation of a *phase-amplitude pattern*, given by $e^{-\Gamma_y y}$, in the negative z-direction, with the propagation constant Γ_z. The impedance in the direction normal to the interface is

$$Z_z^- = \frac{E_y}{H_x} = \frac{E_0 \cos \vartheta}{H_0} = \eta \cos \vartheta.$$

Let the impedance looking into the second medium be Z. If Z is equal to $\eta \cos \vartheta$, the boundary conditions are satisfied by the incident wave and no reflection takes place; otherwise a reflected wave originates at the interface. If Z is constant throughout the interface, the phase amplitude pattern of the reflected wave, in a plane parallel to the interface, must be the same as that of the incident wave or else the resultant wave cannot satisfy the boundary conditions over the entire plane. Thus for the reflected wave we have*

$$H_x^r = H^r e^{-\Gamma_y y - \Gamma_z z}, \quad E_y^r = \frac{1}{g + i\omega\epsilon} \frac{\partial H_x^r}{\partial z} = E_t^r e^{-\Gamma_y y - \Gamma_z z},$$

$$E_z^r = -\frac{1}{g + i\omega\epsilon} \frac{\partial H_x^r}{\partial y} = E_n^r e^{-\Gamma_y y - \Gamma_z z},$$

* See equations (4.12–16) which connect H_x, E_y and E_z.

where the tangential E_t^r and the normal E_n^r are

$$E_t^r = -\frac{\Gamma_z H^r}{g + i\omega\epsilon} = -(\eta \cos \vartheta) H^r, \quad E_n^r = \frac{\Gamma_y H^r}{g + i\omega\epsilon} = (\eta \sin \vartheta) H^r. \qquad (4\text{–}2)$$

That the propagation constant is the same (except for sign) in the positive and negative z-directions follows from equation (4.10–3). The impedance looking in the positive z-direction is

$$Z_z^+ = -\frac{E_y^r}{H_x^r} = \eta \cos \vartheta = Z_z^-;$$

hence by (7.13–3) and (7.13–4) the reflection and transmission coefficients for the tangential components are

$$q_H = \frac{H^r}{H_0} = \frac{\eta \cos \vartheta - Z}{\eta \cos \vartheta + Z}, \quad q_{E_t} = \frac{E_t^r}{E_0 \cos \vartheta} = \frac{Z - \eta \cos \vartheta}{Z + \eta \cos \vartheta},$$

$$p_H = \frac{H^t}{H_0} = \frac{2\eta \cos \vartheta}{\eta \cos \vartheta + Z}, \quad p_{E_t} = \frac{E^t}{E_0 \cos \vartheta} = \frac{2Z}{Z + \eta \cos \vartheta}.$$

For the normal components we have from (1) and (2)

$$q_{E_n} = \frac{E_n^r}{E_0 \sin \vartheta} = \frac{\eta \sin \vartheta H^r}{\eta \sin \vartheta H_0} = q_H, \quad p_{E_n} = 1 + q_{E_n} = p_H.$$

The reflected wave is evidently a uniform plane wave moving in the direction making an angle with the z-axis which is equal to the angle of incidence. This angle is called the *angle of reflection.*

If the xy-plane is a perfect conductor, $Z = 0$ and the magnetic intensity is doubled at the plane. The normal component of E is also doubled but the tangential component is reduced to zero. Except for its direction the total electric vector of the reflected wave is equal to the incident electric vector.

If the medium is nondissipative and if Z is real and less than η, there exists an angle of incidence ϑ_0 for which the impedances are matched

$$Z = \eta \cos \vartheta_0,$$

and there is no reflection. This angle is called the *Brewster angle.* If the absolute value of Z is less than that of η, we can find an angle ϑ_0 for which

$$|Z| = |\eta| \cos \vartheta_0 \quad \text{or} \quad \cos \vartheta_0 = \left|\frac{Z}{\eta}\right|.$$

For this angle the absolute values of the impedances are matched, the amplitude of the reflection coefficient is a minimum, and the phase of the reflection coefficient is $\pm 90°$. This angle is called the "pseudo" Brewster angle; however, we do not find it necessary to distinguish between the two cases and shall refer to either angle as the Brewster angle.

If the medium is nondissipative and if the xy-plane is a perfect conductor, the total magnetic intensity is

$$H_x = H_0 e^{-i\beta y \sin \vartheta} \left(e^{i\beta z \cos \vartheta} + e^{-i\beta z \cos \vartheta}\right) = 2H_0 \cos\left(\beta z \cos \vartheta\right) e^{-i\beta y \sin \vartheta}.$$

The equiphase planes are normal to the xy-plane and they travel parallel to it with the phase velocity $v_y = v/\sin \vartheta$. The components of E are obtained either by adding the incident and the reflected components or directly from (4.12–16); thus

$$E_y = 2i\eta H_0 \cos \vartheta \sin\left(\beta z \cos \vartheta\right) e^{-i\beta y \sin \vartheta},$$
$$E_z = 2\eta H_0 \sin \vartheta \cos\left(\beta z \cos \vartheta\right) e^{-i\beta y \sin \vartheta}.$$

The wave impedances associated with the total wave are

$$Z_y^+ = \eta \sin \vartheta, \quad Z_z^+ = -i\eta \cos \vartheta \tan\left(\beta z \cos \vartheta\right).$$

The impedance looking in the z-direction is imaginary and on the average there is no flow of power in this direction.

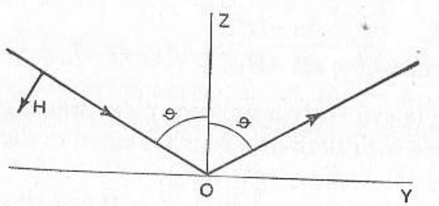


FIG. 8.10. Oblique incidence; E is parallel to the boundary.

If the electric vector is parallel to the xy-plane and the magnetic vector is in the plane of incidence (Fig. 8 10), then for the incident wave we have

$$E_x = E_0 e^{-\Gamma_y y + \Gamma_z z}, \quad H_y = -\frac{E_0}{\eta} \cos \vartheta \, e^{-\Gamma_y y + \Gamma_z z},$$
$$H_z = -\frac{E_0}{\eta} \sin \vartheta \, e^{-\Gamma_y y + \Gamma_z z}, \quad Z_z^- = \eta \sec \vartheta.$$

The impedance associated with the reflected wave is also $\eta \sec \vartheta$. Thus we obtain the following reflection and transmission coefficients for the tangential components of E and H

$$q_E = \frac{Z - \eta \sec \vartheta}{Z + \eta \sec \vartheta}, \quad q_{H_t} = \frac{\eta \sec \vartheta - Z}{\eta \sec \vartheta + Z},$$
$$p_E = \frac{2Z}{Z + \eta \sec \vartheta}, \quad p_{H_t} = \frac{2\eta \sec \vartheta}{\eta \sec \vartheta + Z}.$$

For the normal component of H we have $q_{H_n} = q_E$, $p_{H_n} = p_E$. It is now evident that the reflection coefficient depends on the state of polarization. The impedance Z_z is never greater than η when H is parallel to the xy-plane and it is never smaller than η when E is parallel to the xy-plane. When the angle of incidence ϑ is nearly 90 degrees, the component of E parallel to the xy-plane is very small for the polarization in Fig. 8.9 and hence the impedance is also very small. No matter how small Z may be, for angles sufficiently near 90 degrees the impedance associated with the incident wave will be much smaller than Z so that the total H and the total normal component of E will nearly vanish at the xy-plane, while for most values of ϑ these quantities are nearly doubled. On the other hand for the state of polarization shown in Fig. 8.10, it is the component of H which is small when ϑ is near 90 degrees; the wave impedance is then very large. If Z is smaller than η, then Z is smaller than Z_z for all angles of incidence; and as the angle of incidence increases, reflection only becomes more nearly complete.

The preceding equations apply either to the special case in which the medium below the plane is homogeneous or to the more general case in which the medium below consists of homogeneous layers with their boundaries parallel to the xy-plane. Let us now consider the special case in detail. For the wave below the xy-plane the propagation constant γ_y in the direction parallel to the y-axis must be equal to the corresponding propagation constant in the upper medium or else the tangential E and H cannot possibly be continuous everywhere; thus

$$\gamma_y = \Gamma_y = \sigma \sin \vartheta. \tag{4–3}$$

Since $\gamma_x = \Gamma_x = 0$ and since $\gamma_x^2 + \gamma_y^2 + \gamma_z^2 = \hat{\sigma}^2$, where $\hat{\sigma}$ is the propagation constant characteristic of the lower medium, we have

$$\gamma_z = \sqrt{\hat{\sigma}^2 - \sigma^2 \sin^2 \vartheta}. \tag{4–4}$$

If both media are nondissipative, equation (3) may have another interpretation besides the obvious one that the velocities along the y-axis of the wave above the xy-plane and of that below it are the same. Let us assume that the transmitted wave (or the *refracted* wave in the terminology of optics) is a uniform plane wave and that the *angle of refraction* is $\hat{\vartheta}$ (Fig. 8.11); then just as in the case of the incident wave we have

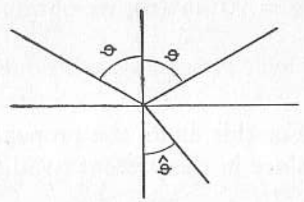


FIG. 8.11. Angles of incidence, reflection, and refraction.

$$\gamma_y = i\hat{\beta} \sin \hat{\vartheta}, \quad \gamma_z = i\hat{\beta} \cos \hat{\vartheta}, \tag{4–5}$$

and equation (3) becomes

$$\beta \sin \vartheta = \hat{\beta} \sin \hat{\vartheta}, \quad \text{or} \quad \frac{\sin \hat{\vartheta}}{\sin \vartheta} = \frac{\beta}{\hat{\beta}} = \frac{\hat{v}}{v} = \frac{\sqrt{\mu\epsilon}}{\sqrt{\hat{\mu}\hat{\epsilon}}}; \tag{4-6}$$

that is, the sines of the angles of incidence and of refraction are proportional to the characteristic phase velocities of the media.

When a wave passes from a medium with higher characteristic velocity into a medium with lower velocity, the equiphase planes tend to become

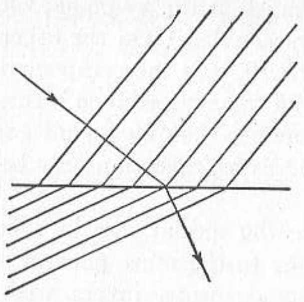


FIG. 8.12. Refraction of waves passing from a medium with high characteristic velocity into a medium with low velocity.

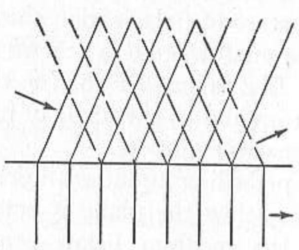


FIG. 8.13. Conditions existing when the angle of incidence is equal to the angle of total internal reflection.

more nearly parallel to the interface (Fig. 8.12); in passing the other way, they tend to become more nearly perpendicular to the interface. This means that if $v < \hat{v}$ or $\sqrt{\mu\epsilon} > \sqrt{\hat{\mu}\hat{\epsilon}}$, there will exist an angle of incidence $\bar{\vartheta}$ for which the angle of refraction is equal to 90 degrees and the equiphase planes in the lower medium are normal to the interface (Fig. 8.13). This critical angle of incidence is called the angle of *total internal reflection* because, as we shall presently see, the wave is completely reflected. Setting $\hat{\vartheta} = 90°$ in (6), we obtain

$$\sin \bar{\vartheta} = \frac{\hat{\beta}}{\beta} = \frac{v}{\hat{v}} = \frac{\sqrt{\hat{\mu}\hat{\epsilon}}}{\sqrt{\mu\epsilon}}.$$

For this angle the propagation constant γ_z in the lower medium vanishes since in the present nondissipative case we have

$$\gamma_z = i\hat{\beta} \cos \hat{\vartheta}, \quad \cos \hat{\vartheta} = \sqrt{1 - \frac{\beta^2}{\hat{\beta}^2} \sin^2 \vartheta}.$$

For $\vartheta < \bar{\vartheta}$, γ_z is imaginary and the angle of refraction is real; but for $\vartheta > \bar{\vartheta}$, γ_z becomes real and the equiphase planes become normal to the

interface. The field in the lower medium is then attenuated exponentially with the distance from the interface, which indicates that the average flow of power across the interface is zero. Since the free space wave velocity is higher than the wave velocity in any other dielectric, the phenomenon of total internal reflection can always occur at a boundary between free space and a dielectric when the incident wave is in the dielectric. In the case of water and free space $\beta = 9\hat{\beta}$ and $\bar{\vartheta} = 6° 23'$. When ϑ is sufficiently greater than $\bar{\vartheta}$, γ_z is given approximately by

$$\gamma_z = \beta \sin \vartheta = \frac{2\pi \sin \vartheta}{\lambda}, \quad \gamma_z \lambda = 2\pi \sin \vartheta.$$

For short waves the attenuation becomes substantial.

We have seen that if the incident wave is uniform, the reflected wave is also uniform; on the other hand, the transmitted wave is not necessarily uniform even in the special case of nondissipative media.

The foregoing properties of transmitted waves are independent of the state of polarization. This state has to be specified if the values of the transmission and reflection coefficients are sought. Let us start with the case in which *H is parallel to the boundary*. Inasmuch as the wave is generally nonuniform we shall write our equations in terms of the propagation constants γ_y and γ_z and use the "angle" of refraction $\hat{\vartheta}$ only for the sake of attaining formal symmetry in the results. We simply define the complex angle $\hat{\vartheta}$ by the following equations

$$\gamma_y = \hat{\sigma} \sin \hat{\vartheta}, \quad \gamma_z = \hat{\sigma} \cos \hat{\vartheta},$$

whereas the propagation constants themselves are given by (3) and (4) in terms of known quantities. For the transmitted wave we then have

$$H_x^t = H^t e^{-\gamma_y y + \gamma_z z}, \quad E_y = \frac{\gamma_z}{\hat{g} + i\omega\hat{\epsilon}} H_x^t, \quad E_z = \frac{\gamma_y}{\hat{g} + i\omega\hat{\epsilon}} H_x^t.$$

Hence we have

$$\hat{Z}_z^- = \frac{\gamma_z}{\hat{g} + i\omega\hat{\epsilon}} = \hat{\eta} \cos \hat{\vartheta}, \quad \hat{Z}_y^+ = \frac{\gamma_y}{\hat{g} + i\omega\hat{\epsilon}} = \hat{\eta} \sin \hat{\vartheta}.$$

These expressions are of the same form as for uniform waves except that $\hat{\vartheta}$ is no longer real. We now let $Z = \hat{Z}_z^-$ and obtain

$$q_H = \frac{\eta \cos \vartheta - \hat{\eta} \cos \hat{\vartheta}}{\eta \cos \vartheta + \hat{\eta} \cos \hat{\vartheta}} = \frac{1 - k}{1 + k}, \quad q_{E_t} = -q_H,$$

$$p_H = \frac{2\eta \cos \vartheta}{\eta \cos \vartheta + \hat{\eta} \cos \hat{\vartheta}} = \frac{2}{1 + k}, \quad q_{E_n} = q_H, \tag{4-7}$$

$$p_{E_t} = \frac{2\hat{\eta} \cos \hat{\vartheta}}{\eta \cos \vartheta + \hat{\eta} \cos \hat{\vartheta}} = \frac{2k}{1 + k}, \quad p_{E_n} = p_H,$$

where k is the following impedance ratio

$$k = \frac{\hat{\eta} \cos \hat{\vartheta}}{\eta \cos \vartheta}.$$

From (3) we have

$$\sigma \sin \vartheta = \hat{\sigma} \sin \hat{\vartheta}, \quad \frac{\sigma}{\hat{\sigma}} = \frac{\sin \hat{\vartheta}}{\sin \vartheta}.$$

If both media are nonmagnetic (or, more generally, if they have the same permeabilities), then

$$\frac{\hat{\eta}}{\eta} = \frac{\sigma}{\hat{\sigma}} = \frac{\sin \hat{\vartheta}}{\sin \vartheta}.$$

Substituting in the above equations, we have

$$q_H = \frac{\sin 2\vartheta - \sin 2\hat{\vartheta}}{\sin 2\vartheta + \sin 2\hat{\vartheta}} = \frac{\sin(\vartheta - \hat{\vartheta}) \cos(\vartheta + \hat{\vartheta})}{\cos(\vartheta - \hat{\vartheta}) \sin(\vartheta + \hat{\vartheta})} = \frac{\tan(\vartheta - \hat{\vartheta})}{\tan(\vartheta + \hat{\vartheta})},$$

$$p_H = \frac{2 \sin 2\vartheta}{\sin 2\vartheta + \sin 2\hat{\vartheta}} = \frac{\sin 2\vartheta}{\cos(\vartheta - \hat{\vartheta}) \sin(\vartheta + \hat{\vartheta})},$$

$$p_{E_t} = \frac{2 \sin 2\hat{\vartheta}}{\sin 2\vartheta + \sin 2\hat{\vartheta}} = \frac{\sin 2\hat{\vartheta}}{\cos(\vartheta - \hat{\vartheta}) \sin(\vartheta + \hat{\vartheta})}.$$

If the *E-vector is parallel to the interface* between the two media, then using (4.12–15) we have

$$E_x^t = E^t e^{-\gamma_y y + \gamma_z z}, \quad H_y^t = -\frac{\gamma_z}{i\omega\hat{\mu}} E_x^t, \quad H_z^t = -\frac{\gamma_y}{i\omega\hat{\mu}} E_x^t;$$

$$\hat{Z}_z^- = \frac{i\omega\hat{\mu}}{\gamma_z} = \hat{\eta} \sec \hat{\vartheta}, \quad \hat{Z}_y^+ = \frac{i\omega\hat{\mu}}{\gamma_y} = \hat{\eta} \csc \hat{\vartheta},$$

and the impedance ratio becomes

$$k = \frac{\hat{\eta} \sec \hat{\vartheta}}{\eta \sec \vartheta} = \frac{\hat{\eta} \cos \vartheta}{\eta \cos \hat{\vartheta}}.$$

Expressions (7) for the reflection and transmission coefficients in terms of the impedance ratio k are, of course, independent of the state of polarization. For *nonmagnetic media* the above equation becomes $k = \tan \hat{\vartheta} \cot \vartheta$ and the coefficients assume the following special form

$$q_E = \frac{\sin(\hat{\vartheta} - \vartheta)}{\sin(\hat{\vartheta} + \vartheta)}, \quad p_E = \frac{2 \sin \hat{\vartheta} \cos \vartheta}{\sin(\hat{\vartheta} + \vartheta)}, \quad p_{H_t} = \frac{2 \sin \vartheta \cos \hat{\vartheta}}{\sin(\hat{\vartheta} + \vartheta)}.$$

The expressions involving the angle of refraction $\hat{\vartheta}$ are convenient if

this angle is real; otherwise in actual computations it is best to eliminate it with the aid of the defining equation

$$\cos \hat{\vartheta} = \sqrt{1 - \frac{\sigma^2}{\hat{\sigma}^2} \sin^2 \vartheta}.$$

In the case of wave propagation over a plane ground the absolute value of σ is considerably smaller than the absolute value of $\hat{\sigma}$ and the above equation becomes approximately

$$\cos \hat{\vartheta} \simeq 1 - \frac{\sigma^2}{2\hat{\sigma}^2} \sin^2 \vartheta \simeq 1.$$

The accuracy of this approximation may be gauged by expressing the ratio of the squares of the propagation constants in the following form

$$\frac{\hat{\sigma}^2}{\sigma^2} = \hat{\mu}_r \hat{\epsilon}_r - i60\hat{\mu}_r \hat{g}\lambda = \hat{\mu}_r \hat{\epsilon}_r \left(1 - \frac{i}{Q}\right),$$

where $\hat{\mu}_r$, $\hat{\epsilon}_r$ are the permeability and the dielectric constant of ground relative to free space and Q is the Q of the ground. For the actual earth $\hat{\epsilon}_r$ is often greater than 10 and then $\cos \hat{\vartheta}$ will not deviate from unity by more than 5 per cent and the impedance looking into ground is substantially $\hat{\eta}$.

The intrinsic impedance of the ground is usually smaller than that of free space and the ratio of the two impedances is

$$\frac{\hat{\eta}}{\eta} = \frac{\sigma \hat{\mu}_r}{\hat{\sigma}} = \sqrt{\frac{\hat{\mu}_r}{\hat{\epsilon}_r}} \left(1 - \frac{i}{Q}\right)^{-1/2}.$$

It is only if the relative permeability of the ground is higher than the relative dielectric constant — an unlikely case — that this impedance ratio may be greater than unity.

When the E-vector is parallel to the ground the ratio of the impedances normal to the ground is approximately

$$k = \frac{\hat{\eta}}{\eta} \cos \vartheta.$$

The amplitude of k is substantially less than unity even at normal incidence and becomes very small indeed as ϑ approaches 90 degrees; hence as the angle of incidence increases the amplitude of the E-reflection coefficient steadily approaches -1.

On the other hand if the H-vector is parallel to the ground, the impedance ratio is

$$k = \frac{\hat{\eta}}{\eta} \sec \vartheta.$$

When ϑ is small, the amplitude of k is less than unity; but near grazing incidence it becomes very large. The amplitude of k is unity when the angle of incidence is approximately

$$\cos \vartheta_0 = \left| \frac{\hat{\eta}}{\eta} \right| = \hat{\mu}_r[(\hat{\mu}_r\hat{\epsilon}_r)^2 + (60\hat{\mu}_r\hat{g}\lambda)^2]^{-1/4}.$$

This is the Brewster angle; at this angle of incidence the amplitude of the reflection coefficient is minimum and the phase is $\pm 90°$. At this angle if the ground were nondissipative the wave with the H-vector parallel to the reflecting plane would not be reflected at all. For high angle waves ($\vartheta \ll \vartheta_0$), k is well inside the unit circle and H as well as the normal component of E may be almost doubled; but for low angle waves ($\vartheta \gg \vartheta_0$), k is well outside the unit circle and these components are nearly annihilated. In the latter case the tangential component of E is nearly doubled; but this component is small to begin with. Thus near grazing incidence the entire field at the ground is nearly annihilated by the reflected wave.

8.5. *Uniform Cylindrical Waves*

A wave is *cylindrical* if its equiphase surfaces form a family of coaxial cylinders; it is uniform if the amplitude is the same at all points of a given equiphase surface. Choosing the axis of such waves as the z-axis, and assuming $\partial/\partial\varphi = 0$, $\partial/\partial z = 0$ in the general equations, we have $E_\rho = 0$, $H_\rho = 0$, and

$$\frac{dE_z}{d\rho} = i\omega\mu H_\varphi, \quad \frac{d}{d\rho}(\rho H_\varphi) = (g + i\omega\epsilon)\rho E_z; \tag{5–1}$$

$$\frac{d}{d\rho}(\rho E_\varphi) = -i\omega\mu\rho H_z, \quad \frac{dH_z}{d\rho} = -(g + i\omega\epsilon)E_\varphi. \tag{5–2}$$

Thus uniform cylindrical waves are transverse electromagnetic and they may be of two types: (1) waves with the E-vector parallel to the axis, (2) waves with the H-vector parallel to the axis.

If H_φ is eliminated from (1), we obtain

$$\rho \frac{d^2E_z}{d\rho^2} + \frac{dE_z}{d\rho} - \sigma^2\rho E_z = 0.$$

This is the modified Bessel equation of order zero, with the independent variable $\sigma\rho$, and it has two independent solutions

$$E_z^+(\rho) = AK_0(\sigma\rho), \quad E_z^-(\rho) = BI_0(\sigma\rho). \tag{5–3}$$

By (3.4–8) we have at great distances

$$E_z^+(\rho) = A\sqrt{\frac{\pi}{2\sigma\rho}}\, e^{-\sigma\rho};$$

this function vanishes at infinity while the second solution becomes exponentially infinite. Thus $E_z^+(\rho)$ represents an outward bound wave. For large values of ρ it is very similar to a plane wave except that the amplitude is steadily decreasing, as indicated by the factor $\rho^{-1/2}$. While the first solution is infinite at $\rho = 0$, the second is finite for all finite values of ρ; hence it is appropriate for source-free regions for which $\rho < a$.

The H-wave functions corresponding to (3) are obtained immediately from (1); thus

$$H_\varphi^+(\rho) = -\frac{A}{\eta} K_1(\sigma\rho), \quad H_\varphi^-(\rho) = \frac{B}{\eta} I_1(\sigma\rho). \tag{5–4}$$

From (3) and (4) we obtain the radial impedances

$$Z_\rho^+ = \eta \frac{K_0(\sigma\rho)}{K_1(\sigma\rho)}, \quad Z_\rho^- = \eta \frac{I_0(\sigma\rho)}{I_1(\sigma\rho)}. \tag{5–5}$$

Equations (1) apply either to cylindrical waves in an unlimited medium or to waves between two perfectly conducting planes perpendicular to the E-vector. Let one of these planes be $z = 0$ and the other $z = h$. If V is the transverse voltage from the lower plane to the upper (Fig. 8.14) at a distance ρ from the cylindrical axis and if I is the total radial current in the lower plane, then $V = hE_z$, $I = -2\pi\rho H_\varphi$. The second equation can be derived in several ways. Thus the outward radial current $I(\rho)$ is equal to the downward transverse current inside the cylinder of radius ρ; since this transverse current is equal to the magnetomotive force, we have the desired equation. Substituting in (1), we have

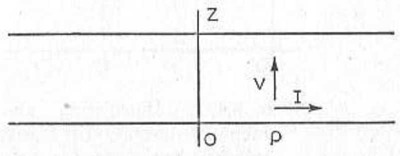


FIG. 8.14. Two parallel conducting planes supporting uniform cylindrical waves whose axis is OZ.

$$\frac{dV}{d\rho} = -i\omega LI, \quad \frac{dI}{d\rho} = -(G + i\omega C)V,$$

where the distributed constants per unit length of the "disc transmission line" are

$$L = \frac{\mu h}{2\pi\rho}, \quad G = \frac{2\pi g\rho}{h}, \quad C = \frac{2\pi\epsilon\rho}{h}.$$

Since the electric lines are straight lines normal to the two planes, we could have obtained G and C directly by considering the conductance and the capacitance between annular rings of width $d\rho$, one in each plane, and dividing the result by $d\rho$. The inductance per unit length along a radius could be similarly obtained from the magnetic flux passing through the rectangle $ABCD$ shown in Fig. 8.15 and letting $AB = d\rho$.

For the wave impedances of the " disc transmission line " we have

$$K^+ = \frac{V^+(\rho)}{I^+(\rho)} = -\frac{hE_z^+}{2\pi\rho H_\varphi^+} = \frac{h}{2\pi\rho} Z_\rho^+, \quad K^- = \frac{h}{2\pi\rho} Z_\rho^-.$$

The complex power carried by a progressive wave traveling outward is then $\Psi^+ = \frac{1}{2}K^+II^*$. The wave E_z^-, H_φ^- is not progressive; it is strictly stationary when the medium is nondissipative. If the medium is homogeneous within the cylinder of radius ρ, the radial current in the planes must vanish at $\rho = 0$; hence the disc line must behave as electrically open at $\rho = 0$ and the energy will be completely reflected. Some energy will travel

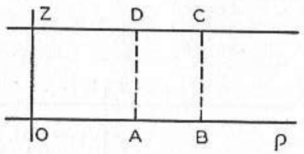


FIG. 8.15. Illustrating elementary derivation of transmission equations for uniform cylindrical waves.

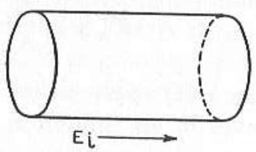


FIG. 8.16. A section of an infinitely long wire.

inward only if the medium is dissipative or at least if a dissipative wire connects the two planes along the axis. The wave functions (3), each corresponding to a homogeneous region seem to be more suitable for practical purposes than other possible sets. If the region is homogeneous and source-free only between two cylindrical surfaces $\rho = a$ and $\rho = b$, then the field is expressed in terms of both wave functions. The complex power flow in the stationary wave is $\Psi^- = \frac{1}{2}K^-II^*$. These expressions for the complex power represent the total power flow across a cylindrical surface between two parallel planes. The radial flow per unit area depends on the radial impedances (5); thus

$$\Psi^+ = \tfrac{1}{2}Z_\rho^+H_\varphi^+H_\varphi^{+*}, \quad \Psi^- = \tfrac{1}{2}Z_\rho^-H_\varphi^-H_\varphi^{-*}.$$

There is another aspect to K^+ and K^-. Consider an infinitely long wire and let an electric intensity E_i be applied uniformly round the surface of the wire (Fig. 8.16). Let I be the current in the wire in response to E_i. If $E^-(\rho)$ and $E^+(\rho)$ are respectively the field intensities in the wire and outside it, then

$$E_i = E^-(a) - E^+(a).$$

The intensity driving the return current, external to the wire, acts of course

in the direction opposite to the intensity driving the current in the wire. Dividing by I, we have

$$\frac{E_i}{I} = \frac{E^-(a)}{I} + \frac{-E^+(a)}{I} = Z_i + Z_e, \qquad (5\text{–}6)$$

where Z_i and Z_e are respectively the internal and the external impedances of the wire per unit length; they are equal respectively to K^-/h and K^+/h.

We shall now consider the numerical magnitudes of the impedances under various conditions. If the frequency is so low that $|\sigma a| \ll 1$, then we take the first two terms of the power series for the I-functions and at $\rho = a$ we have approximately

$$Z_\rho^- = \frac{2}{(g + i\omega\epsilon)a} + \tfrac{1}{4}i\omega\mu a, \quad K^- = \frac{h}{(g + i\omega\epsilon)\pi a^2} + \frac{i\omega\mu h}{8\pi}.$$

If $\epsilon = 0$ or if $\omega\epsilon$ can be neglected in comparison with g, then

$$K^- = R + i\omega L, \quad R = \frac{h}{g\pi a^2}, \quad L = \frac{\mu h}{8\pi}.$$

Thus we have the low frequency resistance and internal inductance of a wire of radius a and of length h. If on the other hand $g = 0$, then

$$K^- = \frac{1}{i\omega C} + i\omega L, \quad C = \frac{\epsilon\pi a^2}{h}, \qquad (5\text{–}7)$$

and we have an expression for the low frequency capacitance of a capacitor formed by two parallel circular discs of radius a separated by distance h. Inasmuch as these expressions have been obtained on the assumption that the electric lines are normal to the metal discs, in practical applications of (7) we must assume that h is small compared with a; then the formula furnishes an approximate value for the " internal capacitance " of the two discs. The external capacitance between the outer surfaces of the discs is more difficult to calculate; but it is, of course, considerably smaller than the internal capacitance.

When $|\sigma a| \ll 1$, the outward looking impedances (in a nondissipative medium) become

$$Z_\rho^+ = i\omega\mu a \left(\log \frac{\lambda}{2\pi a} + 0.116\right) + \frac{\pi^2 a\eta}{\lambda},$$

$$K^+ = \frac{\pi\eta h}{2\lambda} + \frac{i\omega\mu h}{2\pi}\left(\log \frac{\lambda}{2\pi a} + 0.116\right).$$

The external inductance depends on the frequency and becomes infinite at $f = 0$; but the external impedance vanishes at $f = 0$.

At high frequencies in good conductors we obtain from the asymptotic expansions of the modified Bessel functions the following expressions

$$Z_\rho^+ = \eta\left(1 - \frac{1}{2\sigma a}\right) = \eta - \frac{1}{2ga}, \quad Z_\rho^- = \eta\left(1 + \frac{1}{2\sigma a}\right) = \eta + \frac{1}{2ga},$$

$$K^+ = \frac{\eta h}{2\pi a} - \frac{h}{4g\pi a^2}, \quad K^- = \frac{\eta h}{2\pi a} + \frac{h}{4g\pi a^2}.$$

Since in this case $\eta = \mathcal{R}(1 + i)$, the high frequency resistances of a wire of radius a and of length h and of a metallic conductor extending to infinity in the radial direction are respectively

$$R = \frac{\mathcal{R}h}{2\pi a} + \frac{h}{4g\pi a^2}, \quad R = \frac{\mathcal{R}h}{2\pi a} - \frac{h}{4\pi g a^2}.$$

The inductive reactances are equal in magnitude to the first terms in these formulae.

The exact expression for the internal impedance per unit length of a conducting wire is

$$Z_i = \frac{Z_\rho^-}{2\pi a} = \frac{\eta I_0(\sigma a)}{2\pi a I_1(\sigma a)}. \tag{5–8}$$

The phase of σ is 45° and, in order to separate the real and imaginary parts, the following auxiliary functions are introduced

$$I_0(u\sqrt{i}) = \text{ber } u + i \text{ bei } u.$$

The power series for these functions can be readily obtained from the power series for the I-functions; thus

$$\text{ber } u = \sum_{n=0}^{\infty} \frac{(-)^n u^{4n}}{2^{4n}[(2n)!]^2}, \quad \text{bei } u = \sum_{n=0}^{\infty} \frac{(-)^n u^{4n+2}}{2^{4n+2}[(2n+1)!]^2}.$$

Hence if we let $u = a\sqrt{\omega\mu g}$ in equation (8) and separate the real and imaginary parts, we obtain

$$\frac{Z_i(f)}{Z_i(0)} = \frac{u[\text{ber } u \text{ bei}' u - \text{bei } u \text{ ber}' u]}{2[(\text{ber}' u)^2 + (\text{bei}' u)^2]} + i\,\frac{u[\text{ber } u \text{ ber}' u + \text{bei } u \text{ bei}' u]}{2[(\text{ber}' u)^2 + (\text{bei}' u)^2]},$$

where $Z_i(0)$ is the d-c resistance of the wire. The ratio of a-c to d-c resistance is represented by the solid curve in Fig. 8.17; the dotted curve represents the ratio of a-c reactance to d-c resistance.

Let us consider the field external to an electric current filament of radius a. For the magnetic intensity at distance ρ we have

$$H_\varphi(\rho) = \frac{IK_1(\sigma\rho)}{2\pi a K_1(\sigma a)} \simeq \frac{\sigma I}{2\pi} K_1(\sigma\rho).$$

The electric intensity is

$$E_z(\rho) = -Z_\rho^+(\rho)H_\varphi(\rho) = -\frac{\eta IK_0(\sigma\rho)}{2\pi a K_1(\sigma a)} \simeq -\frac{i\omega\mu I}{2\pi} K_0(\sigma\rho).$$

When $\sigma\rho$ is small, then the field is approximately

$$H_\varphi(\rho) = \frac{I}{2\pi\rho}, \quad E_z(\rho) = \frac{i\omega\mu I}{2\pi}(\log \sigma\rho - 0.116);$$

and in nondissipative media we have

$$E_z(\rho) = -\tfrac{1}{4}\omega\mu I - \frac{i\omega\mu I}{2\pi}\left(\log\frac{\lambda}{2\pi\rho} + 0.116\right).$$

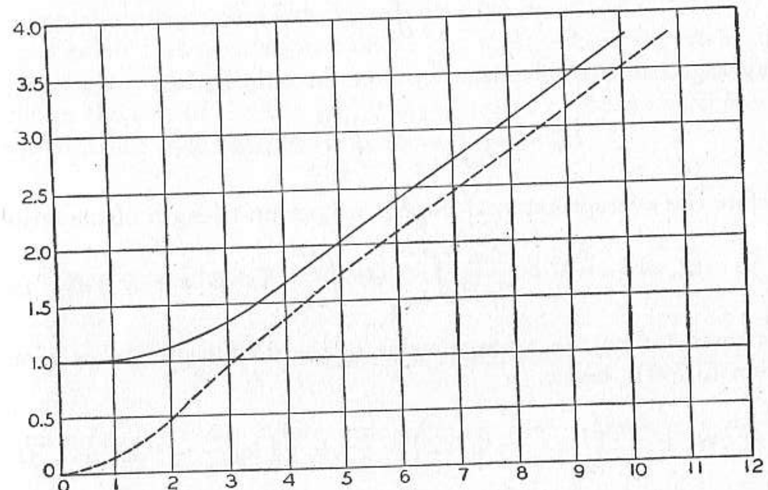


FIG. 8.17. "Skin effect" in cylindrical wires. The solid curve represents the ratio of the a-c resistance of the wire to the d-c resistance and the dotted curve the ratio of the a-c reactance to the d-c resistance. The parameter $u = a\sqrt{\omega\mu g}$.

Thus the magnetic intensity is in phase with I and it varies inversely as the distance from the axis of the wire. The electric intensity has a constant component 180 degrees out of phase with the current and a component in quadrature with I which varies logarithmically with the distance. At large distances in nondissipative media the amplitudes of both field intensities vary inversely as the square root of ρ (see 3.4–8)

$$E_z(\rho) = -\frac{i\omega\mu I}{2\pi}\sqrt{\frac{\pi}{2i\beta\rho}}\,e^{-i\beta\rho} = \frac{\eta I e^{-i\beta\rho - 3i\pi/4}}{2\sqrt{\lambda\rho}},$$

$$H_\varphi(\rho) = -\frac{E_z(\rho)}{\eta} = \frac{Ie^{-i\beta\rho + i\pi/4}}{2\sqrt{\lambda\rho}}. \tag{5–9}$$

Let us now consider an infinitely long and infinitely thin plane strip of width s, carrying a uniformly distributed current I. Assume that the strip is in the xz-plane and that its axis coincides with the z-axis. The density of the electric current is $J_z = I/s$. Let $E_z(x,\hat{x})$ be the intensity produced by an infinitely thin filament passing through the point $(\hat{x},0,0)$ along the line $(x,0,z)$; then

$$E_z(x,\hat{x}) = -\frac{i\omega\mu}{2\pi} K_0(i\beta\,|\,x - \hat{x}\,|).$$

The intensity $E_z(x)$ due to the current in the entire strip will be

$$E_z(x) = \frac{I}{s}\int_{-s/2}^{s/2} E_z(x,\hat{x})\,d\hat{x}.$$

The average value of this intensity over the strip is then

$$E_{\text{av}} = \frac{I}{s^2}\int_{-s/2}^{s/2} dx \int_{-s/2}^{s/2} E_z(x,\hat{x})\,d\hat{x}. \qquad (5\text{–}10)$$

Therefore the average external impedance per unit length of the strip is

$$Z_e = -\frac{E_{\text{av}}}{I} = \frac{i\omega\mu}{2\pi s^2}\int_{-s/2}^{s/2} dx \int_{-s/2}^{s/2} K_0(i\beta\,|\,x - \hat{x}\,|)\,d\hat{x}. \qquad (5\text{–}11)$$

If s is small,* then the average value of the double integral is given by equation (3.7–8); hence

$$Z_e = \tfrac{1}{4}\omega\mu + \frac{i\omega\mu}{2\pi}\left(\log\frac{\lambda}{s} - 0.222\right) = \frac{\pi\eta}{2\lambda} + \frac{i\eta}{\lambda}\left(\log\frac{\lambda}{s} - 0.222\right). \qquad (5\text{–}12)$$

If now a plane wave strikes a perfectly conducting strip of the above dimensions, the approximate current in the strip will be $I = E_0/Z_e$. The distant field of this current is then obtained from (9). The electric intensity, for example, is

$$E_z^r(\rho) = \frac{E_0 e^{-i\beta\rho - 3i\pi/4}}{\pi + 2i\left(\log\frac{\lambda}{s} - 0.222\right)}\sqrt{\frac{\lambda}{\rho}},$$

where E_0 is the intensity of the incident wave. This is the field scattered by a narrow perfectly conducting strip. In optics this field is called the field *diffracted* by the strip. If the strip is not perfectly conducting we should add its internal impedance to Z_e and then compute the induced current.

* Remembering that this means βs is small compared with unity.

8.6 *Cylindrical Cavity Resonators*

Consider a perfectly conducting cylindrical box of radius a (Fig. 8.18) and assume that the medium inside is nondissipative. An electric disturbance, once started inside this box, will continue indefinitely since no energy can escape through the conducting walls. Thus there may exist free oscillations similar to those in a simple circuit containing an inductor and a capacitor or to those in a transmission line short-circuited at both ends. Even in the latter case there are infinitely many oscillation modes and corresponding natural frequencies; the box, having three dimensions, may be expected to have a triple infinity of oscillation modes. In this section we shall confine our attention to the particular oscillation modes in which the E-vector is parallel to the axis of the box and is independent of the φ-coordinate.

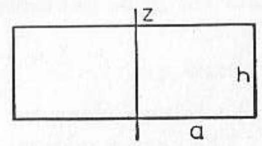


FIG. 8.18. A cross-section of a cylindrical cavity resonator.

In accordance with the preceding section we have

$$E_z = EJ_0(\beta\rho), \quad H_\varphi = \frac{iE}{\eta} J_1(\beta\rho).$$

Since E_z must vanish on the boundary $\rho = a$, βa must be a zero of $J_0(x)$ and

$$\beta a = \frac{2\pi a}{\lambda} = 2.40,\ 5.52,\ 8.65,\ 11.79, \cdots.$$

The consecutive values differ approximately by π. For the mode corresponding to the lowest natural frequency we have

$$d = 2a = \frac{2.40}{\pi}\lambda = 0.764\lambda, \quad \lambda = 1.31d,$$

where λ is the wavelength characteristic of the medium. The corresponding frequency f is then $v/\lambda = 1/\lambda\sqrt{\mu\epsilon}$.

On the axis the electric intensity has the greatest amplitude and the magnetic intensity is zero at all times. In general E_z and H_φ are in quadrature. The charge density on the bottom face of the cavity is $q_S = \epsilon EJ_0(\beta\rho)$, and the total charge is

$$q = 2\pi\epsilon E\int_0^a \rho J_0(\beta\rho)\,d\rho = \frac{2J_1(\beta a)}{\beta a}\,\epsilon\pi a^2 E.$$

For the lowest mode this becomes $q = 0.433\epsilon\pi a^2 E$. On the top face we have an equal but opposite charge and there is no charge on the cylinder.

The charge fluctuates between the top and bottom and the electric current in the cylinder is

$$I = -2\pi a H_\varphi(a) = -i\frac{2\pi a J_1(\beta a)E}{\eta} = -\frac{1.25i\lambda E}{\eta}.$$

For the principal mode (the lowest) in an empty cavity we have

$$I = -i8.65aE = -i3.30\lambda E \text{ milliamperes.}$$

The voltage along the axis is, of course, $V = hE$.

The energy content may be calculated by integrating $\frac{1}{2}\epsilon E_z^2$; thus

$$W = 0.135\epsilon\pi a^2 hE^2 = 0.135\frac{\epsilon\pi a^2}{h}V^2 = \frac{\mu h}{8\pi}|I|^2.$$

For a given E the energy content is proportional to the volume of the cavity. Incidentally if we were to use $E/2$ as the average value of E_z and then assume E_z to be uniform, we should find the numerical factor 0.125 instead of 0.135.

We have seen that the maximum electric intensity is along the z-axis and, for the principal mode, it diminishes steadily and vanishes at $\rho = a$. The magnetic intensity vanishes on the axis and then increases; its maximum, however, is not at the surface of the cylinder. It is the magnetomotive force that reaches a maximum there since the vertical displacement current increases steadily with the radius. The maximum H_φ occurs at a distance ρ_1 defined by $J_1'(\beta\rho_1) = 0$. The first root of J_1' is 1.84 and therefore

$$\beta\rho_1 = 1.84, \quad \frac{\rho_1}{a} = \frac{1.84}{2.40} = 0.766.$$

If the walls of the cavity are not quite perfectly conducting, the above results become first approximations. For good conductors the tangential component of E is small but not zero and the E/H ratio is equal to the intrinsic impedance $\hat{\eta}$ of the metal walls. The principal correction consists in taking account of the energy absorption by the walls since the reactive flow of energy in and out of the walls will constitute a negligible part of the total energy of the cavity. The power $\hat{\mathcal{W}}_1$ absorbed by the two flat faces of the cavity is

$$\hat{\mathcal{W}}_1 = \mathcal{R}\int_0^{2\pi}\int_0^a H_\varphi H_\varphi^* \rho\, d\rho\, d\varphi = \frac{[2.40J_1(2.40)]^2\lambda^2\mathcal{R}V^2}{4\pi h^2\eta^2} = \frac{1}{4\pi}\mathcal{R}I^2.$$

The power $\hat{\mathcal{W}}_2$ absorbed by the cylinder can be obtained without integration since the current in the cylinder is uniform and the resistance of the

cylinder is $\mathcal{R}h/2\pi a$; thus

$$\hat{\mathcal{W}}_2 = \frac{\mathcal{R}hI^2}{4\pi a}.$$

The total power loss is then

$$\hat{\mathcal{W}} = \frac{\mathcal{R}}{4\pi}\left(1 + \frac{h}{a}\right)I^2 \text{ watts, or } \hat{\mathcal{W}} = 0.872\mathcal{R}V^2\left(\frac{\lambda^2}{h^2} + 2.62\frac{\lambda}{h}\right) \text{ microwatts.}$$

Hence for the Q of the cavity we have

$$Q = \frac{\omega W}{\hat{\mathcal{W}}} = \frac{\omega\mu a}{2\mathcal{R}\left(1 + \frac{a}{h}\right)} = \frac{1.20\eta}{\mathcal{R}\left(1 + \frac{a}{h}\right)}.$$

In dealing with uniform cylindrical waves in nondissipative media bounded by two cylinders $\rho = a$ and $\rho = b$, where $b > a$, it is more convenient to employ the following wave functions

$$E_z^-(\rho) = J_0(\beta\rho), \quad E_z^+(\rho) = N_0(\beta\rho);$$
$$\eta H_\varphi^-(\rho) = iJ_1(\beta\rho), \quad \eta H_\varphi^+(\rho) = iN_1(\beta\rho).$$

The K-function which is more suitable for waves traveling to infinity is now replaced by the N-function which represents a stationary wave with a singularity at $\rho = 0$. The radial impedance looking from $\rho = a$ to $\rho = b$ may be obtained from (7.10–8); thus for a perfect conductor at $\rho = b$ we have

$$Z_\rho(a) = i\eta\frac{J_0(\beta a)N_0(\beta b) - N_0(\beta a)J_0(\beta b)}{J_1(\beta a)N_0(\beta b) - N_1(\beta a)J_0(\beta b)}.$$

This impedance either vanishes or becomes infinite, according as

$$\frac{J_0(\beta a)}{N_0(\beta a)} = \frac{J_0(\beta b)}{N_0(\beta b)}, \quad \text{or} \quad \frac{J_1(\beta a)}{N_1(\beta a)} = \frac{J_0(\beta b)}{N_0(\beta b)}. \qquad (6\text{–}1)$$

The first case corresponds to the natural oscillations when there is another perfectly conducting cylinder at $\rho = a$ (Fig. 8.19) so that we have a toroidal cavity. The second case corresponds to a screen of infinite impedance at $\rho = a$; it approximates a cavity with a small hole through the center of each of its flat faces (Fig. 8.20).

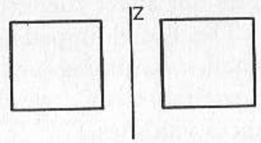


FIG. 8.19. A toroidal cavity bounded by two coaxial cylinders and two parallel planes.

When a and b are large, the roots of the above equations are easy to calculate since the disc line becomes nearly uniform and in the first approximation $b - a = \lambda/2$ (for the gravest mode, of course) in the case of the two conducting cylinders and $b - a = \lambda/4$ in the case of the one perforated cylinder. In this case only the small deviations from these values need to be calculated. For this purpose the Bessel functions are replaced by their asymptotic expansions Approximate formulae for the roots of (1) are available in books on Bessel functions.

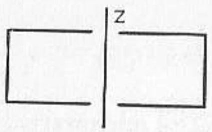


FIG. 8.20. A perforated cylindrical cavity.

When a and b are fairly small which is the case for the principal oscillation mode, it is more expedient to compute the roots graphically. The ratio $J_0(x)/N_0(x)$ is plotted as a function of x. This graph consists of an infinite number of branches and it cuts the x-axis when x is a zero of J_0 and goes off to infinity when x is a zero of N_0. Then we pick pairs (x_1,x_2) corresponding to the same ordinates, and thus obtain pairs of values of βa and βb which satisfy the equation. Starting with $\beta a = 0$ and selecting the smallest corresponding value of βb, we plot the latter against βa. Such a curve makes it possible to compute the dimensions of the resonator or the resonant wavelength, as may be seen from Fig. 8.21. In this figure

$$\bar{a} = \frac{2\pi a}{\lambda}, \quad \bar{b} = \frac{2\pi b}{\lambda}, \quad k = \frac{b}{a},$$

and the curve is the locus of

$$S(\bar{a},\bar{b}) = J_0(\bar{a})N_0(\bar{b}) - N_0(\bar{a})J_0(\bar{b}) = 0.$$

Similarly the curve in Fig. 8.22 is the locus of

$$U(\bar{a},\bar{b}) = N_1(\bar{a})J_0(\bar{b}) - J_1(\bar{a})N_0(\bar{b}) = 0.$$

From this curve we can obtain data on the approximate resonant frequencies of the cylindrical cavity shown in Fig. 8.20. Thus the small holes do not affect appreciably the principal resonant frequency; on the other hand a perfectly conducting cylinder, even if quite thin, changes the resonant frequency by a substantial percentage. However an *infinitely* thin wire does not affect the resonance conditions.

The radial impedance $Z_\rho(a)$ is positive imaginary if b is sufficiently small. Assuming a capacitance sheet over $\rho = a$ (Fig. 8.23), whose radial capacitance is C_ρ, we shall have resonance when the sum of the two impedances vanishes

$$Z_\rho(a) + \frac{1}{i\omega C_\rho} = 0.$$

This condition may also be expressed as follows

$$Y_\rho(a) + i\omega C_\rho = 0, \quad \text{or} \quad iY_\rho(a) = \omega C_\rho.$$

Plotting $iY_\rho(a)$ for different values of $k = b/a$, we obtain the family of

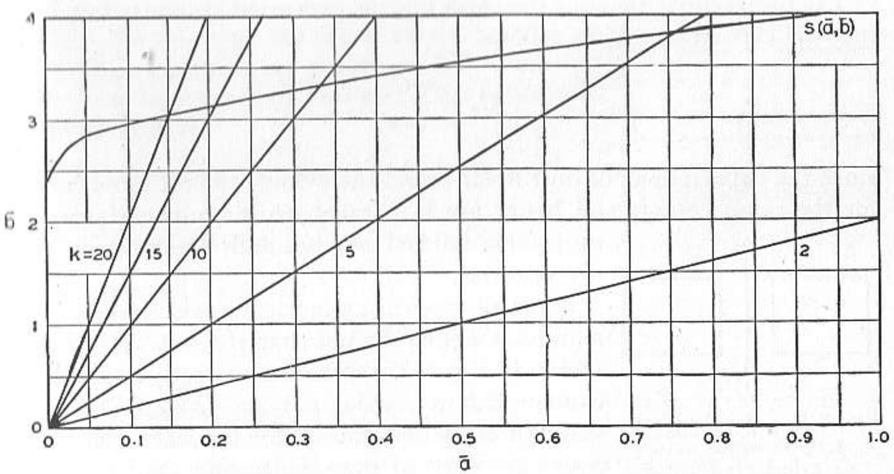


FIG. 8.21. Curves pertaining to resonance in the cavity shown in Fig. 8.19 when the walls of the cavity are of zero impedance.

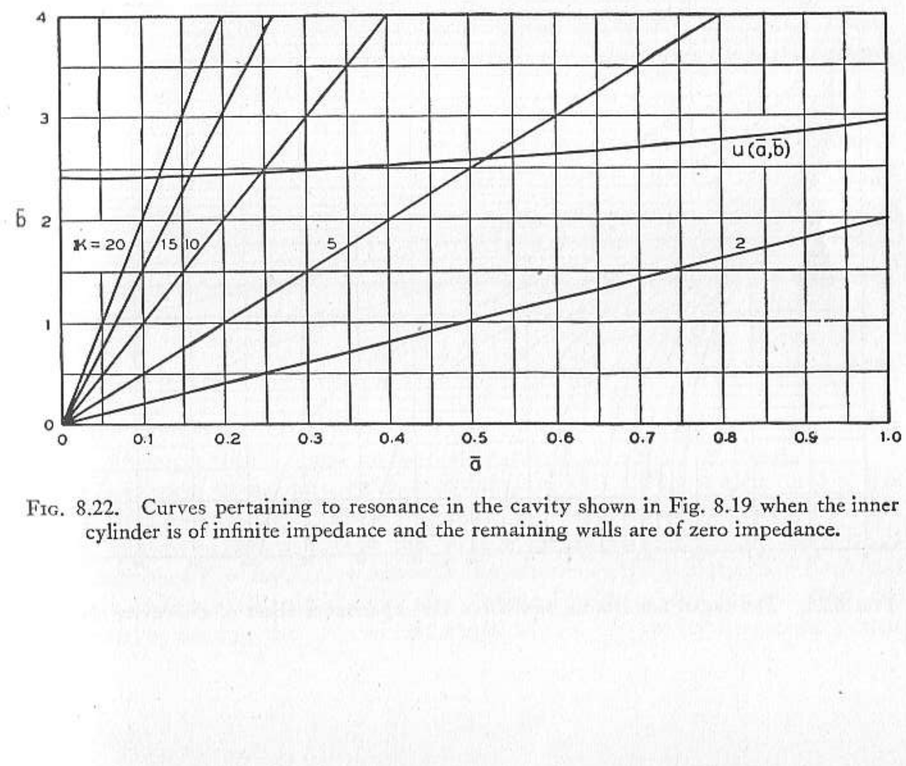


FIG. 8.22. Curves pertaining to resonance in the cavity shown in Fig. 8.19 when the inner cylinder is of infinite impedance and the remaining walls are of zero impedance.

curves shown in Fig. 8.24 and from these the dimensions of the resonant cavity can be determined for various values of C_ρ.

The radial capacitance of the sheet may be expressed in terms of the total internal capacitance C_i as follows

$$C_\rho = \frac{hC_i}{2\pi a},$$

since the capacitances of unit areas round the cylinder admit more current for the same voltage and hence are in parallel while the capacitances of unit areas stacked longitudinally along the cylinder are in series.

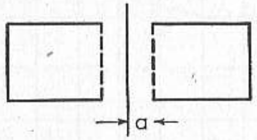


FIG. 8.23. A cylindrical cavity in which the inner cylinder is a capacitance sheet and the remaining walls are of zero impedance.

After all types of cylindrical waves have been examined it will be obvious that, if $h < \lambda/2$, the cavity shown in Fig. 8.23 represents correctly a cylindrical cavity with a coaxial plunger (Fig. 8.25). The value of C_i is determined by the capacitance between the base of the plunger and the base of the cavity, including the "fringing capacitance." The latter may comprise a substantial fraction of the total capacitance C_i.

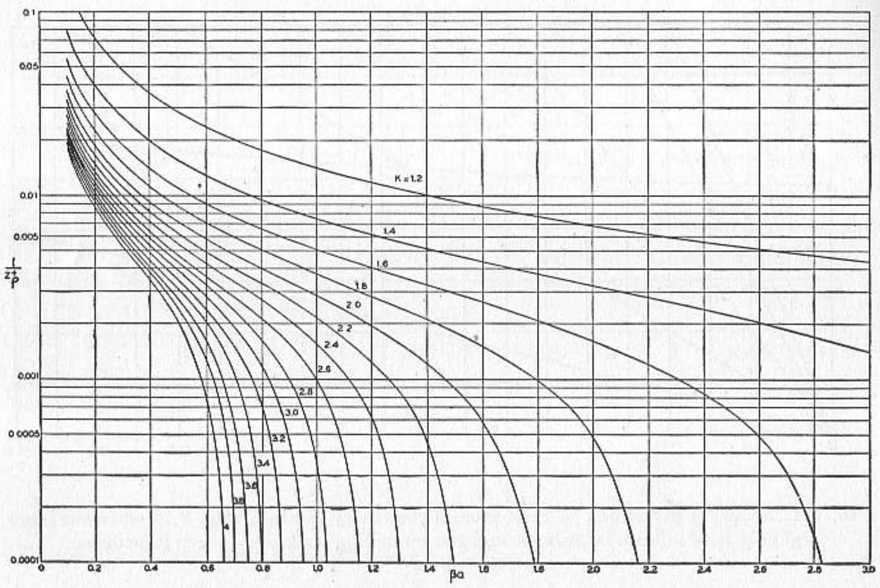


FIG. 8.24. The radial admittance seen from the capacitance sheet of the cavity shown in Fig. 8.23.

8.7. *Solenoids and Wedge Transmission Lines*

We shall now turn our attention to the converse type of uniform cylindrical waves in which the magnetic lines are parallel to a given axis while the electric lines are circular. For this type the transmission equations are (5–2); these equations are similar to (5–1) and their solution can be obtained by analogy. Thus for the field intensities we have

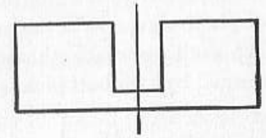


FIG. 8.25. A cross-section of a cylindrical cavity with a coaxial plunger.

$$H_z^+(\rho) = AK_0(\sigma\rho), \quad H_z^-(\rho) = BI_0(\sigma\rho),$$
$$E_\varphi^+(\rho) = A\eta K_1(\sigma\rho), \quad E_\varphi^-(\rho) = -B\eta I_1(\sigma\rho), \tag{7–1}$$

and for the radial impedances

$$Z_\rho^+(\rho) = \frac{\eta K_1(\sigma\rho)}{K_0(\sigma\rho)}, \quad Z_\rho^-(\rho) = \frac{\eta I_1(\sigma\rho)}{I_0(\sigma\rho)}. \tag{7–2}$$

Comparing these expressions with (5–5), we find that the products of the corresponding radial impedances for the two types of waves are equal to the square of the intrinsic impedance; the radial admittances of the present waves are obtained if we divide the impedances of the other type by η^2.

In nondissipative media, for small values of $\rho = a$, we have approximately

$$Z_\rho^-(a) = \tfrac{1}{2}i\omega\mu a, \quad Y_\rho^+(a) = i\omega\epsilon a\left(\log\frac{\lambda}{2\pi a} + 0.116\right) + \frac{\pi^2 a}{\eta\lambda}.$$

For large values of ρ the outward looking impedance approaches η while the inward looking impedance fluctuates between $-\infty$ and $+\infty$ if the medium is nondissipative and approaches η otherwise.

Consider now a circulating current sheet of density $J_\varphi = J$ per unit length on a cylinder of radius a; that is, a "coil" with one turn per unit length. The electric intensity is continuous across the sheet but the magnetic intensity increases by an amount J; thus

$$E_\varphi^+(a) = E_\varphi^-(a) = -E, \quad H_z^-(a) - H_z^+(a) = J,$$

where E is the driving electric intensity. Dividing the second equation by the corresponding terms of the first, we have

$$\frac{J}{E} = Y_\rho^+ + Y_\rho^-.$$

Thus the internal and external media are in parallel. For small values of a, Y^- is very much larger than Y^+ and we have substantially $E = Z_\rho^- J = \tfrac{1}{2}i\omega\mu a J$.

For a solenoid wound on a cylinder of radius a, with closely spaced turns of fine wire, n turns per unit length, we have $J = nI$. If V is the voltage applied to a portion of the solenoid of length l, then the voltage E per unit length of the wire is $V/2\pi anl$; hence $V = i\omega LI$, $L = \mu\pi a^2 n^2 l$. Thus we have the inductance of a solenoid of length l when it is a part of an infinitely long solenoid, or when the end effects are eliminated by bending the solenoid into a toroidal coil. In the latter case there exists some curvature effect, of course; but one may expect it to be small if the radius of the cross-section of the coil is small compared with the mean radius of the torus itself.

The equations of this section may be used in the solution of problems which have no apparent connection with solenoids. One such problem is that of the diffraction of plane waves by a narrow slit in an infinite perfectly conducting plane (Fig. 8.26). We shall approach this problem by considering first a " wedge transmission line " formed by two half planes issuing from the same axis or nearly so (Fig. 8.27). Let V

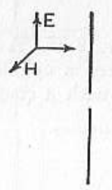


FIG. 8.26. A uniform plane wave incident on a perfectly conducting plane with an infinitely long, narrow slit.

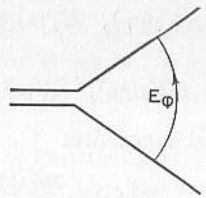


FIG. 8.27. A cross-section of a wedge by a plane normal to the axis of the wedge.

be the voltage impressed on this line and let J be the input radial current in the planes (per unit length along the axis). If the half planes terminate at distance $\rho = a$, then the input voltage is given by $V = -\psi a E_\varphi(a)$, where ψ is the wedge angle. Hence the approximate input admittance per unit length is

$$Y = \frac{1}{\psi a} Y_\rho^+(a) = \frac{i\omega\epsilon}{\psi} K_0(i\beta a). \qquad (7\text{–}3)$$

In Fig. 8.26 the input admittance of each wedge line, one looking to the left and the other to the right is given by the above expression with $\psi = \pi$. The two lines are in parallel and the total input admittance is $Y = (2i\omega\epsilon/\pi)K_0(i\beta a)$. This expression becomes more accurate as a becomes smaller. For a finite slit of width s a better approximation to the input admittance is obtained by assuming $a = |x - \hat{x}|$ and averaging the admittance over the slit just as we have done in the case of a metal strip.* Comparing the above expression with (5–11), and using (5–12), we obtain the following average value

$$Y = \omega\epsilon + \frac{2i\omega\epsilon}{\pi}\left(\log\frac{\lambda}{s} - 0.222\right) = \frac{2\pi}{\eta\lambda} + \frac{4i}{\eta\lambda}\left(\log\frac{\lambda}{s} - 0.222\right).$$

Consider now a uniform plane wave incident on a screen with a narrow slit through it and let E be perpendicular to the slit. If there were no slit, the wave would be completely reflected and H would be doubled; electric current of density $2H$ would flow upward in the plane (Fig. 8.26). With the slit present we should still have complete reflection except in the region surrounding the slit. The electric current between the edges of the slit is reduced to zero and the voltage necessary to reduce the

* See equation (5–10).

current density $2H$ to zero is

$$V = -\frac{2H}{Y} = -\frac{\eta\lambda H}{\pi + 2i\left(\log\frac{\lambda}{s} - 0.222\right)}. \qquad (7\text{–}4)$$

This is the counter-electromotive force produced by charge concentrations on the edges of the slit.

Expressions (1) for E_φ^+, H_z^+ to the right of the screen may be found in terms of the voltage V applied over half the circumference near $\rho = 0$; for in this neighborhood $E_\varphi^+(\rho_1) = -V/\pi\rho_1$. Hence at all distances

$$E_\varphi^+(\rho) = -\frac{V}{\pi\rho_1}\cdot\frac{K_1(i\beta\rho)}{K_1(i\beta\rho_1)} = -\frac{i\beta V}{\pi}K_1(i\beta\rho),$$

$$H_z^+(\rho) = Y_\rho^+(\rho)E_\varphi^+(\rho) = -\frac{i\beta V}{\pi\eta}K_0(i\beta\rho).$$

Substituting from (4), we have

$$H_z^+(\rho) = \frac{2iHK_0(i\beta\rho)}{\pi + 2i\left(\log\frac{\lambda}{s} - 0.222\right)} = \frac{He^{-i\beta\rho + i\pi/4}}{\pi + 2i\left(\log\frac{\lambda}{s} - 0.222\right)}\sqrt{\frac{\lambda}{\rho}}, \text{ as } \rho \to \infty.$$

8.8. *Wave Propagation along Coaxial Cylinders*

Consider a pair of perfectly conducting coaxial cylinders. If we apply a transverse voltage between these conductors we expect that longitudinal currents will be generated. If the voltage is so applied that circular symmetry is preserved, we expect that the resultant field will be independent of the φ-coordinate. One such field is described by equations (4.12–8) and the other by (4.12–9). The first of these has no radial electric intensity; hence we need consider only equations (4.12–9). In section 6.11 we have found that if E_z vanishes on the surfaces of the cylinders but not between them, waves will travel along the coaxial pair only if the wavelength is comparable to the transverse dimensions. In Chapter 10 we shall deal with such waves in detail; but at present we are concerned with a type of wave propagation which is possible at low frequencies as well as at high frequencies. This leads us to assume that the wave is transverse electromagnetic ($E_z = 0$). Our equations now become

$$\frac{\partial H_\varphi}{\partial z} = -(g + i\omega\epsilon)E_\rho, \quad \frac{\partial E_\rho}{\partial z} = -i\omega\mu H_\varphi, \quad \frac{\partial}{\partial\rho}(\rho H_\varphi) = 0. \qquad (8\text{–}1)$$

The last equation implies that the magnetomotive force round any circle coaxial with the cylinders is independent of the distance from the axis. This is natural since in the absence of longitudinal displacement currents this magnetomotive force must equal the conduction current $I(z)$ in the inner cylinder

$$2\pi\rho H_\varphi = I(z), \quad H_\varphi = \frac{I(z)}{2\pi\rho}. \tag{8–2}$$

Substituting from (2) in (1), integrating from $\rho = a$ to $\rho = b$, and introducing the transverse voltage $V(z)$ along a radius from the inner cylinder to the outer, we have

$$\frac{dI}{dz} = -YV, \quad \frac{dV}{dz} = -ZI,$$
$$Y = \frac{2\pi(g + i\omega\epsilon)}{\log \frac{b}{a}}, \quad Z = \frac{i\omega\mu}{2\pi}\log\frac{b}{a}. \tag{8–3}$$

Except at very high frequencies it is easier to measure voltages and currents than field intensities, and equations (3) are preferable to the original equations (1).

The propagation constant Γ and the characteristic impedance K are

$$\Gamma = \sigma, \quad K = \frac{\eta}{2\pi}\log\frac{b}{a}.$$

Thus the propagation constant is the same as for uniform plane waves and the characteristic impedance is modified only by a factor depending on the ratio of the radii. If the dielectric between the cylinders is air, then

$$K = 60 \log\frac{b}{a} = 138 \log_{10}\frac{b}{a} \text{ ohms.}$$

The following table gives an idea of the order of magnitude of this impedance:

$\frac{b}{a} =$	1.5,	2,	2.5,	3,	3.5,	4,	5;	5.29.
$K =$	24.3,	41.6,	55.0,	65.9,	75.2,	83.2,	96.6;	100.

K increases very slowly with the ratio of the radii and it is impossible to obtain really high values in coaxial pairs of practical dimensions; nevertheless if b is infinitely large, K is also infinitely large and it is impossible to set up a wave of the present type on a single wire except with an infinite power properly supplied over an *entire* plane perpendicular to the wire. By "properly supplied" we mean that "most" of this power should be supplied at large distances from the wire. The electric intensity E_ρ is inversely proportional to ρ and, while the power carried by the wave per unit area diminishes as ρ increases, an infinite amount is carried outside any cylinder of finite radius. If, for example, we break an infinitely long wire and connect its free ends to an electric generator, we should not expect to obtain a plane wave of the type here considered.

It is evident from (1) that for two progressive waves $E_\rho^+ = \eta H_\varphi^+$, $E_\rho^- = -\eta H_\varphi^-$, and that consequently $E_\rho = \eta I(z)/2\pi\rho$. In an air-filled coaxial pair we have $E_\rho = (60/\rho)\, I(z)$.

If the conductivity of the coaxial conductors is not infinite E_z does not vanish on the boundaries but it is so small that its effect on the magnetic field is entirely negligible. Thus, starting with the magnetic field as given in the above equations, we have

$$E_z(a) = Z_\rho^-(a)\frac{I}{2\pi a}, \quad E_z(b) = -Z_\rho^+(b)\frac{I}{2\pi b},$$

where the radial impedances are determined by the properties of the conductors. The general picture of wave propagation along imperfectly conducting cylinders is that most of the energy is carried longitudinally by the wave between the cylinders and a very small fraction is diverted into the cylinders where it is traveling radially.

The nonvanishing surface voltages $E_z(a)$ and $E_z(b)$ modify the second equation of the set (3) which now becomes

$$\frac{dV}{dz} + E_z(a) - E_z(b) = -ZI. \tag{8–4}$$

The ratio of the voltage on the surface of the wire to the current in it has been called the internal impedance of the wire or the surface impedance. The latter name, particularly, conveys the idea that imperfectly conducting wires provide a surface drag on the wave between them. The surface impedances of the conductors per unit length are then

$$Z_{aa} = \frac{Z_\rho^-(a)}{2\pi a}, \quad Z_{bb} = \frac{Z_\rho^+(b)}{2\pi b},$$

and (4) becomes

$$\frac{dV}{dz} = -(Z_{aa} + Z_{bb} + Z)I.$$

The propagation constant and the impedance now become

$$\Gamma = \sqrt{(Z_{aa} + Z_{bb} + Z)Y}, \quad K = \sqrt{\frac{Z + Z_{aa} + Z_{bb}}{Y}}.$$

In practice the thickness of the outer cylindrical shell is usually small compared with its radius and the radial impedance may be computed from the plane wave formulae; thus $Z_p^+(b) = \hat{\eta} \coth \hat{\sigma}t$, where t is the thickness of the shell and $\hat{\eta}$ and $\hat{\sigma}$ refer to the metal of the shell. The surface impedance of the outer shell is therefore

$$Z_{bb} = \frac{\hat{\eta}}{2\pi b} \coth \hat{\sigma}t.$$

If the frequency is so high that $\hat{\sigma}t$ is large compared with unity, $\coth \hat{\sigma}t$ is nearly equal to unity and

$$Z_{bb} = \frac{\hat{\eta}}{2\pi b} = \frac{\mathcal{R}}{2\pi b}(1 + i), \tag{8–5}$$

where $\mathcal{R} = \sqrt{\pi\hat{\mu}f/\hat{g}}$ is the intrinsic resistance of the outer cylinder. If on the other hand the frequency is very low and $\hat{\sigma}t$ is small compared with unity, then the surface impedance $Z_{bb} = \hat{\eta}/2\pi bt\hat{\sigma} = 1/2\pi btg$ is the d–c resistance of the shell. For intermediate frequencies we separate the real and imaginary parts of $\coth \hat{\sigma}t$ and obtain

$$R_{bb} = \frac{\mathcal{R}}{2\pi b}\frac{\sinh 2\hat{\alpha}t + \sin 2\hat{\alpha}t}{\cosh 2\hat{\alpha}t - \cos 2\hat{\alpha}t}, \quad \omega L_{bb} = \frac{\mathcal{R}}{2\pi b}\frac{\sinh 2\hat{\alpha}t - \sin 2\hat{\alpha}t}{\cosh 2\hat{\alpha}t - \cos 2\hat{\alpha}t},$$

where $\hat{\alpha} = \sqrt{\pi\hat{\mu}f\hat{g}}$ is the intrinsic attenuation constant of the conductor.

It is evident that the surface resistance of the shell is a fluctuating function of the thickness. Thus, using $\coth \hat{\sigma}t = \coth(\hat{\alpha}t + i\hat{\alpha}t)$, it can be shown that R will have a minimum value when $\hat{\alpha}t = \pi/2$, $t = \pi/2\hat{\alpha}$. With this value of t we have $\coth \hat{\sigma}t = \tanh \hat{\alpha}t$, and consequently

$$Z_{bb} = \frac{\mathcal{R}(1 + i)}{2\pi b} \tanh \frac{\pi}{2} = \frac{0.92\mathcal{R}(1 + i)}{2\pi b}.$$

The resistance of this shell is actually lower than that of a thicker shell. At one megacycle the optimum thickness for copper is about 0.104 mm. The optimum thickness is about 57 per cent greater than the skin depth, defined by equation (4.10–13). When the thickness of the shell is twice the optimum thickness, then $\coth \hat{\sigma}t = \coth \hat{\alpha}t = \coth \pi = 1.004$ and further increase in the thickness has a negligible effect on the surface impedance.

The *surface transfer impedance* Z_{ab} is defined as the ratio of the intensity at the outer surface of the shell to the electric current when the current returns internally, or the corresponding ratio for the intensity at the inner surface if the current returns externally. Thus if a part I_a of the total current I returns externally and the other part I_b internally, then, in accordance with our definitions, we have

$$E_z(a) = Z_{aa}I_a + Z_{ab}I_b, \quad E_z(b) = Z_{ba}I_a + Z_{bb}I_b.$$

The transfer impedance is approximately

$$Z_{ab} = Z_{ba} = \frac{\hat{\eta}}{2\pi\sqrt{ab}\sinh \hat{\sigma}t}.$$

The transfer impedance is important in computing fields external to the coaxial pair. Of course, such fields are very feeble; they are nevertheless significant in problems of interference or "crosstalk" between telephone transmission lines. The transfer impedance determines the series electromotive forces induced in external transmission lines of which the outer conductor of the given coaxial pair may form a part.

If a cylindrical conductor consists of coaxial homogeneous conducting layers in electrical contact, the surface self-impedances and the transfer impedance may be calculated step-by-step from the surface impedances of the separate layers by regarding these layers as transducers in series and using equations* (7.26–1). The rule for calculation may be summarized as follows: Let two conductors, each of which may be composed of coaxial layers, fit tightly one inside the other. Any surface self-impedance of the compound conductor equals the corresponding self-impedance of the conductor nearest to the return path diminished by a fraction whose numerator is the square of the transfer impedance across this conductor and whose denominator is the sum of the surface impedances of the two component conductors if each is regarded as the return path for the other. The transfer impedance of the compound conductor is a fraction whose numerator is the product of the transfer impedances of the component conductors and whose denominator is the same as that for each self-impedance. If two coaxial conductors are short circuited at frequent intervals the above rule holds even if the conductors do not fit tightly one over the other, provided we add to the denominators a third term representing the inductive reactance of the space between the conductors.

At sufficiently high frequencies the surface self-impedances are given by the simple expression (5) even for compound conductors; only the intrinsic resistances of the layers nearest to the dielectric between the conductors need be considered. For most purposes we may ignore the effect of dissipation on the characteristic impedance. Likewise we may ignore the effect of dissipation on the phase constant and the velocity; thus the

* The difference between the algebraic signs of Z_{12} and Z_{ab} is the result of convention with respect to the positive directions of I_1 and I_2 in one case and of I_a and I_b in the other; I_1 and I_2 flow in opposite directions through the mutual impedance while I_a and I_b flow in the same direction.

propagation constant becomes

$$\Gamma = \alpha + i\beta, \quad \beta = \omega\sqrt{\mu\epsilon}, \quad \alpha = \frac{\mathcal{R}}{4\pi K}\left(\frac{1}{a} + \frac{1}{b}\right) + \tfrac{1}{2}g\sqrt{\frac{\mu}{\epsilon}}.$$

The last term in α is due to the conductivity of the dielectric.

A half wavelength section of a coaxial pair may be used as a resonant circuit. The best way to short circuit the ends is with plane metal caps which completely separate the dielectric between the conductors from external space. Power losses in the caps can be calculated as follows. The cap is a coaxial transmission line in which metal takes the place of dielectric; the characteristic impedance of this line is

$$\frac{\hat{\eta}}{2\pi}\log\frac{b}{a} = \frac{\mathcal{R}(1+i)}{2\pi}\log\frac{b}{a},$$

and its real part is the resistance introduced at each end by the "short-circuiting" caps. The effective resistance of the line itself is $(\lambda/4)(\mathcal{R}/2\pi a + \mathcal{R}/2\pi b)$, since the sinusoidal distribution of current cuts the power loss in half as compared with uniform distribution. The effective total inductance is also cut in half

$$\frac{\lambda}{4}\frac{\mu}{2\pi}\log\frac{b}{a} = \frac{1}{4\omega}\eta\log\frac{b}{a} = \frac{\pi K}{2\omega}.$$

Thus the Q of the resonant section is

$$Q = \frac{4\pi^2 K}{\mathcal{R}\left(\frac{\lambda}{a} + \frac{\lambda}{b} + 8\log\frac{b}{a}\right)}.$$

The assumption that the field is transverse and circularly symmetric has led to the determination of a specific field distribution in a typical transverse plane; thus the electric intensity is radial and it varies inversely as the distance from the axis; the magnetic intensity is circular and it also varies inversely as the distance from the axis. In order to generate this wave in pure form it is necessary, therefore, to apply a transverse voltage with its radial gradient varying inversely as the distance from the axis. No assumptions have been made with regard to the frequency, at least when the cylindrical conductors are perfect; hence such waves can presumably be set up at all frequencies. Imperfect conductivity restricts the frequency to some extent but not until we reach the "optical" frequency range. The terminal conditions are much more important. Usually it is not practicable to apply the transverse voltage in the manner specified above; we can control the total voltage much more readily than its distribution. It is in the terminal conditions that the real frequency limitation is introduced. Unless the applied voltage varies inversely as the distance, other types of waves besides the one we have been considering will be set up. These waves are rapidly attenuated at frequencies below a certain critical frequency* and then they represent merely an end effect. For this reason the present wave may be called the *principal* or the *dominant* wave. But above the critical frequency these other waves will make themselves felt just as far from the end of the line as the dominant wave and the character of transmission changes.

8.9. *Transverse Electromagnetic Plane Waves* (TEM-waves)

For transverse electromagnetic waves $E_z = H_z = 0$ and the electromagnetic equations are considerably simplified; thus we have

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} = 0, \quad \frac{\partial H_x}{\partial x} + \frac{\partial H_y}{\partial y} = 0; \quad \frac{\partial E_x}{\partial y} = \frac{\partial E_y}{\partial x}, \quad \frac{\partial H_x}{\partial y} = \frac{\partial H_y}{\partial x}; \qquad (9\text{–}1)$$

so that either E or H or both may be derived from stream functions or from scalar potentials. The remaining equations are

$$H_x = \frac{1}{i\omega\mu}\frac{\partial E_y}{\partial z}, \quad H_y = -\frac{1}{i\omega\mu}\frac{\partial E_x}{\partial z},$$
$$E_x = -\frac{1}{g + i\omega\epsilon}\frac{\partial H_y}{\partial z}, \quad E_y = \frac{1}{g + i\omega\epsilon}\frac{\partial H_x}{\partial z}. \qquad (9\text{–}2)$$

While equations (1) impose restrictions on the field distribution in transverse planes, equations (2) govern the propagation of these transverse field patterns in the z-direction. Eliminating H_y from (2) we obtain

$$\frac{\partial^2 E_x}{\partial z^2} = \sigma^2 E_x. \qquad (9\text{–}3)$$

The other field intensities satisfy the same equation. Thus the *propagation constant of all transverse electromagnetic waves is equal to the intrinsic propagation constant of the medium.*

Consider now a progressive wave moving in the positive z-direction. The constants introduced in integrating equation (3) will not in general be independent of x and y; thus we may write

$$E_x = E_0(x,y)e^{-\sigma z}, \quad H_y = \frac{1}{\eta}E_0(x,y)e^{-\sigma z}, \quad E_x = \eta H_y.$$

* The critical wavelength is equal approximately to the mean circumference of the coaxial pair.

Hence the wave impedance of all transverse electromagnetic waves is equal to the intrinsic impedance of the medium.

Since any cartesian component of either E or H satisfies the wave equation, we have in view of (3)

$$\frac{\partial^2 E_x}{\partial x^2} + \frac{\partial^2 E_x}{\partial y^2} = 0.$$

Hence the distribution of electric and magnetic fields in a typical transverse plane is governed by static field equations. If we imagine a current distribution conforming to a given electrostatic field pattern in two dimensions and if we assume that the current density is proportional to E and is a harmonic function of time, then we shall obtain two waves traveling in opposite directions normal to the plane and the relative field distribution in all equiphase planes will conform to the assumed pattern.

The wave equation (4.10–1) imposes a restriction on all exponential plane waves. Thus if the propagation constant in the z-direction is Γ_z, then the field distribution in a transverse plane must satisfy the following equation

$$\frac{\partial^2 E_x}{\partial x^2} + \frac{\partial^2 E_x}{\partial y^2} = (\sigma^2 - \Gamma_z^2)E_x.$$

Conversely if the field distribution satisfies this equation, the field pattern represented by it will be propagated in the z-direction and the propagation constant will be Γ_z. If we were to impress over a given plane transverse electric forces not satisfying the above partial differential equation for a constant value of $\sigma^2 - \Gamma_z^2$, then the field distribution over planes parallel to the plane of the impressed forces would not conform to the impressed field.

As we have already stated, the transverse electric field may be derived either from a potential or from a stream function. Choosing the former method, we have

$$E_t = -\text{grad}_t\, V, \quad E_x = -\frac{\partial V}{\partial x}, \quad E_y = -\frac{\partial V}{\partial y}. \tag{9–4}$$

Substituting in (2), we obtain

$$H_x = \frac{\partial \Psi}{\partial y}, \quad H_y = -\frac{\partial \Psi}{\partial x}, \tag{9–5}$$

where the stream function Ψ is

$$\Psi = -\frac{1}{i\omega\mu}\frac{\partial V}{\partial z}. \tag{9–6}$$

Except for a coefficient depending on z the stream function Ψ has the same form as the potential V. Substituting from (5) in the last two equations of the set (2) and comparing with (4), we obtain

$$V = -\frac{1}{g + i\omega\epsilon}\frac{\partial \Psi}{\partial z}. \tag{9–7}$$

This and the preceding equation form the familiar set of transmission equations.

Potential and stream functions are also solutions of the two dimensional Laplace's equation and to any solution of this equation there corresponds a transverse electromagnetic wave.

8.10. *Transverse Electromagnetic Waves on Parallel Wires*

Equation (6.23–14) represents the stream function of two parallel current filaments carrying steady currents. In accordance with the preceding section it is also the stream function associated with transverse electromagnetic waves along the wires. Thus substituting from (6.23–14) in (9–6) and (9–7) we have

$$\frac{\partial \hat{V}}{\partial z} = \frac{i\omega\mu}{2\pi}\log\frac{\rho_1}{\rho_2}I, \quad \hat{V} = \frac{\log\dfrac{\rho_1}{\rho_2}}{2\pi(g + i\omega\epsilon)}\frac{\partial I}{\partial z}, \tag{10–1}$$

where ρ_1 and ρ_2 are the distances from the filaments carrying currents I and $-I$ respectively. The cylinders $u = \log \rho_2/\rho_1 =$ constant are equipotential surfaces. In cartesian coordinates their equation is

$$(x - c \coth u)^2 + y^2 = c^2 \operatorname{csch}^2 u. \tag{10–2}$$

Let two perfectly conducting cylinders be introduced along two of these surfaces, $u = u_1$ and $u = u_2$. We can remove the original filaments without disturbing the wave between the cylinders and will be left with a transverse electromagnetic wave along a pair of cylinders, either external to each other or one inside the other. Equations (1) apply to any line parallel to the cylinders. Let $\hat{V}_1$ and $\hat{V}_2$ be the potentials on the cylinders; then the difference $V = \hat{V}_1 - \hat{V}_2$ will satisfy the transmission equations in which

$$L = \frac{\mu}{2\pi}(u_1 - u_2), \quad C = \frac{2\pi\epsilon}{u_1 - u_2}, \quad G = \frac{2\pi g}{u_1 - u_2}. \tag{10–3}$$

Let the distance between the axes of the cylindrical conductors be l and let a and b be their radii. Then from (2) we have

$$a^2 = c^2 \operatorname{csch}^2 u_1, \quad b^2 = c^2 \operatorname{csch}^2 u_2,$$
$$l^2 = c^2 (\coth u_1 - \coth u_2)^2. \tag{10–4}$$

Expanding the last equation and substituting $\coth^2 u = \operatorname{csch}^2 u + 1$, we obtain

$$l^2 = c^2 \operatorname{csch}^2 u_1 + c^2 \operatorname{csch}^2 u_2 + 2c^2 (1 - \coth u_1 \coth u_2)$$
$$= c^2 \operatorname{csch}^2 u_1 + c^2 \operatorname{csch}^2 u_2 - 2c^2 \operatorname{csch} u_1 \operatorname{csch} u_2 \cosh (u_1 - u_2).$$

In substituting the radii a and b from (4) into this equation we should bear in mind that u_1 and u_2 may be either positive or negative while the radii are essentially positive; thus we have

$$l^2 = a^2 + b^2 \pm 2ab \cosh (u_1 - u_2).$$

The upper sign corresponds to the case in which u_1 and u_2 have opposite signs and the cylinders are external to each other (see Fig. 1.6); the lower sign corresponds to the case in which u_1 and u_2 have like signs and one cylinder is inside the other.

Thus, depending upon whether the cylinders are external to each other or one inside the other, we have

$$u_1 - u_2 = \cosh^{-1} \frac{l^2 - a^2 - b^2}{2ab} \quad \text{or} \quad \cosh^{-1} \frac{a^2 + b^2 - l^2}{2ab}.$$

Substituting in (3) we obtain L, C, G.

The foregoing formulae have been obtained on the assumption of perfectly conducting cylinders and if the conductivity is finite the above results have to be modified. At very low frequencies, for example, for the case of two solid cylindrical wires, the magnetic flux penetrates the wires and the inductance is

$$L = \frac{\mu}{\pi} \log \frac{l}{\sqrt{ab}} + \frac{\mu}{4\pi}.$$

This inductance is larger than that given by (3); the capacity on the other hand remains the same. Hence the wave velocity on the wires is smaller at low frequencies than at high frequencies. Furthermore we should include the resistance of the wires in series with the inductance.

At very high frequencies the flux is largely forced out of the conductors and the resistance per unit length is given by $R = \mathcal{R} \int H_s H_s^* \, ds$, where H_s is the component of H tangential to the wires and the integration is taken round both wires. The field H is that produced by a unit current in each wire. If the radii of the conductors are changed by an infinitesimal amount δn in such a way as to *increase* the inductance, then the increment in the inductance is $\delta L = \mu \delta n \int H_s H_s^* \, ds$. Hence we obtain the following

simple principle*

$$R = \frac{\mathcal{R}}{\mu} \frac{\delta L}{\delta n},$$

where $\delta L/\delta n$ is the variational derivative of the inductance.

For two wires external to each other we have therefore

$$R = \frac{\mathcal{R}}{\mu} \left(- \frac{\partial L}{\partial a} - \frac{\partial L}{\partial b} \right).$$

Calculating the derivatives, we obtain

$$R = \frac{\mathcal{R}}{2\pi} \frac{2(a+b) + \left(\frac{1}{a} + \frac{1}{b}\right)(l^2 - a^2 - b^2)}{\sqrt{[l^2 - (a+b)^2][l^2 - (a-b)^2]}},$$
$$= \frac{\mathcal{R}}{\pi a \sqrt{1 - \frac{4a^2}{l^2}}}, \quad \text{if } b = a.$$

Similarly when one wire is inside a cylindrical shell, then (for $b > a$)

$$R = \frac{\mathcal{R}}{\mu} \left(- \frac{\partial L}{\partial a} + \frac{\partial L}{\partial b} \right);$$

hence

$$R = \frac{\mathcal{R}}{2\pi} \frac{2(b-a) + \left(\frac{1}{a} - \frac{1}{b}\right)(a^2 + b^2 - l^2)}{\sqrt{[(a+b)^2 - l^2][(b-a)^2 - l^2]}}.$$

The high frequency resistance of parallel wires increases as the wires approach each other. This phenomenon is called the "proximity effect." If a wire is inside a cylindrical shell, the resistance is minimum when they are coaxial; in this case the proximity effect is sometimes called the "eccentricity effect."

8.11. *Transverse Electromagnetic Spherical Waves* (TEM-waves)

For transverse electromagnetic spherical waves we have $E_r = H_r = 0$. The theory of these waves is similar to the theory of transverse electromagnetic plane waves. We have two divergence equations

$$\frac{\partial}{\partial v} (\sin \theta \, E_\theta) + \frac{\partial}{\partial \varphi} E_\varphi = 0, \quad \frac{\partial}{\partial \theta} (\sin \theta \, H_\theta) + \frac{\partial}{\partial \varphi} H_\varphi = 0, \qquad (11\text{–}1)$$

* Harold A. Wheeler, "Formulas for the Skin Effect," I.R.E. Proc., 30, pp. 412–424, Sept. 1942.

signifying that either E or H or both can be expressed in terms of stream functions. Then we have another pair of equations

$$\frac{\partial}{\partial\theta}(\sin\theta\, E_\varphi) = \frac{\partial}{\partial\varphi}E_\theta, \quad \frac{\partial}{\partial\theta}(\sin\theta\, H_\varphi) = \frac{\partial}{\partial\varphi}H_\theta, \tag{11–2}$$

signifying that the field intensities may be expressed in terms of potential functions. All four equations impose restrictions on the form of the field distribution in equiphase surfaces. Besides these equations we have four equations which describe wave propagation in the radial direction

$$i\omega\mu r H_\theta = \frac{\partial}{\partial r}(rE_\varphi), \quad i\omega\mu r H_\varphi = -\frac{\partial}{\partial r}(rE_\theta),$$
$$(g + i\omega\epsilon)rE_\theta = -\frac{\partial}{\partial r}(rH_\varphi), \quad (g + i\omega\epsilon)rE_\varphi = \frac{\partial}{\partial r}(rH_\theta). \tag{11–3}$$

We choose to represent the electric intensity in terms of a potential function V and the magnetic intensity in terms of a stream function A

$$E_t = -\text{grad } V, \quad rE_\theta = -\frac{\partial V}{\partial\theta}, \quad r\sin\theta\, E_\varphi = -\frac{\partial V}{\partial\varphi},$$
$$r\sin\theta\, H_\theta = \frac{\partial A}{\partial\varphi}, \quad rH_\varphi = -\frac{\partial A}{\partial\theta}. \tag{11–4}$$

If A is regarded as a radial vector, then $H = \text{curl } A$. Our choice of auxiliary functions satisfies the second equation of the set (1) and the first equation of the set (2). Substituting from (4) in the remaining equations of these two sets, we find that V and A satisfy equation (3.6–14).

Equations (3) show that rE_θ and rH_φ vary as E_x and H_y in a uniform plane wave; rE_φ and rH_θ behave as E_y and H_x. Thus the propagation constant of all transverse electromagnetic spherical waves is equal to the intrinsic propagation constant of the medium and the wave impedance is equal to the intrinsic impedance. Substituting from (4) in (3) and integrating with respect to θ and φ, we have

$$\frac{\partial V}{\partial r} = -i\omega\mu A, \quad \frac{\partial A}{\partial r} = -(g + i\omega\epsilon)V.$$

Consider, for example, a progressive wave traveling outward and let

$$A = T(\theta,\varphi)e^{-\sigma r},$$

where $T(\theta,\varphi)$ is a solution of (3.6–14); then we have

$$rH_\varphi = -\frac{\partial T}{\partial\theta}e^{-\sigma r}, \quad r\sin\theta\, H_\theta = \frac{\partial T}{\partial\varphi}e^{-\sigma r}, \quad E_\theta = \eta H_\varphi, \quad E_\varphi = -\eta H_\theta. \tag{11–5}$$

8.12. *Transverse Electromagnetic Waves on Coaxial Cones*

If two infinite cones (Fig. 8.28) are insulated at their common apex and if an a-c voltage V is maintained between their apices, waves will be generated with electric lines coinciding with meridians and magnetic lines along circles coaxial with the axis. This field is independent of φ and the T-function* is given by (3.6–17); thus

$$T = P\log\cot\frac{\theta}{2}. \tag{12–1}$$

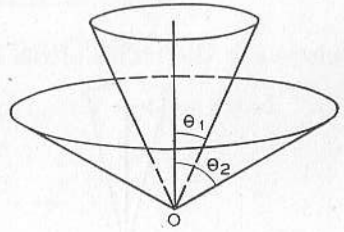


FIG. 8.28. Two coaxial cones.

Hence for outward bound progressive waves in nondissipative media, we have

$$H_\varphi = \frac{Pe^{-i\beta r}}{r\sin\theta}, \quad E_\theta = \frac{\eta Pe^{-i\beta r}}{r\sin\theta}. \tag{12–2}$$

Let $I(r)$ be the total current in the inner cone and $V(r)$ be the voltage from the inner cone to the outer along a meridian (or along any path which is contained in a typical equiphase sphere); then

$$I(r) = 2\pi r\sin\theta\, H_\varphi = 2\pi Pe^{-i\beta r}; \quad I(0) = 2\pi P.$$

The voltage is the integral of rE_θ along a meridian and from (11–5) we find

$$V(r) = \eta e^{-i\beta r}[T(\theta_1) - T(\theta_2)]. \tag{12–3}$$

Therefore,

$$I(r) = I(0)e^{-i\beta r}, \quad V(r) = KI(r),$$
$$K = \frac{\eta}{2\pi}\log\left(\cot\frac{\theta_1}{2}\tan\frac{\theta_2}{2}\right). \tag{12–4}$$

That a "cone transmission line" should turn out to be a uniform line might have been expected. As the distance from the origin increases, the length of the electric lines of force increases in proportion to the distance but the circumference of the conductors also increases in the same proportion; hence the capacity remains constant.

If $\theta_2 = 90°$, one cone becomes a plane and

$$K = \frac{\eta}{2\pi}\log\cot\frac{\theta_1}{2} = 60\log\cot\frac{\theta_1}{2},$$

the numerical coefficient being for free space. When θ_1 is small, then $K = 60\log(2/\theta_1)$. If $\theta_2 = \pi - \theta_1$, the cones become equal and oppositely directed (Fig. 8.29); then

$$K = \frac{\eta}{\pi}\log\cot\frac{\psi}{2} = 120\log\cot\frac{\psi}{2}, \tag{12–5}$$

* The present case can be treated either directly or as a special case of (11–5).

where $\theta_1 = \psi$ is the cone angle, defined as the angle between the axis and the generators of the cone. For small cone angles, we have

$$K = 120 \log \frac{2}{\psi} = 120 \log \frac{2l}{a}, \tag{12-6}$$

where a is the radius of the cone at some fixed distance l from the apex.

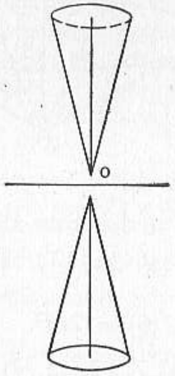


FIG. 8.29. Coaxial cones of equal angles.

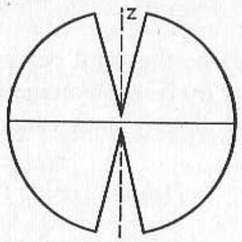


FIG. 8.30. A cavity resonator bounded by two cones and a sphere.

A resonator is formed by two cones inside a perfectly conducting sphere or hemisphere (Fig. 8.30). If the apices of the cones are insulated so that the cone line is open at the origin, the first resonance occurs when the radius l of the sphere equals a quarter wavelength. Then we have

$$V(r) = V_0 \cos \beta r, \quad I(r) = I_0 \sin \beta r, \quad V_0 = iKI_0,$$

$$H_\varphi = \frac{I_0 \sin \beta r}{2\pi r \sin \theta}, \quad E_\theta = i\eta H_\varphi,$$

where I_0 is the maximum current occurring at $r = l$ and V_0 is the maximum voltage at $r = 0$.

The power absorbed by the conductors when they are not perfect may be obtained (at high frequencies) by integrating $\frac{1}{2}\mathcal{R} H_\varphi H_\varphi^*$ where $\mathcal{R}$ is the intrinsic resistance. Thus the power loss $\mathcal{W}_1$ in the spherical portion of the resonator and the loss $\mathcal{W}_2$ in the two cones are

$$\mathcal{W}_1 = \frac{\mathcal{R} I_0^2}{4\pi} \log \left(\cot \frac{\theta_1}{2} \tan \frac{\theta_2}{2}\right), \quad \mathcal{W}_2 = \frac{p\mathcal{R} I_0^2}{4\pi} (\csc \theta_1 + \csc \theta_2),$$

$$p = \tfrac{1}{2}(C + \log \pi - \mathrm{Ci}\, \pi) = 0.824.$$

Since the energy stored in the resonator is

$$W = \frac{1}{16} L\lambda I_0^2 = \frac{\pi K I_0^2}{8\omega},$$

we have the following expression for Q

$$Q = \frac{30\pi^2}{\mathcal{R}} \frac{\log \left(\cot \frac{\theta_1}{2} \tan \frac{\theta_2}{2}\right)}{\log \left(\cot \frac{\theta_1}{2} \tan \frac{\theta_2}{2}\right) + p(\csc \theta_1 + \csc \theta_2)}.$$

For two equal and oppositely directed cones $\theta_1 = \pi - \theta_2 = \psi$ and

$$Q = \frac{30\pi^2}{\mathcal{R}} \frac{\log \cot \frac{\psi}{2}}{\log \cot \frac{\psi}{2} + p \csc \psi}.$$

This Q is maximum when $\psi = 33°.5$ and then

$$Q = \frac{132}{\mathcal{R}}.$$

For a cone of angle $\theta_1 = \psi$ inside a hemisphere, $\theta_2 = 90°$ and

$$Q = \frac{30\pi^2}{\mathcal{R}} \frac{\log \cot \frac{\psi}{2}}{\log \cot \frac{\psi}{2} + p(1 + \csc \psi)}.$$

The Q is maximum when $\psi = 24°.1$ and then

$$Q = \frac{104}{\mathcal{R}}.$$

The input impedance at resonance is

$$Z_i = \frac{VV^*}{2(\mathcal{W}_1 + \mathcal{W}_2)} = \frac{K^2 I_0^2}{2(\mathcal{W}_1 + \mathcal{W}_2)}.$$

Hence we obtain

$$Z_i = \frac{7200\pi}{\mathcal{R}} \frac{\left[\log \left(\cot \frac{\theta_1}{2} \tan \frac{\theta_2}{2}\right)\right]^2}{\log \left(\cot \frac{\theta_1}{2} \tan \frac{\theta_2}{2}\right) + p(\csc \theta_1 + \csc \theta_2)}.$$

For two equal and oppositely directed cones we have

$$Z_i = \frac{14400\pi}{\mathcal{R}} \frac{\left(\log \cot \frac{\psi}{2}\right)^2}{\log \cot \frac{\psi}{2} + p \csc \psi}.$$

This impedance is maximum when $\psi = 9°.2$ and then

$$Z_i = \frac{3.74 \times 10^4}{\mathcal{R}} \text{ ohms.}$$

For a cone in a hemispherical cavity we have

$$Z_i = \frac{7200\pi}{\mathcal{R}} \frac{\left(\log \cot \frac{\psi}{2}\right)^2}{\log \cot \frac{\psi}{2} + p(1 + \csc \psi)}.$$

This is maximum when $\psi = 7°.5$ and then

$$Z_i = \frac{1.70 \times 10^4}{\mathcal{R}} \text{ ohms.}$$

When conical conductors such as those shown in Fig. 8.29 are terminated at some distance l from the apex, the transmission problem is complicated by a sudden change in the physical character of the transmitting medium. Transverse electromagnetic waves require longitudinal conductors and when these are absent such waves are no longer possible. Thus the discontinuity at the " boundary sphere " $r = l$ is more than just a discontinuity in the characteristic impedance of the radial transmission line; the set of transmission modes for $r > l$ is different from the set for $r < l$. The theory of wave transmission on such terminated cones and on wires of other shapes will be considered in Chapter 11.

8.13. *Transverse Electromagnetic Waves on a Cylindrical Wire*

The theory of cone transmission lines can be extended to cylindrical wires (Fig. 8.31) of sufficiently small radius a since such wires will support nearly spherical waves. In this case the inductance and capacitance vary

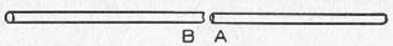


FIG. 8.31. A cylindrical wire energized between points A and B.

with the distance r from the origin and by (12–6) may be expressed approximately as follows

$$L = \frac{\mu}{\pi} \log \frac{2r}{a}, \quad C = \frac{\pi\epsilon}{\log \frac{2r}{a}}. \tag{13–1}$$

When L and C are varying slowly with r, we have approximately

$$K(r) = \sqrt{\frac{L}{C}} = \frac{\eta}{\pi} \log \frac{2r}{a} = 120 \log \frac{2r}{a}. \tag{13–2}$$

Since $Z(r)Y(r)$ is constant, we can use (7.12–5) for obtaining the functions involved in the second approximation to the voltage and current in the present transmission line; thus

$$K_0 = \frac{1}{l}\int_0^l K(r)\,dr = 120 \log \frac{2l}{a} - 120,$$

$$A(r) = \frac{i\beta}{K_0}\int_0^r [K(r) - K_0] \cos 2\beta r\,dr, \tag{13–3}$$

$$B(r) = \frac{\beta}{K_0}\int_0^r [K_0 - K(r)] \sin 2\beta r\,dr.$$

For further convenience we introduce the following functions

$$M(r) = K_0 B(r), \quad N(r) = iK_0 A(r).$$

Then we have

$$M(r) = 60(1 - \cos 2\beta r)(\log 2\beta l - 1) - 60\int_0^{2\beta r} \sin x \log x\,dx$$

$$= 60\,(1 - \cos 2\beta r)\left(\log \frac{l}{r} - 1\right) + 60(\log 2\beta r - \text{Ci}\,2\beta r + C),$$

$$N(r) = 60(\log 2\beta l - 1) \sin 2\beta r - 60\int_0^{2\beta r} \cos x \log x\,dx$$

$$= 60\left(\log \frac{l}{r} - 1\right) \sin 2\beta r + 60\,\text{Si}\,2\beta r;$$

$$M(l) = 60(\log 2\beta l - \text{Ci}\,2\beta l + C - 1 + \cos 2\beta l),$$

$$N(l) = 60\,(\text{Si}\,2\beta l - \sin 2\beta l).$$

The voltage-current equations become

$$V(r) = V(0)\left[\cos \beta r + \frac{M(r)}{K_0} \cos \beta r - \frac{N(r)}{K_0} \sin \beta r\right]$$

$$-iK_0 I(0)\left[\sin \beta r - \frac{M(r)}{K_0} \sin \beta r - \frac{N(r)}{K_0} \cos \beta r\right], \tag{13–4}$$

$$I(r) = -i\frac{V(0)}{K_0}\left[\sin \beta r + \frac{M(r)}{K_0} \sin \beta r + \frac{N(r)}{K_0} \cos \beta r\right]$$

$$+I(0)\left[\cos \beta r - \frac{M(r)}{K_0} \cos \beta r + \frac{N(r)}{K_0} \sin \beta r\right].$$

We shall now calculate the approximate input impedance of an infinitely long wire. Taking $l = \lambda/2$ we obtain from (2) the impedance looking outward from $r = l$; thus $Z(\lambda/2) = K(\lambda/2) = 120 \log \lambda/a$. On the other hand from (4) we obtain approximately

$$Z(0) = \frac{V(0)}{I(0)} \simeq 120 \log \frac{\lambda}{a} - 2M\left(\frac{\lambda}{2}\right) - 2iN\left(\frac{\lambda}{2}\right)$$

$$= 120\left(\log \frac{\lambda}{2\pi a} + \mathrm{Ci}\, 2\pi - C - i\, \mathrm{Si}\, 2\pi\right).$$

For a spherical or a hemispherical resonator with a cylinder running from the center to the periphery (Fig. 8.32) we can find the resonant length by setting $I(0)$ and $V(l)$ equal to zero. Thus at resonance we have

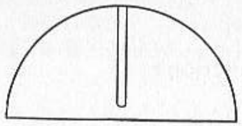


FIG. 8.32. A hemispherical cavity with a conducting cylinder inside.

$$\cos \beta l + \frac{M(l)}{K_0} \cos \beta l - \frac{N(l)}{K_0} \sin \beta l = 0,$$

$$\cot \beta l = \frac{N(l)}{K_0 + M(l)}.$$

From this we obtain approximately

$$\beta l = \frac{\pi}{2} - \frac{1.85}{2 \log \frac{\lambda}{2a} - 0.352}, \quad \frac{l}{\lambda} = \frac{1}{4} - \frac{0.295}{2 \log \frac{\lambda}{2a} - 0.352}.$$

In practice there is usually some capacitance C_0 at the center and the resonance conditions become

$$V(l) = 0, \quad \frac{I(0)}{V(0)} = -i\omega C_0 = -i\beta v C_0,$$

where v is the characteristic velocity. Substituting these values in (4) we obtain equations from which l/λ can be determined.

8.14. *Waves on Inclined Wires*

The results of the preceding sections can be generalized to cover the case of wires diverging from a common point and making an angle less than 180 degrees with each other. We start with two diverging conical conductors (Fig. 8.33). Using (12–1) and (3.6–15), we construct the following stream function for two infinitely thin wires along arbitrary radii $\theta = \theta_1, \varphi = \varphi_1$ and $\theta = \theta_2, \varphi = \varphi_2$

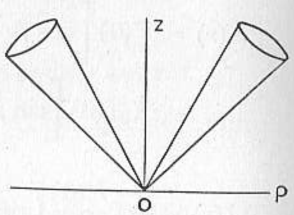


FIG. 8.33. Two diverging cones.

$$A = \frac{I}{4\pi} \log \frac{\left(e^{i\varphi} \cot \frac{\theta}{2} - e^{i\varphi_1} \cot \frac{\theta_1}{2}\right)\left(e^{-i\varphi} \cot \frac{\theta}{2} - e^{-i\varphi_1} \cot \frac{\theta_1}{2}\right)}{\left(e^{i\varphi} \cot \frac{\theta}{2} - e^{i\varphi_2} \cot \frac{\theta_2}{2}\right)\left(e^{-i\varphi} \cot \frac{\theta}{2} - e^{-i\varphi_2} \cot \frac{\theta_2}{2}\right)}.$$

The only singularities of this function occur at points for which either the numerator or the denominator vanishes; that is, along the above mentioned radii. The coefficient can be verified by obtaining the magnetic intensity and integrating it round an infinitely small circle surrounding either of the singular lines. The stream function A may be expressed in a form free of complex quantities

$$A = \frac{I}{4\pi} \log \frac{\cot^2 \frac{\theta}{2} - 2 \cot \frac{\theta}{2} \cot \frac{\theta_1}{2} \cos (\varphi - \varphi_1) + \cot^2 \frac{\theta_1}{2}}{\cot^2 \frac{\theta}{2} - 2 \cot \frac{\theta}{2} \cot \frac{\theta_2}{2} \cos (\varphi - \varphi_2) + \cot^2 \frac{\theta_2}{2}}.$$

It is now easy to determine the characteristic impedance of two diverging conical wires whose cone angles ψ_1 and ψ_2 are small compared with the angle ϑ between the axes of the cones. The restriction is imposed to assure that the proximity effect is small.* The impedance is

$$K = \frac{\eta}{\pi} \log \frac{2 \sin \frac{\vartheta}{2}}{\sqrt{\psi_1 \psi_2}} = \frac{\eta}{\pi} \log \frac{d(r)}{\sqrt{a_1(r) a_2(r)}}, \tag{14–1}$$

where $d(r)$ is the distance between the points on the axes of the conical wires which are at distance r from point O (Fig. 8.34) and $a_1(r)$ and $a_2(r)$ are the radii of the cones at these points. Only the distance between the elements of the wires and the radii of the wires occur in this expression which may, therefore, be taken as the "nominal" characteristic impedance (that is $\sqrt{L/C}$) of wires of other than conical shapes.

In the case of cylindrical wires of radius a we have

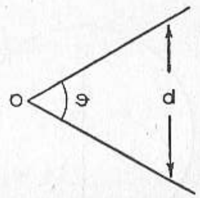


FIG. 8.34. Two diverging wires.

$$K = \frac{\eta}{\pi} \log \frac{2r \sin \frac{\vartheta}{2}}{a}, \quad K_0 = \frac{\eta}{\pi}\left(\log \frac{2l}{a} - 1 + \log \sin \frac{\vartheta}{2}\right),$$

where K_0 is the average characteristic impedance of a section of length l.

Two parallel wires of radius a, with interaxial separation s fairly large compared with a, carrying currents in the same direction, are approximately equivalent to a single wire of radius $\sqrt{as}$ in so far as the inductance and capacitance are concerned. This approximate equivalence applies also to a pair of diverging wires. If two such pairs are used for a transmission line of the type shown in Fig. 8.35, the nominal characteristic impedance is made more nearly constant.

Consider now a wire of radius a bent into a circle of radius b (Fig. 8.36). If η is replaced by μ in (1), we have a formula for the inductance per unit length. Integrating over the circle, we obtain an approximate value for the inductance of a circular turn of wire when the length of the circumference is small so that we have to deal with

* The proximity effect can be included without too much added labor.

a short transmission line. Thus we have

$$L = \frac{\mu b}{\pi}\int_0^{\pi}\left(\log\frac{2b}{a} + \log\sin\psi\right)d\psi = \mu b\log\frac{b}{a}. \tag{14–2}$$

If the length of the circumference is large, then we should treat the wire as a nonuniform transmission line in order to obtain its input reactance. In this case, however, further corrections should be introduced on account of radiation and the problem belongs to antenna theory.

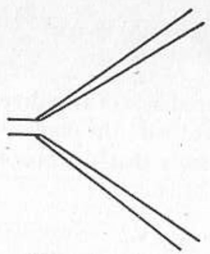


FIG. 8.35. Two pairs of diverging wires.

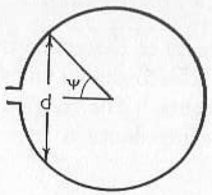


FIG. 8.36. A wire loop.

8.15. *Circular Magnetic Waves Inside a Hollow Metal Sphere*

A wave is called *circular magnetic* if the magnetic lines are circles. Waves generated by an electric current element or by currents in a straight wire are circular magnetic. In this section we shall consider such waves inside a hollow metal sphere, assuming that they are generated by an antenna along a diameter (Fig. 8.37). Let the length $2l$ of the antenna be small; then the current distribution is nearly linear and equations (6.2–11) represent the free-space field generated by this antenna. In these equations I is to be interpreted as the input current at the center of the antenna.*

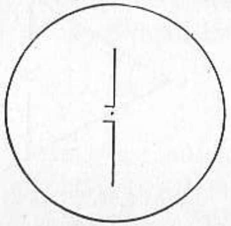


FIG. 8.37. A spherical cavity with a doublet antenna along its axis.

The metal sphere at $r = a$ gives rise to reflection. The reflected wave depends on θ in the same way as the incident wave. If we assume that the reflected wave is of stationary type without singularities at the origin, there will be no need to consider multiple reflections from the center; the field of this wave may be obtained from the following vector potential: $A_z = \sin\beta r/r$. Thus

$$E_\theta^- = -\frac{\eta P}{\lambda r}\left(\sin\beta r + \frac{\cos\beta r}{\beta r} - \frac{\sin\beta r}{\beta^2 r^2}\right)\sin\theta,$$

$$E_r^- = \frac{\eta P}{\pi r^2}\left(\frac{\sin\beta r}{\beta r} - \cos\beta r\right)\cos\theta,\quad H_\varphi^- = \frac{iP}{\lambda r}\left(\frac{\sin\beta r}{\beta r} - \cos\beta r\right)\sin\theta, \tag{15–1}$$

* The present antenna of length $2l$ is equivalent to a uniform electric current element of moment Il.

where the constant P is determined by the boundary condition at the surface of the sphere. If the latter is a perfect conductor, then the total E_θ should vanish at $r = a$ and therefore

$$P = \tfrac{1}{2}iIle^{-i\beta a}\frac{(\beta^2a^2 - 1) - i\beta a}{(\beta^2a^2 - 1)\sin\beta a + \beta a\cos\beta a}.$$

On the axis of the antenna near its center, we have

$$E_r^-(r) = \frac{\eta\beta^2 P}{3\pi}(1 - 0.1\beta^2r^2). \tag{15–2}$$

Hence the mutual impedance between the antenna and the sphere is approximately

$$Z_M = -\frac{E_r(0)l}{I} = -\frac{\eta\beta^2 lP}{3\pi I}.$$

Adding this mutual impedance to the self-impedance of the antenna, we obtain the total input impedance of the antenna inside a perfectly conducting metal sphere. The input impedance must be a pure reactance since under the assumed conditions no power is either dissipated or radiated. The radiation resistance of the antenna in free space must be canceled by an equal but negative resistance component of Z_M. Separating the real and the imaginary parts of P, we obtain

$$P = \tfrac{1}{2}iIl\frac{(\beta^2a^2 - 1)\cos\beta a - \beta a\sin\beta a}{(\beta^2a^2 - 1)\sin\beta a + \beta a\cos\beta a} + \tfrac{1}{2}Il.$$

Hence the real component of Z_M is $R_M = -(\eta/6\pi)\beta^2l^2$, as we have already anticipated.

The self-reactance X_0 of the antenna is largely capacitive and it may be obtained from (13–3); thus*

$$X_0 = -\frac{\eta}{\pi\beta l}\left(\log\frac{2l}{b} - 1\right),$$

where b is the radius of the antenna. Hence the total input impedance of the antenna is

$$Z = \frac{i\eta\beta^2l^2}{6\pi}\frac{\beta a\sin\beta a - (\beta^2a^2 - 1)\cos\beta a}{\beta a\cos\beta a + (\beta^2a^2 - 1)\sin\beta a} - \frac{i\eta}{\pi\beta l}\left(\log\frac{2l}{b} - 1\right). \tag{15–3}$$

This input impedance becomes infinite when

$$\tan\beta a = \frac{\beta a}{1 - \beta^2a^2}\quad\text{or}\quad\cot\beta a = \frac{1}{\beta a} - \beta a. \tag{15–4}$$

* The distributed impedance and admittance per unit length of the line are $i\omega L = i\beta K$ and $i\omega C = i\beta/K$.

This equation determines the frequencies for resonance and also for natural oscillations inside a hollow metal sphere; at these frequencies a field of type (1) can exist without a continuous impressed force. The smallest root of (4) and the corresponding natural wavelength are

$$\beta a = 2.744, \quad \hat{\lambda} = 2.290a. \tag{15-5}$$

The larger roots of (4) are nearly equal to $n\pi$ and hence are approximately given by

$$\beta a = n\pi - \frac{n\pi}{n^2\pi^2 - 1}, \quad \frac{2a}{\lambda} = n - \frac{n}{n^2\pi^2 - 1}. \tag{15-6}$$

The Q of the spherical cavity might be calculated by the method we have used so often in preceding sections. However, there exists another method which we shall now illustrate. If there is no antenna in the cavity the field must be of the form (1). At the surface of the sphere the sum of the inward looking radial impedance and the surface impedance of the sphere must vanish; thus $Z_r^-(a) + \hat{\eta} = 0$, where the surface impedance of the sphere has been assumed equal to the intrinsic impedance $\hat{\eta} = \mathcal{R}(1 + i)$. In view of (1) this condition becomes

$$\frac{(\beta^2a^2 - 1)\sin\beta a + \beta a\cos\beta a}{\beta a(\sin\beta a - \beta a\cos\beta a)} = -\frac{i\hat{\eta}}{\eta}. \tag{15-7}$$

When $\hat{\eta} = 0$, the principal solution of this equation is given by (5) and in general by (6). Since $\hat{\eta}/\eta$ is very small, an excellent approximation to any solution of (7) can be obtained by assuming

$$\beta a = k + \Delta, \tag{15-8}$$

where k is the corresponding solution of (4) and Δ is a small quantity. Retaining only the first powers of Δ, we reduce (7) to

$$\frac{\Delta N'(k)}{D(k)} = -\frac{i\hat{\eta}}{\eta},$$

where $N'(k)$ is the derivative of the numerator and $D(k)$ is the denominator; thus

$$\Delta = -\frac{i\hat{\eta}D(k)}{\eta N'(k)} = \frac{\mathcal{R}D(k)}{\eta N'(k)} - i\frac{\mathcal{R}D(k)}{\eta N'(k)}.$$

The natural frequency is no longer real and it is better to introduce the natural oscillation constant $p = i\omega$ in (8); thus we have

$$p = \frac{ik + i\Delta}{a\sqrt{\mu\epsilon}}.$$

On the other hand the oscillation constant may be expressed in terms of Q and the new real natural frequency ω

$$p = \xi + i\omega = -\frac{\omega}{2Q} + i\omega = i\omega\left(1 + \frac{i}{2Q}\right).$$

From the above equations we observe that the natural frequency is slightly affected by the imperfect conductivity of the sphere. For Q we obtain

$$\frac{1}{2Q} = -\frac{\mathcal{R}D(k)}{k\eta N'(k)}, \quad Q = -\frac{k\eta N'(k)}{2\mathcal{R}D(k)}.$$

Since $N'(k) = k\sin k + k^2\cos k$, we have in view of (4)

$$\frac{N'(k)}{D(k)} = \frac{\sin k + k\cos k}{\sin k - k\cos k} = \frac{2 - k^2}{k^2}; \quad Q = \tfrac{1}{2}k\left(1 - \frac{2}{k^2}\right)\frac{\eta}{\mathcal{R}}.$$

For the principal resonance we have

$$Q = 1.0076\frac{\eta}{\mathcal{R}} \simeq \frac{\eta}{\mathcal{R}}, \quad \text{and in air} \quad Q = \frac{380}{\mathcal{R}}.$$

At resonance the electric intensity is maximum at the center of the sphere. The voltage V along the diameter can be expressed in terms of this intensity E_0; thus

$$V = 2\int_0^a E_r\,dr = \frac{3E_0\lambda}{2\pi}\left(\mathrm{Si}\,k + \frac{\cos k}{k} - \frac{\sin k}{k^2}\right).$$

For the principal resonance we have $V = 0.686E_0\lambda$. At the boundary of the sphere the field intensities are

$$E_r(a) = \frac{3E_0}{k^2}\left(\frac{\sin k}{k} - \cos k\right)\cos\theta,$$

$$H_\varphi(a) = i\frac{1.5E_0}{\eta k}\left(\frac{\sin k}{k} - \cos k\right)\sin\theta.$$

For the principal resonance we obtain $E_r(a) = 0.423E_0\cos\theta$.

Returning to the impedance function (3), we shall now calculate its zeros. Since the second term (the self-impedance of the antenna) is rather large and the mutual impedance is rather small except near resonant points of the spherical resonator, Z should vanish at frequencies not very different from resonant frequencies of the sphere. The equation which we have to solve is

$$\frac{(\beta a)^3[\beta a\sin\beta a - (\beta^2a^2 - 1)\cos\beta a]}{\beta a\cos\beta a + (\beta^2a^2 - 1)\sin\beta a} = 6\left(\frac{a}{l}\right)^3\left(\log\frac{2l}{b} - 1\right).$$

Writing $\beta a = k + \delta$, where k is a solution of (4), substituting in the above equation, and retaining only the first powers of δ, we obtain

$$\delta = -\frac{k(k^4 - k^2 + 1)l^3}{6(k^2 - 2)a^3\left(\log \dfrac{2l}{b} - 1\right)}.$$

This approximation is not good when k becomes so large that δ is no longer small. The large zeros of Z would seem to be nearer to the zeros of the numerator than to those of the denominator. The smallest zero of Z is somewhat smaller than the smallest infinity (except for that at the origin). This was to be expected because the zeros and infinities of a reactance function separate each other. The fractional deviation of the first zero from the nearest infinity is δ/k. If $l = 0.1a$ and $b = 0.1l$, the fractional deviation is 0.0007576 or 0.076 per cent.

8.16. *Circular Electric Waves Inside a Hollow Sphere*

Circular electric waves are waves with circular electric lines of force. Such waves can be generated by a uniform circular current filament. The current distribution in a small electric current loop is approximately uniform even if the loop is energized at one point only; thus small loops can be used to generate circular electric waves. In free space the field of a loop of area S carrying current I is given by equations (6.17–5).

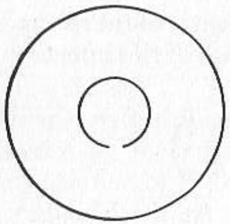


FIG. 8.38. A spherical cavity with a loop inside it.

Consider now a loop concentric with a sphere of radius a (Fig. 8.38). If we choose the plane of the loop as the equatorial plane of our system of coordinates, the field reflected from the sphere can be obtained from the following electric vector potential: $F_z = \sin \beta r/r$. Thus we have

$$H_{\bar{\theta}} = -\frac{P}{\eta\lambda r}\left(\sin \beta r + \frac{\cos \beta r}{\beta r} - \frac{\sin \beta r}{\beta^2 r^2}\right)\sin \theta,$$

$$H_{\bar{r}} = \frac{P}{\pi\eta r^2}\left(\frac{\sin \beta r}{\beta r} - \cos \beta r\right)\cos \theta, \quad E_{\bar{\varphi}} = \frac{iP}{\lambda r}\left(\cos \beta r - \frac{\sin \beta r}{\beta r}\right)\sin \theta. \tag{16–1}$$

If the sphere is a perfect conductor, then the total tangential electric intensity vanishes at the surface and

$$P = \tfrac{1}{2}\eta\beta S I e^{-i\beta a}\frac{1 + i\beta a}{\beta a \cos \beta a - \sin \beta a}$$

$$= \tfrac{1}{2}\eta\beta S I\frac{\beta a \sin \beta a + \cos \beta a}{\beta a \cos \beta a - \sin \beta a} + \tfrac{1}{2}i\eta\beta S I.$$

Near the center the electric intensity in the plane of the loop is

$$E_{\bar{\varphi}}(r) = -\frac{i}{6\pi}\beta^3 r P(1 - 0.1\beta^2 r^2).$$

Hence the mutual impedance between the loop and the sphere is

$$Z_M = -\frac{2\pi b E_{\bar{\varphi}}(b)}{I} = \frac{i\beta^3 b^2 P}{3I},$$

where b is the radius of the loop. The resistance component of Z_M is the negative of the radiation resistance of the loop. Thus the input impedance of the loop is a pure reactance

$$Z = \frac{i\eta}{6\pi}\beta^4 S^2 \frac{\beta a \sin \beta a + \cos \beta a}{\beta a \cos \beta a - \sin \beta a} + i\eta\beta b \log \frac{b}{c}, \tag{16–2}$$

where the second term represents the reactance of the loop in free space (c is the radius of the wire).

The input impedance becomes infinite when

$$\tan \beta a = \beta a, \quad \beta a \neq 0. \tag{16–3}$$

When the frequency is zero, Z is zero and not infinite; hence the exception. The lowest root of this equation is $\beta a = 4.493\cdots$, and $\hat{\lambda} = 1.398a$, $a = 0.715\hat{\lambda}$. The first zero of the input impedance is at $f = 0$. The remaining zeros are obtained from the following equation

$$\frac{\beta^3 a^3(\beta a \sin \beta a + \cos \beta a)}{\beta a \cos \beta a - \sin \beta a} = -\frac{6a^3}{\pi b^3}\log \frac{b}{c}. \tag{16–4}$$

For small loops the right side of the equation is large and the zeros of Z are near its infinities. It is not surprising that the zeros and infinities of Z should be close to each other; the former are the natural frequencies of the spherical cavity with a small perfectly conducting ring at the center while the latter are the natural frequencies of the cavity when the ring is open. If the ring is small it should make little difference whether it is open or shorted.

We leave it to the reader to show that the approximate solution of (4) is

$$\beta a = k\left[1 + \frac{\pi(k^2 + 1)b^3}{6a^3 \log \dfrac{b}{c}}\right],$$

where k is a solution of (3). It is also left to the reader to show that the equation for natural oscillations inside an imperfectly conducting sphere and the Q of the cavity are

$$\frac{\beta a(\beta a \cos \beta a - \sin \beta a)}{(\beta^2 a^2 - 1)\sin \beta a + \beta a \cos \beta a} = -\frac{i\hat{\eta}}{\eta}, \quad Q = \frac{k\eta}{2\mathcal{R}}.$$

8.17. *Two-Dimensional Fields*

In this section we shall obtain basic wave functions independent of the z-coordinate. Such functions are solutions of (4.12–6) and (4.12–7). The first set represents transverse electric cylindrical waves with the E-vector parallel and the H-vector normal to the axis of the waves; the second set represents transverse magnetic cylindrical waves with the H-vector parallel and the E-vector normal to the axis. Instead of solving the equations directly, we may obtain the complete set of wave functions from the wave functions considered in sections 5 and 7.

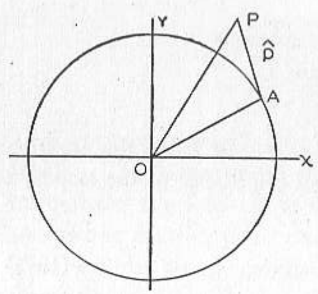


FIG. 8.39. Illustrating translation of Bessel functions from one origin to another.

Let us start with a uniform infinitely thin electric current filament parallel to the z-axis and passing through the point $A(\rho_0,\varphi_0)$ (Fig. 8.39). The field of this filament is

$$E_z(\hat{\rho}) = -\frac{i\omega\mu I}{2\pi}K_0(\sigma\hat{\rho}), \quad H_\varphi(\hat{\rho}) = \frac{\sigma I}{2\pi}K_1(\sigma\hat{\rho}), \quad (17\text{–}1)$$

where $\hat{\rho}$ is the distance from the filament. Let this field be referred to the coordinate system whose origin is at O. Evidently E_z is a periodic function of φ and we should be able to represent it as a Fourier series in φ. Indeed, in the theory of Bessel functions the following expansions have been established

$$\begin{aligned} K_0(\sigma\hat{\rho}) &= K_0(\sigma\rho_0)I_0(\sigma\rho) + 2\sum_{n=1}^{\infty} K_n(\sigma\rho_0)I_n(\sigma\rho)\cos n(\varphi - \varphi_0), \quad \rho < \rho_0, \\ &= I_0(\sigma\rho_0)K_0(\sigma\rho) + 2\sum_{n=1}^{\infty} I_n(\sigma\rho_0)K_n(\sigma\rho)\cos n(\varphi - \varphi_0), \quad \rho > \rho_0. \end{aligned} \quad (17\text{–}2)$$

Each term of the first expansion represents a cylindrical wave of the stationary type and the second expansion is composed of progressive waves traveling outward. The magnetic intensity in terms of the new coordinates may be obtained from

$$H_\rho = -\frac{1}{i\omega\mu\rho}\frac{\partial E_z}{\partial\varphi}, \quad H_\varphi = \frac{1}{i\omega\mu}\frac{\partial E_z}{\partial\rho}. \quad (17\text{–}3)$$

Thus let

$$E_z(\rho_0,\varphi) = \sum_{n=0}^{\infty} E_n \cos n(\varphi - \varphi_0), \quad (17\text{–}4)$$

where

$$E_0 = -\frac{i\omega\mu I}{2\pi}I_0(\sigma\rho_0)K_0(\sigma\rho_0), \quad E_n = -\frac{i\omega\mu I}{\pi}I_n(\sigma\rho_0)K_n(\sigma\rho_0), \quad n \neq 0; \quad (17\text{–}5)$$

then the complete expressions for the electric intensity become

$$\begin{aligned} E_z(\rho,\varphi) &= \sum_{n=0}^{\infty} E_n \frac{I_n(\sigma\rho)}{I_n(\sigma\rho_0)}\cos n(\varphi - \varphi_0), \; \rho < \rho_0, \\ &= \sum_{n=0}^{\infty} E_n \frac{K_n(\sigma\rho)}{K_n(\sigma\rho_0)}\cos n(\varphi - \varphi_0), \; \rho > \rho_0. \end{aligned} \quad (17\text{–}6)$$

From (3) and (6) we now obtain the radial component of the magnetic intensity

$$\begin{aligned} H_\rho(\rho,\varphi) &= \frac{1}{\eta}\sum_{n=1}^{\infty} nE_n \frac{I_n(\sigma\rho)}{\sigma\rho I_n(\sigma\rho_0)}\sin n(\varphi - \varphi_0), \; \rho < \rho_0, \\ &= \frac{1}{\eta}\sum_{n=1}^{\infty} nE_n \frac{K_n(\sigma\rho)}{\sigma\rho K_n(\sigma\rho_0)}\sin n(\varphi - \varphi_0), \; \rho > \rho_0; \end{aligned} \quad (17\text{–}7)$$

and the circular component

$$\begin{aligned} H_\varphi(\rho,\varphi) &= \frac{1}{\eta}\sum_{n=0}^{\infty} E_n \frac{I_n'(\sigma\rho)}{I_n(\sigma\rho_0)}\cos n(\varphi - \varphi_0), \; \rho < \rho_0, \\ &= \frac{1}{\eta}\sum_{n=0}^{\infty} E_n \frac{K_n'(\sigma\rho)}{K_n(\sigma\rho_0)}\cos n(\varphi - \varphi_0), \; \rho > \rho_0, \end{aligned} \quad (17\text{–}8)$$

where the prime indicates differentiation with respect to $\sigma\rho$.

The radial impedances for the cylindrical waves of order n are

$$Z^+_{\rho,n}(\rho) = -\eta\frac{K_n(\sigma\rho)}{K_n'(\sigma\rho)}, \quad Z^-_{\rho,n}(\rho) = \eta\frac{I_n(\sigma\rho)}{I_n'(\sigma\rho)}. \quad (17\text{–}9)$$

When $\sigma\rho$ and $\sigma\rho_0$ are small, as is the case in nondissipative media at short distances from the source then from (3.4-7) we obtain the following approximations

$$E_0 = -\frac{i\omega\mu I}{2\pi}K_0(\sigma\rho_0), \quad E_n = -\frac{i\omega\mu I}{2\pi n}. \quad (17\text{–}10)$$

The electric intensity becomes

$$\begin{aligned} E_z(\rho,\varphi) &= \sum_{n=0}^{\infty} E_n \frac{\rho^n}{\rho_0^n}\cos n(\varphi - \varphi_0), \; \rho < \rho_0, \\ &= \sum_{n=1}^{\infty} E_n \frac{\rho_0^n}{\rho^n}\cos n(\varphi - \varphi_0) + \frac{i\omega\mu I}{2\pi}(\log\sigma\rho + C - \log 2), \; \rho > \rho_0. \end{aligned} \quad (17\text{–}11)$$

For the radial component of H we have

$$\begin{aligned} H_\rho(\rho,\varphi) &= -\frac{I}{2\pi}\sum_{n=1}^{\infty}\frac{\rho^{n-1}}{\rho_0^n}\sin n(\varphi - \varphi_0), \; \rho < \rho_0, \\ &= -\frac{I}{2\pi}\sum_{n=1}^{\infty}\frac{\rho_0^n}{\rho^{n+1}}\sin n(\varphi - \varphi_0), \; \rho > \rho_0. \end{aligned} \quad (17\text{–}12)$$

And finally for the circular component of H, we obtain

$$\begin{aligned} H_\varphi(\rho,\varphi) &= -\frac{I}{2\pi}\sum_{n=1}^{\infty}\frac{\rho^{n-1}}{\rho_0^n}\cos n(\varphi - \varphi_0), \; \rho < \rho_0, \\ &= \frac{I}{2\pi}\sum_{n=0}^{\infty}\frac{\rho_0^n}{\rho^{n+1}}\cos n(\varphi - \varphi_0), \; \rho > \rho_0. \end{aligned} \quad (17\text{–}13)$$

The approximate values of the radial impedances for $n > 0$ are then

$$Z^+_{\rho,n}(\rho) = Z^-_{\rho,n}(\rho) = \frac{i\omega\mu\rho}{n}. \tag{17-14}$$

Let $2m$ current filaments, carrying alternately equal and opposite currents, be equispaced on the surface of the cylinder $\rho = \rho_0$, and let the first filament carrying current I/m pass through the point $(\rho_0,0)$. Since

$$\sum_{p=0}^{2m-1} (-)^p \exp\left[in\left(\varphi - \frac{p\pi}{m}\right)\right] = 0, \quad \text{if } n \neq (2k+1)m,$$
$$= 2me^{in\varphi}, \quad \text{if } n = (2k+1)m, \tag{17-15}$$

the electric intensity of this set of filaments becomes

$$E_z(\rho,\varphi) = 2E_m \frac{I_m(\sigma\rho)}{I_m(\sigma\rho_0)} \cos m\varphi + 2E_{3m} \frac{I_{3m}(\sigma\rho)}{I_{3m}(\sigma\rho_0)} \cos 3m\varphi + \cdots,$$
$$= 2E_m \frac{K_m(\sigma\rho)}{K_m(\sigma\rho_0)} \cos m\varphi + 2E_{3m} \frac{K_{3m}(\sigma\rho)}{K_{3m}(\sigma\rho_0)} \cos 3m\varphi + \cdots. \tag{17-16}$$

Only the terms corresponding to the odd multiples of m remain.

In some circumstances only the first terms in (16) are significant. Thus if ρ_0 is small, then for $\rho \gg \rho_0$ we have approximately

$$E_z(\rho,\varphi) = -\frac{i\omega\mu I}{m\pi} \frac{K_m(\sigma\rho)}{K_m(\sigma\rho_0)} \cos m\varphi. \tag{17-17}$$

If ρ is also small while remaining substantially larger than ρ_0, we have the quasistatic case

$$E_z(\rho,\varphi) = -\frac{i\omega\mu I \rho_0^m}{m\pi\rho^m} \cos m\varphi. \tag{17-18}$$

Thus just as equations (1) represent a progressive cylindrical wave generated by a single electric current filament, various terms in the second line of equation (6) represent progressive cylindrical waves generated by groups of $2n$ current filaments, carrying alternately equal and opposite currents, equispaced on the surface of a cylinder of an infinitely small radius. Equation (6) shows that the field of a simple line source at $\rho = \rho_0$, at distance $\rho > \rho_0$, may be regarded as the resultant of an infinite number of multiple line sources at $\rho = 0$.

In the special case of two filaments separated by distance l and carrying equal and opposite currents, we have

$$E_z(\rho,\varphi) = -\frac{i\omega\mu\sigma Il}{2\pi} K_1(\sigma\rho) \cos\varphi; \tag{17-19}$$

in the quasistatic case this electric intensity and the corresponding magnetic intensities become

$$E_z(\rho,\varphi) = -\frac{i\omega\mu Il}{2\pi\rho} \cos\varphi, \quad H_\varphi = \frac{Il \cos\varphi}{2\pi\rho^2}, \quad H_\rho = -\frac{Il \sin\varphi}{2\pi\rho^2}. \tag{17-20}$$

The case of transverse magnetic cylindrical waves is very similar; we need only replace the electric current I by the magnetic current K, E_z by H_z, H_φ by $-E_\varphi$, and interchange $i\omega\mu$ and $(g + i\omega\epsilon)$.

Transverse magnetic waves may be generated by appropriate electric current distributions. Consider for example a uniform electric current " strip " of width s; on the edges of the strip there exist two equal line charges of opposite signs. Assuming that the strip is situated in the xz-plane along the z-axis, the magnetic vector potential of the field is

$$A_z = \frac{Js}{2\pi} K_0(\sigma\rho), \tag{17-21}$$

where J is the current and Js the moment per unit length of the strip. The form of this expression is obviously correct since the vector potential of the current strip is independent of φ and is parallel to the x-axis; that the constant coefficient is correct may be shown by calculating the field and studying it in the vicinity of $\rho = 0$. For the field we obtain

$$H_z = -\frac{\partial A_z}{\partial y} = \frac{\sigma Js}{2\pi} K_1(\sigma\rho) \sin\varphi,$$
$$E_\rho = \frac{1}{(g + i\omega\epsilon)\rho} \frac{\partial H_z}{\partial\varphi} = \frac{\eta Js}{2\pi\rho} K_1(\sigma\rho) \cos\varphi, \tag{17-22}$$
$$E_\varphi = -\frac{1}{g + i\omega\epsilon} \frac{\partial H_z}{\partial\rho} = -\frac{i\omega\mu Js}{2\pi} K_1'(\sigma\rho) \sin\varphi.$$

In the vicinity of $\rho = 0$ we have

$$E_\rho = \frac{Js \cos\varphi}{2\pi(g + i\omega\epsilon)\rho^2}, \quad E_\varphi = \frac{Js \sin\varphi}{2\pi(g + i\omega\epsilon)\rho^2}. \tag{17-23}$$

This intensity is the negative of the gradient of the following electric potential

$$V = \frac{Js \cos\varphi}{2\pi(g + i\omega\epsilon)\rho} = \frac{Js}{2\pi(g + i\omega\epsilon)} \frac{\partial}{\partial x} \log\rho. \tag{17-24}$$

Similarly for a magnetic current strip, we have

$$F_z = \frac{Ms}{2\pi} K_0(\sigma\rho), \quad E_z = \frac{\partial F_z}{\partial y} = -\frac{\sigma Ms}{2\pi} K_1(\sigma\rho) \sin\varphi,$$
$$H_\rho = \frac{Ms}{2\pi\eta\rho} K_1(\sigma\rho) \cos\varphi, \quad H_\varphi = -\frac{\sigma Ms}{2\pi\eta} K_1'(\sigma\rho) \sin\varphi, \tag{17-25}$$

where M is the magnetic current per unit length.

8.18. *Shielding Theory*

In order to shield an electric circuit from external influences the circuit is enclosed in a metal box. Cylindrical and spherical shields are the simplest from the analytical point of view and in practice the results obtained for these shields are found to be applicable to shields of other shapes. First we shall deal with shielding at fairly high frequencies when the attenuation of fields in passing through metallic substances plays an important role and later we shall see what happens at low frequencies where the major factor is the impedance mismatch at the shield and the resulting reflection loss.

First we shall consider a cylindrical shield and a pair of closely spaced parallel current filaments (Fig. 8.40). Let the inner radius of the shield be a and the outer b; let $b - a = t$. Assume that a is small; then the radial impedances in the dielectric are (see 17–14)

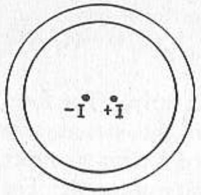


FIG. 8.40. Two parallel current filaments and a cylindrical shield.

$$Z_\rho^+ = Z_\rho^- = i\omega\mu_0\rho, \tag{18–1}$$

where μ_0 is the permeability of the dielectric. If the thickness of the shield is small compared with the radius, the divergence of the wave in passing through the shield may be ignored and the wave propagation in the shield may be regarded as substantially plane. Hence in the shield we have

$$Z_\rho^+ = Z_\rho^- = \eta = \sqrt{\frac{i\omega\mu}{g}}, \quad \sigma = \alpha(1 + i), \quad \alpha = \sqrt{\pi\mu f g}. \tag{18–2}$$

The attenuation of a wave passing through the shield once is then

$$A = \alpha t \text{ nepers} = 8.686\ \alpha t \text{ decibels}. \tag{18–3}$$

This wave is partially reflected at the outer surface of the shield and then partially re-reflected at the inner surface. The level of the re-reflected wave is lowered by $2\alpha t$ nepers and at high frequencies the re-reflected wave may be ignored. Let us suppose, for example, $f = 10^6$ and $t = 0.2$ mm; then the attenuation constant for copper is over 130 db per millimeter and the level of the second reflected wave is below the level of the first by 52 db (which means that the amplitude is reduced to 0.0024 of its original value). Thus for most purposes the shielding is due to a reflection loss at the boundaries of the shield and an attenuation loss expressed by (3). The reflection loss is the product of the transmission coefficients across the two boundaries of the shield; hence from (7.13–4) we obtain

$$p_H = p_E = \frac{4k}{(k + 1)^2}, \tag{18–4}$$

where k is the impedance ratio* $k = \eta/i\omega\mu_0 a$. In decibels the reflection

* We ignore the slight variation in the radial impedance (1) in passing from one surface to the other.

loss R becomes

$$R = 20 \log_{10} \frac{|k + 1|^2}{4|k|}. \tag{18–5}$$

When k is either small or large, we have respectively

$$R = 20 \log_{10} \frac{1}{4|k|}, \quad R = 20 \log_{10} \frac{|k|}{4}.$$

The total shielding improvement is the sum of the reflection and attenuation losses; thus $S = R + A$.

The reflection loss may be considerable. For a shield whose radius is 1 cm the radial impedance in air for $f = 10^6$ is $Z_\rho = 0.079i$ ohms. On the other hand for copper we have $Z_\rho = 0.000369\sqrt{i}$ ohms. Hence $1/|k| = 214$, and the reflection loss is

$$R = 20 \log_{10} \frac{214}{4} = 35 \text{ db}.$$

We have seen that at this frequency the attenuation loss in a 0.2 mm shield is about 26 db; thus the reflection loss is the larger part of the total shielding efficiency of 61 db.

For magnetic shields the reflection loss may be small. Consider, for example, an iron shield with relative permeability 100 and with conductivity only a fifth of that for copper. The radial impedance in the shield is then increased by a factor $\sqrt{500} \simeq 22$ which brings it up to $0.008\sqrt{i}$. The impedance ratio becomes nearly 0.1 and the reflection loss is comparatively small. In air the radial impedance is proportional to f but in metals it is proportional to $\sqrt{f}$: hence the magnitude of the foregoing impedance ratio becomes unity at about $f = 10^4$ and the reflection loss, due only to a 45° phase difference in impedances, is negligible from the shielding point of view. But the attenuation constant is larger in magnetic shields than in nonmagnetic. Thus for the iron shield just considered the attenuation constant is increased by a factor $\sqrt{20} = 4.47$. When the shield is thick this will more than compensate for the lack of reflection loss. However, if the shield is thin, the reflection loss may become more important than the attenuation loss and the nonmagnetic shield will be more effective than the magnetic shield. For two metals the impedance ratio

$$k = \frac{\sqrt{\mu_1 g_2}}{\sqrt{\mu_2 g_1}}$$

is independent of the frequency. The impedance mismatch, as between copper and iron, may be considerable and it is advantageous to use a shield consisting of composite layers of copper and iron, *with copper layers on the outside*, to utilize the substantial reflection losses at air-copper surfaces.

The complete formula for the shielding efficiency is obtained from equation (7.19–2); in decibels it becomes

$$S = -20 \log_{10} | T | = R + A + 20 \log_{10} \left| 1 - \frac{(k-1)^2}{(k+1)^2} e^{-2\sigma t} \right|. \quad (18\text{–}6)$$

This represents the difference in the amplitude levels of the incident wave just inside the shield and the transmitted wave just outside the shield and it is substantially equal to the difference in field levels before and after introducing the shield. In order to obtain the latter difference in levels we should multiply T by b/a for the electric intensity and by b^2/a^2 for the magnetic intensity.

If the line source is not along the axis, the shielding efficiency varies with position round the shield. The actual source may be replaced by an infinite number of virtual sources along the axis in accordance with the results obtained in the preceding section. One of the sources is the actual source translated to the axis. The remaining sources emit higher order cylindrical waves whose radial impedances vary somewhat with the order of the wave in air and are nearly constant in metals. Hence the field distribution is altered by the shield, though not very greatly. Except for this unevenness in the field reduction, the shielding efficiency is the same as for an axial source.

Let us now consider a spherical shield with a current loop at the center. The radial impedance in the shield is the same as for a cylindrical shield; but the radial impedances in air are

$$Z_r^+ = i\omega\mu_0 r, \quad Z_r^- = \tfrac{1}{2} i\omega\mu_0 r. \quad (18\text{–}7)$$

The external wave impedance is calculated from (6.17–5) and the internal from (16–1) assuming that r is small. The product of two transmission coefficients is now obtained from (7.20–2); thus

$$p = \frac{(Z_r^- + Z_r^+)2\eta}{(Z_r^- + \eta)(\eta + Z_r^+)} = \frac{3k}{(1+k)(1+\frac{1}{2}k)},$$

where $k = \eta / i\omega\mu_0 a$. When k is small, then $p = 3k$ and the reflection loss becomes

$$R = 20 \log_{10} \frac{1}{3\,|\,k\,|} = 20 \log_{10} \frac{1}{4\,|\,k\,|} + 2.5 \text{ db}.$$

The difference between the formulae for cylindrical and spherical shields consists in a 2.5 db improvement in shielding efficiency for the latter.

A cubical shield may be replaced by a spherical shield in estimates of the shielding efficiency. Since, in practice, high shielding efficiencies are required, one is not justified in looking for exact solutions, especially as such solutions show that in general

no shield has a single figure of merit, independent of the position of the point at which the shielding efficiency is measured.

In the two preceding examples the fields are largely "magnetic" since the wave impedances are very small quantities and, in elementary discussions of such fields, electric intensities are usually excluded from explicit consideration. Another type of field is the so-called "electrostatic" field which is produced, for example, by electric charges on the plates of a capacitor. Such "electrostatic" fields may vary at the rate of one million cycles per second and perhaps a different name would be more appropriate. What really matters is that some fields in the neighborhood of small structures are low impedance fields so that magnetic components are large and electric components small and other fields are high impedance fields with large electric and small magnetic components. The terms "low impedance" and "high impedance" are used here with reference to the intrinsic impedance of the dielectric in which the field happens to reside. Thus the field (17–23) of two charged filaments is a high impedance field since in a good dielectric

$$Z_\rho^+ = Z_\rho^- = \frac{1}{i\omega\epsilon\rho} = \frac{\eta}{i\beta\rho} = \frac{\eta\lambda}{2\pi i\rho},$$

which is large compared with η. A simple calculation shows that at a metal boundary a high impedance field is reflected so completely that there is no need to compute the shielding efficiency. The reflection loss tends to infinity as the frequency approaches zero.

The theory of shielding thus far developed may be called the "transmission theory of shielding" since our attention has been fixed on cylindrical waves passing through the shield. A different physical picture of shielding is possible. The fields are produced by electric currents, the original field by currents in the source along the axis and the final field by the currents induced in the shield. It may be said that the shielding action is caused by currents in the two halves of the shield flowing in directions opposite to the currents in the source. In Fig. 8.40 the current in the right half of the shield is negative and in the left half positive. The principal weakness in this theory is that frequently it is difficult to compute the induced currents without first solving the shielding problem itself. In order to find the induced currents we have to calculate the tangential magnetic intensity at both surfaces of the shield; but the shielding efficiency is then determined. On the other hand the second method suggests some deductions not immediately indicated by the first; thus from the second picture of the way in which the shielding is effected we conclude at once that a greater impairment in shielding will occur if the shield is cut so as to interfere with the induced current flow than if it is cut along the lines of flow. For example, if a plane wave is falling normally on a perfectly conducting screen with an infinitely long slit, more power will be transmitted when the E-vector is perpendicular to the slit than when it is parallel to it. The two physical pictures of shielding are complementary. In some cases it is more natural to use the second theory even in analytical work.

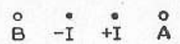


FIG. 8.41. Normal cross-section of two parallel wires at A and B which act as a shield for two current filaments.

An example is furnished by a "two-wire shield" (Fig. 8.41), for which it is easy to obtain the currents induced in the wires A and B and to show that these currents tend to weaken the original field at great distances.

We now turn our attention to magnetostatic shielding. At $f = 0$ the electric field vanishes and we lose our principal tool, the radial impedance. The transmission theory of shielding could be reformulated in terms of properly defined "radial reluctances"; but it is more convenient to regard steady magnetic fields as fields of vanishingly small frequency. The field of two closely spaced equal and opposite current filaments is then given by (17–20). The field reflected from the shield must be of the form

$$E_z^r = Q\rho \cos\varphi, \quad H_\varphi^r = \frac{Q}{i\omega\mu}\cos\varphi, \quad H_\rho^r = \frac{Q}{i\omega\mu}\sin\varphi. \tag{18–8}$$

The proportionality of E to $\cos\varphi$ follows from the boundary conditions; the proper exponent for ρ comes from (17–11); and then H is obtained from (17–3). In cartesian coordinates $H_x = 0$, $H_y = Q/i\omega\mu$; hence H is uniform. Thus the radial impedances are $Z_\rho^+ = Z_\rho^- = i\omega\mu\rho$ and the impedance ratio

$$k = \frac{\mu_2}{\mu_1} \tag{18–9}$$

is independent of the frequency.

The "attenuation" is due solely to the cylindrical divergence of the field and is independent of the medium; hence the shielding efficiency is due entirely to the reflection loss. The transfer ratio χ which enters the general expression (7.20–4) for the transmission coefficient T is obtained as follows. From the definitions in section 7.10 and from (17–20) and (8) we have

$$\chi_E^+(\rho_1,\rho_2) = \chi_E^-(\rho_2,\rho_1) = \frac{\rho_1}{\rho_2}, \quad \chi_H^+(\rho_1,\rho_2) = \frac{\rho_1^2}{\rho_2^2}, \quad \chi_H^- = 1.$$

If the inner radius of the shield is a and the outer b, the transfer ratio χ for a wave passing through the shield once outward and once inward is defined by (7.20–3) and is the product of the above transfer ratios; this total transfer ratio is the same for E and H

$$\chi = \frac{a^2}{b^2}. \tag{18–10}$$

Hence by (7.20–4) the shielding ratio s, defined as the ratio of the field outside the shield to the field which would exist there if the shield were removed is

$$s = \frac{p}{1 - q\chi}. \tag{18–11}$$

The factor $\chi^+(x_1,x_2)$ in the expression for T represents the transmission ratio across (x_1,x_2) for the original field as well as for the field modified by the shield; since s represents the "insertion" loss, this factor does not appear in (11). For the products p and q of the transmission and reflection coefficients across the two boundaries of the shield we have

$$p = \frac{4k}{(1+k)^2} = \frac{4\mu_1\mu_2}{(\mu_1+\mu_2)^2}, \quad q = \frac{(1-k)^2}{(1+k)^2} = \frac{(\mu_1-\mu_2)^2}{(\mu_1+\mu_2)^2}.$$

Thus we have determined all factors in the shielding ratio s.

When the difference $\mu_1 - \mu_2$ is large, then q is approximately equal to unity and $s = p/(1-\chi)$. The effectiveness of the shield depends largely on the ratio of the permeabilities. The thickness of the shield when increased beyond a certain value has relatively little effect. Thus if χ becomes small, we have simply $s = p$.

When shielding depends on the reflection loss the shielding is improved by using less, not more, shielding material. Thus if the shielding material extends from $\rho = 1$ to $\rho = 1000$, then $\chi = 10^{-6}$ and s is practically equal to p. Suppose now that the material between $\rho = 10$ and $\rho = 100$ is removed so that in effect we have two shields. For each of these shields $\chi = 0.01$ and the shielding ratio is still practically equal to p; hence the shielding ratio for both shields is p^2. Assuming shielding material extending from $\rho = 1$ to $\rho = 10^{2n+1}$ in layers, each layer extending from $\rho = 10^{2m}$ to $\rho = 10^{2m+1}$, we obtain the shielding ratio p^{n+1}. If the space between the above layers is also filled with the shielding material the shielding ratio becomes p. These figures illustrate the principle. It is not practicable to use shields with radii whose ratio is so large that the full benefit of reflection is obtained from each layer. A formula for a practical case may be obtained from (7.19–10) in which the exponential factors should be replaced by the transfer ratios. For a shield with only two layers calculations are not too laborious but for a large number of regularly arranged layers it would be preferable to consider first the problem of wave transmission through the corresponding iterated structure.

In the case of spherical shields equations (6.17–7) give the field of a current loop at distances which are large compared with the radius of the loop as long as the frequency is so small that the product βr is small, and from (16–1) we obtain the following form of the field reflected from a spherical shield concentric with the loop*

$$E_\varphi^- = i\omega\mu P r \sin\theta, \quad H_\theta^- = 2P\sin\theta, \quad H_r^- = -2P\cos\theta. \tag{18–12}$$

The magnetic intensity of the reflected field is parallel to the z-axis since $H_\rho = 0$,

* The numerical factor P has been changed. Once E_φ is obtained, H_θ and H_r can be calculated from (4.12–10).

$H_z = -2P$. The radial impedances are given by (7) and the transfer ratios are

$$\chi_E^+(r_1,r_2) = \frac{r_1^2}{r_2^2}, \quad \chi_E^-(r_2,r_1) = \frac{r_1}{r_2},$$
$$\chi_H^+(r_1,r_2) = \frac{r_1^3}{r_2^3}, \quad \chi_H^-(r_2,r_1) = 1.$$

The product which enters the shielding ratio (11) is then

$$\chi_E = \chi_E^+(r_1,r_2)\chi_E^-(r_2,r_1) = \frac{r_1^3}{r_2^3} = \chi_H.$$

For the products of the transmission and reflection coefficients across the two boundaries of the shield we have (with $k = \mu_2/\mu_1$)

$$p_H = \frac{(\frac{1}{2}+1)(\frac{1}{2}k+k)}{(\frac{1}{2}+k)(\frac{1}{2}k+1)} = \frac{9k}{(1+2k)(2+k)} = p_E.$$
$$-q_H = q_E = \frac{(\frac{1}{2}k-\frac{1}{2})(k-1)}{(k+\frac{1}{2})(\frac{1}{2}k+1)} = \frac{2(k-1)^2}{(1+2k)(2+k)}.$$

If k is large, then $q \simeq 1$; and the shielding ratio becomes

$$s = \frac{p}{1-\chi} = \frac{9}{2k(1-\chi)} = \frac{9\mu_1 b^3}{2\mu_2(b^3-a^3)},$$

where a is the inner radius of the shield and b the outer radius. Comparing this with the corresponding shielding ratio for cylindrical shields, we find that when b is much larger than a, the spherical shield is somewhat less effective than the cylindrical shield; but in practical cases, in which the thickness of the shields is moderate, the spherical shield may be the more effective.

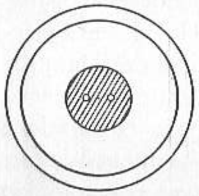


FIG. 8.42. A current loop surrounded by a magnetic core and a spherical shield.

The preceding specific formulae for the shielding ratio apply to the case when the permeability of the medium inside the shield is the same at all points. If the current loop is surrounded by some magnetic material (Fig. 8.42) as is the case for coils wound on magnetic cores, then the radial impedance looking inward from the inner surface of the shield is modified and the shielding ratio is altered. In order to obtain more general expressions for p and q in (11) we have to distinguish between radial impedances for waves in homogeneous media and radial impedances affected by reflection. We shall now designate the former by the letter K and the latter by Z. From equations (7.20–2) and (7.20–3), paying due attention to Fig. 7.24 which explains successive reflections within the shield, we obtain

$$p = \frac{[Z_1^-(a) + K_1^+(a)][K_2^-(b) + K_2^+(b)]}{[Z_1^-(a) + K_2^+(a)][K_2^-(b) + K_1^+(b)]},$$
$$q = \frac{[K_2^-(a) - Z_1^-(a)][K_2^+(b) - K_1^+(b)]}{[K_2^+(a) + Z_1^-(a)][K_2^-(b) + K_1^+(b)]}, \quad (18\text{–}13)$$

where the subscript 2 refers to the shield.

Let the radius of the spherical "core" surrounding the current loop be c. The radial impedance $Z_1^-(a)$ may be obtained from (7.10–8) if we let $x_1 = a$, $x_2 = c$ and $Z(c) = \frac{1}{2}i\omega\mu_3 c$, where μ_3 is the permeability of the core; thus

$$Z_1^-(a) = \tfrac{1}{2}i\omega\mu_1 a\frac{2(1-\hat{\chi}) + \hat{k}(1+2\hat{\chi})}{(2+\hat{\chi}) + \hat{k}(1-\hat{\chi})}, \quad \hat{\chi} = \frac{c^3}{a^3}, \quad \hat{k} = \frac{\mu_3}{\mu_1}. \quad (18\text{–}14)$$

The magnetic core has altered the impedance looking inward at the inner surface of the shield in the following ratio

$$\bar{k} = \frac{Z_1^-(a)}{K_1^-(a)} = \frac{2(1-\hat{\chi}) + \hat{k}(1+2\hat{\chi})}{(2+\hat{\chi}) + \hat{k}(1-\hat{\chi})}.$$

In practice μ_3 is usually much larger than μ_1 since the purpose of the core is to increase the inductance of the coil. Hence we have approximately

$$\bar{k} = \frac{1+2\hat{\chi}}{1-\hat{\chi}};$$

thus the impedance $Z_1^-(a)$ is increased by the presence of the core. This approximation, however, should not be used when $\hat{\chi}$ is near unity.

Expressing (13) in terms of various ratios, we obtain

$$p = \frac{(\frac{1}{2}\bar{k}+1)(\frac{1}{2}k+k)}{(\frac{1}{2}\bar{k}+k)(\frac{1}{2}k+1)} = \frac{3k(2+\bar{k})}{(k+2)(2k+\bar{k})},$$
$$q = \frac{(\frac{1}{2}k-\frac{1}{2}\bar{k})(k-1)}{(k+\frac{1}{2}\bar{k})(\frac{1}{2}k+1)} = \frac{2(k-1)(k-\bar{k})}{(k+2)(2k+\bar{k})},$$

where k and $\bar{k}$ retain their previous meanings. Without the magnetic core $\bar{k} = 1$ and the core makes $\bar{k}$ larger. Since k is large for an effective shield, an increase in $\bar{k}$ makes p larger while q is affected only slightly; hence the shield is less effective for a coil wound on a magnetic core than for a coil wound on a nonmagnetic core. If the core is made of the same material as the shield and if $c = a$, the shield loses its effectiveness completely.

The above examples show that transmission theory provides a method for solving problems on cylindrical and spherical shields under a variety of circumstances. This method is particularly useful from the computational point of view when the structure is fairly complex. Even in the case of a single layer shield for a coil with a magnetic core the solution of the boundary value problem would require writing expressions for the fields in the core, in the region between the core and the shield, in the shield and outside the shield. These equations would contain six arbitrary constants and we should have to use the six equations of continuity of the tangential components of E and H to determine these constants; and the labor involved in solving these equations is considerable except when the method used is essentially equivalent to the transmission method of handling the problem. From the mathematical point of view the transmission method consists in solving the boundary value problem for a single section, then applying the results to each additional section. Besides, the transmission theory enables us to make qualitative appraisals of the effectiveness of shields once we have mastered the simple but very fundamental ideas concerning the effect of impedance mismatch on reflection.

We shall now examine the effect of the presence of the shield on the source itself. For example, for the current loop at the center of a spherical shield the incident field is expressed by (6.17–7); the reflected electric intensity is then

$$E_\varphi^-(r) = q_E E_\varphi^+(a)\frac{r}{a} = -q_E\frac{i\omega\mu_1 SIr\sin\theta}{4\pi a^3},$$

where q_E is the E-reflection coefficient at $r = a$ which takes into account the shield and the medium beyond it. Hence by (7.20–2) we have*

$$q_E = \frac{K_r^-(a)[Z_r^+(a) - K_r^+(a)]}{K_r^+(a)[Z_r^+(a) + K_r^-(a)]} = \frac{Z_r^+(a) - i\omega\mu_1 a}{2[Z_r^+(a) + \frac{1}{2}i\omega\mu_1 a]}.$$

The reflected field is impressed on the current loop and introduces an impedance

$$Z_M = -\frac{2\pi c E_\varphi^-(c)}{I} = q_E\frac{i\omega\mu_1 Sc^2}{2a^3} = q_E\frac{i\omega\mu_1 S^2}{2\pi a^3},$$

where c is the radius of the loop.

In the case of nonmagnetic metal shields at frequencies which are not too low, $Z_r^+(a)$ is small compared with $i\omega\mu_1 a$ and for perfect conductors $Z_r^+(a)$ is always zero; then $q_E = -1$ and the impedance added in series with the loop becomes

$$Z_M = -\frac{i\omega\mu_1 S^2}{2\pi a^3}, \quad L_M = -\frac{\mu_1 S^2}{2\pi a^3}.$$

Thus the inductance of the loop is made smaller by the presence of the shield.

8.19 *Theory of Laminated Shields*

In the preceding section we have shown that a shield made of alternate layers with different permeabilities may be more effective than a solid shield. Laminated shields made of regularly spaced layers may be treated as special cases of iterative structures. As an illustration we shall consider a cylindrical shield made of coaxial layers with the permeabilities of the adjacent layers equal to μ_1 and μ_2. Let the common ratio of the inner radius of a layer to its outer radius be $\sqrt{\chi}$ and let the permeability of the layer between $\rho_1 = a$ and $\rho_2 = b = a/\sqrt{\chi}$ be μ_1.

First of all we calculate the transducer impedances

$$-E_1 = Z_{11}'H_1 + Z_{12}'H_2, \quad E_z(a) = E_1\cos\varphi, \quad E_z(b) = E_2\cos\varphi,$$
$$E_2 = Z_{21}'H_1 + Z_{22}'H_2, \quad H_\varphi(a) = H_1\cos\varphi, \quad H_\varphi(b) = H_2\cos\varphi, \tag{19–1}$$

from the general equations for cylindrical waves of order one

$$E_z = i\omega\mu_1\left(\frac{P}{\rho} + Q\rho\right)\cos\varphi, \quad H_\varphi = \left(-\frac{P}{\rho^2} + Q\right)\cos\varphi. \tag{19–2}$$

* Expressing $q_V = q_E$ in terms of impedances instead of admittances.

By assigning values to P and Q which make either H_1 or H_2 vanish we have

$$Z_{11}' = i\omega\mu_1 a\frac{1+\chi}{1-\chi}, \quad Z_{12}' = -i\omega\mu_1 a\frac{2}{1-\chi},$$
$$Z_{21}' = -i\omega\mu_1 b\frac{2\chi}{1-\chi}, \quad Z_{22}' = i\omega\mu_1 b\frac{1+\chi}{1-\chi}. \tag{19–3}$$

The next layer from $\rho_1 = b = a/\sqrt{\chi}$ to $\rho_2 = b/\sqrt{\chi} = a/\chi$ has the permeability μ_2; the impedance coefficients can be obtained by analogy with (3):

$$Z_{11}'' = i\omega\mu_2 b\frac{1+\chi}{1-\chi}, \quad Z_{12}'' = -i\omega\mu_2 b\frac{2}{1-\chi},$$
$$Z_{21}'' = -i\omega\mu_2 a\frac{2}{1-\chi}, \quad Z_{22}'' = i\omega\mu_2 a\frac{1+\chi}{\chi(1-\chi)}. \tag{19–4}$$

The coefficients of the transducer comprised of these two layers may be obtained directly from (7.26–1); thus if $k = \mu_2/\mu_1$, we have

$$Z_{11} = i\omega\mu_1 a\frac{(1+k)(1+\chi)^2 - 4\chi}{(1+k)(1-\chi^2)}, \quad Z_{12} = -i\omega\mu_1 a\frac{4k}{(1+k)(1-\chi^2)},$$
$$Z_{21} = -i\omega\mu_2 a\frac{4\chi}{(1+k)(1-\chi^2)}, \quad Z_{22} = i\omega\mu_2 a\frac{(1+k)(1+\chi)^2 - 4k\chi}{(1+k)\chi(1-\chi^2)}. \tag{19–5}$$

The impedance coefficients of the transducer consisting of the next pair of layers are obtained if we multiply (5) by $1/\chi$. Equations for the iterative structure are then obtained from the first and third equations of the set (7.26–0); thus

$$Z_{11}H_1 + Z_{12}H_2 = -E_1,$$
$$Z_{21}H_1 + \left(Z_{22} + \frac{Z_{11}}{\chi}\right)H_2 + \frac{Z_{12}}{\chi}H_3 = 0. \tag{19–6}$$

Let the H-transfer ratio be

$$\frac{H_3}{H_2} = \frac{H_2}{H_1} = \hat{\chi}. \tag{19–7}$$

Substituting from (5) and (7) in (6), we obtain

$$\frac{\hat{\chi}}{\chi} - 2A + \frac{\chi}{\hat{\chi}} = 0, \quad A = u(1 + vv),$$
$$u = \frac{(1+\chi)^2}{4\chi}, \quad v = \frac{(1-\chi)^2}{(1+\chi)^2}, \quad \nu = \frac{1}{2}\left(k + \frac{1}{k}\right) = \frac{1}{2}\left(\frac{\mu_2}{\mu_1} + \frac{\mu_1}{\mu_2}\right).$$

Solving, we have

$$\frac{\hat{\chi}}{\chi} = A \pm \sqrt{A^2 - 1}.$$

If $\mu_1 = \mu_2$, then $\nu = 1$ and

$$\frac{\hat{\chi}}{\chi} = \frac{1+\chi^2}{2\chi} \pm \frac{1-\chi^2}{2\chi} = \frac{1}{\chi}, \chi; \quad \hat{\chi} = 1, \chi^2.$$

The first value is the H-transfer ratio for the converging cylindrical wave in a homogeneous medium and the second is that for the diverging wave. Thus in the laminated medium we have

$$\hat{\chi}^+ = \chi(A - \sqrt{A^2-1}) = \frac{\chi}{A+\sqrt{A^2-1}},$$

$$\hat{\chi}^- = \chi(A + \sqrt{A^2-1}) = \frac{\chi}{A-\sqrt{A^2-1}}.$$

Dividing these transfer ratios by the corresponding ratios for a homogeneous medium we have

$$\bar{\chi} = \frac{\hat{\chi}^+}{\chi^2} = \frac{A-\sqrt{A^2-1}}{\chi}, \quad \frac{\hat{\chi}^-}{1} = \frac{\chi}{A-\sqrt{A^2-1}} = \frac{1}{\bar{\chi}}.$$

The quantity $\bar{\chi}$ represents the shielding attenuation ratio for each pair of layers; for n pairs the attenuation ratio will be $\bar{\chi}^n$.

If the layers are thin, χ is near unity and we may write $\chi = 1 - \delta$. Making the necessary substitutions, we obtain

$$A = 1 + \tfrac{1}{4}(\nu+1)\delta^2, \quad \sqrt{A^2-1} = \delta\sqrt{\frac{\nu+1}{2}},$$

$$\bar{\chi} = 1 - \left(\sqrt{\frac{\nu+1}{2}} - 1\right)\delta.$$

Supposing that the shielding space between $\rho_1 = a$ and $\rho_2 = b$ has been divided into $2n$ layers, we have $\delta = 1 - \sqrt[n]{a/b}$. Let

$$\frac{a}{b} = e^{-\Theta}, \quad \Theta = \log\frac{b}{a},$$

and substitute in the preceding equation; then

$$\delta = 1 - e^{-\Theta/n} = \frac{\Theta}{n} - \frac{\Theta^2}{2n^2} + \cdots.$$

As n increases, the attenuation across the shield becomes

$$\lim \bar{\chi}^n = \left(1 - \frac{\Gamma\Theta}{n}\right)^n = e^{-\Gamma\Theta},$$

$$\Gamma = \sqrt{\frac{\nu+1}{2}} - 1 = \frac{\mu_1+\mu_2}{2\sqrt{\mu_1\mu_2}} - 1.$$

Thus the diverging H-wave is attenuated by $\Gamma\Theta$ nepers besides being multiplied by the factor a^2/b^2. The converging wave is also attenuated at the same rate. Hence the cumulative effect of reflections has resulted in exponential attenuation.

The characteristic radial impedance in the positive ρ-direction in the laminated medium may be obtained from (6) and (7); thus

$$K^+ = -\frac{E_1}{H_1} = Z_{11} + Z_{12}\frac{H_2}{H_1} = Z_{11} + \hat{\chi}^+ Z_{12}, \quad \text{or} \quad K^+ = Z_{11} + \chi^2\bar{\chi}Z_{12}.$$

Similarly the characteristic impedance in the negative ρ-direction is obtained from the analog of equations (7.26–0); thus

$$K^- = Z_{22} + \frac{Z_{21}}{\hat{\chi}^-} = Z_{22} + \bar{\chi}Z_{21}.$$

In the limit as the thickness of each layer approaches zero we have

$$K^+ = K^- = i\omega\rho\sqrt{\mu_1\mu_2}.$$

If a shield with infinitely thin laminations is used in a medium with permeability μ, the shielding ratio is

$$s = \frac{pe^{-\Gamma\Theta}}{1 - qe^{-2(\Gamma+1)\Theta}}, \quad p = \frac{4k}{(1+k)^2}, \quad q = \frac{(1-k)^2}{(1+k)^2},$$

$$k = \frac{\sqrt{\mu_1\mu_2}}{\mu}, \quad \Gamma = \frac{\mu_1+\mu_2}{2\sqrt{\mu_1\mu_2}} - 1.$$

8.20. *A Diffraction Problem*

In section 5 we have calculated the field scattered by a wire of radius a when the E-vector of the incident plane wave is parallel to the wire. Let us now suppose that the E-vector is perpendicular to the wire (Fig. 8.43). If the radius of the wire is small, the electric intensity impressed on the wire is $E^i_\varphi = E_0 \cos\varphi$, where E_0 is the intensity of the incident wave at the axis of the wire. The reflected wave is of the type discussed in section 17; thus

$$H^r_z = AK_1(i\beta\rho)\cos\varphi, \quad E^r_\varphi = -\eta AK_1'(i\beta\rho)\cos\varphi, \quad E^r_\rho = -\frac{A}{i\omega\epsilon\rho}K_1(i\beta\rho)\sin\varphi.$$

Since at the surface of a perfectly conducting cylinder we must have $E^i_\varphi + E^r_\varphi = 0$, the coefficient A must be

$$A = \frac{E_0}{\eta K_1'(i\beta a)} \simeq \frac{\beta^2 a^2 E_0}{\eta} = \beta^2 a^2 H_0.$$

Thus the reflected field is proportional to the square of the radius of the wire in wavelengths and is considerably smaller than in the case in which E is parallel to the wire. In the latter case the impedance to the current flow in the wire is small whereas when the E-vector is perpendicular to the wire the impedance is large.

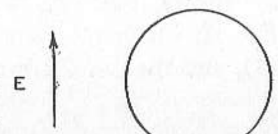


FIG. 8.43. A plane wave incident on a cylinder; E is normal to the axis.

8.21. *Dominant Waves in Wave Guides of Rectangular Cross-Section* ($TE_{1,0}$-mode*)

Consider a wave guide of rectangular cross-section (Fig. 8.44) and let the electric vector be parallel to the shorter side. That this assumption is consistent with the electromagnetic equations is shown by assuming $E_x = E_z = 0$, and substituting in (4.12–1) to obtain

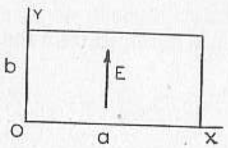


FIG. 8.44. Cross-section of a rectangular wave guide supporting a wave in which E is parallel to the yz-plane ($TE_{1,0}$-wave).

$$H_x = \frac{1}{i\omega\mu}\frac{\partial E_y}{\partial z}, \quad H_z = -\frac{1}{i\omega\mu}\frac{\partial E_y}{\partial x}, \quad H_y = 0,$$

$$\frac{\partial H_z}{\partial y} = 0, \quad \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = (g + i\omega\epsilon)E_y, \quad \frac{\partial H_x}{\partial y} = 0. \tag{21–1}$$

These equations define a realizable field and their solution is obtained as follows. First eliminating the magnetic intensities, we have

$$\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial z^2} = \sigma^2 E_y. \tag{21–2}$$

It has been shown in section 3.1 that this equation possesses solutions which are exponential functions of z and which represent, therefore, waves traveling along the tube. Assuming

$$E_y(x,z) = \hat{E}_y(x)e^{-\Gamma z} \tag{21–3}$$

and substituting in (2), we have

$$\frac{d^2\hat{E}_y}{dx^2} = -\chi^2\hat{E}_y, \quad \chi^2 = \Gamma^2 - \sigma^2, \quad \Gamma = \sqrt{\chi^2 + \sigma^2}. \tag{21–4}$$

The general solution is $\hat{E}_y = A \sin \chi x + B \cos \chi x$; but if the tube is perfectly conducting, B must be equal to zero since E_y must vanish at the face $x = 0$; furthermore we must have

$$\chi a = n\pi, \quad n = 1, 2, \cdots, \tag{21–5}$$

to ensure that E_y vanishes at the opposite face. Thus we have $\hat{E}_y = E \sin n\pi x/a$, where E is the maximum amplitude. Substituting in (3), and then in (1), we obtain

$$E_y = E \sin \frac{n\pi x}{a} e^{-\Gamma z}, \quad H_z = -\frac{n\pi E}{i\omega\mu a} \cos \frac{n\pi x}{a} e^{-\Gamma z}, \quad H_x = -\frac{E_y}{K_z}, \tag{21–6}$$

where the wave impedance in the z-direction is $K_z = i\omega\mu/\Gamma$.

* The meaning of subscripts is explained in Chapter 10.

In nondissipative media there is a sharp separation between the frequencies at which the power is carried by the wave along the tube and the frequencies at which the wave is attenuated. This may be seen from the expressions for the propagation constant and for the impedance

$$\Gamma = \sqrt{\frac{n^2\pi^2}{a^2} - \frac{4\pi^2}{\lambda^2}}, \quad K_z = \frac{i\omega\mu}{\Gamma}. \tag{21–7}$$

When the wavelength is large, Γ is real and the wave is attenuated; the wave impedance is a pure reactance and energy merely fluctuates back and forth across any given cross-section of the tube. On the other hand, when λ is small, Γ is imaginary and the amplitude of the wave is independent of z; the wave impedance is real and energy is carried by the wave along the tube.

The cut-off frequency and wavelength are obtained from the condition $\Gamma = 0$; thus

$$\lambda_c = \frac{2a}{n}, \quad f_c = \frac{v}{\lambda_c} = \frac{nv}{2a} = \frac{n}{2a\sqrt{\mu\epsilon}}. \tag{21–8}$$

The lowest cut-off frequency corresponds to $n = 1$ and the corresponding wavelength (characteristic of the medium) is twice the length of the side to which E is perpendicular

$$\lambda_c = 2a. \tag{21–9}$$

This wave is called the *dominant* wave because, having the lowest cut-off frequency, it is the only wave which will exist at a distance from the source of energy when the frequency is in the interval between the absolute cut-off defined by (9) and the next higher cut-off.

Let the ratio of the operating wavelength to the cut-off wavelength be

$$\nu = \frac{\lambda}{\lambda_c} = \frac{f_c}{f} = \frac{\omega_c}{\omega} = \frac{n\lambda}{2a}. \tag{21–10}$$

Introducing this ratio in (7) we have

$$\Gamma = i\beta_z, \quad \beta_z = \beta\sqrt{1 - \nu^2}; \quad K_z = \frac{\eta}{\sqrt{1 - \nu^2}}, \quad \lambda < \lambda_c,$$

$$\Gamma = \frac{n\pi}{a}\sqrt{1 - \frac{1}{\nu^2}}, \quad K_z = \frac{i\omega\mu a}{n\pi}\left(1 - \frac{1}{\nu^2}\right)^{-1/2}, \quad \lambda > \lambda_c. \tag{21–11}$$

The wavelength λ_z and the wave velocity v_z in the tube are

$$\lambda_z = \frac{2\pi}{\beta_z} = \frac{\lambda}{\sqrt{1 - \nu^2}}, \quad v_z = \frac{v}{\sqrt{1 - \nu^2}}. \tag{21–12}$$

At the cut-off the wavelength and the wave velocity are infinite; as the frequency increases, they approach the values characteristic of the medium.

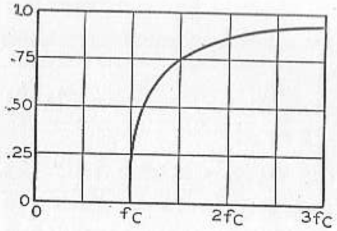


FIG. 8.45. Ratio of the characteristic velocity to the velocity in the wave guide.

In Fig. 8.45 the ratio v/v_z is shown as a function of the frequency. The wave impedance is infinite at the cut-off and approaches the intrinsic impedance as the frequency increases. Below the cut-off the wave impedance is a positive reactance and sufficiently below the cut-off it is substantially an inductance. At low frequencies the propagation constant tends to a constant value $n\pi/a$ so that the attenuation per length equal to a is $n\pi$ nepers. For dominant waves this attenuation is π nepers or about 27 db.

In dissipative dielectrics no sharp cut-off exists since power is dissipated at all frequencies wherever there is a field and this power must be carried by the wave to the place of dissipation. However, in high Q dielectrics, the conditions approximate those in nondissipative media and if we define the frequency ratio ν as above, by ignoring g, we may express the propagation constant as follows

$$\Gamma = i\beta\sqrt{(1-\nu^2) - i\frac{g}{\omega\epsilon}}. \qquad (21\text{–}13)$$

Since we have assumed that $Q = \omega\epsilon/g$ is large, we have approximately

$$\Gamma = \alpha_z + i\beta_z, \quad \alpha_z = \tfrac{1}{2}g\sqrt{\frac{\mu}{\epsilon}}(1-\nu^2)^{-1/2}, \qquad (21\text{–}14)$$

except near the cut-off. The effect of dissipation on the phase constant is of the second order.

We now turn to a more detailed discussion of the dominant wave; its field is

$$E_y = E\sin\frac{\pi x}{a}e^{-i\bar\beta z}, \quad H_x = -\frac{\bar\beta}{\eta\beta}E_y, \quad H_z = -\frac{\pi E}{i\omega\mu a}\cos\frac{\pi x}{a}e^{-i\bar\beta z}, \qquad (21\text{–}15)$$

where $\bar\beta$ is the longitudinal phase constant. The maximum voltage V is

$$V = bE_y\left(\frac{a}{2}\right) = bEe^{-i\bar\beta z}. \qquad (21\text{–}16)$$

The total longitudinal current I in the lower face of the tube (Fig. 8.44) is obtained by integrating $-H_x$; thus

$$I = -\int_0^a H_x\,dx = \frac{2a\bar\beta E}{\eta\beta\pi}e^{-i\bar\beta z} = \frac{2aE}{\pi K_z}e^{-i\bar\beta z}. \qquad (21\text{–}17)$$

The power W carried by the wave is

$$W = \frac{1}{2}\int_0^b\int_0^a \frac{E_yE_y^*}{K_z}\,dx\,dy = \frac{abE^2}{4K_z}. \qquad (21\text{–}18)$$

The "integrated" characteristic impedance may be defined in several ways. Thus we may define it either as the ratio of the maximum voltage to the total longitudinal current or in such a way that the formulae for the power in terms of the longitudinal current or in terms of the maximum voltage remain the same as the corresponding formulae for transmission lines consisting of parallel wires; thus

$$K_{V,I} = \frac{V}{I}, \quad K_{W,I} = \frac{2W}{II^*}, \quad K_{W,V} = \frac{VV^*}{2W}. \qquad (21\text{–}19)$$

These integrated characteristic impedances satisfy the following equation

$$K_{V,I} = \sqrt{K_{W,I}K_{W,V}}. \qquad (21\text{–}20)$$

Other definitions are possible; but at present these satisfy all practical requirements.

From these definitions and from the expressions for V, I, and W, we obtain the following expressions for dominant waves

$$K_{V,I} = \frac{\pi b}{2a}K_z, \quad K_{W,V} = \frac{2b}{a}K_z, \quad K_{W,I} = \frac{\pi^2 b}{8a}K_z. \qquad (21\text{–}21)$$

All these integrated impedances are proportional to the wave impedance at a point and involve the dimensions of the guide in the same way. They are proportional to b, the dimension parallel to the electric lines, and inversely proportional to a, the width of the faces which carry the longitudinal current. These expressions do not differ greatly from the approximate characteristic impedance $(b/a)\eta$ for transverse electromagnetic waves in a pair of parallel metal strips. The integrals in (17) and (18) may have obscured simple considerations by which the values of the integrated impedances may be obtained from the wave impedance K_z. The longitudinal power flow per unit area is proportional to K_z; if the transverse field were uniform throughout the cross-section of the guide all the integrated impedances would be equal to $(b/a)K_z$. Since the longitudinal current distribution is sinusoidal the total current is only $2/\pi$ times the maximum current density and the voltage/current ratio becomes $(\pi b/2a)K_z$. Similarly, for a sinusoidal distribution of the field over the cross-section, the power flow is only half what it would be in the case of uniform distribution with the same maximum voltage along electric lines; hence the impedance on the power-voltage basis is twice that for the uniform field.

As we have already seen, the impedance concept plays an important role in the theory of reflection. When the field distributions in equiphase surfaces are the same for the incident, reflected and transmitted waves, the reflection and transmission coefficients depend solely on the ratio of two wave impedances, taken in the direction of wave propagation. This is the case when the boundary conditions over the entire interface between two media are satisfied automatically as soon as the boundary conditions at any one point of the interface are satisfied. In some cases the incident wave may be resolved into components which are thus simply reflected. Examples of this type of reflection have been encountered in the theory of cylindrical and spherical shields when the sources were not axially or centrally disposed. In such cases we have to consider an infinite number of wave impedances, one for each wave component, and the reflected wave will depend on an infinite number of impedance ratios. In wave guides the exact theory of reflection is often even more complicated. Nevertheless there are instances in which the ratio of integrated impedances may be expected to give a satisfactory indication of the amount of reflection. For example, in two conductor wave guides (parallel wires or coaxial pairs) at frequencies so low that only transverse electromagnetic waves are transmitted while other waves are rapidly attenuated and contribute only negligible end effects, the integrated characteristic impedances* determine quite accurately the reflection and transmission coefficients at a junction between two lines. In Chapter 12 we shall prove that for frequencies between the lowest cut-off and the next higher, the above mentioned end effects may be represented by proper reactances either in shunt or in series with the wave guides.

If a and b are varied simultaneously and fairly slowly, it is possible to keep the characteristic impedances constant and thus eliminate reflections which would ordinarily occur when the dimensions are altered. Reflection losses may be avoided even when the change in the dimensions is abrupt provided we introduce compensating discontinuities.

If the conductivity of the tube is finite but large, the tangential component of E is very small. The above expressions for the field become first approximations to the exact expressions. The magnetic field tangential to the tube is large and is not appreciably affected by the change in the boundary condition; hence the tangential electric intensity is obtained if we multiply the magnetic intensity by $\eta_c = \mathcal{R}(1 + i)$, where $\mathcal{R}$ is the intrinsic resistance of the tube. The power absorption of the two faces parallel to the E-vector is then

$$\hat{W}_1 = \mathcal{R} \int_0^b H_z H_z^* \, dy = \frac{\mathcal{R}\pi^2 b E^2}{\eta^2 \beta^2 a^2}. \qquad (21\text{–}22)$$

* For transverse electromagnetic waves the integrated impedances $K_{V,I}$, $K_{W,V}$ and $K_{W,I}$ are equal.

The power absorbed by the faces perpendicular to E is

$$\hat{W}_2 = \mathcal{R} \int_0^a (H_x H_x^* + H_z H_z^*) \, dx = \frac{\mathcal{R} a E^2}{2\eta^2}. \qquad (21\text{–}23)$$

The power absorbed by these faces depends on the frequency only through $\mathcal{R}$ and hence is proportional to the square root of the frequency. On the other hand, $\hat{W}_1$ has the square of the frequency in the denominator and hence becomes small at sufficiently high frequencies. This is not surprising since with increasing frequency the displacement current between the faces carrying the longitudinal current will predominate and the transverse conduction current will become smaller.

In the yz-plane the transverse conduction current is

$$J_c = \frac{\pi E}{i\omega\mu a} e^{-i\beta z} = \frac{\pi V}{i\omega\mu ab}.$$

An equal current flows in the opposite face. For the transverse displacement current we have

$$J_d = i\omega\epsilon \int_0^a E_y \, dx = \frac{2}{\pi} i\omega\epsilon a E e^{-i\beta z} = \frac{2i\omega\epsilon a}{\pi b} V.$$

From these equations we obtain the shunt capacitance per unit length and the shunt inductance on the voltage current basis

$$C_{V,I} = \frac{2\epsilon a}{\pi b}, \quad \hat{L}_{V,I} = \frac{\mu ab}{2\pi}.$$

The longitudinal inductance L per unit length may be obtained from its definition

$$\frac{dV}{dz} = -i\omega L_{V,I} I, \quad L_{V,I} = -\frac{1}{i\omega I}\frac{dV}{dz},$$

and from (16) and (17); thus $L_{V,I} = \pi\mu b/2a$. These expressions provide another method of obtaining $K_{V,I}$.

The integrated distributed constants of the guide may also be defined in terms of the energies associated with currents and voltages. But from the practical point of view the primary constants of wave guides are of lesser importance than the secondary constants.

The total power dissipation per unit length of the guide is

$$\hat{W} = \hat{W}_1 + \hat{W}_2 = \frac{\mathcal{R} a E^2}{2\eta^2}\left[1 + \frac{2b}{a}\left(\frac{\lambda}{2a}\right)^2\right]. \qquad (21\text{–}24)$$

From this equation and from (18) we obtain the attenuation constant

$$\alpha = \frac{\mathcal{R}}{\eta b\sqrt{1 - \nu^2}}\left(1 + \frac{2b\nu^2}{a}\right) = \frac{\mathcal{R}}{\eta\sqrt{1 - \nu^2}}\left(\frac{1}{b} + \frac{2\nu^2}{a}\right). \qquad (21\text{–}25)$$

In the case of imperfect dielectrics the total attenuation constant is obtained by adding (14) and (25). The attenuation constant in an air-filled copper tube with $a = 20$ cms and $b = 10$ cms varies as shown in Fig. 8.46. The wavelength is expressed in centimeters and the attenuation in decibels per meter.

In order to produce waves of the type considered in this section the electric field impressed over the cross-section of the guide must conform to the

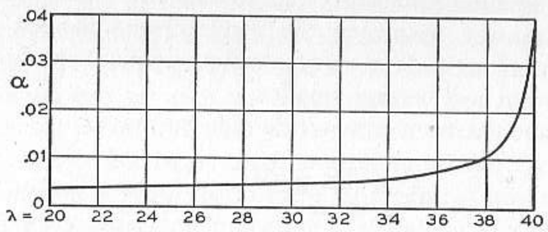


Fig. 8.46. The attenuation constant for the dominant wave ($TE_{1,0}$-wave) in a rectangular wave guide; $a = 20$ cms, $b = 10$ cms, copper walls.

electric lines and to the intensity distribution; thus the impressed intensity must be independent of the y-coordinate and be a sinusoidal function of the x-coordinate. However, from general equations of Chapter 10 it will be apparent that if $b < a$ and $a < \lambda < 2a$, then the field of any source will differ from the one here considered only in the vicinity of the source.

8.22. *Dominant Waves in Circular Wave Guides* ($TE_{1,1}$-mode)

A wave similar to the one considered in the preceding section may exist in tubes of circular cross-section. Electric lines cannot be parallel, of course, since they must end at right angles to the tube; but they show an unmistakable resemblance (Fig. 8.47) to electric lines for the corresponding transmission mode in rectangular guides. From the equations of section 21 we find that for the dominant wave the longitudinal magnetic current flows in one direction in one half of the guide and in the opposite direction in the other half. This suggests that in the present case we should seek a solution of the following type

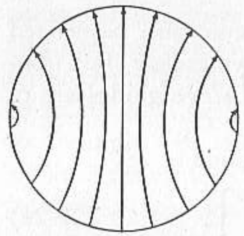


Fig. 8.47. Electric lines associated with the dominant waves in metal tubes of circular cross-section ($TE_{1,1}$-waves).

$$H_z = \hat{H}_z(\rho) \cos \varphi \, e^{-\Gamma z}. \qquad (22\text{–}1)$$

Since H_z satisfies the wave equation, we have

$$\rho^2 \frac{d^2\hat{H}_z}{d\rho^2} + \rho \frac{d\hat{H}_z}{d\rho} + (\chi^2\rho^2 - 1)\hat{H}_z = 0,$$

$$\chi^2 = \Gamma^2 + \beta^2, \quad \Gamma = \sqrt{\chi^2 - \beta^2}.$$

This is the Bessel equation of order 1 and its solution for the case when the medium inside the tube of radius $\rho = a$ is homogeneous must be $\hat{H}_z(\rho) = HJ_1(\chi\rho)$. The second solution becomes infinite at $\rho = 0$ and should be included only when the axis is excluded from the region under consideration. Thus (1) becomes

$$H_z = HJ_1(\chi\rho) \cos \varphi \, e^{-\Gamma z}. \qquad (22\text{–}2)$$

The remaining field components are now obtained from the general electromagnetic equations by setting $E_z = 0$ and using (2); thus

$$H_\rho = -\frac{\Gamma H}{\chi} J_1'(\chi\rho) \cos \varphi \, e^{-\Gamma z}, \quad E_\varphi = -K_z H_\rho,$$

$$H_\varphi = \frac{\Gamma H}{\chi^2 \rho} J_1(\chi\rho) \sin \varphi \, e^{-\Gamma z}, \quad E_\rho = K_z H_\varphi,$$

where the wave impedance in the z-direction is

$$K_z = \frac{i\omega\mu}{\Gamma} = \frac{i\omega\mu}{\sqrt{\chi^2 - \beta^2}} = \frac{\eta}{\sqrt{1 - \nu^2}}.$$

Since E_φ must vanish for $\rho = a$, we have $J_1'(\chi a) = 0$. The roots of this equation are

$$\chi a = k = 1.841, \quad 5.331, \quad 8.536, \cdots;$$

hence the longest cut-off wavelength is

$$\lambda_c = \frac{2\pi}{\chi} = \frac{2\pi a}{1.84} = 3.41a = 1.7d.$$

In Chapter 10 we shall examine all possible transmission modes in circular tubes and we shall find that $1.7d$ is the largest value for the cut-off wavelength.

The longitudinal electric current consists of conduction current only; this current is determined by H_φ and it flows in opposite directions in the

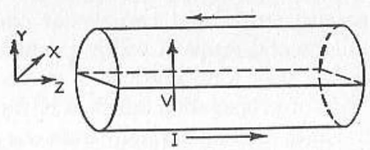


Fig. 8.48. Longitudinal conduction currents associated with a dominant wave flow in opposite directions in the two halves of the guide.

two halves of the tube (Fig. 8.48). Thus the total current I in the lower half of the tube is

$$I = -a \int_\pi^{2\pi} H_\varphi(a,\varphi) \, d\varphi = \frac{2\Gamma H a^2 J_1(k)}{k^2} e^{-\Gamma z}.$$

Calculating the maximum transverse voltage V and dividing by I, we obtain $K_{V,I}$

$$V = i\omega\mu \int_{-\pi/2}^{\pi/2} \int_0^a H_z\rho \, d\rho \, d\varphi = \frac{2i\omega\mu a^2 H e^{-\Gamma z}}{k^2} \left[\int_0^k J_0(x)\, dx - J_1(k) \right],$$

$$K_{V,I} = \frac{V}{I} = \frac{\int_0^k J_0(x)\, dx - J_1(k)}{J_1(k)} K_z.$$

The power flow above the cut-off is

$$W = \tfrac{1}{4}\pi k^{-4}(k^2 - 1) J_1^2(k) \beta_z^2 a^4 K_z H^2;$$

therefore

$$K_{W,I} = \frac{\pi}{8}(k^2 - 1) K_z.$$

For the dominant mode the above expressions become

$$K_{V,I} = 1.38 K_z = \frac{520}{\sqrt{1-\nu^2}}, \quad K_{W,I} = \frac{354}{\sqrt{1-\nu^2}}, \quad K_{W,V} = \frac{764}{\sqrt{1-\nu^2}}.$$

The expressions for the attenuation constant and the power absorbed by the tube are

$$\alpha = \frac{\mathcal{R}}{\eta a}\left(\frac{1}{k^2-1} + \nu^2\right)(1-\nu^2)^{-1/2}, \quad \hat{W} = 2\alpha W.$$

For dominant waves in air-filled guides the attenuation constant becomes

$$\alpha = \frac{\mathcal{R}}{a}[3.76(1-\nu^2)^{-1/2} - 2.65\sqrt{1-\nu^2}]10^{-3}.$$

8.23. *The Effect of Curvature on Wave Propagation*

Heretofore we have considered only straight wave guides. In this section we shall study the effect of bending the wave guide by solving four simple problems; two concerning the bending of parallel strips, and two similar ones for rectangular wave guides. Imagine two parallel metal strips of width a, separated by distance $b \ll a$, and assume that after bending they form portions of two coaxial circular cylinders. We shall consider that mode of propagation which was transverse electromagnetic before bending, with electric lines running normally from one strip to the other. After bending these lines become approximately radial as shown in Fig. 8.49. If the strips were straight we would use cartesian coordinates and call the field intensities perhaps E_x and H_y for a wave moving in the z-direction. In the present case, however, cylindrical coordinates are more suitable. The principal component of the electric intensity is E_ρ and the magnetic intensity is H_z; the wave propagation is in the direction of the φ-coordinate. We shall ignore the edge effect. It is possible to state a clear-cut mathematical problem in which the edge effect is absent. We need only consider a rectangular wave guide in which the impedance of two opposite faces is zero and the impedance of the two remaining faces infinite. The tangential electric intensity and the normal magnetic intensity vanish at a surface of zero impedance; similarly the tangential magnetic intensity and the normal electric intensity vanish at a surface of infinite impedance. Thus in the present case we assume zero impedance boundaries at $\rho = \rho_1$ and $\rho = \rho_2$ where $\rho_2 - \rho_1 = b$ and infinite impedance boundaries at $z = 0$ and $z = a$.

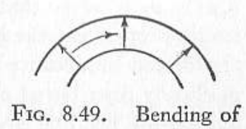


FIG. 8.49. Bending of waves.

The field under consideration has at least two components E_ρ and H_z. If we were to assume only these two components we should find from (4.12–2) that the assumption is inconsistent with the equations. We then make a less stringent assumption $\partial/\partial z = 0$ which leads us to equations (4.12–7), connecting E_ρ, H_z, and E_φ. It appears therefore that one consequence of bending as indicated in Fig. 8.49 is to introduce an electric intensity in the direction of wave propagation. For the type of wave under consideration this intensity must be small, of course, but it is required by the induction laws.

For progressive waves traveling in the negative φ-direction (clockwise in Fig. 8.49) we assume the following expression for the magnetic intensity

$$H_z(\rho,\varphi) = \hat{H}(\rho)e^{iq\varphi}.$$

Substituting in (4.12–7), we have (assuming $g = 0$) the remaining intensities and the equation for $\hat{H}(\rho)$

$$E_\rho(\rho,\varphi) = \frac{q}{\omega\epsilon\rho}\hat{H}(\rho)e^{iq\varphi}, \quad E_\varphi(\rho,\varphi) = -\frac{1}{i\omega\epsilon}\frac{d\hat{H}}{d\rho}e^{iq\varphi},$$

$$\rho^2\frac{d^2\hat{H}}{d\rho^2} + \rho\frac{d\hat{H}}{d\rho} + (\beta^2\rho^2 - q^2)\hat{H} = 0.$$

This is the Bessel equation of order q; hence its solution and the φ-component of E are

$$\hat{H}(\rho) = AJ_q(\beta\rho) + BN_q(\beta\rho),$$

$$E_\varphi(\rho,\varphi) = i\eta[AJ_q'(\beta\rho) + BN_q'(\beta\rho)]e^{iq\varphi}.$$

The boundary conditions are $E_\varphi(\rho_1,\varphi) = 0$, $E_\varphi(\rho_2,\varphi) = 0$; therefore

$$AJ_q'(\beta\rho_1) + BN_q'(\beta\rho_1) = 0,$$
$$AJ_q'(\beta\rho_2) + BN_q'(\beta\rho_2) = 0. \tag{23–1}$$

Nonvanishing values of the constants A and B will be obtained only if q is a root of the following equation

$$-\frac{B}{A} = \frac{J_q'(\beta\rho_1)}{N_q'(\beta\rho_1)} = \frac{J_q'(\beta\rho_2)}{N_q'(\beta\rho_2)}. \tag{23–2}$$

One method of solving this equation would be to plot both sides against q and obtain the points of intersection. Another method is to use the series expansions for the Bessel functions and solve for q by successive approximations, this method suggests itself when $\beta\rho_2$ and q are small. In any case it is evident that the solution of (2) is more difficult than that of the corresponding equation for a straight guide, and it is more practical to look at the problem from a physical point of view. Thus if $\rho_1 = \rho_2 = \infty$ so that the guide is straight, the wave under consideration has very simple properties; the transmission equations are the equations of a uniform line with distributed inductance and capacity given by $L = \mu b/a$, $C = \epsilon a/b$; E_ρ and H_z are uniformly distributed and E_φ is zero. Assuming therefore that in the first approximation the bending does not alter the uniformity of H_z and that the displacement current density $i\omega\epsilon E_\varphi$ may be ignored, we can calculate the new values of the distributed inductance and capacity. The inductance $\hat{L}$ and the capacitance $\hat{C}$ per radian are then

$$\hat{L} = \frac{\mu(\rho_2^2 - \rho_1^2)}{2a} = \frac{\mu bc}{a}, \quad \hat{C} = \frac{\epsilon a}{\log \frac{\rho_2}{\rho_1}}, \quad c = \frac{\rho_1 + \rho_2}{2}.$$

Hence the characteristic impedance, the phase constant, and the wave velocity become

$$K = \sqrt{\frac{\hat{L}}{\hat{C}}} = \frac{\eta b}{a}\sqrt{\frac{c}{b}\log\frac{\rho_2}{\rho_1}}, \quad q = \beta_\varphi = \omega\sqrt{\hat{L}\hat{C}} = \beta\frac{\sqrt{bc}}{\sqrt{\log\frac{\rho_2}{\rho_1}}}, \quad v_\varphi = \frac{v\sqrt{\log\frac{\rho_2}{\rho_1}}}{\sqrt{bc}}.$$

If b is small compared with c we have approximately

$$K = \frac{\eta b}{a}\left(1 + \frac{b^2}{24c^2}\right), \quad cv_\varphi = v\left(1 + \frac{b^2}{24c^2}\right), \quad \beta_\varphi = \beta c\left(1 - \frac{b^2}{24c^2}\right).$$

Thus the characteristic impedance has been increased by bending; likewise the wave velocity along the mean circle has been increased. The characteristic impedance can be reduced to the former value either by decreasing the distance between the parallel strips or by making them wider.

The foregoing extremely elementary calculations may have created an erroneous impression that the fundamental electromagnetic equations have been ignored. In reality these calculations are based on equations (4.12–7). Thus dividing the first of these equations by ρ and integrating the result as well as the remaining equation from $\rho = \rho_1$ to $\rho = \rho_2$, we obtain

$$\frac{\partial}{\partial\varphi}\int_{\rho_1}^{\rho_2}\frac{1}{\rho}H_z\,d\rho = i\omega\epsilon V, \quad H_z(\rho_2) - H_z(\rho_1) = -i\omega\epsilon\int_{\rho_1}^{\rho_2}E_\varphi\,d\rho, \tag{23–3}$$

$$\rho E_\varphi(\rho)\Big|_{\rho_1}^{\rho_2} - \frac{\partial V}{\partial\varphi} = -i\omega\mu\int_{\rho_1}^{\rho_2}\rho H_z\,d\rho, \quad \int_{\rho_1}^{\rho_2}E_\rho\,d\rho = V.$$

The first term in the third equation vanishes on account of the boundary conditions. The second equation shows that the variation in H is small because $\rho_1 - \rho_2$ is small compared with the wavelength and E_φ is small for the wave under consideration; hence $H_z(\rho) = -I/a$, where I is the conduction current in the strip of radius ρ_1.

Substituting in (3) and integrating, we have

$$\frac{dV}{d\varphi} = -\frac{i\omega\mu(\rho_2^2 - \rho_1^2)}{2a}I, \quad \frac{dI}{d\varphi} = -\frac{i\omega\epsilon a}{\log\frac{\rho_2}{\rho_1}}V. \tag{23–4}$$

Thus in following physical intuition we have not strayed very far from the fundamental equations.

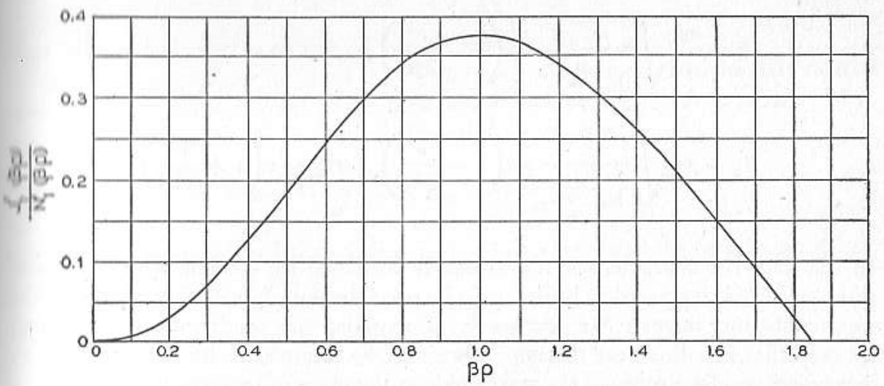


FIG. 8.50.

Returning to the exact formulation of the boundary value problem and to equation (2), we note a curious fact that this fundamental equation has been derived from the boundary conditions imposed on E_φ, that is, on a function which nearly vanishes for small curvatures. Equation (2) is not easy to handle numerically. However some values satisfying this equation can be obtained indirectly; thus if $q = 1$, we can plot J_1'/N_1' against $\beta\rho$ from the available tables (Fig. 8.50), select pairs of values $\beta\rho_1$ and $\beta\rho_2$ satisfying (2), and plot $\beta c = \frac{1}{2}\beta(\rho_1 + \rho_2)$ against ρ_1/ρ_2 (Fig. 8.51). In this special case the phase $q\varphi$ changes exactly by 2π per revolution and the field between the cylinders is periodic. Two waves of this kind traveling in opposite directions will form stationary waves and Fig. 8.51 gives the ratio of the average circumference to the natural wavelength as a function of the ratio of the radii.

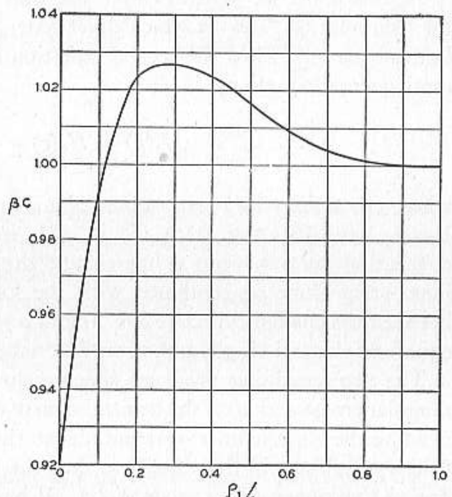


FIG. 8.51. The ratio of the mean circumference of two coaxial cylinders to the natural wavelength for the gravest oscillation mode.

The second problem concerns parallel strips bent in the plane perpendicular to the electric lines, so that the strips become two annular rings of width a, separated by distance b. Using the same coordinates we now have to consider the field components H_ρ, E_z, and H_φ. The equations to be solved are (4.12–6). The boundary value problem reduces to equation (2) and just as in the preceding case it is simpler to find an approximate solution by integrating the equations with respect to ρ and then making suitable approximations; thus letting $c = \frac{1}{2}(\rho_1 + \rho_2)$ and $\rho_2 - \rho_1 = a$ we obtain

$$K = \frac{\eta b}{a}\sqrt{\frac{a}{c \log \frac{\rho_2}{\rho_1}}} = \frac{\eta b}{a}\left(1 - \frac{a^2}{24c^2}\right),$$

$$\beta_\varphi = \beta c\sqrt{\frac{a}{c \log \frac{\rho_2}{\rho_1}}} = \beta c\left(1 - \frac{a^2}{24c^2}\right), \quad cv_\varphi = v\left(1 + \frac{a^2}{24c^2}\right).$$

In this case the characteristic impedance is decreased by bending while the mean wave velocity is increased as before. To increase the impedance to its original value, we should either increase b or decrease a. Comparing this result with the preceding, we note that in either case the impedance may be returned to its original value by decreasing the dimensions of the wave guide in the plane of bending or increasing the dimensions perpendicular to this plane.

At this point we wish to emphasize that there exist exact transmission equations for "circulating" waves which differ from (4) only in the values of the distributed line constants. Thus the second equation in (3) may be replaced by the following more general equation,

$$H_z(\rho) = H_z(c) - i\omega\epsilon \int_c^\rho E_\varphi \, d\rho, \tag{23–5}$$

where c is either the mean radius, or ρ_1, or some other radius between $\rho = \rho_1$ and $\rho = \rho_2$. Substituting from (5) in (3), we obtain exact transmission equations, connecting the transverse voltage V and the longitudinal electric current either in the inner strip alone or combined with the longitudinal displacement current flowing between the coaxial cylinders $\rho = \rho_1$ and $\rho = c$. The constants in these transmission equations depend on the unknown function, $E_\varphi(\rho)$.

The two remaining problems concern the effect of bending of metal tubes of rectangular cross-section on the transmission of dominant waves. We assume that before bending the electric lines were parallel to the shorter side b of the rectangle, that the field was uniform in this direction, that the transverse intensities varied in the direction of the longer side a as $\sin \pi s/a$. If bending takes place in the plane perpendicular to the electric intensity the appropriate equations are (4.12–6), assuming that wave propagation takes place in the φ-direction. The solution of interest to us is the one which approaches

$$E_z = E \sin \frac{\pi(\rho - \rho_1)}{a} e^{-i\beta\rho\varphi},$$

as ρ increases indefinitely. As in previous cases it is assumed that ρ_1 is the smaller radius of the bent wave guide and ρ_2 is the larger radius.

If the bending takes place in the plane determined by the electric intensity and the direction of wave propagation, then the proper equations are obtained from (4.12–2) by equating E_z to zero. In this instance the field has three components before bending, namely E_ρ, H_φ, H_z and the solution in which we are interested is

$$E_\rho = E \sin \frac{\pi z}{a} e^{-i\beta\rho\varphi}.$$

Thus bending has introduced two additional components E_φ and H_ρ.

We shall now consider in detail the former case and assume that the field distribution over the plane normal to the direction of wave propagation is unaltered by bending; hence taking (21–15) into consideration we assume

$$E_z \sim E \sin \frac{\pi(\rho - \rho_1)}{a}, \quad H_\varphi \sim \frac{\pi E}{i\omega\mu a} \cos \frac{\pi(\rho - \rho_1)}{a}.$$

If V and I are respectively the maximum transverse voltage and the total longitudinal current in the face $z = 0$ and if

$$V = bE, \quad I = \int_{\rho_1}^{\rho_2} H_\rho \, d\rho,$$

then dividing the first equation of the set (4.12–6) by ρ and integrating it and the third equation from $\rho = \rho_1$ to $\rho = \rho_2$ and from $z = 0$ to $z = b$ we obtain

$$\frac{dV}{ds} = -\frac{i\omega\mu\pi b}{2Aa} I, \quad \frac{dI}{ds} = -\left(\frac{2i\omega\epsilon aB}{\pi b} + \frac{2\pi}{i\omega\mu ab}\right) V, \tag{23–6}$$

where $s = c\varphi$, $c = (\rho_1 + \rho_2)/2$, and

$$A = \frac{\pi}{2a}\int_{\rho_1}^{\rho_2} \frac{c}{\rho} \sin \frac{\pi(\rho - \rho_1)}{a} d\rho, \quad B = \frac{\pi}{2a}\int_{\rho_1}^{\rho_2} \frac{\rho}{c} \sin \frac{\pi(\rho - \rho_1)}{a} d\rho. \tag{23–7}$$

As ρ_1 and ρ_2 increase, A and B approach unity and equations (6) approach the transmission equations for a straight wave guide, written in terms of the maximum transverse voltage V and the total longitudinal current I.

The cut-off wavelength and the characteristic impedance become

$$\lambda_c = 2a\sqrt{B}, \quad K_{V,I} = \frac{\pi\eta b}{2a\sqrt{AB}} (1 - \nu^2)^{-1/2}.$$

Evaluating (7), we have $B = 1$ and

$$A = \frac{\pi c}{2a}\left[\left(\text{Si}\frac{\pi\rho_2}{a} - \text{Si}\frac{\pi\rho_1}{a}\right)\cos\frac{\pi\rho_1}{a} - \left(\text{Ci}\frac{\pi\rho_2}{a} - \text{Ci}\frac{\pi\rho_1}{a}\right)\sin\frac{\pi\rho_1}{a}\right].$$

Thus in the first approximation the cut-off wavelength is unaltered by bending. For large values of $\pi\rho_1/a$ we have approximately

$$A = \frac{c^2}{\rho_1\rho_2}, \quad \frac{1}{A} = 1 - \frac{a^2}{4c^2}, \quad \frac{1}{\sqrt{A}} = 1 - \frac{a^2}{8c^2};$$

hence the impedance is reduced by bending. To bring the impedance back to its original value a should be made smaller or b larger.

The exact equations for the above case are

$$E_z = [AJ_q(\beta\rho) + BN_q(\beta\rho)]e^{iq\varphi},$$

$$H_\rho = -\frac{q}{\omega\mu\rho}[AJ_q(\beta\rho) + BN_q(\beta\rho)]e^{iq\varphi},$$

$$H_\varphi = \frac{1}{i\eta}[AJ_q'(\beta\rho) + BN_q(\beta\rho)]e^{iq\varphi}.$$

Since $E_z(\rho_1)$ and $E_z(\rho_2)$ should vanish we must have

$$-\frac{B}{A} = \frac{J_q(\beta\rho_1)}{N_q(\beta\rho_1)} = \frac{J_q(\beta\rho_2)}{N_q(\beta\rho_2)}.$$

Again the exact treatment of the problems depends on Bessel functions regarded as functions of their order.

CHAPTER IX

RADIATION AND DIFFRACTION

9.0. *Introduction*

If the current distribution producing an electromagnetic field is known, the power radiated in a nondissipative medium can be calculated by obtaining either the power contributed to the field by the sources or the power flowing through any closed surface surrounding the sources. The second method is based on the energy theorem (4.8–7) which states that when $g = 0$ the average power contributed to the field is equal to the real part of a certain surface integral (4.8–8). The surface (S) of integration may be chosen to suit our convenience. Since in general the distant field is much simpler than the local field, a sphere of infinite radius is a particularly suitable surface of integration. For the same reason the second method is usually simpler than the first from the computational point of view. On the other hand from the theoretical point of view the first method is more direct and fundamental than the second; and in some instances, when the local field can be obtained in a simple form, this method is also preferable from the computational point of view. Furthermore the first method makes it possible to determine the reactive power while the second method is useless for this purpose. In computing the power radiated in dissipative media the second method does not apply; in this case we must add to the surface integral the volume integral which represents the dissipated power.

9.1. *The Distant Field*

Equations (6.1–10) represent the intensities of the field produced by a given distribution of electric and magnetic currents in terms of two vector potentials A and F and two scalar potentials V and U. In nondissipative media the vector potentials are

$$A = \int \frac{e^{-i\beta\bar{r}}}{4\pi\bar{r}}\,dp_e, \quad F = \int \frac{e^{-i\beta\bar{r}}}{4\pi\bar{r}}\,dp_m, \tag{1–1}$$

where dp_e and dp_m are the moments of typical electric and magnetic current elements, $\bar{r}$ is the distance between an element at $P(\hat{x},\hat{y},\hat{z})$ or $P(\hat{r},\hat{\theta},\hat{\varphi})$ and a typical point in space $Q(x,y,z)$ or $Q(r,\theta,\varphi)$ and the integration is extended over the entire current distribution. Thus for $\bar{r}$ we have

$$\bar{r} = \sqrt{(x-\hat{x})^2 + (y-\hat{y})^2 + (z-\hat{z})^2}. \tag{1–2}$$

From the triangle OPQ joining the origin, the element, and the typical point in space, we have

$$\bar{r} = \sqrt{r^2 - 2r\hat{r}\cos\psi + \hat{r}^2}, \tag{1–3}$$

where ψ is the angle between OP and OQ, that is, between the directions $(\hat{\theta},\hat{\varphi})$ and (θ,φ); therefore

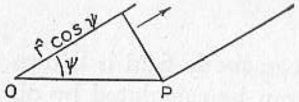


FIG. 9.1. The difference between the distances from two points to an infinitely distant point.

$$\cos\psi = \cos\theta\cos\hat{\theta} + \sin\theta\sin\hat{\theta}\cos(\varphi - \hat{\varphi}). \tag{1–4}$$

As r increases indefinitely we have

$$\bar{r} = r - \hat{r}\cos\psi + O\left(\frac{1}{r}\right); \tag{1–5}$$

that is, the difference $r - \bar{r}$ approaches the projection of the radius drawn from the origin to the element upon the radius in the typical direction (θ,φ) (Fig. 9.1). Since

$$\frac{1}{\bar{r}} = \frac{1}{r - \hat{r}\cos\psi + O\left(\frac{1}{r}\right)} = \frac{1}{r} + O\left(\frac{1}{r^2}\right),$$

the expression for A may be written as

$$A = \frac{e^{-i\beta r}}{4\pi r}\int e^{i\beta\hat{r}\cos\psi}\,dp_e + O\left(\frac{1}{r^2}\right). \tag{1–6}$$

We now define the *magnetic radiation vector* N as follows

$$N = \int e^{i\beta\hat{r}\cos\psi}\,dp_e. \tag{1–7}$$

This vector is determined solely by the current distribution and it determines the principal term in the expression for the magnetic vector potential at large distances

$$A = \frac{Ne^{-i\beta r}}{4\pi r} + O\left(\frac{1}{r^2}\right). \tag{1–8}$$

Similarly we define the *electric radiation vector* L and then express the electric vector potential at great distances as follows

$$L = \int e^{i\beta\hat{r}\cos\psi}\,dp_m, \quad F = \frac{Le^{-i\beta r}}{4\pi r} + O\left(\frac{1}{r^2}\right). \tag{1–9}$$

Equations (6.2–11) show that at great distances from a current element the radial component of the field intensity varies inversely as the square of the distance and that the principal components of the field are therefore normal to the radius. From the foregoing discussion it is evident that, for any distribution of current sources which can be enclosed within a surface of finite area, the principal field components are normal to the radius drawn from some origin in the region of the sources. These principal components can be expressed in terms of the radiation vectors, using the general formula (6.1–10). The divergence of A varies at least as $1/r$ and the θ and φ-components of the gradient contain a factor $1/r$; hence the only contributions to the principal terms come from the vector potentials. With this consideration in mind we obtain the following expressions for the intensities of the distant field

$$E_\theta = \eta H_\varphi = -\frac{i}{2\lambda r}(\eta N_\theta + L_\varphi)e^{-i\beta r},$$
$$E_\varphi = -\eta H_\theta = \frac{i}{2\lambda r}(-\eta N_\varphi + L_\theta)e^{-i\beta r}. \tag{1–10}$$

9.2. *A General Radiation Formula*

The complex flow of power (4.8–8) across a sphere of radius r is

$$\Psi = \tfrac{1}{2}r^2\int (E_\theta H_\varphi^* - E_\varphi H_\theta^*)\,d\Omega,$$

where $d\Omega$ is an elementary solid angle, $d\Omega = \sin\theta\,d\theta\,d\varphi$. Substituting from (1–10) and assuming that r becomes infinite, we find that Ψ is real and we obtain the following expression for the radiated power W

$$W = \int \Phi\,d\Omega, \quad \Phi = \Phi_{11} + 2\Phi_{12} + \Phi_{22}, \quad \Phi_{11} = \frac{\eta}{8\lambda^2}(N_\theta N_\theta^* + N_\varphi N_\varphi^*), \tag{2–1}$$

$$\Phi_{22} = \frac{1}{8\eta\lambda^2}(L_\theta L_\theta^* + L_\varphi L_\varphi^*), \quad \Phi_{12} = \frac{1}{8\lambda^2}\,\mathrm{re}(N_\theta L_\varphi^* - N_\varphi L_\theta^*).$$

The integrand Φ is called the *radiation intensity* in the direction (θ,φ) and may be interpreted as the power flow per unit solid angle in that direction. Since the power flow per unit area normal to that direction may then be expressed in the following forms

$$W_S = \frac{\Phi}{r^2} = \tfrac{1}{2}\eta HH^* = \frac{EE^*}{2\eta}, \tag{2–2}$$

the amplitudes of the electric and magnetic intensities are

$$|E| = \sqrt{2\eta W_S} = \frac{\sqrt{2\eta\Phi}}{r}, \quad |H| = \sqrt{\frac{2W_S}{\eta}} = \frac{1}{r}\sqrt{\frac{2\Phi}{\eta}}. \tag{2–3}$$

The total radiated power is seen to consist of three terms

$$W = W_{11} + 2W_{12} + W_{22}, \tag{2-4}$$

where W_{11} is the power radiated by the electric currents alone, W_{22} is the power radiated by the magnetic currents alone, and $2W_{12}$ is the mutual radiated power.

9.3. *On Calculation of Radiation Vectors*

The general formulae (1–7) and (1–9) are adapted for calculation of the cartesian components of radiation vectors. For a curvilinear electric current filament the moment of a typical current element is $I(s)\,\overline{ds}$, where $\overline{ds}$ is an elementary vector tangential to the filament and $I(s)$ is the current in the filament; hence a typical cartesian component of N is

$$N_x = \int I(\hat{x},\hat{y},\hat{z})e^{i\beta\hat{r}\cos\psi}\,d\hat{x}. \tag{3-1}$$

For a magnetic current filament we have a similar expression with the magnetic current K in place of the electric current I. The spherical components which enter the expression for Φ may now be obtained from

$$N_\theta = N_x \cos\theta\cos\varphi + N_y\cos\theta\sin\varphi - N_z\sin\theta,$$
$$N_\varphi = -N_x\sin\varphi + N_y\cos\varphi. \tag{3-2}$$

There are certain properties of the radiation vectors which simplify practical applications of the above equations. If an electric current element of moment Il is situated at the origin, the radiation vector is simply the moment of the element

$$N = Il. \tag{3-3}$$

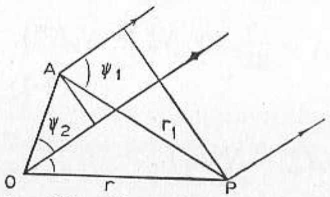


FIG. 9.2. Successive translations of the source.

If the element is at (r,θ,φ), then the radiation vector

$$N = Ile^{i\beta\hat{r}\cos\psi} \tag{3-4}$$

differs from (3) only by a "translation factor" $\exp(i\beta\hat{r}\cos\psi)$, as if we had a virtual source at the origin with a moment of the same amplitude. The wave from this virtual source is advanced in phase by $\beta\hat{r}\cos\psi$ in order to correct for the phase delay introduced while it travels through the distance $\hat{r}\cos\psi$.

More generally we have the following successive translation formulae (Fig. 9.2)

$$N_P = Il,\quad N_A = N_P e^{i\beta r_1\cos\psi_1},$$
$$N_O = N_A e^{i\beta r_2\cos\psi_2} = Ile^{i\beta r_1\cos\psi_1 + i\beta r_2\cos\psi_2}, \tag{3-5}$$

where the subscripts designate the successive positions of virtual sources. The formula for the translation of a virtual source from one point to another follows from the theorem that the projection of a compound vector on any straight line equals the sum of the projections of its components (Fig. 9.2)

$$\hat{r}\cos\psi = r_1\cos\psi_1 + r_2\cos\psi_2. \tag{3-6}$$

If we have *an array* of several identical and similarly oriented radiators, located at a number of points $P_1(r_1,\theta_1,\varphi_1)$, $P_2(r_2,\theta_2,\varphi_2)$, etc., then the radiation vector of the array is evidently

$$N = N_1 \sum A_n e^{i\beta r_n\cos\psi_n}, \tag{3-7}$$

where N_1 is the radiation vector of any element of the array and A_n is the complex strength of each source, having amplitude a_n and phase ϑ_n,

$$A_n = a_n e^{i\vartheta_n}. \tag{3-8}$$

This theorem follows immediately from the translation formula. The summation factor is called the *complex space factor* of the array and its absolute value simply the *space factor*. The space factor represents more than just the effect of the arrangement in space of individual elements of the array since it includes the effect of the amplitudes and phases of the elements. Only when all the A_n's are equal to unity does the space factor represent the effect of the spatial arrangement. Since the cartesian and hence the spherical components of N contain the space factor, the radiation intensity of the array is the product of the radiation intensity of an individual element and the square of the space factor

$$\Phi = S^2\Phi_1,\quad S = \left|\sum A_n e^{i\beta r_n\cos\psi_n}\right|. \tag{3-9}$$

9.4. *Directivity*

The *radiation pattern* or the *directive pattern* of a source or an array of sources is represented by Φ or $\sqrt{\Phi}$. An overall measure of directivity is defined as follows. A uniform radiator is taken as a standard. In the case of acoustic waves this standard is represented by a sphere pulsating radially. In the case of electromagnetic waves such a standard cannot be realized; but if figures of merit relative to this standard are known, other relative figures of merit can be immediately obtained. Secondary standards may be varied as convenience demands. Let the strengths of the given source and of the standard be adjusted for a unit power output

$$W = \int \Phi\,d\Omega = 1,\quad W_0 = \int \Phi_0\,d\Omega = 1; \tag{4-1}$$

then the *directive gain* or the *directivity* of the given source is defined as the ratio of the maximum radiation intensity to the radiation intensity of the standard. This gain is expressed also in logarithmic units; thus

$$g = \frac{\Phi_{\max}}{\Phi_0}, \quad G = 10 \log_{10} g. \tag{4–2}$$

Since $\Phi_0 = 1/4\pi$, equation (2) becomes

$$g = 4\pi\Phi_{\max}. \tag{4–3}$$

Another way of defining the directivity is to make the maximum radiation intensities of the source and of the standard equal to unity and compute the ratio of the power radiated by the standard to that radiated by the source

$$g = \frac{W_0}{W} = \frac{4\pi}{W} = \frac{4\pi}{\int \Phi \, d\Omega}. \tag{4–4}$$

The directivity of an antenna and its radiation pattern describe completely the directive properties of the antenna. In order to obtain the gain (or loss) in some direction other than that of maximum radiation, we simply multiply g by $\Phi/\Phi_{\max}$ or add $10 \log_{10} \Phi/\Phi_{\max}$ to G.

In practice we are interested in radiation properties in the presence of the ground; when the ground is assumed to be perfectly conducting the field can be obtained by the image theory. Equations (3) and (4) are still applicable if W represents the power radiated by the given sources and their images. Of course, the actual power radiated above ground is only $\frac{1}{2}W$ and the power radiated by the nondirective radiator is 2π instead of 4π.

9.5. *Directive Properties of an Electric Current Element*

Assume that an electric current element of moment Il is situated along the z-axis at the origin; then its radiation vector is

$$N_z = Il, \quad N_\theta = -Il \sin\theta,$$

and the radiation intensity becomes*

$$\Phi = \frac{\eta I^2 l^2}{8\lambda^2} \sin^2\theta = \frac{15\pi I^2 l^2}{\lambda^2} \sin^2\theta, \quad \sqrt{\Phi} = \frac{\sqrt{15\pi} Il}{\lambda} \sin\theta. \tag{5–1}$$

In the equatorial plane the element is nondirective and the radiation pattern is a circle with its center at the element. In any meridian plane the plot of $\sqrt{\Phi}$ is a circle tangential to the element.

* As usual the numerical coefficient refers to free space.

In order to obtain the directivity of the element we make the maximum radiation intensity unity and calculate the radiated power

$$\hat{\Phi} = \sin^2\theta, \quad W = \int \hat{\Phi}\, d\Omega = 2\pi \int_0^\pi \sin^3\theta \, d\theta = \frac{8\pi}{3};$$

hence the gain is

$$g = 1.5, \quad G = 1.76 \text{ db}. \tag{5–2}$$

The above power W could also have been obtained from equation (6.3–1); to make $\Phi_{\max}$ unity we let $I^2 l^2/\lambda^2 = 1/15\pi$.

Next we shall consider the effect of perfectly conducting ground (the equatorial plane) on an element at height $z = h$ above ground; the axis of the element is supposed to be perpendicular to the ground. The image of the element is positive and therefore the radiation vector and $\sqrt{\Phi}$ are*

$$N_z = Il(e^{i\beta h \cos\theta} + e^{-i\beta h \cos\theta}) = 2Il \cos(\beta h \cos\theta),$$

$$N_\theta = -2Il \cos(\beta h \cos\theta) \sin\theta,$$

$$\sqrt{\Phi} = \frac{2\sqrt{15\pi} Il}{\lambda} \cos(\beta h \cos\theta) \sin\theta.$$

The "horizontal" pattern is still a circle, but the "vertical" pattern is affected by the height. If h is sufficiently large, the radiation intensity may vanish for values of the angle θ other than zero; these angles of the *cones of silence* are obtained from

$$\beta h \cos\bar{\theta} = n\pi + \frac{\pi}{2}, \quad \cos\bar{\theta} = \frac{(2n+1)\lambda}{4h}.$$

The height must be greater than a quarter wavelength before other nulls than in the direction $\theta = 0$ make their appearance. Cones of silence are produced because in some directions the direct wave from the source is canceled by the ground reflected wave.

The power radiated by the element in the presence of ground may be obtained from the equations of section 6.4. The mutual radiation resistance of the element and its image is obtained from (6.4–24) if we assume $z_1 - z_2 = 2h$; thus

$$R_{12} = \frac{15 l^2}{h^2} \left(\frac{\sin 2\beta h}{2\beta h} - \cos 2\beta h \right),$$

and the total radiated power is

$$W = 2(W_{11} + W_{12}) = (R_{11} + R_{12})I^2 = \left(\frac{80\pi^2 l^2}{\lambda^2} + R_{12} \right) I^2.$$

* For the upper element $\theta = 0$ in equation (1–4) and for the lower element $\theta = \pi$; for both elements $r = h$.

We remind the reader that this is the power radiated by the source and its image in free space; only one-half of this power is radiated above ground. The maximum radiation intensity is unity when $I^2l^2/\lambda^2 = 1/60\pi$, and the corresponding radiated power and directivity become

$$W = \frac{4\pi}{3}\left[1 + \frac{3}{4\beta^2h^2}\left(\frac{\sin 2\beta h}{2\beta h} - \cos 2\beta h\right)\right],$$

$$g = 3\left[1 + \frac{3}{4\beta^2h^2}\left(\frac{\sin 2\beta h}{2\beta h} - \cos 2\beta h\right)\right]^{-1}.$$

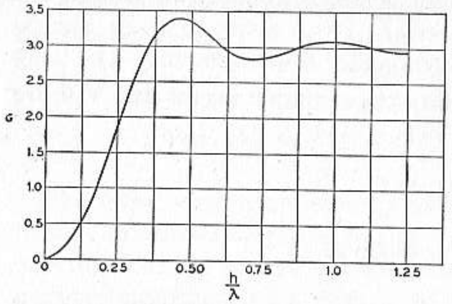


FIG. 9.3. The gain in decibels of a current element normal to a perfectly conducting ground.

If the element is just above ground, g becomes equal to 1.5 and the ground has no effect on the gain. If the element is very high above ground, then $g = 3$ and the ground adds 3 db gain. When

$$\frac{\sin 2\beta h}{2\beta h} - \cos 2\beta h = 0, \quad h = .357\lambda,$$

the added gain is also 3 db. The added gain is greater than 3 db when the mutual radiation resistance R_{12} is negative; otherwise the gain is less than 3 db. Figure 9.3 shows how the gain in decibels varies with the height in wavelengths.

9.6. *Directive Properties of a Small Electric Current Loop*

In free space the radiation pattern of a small loop carrying substantially uniform current is the same as that of a current element. In the plane of the loop the diagram is a circle with its center at the center of the loop; in any perpendicular half plane the diagram is a circle tangential to the axis of the loop. This is not surprising since the loop is equivalent to a magnetic current element. If the plane of the loop is parallel to ground, the image source is negative and the radiation intensity in the ground plane vanishes. If the loop is near ground and the current is kept constant, the radiated power is considerably reduced.

If the plane of the loop is perpendicular to ground, the image is positive. Let the center of the loop be at height $z = h$ above ground and let the axis of the loop be parallel to the x-axis. If S is the area of the loop, then the moment Kl of the equivalent magnetic current element equals $i\omega\mu SI$ and the magnetic radiation vector of the loop and its image is

$$L_x = 2i\omega\mu SI \cos(\beta h \cos\theta).$$

For the radiation intensity we obtain

$$\Phi = \frac{1}{8\eta\lambda^2}L_xL_x^* (\cos^2\theta\cos^2\varphi + \sin^2\varphi)$$
$$= \frac{240\pi^3S^2I^2}{\lambda^4}\cos^2(\beta h\cos\theta)(\cos^2\theta\cos^2\varphi + \sin^2\varphi).$$

Since

$$\cos^2\theta\cos^2\varphi + \sin^2\varphi = 1 - \sin^2\theta\cos^2\varphi,$$

Φ is maximum for all values of h when $\theta = 90°$, $\varphi = \pm 90°$; that is, the maximum intensity occurs all along the intersection of the plane of the loop and the ground plane. Φ_{max} is unity when $S^2I^2/\lambda^4 = 1/240\pi^3$.

The radiation resistance R_{11} of the loop is given by (6.10–3). The mutual radiation resistance between the loop and its image may be obtained from (6.17–5) if we compute the electromotive force induced in the loop by its image; thus

$$R_{12} = \text{re}\left[\frac{i\omega\mu\beta^2S^2}{4\pi r}\left(1 + \frac{1}{i\beta r} - \frac{1}{\beta^2r^2}\right)e^{-i\beta r}\right],$$

or

$$R_{12} = \frac{240\pi^3S^2}{\lambda^3r}\left[\left(1 - \frac{1}{\beta^2r^2}\right)\sin\beta r + \frac{\cos\beta r}{\beta r}\right], \quad r = 2h.$$

The power radiated by the loop and its image is then

$$W = (R_{11} + R_{12})I^2 = \frac{(R_{11} + R_{12})\lambda^4}{240\pi^3S^2},$$

and the directivity may be expressed as follows

$$g = 3\left[1 + \left(1 - \frac{1}{4\beta^2h^2}\right)\frac{3\sin 2\beta h}{4\beta h} + \frac{3\cos 2\beta h}{8\beta^2h^2}\right]^{-1}.$$

When the loop is just above ground, it radiates uniformly in its own plane ($\varphi = 90°$). If however $\beta h = \pi/2$, $h = \lambda/4$, then the vertical radiation from the loop is canceled by that from the image and we expect a substantial improvement in gain. In this case we have

$$g = 3\left(1 - \frac{3}{2\pi^2}\right)^{-1}, \quad G = 5.49 \text{ db}.$$

The gain of a vertical current element, a quarter wavelength above ground is only 3.62 db; hence the loop has a gain of 1.87 db over the element.

9.7. *Directive Properties of a Vertical Antenna*

In the case of a vertical antenna above a perfectly conducting ground the ground may again be replaced by an image antenna (Fig. 9.4). As in preceding sections we suppose the ground removed and consider that the antenna and its image constitute a wire in free space energized from the center. If the wire is perfectly conducting and if its radius is vanishingly small, the current distribution is given by (6.8–1). In the next section we shall study the effect of finite radius on the assumption that the current distribution remains sinusoidal and in Chapter 11 we shall determine the extent to which the actual current distribution deviates from the assumed distribution. The finite radius affects principally the input impedance and the radiation pattern in those directions in which the radiation intensity is small.

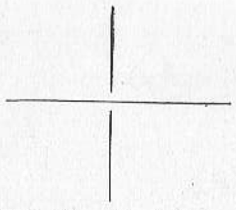


FIG. 9.4. Vertical antenna above a perfect ground and its image.

The radiation vector in the present case is

$$N_z = 2I \int_0^l \sin \beta(l - z) \cos (\beta z \cos \theta)\, dz$$

$$= \frac{2I[\cos (\beta l \cos \theta) - \cos \beta l]}{\beta \sin^2 \theta}, \qquad (7\text{–}1)$$

where l is the length of the wire above ground and $2l$ is the total length of the antenna in free space. Hence for the radiation intensity we have

$$\Phi = \frac{60I^2 \sin^2 \left[\frac{\pi l}{\lambda} (1 - \cos \theta)\right] \sin^2 \left[\frac{\pi l}{\lambda} (1 + \cos \theta)\right]}{\pi \sin^2 \theta}. \qquad (7\text{–}2)$$

In the special case when $l = \lambda/4$, $\beta l = \pi/2$ and the radiation intensity becomes

$$\Phi = \frac{15I^2 \cos^2 (\frac{1}{2}\pi \cos \theta)}{\pi \sin^2 \theta}, \quad \Phi_{max} = \frac{15I^2}{\pi}.$$

The radiated power may be obtained by integrating Φ. If a new variable $t = \cos \theta$ is introduced, the integral may be reduced to Si and Ci functions; thus we obtain

$$W = 15 (\log 2\pi + C - \text{Ci } 2\pi)I^2 = 36.56I^2.$$

This is the total radiated power in free space; one half of it is radiated above ground. The voltage between the upper terminal (Fig. 9.4) and the ground is one half of the total voltage applied at the center of the antenna in free space; thus the radiation resistance of a quarter-wave vertical antenna (of an infinitely small radius) just above a perfect ground is 36.56 ohms while the radiation resistance of the half-wave antenna in free space is 73.129 ohms.

The directivity of our antenna is

$$G = 10 \log_{10} \frac{60}{36.56} = 2.15 \text{ db}.$$

Comparing this with $G = 1.76$ db for an electric current element, we find that the gain of the half-wave antenna over the element is only 0.39 db. The principal difference between a very short antenna and a half-wave antenna is in the input impedance; the short antenna has a high reactance and a very small radiation resistance while the half wave antenna has a comparatively large radiation resistance. Figure 9.5 compares the vertical radiation patterns of the element and the half-wave antenna.

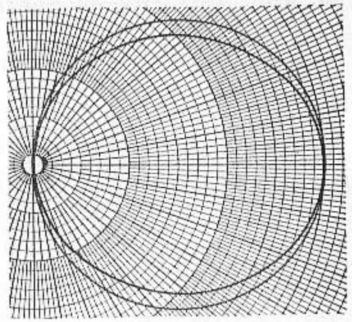


FIG. 9.5. Vertical radiation patterns of a short doublet (the outside circle) and a half-wave antenna (the inside oval).

It is evident that the length of the antenna above ground can be increased to $\lambda/2$ and the maximum radiation will still be in the ground plane. In the ground plane the radiation vector is equal to the moment of the current distribution

$$N_z = 2I \int_0^l \sin \beta(l - z)\, dz = \frac{\lambda I}{\pi} (1 - \cos \beta l);$$

this becomes maximum when $\beta l = \pi$. Further increase in l reduces the intensity in the ground plane and shifts the maximum toward the normal to the ground plane.

9.8. *The Effect of the Radius of the Wire on the Radiated Power*

In this section we shall consider the effect of the finite radius of the wire, assuming that the current distribution remains sinusoidal; a more complete discussion is deferred to Chapter 11. If we assume that the current is distributed on the surface of a wire of radius a, we may regard the wire as a circular array of sources of which each source is an elementary current filament. The plane of the array is the equatorial plane. From (7–1) we obtain the radiation vector of a typical element of the array

$$d\bar{N}_z = \frac{d\hat{\varphi}}{2\pi} N_z e^{i\beta a \sin \theta \cos (\varphi - \hat{\varphi})},$$

where $\hat{\varphi}$ defines the angular position of the elementary filament. Hence the space factor is

$$S = \frac{1}{2\pi} \int_0^{2\pi} e^{i\beta a \sin \theta \cos (\varphi - \hat{\varphi})}\, d\hat{\varphi} = J_0(\beta a \sin \theta),$$

and the radiation intensity $\bar{\Phi}$ becomes $\bar{\Phi} = S^2\Phi$, where Φ is given by (7–2). If $\beta a \ll 1$, $S \simeq 1$ and the radius has a negligible effect on the radiated power.

9.9. *Linear Arrays with Uniform Amplitude Distribution*

Consider an array of n like radiators equispaced on a straight line (Fig. 9 6). Let the amplitudes of the sources be equal and let ϑ be the phase lag as we pass from one source to the next from left to right. If Φ is the radiation intensity of each individual radiator, the radiation intensity $\hat{\Phi}$ of the array is obtained from (3–9); thus

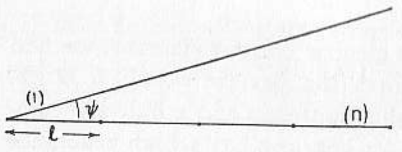


FIG. 9.6. Linear array of equispaced sources.

$$A_1 = 1, \quad A_2 = e^{-i\vartheta}, \quad A_3 = e^{-2i\vartheta}, \cdots$$

$$r_1 = 0, \quad r_2 = l\cos\psi, \quad r_3 = 2l\cos\psi, \cdots$$

and consequently

$$\hat{\Phi} = S^2\Phi, \tag{9–1}$$

where

$$S = \left| 1 + e^{i\xi} + e^{2i\xi} + \cdots + e^{i(n-1)\xi} \right|, \quad \xi = \beta l \cos\psi - \vartheta. \tag{9–2}$$

Hence the space factor of the array is

$$S = \frac{|e^{in\xi} - 1|}{|e^{i\xi} - 1|} = \left| \frac{\sin\frac{n\xi}{2}}{\sin\frac{\xi}{2}} \right|. \tag{9–3}$$

S is maximum when $\xi = 0$; then $S = n$. The direction $\hat{\psi}$ for which the space factor is maximum is determined from (2)

$$\beta l \cos\hat{\psi} = \vartheta, \quad \cos\hat{\psi} = \frac{\vartheta}{\beta l} = \frac{\vartheta\lambda}{2\pi l}.$$

Thus if the phase delay ϑ along the array is zero, $\hat{\psi} = 90°$ and the space factor is maximum in the plane normal to the line of sources. Such an array is called a *broadside array*. The direction of maximum space factor is parallel to the line of sources if $\hat{\psi} = 0$, that is, if

$$\vartheta = \beta l = \frac{2\pi l}{\lambda};$$

such an array is called an *end-on array* or an *end-fire array*. The requisite phase delay may be obtained automatically if the sources are fed from a transmission line starting with the antenna at the extreme left. For intermediate values of the phase delay ϑ the space factor is maximum in some direction making an angle between 0° and 90° with the line of sources.

Null directions for the space factor are given by

$$e^{in\xi} = 1, \quad n\xi = \pm 2k\pi,$$

where k is an integer, excluding zero. Substituting from (2), we have

$$\beta l \cos\bar{\psi} - \vartheta = \pm\frac{2k\pi}{n}.$$

For an end-fire array this becomes

$$\beta l(1 - \cos\bar{\psi}) = \frac{2k\pi}{n}, \quad 1 - \cos\bar{\psi} = \frac{k\lambda}{nl}, \quad \sin\frac{\bar{\psi}}{2} = \sqrt{\frac{k\lambda}{2nl}}.$$

If nl is large, the first few nulls are small angles and we have approximately

$$\bar{\psi} = \sqrt{\frac{2k\lambda}{nl}}, \quad \bar{\psi}_1 = \sqrt{\frac{2\lambda}{nl}}.$$

The total width of the *major radiation lobe* is

$$\Delta = 2\bar{\psi}_1 = 2\sqrt{\frac{2\lambda}{nl}}.$$

It should be noted that nl equals the total length of the array, augmented by l.

For a broadside array we have the following equation for null directions

$$\beta l \cos\bar{\psi} = \pm\frac{2k\pi}{n}, \quad \cos\bar{\psi} = \pm\frac{k\lambda}{nl}.$$

When the total length of the array is large, then for the first few nulls we have approximately

$$\sin\left(\frac{\pi}{2} - \bar{\psi}\right) = \pm\frac{k\lambda}{nl}, \quad \frac{\pi}{2} - \bar{\psi} = \pm\frac{k\lambda}{nl}.$$

The first nulls, one on each side of the direction of maximum radiation, are

$$\frac{\pi}{2} - \bar{\psi}_1 = \frac{\lambda}{nl}, \quad \frac{\pi}{2} - \bar{\psi}_{-1} = -\frac{\lambda}{nl}.$$

In this case the width of the major radiation lobe is

$$\Delta = \frac{2\lambda}{nl}.$$

For the broadside array the major lobe is narrower than for the end-fire array.

Between successive null directions there will exist secondary maxima of the space factor. These maxima coincide approximately with the directions for which the numerator in (3) is maximum; thus

$$\frac{n\xi}{2} = \pm(2k+1)\frac{\pi}{2}, \quad \xi = \pm\frac{(2k+1)\pi}{n}, \quad \beta l \cos\hat{\psi} - \vartheta = \pm\frac{(2k+1)\pi}{n},$$

where k is an integer. The amplitudes of the successive maxima, beginning with the second, vary approximately as

$$\frac{1}{\sin\frac{\xi}{2}} = \frac{1}{\sin\frac{(2k+1)\pi}{2n}}.$$

For large n and small k the maximum amplitudes, beginning with the principal maximum, are

$$n, \quad \frac{2n}{3\pi}, \quad \frac{2n}{5\pi}, \quad \frac{2n}{7\pi}, \cdots,$$

or

$$1 : \ 0.212 : \ 0.127 : \ 0.091 : \cdots.$$

The level of the second maximum is about 13.5 db below the principal maximum, independently of the actual value of n as long as n is large.

Figure 9.7 illustrates how the space factor varies with ξ for the case $n = 10$. As the angle ψ varies from 0° to 180°, which is the maximum span for ψ, ξ varies from $\beta l - \vartheta$ to $\beta l + \vartheta$; if this range is within $(-\pi, +\pi)$,

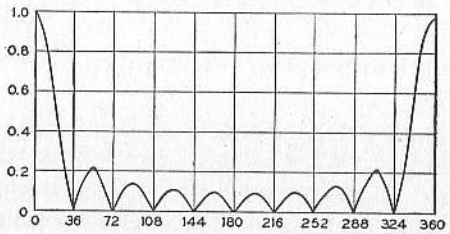


FIG. 9.7. Universal radiation pattern for linear arrays of ten sources with equal amplitudes.

then the space factor contains only one major lobe and the minor lobes diminish with increasing angle from the direction of maximum S. For a broadside array the span of ξ is $(-\beta l, \beta l)$ and so long as l does not exceed one-half wavelength, the above condition prevails. As soon as l becomes larger than $\lambda/2$, however, the secondary maxima beyond a certain point begin to grow more prominent; for $l = \lambda$, for instance, there will be just as much radiation in directions parallel to the array as at right angles to it.

For an end-fire array the absolute value* of ξ varies from 0 to $2\beta l$. In this case l should not exceed $\lambda/4$ if we wish to exclude growing secondary maxima. If $l = \lambda/2$, there is just as much radiation in the direction $\theta = 180°$ as in the direction $\theta = 0°$.

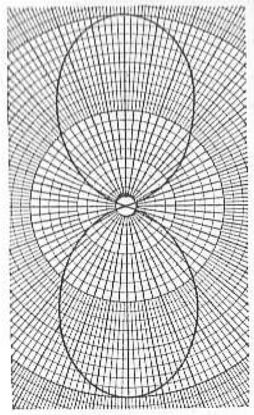


FIG. 9.8. Radiation pattern of a broadside array of two elements; $l = \frac{\lambda}{2}$.

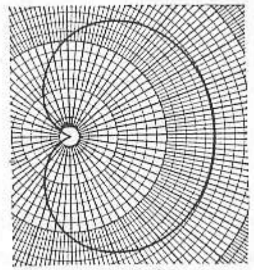


FIG. 9.9. Radiation pattern of an end-fire array of two elements; $l = \frac{\lambda}{4}$.

We shall now consider two special cases. The first is a broadside array of two elements, one-half wavelength apart. The space factor of this array is (reducing the principal maximum to unity)

$$S = \cos\frac{\xi}{2} = \cos\left(\frac{\pi}{2}\cos\psi\right).$$

S vanishes along the line joining the sources; the polar diagram is shown in Fig. 9.8. The second case is the end-fire array of two sources, one-quarter wavelength apart. In this case

$$S = \cos\frac{\xi}{2} = \cos\left[\frac{\pi}{4}(1 - \cos\theta)\right].$$

The polar diagram is heart-shaped (Fig. 9.9).

9.10. *The Gain of End-Fire Arrays of Current Elements*

First we shall consider an end-fire array of n elements, perpendicular to the line of the array, a quarter wavelength apart. Let us assume that the z-axis is the line of the array and that the elements are parallel to the

* Only the absolute value is important in the expression for S.

x-axis; then the space factor and the radiation intensity become

$$S = \frac{\left|\sin\left(\frac{n\pi}{2}\sin^2\frac{\theta}{2}\right)\right|}{\sin\left(\frac{\pi}{2}\sin^2\frac{\theta}{2}\right)};$$

$$\Phi = \frac{15\pi I^2 l^2}{\lambda^2}(\cos^2\theta\cos^2\varphi + \sin^2\varphi)S^2. \tag{10–1}$$

If we reduce the maximum radiation intensity of the array to unity, we obtain

$$\hat{\Phi} = \frac{\sin^2\left(\frac{n\pi}{2}\sin^2\frac{\theta}{2}\right)}{n^2\sin^2\left(\frac{\pi}{2}\sin^2\frac{\theta}{2}\right)}(\cos^2\theta\cos^2\varphi + \sin^2\varphi). \tag{10–2}$$

The radiated power may be obtained either by integrating $\hat{\Phi}$ or from the mutual impedances between the elements. In the case of a few elements the second method is preferable; for a large number of elements, it is easy to obtain an asymptotic expression by the first method. When n is large, the principal radiation lobe is within a small angle θ where we have approximately $\sin\theta = \theta$ and $\cos\theta = 1$; thus

$$\hat{\Phi} = \frac{64\sin^2\left(\frac{1}{8}n\pi\theta^2\right)}{n^2\pi^2\theta^4}, \quad W = \frac{128}{n^2\pi}\int_0^\pi \frac{\sin^2\left(\frac{1}{8}n\pi\theta^2\right)}{\theta^3}\,d\theta.$$

Introducing a new variable $t = \frac{1}{2}n\pi\theta^2$, we have asymptotically

$$W = \frac{8}{n}\int_0^\infty \frac{\sin^2 t}{t^2}\,dt = \frac{4}{n}\int_0^\infty \frac{1-\cos 2t}{t^2}\,dt$$

$$= \frac{4}{n}\left[-\frac{1-\cos 2t}{t}\Big|_0^\infty + 2\int_0^\infty \frac{\sin 2t}{t}\,dt\right].$$

From (3.7–18) and (3.7–19) we have $W = (8/n)\,\mathrm{Si}\,\infty = 4\pi/n$. Hence the directivity of the array is

$$g = n, \quad G = 10\log_{10} n. \tag{10–3}$$

If the separation between the elements is a, then instead of (2) we have

$$\hat{\Phi} = \frac{\sin^2\left(n\beta a\sin^2\frac{\theta}{2}\right)}{n^2\sin^2\left(\beta a\sin^2\frac{\theta}{2}\right)}(\cos^2\theta\cos^2\varphi + \sin^2\varphi).$$

Let a decrease and n increase in such a way that the total length l of the array remains constant; then

$$\hat{\Phi} = \frac{\sin^2\left(\beta l\sin^2\frac{\theta}{2}\right)}{\beta^2 l^2\sin^4\frac{\theta}{2}}(1 - \sin^2\theta\cos^2\varphi),$$

and the radiated power is

$$W = \frac{4\pi}{\beta^2 l^2}\left[\int_0^\pi \frac{\sin^2\left(\beta l\sin^2\frac{\theta}{2}\right)}{\sin^4\frac{\theta}{2}}\,d\left(\sin^2\frac{\theta}{2}\right) - 2\int_0^\pi \frac{\sin^2\left(\beta l\sin^2\frac{\theta}{2}\right)}{\sin^2\frac{\theta}{2}}\cos^2\frac{\theta}{2}\,d\left(\sin^2\frac{\theta}{2}\right)\right].$$

The first term represents the power which would be radiated by a continuous array of nondirective elements; the second term expresses the modification due to the directivity of the elements. As l becomes larger the second term becomes smaller in comparison with the first. Evaluating, we have

$$W = \frac{4\pi}{\beta l}\left(\mathrm{Si}\,2\beta l - \frac{\log 2\beta l - \mathrm{Ci}\,2\beta l + C - \cos^2\beta l}{\beta l} - \frac{\sin 2\beta l}{2\beta^2 l^2}\right).$$

As l increases, W approaches $W = 2\pi^2/\beta l = \pi\lambda/l$. Hence the directivity of the array is

$$g = \frac{4l}{\lambda}, \quad G = 10\log_{10}\frac{l}{\lambda} + 6. \tag{10–4}$$

If (3) is expressed in terms of the total length of the array in wavelengths, then we also obtain (4).

The power W_1 radiated by a long continuous array within the major lobe is $W_1 = (4\pi/\beta l)\,\mathrm{Si}\,2\pi$, hence the ratio of W_1 to the total radiated power is

$$\frac{W_1}{W} = \frac{2\,\mathrm{Si}\,2\pi}{\pi} \simeq 1 - \frac{1}{\pi^2}.$$

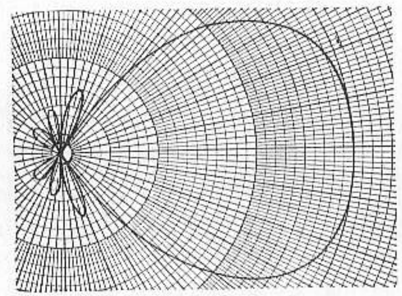


Fig. 9.10. Radiation pattern of an end-fire array of eight elements, spaced one-quarter wavelength apart.

If we had ignored all secondary lobes in computing the directive gain, our figure would have been 0.46 db too high.

The first null direction of the continuous array and the width of the major lobe are

$$\bar{\theta} = \sqrt{\frac{2\lambda}{l}}, \quad \Delta = 2\bar{\theta} = 2\sqrt{\frac{2\lambda}{l}}. \tag{10–5}$$

The total width of the major lobe is inversely proportional to the square root of the length of the array in wavelengths. Figure 9.10 represents the radiation pattern (S) of an end-fire array of eight nondirective elements, spaced one-quarter wavelength apart.

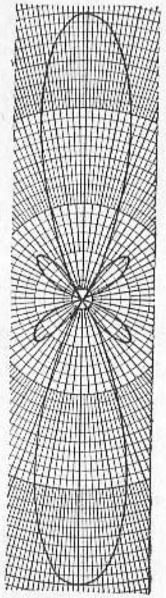


FIG. 9.11. Radiation pattern of a continuous broadside array, two wavelengths long.

9.11. *The Gain of Broadside Arrays of Current Elements*

For a continuous broadside array of electric current elements parallel to the x-axis, the radiation intensity with its maximum reduced to unity is

$$\Phi = \frac{4\sin^2(\frac{1}{2}\beta l\cos\theta)}{\beta^2 l^2\cos^2\theta}(1-\sin^2\theta\cos^2\varphi). \qquad (11\text{–}1)$$

The first two null directions, one on each side of the maximum direction are obtained from $\frac{1}{2}\beta l\cos\bar{\theta} = \pm\pi$. The total width of the major lobe is

$$\Delta = 2\left(\frac{\pi}{2}-\bar{\theta}\right) = \frac{4\pi}{\beta l} = \frac{2\lambda}{l}. \qquad (11\text{–}2)$$

The directivity of the array is calculated to be

$$g = \left(\frac{\operatorname{Si}\beta l}{\beta l} + \frac{\cos\beta l}{\beta^2 l^2} - \frac{\sin\beta l}{\beta^3 l^3}\right)^{-1}. \qquad (11\text{–}3)$$

As l increases, we have asymptotically

$$g = \frac{2\beta l}{\pi} = \frac{4l}{\lambda}, \quad G = 10\log_{10}\frac{l}{\lambda} + 6 \text{ db}. \qquad (11\text{–}4)$$

Figure 9.11 represents the radiation pattern of a continuous broadside array of nondirective elements, two wavelengths long.

9.12. *Radiation from Progressive Current Waves on a Wire*

Let us now suppose that a progressive wave is traveling in a wire of length l with velocity $\bar{v}$, in general different from the velocity v characteristic of the medium. If the wire extends from $z = 0$ to $z = l$, the radiation vector and the radiation intensity are

$$N_z = I\int_0^l e^{-i\bar{\beta}z+i\beta z\cos\theta}\,dz = \frac{2I\sin\frac{1}{2}(\bar{\beta}-\beta\cos\theta)l}{\bar{\beta}-\beta\cos\theta}\,e^{-i(\bar{\beta}-\beta\cos\theta)l/2}$$

$$\Phi = \frac{30\pi I^2[1-\cos(\bar{\beta}-\beta\cos\theta)l]}{\lambda^2(\bar{\beta}-\beta\cos\theta)^2}\sin^2\theta.$$

When integrating Φ in order to obtain W, we introduce a new variable $t = (\bar{\beta}-\beta\cos\theta)l$; thus we obtain

$$W = \frac{30\pi l}{\lambda}I^2 \times$$
$$\left[\left(1-\frac{\bar{\beta}^2}{\beta^2}\right)\int_{(\bar{\beta}-\beta)l}^{(\bar{\beta}+\beta)l}\frac{1-\cos t}{t^2}\,dt + \frac{2\bar{\beta}}{\beta^2 l}\int_{(\bar{\beta}-\beta)l}^{(\bar{\beta}+\beta)l}\frac{1-\cos t}{t}\,dt - \frac{1}{\beta^2 l^2}\int_{(\bar{\beta}-\beta)l}^{(\bar{\beta}+\beta)l}(1-\cos t)\,dt\right].$$

If the phase velocity along the wire is equal to the phase velocity characteristic of the surrounding medium, then $\bar{\beta} = \beta$ and

$$W = 30I^2\left(\log 2\beta l - \operatorname{Ci}2\beta l + C - 1 + \frac{\sin 2\beta l}{2\beta l}\right). \qquad (12\text{–}1)$$

The direction of maximum radiation is found by equating to zero the derivative of Φ with respect to θ; thus we obtain the following equation

$$\frac{\tan u}{u} = 2\left(1-\frac{u}{\beta l}\right), \quad \frac{l}{\lambda} = \frac{u}{\pi(1-\cos\hat{\theta})}.$$

The greatest maximum corresponds to the smallest root of this equation. When the wire is long, this root is substantially independent of l and it occurs for a value of u about halfway between 1.16 and 1.17. The maximum radiation intensity is unity if

$$I^2 = \frac{1.16\lambda}{30l\sin^2 1.16} = 0.046\frac{\lambda}{l}.$$

Substituting this in (1) and calculating the gain, we obtain

$$G = 10\log_{10}\frac{l}{\lambda} + 5.97 - 10\log_{10}\left(\log_{10}\frac{l}{\lambda} + 0.915\right). \qquad (12\text{–}2)$$

9.13. *Arrays with Nonuniform Amplitude Distribution*

In all arrays with uniform amplitude distribution the shape of the axial cross-section (by a plane passing through the line of radiators) of the space factors is more or less the same, regardless of the number of elements, so long as this number is not too small. If, however, the amplitudes of the individual elements are varied, the shape of the major lobe may be altered. Consider, for example, an array in which the amplitudes are proportional to the coefficients in the binomial expansion

$$1,\ n-1,\ \frac{(n-1)(n-2)}{1\cdot 2},\ \frac{(n-1)(n-2)(n-3)}{1\cdot 2\cdot 3},\ \cdots n-1,\ 1.$$

In the notation of section 9, we have

$$S = |1 + e^{i\xi}|^{n-1} = 2^{n-1}\cos^{n-1}\frac{\xi}{2}.$$

Thus the space factor of an array of n elements with binomial amplitude distribution is the $(n-1)$th power of the space factor of a pair of elements of equal amplitude. While in the case of uniform distribution the number of radiation lobes increases with the number of elements, in the "binomial" array the number of lobes remains the same as for a pair of elements.

Similarly we can start with an array of three elements with uniform amplitude distribution so that the space factor is

$$S_0 = |1 + e^{i\xi} + e^{2i\xi}|,$$

and design another array with the space factor equal to the square of S_0

$$S = S_0^2 = |1 + 2e^{i\xi} + 3e^{2i\xi} + 2e^{3i\xi} + e^{4i\xi}|. \tag{13–1}$$

This array will contain five elements with amplitudes proportional to the coefficients in equation (1). We have seen that in a uniform array the levels of the secondary radiation lobes are substantially independent of the number of elements if this number is fairly large. In an array of type (1) the secondary lobes are considerably reduced in size. On the other hand the width of the major lobe is larger than in the case of a uniform array of the same number of elements.

It is possible to assign the disposition of null directions at will, thus let

$$S = |(e^{i\xi} - e^{i\xi_1})(e^{i\xi} - e^{i\xi_2})\cdots(e^{i\xi} - e^{i\xi_{n-1}})|,$$

where $\xi_1, \xi_2, \cdots \xi_{n-1}$ correspond to the chosen null directions. When S is expanded, it becomes a polynomial of $(n-1)$th degree in $e^{i\xi}$ and it has n terms. The coefficients of this polynomial represent the relative amplitudes and phases of an array with the preassigned null directions. This method can be used to design highly directive arrays; but nonuniform amplitude distributions require larger currents in the individual elements so that ohmic losses are increased and may easily become prohibitive.

9.14. *The Solid Angle of the Major Radiation Lobe, the Form Factor, and the Gain*

The expressions for the gain of various long arrays show that the gain increases by 3 db when the length of the array is doubled. If we compute the solid angle occupied by the major lobe, we shall find that the solid angle is inversely proportional to the length of the array; and it is easy to see that g should be inversely proportional to the solid angle if the shape of the major lobe remains the same.

For an idealized radiator which sends all its energy uniformly within a given solid angle Ω, the directivity is simply

$$g = \frac{4\pi}{\Omega}, \quad G = 10\log_{10}\frac{1}{\Omega} + 10.99. \tag{14–1}$$

If the solid angle is formed by a cone of angle $\bar\theta$, then

$$\Omega = 2\pi\int_0^{\bar\theta}\sin\theta\,d\theta = 2\pi(1 - \cos\bar\theta).$$

When $\bar\theta$ is small, this becomes $\Omega = \pi\bar\theta^2$. In terms of this angle the gain is

$$g = \frac{4}{\bar\theta^2}, \quad G = 20\log_{10}\frac{1}{\bar\theta} + 6. \tag{14–2}$$

If the radiation intensity is

$$\Phi(\theta) = \left(1 - \frac{\theta^2}{\bar\theta^2}\right)^{2n}, \quad 0 \le \theta \le \bar\theta,$$
$$= 0, \qquad \bar\theta \le \theta \le \pi, \tag{14–3}$$

and if $\bar\theta$ is small, then $W = \pi\bar\theta^2/(2n+1) = \Omega/(2n+1)$. In this case the gain is

$$G = 10\log_{10}\frac{1}{\Omega} + 10.99 + 10\log_{10}(2n+1), \tag{14–4}$$

where the last term represents the effect of the form of the major lobe.

For a continuous end-fire array of length l we have from (10–5)

$$\bar\theta^2 = \frac{2\lambda}{l}, \quad \Omega = \frac{2\pi\lambda}{l},$$

and the gain (10–4) of the array becomes

$$g = \frac{8\pi}{\Omega}, \quad G = 10\log_{10}\frac{1}{\Omega} + 14.00. \tag{14–5}$$

If $n = 1$ in equation (4), the gain becomes

$$G = 10\log_{10}\frac{1}{\Omega} + 15.76.$$

This differs from (5) by less than 2 db and the array whose radiation intensity is given by (3) with $n = 1$ approximates fairly well the end-fire array. In making estimates of the ground effect or in estimating the gain of arrays made of end-fire arrays calculations may be simplified if we replace the original Φ by the simplified form (3).

Various alternative approximations for Φ might be used instead of (3). For instance, from the form of (10–2) for small values of θ we might assume

$$\Phi(\theta) = \left(1 - \frac{\theta^4}{\bar{\theta}^4}\right)^2, \quad 0 \leqq \theta \leqq \bar{\theta},$$
$$= 0, \qquad \bar{\theta} \leqq \theta \leqq \pi.$$

Integrating, we find

$$W = \frac{8\Omega}{15}, \quad G = 10 \log_{10} \frac{1}{\Omega} + 13.72.$$

This value is very close to that found in (5), but the simplicity of the approximation (3) may at times outweigh the advantage of greater accuracy obtained by using this particular form of Φ.

9.15. *Broadside Arrays of Highly Directive Elements*

Consider a pair of highly directive radiators located at $(-d/2,0,0)$ and $(d/2,0,0)$. Assume that the radiation intensity of each radiator is

$$\Phi_0 = 1, \quad 0 \leq \theta \leq \bar{\theta}, \quad 0 \leq \varphi \leq 2\pi,$$
$$= 0, \quad \bar{\theta} \leq \theta \leq 2\pi, \; 0 \leq \varphi \leq 2\pi.$$

The radiation intensity of the array is then (with the maximum reduced to unity)

$$\Phi = \tfrac{1}{2}\Phi_0 + \tfrac{1}{2}\Phi_0 \cos\left(\frac{2\pi d}{\lambda} \sin\theta \cos\varphi\right).$$

The second term represents the mutual radiation of the two sources. If d is small, the mutual radiation is almost equal to $\frac{1}{2}\Phi_0$ and there is little gain from the array over a single radiator; if d is sufficiently large, the gain is 3 db. If $\bar{\theta}$ is small, we have approximately

$$\Phi = \tfrac{1}{2}\Phi_0 + \tfrac{1}{2}\Phi_0 \cos\left(\frac{2\pi d\theta}{\lambda} \cos\varphi\right),$$

$$W = \tfrac{1}{2}\Omega + \pi \int_0^{\bar{\theta}} J_0\left(\frac{2\pi d}{\lambda}\theta\right) \theta \, d\theta = \frac{\Omega}{2}\left[1 + \frac{\lambda}{\pi d\bar{\theta}} J_1\left(\frac{2\pi d\bar{\theta}}{\lambda}\right)\right].$$

As the directivity of the individual radiators increases, they must be set further apart in order to increase the gain. If $d = \lambda/2\bar{\theta}$, then $W = 0.59\Omega$ and the gain of the array over a single radiator is 2.3 db. If a continuous end-fire array is arbitrarily approximated by a radiator of the type postulated in this section and if $\bar{\theta}$ is taken to be one-half of $\bar{\theta}$ corresponding to the end-fire array, then $d = \sqrt{\frac{1}{2}\lambda l}$. Thus two end-fire arrays, eight wavelengths long, would have to be separated by about two wavelengths in

order to obtain an additional 2.3 db gain. These figures give a general indication of the relationship between the directivity of the individual radiators, the separation between them, and the directivity of the array. Better numerical estimates may be made if Φ_0 is chosen to conform more accurately to the actual radiation intensity of the elements of the array; but the general trends will remain the same.

9.16. *Ground Effect*

The effect of a perfectly conducting ground on the directivity of a given radiator is obtained if we consider the array consisting of the radiator and its image. If the image is positive we have exactly the case analyzed in the preceding section, where $d = 2h$ and h is the height of the radiator above ground. If the image is negative, then the sign of the mutual radiation intensity becomes negative. In this case the maximum Φ is $\sin^2 \beta h$ so long as h does not exceed a quarter wavelength; for larger values of h the maximum intensity is unity. When the image is negative, the radiation intensity in the original maximum direction (parallel to ground) is canceled.

If the ground is not a perfect conductor, then its approximate effect may be obtained by assuming a reflection coefficient equal to that for uniform plane waves; thus

$$\Phi = |\, 1 + q \exp(2i\beta h \sin\theta \cos\varphi) \,|^2 \; \Phi_0,$$

where q is a function of θ and φ. This assumption is justified when h is not too small as will be shown in Chapter 10.

9.17. *Rectangular Arrays*

Consider now a broadside rectangular array of identical radiators with equal amplitudes. If a_1 is the separation between the radiators in the x-direction and b_1 the separation in the y-direction (Fig. 9.12), the space factor of the array is

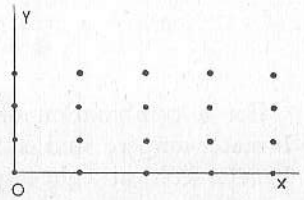


FIG. 9.12. Rectangular array of equispaced sources.

$$\bar{S} = \left| \sum_{p=0}^{m-1} \exp(ip\beta a_1 \sin\theta \cos\varphi) \sum_{q=0}^{n-1} \exp(iq\beta b_1 \sin\theta \sin\varphi) \right|. \qquad (17\text{–}1)$$

This space factor is the product of two space factors, one corresponding to a linear array parallel to the x-axis and the other to a linear array parallel to the y-axis. From (1) we obtain

$$\bar{S} = \left| \frac{\sin\left(\dfrac{m\beta a_1}{2} \sin\theta \cos\varphi\right)}{\sin\left(\dfrac{\beta a_1}{2} \sin\theta \cos\varphi\right)} \; \frac{\sin\left(\dfrac{n\beta b_1}{2} \sin\theta \sin\varphi\right)}{\sin\left(\dfrac{\beta b_1}{2} \sin\theta \sin\varphi\right)} \right|. \qquad (17\text{–}2)$$

Let a_1 and b_1 become infinitely small while m and n become infinitely large, in such a way that the dimensions $a = (m-1)a_1$ and $b = (n-1)b_1$ of the array remain constant; let the amplitude of the elementary radiator be $A\,dS (= Aa_1b_1)$ where dS is an element of area; then in the limit we have

$$\bar{S} = AS \left| \frac{\sin\left(\frac{\pi a}{\lambda}\sin\theta\cos\varphi\right)}{\frac{\pi a}{\lambda}\sin\theta\cos\varphi} \; \frac{\sin\left(\frac{\pi b}{\lambda}\sin\theta\sin\varphi\right)}{\frac{\pi b}{\lambda}\sin\theta\sin\varphi} \right|, \qquad (17\text{–}3)$$

where S is the area of the array. The space factor is maximum and equal to AS when $\theta = 0$.

9.18. *Radiation from Plane Electric and Magnetic Current Sheets*

Consider a plane electric current sheet of density J_x and a plane magnetic current sheet of density M_y. The radiation vectors are

$$N_x = \iint J_x(\hat{x},\hat{y})e^{i\beta\hat{\rho}\cos\psi}\,dS, \quad L_y = \iint M_y(\hat{x},\hat{y})e^{i\beta\hat{\rho}\cos\psi}\,dS,$$

where $\hat{\rho}$ is the distance of a typical current element from the origin and

$$\cos\psi = \sin\theta\cos(\varphi - \hat{\varphi}), \quad \hat{\rho}\cos\psi = (\hat{x}\cos\varphi + \hat{y}\sin\varphi)\sin\theta.$$

If $M_y(\hat{x},\hat{y}) = \eta J_x(\hat{x},\hat{y})$ and if the two sheets are superimposed on each other, then the radiation intensity becomes

$$\Phi = \frac{\eta}{8\lambda^2}(1+\cos\theta)^2\,|N_x|^2. \qquad (18\text{–}1)$$

For a combination of an electric current element whose moment is 1 meter-ampere and a magnetic current element of moment equal to 1 meter-volt, at right angles to each other, we have

$$\Phi_0 = \frac{\eta}{8\lambda^2}(1+\cos\theta)^2.$$

The radiation pattern of this particular source, called the *Huygens source*, is similar to the radiation pattern of an end-fire couplet of nondirective sources, a quarter wavelength apart (Fig. 9.9). Both patterns are symmetric about the axis of the source (the straight line connecting the elements in one case and the perpendicular to the plane of the elements in the other) and there is no radiation in the direction opposite to that of maximum radiation. The two patterns are not exactly the same however; for instance when $\theta = 90°$, the radiation intensity of the Huygens source is 6 db below the maximum while for the end-fire couplet it is only

3 db below the maximum. The directivity of the couplet is $g = 2$ and the directivity of the Huygens source is $g = 3$.

9.19. *Transmission through a Rectangular Aperture in an Absorbing Screen*

Assume that the xy-plane is the plane of a screen and let the boundaries of the aperture be $x = 0$, $x = a$, $y = 0$, $y = b$. Let a uniform plane wave impinge on the screen normally and let its field in the xy-plane be

$$E_x = E_0, \quad H_y = H_0, \quad E_0 = \eta H_0.$$

The effect of the screen is expressed by an added field which, by the Induction Theorem of section 6.13, may be produced by electric and magnetic current sheets of densities $J_x = -H_0$, $M_y = -E_0$.

If the screen is a perfect absorber and if the aperture is large, we may assume that the waves emitted by elementary sources directly in front of the screen are completely absorbed and that the field of the sources over the aperture is substantially unaltered. In this case the radiation intensity of the field transmitted through the aperture is approximately

$$\Phi = \frac{\eta H_0^2}{8\lambda^2}(1+\cos\theta)^2\hat{S}^2 = \frac{E_0^2}{8\eta\lambda^2}(1+\cos\theta)^2\hat{S}^2, \qquad (19\text{–}1)$$

where $\hat{S}$ is equal to $\bar{S}$ in equation (17–3) with $A = 1$.

The maximum Φ is in the direction normal to the aperture and

$$\Phi_{max} = \frac{1}{2}\eta\frac{H_0^2S^2}{\lambda^2} = \frac{E_0^2S^2}{2\eta\lambda^2}, \qquad (19\text{–}2)$$

where S is the area of the aperture. In the xz-plane, $\varphi = 0$ and (1) becomes

$$\Phi = \frac{\eta H_0^2S^2}{8\lambda^2}(1+\cos\theta)^2\frac{\sin^2\left(\frac{\pi a}{\lambda}\sin\theta\right)}{\frac{\pi^2a^2}{\lambda^2}\sin^2\theta}.$$

If θ is small, we have approximately

$$\Phi = \frac{\sin^2 u}{u^2}\Phi_{max}, \quad u = \frac{\pi a\theta}{\lambda}.$$

Figures 9.13 and 9.14 represent $\sqrt{\Phi}$ and Φ as functions of u. In the optical frequency range the pattern observed on a remote screen parallel to the aperture would consist of alternate bands of high and low illumination. The distribution of Φ in the yz-plane is similar to that in the xz-plane. In optics the field transmitted through the aperture is called the *diffracted field*.

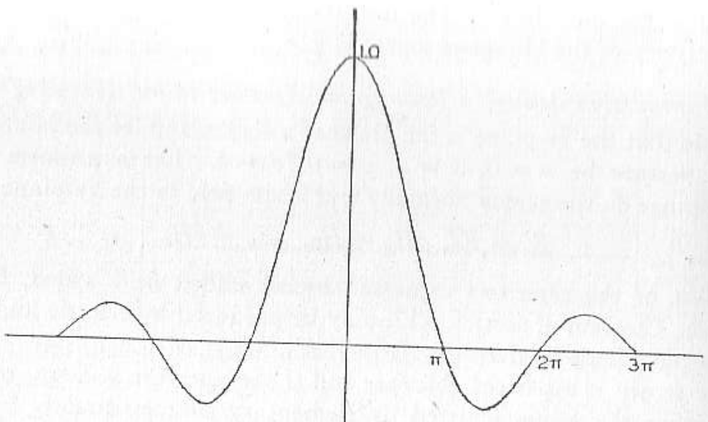


FIG. 9.13. $\sqrt{\Phi}$ vs. $u = \pi a\theta/\lambda$ for a rectangular aperture in an absorbing screen.

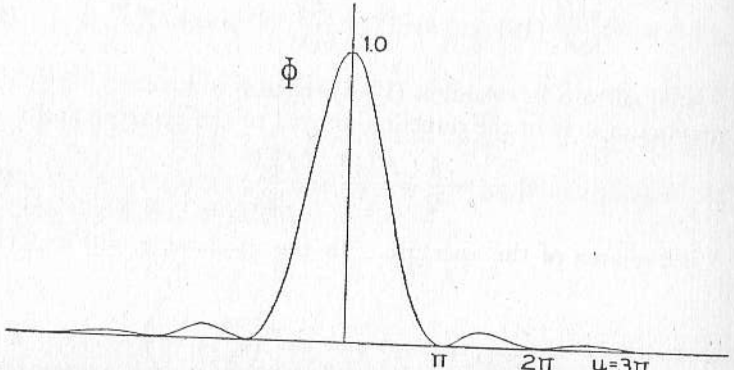


FIG. 9.14. Φ vs. $u = \pi a\theta/\lambda$ for a rectangular aperture in an absorbing screen.

9.20. *Transmission through a Circular Aperture and Reflection from a Circular Plate*

For a circular aperture of radius a equation (18–1) becomes

$$\Phi = \frac{\eta\pi^2 H_0^2}{2\lambda^2}(1+\cos\theta)^2\left|\int_0^a J_0(\beta\hat{\rho}\sin\theta)\hat{\rho}\,d\hat{\rho}\right|^2$$

$$= \frac{\eta\pi^2 a^4 H_0^2}{2\lambda^2}\left[\frac{J_1(\beta a\sin\theta)}{\beta a\sin\theta}\right]^2(1+\cos\theta)^2. \qquad (20\text{–}1)$$

The radiation intensity vanishes when

$$\beta a\sin\theta = k_n,\quad \sin\theta = \frac{k_n\lambda}{2\pi a}\quad J_1(k_n) = 0,$$

the zero root being excluded. In the case of light waves we should see on a distant screen alternate light and dark rings.

It will be remembered that equation (1) has been obtained on the assumption that the screen does not affect the field of the secondary sources over the aperture; to obtain an idea of how large the aperture should be to make this approximation satisfactory, we shall compute the ratio of the power W radiated by the electric and magnetic current sheets in free space to the power $W_0 = \frac{1}{2}\eta\pi a^2 H_0^2$ delivered to the aperture by the incident wave. For this purpose we integrate Φ over the unit sphere. In the form of a power series we have

$$\frac{W}{W_0} = \sum_{n=0}^{\infty}\frac{(-)^n(\beta a)^{2n+2}}{(2n+3)[(n+1)!]^2}; \qquad (20\text{–}2)$$

but when βa is large, the following asymptotic expansion is more convenient for numerical computations

$$\frac{W}{W_0} = 1 - \frac{1+J_1(2\beta a)}{2\beta a} + \frac{J_0(2\beta a)}{(2\beta a)^2} + \frac{J_1(2\beta a)}{(2\beta a)^3} + \frac{1\cdot 3J_2(2\beta a)}{(2\beta a)^4} + \cdots. \qquad (20\text{–}3)$$

In order to obtain this expansion we multiply (2) by βa, differentiate the result with respect to βa, compare the series with the power series for $J_0(x)$, and then integrate with respect to βa; thus we have

$$\frac{d}{d(\beta a)}\left(\beta a\frac{W}{W_0}\right) = 1 - J_0(2\beta a),\quad \frac{W}{W_0} = 1 - \frac{1}{2\beta a}\int_0^{2\beta a} J_0(t)\,dt.$$

Integrating by parts, we obtain (3).

Figure 9.15 shows how W/W_0 varies with the size of the aperture. When the diameter equals one wavelength, $\beta a = \pi$ and the power ratio is about 0.9. It may be expected that for this and larger apertures, the radiation pattern will not differ very much from the actual radiation pattern, particularly in directions in which the radiation is large. The radiation pattern near the null directions is likely to be affected to a much greater extent by the approximations.

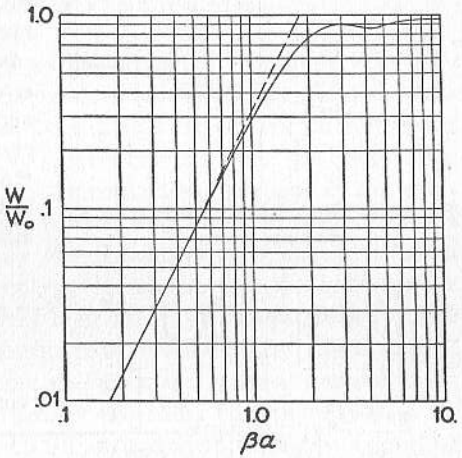


FIG. 9.15. Power ratio vs. $\beta a = 2\pi a/\lambda$; W_0 is the power incident on a circular aperture and W is the power scattered by the aperture if the edge effect is ignored.

Huygens was the first to suggest that any wavefront could be regarded as an array of secondary sources and he postulated that the field of the secondary sources should be entirely in the direction of the advancing wavefront. He offered no explanation of this property and no suggestion of the physical nature of these secondary sources. The Induction Theorem states that each secondary source is a combination of electric and magnetic current elements whose moments are respectively proportional to the magnetic and electric intensities tangential to the wavefront. The Induction Theorem is, of course, not restricted to the wavefront; but its practical usefulness is largely limited either to situations in which we have reason to believe that the field of the secondary sources is not appreciably affected by the surroundings or to situations in which the surroundings are such that the exact field of a typical secondary source may be determined.

The approximate solution of the problem of reflection from a large conducting plate is quite similar to the above. Thus on reflection from an infinite plate the tangential component H_t of the incident field is substantially doubled and the linear current density in the plate will be $2H_t$. In the case of a large plate the current density is assumed to be $2H_t$ in the first approximation. The component of the current density normal to the edge of the plate must vanish, of course, and our assumption is at its worst in the vicinity of the edge; but this "edge effect" depends on the circumference of the plate while the main effect (for *large* plates) depends on the area. If a plane wave of the type considered in the preceding section is incident normally on a circular plate, the radiation intensity of the field reflected from the plate is given by (1) provided we replace the factor $(1 + \cos\theta)^2$ by $4(\cos^2\theta\cos^2\varphi + \sin^2\varphi)$.

9.21. *Transmission through a Rectangular Aperture: Oblique Incidence*

Let us now consider the case in which the incident wave strikes the screen obliquely. Let the angle ψ between the wave normal and the normal to the screen be small (Fig. 9.16) and assume that the wave normal is parallel to the xz-plane. The components of E and H parallel to the screen are nearly equal to the total E and H and the principal difference between this case and the case of normal incidence is in the progressive lag $\beta\hat{x}\psi$ of secondary sources in the positive x-direction. Hence the radiation intensity is

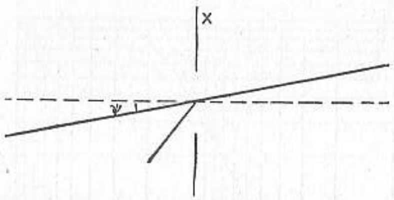


FIG. 9.16. Rectangular aperture in an infinite screen.

$$\Phi = \frac{\eta H_0}{8\lambda^2}(1+\cos\theta)^2\left|\int_0^b\int_0^a e^{i\beta\hat{x}(\sin\theta\cos\varphi-\psi)+i\beta\hat{y}\sin\theta\sin\varphi}\,d\hat{x}\,d\hat{y}\right|^2$$

$$= \frac{\eta a^2b^2H_0^2}{8\lambda^2}(1+\cos\theta)^2\frac{\sin^2\left[\frac{\pi a}{\lambda}(\sin\theta\cos\varphi-\psi)\right]\sin^2\left(\frac{\pi b}{\lambda}\sin\theta\sin\varphi\right)}{\left(\frac{\pi a}{\lambda}\right)^2(\sin\theta\cos\varphi-\psi)^2\left(\frac{\pi b}{\lambda}\right)^2\sin^2\theta\sin^2\varphi}. \quad (21\text{–}1)$$

In the plane of incidence $\varphi = 0$, $\cos\varphi = 1$; hence, for small angles, ϕ is the same function of $\theta - \psi$ for all values of ψ. Figures 9.13 and 9.14 represent the variation in $\sqrt{\Phi}$ and Φ if we take $u = \pi a(\theta - \psi)/\lambda$. The radiation pattern in the yz-plane is independent of ψ.

9.22. *Radiation from an Open End of a Rectangular Wave Guide*

For the dominant wave in a rectangular wave guide open in the plane $z = 0$ we have

$$E_y = E_0\sin\frac{\pi\hat{x}}{a},\quad H_x = -\frac{k}{\eta}E_y,\quad k = \sqrt{1-\nu^2},$$

if we assume that in the first approximation the field is not altered by the sudden discontinuity. On this basis we calculate the radiation vectors and the radiation intensity

$$L_x = E_0\int_0^b\int_0^a \sin\frac{\pi\hat{x}}{a}e^{i\beta(\hat{x}\cos\varphi+\hat{y}\sin\varphi)\sin\theta}\,d\hat{x}\,d\hat{y},\quad N_y = -\frac{k}{\eta}L_x,$$

$$|L_x| = \frac{2a\lambda E_0\sin\left(\frac{\pi b}{\lambda}\sin\theta\sin\varphi\right)\cos\left(\frac{\pi a}{\lambda}\sin\theta\cos\varphi\right)}{\sin\theta\sin\varphi\,(\pi^2-\beta^2a^2\sin^2\theta\cos^2\varphi)},$$

$$\Phi = \frac{a^2E_0^2\sin^2\left(\frac{\pi b}{\lambda}\sin\theta\sin\varphi\right)\cos^2\left(\frac{\pi a}{\lambda}\sin\theta\cos\varphi\right)}{2\eta\sin^2\theta\sin^2\varphi(\pi^2-\beta^2a^2\sin^2\theta\cos^2\varphi)^2}F(\theta,k),$$

$$F(\theta,k) = (k+\cos\theta)^2 + \nu^2\sin^2\theta\sin^2\varphi.$$

Ordinarily there exists backward radiation since $F(\pi,k) = (1-k)^2$. At the cut-off, $k = 0$, and the radiation intensities in the forward and backward directions are equal. What happens is that the electric current sheet over the aperture vanishes and the magnetic current sheet by itself has a radiation pattern which is symmetric with respect to the plane of the aperture. In this case, however, our assumption that the field over the aperture of the wave guide is unaltered by the discontinuity is least justifiable. This assumption implies that no energy is transferred along the guide which would be the case in an infinitely long guide but cannot be true when power is radiated. If we compute the power radiated by the magnetic current sheet alone, we can determine the magnetic intensity which must exist over the aperture in order to deliver the radiated power. Then we can obtain a second approximation to the radiation intensity by including the radiation from the electric current sheet defined by this magnetic intensity. This inclusion will increase the radiation intensity in the forward direction and decrease it in the backward direction.

When the operating frequency is high above the cut-off, ν is small and k is nearly unity; then $F(\theta,k) \simeq (1+\cos\theta)^2$ and the backward radiation intensity is approximately zero. The radiation intensity in the forward direction is

$$\Phi_{max} = \frac{2a^2b^2E_0^2}{\pi^2\eta\lambda^2}.$$

If we also assume that there is no reflection of power in the wave guide at the aperture, then we have the radiated power W from (8.21–18); thus $W = abE_0^2/4\eta$. If the electric intensity is adjusted for unit power output, then $\Phi_{max} = 8ab/\pi^2\lambda^2$, and the directivity of the wave guide as a radiator is

$$g = \frac{32ab}{\pi\lambda^2}, \quad G = 10\log_{10}\frac{ab}{\lambda^2} + 10.08. \tag{22–1}$$

If the field were uniform over the aperture of the wave guide, then we should have

$$\Phi_{max} = \frac{ab}{\lambda^2},$$

for unit power output and the gain would be

$$g = \frac{4\pi ab}{\lambda^2}, \quad G = 10\log_{10}\frac{ab}{\lambda^2} + 10.99. \tag{22–2}$$

Hence the *effective area* of the aperture is $8/\pi^2$ times the actual area or about four-fifths of the actual area.

9.23. *Electric Horns*

Ordinarily the cross-section of a wave guide supporting a dominant wave is comparatively small; the larger dimension will be, perhaps, greater than $\lambda/2$ and less than λ; the smaller dimension may be about one-half the larger dimension. If an open end of such a wave guide is used to radiate energy, the radiation pattern will be comparatively broad. In order to increase the directivity the wave guide is flared out into an "electric horn" (Fig. 9.17); it may be flared out either in the direction of electric lines of force or in the direction of magnetic lines of force or in both directions.

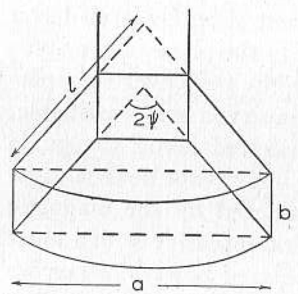


FIG. 9.17. Electric horn, flared out in one plane.

Let l as shown in Fig. 9.17 be defined as the length of the horn and let 2ψ be the angle of the horn. We shall assume that ψ is so small that the ratio of the area of the aperture to the area of the wavefront at the aperture is nearly unity. Since

$$\frac{\sin\psi}{\psi} \simeq 1 - \tfrac{1}{6}\psi^2 \simeq 1 - \frac{a^2}{24l^2},$$

the ratio is nearly unity even for a fairly large angle ψ. We shall assume that the field distribution over the wavefront is that which would exist if the horn were continued. The approximate gain can then be obtained by the method used in the preceding section.

First let us consider a sectorial horn flared out in the magnetic plane (the xz-plane). The radiation intensity in the forward direction (the positive z-axis) is

$$\Phi(0) = \frac{\eta H_0^2}{2\lambda^2}\left|\int_{-b/2}^{b/2}\int_{-a/2}^{a/2}\cos\frac{\pi\hat{x}}{a}\,e^{i\beta\hat{z}}\,d\hat{x}\,d\hat{y}\right|^2, \tag{23–1}$$

where $(\hat{x},\hat{y},\hat{z})$ is a typical point in the wavefront and hence $\hat{z}$ is the distance of a point in the wavefront from the plane of the aperture (Fig. 9.18). This distance is approximately

$$\hat{z} = l\left(\cos\frac{\hat{x}}{l} - \cos\frac{a}{2l}\right) = \frac{a^2}{8l} - \frac{\hat{x}^2}{2l}. \tag{23–2}$$

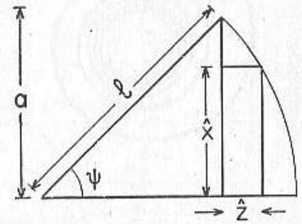


FIG 9.18. Distances connected with the horn in Fig. 9.17.

If the horn flares out in the electric plane we may express the radiation intensity in the forward direction in the following form

$$\Phi(0) = \frac{\eta H_0^2}{2\lambda^2}\left|\int_{-b/2}^{b/2}\int_{-a/2}^{a/2}\cos\frac{\pi\hat{y}}{b}\,e^{i\beta\hat{z}}\,d\hat{x}\,d\hat{y}\right|^2, \tag{23–3}$$

where $\hat{z}$ has the value defined by (2). In both cases the radiated power is approximately

$$W = \frac{abE_0^2}{4\eta} = \tfrac{1}{4}\eta abH_0^2, \tag{23–4}$$

which is equal to unity when

$$\eta abH_0^2 = 4. \tag{23–5}$$

Expression (3) is simpler to evaluate than (1). Substituting from (2) in (3), we obtain

$$\Phi(0) = \frac{\eta H_0^2}{2\lambda^2}\left|\int_{-b/2}^{b/2}\cos\frac{\pi\hat{y}}{b}\,d\hat{y}\right|^2\left|\int_{-a/2}^{a/2}\exp\left(-\frac{i\pi\hat{x}^2}{\lambda l}\right)d\hat{x}\right|^2.$$

The second factor may be expressed in terms of Fresnel integrals as defined by equations (3.7–32). Thus we find

$$\Phi(0) = \frac{4\eta b^2lH_0^2}{\pi^2\lambda}\left[C^2\left(\frac{a}{\sqrt{2\lambda l}}\right) + S^2\left(\frac{a}{\sqrt{2\lambda l}}\right)\right].$$

Substituting for H_0^2 from (5) and using (4–3), we have

$$g = \frac{64bl}{\pi\lambda a}\left[C^2\left(\frac{a}{\sqrt{2\lambda l}}\right) + S^2\left(\frac{a}{\sqrt{2\lambda l}}\right)\right]. \tag{23–6}$$

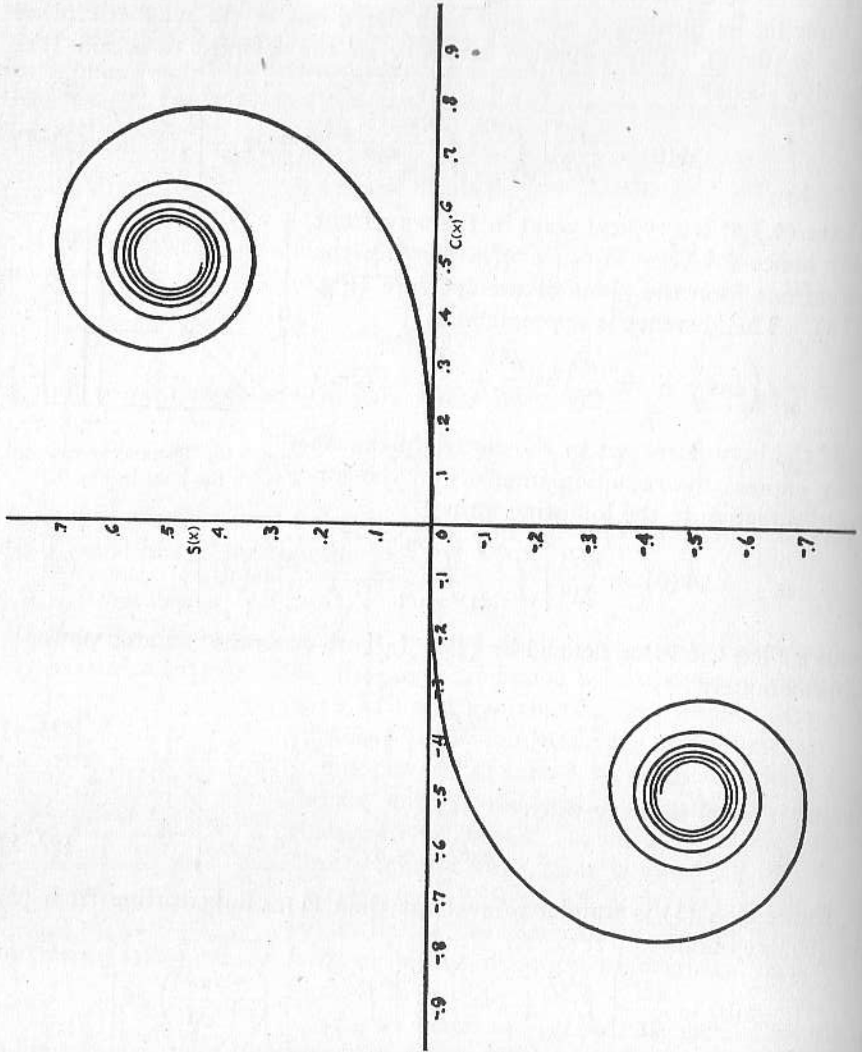


FIG. 9.19. Cornu's spiral.

Since $2l\psi = a$, $l = a/2\psi$, this may be written

$$g = \frac{32b}{\pi\lambda\psi}\left[C^2\left(\sqrt{\frac{\psi a}{\lambda}}\right) + S^2\left(\sqrt{\frac{\psi a}{\lambda}}\right)\right]. \tag{23–7}$$

The plot of $S(x)$ against $C(x)$ is called Cornu's spiral (Fig. 9.19). The distance of a point on the spiral from the origin is $\sqrt{C^2 + S^2}$ and the distance between any two points is $\sqrt{(C_2 - C_1)^2 + (S_2 - S_1)^2}$. Since $C(x)$ and $S(x)$ are odd functions, the sums $C(x_1) + C(x_2)$ and $S(x_1) + S(x_2)$ can always be regarded as differences $C(x_1) - C(-x_2)$ and $S(x_1) - S(-x_2)$. The length s of the spiral between the origin and a point corresponding to $x = x$ is

$$s = \int_0^x \sqrt{[dC(x)]^2 + [dS(x)]^2} = \int_0^x dx = x.$$

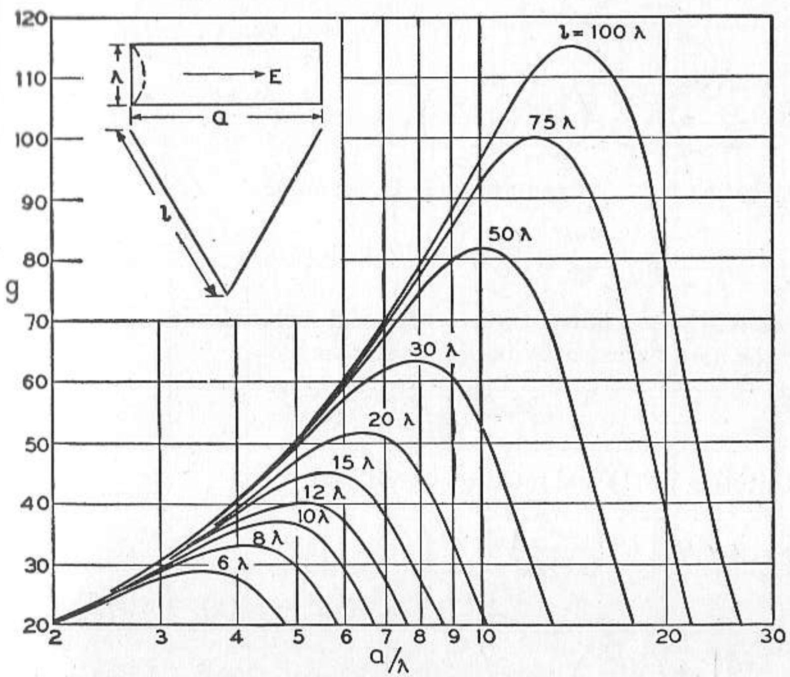


FIG. 9.20. Directive gain of a horn flared out in the electric plane.

Thus we have a simple geometric interpretation for the independent variable.

Cornu's spiral helps to interpret formulae involving Fresnel integrals. Thus it is evident at once that if ψ is kept constant and a is increased, a point is reached at which g of (7) is maximum; for larger values of a the gain in the forward direction will never be as great. What happens is that as the area of the aperture increases, some of the secondary sources on the wavefront at the aperture are out of phase with others and interfere destructively in the forward direction.

Figure 9.20 shows how g varies with the size of the aperture and the length of the horn when $b = \lambda$.

We now turn our attention to the case in which the horn flares out in the magnetic plane. Integrating (1) with respect to $\hat{y}$ and expressing the cosine in terms of exponentials we have

$$\Phi(0) = \frac{\eta b^2 H_0^2}{2\lambda^2} \left| \int_0^{a/2} \left[\exp\left(\frac{i\pi\hat{x}}{a}\right) + \exp\left(-\frac{i\pi\hat{x}}{a}\right) \right] \exp\left(-\frac{i\pi\hat{x}^2}{\lambda l}\right) d\hat{x} \right|^2 .$$

Introducing new variables we reduce this to Fresnel integrals and obtain

$$\Phi(0) = \frac{\eta b^2 l H_0^2}{4\lambda} \{[C(u) - C(v)]^2 + [S(u) - S(v)]^2\},$$

$$u = \frac{1}{\sqrt{2}}\left(\frac{\sqrt{\lambda l}}{a} + \frac{a}{\sqrt{\lambda l}}\right), \quad v = \frac{1}{\sqrt{2}}\left(\frac{\sqrt{\lambda l}}{a} - \frac{a}{\sqrt{\lambda l}}\right).$$

Substituting from (5) and using (4–3), we obtain

$$g = \frac{4\pi b l}{\lambda a} \{[C(u) - C(v)]^2 + [S(u) - S(v)]^2\}. \qquad (23\text{–}8)$$

Figure 9.21 shows how g varies with a and l when $b = \lambda$.

If the horn flares out in both planes, then

$$\hat{z} = \frac{a^2 + b^2}{8l} - \frac{\hat{x}^2 + \hat{y}^2}{2l}.$$

Substituting in (1) and integrating, we have

$$\Phi = \tfrac{1}{2}\eta l^2 H_0^2 \left[C^2\left(\frac{b}{\sqrt{2\lambda l}}\right) + S^2\left(\frac{b}{\sqrt{2\lambda l}}\right) \right] \times \{[C(u) - C(v)]^2 + [S(u) - S(v)]^2\}, \qquad (23\text{–}9)$$

$$g = \frac{8\pi l^2}{ab} \left[C^2\left(\frac{b}{\sqrt{2\lambda l}}\right) + S^2\left(\frac{b}{\sqrt{2\lambda l}}\right) \right] \{[C(u) - C(v)]^2 + [S(u) - S(v)]^2\}.$$

The directivity of such a horn can be obtained from Figs. 9.20 and 9.21 if we multiply the g's corresponding to the two sides of the aperture and divide the result by $32/\pi = 10.02$. For example let $l = 100$, then if the side normal to E is 5λ, from Fig. 9.21 we find $g_1 = 51$; if the side parallel to E is 4λ, then from Fig. 9.20, we find $g_2 = 41$. The directivity of the horn is then $g = 0.1 g_1 g_2 = 209$. Of course, when l is large and a and b small, it does not make much difference whether the horn flares out one way or the other; the curves in Figs. 9.20 and 9.21 begin to differ as they approach the maximum points.

The directivity can be expressed in the following form

$$g = 4\pi \frac{S_{\text{eff}}}{\lambda^2}, \qquad (23\text{–}10)$$

where S_{eff} is the effective area of the aperture. When l is large and a and b are small, we have $S_{\text{eff}} = (8/\pi^2)S$, where S is the actual area. For

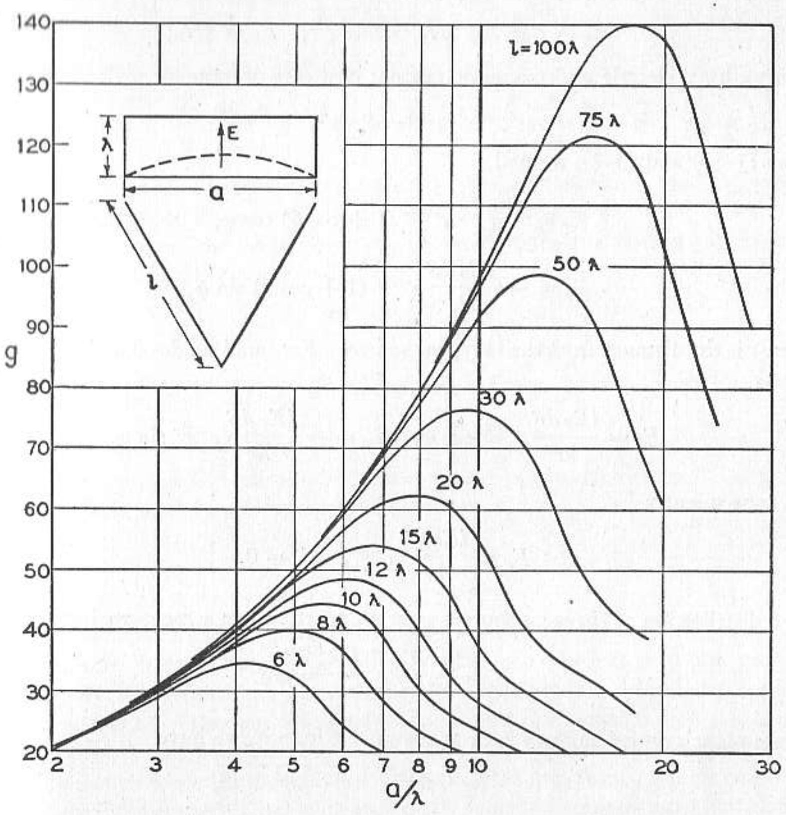


FIG. 9.21. Directive gain of a horn flared out in the magnetic plane.

optimum horns the effective area is about 65 per cent or 63 per cent, depending on whether the horn flares out only in the electric plane or only in the magnetic plane.

9.24. *Fresnel Diffraction*

In sections 19 and 20 we have studied the field transmitted through an aperture in an absorbing screen at distances so great that all the waves "emitted" by the secondary sources over the aperture arrive in phase at a point on the normal to the screen. In optics this is the case of *Fraunhofer diffraction*. In the case of *Fresnel diffraction* we are concerned with transmitted fields at distances large compared with the wavelength and the dimensions of the aperture and yet small enough to require consideration of the effect of the phase differences between secondary wavelets even along the normal to the screen.*

First let us consider the distant field of an elementary Huygens source. Let this source correspond to an element of area dS (in the xy-plane) of a plane wave traveling in the positive z-direction

$$E_x = E_0 e^{-i\beta z}, \quad H_y = H_0 e^{-i\beta z}, \quad E_0 = \eta H_0;$$

then we have electric and magnetic current elements of moment

$$p_{e,x} = -H_0\, dS, \quad p_{m,y} = -E_0\, dS.$$

From (1–10) and (3–2), we find

$$E_\theta = \frac{iE_0\, dS}{2\lambda r} e^{-i\beta r} (1 + \cos\theta) \cos\varphi,$$

$$E_\varphi = -\frac{iE_0\, dS}{2\lambda r} e^{-i\beta r} (1 + \cos\theta) \sin\varphi,$$

where r is the distance from the Huygens source. For small angles θ we have approximately

$$E_\theta = \frac{iE_0\, dS}{\lambda r} e^{-i\beta r} \cos\varphi, \quad E_\varphi = -\frac{iE_0\, dS}{\lambda r} e^{-i\beta r} \sin\varphi,$$

and consequently

$$E_x = \frac{iE_0\, dS}{\lambda r} e^{-i\beta r}, \quad E_y = 0.$$

For a distribution of Huygens sources over an aperture in a screen, we have therefore

$$E_x = \frac{i}{\lambda} \iint \frac{E_0 e^{-i\beta r}}{r}\, dS. \qquad (24\text{–}1)$$

In a plane parallel to the screen at distance z from it, we have

$$r = \sqrt{(x - \hat{x})^2 + (y - \hat{y})^2 + z^2},$$

where $(\hat{x},\hat{y},0)$ is a typical point in the plane of the aperture. When z is large compared with the dimensions of the aperture this is approximately

$$r = z + \frac{(x - \hat{x})^2 + (y - \hat{y})^2}{2z}.$$

If we now assume that z is in the range in which the amplitude factor $1/r$ is substantially constant but the phase factor $e^{-i\beta r}$ is not, then equation (1) becomes

$$E_x = \frac{ie^{-i\beta z}}{\lambda z} \iint E_0 \exp\left[-i\beta \frac{(x - \hat{x})^2 + (y - \hat{y})^2}{2z}\right] dS. \qquad (24\text{–}2)$$

* Assuming that the incident wave is normal to the screen.

At normal incidence E_0 is constant over the aperture; hence for a rectangular aperture bounded by $x = -a/2$, $x = a/2$, $y = -b/2$, $y = b/2$, we have

$$E_x = \tfrac{1}{2} iE_0 e^{-i\beta z}\left[C\left(\frac{a + 2x}{\sqrt{2\lambda z}}\right) + C\left(\frac{a - 2x}{\sqrt{2\lambda z}}\right) - iS\left(\frac{a + 2x}{\sqrt{2\lambda z}}\right) - iS\left(\frac{a - 2x}{\sqrt{2\lambda z}}\right)\right] \times$$
$$\left[C\left(\frac{b + 2y}{\sqrt{2\lambda z}}\right) + C\left(\frac{b - 2y}{\sqrt{2\lambda z}}\right) - iS\left(\frac{b + 2y}{\sqrt{2\lambda z}}\right) - iS\left(\frac{b - 2y}{\sqrt{2\lambda z}}\right)\right].$$

On the z-axis the square of the amplitude is

$$|E_x|^2 = 4E_0^2\left[C^2\left(\frac{a}{\sqrt{2\lambda z}}\right) + S^2\left(\frac{a}{\sqrt{2\lambda z}}\right)\right]\left[C^2\left(\frac{b}{\sqrt{2\lambda z}}\right) + S^2\left(\frac{b}{\sqrt{2\lambda z}}\right)\right].$$

As the size of the aperture increases the electric intensity increases until it reaches the first maximum (see Cornu's spiral).

It is easier to follow the variation in the intensity when the aperture is circular. In this case we have, when $x = y = 0$,

$$E_x = 2iE_0 \sin\frac{\pi a^2}{2\lambda z} \exp\left[-i\beta\left(z + \frac{a^2}{4z}\right)\right].$$

The maximum amplitude of E_x is $2E_0$ and it occurs when

$$a = \sqrt{n\lambda z}, \quad n = 1, 3, 5, \cdots.$$

The minimum value of E_x is zero and it occurs when n is an even integer. The above points correspond to the radii for which the difference

$$\sqrt{z^2 + a^2} - z = \frac{a^2}{2z}$$

between the distances from the point $(0,0,z)$ to the edge of the aperture and to its center equals an integral number of half wavelengths. The plane of the aperture is divided into *Fresnel zones*. The neighboring zones produce equal and opposite intensities at $(0,0,z)$. The first zone produces an intensity which is twice the intensity which would have existed at the point if the screen were removed. The aperture has a focusing effect on the field at distances of the order a^2/λ, provided a is large compared with λ. The intensity at $(0,0,z)$ is increased still more if alternate zones are blocked out; but the successive increments will eventually become smaller as the distance between $(0,0,z)$ and the zones increases.

The case of a rectangular aperture can be discussed qualitatively by the use of Cornu's spiral. Consider, for example, the variation in the electric intensity along a line parallel to the x-axis. In this connection we need to fix our attention only on the factor

$$\sqrt{\left[C\left(\frac{2x + a}{\sqrt{2\lambda z}}\right) - C\left(\frac{2x - a}{\sqrt{2\lambda z}}\right)\right]^2 + \left[S\left(\frac{2x + a}{\sqrt{2\lambda z}}\right) - S\left(\frac{2x - a}{\sqrt{2\lambda z}}\right)\right]^2}.$$

On the Cornu spiral this factor is represented by the chord joining two points separated by distance $2a/\sqrt{2\lambda z}$ measured along the curve; the midpoint of the arc is at distance

$2x/\sqrt{2\lambda z}$ from the origin (along the curve). First let the position of the receiver be fixed; then the midpoint of the arc is fixed. As a increases, the ends of the arc move outward and, up to a certain width of the aperture, the amplitude of the field at the receiver increases. Beyond this point the arc begins to wind itself on the coil of the spiral and the amplitude starts to decrease and finally reaches a minimum. From there on the intensity passes through successive maxima and minima but the fluctuations gradually diminish. On the other hand, if the size of the aperture and the distance between the receiver and the screen are fixed, then the arc is of fixed length while its midpoint will move as the receiver is moved parallel to the x-axis. The spiral is least curved at the origin; hence when the receiver is on the z-axis, the length of the chord is maximum. As the receiver is moved away from the z-axis, the length of the chord decreases and therefore also the received voltage. If the arc is very short (narrow aperture) the diminution is gradual; but if the arc is long enough (a wide aperture), it will wind itself around one of the coils of the spiral and the received field will fluctuate.

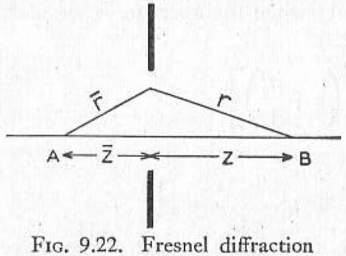


FIG. 9.22. Fresnel diffraction through an aperture.

Consider now a source at some finite distance $\bar z$ from the screen (Fig. 9.22). If the source and the transmitter are on the z-axis, then

$$\bar r = \sqrt{\bar z^2 + \hat x^2 + \hat y^2} = \bar z + \frac{\hat x^2 + \hat y^2}{2\bar z},$$

$$r = \sqrt{z^2 + \hat x^2 + \hat y^2} = z + \frac{\hat x^2 + \hat y^2}{2z}.$$

If the source is an electric current element of moment Il parallel to the screen, then at the aperture we have approximately

$$E_0 = -\frac{i\eta I l e^{-i\beta \bar z}}{2\lambda \bar z} \exp\left(-i\beta \frac{\hat x^2 + \hat y^2}{2\bar z}\right).$$

Substituting in (2) and integrating we obtain the intensity $E(B)$ at the receiver. Thus for a rectangular aperture

$$E(B) = \frac{\eta I l e^{-i\beta(\bar z + z)}}{\lambda(\bar z + z)} [C(u) - iS(u)][C(v) - iS(v)],$$

$$u = a\sqrt{\frac{1}{2\lambda}\left(\frac{1}{\bar z} + \frac{1}{z}\right)}, \quad v = b\sqrt{\frac{1}{2\lambda}\left(\frac{1}{\bar z} + \frac{1}{z}\right)},$$

and if the aperture is circular, then

$$E(B) = \frac{\eta I l e^{-i\beta(\bar z + z)}}{\lambda(\bar z + z)} \sin\left[\frac{\pi a^2}{2\lambda}\left(\frac{1}{\bar z} + \frac{1}{z}\right)\right] \exp\left[-\frac{i\pi a^2}{2\lambda}\left(\frac{1}{\bar z} + \frac{1}{z}\right)\right].$$

These expressions differ from the corresponding expressions for a source at an infinite distance quantitatively but not qualitatively.

9.25. *The Field of Sinusoidally Distributed Currents*

As a rule the calculation of the complete field produced by a given current distribution involves difficult integrations. One important exception is the case of a sinusoidal distribution on a thin straight wire. In section 6.7 it has been shown that such distribution is obtained on an infinitely thin perfectly conducting wire and therefore may be taken as the first approximation to the distribution on a thin wire.

Let the current filament be along the z-axis between $\hat z = z_1$ and $\hat z = z_2$ and let the current $I(\hat z)$ satisfy the following differential equation

$$\frac{d^2 I(\hat z)}{d\hat z^2} = -\beta^2 I(\hat z); \qquad (25\text{–}1)$$

let the current and its derivative be continuous in the interval $z_1 < \hat z < z_2$. The vector potential of a typical element of the filament is

$$dA_z = \Pi I(\hat z)\, d\hat z, \quad \Pi = \frac{e^{-i\beta \hat r}}{4\pi \hat r}, \quad \hat r = \sqrt{\rho^2 + (z - \hat z)^2}. \qquad (25\text{–}2)$$

By (6.1–10) the corresponding electric intensity parallel to the filament is

$$dE_z = \frac{1}{i\omega\epsilon}\left(\frac{\partial^2 \Pi}{\partial z^2} + \beta^2 \Pi\right) I(\hat z)\, d\hat z,$$

$$E_z = \frac{1}{i\omega\epsilon}\int_{z_1}^{z_2} I(\hat z)\frac{\partial^2 \Pi}{\partial z^2}\, d\hat z + \frac{\beta^2}{i\omega\epsilon}\int_{z_1}^{z_2} \Pi I(\hat z)\, d\hat z. \qquad (25\text{–}3)$$

From (2), we have

$$\frac{\partial \hat r}{\partial z} = -\frac{\partial \hat r}{\partial \hat z}, \quad \frac{\partial \Pi}{\partial z} = -\frac{\partial \Pi}{\partial \hat z}, \quad \frac{\partial^2 \Pi}{\partial z^2} = \frac{\partial^2 \Pi}{\partial \hat z^2}.$$

Substituting in (3), we have

$$E_z = \frac{1}{i\omega\epsilon}\int_{z_1}^{z_2} I(\hat z)\frac{\partial^2 \Pi}{\partial \hat z^2}\, d\hat z + \frac{\beta^2}{i\omega\epsilon}\int_{z_1}^{z_2} \Pi I(\hat z)\, d\hat z.$$

Integrating the first term by parts twice, we obtain

$$E_z = \frac{1}{i\omega\epsilon}\left[I(\hat z)\frac{\partial \Pi}{\partial \hat z} - \frac{\partial I}{\partial \hat z}\Pi(\hat z)\right]_{z_1}^{z_2} + \frac{1}{i\omega\epsilon}\int_{z_1}^{z_2} \Pi(\hat z)\left(\frac{\partial^2 I}{\partial \hat z^2} + \beta^2 I\right) d\hat z.$$

In view of (1) the integral vanishes and the integration is completed; thus

$$E_z = \frac{1}{i\omega\epsilon}\left[I(\hat z)\frac{\partial \Pi}{\partial \hat z} - \frac{\partial I}{\partial \hat z}\Pi(\hat z)\right]_{z_1}^{z_2} = -\frac{1}{i\omega\epsilon}\left[I(\hat z)\frac{\partial \Pi}{\partial z} + \frac{\partial I}{\partial \hat z}\Pi\right]_{z_1}^{z_2}. \qquad (25\text{–}4)$$

Let r_1 and r_2 be the distances from the end of the filament to a typical point $P(\rho,\varphi,z)$

$$r_1 = \sqrt{\rho^2 + (z - z_1)^2}, \quad r_2 = \sqrt{\rho^2 + (z - z_2)^2}; \qquad (25\text{–}5)$$

then (4) becomes

$$E_z = \frac{1}{4\pi i\omega\epsilon}\left[I'(z_1)\frac{e^{-i\beta r_1}}{r_1} - I'(z_2)\frac{e^{-i\beta r_2}}{r_2} + I(z_1)\frac{\partial}{\partial z}\frac{e^{-i\beta r_1}}{r_1} - I(z_2)\frac{\partial}{\partial z}\frac{e^{-i\beta r_2}}{r_2}\right]. \qquad (25\text{–}6)$$

Thus the electric intensity parallel to the filament is expressed in terms of the current and its derivatives at the ends of the filament.

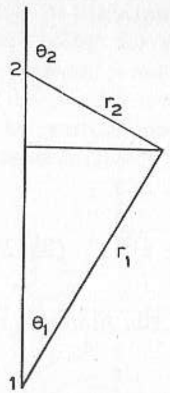


FIG. 9.23. Distances involved in the expressions for the field of a sinusoidal current in a wire of finite length.

By (4.12–9) the magnetic intensity is expressed in terms of E_z as follows

$$\frac{\partial}{\partial\rho}(\rho H_\varphi) = i\omega\epsilon\rho E_z, \quad \rho H_\varphi = i\omega\epsilon\int^\rho \rho E_z\,d\rho + F(z), \qquad (25\text{–}7)$$

where $F(z)$ may be determined from the condition

$$2\pi\rho H_\varphi = I(z), \quad z_1 < z < z_2, \quad \text{as } \rho \to 0,$$
$$= 0, \qquad z < z_1 \quad \text{or} \quad z > z_2. \qquad (25\text{–}8)$$

From (5) we have

$$\rho\,d\rho = r_1\,dr_1, \quad \rho\,d\rho = r_2\,dr_2. \qquad (25\text{–}9)$$

Substituting from (6) and (9) in (7) and integrating, we have

$$-4\pi i\beta\rho H_\varphi$$
$$= I'(z_1)e^{-i\beta r_1} - I'(z_2)e^{-i\beta r_2} + I(z_1)\frac{\partial}{\partial z}e^{-i\beta r_1} - I(z_2)\frac{\partial}{\partial z}e^{-i\beta r_2}, \qquad (25\text{–}10)$$

except, perhaps, for a function of z alone. In order to evaluate the derivatives we note that

$$\frac{\partial}{\partial z}e^{-i\beta r_1} = \frac{\partial}{\partial r_1}e^{-i\beta r_1}\frac{\partial r_1}{\partial z} = -i\beta e^{-i\beta r_1}\frac{z - z_1}{r_1} = -i\beta e^{-i\beta r_1}\cos\theta_1,$$

where θ_1 is defined in Fig. 9.23. A similar expression is obtained for the derivative in the last term. Thus (10) becomes

$$4\pi\rho H_\varphi = \frac{1}{i\beta}I'(z_2)e^{-i\beta r_2} - \frac{1}{i\beta}I'(z_1)e^{-i\beta r_1}$$
$$+ I(z_1)e^{-i\beta r_1}\cos\theta_1 - I(z_2)e^{-i\beta r_2}\cos\theta_2. \qquad (25\text{–}11)$$

Assuming a general solution of (1), substituting in the above equation, and evaluating for $\rho = 0$, we can verify that conditions (8) are satisfied and

that $F(z)$ in (7) vanishes as we have already anticipated in the subsequent equations.

The radial component of the electric intensity is obtained from (4.12–9); thus

$$4\pi i\omega\epsilon\rho E_\rho = I'(z_2)e^{-i\beta r_2}\cos\theta_2 - I'(z_1)e^{-i\beta r_1}\cos\theta_1$$
$$+I(z_1)\left(i\beta\cos^2\theta_1 - \frac{\sin^2\theta_1}{r_1}\right)e^{-i\beta r_1} \qquad (25\text{–}12)$$
$$-I(z_2)\left(i\beta\cos^2\theta_2 - \frac{\sin^2\theta_2}{r_2}\right)e^{-i\beta r_2}.$$

When an infinitely thin wire, extending from $\hat{z} = -l$ to $\hat{z} = l$, is energized at the center, then

$$I(\hat{z}) = I\sin\beta(l - \hat{z}), \quad 0 \le \hat{z} \le l,$$
$$= I\sin\beta(l + \hat{z}), \quad -l \le \hat{z} \le 0, \qquad (25\text{–}13)$$

where I is the maximum amplitude. In this case

$$I(-l) = I(l) = 0, \quad I(0) = I\sin\beta l, \quad I'(-l) = \beta I,$$
$$I'(-0) = \beta I\cos\beta l, \quad I'(+0) = -\beta I\cos\beta l, \quad I'(l) = -\beta I.$$

Since the current vanishes at the ends of the wire and is continuous at the center, the only contributions to E_z come from the first two terms in (6). The derivative $I'(\hat{z})$ is discontinuous at $\hat{z} = 0$, so it is necessary to apply (6) to each half of the wire separately; thus in free space we have

$$E_z = 30iI\left(2\frac{e^{-i\beta r}}{r}\cos\beta l - \frac{e^{-i\beta r_1}}{r_1} - \frac{e^{-i\beta r_2}}{r_2}\right),$$

$$H_\varphi = \frac{iI}{4\pi\rho}(e^{-i\beta r_1} + e^{-i\beta r_2} - 2e^{-i\beta r}\cos\beta l), \qquad (25\text{–}14)$$

$$E_\rho = \frac{30iI}{\rho}(e^{-i\beta r_1}\cos\theta_1 + e^{-i\beta r_2}\cos\theta_2 - 2\cos\beta l\,e^{-i\beta r}\cos\theta).$$

In the case of a perfectly conducting wire E_z should vanish on the surface except in the vicinity of the center where it should be equal and opposite to the applied electric intensity. Hence on a wire of finite radius, no matter how small, the current will deviate from the sinusoidal distribution although the deviation diminishes as the radius of the wire approaches zero. In Chapter 11 we shall obtain a quantitative idea of the magnitude of this deviation.

9.26. *The Mutual Power Radiated by Two Parallel Wires*

Consider two parallel wires (Fig. 9.24) of equal length $2l$. Let each wire be energized at the center so that

$$\begin{aligned} I_1(z) &= I_1 \sin \beta(l - z), \quad I_2(z) = I_2 \sin \beta(l - z), \quad 0 \le z \le l, \\ &= I_1 \sin \beta(l + z), \qquad\quad = I_2 \sin \beta(l + z), \quad -l \le z \le 0, \end{aligned} \tag{26–1}$$

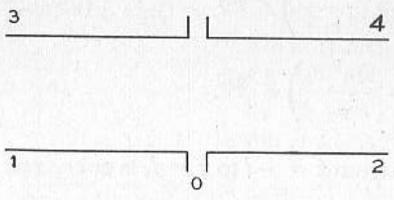


FIG. 9.24. Parallel wires, each energized at the center.

where I_1 and I_2 are the maximum amplitudes. Without loss of generality, we may assume I_1 and I_2 to be in phase. The complex power contributed to the field in virtue of the electromotive force in one wire sustaining the current in it against the field of the other wire is

$$\Psi_{12} = -\tfrac{1}{2}\int_{-l}^{l} E_{1,z}\, I_2^*(z)\, dz = -\tfrac{1}{2}\int_{-l}^{l} E_{2,z}\, I_1^*(z)\, dz. \tag{26–2}$$

Substituting from (25–14), we have

$$\Psi_{12} = 30iI_1I_2\int_0^l \left(\frac{e^{-i\beta r_1}}{r_1} + \frac{e^{-i\beta r_2}}{r_2} - 2\frac{e^{-i\beta r_0}}{r_0}\cos \beta l\right) \sin \beta(l - z)\, dz, \tag{26–3}$$

where r_0 is the distance from the center of one wire to a typical point on the second wire.

Defining the mutual impedance Z_{12} with reference to current antinodes

$$Z_{12} = R_{12} + iX_{12} = \frac{2\Psi_{12}}{I_1I_2}, \tag{26–4}$$

and integrating* (3), we obtain

$$\begin{aligned} R_{12} = {} & 60[2 \operatorname{Ci} \beta\rho - \operatorname{Ci} \beta(r_{04} + l) - \operatorname{Ci} \beta(r_{04} - l)] \\ & + 30[2 \operatorname{Ci} \beta\rho - 2 \operatorname{Ci} \beta(r_{04} + l) - 2 \operatorname{Ci} \beta(r_{04} - l) \\ & + \operatorname{Ci} \beta(r_{14} + 2l) + \operatorname{Ci} \beta(r_{14} - 2l)] \cos 2\beta l \\ & + 30[2 \operatorname{Si} \beta(r_{04} - l) - 2 \operatorname{Si} \beta(r_{04} + l) \\ & + \operatorname{Si} \beta(r_{14} + 2l) - \operatorname{Si} \beta(r_{14} - 2l)] \sin 2\beta l, \end{aligned} \tag{26–5}$$

$$\begin{aligned} X_{12} = {} & 60[\operatorname{Si} \beta(r_{04} + l) + \operatorname{Si} \beta(r_{04} - l) - 2 \operatorname{Si} \beta\rho] \\ & + 30[2 \operatorname{Si} \beta(r_{04} + l) + 2 \operatorname{Si} \beta(r_{04} - l) - 2 \operatorname{Si} \beta\rho \\ & - \operatorname{Si} \beta(r_{14} + 2l) - \operatorname{Si} \beta(r_{14} - 2l)] \cos 2\beta l \\ & + 30[2 \operatorname{Ci} \beta(r_{04} - l) - 2 \operatorname{Ci} \beta(r_{04} + l) \\ & + \operatorname{Ci} \beta(r_{14} + 2l) - \operatorname{Ci} \beta(r_{14} - 2l)] \sin 2\beta l. \end{aligned} \tag{26–6}$$

* See equations (3.7–39) to (3.7–42).

The principal advantage of the present method of calculating the radiated power over the method based on the power flow across an infinite sphere is that it enables us to obtain the reactive power. However, in Chapter 11 we shall show that, in practical situations in which the radii of the wires are small but not infinitely small, the reactance as given by (6) can be used only for a pair of parallel wires so energized that they carry equal and opposite currents, and that in the case of an isolated wire certain modifications are needed. It should also be noted that in the above expressions Z_{12} refers to maximum current amplitudes and not to input currents; just how these expressions can be used to obtain the input impedance will be explained in Chapter 11.

9.27. *Power Radiated by a Straight Antenna Energized at the Center*

The self-impedance (referred to the maximum current amplitude) of an isolated wire of radius a is obtained from the mutual impedance formula of the preceding section if we assume that the distance ρ between the axes of the filaments is equal to the radius of the wire. When the radius is very small we have approximately

$$\sqrt{l^2 + a^2} = l + \frac{a^2}{2l}, \quad \sqrt{4l^2 + a^2} = 2l + \frac{a^2}{4l}. \tag{27–1}$$

Substituting in (26–5) and (26–6) and evaluating, we have

$$\begin{aligned} R = {} & 60(C + \log 2\beta l - \operatorname{Ci} 2\beta l) + 30(\operatorname{Si} 4\beta l - 2 \operatorname{Si} 2\beta l) \sin 2\beta l \\ & + 30(C + \log \beta l - 2 \operatorname{Ci} 2\beta l + \operatorname{Ci} 4\beta l) \cos 2\beta l, \end{aligned} \tag{27–2}$$

$$\begin{aligned} X = {} & 60 \operatorname{Si} 2\beta l + 30(2 \operatorname{Si} 2\beta l - \operatorname{Si} 4\beta l) \cos 2\beta l \\ & - 30\left(\log \frac{l\lambda}{a^2} - C - \log 2\pi - \operatorname{Ci} 4\beta l + 2 \operatorname{Ci} 2\beta l\right) \sin 2\beta l. \end{aligned} \tag{27–3}$$

If I is the maximum current amplitude, the radiated power is $W = \frac{1}{2}RI^2$. When expressed in this form the radiated power is independent of the radius of the wire. If $I(0)$ is the input current, then as the radius of the wire approaches zero we have $I(0) = I \sin \beta l$; hence the asymptotic expression for the input resistance is

$$R_i = \frac{R}{\sin^2 \beta l}. \tag{27–4}$$

If we plot this R as a function of βl, we obtain a curve which is approached by the input resistance curves for finite radii as the latter approach zero.

9.28. *Power Radiated by a Pair of Parallel Wires*

Consider now two parallel wires of length $2l$ (Fig. 9.24) carrying equal and opposite currents. The complex power Ψ is

$$\Psi = (Z - Z_{12})I^2, \tag{28–1}$$

where Z is the self-impedance of either wire.

Evaluating Z_{12}, we find that when the interaxial separation d is small, then

$$Z_{12} = Z - \tfrac{1}{2}\hat{Z}, \quad \hat{Z} = \hat{R} + i\hat{X}, \tag{28-2}$$

$$\hat{R} = \frac{240\pi^2 d^2}{\lambda^2}\left(1 + \tfrac{1}{2}\cos 2\beta l - \frac{3\sin 2\beta l}{4\beta l} + \frac{1}{2\beta^2 l^2} + \frac{7\cos 2\beta l}{16\beta^2 l^2}\right), \tag{28-3}$$

$$\hat{X} = \frac{480\pi d}{\lambda}(1 + \tfrac{1}{2}\cos 2\beta l) - 120\log\frac{d}{a}\sin 2\beta l, \tag{28-4}$$

and the power radiated by the two wires is $W = \frac{1}{2}\hat{R}I^2$. It should be noted that as the length of the wires increases, the radiated power fluctuates between two small limits.

CHAPTER X

WAVES, WAVE GUIDES, AND RESONATORS — 2

10.1. *Transverse Magnetic Plane Waves* (TM–waves)

Assuming that transverse magnetic waves are traveling in the z-direction, we have by definition $H_z = 0$. Since there is no longitudinal magnetic current, the transverse electric intensity can be expressed as the gradient of a scalar potential

$$E_t = -\text{grad } V. \tag{1-1}$$

The divergence equation for H becomes

$$\frac{\partial H_x}{\partial x} + \frac{\partial H_y}{\partial y} = 0, \tag{1-2}$$

and the magnetic intensity can be derived from a stream function Π

$$H_x = \frac{\partial \Pi}{\partial y}, \quad H_y = -\frac{\partial \Pi}{\partial x}. \tag{1-3}$$

Hence we may write

$$H = \text{curl } A, \quad A_x = A_y = 0, \qquad A_z = \Pi, \tag{1-4}$$

where Π satisfies the wave equation

$$\frac{\partial^2 \Pi}{\partial x^2} + \frac{\partial^2 \Pi}{\partial y^2} + \frac{\partial^2 \Pi}{\partial z^2} = \sigma^2 \Pi. \tag{1-5}$$

For the electric intensity we have

$$E = \frac{\text{curl } H}{g + i\omega\epsilon} = \frac{\text{grad div } A - \Delta A}{g + i\omega\epsilon} = \frac{\text{grad div } A - \sigma^2 A}{g + i\omega\epsilon}.$$

Since A has only a z-component the transverse electric intensity is

$$E_t = \frac{\text{grad div } A}{g + i\omega\epsilon} = \frac{1}{g + i\omega\epsilon}\,\text{grad}\,\frac{\partial \Pi}{\partial z},$$

and in equation (1) we may assume

$$V = -\frac{1}{g + i\omega\epsilon}\frac{\partial \Pi}{\partial z}. \tag{1-6}$$

The longitudinal electric intensity is then

$$E_z = \frac{1}{g + i\omega\epsilon}\left(\frac{\partial^2 \Pi}{\partial z^2} - \sigma^2 \Pi\right) = -\frac{1}{g + i\omega\epsilon}\left(\frac{\partial^2 \Pi}{\partial x^2} + \frac{\partial^2 \Pi}{\partial y^2}\right). \quad (1\text{–}7)$$

Thus the entire field has been expressed in terms of one scalar wave function Π.

The foregoing expressions are general for any field in which $H_z = 0$. For waves in which the field pattern in planes parallel to the xy-plane is the same, we have

$$\Pi = T(x,y)\hat{T}(z). \quad (1\text{–}8)$$

In general $T(x,y)$ may be complex

$$T(x,y) = T_1(x,y) + iT_2(x,y); \quad (1\text{–}9)$$

but for plane waves there should be no phase change in any direction parallel to the xy-plane (except for a complete reversal) and $T(x,y)$ must be real except for a possibly complex constant factor. This factor can always be included in $\hat{T}(z)$, and hence for plane waves $T(x,y)$ may be taken as real.

Substituting from (8) in (5) and dividing by Π, we have

$$\frac{1}{T}\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right) + \frac{1}{\hat{T}}\frac{d^2\hat{T}}{dz^2} = \sigma^2. \quad (1\text{–}10)$$

The first term is independent of z and the second is independent of x and y, while the sum is a constant; hence each term is a constant and

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = -\chi^2 T. \quad (1\text{–}11)$$

For plane waves χ^2 is real and χ itself is either real or a pure imaginary.

Substituting from (11) in (10), we have

$$\frac{d^2\hat{T}}{dz^2} = \Gamma^2\hat{T}, \quad \Gamma = \sqrt{\sigma^2 + \chi^2}; \quad \hat{T}(z) = Pe^{-\Gamma z} + Qe^{\Gamma z}. \quad (1\text{–}12)$$

Substituting from (8) and (11) in (7), we find that the longitudinal electric intensity and the electric current density differ from the stream function Π by constant factors

$$E_z = \frac{\chi^2}{g + i\omega\epsilon}\Pi, \quad (g + i\omega\epsilon)E_z = \chi^2\Pi. \quad (1\text{–}13)$$

In cartesian coordinates the transverse field intensities are

$$H_x = \frac{\partial T}{\partial y}\hat{T}, \quad H_y = -\frac{\partial T}{\partial x}\hat{T}, \quad V = -\frac{T}{g + i\omega\epsilon}\frac{d\hat{T}}{dz},$$

$$E_x = Z_z^+ H_y, \quad E_y = -Z_z^+ H_x, \quad Z_z^+ = -\frac{1}{(g + i\omega\epsilon)\hat{T}}\frac{d\hat{T}}{dz}. \quad (1\text{–}14)$$

In cylindrical coordinates we have

$$H_\rho = \frac{\partial T}{\rho\partial\varphi}\hat{T}, \quad H_\varphi = -\frac{\partial T}{\partial\rho}\hat{T}, \quad E_\rho = Z_z^+ H_\varphi, \quad E_\varphi = -Z_z^+ H_\rho. \quad (1\text{–}15)$$

For progressive waves traveling in the positive z-direction, we obtain

$$\hat{T} = Pe^{-\Gamma z}, \quad Z_z^+ = K_z = \frac{\Gamma}{g + i\omega\epsilon}. \quad (1\text{–}16)$$

Then in cartesian and cylindrical coordinates we have

$$H_x = P\frac{\partial T}{\partial y}e^{-\Gamma z}, \quad H_y = -P\frac{\partial T}{\partial x}e^{-\Gamma z},$$

$$E_x = K_z H_y, \quad E_y = -K_z H_x;$$

$$H_\rho = P\frac{\partial T}{\rho\partial\varphi}e^{-\Gamma z}, \quad H_\varphi = -P\frac{\partial T}{\partial\rho}e^{-\Gamma z}, \quad (1\text{–}17)$$

$$E_\rho = K_z H_\varphi, \quad E_\varphi = -K_z H_\rho.$$

In nondissipative media the propagation constant Γ and the wave impedance K_z for progressive waves become

$$\Gamma = \sqrt{\chi^2 - \beta^2}, \quad K_z = \frac{\Gamma}{i\omega\epsilon} = \frac{\eta\Gamma}{i\beta}. \quad (1\text{–}18)$$

When the frequency is such that

$$\beta = \frac{\omega}{v} = \frac{2\pi}{\lambda} = \chi,$$

then $\Gamma = 0$. The frequency so defined is the cut-off frequency, since for lower frequencies Γ is real and on the average no energy is transmitted in the z-direction. The cut-off frequency and the corresponding wavelength are

$$\omega_c = 2\pi f_c = \chi v = \frac{\chi}{\sqrt{\mu\epsilon}}, \quad \lambda_c = \frac{2\pi}{\chi}. \quad (1\text{–}19)$$

In terms of the frequency ratio

$$\nu = \frac{f_c}{f} = \frac{\lambda}{\lambda_c} = \frac{\chi}{\beta}, \tag{1–20}$$

equations (18) become, above the cut-off,

$$\Gamma = i\beta\sqrt{1-\nu^2}, \quad K_z = \eta\sqrt{1-\nu^2}. \tag{1–21}$$

Below the cut-off Γ is positive real and the wave impedance is a negative reactance. Sufficiently below the cut-off, we have approximately

$$\Gamma = \chi, \quad K_z = \frac{\chi}{i\omega\epsilon}, \tag{1–22}$$

and the wave impedance is substantially a capacitance ϵ/χ.

Equations (6) and (12) imply that V and Π satisfy the following transmission equations

$$\frac{\partial V}{\partial z} = -\left(i\omega\mu + \frac{\chi^2}{g+i\omega\epsilon}\right)\Pi, \quad \frac{\partial \Pi}{\partial z} = -(g+i\omega\epsilon)V. \tag{1–23}$$

These are the equations of a transmission line with series distributed constants per unit length equivalent to an inductance μ plus a capacitance ϵ/χ^2 in parallel with a conductance g/χ^2, and with shunt distributed constants per unit length consisting of a conductance g and a capacitance ϵ.

Equipotential lines in transverse planes are defined by the following family of curves

$$T(x,y) = \text{const.} \quad \text{or} \quad T(\rho,\varphi) = \text{const.} \tag{1–24}$$

Differentiating along each curve, we have

$$\frac{\partial T}{\partial x}dx + \frac{\partial T}{\partial y}dy = 0.$$

Since the partial derivatives are proportional to $-H_y$ and H_x (see eq. 14), we obtain $dy/dx = H_y/H_x$; thus the magnetic lines coincide with the equipotential lines. On the other hand we have $E_xH_x + E_yH_y = 0$; hence the transverse electric intensity is perpendicular to the magnetic intensity.

The magnetic intensity is linearly polarized at each point. For progressive waves in nondissipative media the transverse electric intensity is in phase with the magnetic intensity while the longitudinal electric intensity is in quadrature. Thus in the case of progressive waves the electric intensity is elliptically polarized and the plane of the ellipse is perpendicular to H.

Consider a cylindrical strip of unit length in the direction of its generators (the z-axis) and suppose it intersects the xy-plane in a curve PQ (Fig. 10.1). Then the magnetic flux Φ crossing the strip in the counterclockwise direction depends only on the end points of the curve; thus

$$\Phi = \mu\int_{(PQ)}(H_y\,dx - H_x\,dy) = -\mu\int_{(PQ)}\left(\frac{\partial \Pi}{\partial x}dx + \frac{\partial \Pi}{\partial y}dy\right),$$

or

$$\Phi = \mu[\Pi(P) - \Pi(Q)].$$

In an unlimited homogeneous medium plane waves can be generated only by current sheets of infinite extent. Consider such a sheet in the xy-plane. The lines of flow should be perpendicular to the magnetic intensity. The electric intensity should be continuous across the sheet while the magnetic intensity should be discontinuous and the discontinuity should equal the current density. This sheet will generate two progressive waves of equal amplitudes, traveling in opposite directions; hence from (17) we find the current density in the sheet

$$J_x = 2P\frac{\partial T}{\partial x}, \quad J_y = 2P\frac{\partial T}{\partial y}.$$

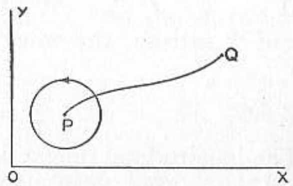


FIG. 10.1. Cross-section PQ of a cylindrical strip whose generators are parallel to the z-axis.

The current density J is subject to two restrictions. The less stringent one is obtained by differentiating the above equations

$$\frac{\partial J_x}{\partial y} = \frac{\partial J_y}{\partial x}.$$

The more stringent restriction demands that T be a solution of (11). Let us suppose that the current density is

$$J_x = 2P_1\frac{\partial T_1}{\partial x} + 2P_2\frac{\partial T_2}{\partial x}, \quad J_y = 2P_1\frac{\partial T_1}{\partial y} + 2P_2\frac{\partial T_2}{\partial y},$$

where T_1 is a solution of (11) with $\chi = \chi_1$ and T_2 is a solution with $\chi = \chi_2$. In this case two waves will be produced on each side of the current sheet. These waves will travel with different propagation constants. At the current sheet the field pattern will conform to the impressed pattern $T_1 + T_2$; but this pattern will not be maintained as the waves travel away from the sheet. Complicated fields can be constructed by superposing transverse magnetic waves with finite amplitudes or with infinitely small amplitudes as for instance

$$J_x = \int \frac{\partial}{\partial x}T(x,y;\chi)\,d\chi, \quad J_y = \int \frac{\partial}{\partial y}T(x,y;\chi)\,d\chi.$$

10.2. *Transverse Electric Plane Waves* (TE–waves)

Transverse electric waves represent a counterpart of transverse magnetic waves; the theories of these two types of plane waves are closely analogous. Thus by definition we have $E_z = 0$. Since there is no longitudinal electric current, the transverse magnetic intensity can be expressed as the gradient of a scalar potential

$$H_t = -\text{grad } U. \tag{2–1}$$

The divergence equation for E becomes

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} = 0, \tag{2–2}$$

and the electric intensity can be derived from a stream function Ψ

$$E_x = -\frac{\partial \Psi}{\partial y}, \quad E_y = \frac{\partial \Psi}{\partial x}. \tag{2–3}$$

Hence

$$E = -\text{ curl } F, \quad \text{where} \quad F_x = F_y = 0, \quad F_z = \Psi, \tag{2–4}$$

and Ψ satisfies the wave equation. As in the preceding section we find

$$U = -\frac{1}{i\omega\mu}\frac{\partial \Psi}{\partial z}. \tag{2–5}$$

The longitudinal magnetic intensity is then

$$H_z = \frac{1}{i\omega\mu}\left(\frac{\partial^2 \Psi}{\partial z^2} - \sigma^2\Psi\right) = -\frac{1}{i\omega\mu}\left(\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2}\right). \tag{2–6}$$

Thus the entire field has been expressed in terms of one scalar wave function Ψ.

The foregoing expressions are general for any field in which $E_z = 0$. For waves in which the field pattern in planes parallel to the xy-plane is the same we have $\Psi = T(x,y)\hat{T}(z)$. In general $T(x,y)$ may be complex as in (1–9); but for plane waves it is real. The functions T and $\hat{T}$ satisfy the same equations as in the case of transverse magnetic waves. Substituting in (6), we find that the longitudinal magnetic intensity and magnetic current density are constant multiples of Ψ

$$H_z = \frac{\chi^2}{i\omega\mu}\Psi, \quad i\omega\mu H_z = \chi^2\Psi. \tag{2–7}$$

In cartesian coordinates the transverse field intensities are

$$E_x = -\frac{\partial T}{\partial y}\hat{T}, \quad E_y = \frac{\partial T}{\partial x}\hat{T}, \quad U = -\frac{T}{i\omega\mu}\frac{d\hat{T}}{dz},$$
$$H_x = -Y_z^+ E_y, \quad H_y = Y_z^+ E_x, \quad Y_z^+ = \frac{1}{Z_z^+} = -\frac{1}{i\omega\mu\hat{T}}\frac{d\hat{T}}{dz}. \tag{2–8}$$

In cylindrical coordinates we have

$$E_\rho = -\frac{\partial T}{\rho\partial\varphi}\hat{T}, \quad E_\varphi = \frac{\partial T}{\partial\rho}\hat{T}, \quad H_\rho = -Y_z^+ E_\varphi, \quad H_\varphi = Y_z^+ E_\rho. \tag{2–9}$$

For progressive waves traveling in the positive z-direction, we obtain

$$\hat{T} = Pe^{-\Gamma z}, \quad Z_z^+ = K_z = \frac{1}{M_z} = \frac{i\omega\mu}{\Gamma}. \tag{2–10}$$

Then in cartesian and cylindrical coordinates we have

$$E_x = -P\frac{\partial T}{\partial y}e^{-\Gamma z}, \quad E_y = P\frac{\partial T}{\partial x}e^{-\Gamma z}, \quad H_x = -M_zE_y, \quad H_y = M_zE_x,$$
$$E_\rho = -P\frac{\partial T}{\rho\partial\varphi}e^{-\Gamma z}, \quad E_\varphi = P\frac{\partial T}{\partial\rho}e^{-\Gamma z}, \quad H_\rho = -M_zE_\varphi, \quad H_\varphi = M_zE_\rho. \tag{2–11}$$

In nondissipative media the propagation constant Γ and the wave impedance K_z for progressive waves are

$$\Gamma = \sqrt{\chi^2 - \beta^2}, \quad K_z = \frac{i\omega\mu}{\Gamma} = \frac{i\eta\beta}{\Gamma}. \tag{2–12}$$

The cut-off frequency is expressed in terms of χ by the same formula (1–19) as in the case of transverse magnetic waves. In terms of the frequency ratio defined by (1–20), Γ and K_z may be written

$$\Gamma = i\beta\sqrt{1 - \nu^2}, \quad K_z = \frac{\eta}{\sqrt{1 - \nu^2}}. \tag{2–13}$$

Below the cut-off Γ is positive real and the wave impedance is a positive reactance. Sufficiently below the cut-off, we have approximately

$$\Gamma = \chi, \quad K_z = \frac{i\omega\mu}{\chi}, \tag{2–14}$$

and the wave impedance is substantially an inductance μ/χ.

As in the case of *TM*-waves U and Ψ satisfy the following transmission equations

$$\frac{\partial \Psi}{\partial z} = -i\omega\mu U, \quad \frac{\partial U}{\partial z} = -\left(g + i\omega\epsilon + \frac{\chi^2}{i\omega\mu}\right)\Psi. \tag{2–15}$$

These equations are the equations of a transmission line with series distributed inductance μ, shunt conductance g, shunt capacitance ϵ, and shunt inductance μ/χ^2, all per unit length.

Equation (1–24) now represents the family of magnetic equipotential lines. Electric lines coincide with these equipotential lines and the transverse component of H is perpendicular to E. The electric intensity is linearly polarized, although the direction of polarization varies, in general, from point to point. For progressive waves in nondissipative media, the transverse magnetic intensity is in phase with the electric intensity, while the longitudinal intensity is in quadrature. Thus in the case of progressive waves the magnetic intensity is elliptically polarized and the plane of the ellipse is perpendicular to E.

The electric displacement crossing counterclockwise the cylindrical strip of Fig. 10.1 depends only on the end points of the curve; thus

$$q = \epsilon \int_{(PQ)} (E_y\,dx - E_x\,dy) = \epsilon[\Psi(Q) - \Psi(P)].$$

In an unlimited medium progressive waves can be generated by an infinite plane current sheet. If the generating current sheet is in the xy-plane, then the current density is

$$J_x = -2H_y(x,y,+0) = 2PM_z\frac{\partial T}{\partial y}, \quad J_y = 2H_x(x,y,+0) = -2PM_z\frac{\partial T}{\partial x},$$
$$\frac{\partial J_x}{\partial x} + \frac{\partial J_y}{\partial y} = 0.$$

Let us now suppose that we have an arbitrary current sheet. If we can find two functions T_1 and T_2 satisfying the following equations

$$\frac{\partial T_1}{\partial x} + \frac{\partial T_2}{\partial y} = J_x, \quad \frac{\partial T_1}{\partial y} - \frac{\partial T_2}{\partial x} = J_y,$$

we shall be able to separate that part of the total field for which $H_z = 0$ from the other part for which $E_z = 0$. Eliminating either T_2 or T_1 from the above equations, we have

$$\frac{\partial^2 T_1}{\partial x^2} + \frac{\partial^2 T_1}{\partial y^2} = \frac{\partial J_x}{\partial x} + \frac{\partial J_y}{\partial y}, \quad \frac{\partial^2 T_2}{\partial x^2} + \frac{\partial^2 T_2}{\partial y^2} = \frac{\partial J_x}{\partial y} - \frac{\partial J_y}{\partial x}.$$

Thus we have a pair of partial differential equations for the unknown functions T_1 and T_2

10.3. *General Expressions for Electromagnetic Fields in Terms of Two Scalar Wave Functions*

The most general electromagnetic field in a source-free region can be expressed in terms of two scalar wave functions Π and Ψ. Suppose we start with a given general field and determine a particular Π from

$$\frac{\partial^2 \Pi}{\partial z^2} - \sigma^2 \Pi = (g + i\omega\epsilon)E_z.$$

If we calculate the field defined by Π and subtract it from the total field, the remainder will be a field in which $E_z = 0$. From H_z we can find another scalar function satisfying

$$\frac{\partial^2 \Psi}{\partial z^2} - \sigma^2 \Psi = i\omega\mu H_z.$$

If the sum of the fields obtained from Π and Ψ is subtracted from the original field, the remainder is a transverse electromagnetic field. It has been shown in section 8.9 that such fields may be expressed in terms of wave functions satisfying the two-dimensional form of Laplace's equation; these functions may be included in Π and Ψ. Hence the two scalar wave functions are sufficient for the complete expression of any electromagnetic field in a source-free region.

10.4. *Natural Waves in Cylindrical Wave Guides*

For practical purposes the best procedure is to start with the theory of wave transmission in regions bounded by perfectly conducting cylinders. Later the first order correction due to finite conductivity which manifests itself primarily in the attenuation will be considered. On the boundary of any perfectly conducting wave guide the tangential electric intensity should vanish. For transverse magnetic plane waves E_z is proportional to the amplitude distribution function $T(x,y)$ and therefore, on the boundary,

$$T(x,y) = 0. \tag{4–1}$$

Then the electric potential V, and hence the tangential component of the transverse electric intensity will also vanish on the boundary and the above equation will represent the only added restriction on T.

The vanishing of the tangential electric intensity is equivalent to the vanishing of the normal component of the magnetic intensity. In the case of transverse electric waves the latter is the normal derivative of the magnetic potential U; since U is proportional to T, we have as the boundary condition

$$\frac{\partial T}{\partial n} = 0. \tag{4–2}$$

The above boundary conditions impose restrictions on the constant χ. We have already shown that for plane waves χ^2 is real. For waves in homogeneous regions bounded by perfectly conducting cylinders χ itself is real. Thus if U and V are any two wave functions then by the two-dimensional Green's theorem we have

$$\iint_{(S)} \left(\frac{\partial U}{\partial x}\frac{\partial V}{\partial x} + \frac{\partial U}{\partial y}\frac{\partial V}{\partial y}\right) dS = -\iint_{(S)} U\left(\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2}\right) dS + \int_{(s)} U\frac{\partial V}{\partial n}\, ds, \quad (4\text{–}3)$$

where the double integration is extended over an area S and the simple integration over its periphery. Let $U = V = T(x,y)$; then

$$\iint_{(S)} |\operatorname{grad} T|^2\, dS = \chi^2 \iint_{(S)} T^2\, dS + \int_{(s)} T\frac{\partial T}{\partial n}\, ds. \quad (4\text{–}4)$$

For waves in a guide either T or $\partial T/\partial n$ vanishes on the periphery; hence

$$\chi^2 = \frac{\iint |\operatorname{grad} T|^2\, dS}{\iint T^2\, dS}. \quad (4\text{–}5)$$

The right side is positive and χ must be real; there is no loss of generality if we assume it positive.

If the wave is not plane T may be complex. By substituting $U = T$ and $V = T^*$ in (3), we can show that χ is still real. Hence if T is of the form (1–9), T_1 and T_2 satisfy (1–11) and any wave in the wave guide may be resolved into plane waves.

At the cut-off $\chi = \omega/v$. Since equation (1–11) is satisfied also by the displacement of a vibrating membrane, the cut-off frequencies in a cylindrical wave guide are proportional to the natural frequencies of membranes equal in shape and area to the cross-section of the wave guide. The boundary condition for transverse magnetic waves corresponds to a fixed edge of the membrane.

The following transformation

$$u = \frac{x\sqrt{\pi}}{\sqrt{S}}, \quad v = \frac{y\sqrt{\pi}}{\sqrt{S}} \quad (4\text{–}6)$$

preserves the shape of the cross-section of the wave guide and changes the area S, making it equal to the area of a circle of unit radius. Applying this transformation to (5), we obtain

$$\chi = \frac{k\sqrt{\pi}}{\sqrt{S}}, \quad k^2 = \frac{\iint \left[\left(\frac{\partial T}{\partial u}\right)^2 + \left(\frac{\partial T}{\partial v}\right)^2\right] du\, dv}{\iint T^2\, du\, dv}, \quad (4\text{–}7)$$

where the *modular constant* k depends only on the shape of the cross-section and the particular transmission mode.

In a progressive transverse magnetic wave at frequencies above the cut-off, the average magnetic energy W^m per unit length of the guide, the average energy W_t^e associated with the transverse electric field and the average energy W_l^e associated with the longitudinal electric field may be expressed in terms of T as follows

$$W^m = \tfrac{1}{4}\mu \iint HH^*\, dS = \tfrac{1}{4}\mu \iint |\operatorname{grad} T|^2\, dS,$$

$$W_t^e = \tfrac{1}{4}\epsilon \iint E_t E_t^*\, dS = \tfrac{1}{4}\mu(1 - \nu^2) \iint |\operatorname{grad} T|^2\, dS,$$

$$W_l^e = \tfrac{1}{4}\epsilon \iint E_z E_z^*\, dS = \tfrac{1}{4}\mu\nu^2\chi^2 \iint T^2\, dS.$$

Using (5) the first two expressions may be written

$$W^m = \tfrac{1}{4}\mu\chi^2 \iint T^2\, dS, \quad W_t^e = \tfrac{1}{4}\mu\chi^2(1 - \nu^2) \iint T^2\, dS.$$

Hence the average magnetic energy in a progressive transverse magnetic wave is equal to the average electric energy. At the cut-off the electric energy is associated exclusively with the longitudinal field and sufficiently above the cut-off most of it is associated with the transverse field.

For transverse electric waves we have

$$W^e = \tfrac{1}{4}\epsilon\chi^2 \iint T^2\, dS, \quad W_l^m = \tfrac{1}{4}\epsilon\chi^2\nu^2 \iint T^2\, dS,$$

$$W_t^m = \tfrac{1}{4}\epsilon\chi^2(1 - \nu^2) \iint T^2\, dS.$$

In this case the average electric and magnetic energies are also equal. At the cut-off the magnetic energy is associated exclusively with the longitudinal field and sufficiently above the cut-off it is associated largely with the transverse field.

Above the cut-off the average power carried by progressive *TM*-waves in the direction of the guide is

$$W = \tfrac{1}{2} \iint (E_x H_y^* - E_y H_x^*)\, dS = \tfrac{1}{2}\chi^2 K_z \iint T^2\, dS, \quad (4\text{–}8)$$

while for *TE*-waves we have

$$W = \frac{\chi^2}{2K_z} \iint T^2\, dS. \quad (4\text{–}9)$$

Since T is proportional to the longitudinal current density, electric or magnetic as the case may be, the power transfer is proportional to the mean-square longitudinal current.

The conduction current density in the cylinder is equal to the tangential component of the magnetic intensity. Hence for transverse magnetic waves the conduction current is strictly longitudinal and its density is

$$J_z = \frac{\partial \Pi}{\partial n} = \frac{\partial T}{\partial n}\hat{T}(z), \tag{4–10}$$

where n is the outward normal to the cylinder.

In the case of transverse electric waves the longitudinal density is the tangential derivative

$$J_z = \frac{\partial U}{\partial s} = -\frac{1}{i\omega\mu}\frac{\partial^2 \Psi}{\partial s\,\partial z} = -\frac{1}{i\omega\mu}\frac{\partial T}{\partial s}\frac{d\hat{T}}{dz} \tag{4–11}$$

of the magnetic potential, taken in the counterclockwise direction as seen from the positive side of the xy-plane; the counterclockwise transverse conduction current density is

$$J_t = \frac{\chi^2}{i\omega\mu}\Psi = \frac{\chi^2}{i\omega\mu}T\hat{T}. \tag{4–12}$$

For progressive TE-waves traveling in the z-direction, we have

$$J_z = \frac{\Gamma}{i\omega\mu}\frac{\partial T}{\partial s}e^{-\Gamma z} = \frac{1}{K_z}\frac{\partial T}{\partial s}e^{-\Gamma z}, \quad J_t = \frac{\chi^2}{i\omega\mu}Te^{-\Gamma z}. \tag{4–13}$$

Below the cut-off J_z and J_t are in phase; above the cut-off they are in quadrature. If J_z is kept constant as the frequency increases, J_t approaches zero. The total longitudinal current is proportional to the total change in T around the periphery; hence this total current equals zero.

In the case of progressive transverse magnetic waves the power absorbed by an imperfectly conducting cylinder is obtained as usual by integrating the square of the tangential magnetic intensity or the square of the conduction current density; thus

$$\mathcal{W} = \tfrac{1}{2}\mathcal{R}\int J_z J_z^*\,ds = \tfrac{1}{2}\mathcal{R}\int\left(\frac{\partial T}{\partial n}\right)^2 ds = \frac{\mathcal{R}}{2s}I^2, \tag{4–14}$$

where s is the length of the periphery and I is the root mean square conduction current. The attenuation constant (due to the losses in the conductor) is therefore

$$\alpha = \frac{\mathcal{W}}{2W} = \frac{\mathcal{R}\int\left(\frac{\partial T}{\partial n}\right)^2 ds}{2\eta\chi^2\int\int T^2\,dS}(1-\nu^2)^{-1/2}. \tag{4–15}$$

On the other hand, for progressive transverse electric waves we have the following expressions for the dissipated power (if the dielectric is nondissipative)

$$\mathcal{W}_1 = \tfrac{1}{2}\mathcal{R}\int J_z J_z^*\,ds = \frac{\mathcal{R}}{2K_z^2}\int\left(\frac{\partial T}{\partial s}\right)^2 ds = \frac{1-\nu^2}{2\eta^2}\mathcal{R}\int\left(\frac{\partial T}{\partial s}\right)^2 ds, \tag{4–16}$$

$$\mathcal{W}_2 = \tfrac{1}{2}\mathcal{R}\int J_t J_t^*\,ds = \frac{\chi^2\nu^2\mathcal{R}}{2\eta^2}\int T^2\,ds.$$

Hence the general expression for the attenuation constant is

$$\alpha = \frac{\mathcal{R}}{2\eta}\left[\frac{\int\left(\frac{\partial T}{\partial s}\right)^2 ds}{\chi^2\int\int T^2\,dS}\sqrt{1-\nu^2} + \frac{\int T^2\,ds}{\int\int T^2\,dS}\frac{\nu^2}{\sqrt{1-\nu^2}}\right]. \tag{4–17}$$

10.5. *Natural Waves in Rectangular Wave Guides*

Let a rectangular wave guide be bounded by the following planes

$$x = 0, \quad x = a, \quad y = 0, \quad y = b.$$

For transverse magnetic waves the boundary conditions are then

$$T(0,y) = T(a,y) = T(x,0) = T(x,b) = 0.$$

Particular solutions satisfying these conditions and the corresponding values of χ are

$$T(x,y) = \sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}, \quad \chi^2 = \left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2,$$

where $m,n = 1,2,3,\cdots$. To each pair of nonvanishing integers there corresponds a definite field pattern and a definite propagation constant. The cut-off wavelength of the "$TM_{m,n}$-wave" is

$$\lambda_{m,n} = \frac{2\pi}{\chi_{m,n}} = \frac{2}{\sqrt{\frac{m^2}{a^2}+\frac{n^2}{b^2}}} = \frac{2ab}{\sqrt{m^2b^2+n^2a^2}}. \tag{5–1}$$

The longest wavelength corresponds to $m = n = 1$

$$\lambda_{1,1} = \frac{2ab}{\sqrt{a^2+b^2}}. \tag{5–2}$$

Thus the cut-off wavelength for the dominant transverse magnetic wave is twice the distance from a vertex of the cross-section to the opposite diagonal (Fig. 10.2). A typical higher mode is the dominant mode in a guide with dimensions a/m and b/n. The cut-off frequencies are inversely proportional to the corresponding wavelengths; hence these frequencies are proportional to the radii from the origin to points (na,mb). The attenuation constant is obtained from the general formula of the preceding section; thus

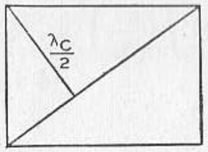


FIG. 10.2. Relation of the cutoff wavelength for the dominant transverse magnetic wave ($TM_{1,1}$-wave) to the dimensions of the guide.

$$\alpha = \frac{2\mathcal{R}(m^2b^3 + n^2a^3)}{\eta ab(m^2b^2 + n^2a^2)}(1 - \nu_{m,n}^2)^{-1/2}. \quad (5\text{–}3)$$

For wave guides of square cross-section this becomes

$$\alpha = \frac{2\mathcal{R}}{\eta a}(1 - \nu_{m,n}^2)^{-1/2}. \quad (5\text{–}4)$$

The field of a $TM_{m,n}$-wave is then obtained in the form

$$H_x = \frac{n\pi}{b}\sin\frac{m\pi x}{a}\cos\frac{n\pi y}{b}\hat{T}_{m,n}(z),$$

$$H_y = -\frac{m\pi}{a}\cos\frac{m\pi x}{a}\sin\frac{n\pi y}{b}\hat{T}_{m,n}(z),$$

$$E_x = Z_{m,n}^+H_y, \quad E_y = -Z_{m,n}^+H_x, \quad Z_{m,n}^+ = -\frac{1}{(g + i\omega\epsilon)\hat{T}_{m,n}}\frac{d\hat{T}_{m,n}}{dz},$$

$$E_z = \frac{\chi_{m,n}^2}{g + i\omega\epsilon}\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}\hat{T}_{m,n}(z),$$

$$\hat{T}_{m,n} = A_{m,n}e^{-\Gamma_{m,n}z} + B_{m,n}e^{\Gamma_{m,n}z}, \quad \Gamma_{m,n} = \sqrt{\chi_{m,n}^2 + \sigma^2}.$$

In the case of transverse electric waves $\partial T/\partial x$ must vanish at $x = 0$ and $x = a$ and $\partial T/\partial y$ must vanish at $y = 0$ and $y = b$; therefore

$$T(x,y) = \cos\frac{m\pi x}{a}\cos\frac{n\pi y}{b},$$

where m and n are integers, not equal to zero simultaneously. Hence the cut-off frequencies are given by (1), but since either m or n may be zero, there exist more transverse electric modes than transverse magnetic. The cut-off wavelengths of the additional modes are

$$\lambda_{m,0} = \frac{2a}{m}, \quad \lambda_{0,n} = \frac{2b}{n}. \quad (5\text{–}5)$$

For these modes the electric lines are parallel to one of the faces of the wave guide. From the general formula the attenuation constant of transverse electric waves is found to be

$$\alpha = \frac{2\mathcal{R}}{\eta b}\left[\frac{p(pm^2 + n^2)}{p^2m^2 + n^2}\sqrt{1 - \nu_{m,n}^2} + \frac{(1 + p)\nu_{m,n}^2}{\sqrt{1 - \nu_{m,n}^2}}\right], \quad m \neq 0, \quad n \neq 0,$$

$$\alpha = \frac{\mathcal{R}}{\eta}\left(\frac{1}{b} + \frac{2}{a}\nu_{m,0}^2\right)(1 - \nu_{m,0}^2)^{-1/2}, \quad m \neq 0, \quad n = 0, \quad (5\text{–}6)$$

$$\alpha = \frac{\mathcal{R}}{\eta}\left(\frac{1}{a} + \frac{2}{b}\nu_{0,n}^2\right)(1 - \nu_{0,n}^2)^{-1/2}, \quad m = 0, \quad n \neq 0,$$

where $p = b/a$. In square tubes this reduces to

$$\alpha = \frac{2\mathcal{R}(1 + \nu^2)}{\eta a\sqrt{1 - \nu^2}}, \text{ if } m, n \neq 0; \; \alpha = \frac{\mathcal{R}(1 + 2\nu^2)}{\eta a\sqrt{1 - \nu^2}}, \text{ if } m \neq 0, n = 0. \quad (5\text{–}7)$$

For the field of a $TE_{m,n}$-wave we obtain

$$E_x = \frac{n\pi}{b}\cos\frac{m\pi x}{a}\sin\frac{n\pi y}{b}\hat{T}_{m,n}(z),$$

$$E_y = -\frac{m\pi}{a}\sin\frac{m\pi x}{a}\cos\frac{n\pi y}{b}\hat{T}_{m,n}(z),$$

$$H_x = -Y_{m,n}^+E_y, \quad H_y = Y_{m,n}^+E_x, \quad Y_{m,n}^+ = -\frac{1}{i\omega\mu\hat{T}_{m,n}}\frac{d\hat{T}_{m,n}}{dz},$$

$$H_z = \frac{\chi_{m,n}^2}{i\omega\mu}\cos\frac{m\pi x}{a}\cos\frac{n\pi y}{b}\hat{T}_{m,n}(z).$$

10.6. *Natural Waves in Circular Wave Guides*

In a circular wave guide solutions of (1–11) in cylindrical coordinates are

$$T(\rho,\varphi) = J_n(\chi\rho)\cos n\varphi, \quad J_n(\chi\rho)\sin n\varphi. \quad (6\text{–}1)$$

The Bessel function $N_n(\chi\rho)$ becomes infinite when $\rho = 0$ and is therefore not an admissible solution when the wave guide is hollow. The boundary condition for transverse magnetic waves is now $J_n(\chi a) = 0$, where a is the radius of the tube. If $\chi a = k_{n,m}$ is the mth nonvanishing root of this equation, then the values of $k_{n,m}$ for small values of n and m are

$$\begin{aligned} &k_{0,1} = 2.40, \quad k_{0,2} = 5.52, \quad k_{0,3} = 8.65, \quad k_{0,4} = 11.79, \cdots \\ &k_{1,1} = 3.83, \quad k_{1,2} = 7.02, \quad k_{1,3} = 10.17, \quad k_{1,4} = 13.32, \cdots \\ &k_{2,1} = 5.14, \quad k_{2,2} = 8.42, \quad k_{2,3} = 11.62, \quad k_{2,4} = 14.80, \cdots \\ &k_{3,1} = 6.38, \quad k_{3,2} = 9.76, \quad k_{3,3} = 13.02, \quad k_{3,4} = 16.22, \cdots . \end{aligned} \quad (6\text{–}2)$$

The boundary condition for transverse electric waves is $J_n'(\chi a) = 0$. In this case if $\chi a = k_{n,m}$ is the mth nonvanishing root of this equation then the first few k values are

$$\begin{aligned} k_{0,1} &= 3.83, & k_{0,2} &= 7.02, & k_{0,3} &= 10.17, \cdots \\ k_{1,1} &= 1.84, & k_{1,2} &= 5.33, & k_{1,3} &= 8.54, \cdots \\ k_{2,1} &= 3.05, & k_{2,2} &= 6.71, & k_{2,3} &= 9.97, \cdots \\ k_{3,1} &= 4.20, & k_{3,2} &= 8.02, & k_{3,3} &= 11.35, \cdots. \end{aligned} \qquad (6\text{–}3)$$

The cut-off wavelength is

$$\lambda_c = \frac{2\pi a}{k}. \qquad (6\text{–}4)$$

Inspecting (2) and (3) we find that there is one transverse electric wave, the $TE_{1,1}$-wave, which has a lower cut-off frequency than any transverse magnetic wave; this $TE_{1,1}$-wave is the dominant wave in a circular wave guide.

Substituting from (1) into the general formula for the attenuation constant, we obtain for transverse magnetic waves

$$\alpha = \frac{\mathcal{R}}{\eta a}(1 - \nu_{n,m}^2)^{-1/2}. \qquad (6\text{–}5)$$

Similarly the attenuation constant for transverse electric waves is

$$\alpha = \frac{\mathcal{R}}{\eta a}\left(\frac{n^2}{k^2 - n^2} + \nu_{n,m}^2\right)(1 - \nu_{n,m}^2)^{-1/2}. \qquad (6\text{–}6)$$

For circular electric waves $n = 0$ and α becomes

$$\alpha = \frac{\mathcal{R}}{\eta a}\nu_{0,m}^2(1 - \nu_{0,m}^2)^{-1/2}. \qquad (6\text{–}7)$$

As the frequency increases the attenuation constant of circular electric waves approaches zero.

10.7. *Natural Waves between Coaxial Cylinders*

When we consider waves between two coaxial cylinders we have no reason for excluding the second Bessel function and the value of T to be considered is

$$T(\rho,\varphi) = [PJ_n(\chi\rho) + QN_n(\chi\rho)] \cos n\varphi. \qquad (7\text{–}1)$$

For transverse magnetic waves the boundary condition is

$$T(a,\varphi) = T(b,\varphi) = 0, \qquad (7\text{–}2)$$

where a is the radius of the inner conductor and b the radius of the outer. Hence

$$-\frac{Q}{P} = \frac{J_n(\chi a)}{N_n(\chi a)} = \frac{J_n(\chi b)}{N_n(\chi b)}. \qquad (7\text{–}3)$$

For transverse electric waves the boundary condition is

$$\frac{\partial T(a,\varphi)}{\partial a} = \frac{\partial T(b,\varphi)}{\partial b} = 0, \quad -\frac{Q}{P} = \frac{J_n'(\chi a)}{N_n'(\chi a)} = \frac{J_n'(\chi b)}{N_n'(\chi b)}. \qquad (7\text{–}4)$$

The smallest roots of these equations can be determined graphically. For the larger roots there exist asymptotic formulae.

When the radii are nearly equal, the curvature effect on the field distribution is small and we have approximately

$$T(\rho,\varphi) = \left[P \cos \frac{m\pi(\rho - a)}{b - a} + Q \sin \frac{m\pi(\rho - a)}{b - a}\right] \cos \frac{ns}{c},$$

where c is the mean radius and s is the distance measured along the circle of this radius. For transverse magnetic waves T should vanish at $\rho = a$ and P must be zero. For transverse electric waves $\partial T/\partial\rho$ should vanish at $\rho = a$ and hence Q must be zero. The constant χ becomes approximately

$$\chi^2 = \frac{m^2\pi^2}{(b - a)^2} + \frac{n^2}{c^2}.$$

In the case of transverse magnetic waves $m = 0$ does not lead to a nonvanishing solution but for transverse electric waves we may have

$$\chi^2 = \frac{n^2}{c^2}, \quad \chi = \frac{n}{c}.$$

The cut-off wavelength for these waves is

$$\lambda_c = \frac{2\pi c}{n}, \quad n = 1,2, \cdots.$$

The longest cut-off wavelength is equal to the average circumference of the coaxial pair $\lambda_c = 2\pi c$. This simple approximation is fairly good even when $a = 0$, when the approximate expression reduces to $\lambda_c = \pi b = 2\pi b/2$, while the exact expression is $\lambda_c = \dfrac{2\pi b}{1.84}$.

Figure 8.51 shows the exact value of the ratio of the mean circumference to the wavelength as a function of the ratio of the radii.

10.8. *Wave Guides of Miscellaneous Cross-Sections*

Exact solutions for the fields and the cut-off frequencies can be found for all wave guides formed by two coaxial cylinders and two axial planes

$$\rho = a, \quad \rho = b; \quad \varphi = 0, \quad \varphi = \psi.$$

For transverse magnetic waves $T(\rho,\varphi)$ should vanish at $\varphi = 0$; hence T is proportional to $\sin p\varphi$, and we write

$$T(\rho,\varphi) = [PJ_p(\chi\rho) + QN_p(\chi\rho)] \sin p\varphi.$$

Since T vanishes for $\varphi = \psi$, we have

$$p\psi = n\pi, \quad p = \frac{n\pi}{\psi}, \quad n = 1,2,3 \cdots. \tag{8-1}$$

In order to satisfy the boundary conditions at the cylindrical surfaces, χ must be a root of (7–3) with p in place of n. If $a = 0$, then $J_p(\chi b) = 0$.

Similarly for transverse electric waves we have

$$T(\rho,\varphi) = [PJ_p(\chi\rho) + QN_p(\chi\rho)] \cos p\varphi,$$

where p is defined by (1) and χ is a root of (7–4) with p in place of n. If $a = 0$, then $J'_p(\chi b) = 0$.

If $a = 0$ and $\psi = 2\pi$, the wave guide becomes a hollow metal tube with a metal baffle along an axial half-plane; the cross-section of this guide is shown in Fig. 10.3.

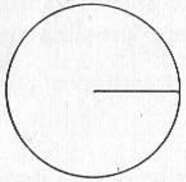


FIG. 10.3. Cross-section of a metal tube with a radial baffle.

In this case $p = n/2$. When $p = \frac{1}{2}$, the boundary equation becomes $\sin \chi a = 0$, where a is now the radius of the metal tube; hence

$$\chi a = m\pi, \quad \lambda_c = \frac{2a}{m}.$$

For transverse electric waves the equation for χ becomes

$$\frac{d}{d(\chi a)} \frac{\sin \chi a}{\sqrt{\chi a}} = 0, \quad \frac{\tan \chi a}{\chi a} = 2.$$

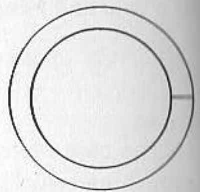


FIG. 10.4. Coaxial cylinders with a radial baffle.

The smallest root of this equation is $\chi a = 1.16$ and the larger roots are given approximately by

$$\chi a = (m + \tfrac{1}{2})\pi - \frac{1}{(2m+1)\pi}.$$

Thus the baffle has increased the cut-off frequency of the dominant transverse magnetic wave and decreased the cut-off for the dominant transverse electric mode.

If $a \neq 0$, but $\psi = 2\pi$, then the wave guide becomes a coaxial pair with a connecting baffle. When a and b are nearly equal, this wave guide is approximately a rectangular wave guide bent into the shape whose cross-section is shown in Fig. 10.4. In this case the equation for χ is approximately

$$\chi^2 = \frac{m^2\pi^2}{(b-a)^2} + \frac{n^2}{4c^2},$$

where c is the mean radius. For the dominant wave the value of χ is $\chi = 1/2c$, $\lambda_c = 4\pi c$.

Another case for which an exact solution is possible is the case of a wave guide whose cross-section is an equilateral triangle. We shall summarize the results without going into the details of their derivation. Let a be the length of the side of the equilateral triangle and b the radius of the inscribed circle, $a = 2b\sqrt{3}$, $b = a/2\sqrt{3}$. Let A, B and C be the vertices of the triangle and O its center (Fig. 10.5) and let α, β, γ be respectively the angles made by AO, BO and CO with the x-axis of a cartesian system whose origin is at O. We introduce the following homogeneous trilinear coordinates

$$u = x\cos\alpha + y\sin\alpha,$$

$$v = x\cos\beta + y\sin\beta, \quad \beta = \alpha + \frac{2\pi}{3},$$

$$w = x\cos\gamma + y\sin\gamma, \quad \gamma = \beta + \frac{2\pi}{3}.$$

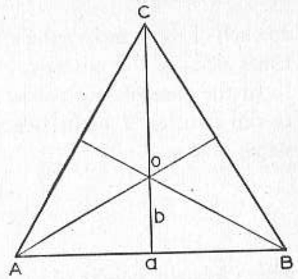


FIG. 10.5. Equilateral triangle representing the normal cross-section of a metal tube.

The line u = constant is perpendicular to AO and hence parallel to BC; similarly v = constant and w = constant are parallel to the remaining sides. If we draw lines parallel to the sides of the triangle through some point (x,y), then u, v and w are respectively the distances from O to these lines. The equations of the sides of the triangle are then $u = b$, $v = b$, $w = b$.

If now l, m, n are three integers whose sum is equal to zero, $l + m + n = 0$, then for transverse magnetic waves the T-function is

$$\begin{aligned} T(x,y) = {} & \sin\frac{2\pi l}{3b}\left(\frac{u}{2}+b\right)\cos\frac{\pi(m-n)(v-w)}{9b} \\ & + \sin\frac{2\pi m}{3b}\left(\frac{u}{2}+b\right)\cos\frac{\pi(n-l)(v-w)}{9b} \\ & + \sin\frac{2\pi n}{3b}\left(\frac{u}{2}+b\right)\cos\frac{\pi(l-m)(v-w)}{9b}. \end{aligned}$$

For transverse electric waves the T-function is

$$\begin{aligned} T(x,y) = {} & \cos\frac{2\pi l}{3b}\left(\frac{u}{2}+b\right)\cos\frac{\pi(m-n)(v-w)}{9b} \\ & + \cos\frac{2\pi m}{3b}\left(\frac{u}{2}+b\right)\cos\frac{\pi(n-l)(v-w)}{9b} \\ & + \cos\frac{2\pi n}{3b}\left(\frac{u}{2}+b\right)\cos\frac{\pi(l-m)(v-w)}{9b}. \end{aligned}$$

Substituting either of these values of T in (4–5), we find

$$\chi = \frac{2\pi}{3b\sqrt{3}}\sqrt{m^2+mn+n^2} = \frac{4\pi}{3a}\sqrt{m^2+mn+n^2}.$$

For transverse electric waves the dominant modes are obtained from one of the sets of values

$$m = 1, \quad n = 0, \quad l = -1;$$
$$m = 0, \quad n = 1, \quad l = -1;$$
$$m = 1, \quad n = -1, \quad l = 0;$$

then

$$\chi = \frac{4\pi}{3a}, \quad \lambda_c = 1.5a.$$

In each of these modes the electric lines are curves roughly perpendicular to one of the three sides of the triangle.

In the case of transverse magnetic waves a zero value for any integer results in a trivial solution $T = 0$; hence the dominant mode corresponds to $m = n = 1$ and the value of χ is

$$\chi = \frac{4\pi}{a\sqrt{3}} = \frac{2\pi}{3b}, \quad \lambda_c = 3b = h,$$

where h is the height of the triangle.

Complete sets of transmission modes can also be found for elliptical wave guides and for wave guides whose cross-sections are isosceles right-angled triangles. In the latter case the T-functions are found among T-functions appropriate to square wave guides; in the former case they are solutions of the Mathieu equation. But in general no convenient exact method exists of solving the problem for wave guides of other than the above-mentioned cross-sections. Some solutions can be found when the cross-section is a regular hexagon and it is always possible to construct isolated solutions by taking some function T which satisfies (1–11) and then determining boundaries which are consistent with this function; thus (4–1) represents the equation of the boundary for transverse magnetic waves and (4–2) the corresponding equation for transverse electric waves. For example we might start with the following function

$$T = J_0(2.40\rho) + 0.63 J_2(2.40\rho) \cos 2\varphi$$

and plot its contour lines which are also magnetic lines of force. The line $T = 0$ can be taken as the boundary of the wave guide supporting the transverse magnetic wave whose field is determined by T. For transverse electric waves the same contour lines are electric lines of force; in this case the boundary should intersect the contour lines at right angles.

Figures 10.6 and 10.7 show E-lines for transverse electric waves and H-lines for transverse magnetic waves for guides with different cross-sections. The arrows along the boundaries indicate the directions of the transverse conduction currents in relation to the directions of the transverse displacement currents; dots and crosses indicate longitudinal currents flowing respectively toward the reader and away from him. The

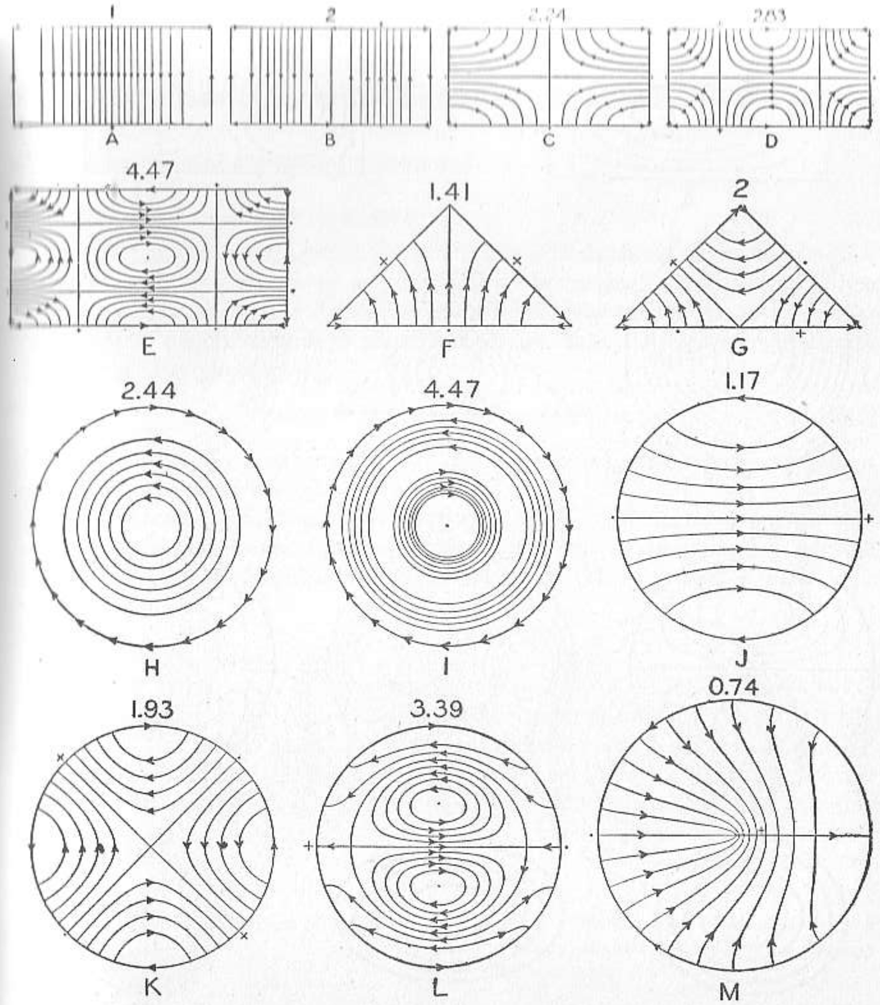


FIG. 10.6. The E-lines for transverse electric waves. The density of the lines indicates the amplitudes of the transverse field components. The transverse component of H is perpendicular to the E-lines; if the E-vector is turned through 90 degrees in the counterclockwise direction so as to coincide with the transverse H-component, then a right-handed screw turned in the same way will advance in the direction of the wave motion. The numbers above the figures refer to the relative cut-off frequencies on the assumption that the largest linear dimensions are the same for all figures. The cut-off wavelength of the transmission mode shown in (A) is twice the longer side. The following transmission modes are shown:

(A) $TE_{1,0}$-mode, $T = \cos \pi x$,
(B) $TE_{2,0}$-mode, $T = \cos 2\pi x$,
(C) $TE_{1,1}$-mode, $T = \cos \pi x \cos 2\pi y$,
(D) $TE_{2,1}$-mode, $T = \cos 2\pi x \cos 2\pi y$,
(E) $TE_{2,2}$-mode, $T = \cos 2\pi x \cos 4\pi y$,
(F) $T = \frac{1}{2}(\cos \pi x - \cos \pi y)$,
(G) $T = \cos \pi x \cos \pi y$,
(H) $TE_{0,1}$-mode, $T = \frac{J_0(3.83\rho)}{J_0(3.83)}$, the boundary is $\rho = 1$,
(I) $TE_{0,2}$-mode, $T = \frac{J_0(7.02\rho)}{J_0(7.02)}$,
(J) $TE_{1,1}$-mode, $T = \frac{J_1(1.84\rho)}{J_1(1.84)} \cos \phi$,
(K) $TE_{2,1}$-mode, $T = \frac{J_2(3.04\rho)}{J_2(3.04)} \cos 2\phi$,
(L) $TE_{1,2}$-mode, $T = \frac{J_1(5.33\rho)}{J_1(1.84)} \cos \phi$,
(M) $TE_{\frac{1}{2},1}$-mode, $T = \frac{J_{\frac{1}{2}}(1.16\rho)}{J_{\frac{1}{2}}(1.16)} \cos \frac{\phi}{2}$.

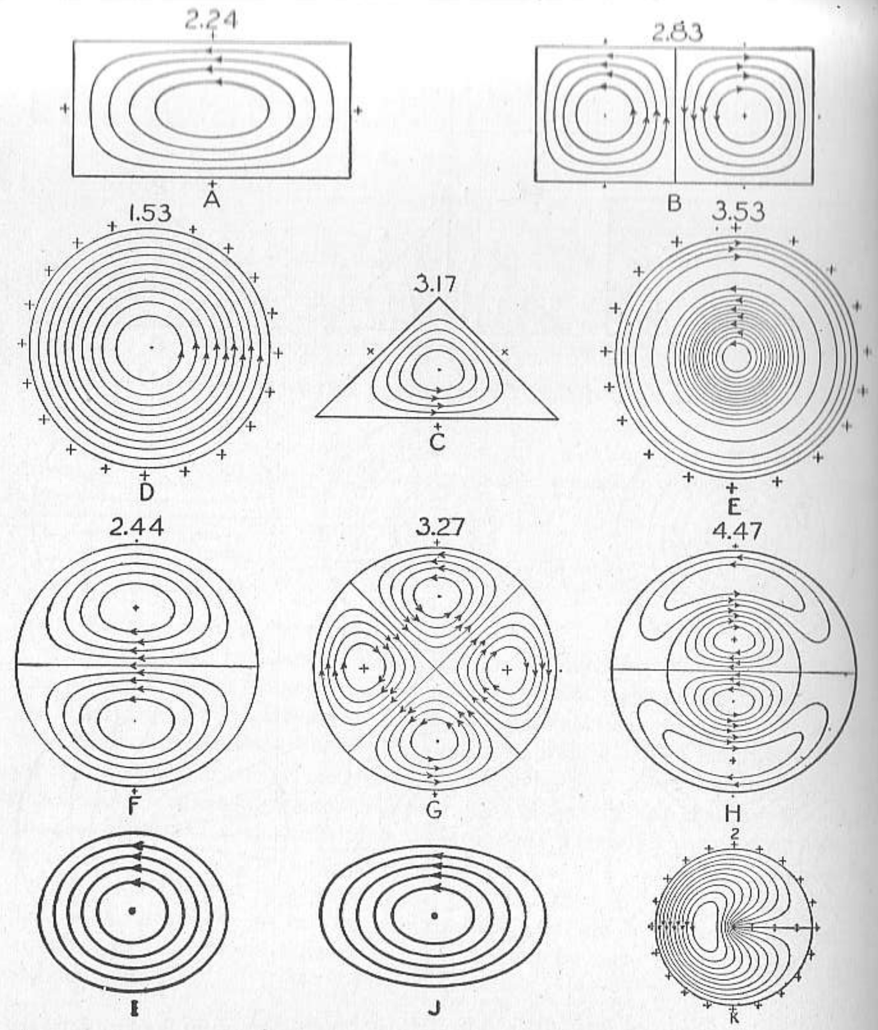


FIG. 10.7. The H-lines for transverse magnetic waves. The density of the lines indicates the amplitudes of the transverse field components. The transverse component of E is perpendicular to the H-lines. The numbers above the figures represent the cut-off frequencies in terms of the cut-off for the transmission mode shown in Fig. 10.6A. The following transmission modes are shown:

(A) $TM_{1,1}$-mode, $T = \sin \pi x \sin 2\pi y$,

(B) $TM_{2,1}$-mode, $T = \sin 2\pi x \sin 2\pi y$,

(C) $T = 2.60 \sin \pi x \sin \pi y \cos \frac{\pi(x-y)}{2} \cos \frac{\pi(x+y)}{2}$ $= 0.65 (\sin 2\pi x \sin \pi y + \sin \pi x \sin 2\pi y)$,

(D) $TM_{0,1}$-mode, $T = J_0(2.40\rho)$, the boundary is $\rho = 1$,

(E) $TM_{0,2}$-mode, $T = J_0(5.52\rho)$,

(F) $TM_{1,1}$-mode, $T = \frac{J_1(3.83\rho)}{J_1(1.84)} \cos \phi$,

(G) $TM_{2,1}$-mode, $T = \frac{J_2(5.14\rho)}{J_2(3.04)} \cos 2\phi$,

(H) $TM_{1,2}$-mode, $T = \frac{J_1(7.02\rho)}{J_1(1.84)} \cos \phi$,

(I) $TM_{0,1}$-mode, $T = J_0(2.40\rho) + 1.35 J_1(2.40\rho) \cos \phi$, the equation of the boundary is $T = 0$,

(J) $TM_{0,1}$-mode, $T = J_0(2.40\rho) + 0.63 J_2(2.40\rho) \cos 2\phi$, the equation of the boundary is $T = 0$,

(K) $TM_{\frac{1}{2},1}$-mode, $T = \frac{J_{\frac{1}{2}}(\pi\rho)}{J_{\frac{1}{2}}(1.16)} \sin \frac{\phi}{2}$.

relative directions as shown in the figures correspond to waves moving toward the reader; for waves moving away from the reader the directions of the longitudinal currents are reversed.

10.9. *Slightly Noncircular Wave Guides*

The effect on wave propagation of a small change in the shape of the cross-section of a wave guide can be obtained very simply by "perturbing" the field distribution function T. For example, in considering the effect of a slight compression of a circular wave guide on the propagation of waves with circular lines of force we might assume the following solution of (1–11)

$$T(\rho,\varphi) = J_0(\chi\rho) + u J_2(\chi\rho) \cos 2\varphi, \qquad (9\text{–}1)$$

where u is small and χ is either a zero of $J_0(x)$ (circular magnetic waves) or a zero of $J_1(x)$ (circular electric waves).

Inasmuch as magnetic lines do not penetrate a perfect conductor we expect that for circular magnetic waves the magnetic lines will simply be compressed to conform to the compression of the guide itself; no more radical change is likely to occur. In

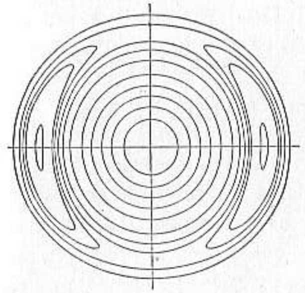


FIG. 10.8. Electric lines of force.

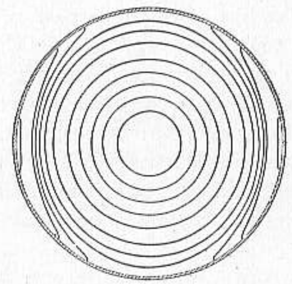


FIG. 10.9. Electric lines associated with a "circular electric wave" in a slightly flattened tube.

the case of circular electric waves, however, the compressed conductor will touch electric lines of nonzero intensity; these lines will break and turn toward the conductor. Thus the appearance of the electric lines in the neighborhood of the conductor may be radically changed even though the quantitative effect of the compression may be small. For example, assuming $\chi = 3.83$ and $u = 0.1$, we obtain Fig. 10.8 for the electric lines of force; inserting a perfect conductor at right angles to these lines we have Fig. 10.9.

Since dT vanishes along a contour line the equation of these lines is

$$\frac{d\varphi}{d\rho} = \frac{\chi[-J_1(\chi\rho) + u J_2'(\chi\rho) \cos 2\varphi]}{2u J_2(\chi\rho) \sin 2\varphi}. \qquad (9\text{–}2)$$

When $u = 0$, $d\varphi/d\rho$ is infinite, the length of the radius drawn from the origin to a point on a particular contour line is independent of φ, and the contour lines are circles coaxial with the origin. When u is small compared with unity $d\varphi/d\rho$ is *usually* large and the contour lines are approximately circular; but in the vicinity of the circle $\rho = 1$, the numerator is small and may even vanish. The vanishing of $d\varphi/d\rho$ marks the points at which the level lines are tangential to the radii and the locus of these points is

$$-J_1(\chi\rho) + uJ_2'(\chi\rho)\cos 2\varphi = 0. \tag{9-3}$$

Let $\rho = 1 + \delta$; then

$$J_1(\chi\rho) = J_1(\chi + \chi\delta) = \chi J_1'(\chi)\delta = \chi J_0(\chi)\delta,$$
$$J_2'(\chi\rho) = J_2'(\chi) = -0.5J_3(\chi). \tag{9-4}$$

Substituting in (3), we have

$$\delta = -\frac{uJ_3(\chi)}{2\chi J_0(\chi)}\cos 2\varphi = 0.136u\cos 2\varphi. \tag{9-5}$$

The numerical constant corresponds to the special case $\chi = 3.83$. The curve (3) crosses the unit circle when $\varphi = \pm 45°, \pm 135°$ and it represents an oval which can be described as a circle slightly flattened in the neighborhood of $\varphi = \pm 90°$. The boundary of the metal tube will approximately coincide with the curve (3). The largest radius corresponds to $\varphi = 0$ and is $1 + 0.136u$; the smallest radius corresponds to $\varphi = 90°$ and is $1 - 0.136u$. The ratio of the largest diameter of the tube to the smallest exceeds unity by an amount $\Delta = 0.272u$. This quantity may be chosen to represent the degree of departure from perfect circularity.

We have seen that the attenuation constant of circular electric waves diminishes with increasing frequency. This happens because $\partial T/\partial s$ vanishes along the boundary of the tube and in the general attenuation formula (4–17) only the second term remains. For a deformed wave guide, however, we have approximately

$$\frac{\partial T}{\partial\varphi} = -2uJ_2(\chi)\sin 2\varphi, \quad \frac{\partial T}{\partial\rho} = 0. \tag{9-6}$$

Thus at the surface of the metal tube $\partial T/\partial s$ is proportional to u and the first term in (4–17) is nearly proportional to u^2. For small departures from perfect circularity the attenuation constant is dominated by the second term of (4–17) in the neighborhood of the cut-off frequency, and as the tube becomes more nearly circular the second term remains dominant within an increasingly large frequency range. In this range the attenuation constant diminishes with increasing frequency.

Substituting from (6) in (4–17), we obtain the following approximate formula for the attenuation constant in a deformed circular wave guide*

$$\alpha = \frac{\mathcal{R}}{\eta}[\nu^2(1-\nu^2)^{-1/2} + p(1-\nu^2)^{1/2}] \quad p = \frac{2[J_2(3.83)]^2\Delta^2}{[J_3(3.83)]^2} = 1.835\Delta^2.$$

Minimizing α with respect to the frequency, we find that, to the same order of approximation, the minimum occurs when $\nu = \sqrt{p/3} = 0.782\Delta$. If for example the largest diameter exceeds the smallest by one per cent, the minimum occurs at a frequency which is 128 times as great as the cut-off.

* The coefficient of the first term in the brackets is also affected by the departure from perfect circularity and a term proportional to Δ^2 should be added to unity. However, this term is of less importance than the correction term involving p.

10.10. *Transverse Magnetic Spherical Waves*

The theory of spherical waves is in all respects similar to the theory of plane waves. Thus there exist transverse magnetic waves for which $H_r = 0$ and transverse electric waves for which $E_r = 0$. In the former case the divergence equation for H becomes

$$\frac{\partial}{\partial\theta}(\sin\theta\, H_\theta) + \frac{\partial H_\varphi}{\partial\varphi} = 0; \tag{10-1}$$

consequently the magnetic intensity may be expressed in terms of a stream function Π

$$r\sin\theta\, H_\theta = \frac{\partial\Pi}{\partial\varphi}, \quad rH_\varphi = -\frac{\partial\Pi}{\partial\theta}. \tag{10-2}$$

Hence we have

$$H = \operatorname{curl} A, \quad A_r = \Pi, \quad A_\theta = A_\varphi = 0. \tag{10-3}$$

Since the radial magnetic current vanishes, the transverse electric intensity may be expressed as the gradient of a potential function; thus

$$rE_\theta = -\frac{\partial V}{\partial\theta}, \quad r\sin\theta\, E_\varphi = -\frac{\partial V}{\partial\varphi}. \tag{10-4}$$

In addition to (1) we have the following field equations

$$rE_\theta = -\frac{1}{g + i\omega\epsilon}\frac{\partial(rH_\varphi)}{\partial r}, \quad rE_\varphi = \frac{1}{g + i\omega\epsilon}\frac{\partial(rH_\theta)}{\partial r}, \tag{10-5}$$

$$\frac{\partial(\sin\theta\, H_\varphi)}{\partial\theta} - \frac{\partial H_\theta}{\partial\varphi} = (g + i\omega\epsilon)r\sin\theta\, E_r, \tag{10-6}$$

$$\frac{\partial E_r}{\partial\varphi} - \sin\theta\,\frac{\partial(rE_\varphi)}{\partial r} = -i\omega\mu r\sin\theta\, H_\theta, \tag{10-7}$$

$$\frac{\partial(rE_\theta)}{\partial r} - \frac{\partial E_r}{\partial\theta} = -i\omega\mu rH_\varphi. \tag{10-8}$$

Substituting from (2) and (4) in (5), we obtain

$$\frac{\partial\Pi}{\partial r} = -(g + i\omega\epsilon)V. \tag{10-9}$$

The radial electric intensity may be obtained from (6) on the one hand and from (7) or (8) on the other; thus

$$E_r = -\frac{1}{(g+i\omega\epsilon)r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Pi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\Pi}{\partial\varphi^2}\right],$$

$$E_r = -i\omega\mu\Pi - \frac{\partial V}{\partial r} = \frac{1}{g+i\omega\epsilon}\left(\frac{\partial^2\Pi}{\partial r^2} - \sigma^2\Pi\right). \qquad (10\text{–}10)$$

Comparing, we have

$$r^2\frac{\partial^2\Pi}{\partial r^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Pi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\Pi}{\partial\varphi^2} = \sigma^2 r^2\Pi. \qquad (10\text{–}11)$$

The complete field has been expressed in terms of a single scalar function Π. In the narrow sense of the term equation (11) for Π is not a wave equation; that is, it is not equation (1–5) in spherical coordinates.

We shall now consider fields for which

$$\Pi(r,\theta,\varphi) = T(\theta,\varphi)\hat{T}(r); \qquad (10\text{–}12)$$

that is, the field pattern is the same on all spheres concentric with the origin. Substituting from (12) in (11), we obtain the following equations

$$\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial T}{\partial\theta}\right) + \frac{\partial^2 T}{\partial\varphi^2} = -n(n+1)\sin^2\theta\, T, \qquad (10\text{–}13)$$

$$\frac{d^2\hat{T}}{dr^2} = \left[\sigma^2 + \frac{n(n+1)}{r^2}\right]\hat{T}, \qquad (10\text{–}14)$$

where n is a constant. For spherical waves T and therefore n must be real; in general, however, n may be complex.

Substituting from (12) and (13) in (10), we have

$$E_r = \frac{n(n+1)\Pi}{(g+i\omega\epsilon)r^2}, \quad (g+i\omega\epsilon)r^2E_r = n(n+1)\Pi; \qquad (10\text{–}15)$$

thus the radial electric current per unit solid angle is a constant multiple of the wave function Π.

From (9) and (14) we have

$$\frac{\partial V}{\partial r} = -\left[i\omega\mu + \frac{n(n+1)}{(g+i\omega\epsilon)r^2}\right]\Pi, \quad \frac{\partial\Pi}{\partial r} = -(g+i\omega\epsilon)V. \qquad (10\text{–}16)$$

Hence V and Π vary with r as the voltage and current in a transmission line whose distributed series constants per unit length are an inductance μ, a capacitance $\dfrac{\epsilon r^2}{n(n+1)}$ in parallel with a conductance $\dfrac{gr^2}{n(n+1)}$, and

whose distributed shunt constants per unit length are a capacitance ϵ and a conductance g. Sufficiently far from the origin the series capacitance and conductance are very large.

When Π is of the form (12), the transverse field intensities are

$$r\sin\theta\, H_\theta = \frac{\partial T}{\partial\varphi}\hat{T}, \quad rH_\varphi = -\frac{\partial T}{\partial\theta}\hat{T},$$

$$E_\theta = Z_r^+ H_\varphi, \quad E_\varphi = -Z_r^+ H_\theta, \qquad (10\text{–}17)$$

where the radial impedance is defined by

$$Z_r^+ = -\frac{1}{(g+i\omega\epsilon)\hat{T}}\frac{d\hat{T}}{dr}. \qquad (10\text{–}18)$$

In terms of the functions defined in section 3.5, we have

$$\hat{T} = A\hat{I}_n(\sigma r) + B\hat{K}_n(\sigma r). \qquad (10\text{–}19)$$

For progressive waves traveling outward we must have $A = 0$ since $\hat{I}$ becomes infinite at $r = \infty$. For stationary waves having no singularity at the origin we must have $B = 0$. Thus the radial impedances for progressive waves and for stationary waves become respectively

$$K_r^+ = -\eta\frac{\hat{K}_n'(\sigma r)}{\hat{K}_n(\sigma r)}, \quad K_r^- = \eta\frac{\hat{I}_n'(\sigma r)}{\hat{I}_n(\sigma r)}. \qquad (10\text{–}20)$$

In nondissipative media the latter becomes

$$K_r^- = -i\eta\frac{\hat{J}_n'(\beta r)}{\hat{J}_n(\beta r)}. \qquad (10\text{–}21)$$

The T-function is a spherical harmonic and is briefly discussed in section 3.6; in general T is either a series of terms of the following type

$$T(\theta,\varphi) = [CP_n^m(\cos\theta) + DP_n^m(-\cos\theta)](P\cos m\varphi + Q\sin m\varphi), \qquad (10\text{–}22)$$

or an integral $\int T(\theta,\varphi)dm\,dn$, where C, D, P, Q are functions of m and n. In free space, however, m must be an integer, since T should return to its original value when φ changes by 2π. Furthermore the Legendre functions become infinite either at $\theta = 0$ or $\theta = \pi$ unless n is an integer; hence for free space n must be an integer. Thus in free space (22) assumes the following form

$$T(\theta,\varphi) = P_n^m(\cos\theta)(C\cos m\varphi + D\sin m\varphi), \qquad (10\text{–}23)$$

where m and n are positive integers. When $m > n$, P_n^m vanishes; hence we may assume $m \leq n$.

The following are the Legendre functions of low order

$$P_0(\cos\theta) = 1, \quad P_1(\cos\theta) = \cos\theta, \quad P_1^1(\cos\theta) = -\sin\theta,$$
$$P_2(\cos\theta) = \tfrac{1}{2}(3\cos^2\theta - 1), \qquad P_2^1(\cos\theta) = -3\sin\theta\cos\theta,$$
$$P_2^2(\cos\theta) = 3\sin^2\theta,$$
$$P_3(\cos\theta) = \tfrac{1}{2}(5\cos^3\theta - 3\cos\theta), \qquad P_3^1(\cos\theta) = -\tfrac{3}{2}\sin\theta(5\cos^2\theta - 1),$$
$$P_3^2(\cos\theta) = 15\sin^2\theta\cos\theta, \qquad P_3^3(\cos\theta) = -15\sin^3\theta, \tag{10–24}$$
$$P_4(\cos\theta) = \tfrac{1}{8}(35\cos^4\theta - 30\cos^2\theta + 3),$$
$$P_4^1(\cos\theta) = -\tfrac{5}{2}\sin\theta(7\cos^3\theta - 3\cos\theta),$$
$$P_4^2(\cos\theta) = \tfrac{15}{2}\sin^2\theta(7\cos^2\theta - 1), \quad P_4^3(\cos\theta) = -105\sin^3\theta\cos\theta,$$
$$P_4^4(\cos\theta) = 105\sin^4\theta.$$

In a region bounded by a perfectly conducting cone $\theta = \psi$, (22) becomes (for $\theta < \psi$)

$$T(\theta,\varphi) = P_n^m(\cos\theta)(P\cos m\varphi + Q\sin m\varphi) \tag{10–25}$$

since the second Legendre function is infinite on the radius $\theta = 0$. In this case m is still an integer, but n is not necessarily integral. Since E_r is proportional to T and must vanish when $\theta = \psi$, n must be a root of the following equation

$$P_n^m(\cos\psi) = 0. \tag{10–26}$$

In a region bounded by two perfectly conducting cones E_r and hence T must vanish on the boundaries $\theta = \theta_1, \theta = \theta_2$; therefore n must be a root of

$$\frac{P_n^m(\cos\theta_1)}{P_n^m(-\cos\theta_1)} = \frac{P_n^m(\cos\theta_2)}{P_n^m(-\cos\theta_2)}. \tag{10–27}$$

If the region is bounded also by two perfectly conducting axial half-planes $\varphi = 0$ and $\varphi = \varphi_0$, then P must vanish in (22) and

$$m = \frac{s\pi}{\varphi_0}, \quad s = 1,2,3,\cdots. \tag{10–28}$$

If the region is bounded by two perfectly conducting spheres $r = a$ and $r = b$, then Z_r^+ and hence $d\hat{T}/dr$ must vanish at these spheres; thus we have

$$-\frac{B}{A} = \frac{\hat{I}_n'(\sigma a)}{\hat{K}_n'(\sigma a)} = \frac{\hat{I}_n'(\sigma b)}{\hat{K}_n'(\sigma b)}, \quad \frac{\hat{J}_n'(\beta a)}{\hat{N}_n'(\beta a)} = \frac{\hat{J}_n'(\beta b)}{\hat{N}_n'(\beta b)}. \tag{10–29}$$

The first equation applies to dissipative media and the second to non-dissipative media.

The average flow of power across a portion (S) of an equiphase surface is

$$\begin{aligned}\Psi &= \tfrac{1}{2}r^2 \iint_{(S)} (E_\theta H_\varphi^* - E_\varphi H_\theta^*)\, d\Omega \\ &= \tfrac{1}{2}Z_r^+\hat{T}\hat{T}^* \iint_{(S)} \left[\left(\frac{\partial T}{\partial\theta}\right)^2 + \frac{1}{\sin^2\theta}\left(\frac{\partial T}{\partial\varphi}\right)^2\right] d\Omega.\end{aligned} \tag{10–30}$$

Multiplying (13) by $T\,d\theta\,d\varphi/\sin\theta$ and integrating over (S), we obtain the following identity for a wave in free space, or for a wave bounded by perfectly conducting conical surfaces including planes emerging from the center of the wave

$$\iint_{(S)} \left[\left(\frac{\partial T}{\partial\theta}\right)^2 + \frac{1}{\sin^2\theta}\left(\frac{\partial T}{\partial\varphi}\right)^2\right] d\Omega = n(n+1)\iint_{(S)} T^2\, d\Omega. \tag{10–31}$$

Hence (30) becomes

$$\Psi = \tfrac{1}{2}n(n+1)Z_r^+\hat{T}\hat{T}^* \iint_{(S)} T^2\, d\Omega. \tag{10–32}$$

Let I be the root mean square radial electric current

$$I = S\sqrt{\frac{1}{S}\iint_{(S)} J_r J_r^*\, dS} = S\sqrt{\frac{n^2(n+1)^2}{r^2 S}\iint_{(S)} \Pi\Pi^*\, d\Omega}; \tag{10–33}$$

then (32) becomes

$$\Psi = \frac{r^2 Z_r^+ I^2}{2n(n+1)S} = \frac{Z_r^+ I^2}{2n(n+1)\Omega}. \tag{10–34}$$

10.11. *Transverse Electric Spherical Waves*

Using the same method as in the preceding section, we obtain the following expressions for transverse electric spherical waves

$$\begin{aligned}
E &= -\operatorname{curl} F, \qquad F_r = \Psi, \quad F_\theta = F_\varphi = 0, \\
H_t &= -\operatorname{grad}_t U, \qquad U = -\frac{1}{i\omega\mu}\frac{\partial\Psi}{\partial r}, \\
r\sin\theta\, E_\theta &= -\frac{\partial\Psi}{\partial\varphi}, \qquad rE_\varphi = \frac{\partial\Psi}{\partial\theta}, \\
rH_\theta &= -\frac{\partial U}{\partial\theta}, \quad r\sin\theta\, H_\varphi = -\frac{\partial U}{\partial\varphi}, \\
H_r &= -\frac{1}{i\omega\mu r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Psi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\Psi}{\partial\varphi^2}\right] \\
&= -(g + i\omega\epsilon)\Psi - \frac{\partial U}{\partial r} = \frac{1}{i\omega\mu}\left(\frac{\partial^2\Psi}{\partial r^2} - \sigma^2\Psi\right).
\end{aligned} \tag{11–1}$$

The stream function Ψ satisfies equation (10–11).

When $\Psi = T(\theta,\varphi)\hat{T}(r)$, the field pattern in spherical surfaces concentric with the origin is independent of r. In this case the expressions for the field become

$$r \sin\theta\, E_\theta = -\frac{\partial T}{\partial \varphi}\hat{T}, \quad rE_\varphi = \frac{\partial T}{\partial \theta}\hat{T},$$

$$H_\theta = -Y_r^+ E_\varphi, \quad H_\varphi = Y_r^+ E_\theta, \quad Y_r^+ = \frac{1}{Z_r^+} = -\frac{1}{i\omega\mu\hat{T}}\frac{d\hat{T}}{dr}, \qquad (11\text{–}2)$$

$$H_r = \frac{n(n+1)}{i\omega\mu r^2}\Psi, \quad i\omega\mu r^2 H_r = n(n+1)\Psi.$$

T and $\hat{T}$ satisfy equations (10–13) and (10–14). The magnetic potential U and the electric stream function Ψ satisfy

$$\frac{\partial \Psi}{\partial r} = -i\omega\mu U, \quad \frac{\partial U}{\partial r} = -\left[(g + i\omega\epsilon) + \frac{n(n+1)}{i\omega\mu r^2}\right]\Psi. \qquad (11\text{–}3)$$

Thus Ψ and U vary with r as the voltage and current in a transmission line with series inductance μ, shunt conductance g, shunt capacitance ϵ, and shunt inductance $\frac{\mu r^2}{n(n+1)}$ all per unit length. Sufficiently far from the origin, the shunt inductance becomes very large.

The radial impedances for progressive waves and for stationary waves without singularities at the origin are respectively

$$K_r^+ = -\eta\frac{\hat{K}_n(\sigma r)}{\hat{K}_n'(\sigma r)}, \quad K_r^- = \eta\frac{\hat{I}_n(\sigma r)}{\hat{I}_n'(\sigma r)}. \qquad (11\text{–}4)$$

In nondissipative media the second impedance becomes

$$K_r^- = i\eta\frac{\hat{J}_n(\beta r)}{\hat{J}_n'(\beta r)}. \qquad (11\text{–}5)$$

Comparing with (10–20) and (10–21), we find that the product of the corresponding impedances for transverse electric and transverse magnetic waves is equal to the square of the intrinsic impedance.

The T-function is, in general, a series of terms of the form (10–22) or it may be expressed in the form of an integral. In free space m and n are integers and $m \le n$. In a region $0 \le \theta \le \psi$ bounded by a perfectly conducting cone $\theta = \psi$, E_φ and therefore $\partial T/\partial\theta$ must vanish on the boundary; thus n must be a root of the following equation

$$\frac{d}{d\psi}P_n^m(\cos\psi) = 0. \qquad (11\text{–}6)$$

For a region between two perfectly conducting cones $\theta = \theta_1$ and $\theta = \theta_2$, n must satisfy

$$\frac{\frac{d}{d\theta_1}P_n^m(\cos\theta_1)}{\frac{d}{d\theta_1}P_n^m(-\cos\theta_1)} = \frac{\frac{d}{d\theta_2}P_n^m(\cos\theta_2)}{\frac{d}{d\theta_2}P_n^m(-\cos\theta_2)}. \qquad (11\text{–}7)$$

If perfectly conducting half-planes $\varphi = 0$ and $\varphi = \varphi_0$ are added to the cones, m need not be an integer. In this case E_θ and hence $\partial T/\partial\varphi$ must vanish at the planes; therefore Q in (10–22) should vanish and m is given by

$$m = \frac{s\pi}{\varphi_0}, \quad s = 0,1,2,3,\cdots. \qquad (11\text{–}8)$$

Finally, if perfectly conducting spheres $r = a$ and $r = b$ are added, then $\hat{T}$ should vanish on the spheres and

$$-\frac{B}{A} = \frac{\hat{I}_n(\sigma a)}{\hat{K}_n(\sigma a)} = \frac{\hat{I}_n(\sigma b)}{\hat{K}_n(\sigma b)}, \quad \frac{\hat{J}_n(\beta a)}{\hat{N}_n(\beta a)} = \frac{\hat{J}_n(\beta b)}{\hat{N}_n(\beta b)}. \qquad (11\text{–}9)$$

The complex power flow across an equiphase sphere in free space, or across a portion of such a sphere bounded by perfectly conducting cones, can be found as in the preceding section; thus we have

$$\Psi = \frac{n(n+1)\hat{T}\hat{T}^*}{2Z_r^*}\iint_{(S)} T^2\, d\Omega. \qquad (11\text{–}10)$$

In terms of the root mean square radial magnetic current K defined by

$$K = S\sqrt{\frac{1}{S}\iint_{(S)}\omega^2\mu^2 H_r H_r^*\, dS}, \qquad (11\text{–}11)$$

the expression for Ψ is

$$\Psi = \frac{r^2K^2}{2n(n+1)SZ_r^*} = \frac{K^2}{2n(n+1)\Omega Z_r^*}. \qquad (11\text{–}12)$$

10.12. *Wave Guides of Variable Cross-Section*

Approximate equations for wave guides in which the cross-section varies in size but not in shape can be obtained as follows. In the case of transverse magnetic waves V and Π satisfy equations (1–23) and χ is given by (4–7) where the modular constant k depends upon the particular transmission mode and on the shape of the cross-section but not on its size. Thus we can write (1–23) in the following form

$$\frac{\partial V}{\partial s} = -\left[i\omega\mu + \frac{\pi k^2}{(g + i\omega\epsilon)S}\right]\Pi, \quad \frac{\partial \Pi}{\partial s} = -(g + i\omega\epsilon)V, \qquad (12\text{–}1)$$

where S is the area of the cross-section and s is the distance measured along wave normals. We now assume that these equations are approximately true when S is a slowly varying function of s. Similarly the equations for transverse electric waves are

$$\frac{\partial \Psi}{\partial s} = -i\omega\mu U, \quad \frac{\partial U}{\partial s} = -\left(g + i\omega\epsilon + \frac{\pi k^2}{i\omega\mu S}\right)\Psi. \tag{12-2}$$

When the cross-section of the wave guide is circular, then $S = \pi a^2(s)$, where $a(s)$ is the variable radius. Furthermore if the wave guide is a cone, then $a(s) = s\psi$, where ψ is the angle between the axis of the cone and its generators. In this case (1) becomes (after s is replaced by r)

$$\frac{\partial V}{\partial r} = -\left[i\omega\mu + \frac{k^2}{(g + i\omega\epsilon)\psi^2 r^2}\right]\Pi, \quad \frac{\partial \Pi}{\partial r} = -(g + i\omega\epsilon)V, \tag{12-3}$$

where k is a zero of $J_m(x)$. The exact equations are (10–16) in which n is a root of (10–26). Let $m = 0$; then according to Hobson the roots of (10–26), when ψ is small, are given approximately by $n = k/(2 \sin \psi/2)$, where k is a zero of $J_0(x)$. Hence we have

$$n(n+1) \simeq n^2 = \frac{k^2}{2(1 - \cos\psi)} \simeq \frac{k^2}{\psi^2}.$$

Thus we again obtain equations (3). It is also apparent that the approximation (3) will be improved if we write $2(1 - \cos\psi)$ in place of ψ^2. This value is obtained from (1) if S is taken to be the area of the equiphase surface intercepted by the guide. This would have been a reasonable initial assumption. Equations (1) and (2) can then be solved approximately by the method of section 7.11.

10.13. *Cylindrical Waves*

By definition a cylindrical wave is a wave whose equiphase surfaces form a family of coaxial circular cylinders. The common axis of the cylinders is the axis of the wave. In general, cylindrical waves are *hybrid waves* in the sense that they are associated with nonvanishing E and H in the direction of wave propagation. If we start with an assumption $E_\rho = 0$, we shall find that E_φ must also vanish and that the field must be independent of the z-coordinate. Further investigation would show that the only transverse electric cylindrical waves are the waves in which the E-vector is parallel to the axis of the wave or the waves having circular symmetry. Similarly the only transverse magnetic cylindrical waves are the waves in which the H-vector is parallel to the axis of the wave or the waves having circular symmetry.

In section 3 we have shown that the most general electromagnetic field may be represented in terms of two scalar functions Π and Ψ satisfying the wave equation. In the case of cylindrical waves these functions should be of the following form

$$\Pi(\rho,\varphi,z) = T(\varphi,z)\hat{T}(\rho), \tag{13-1}$$

where $T(\varphi,z)$ is real. Substituting this in the wave equation, we have

$$\frac{1}{\rho\hat{T}}\frac{d}{d\rho}\left(\rho\frac{d\hat{T}}{d\rho}\right) + \frac{1}{T}\left[\frac{\partial^2 T}{\rho^2\partial\varphi^2} + \frac{\partial^2 T}{\partial z^2}\right] = \sigma^2. \tag{13-2}$$

The sum in the brackets must be a function of ρ alone; but one of the terms of this sum contains ρ and the other does not; therefore, $(1/T)(\partial^2 T/\partial z^2)$ and $(1/T)(\partial^2 T/\partial\varphi^2)$ must be constants and consequently

$$T(\varphi,z) = T(\varphi)\overline{T}(z), \quad \Pi = T(\varphi)\overline{T}(z)\hat{T}(\rho). \tag{13-3}$$

Thus the general form of Π appropriate to cylindrical waves is

$$\Pi = [AI_q(\Gamma\rho) + BK_q(\Gamma\rho)](C \cos \xi z + D \sin \xi z)(P \cos q\varphi + Q \sin q\varphi), \tag{13-4}$$

where the radial propagation constant Γ is

$$\Gamma = \sqrt{\sigma^2 + \xi^2}. \tag{13-5}$$

The field intensities are derived from Π by using equations (1–1), (1–4), and (1–6); thus

$$H_\rho = \frac{1}{\rho}\hat{T}(\rho)\overline{T}(z)\frac{dT(\varphi)}{d\varphi}, \quad H_\varphi = -T(\varphi)\overline{T}(z)\frac{d\hat{T}(\rho)}{d\rho}, \quad H_z = 0,$$
$$E_\rho = \frac{T}{g + i\omega\epsilon}\frac{d\hat{T}}{d\rho}\frac{d\overline{T}}{dz}, \quad E_\varphi = \frac{\hat{T}}{(g + i\omega\epsilon)\rho}\frac{dT}{d\varphi}\frac{d\overline{T}}{dz}, \quad E_z = -\frac{\Gamma^2\Pi}{g + i\omega\epsilon}. \tag{13-6}$$

In general E and H are elliptically polarized. The plane of the H-ellipse is always perpendicular to the axis of the wave but the plane of orientation of the E-ellipse varies from point to point; thus the wave may be called a "magnetically oriented wave." In nondissipative media Γ is either real or imaginary and consequently E_φ and E_z are in phase. In this case the plane of the E-ellipse passes through the wave normal.

In free space q must be an integer; but inside a wedge formed by two perfectly conducting half planes $\varphi = 0$ and $\varphi = \psi$, issuing from the axis of the wave, we must have $P = 0$ and

$$q = \frac{n\pi}{\psi}, \quad n = 1,2,3,\cdots. \tag{13-7}$$

The values of q are restricted as above because E_ρ and E_z and hence $T(\varphi)$ must vanish on the boundaries. In free space ξ may assume any real or imaginary value; but in a region bounded by two perfectly conducting planes $z = 0$ and $z = h$, we must have $D = 0$ and

$$\xi = \frac{m\pi}{h}, \quad m = 0,1,2,\cdots. \tag{13-8}$$

This restriction follows because E_ρ and hence $\partial T/\partial z$ must vanish at the above boundaries.

The field of electrically oriented cylindrical waves is obtained from a function Ψ

of the same general form as Π with the aid of the equations of section 2. Thus we have

$$E_\rho = -\frac{1}{\rho}\hat{T}(\rho)\overline{T}(z)\frac{dT(\varphi)}{d\varphi}, \quad E_\varphi = T(\varphi)\overline{T}(z)\frac{d\hat{T}(\rho)}{d\rho}, \quad E_z = 0,$$
$$H_\rho = \frac{T}{i\omega\mu}\frac{d\hat{T}}{d\rho}\frac{d\overline{T}}{dz}, \quad H_\varphi = \frac{\hat{T}}{i\omega\mu\rho}\frac{dT}{d\varphi}\frac{d\overline{T}}{dz}, \quad H_z = -\frac{\Gamma^2\Psi}{i\omega\mu}. \tag{13-9}$$

In free space q is an integer; but in a region bounded by perfectly conducting half-planes $\varphi = 0$ and $\varphi = \psi$, issuing from the axis of the wave, $dT/d\varphi$ should vanish at these boundaries and consequently

$$Q = 0, \quad q = \frac{n\pi}{\psi}, \quad n = 0,1,2,\cdots. \tag{13-10}$$

In free space ξ may assume any real or imaginary value; but in a region bounded by perfectly conducting planes $z = 0$ and $z = h$, $\overline{T}$ must vanish on the boundaries and therefore

$$C = 0, \quad \xi = \frac{m\pi}{h}, \quad m = 1,2,3,\cdots. \tag{13-11}$$

In the case of circular symmetry $T(\varphi) = 1$ and the field becomes considerably simpler. Thus from (6) we obtain the following expressions for transverse magnetic waves

$$H_\varphi = -\overline{T}(z)\frac{d\hat{T}(\rho)}{d\rho}, \quad E_z = -\frac{\Gamma^2\hat{T}(\rho)\overline{T}(z)}{g + i\omega\epsilon}, \quad E_\rho = \frac{1}{g + i\omega\epsilon}\frac{d\hat{T}}{d\rho}\frac{d\overline{T}}{dz},$$
$$\hat{T}(\rho) = AI_0(\Gamma\rho) + BK_0(\Gamma\rho), \quad \overline{T}(z) = C\cos\xi z + D\sin\xi z. \tag{13-12}$$

The radial impedances for outward bound progressive waves and for stationary waves having no singularity on the axis are

$$K_\rho^+ = \frac{\Gamma K_0(\Gamma\rho)}{(g + i\omega\epsilon)K_1(\Gamma\rho)}, \quad K_\rho^- = \frac{\Gamma I_0(\Gamma\rho)}{(g + i\omega\epsilon)I_1(\Gamma\rho)}. \tag{13-13}$$

The field of transverse electric waves is obtained from (9); thus

$$E_\varphi = \overline{T}(z)\frac{d\hat{T}(\rho)}{d\rho}, \quad H_z = -\frac{\Gamma^2\hat{T}\overline{T}}{i\omega\mu}, \quad H_\rho = \frac{1}{i\omega\mu}\frac{d\hat{T}}{d\rho}\frac{d\overline{T}}{dz},$$
$$K_\rho^+ = \frac{i\omega\mu K_1(\Gamma\rho)}{\Gamma K_0(\Gamma\rho)}, \quad K_\rho^- = \frac{i\omega\mu I_1(\Gamma\rho)}{\Gamma I_0(\Gamma\rho)}. \tag{13-14}$$

If $\overline{T}$ is in the form

$$\overline{T} = Ce^{-i\xi z}, \tag{13-15}$$

we obtain waves whose wave normals, at sufficiently great distances from the axis of the waves, form a cone. A typical equiphase surface will be a cone with its apex on the axis of the waves, except in the neighborhood of the axis itself where the cone is "blunted." Thus, if the medium is nondissipative, and if $\Gamma\rho$ is large we have

$$\hat{T}\overline{T} \sim \frac{1}{\sqrt{\rho}}e^{-\Gamma\rho - i\xi z} = \frac{1}{\sqrt{\rho}}e^{-\rho\sqrt{\xi^2-\beta^2} - i\xi z}. \tag{13-16}$$

When $\xi > \beta$, this function determines a wave traveling parallel to the z-axis; but when $\xi < \beta$, we write (16) as

$$\hat{T}\overline{T} \sim \frac{1}{\sqrt{\rho}}e^{-i\rho\sqrt{\beta^2-\xi^2} - i\xi z}. \tag{13-17}$$

The cone of constant phase is

$$\rho\sqrt{\beta^2 - \xi^2} + \xi z = \text{constant}; \tag{13-18}$$

the wave normals make an angle

$$\vartheta = \tan^{-1}\frac{\sqrt{\beta^2 - \xi^2}}{\xi} \tag{13-19}$$

with the z-axis, and

$$\xi = \beta\cos\vartheta, \quad \sqrt{\beta^2 - \xi^2} = \beta\sin\vartheta. \tag{13-20}$$

Near $\rho = 0$, the phase of $K_0(\Gamma\rho)$ is nearly independent of ρ and the equiphase surfaces are normal to the z-axis. One exception is the case in which $\xi = 0$; then $\vartheta = 90°$, the waves are traveling radially and the equiphase surfaces are cylinders coaxial with the z-axis.

10.14. *Circulating Waves*

A wave is *circulating* if its equiphase surfaces are half-planes issuing from an axis called the axis of circulation. Such waves are possible only in nondissipative media. The field of magnetically oriented circulating waves may be obtained from the following stream function

$$\Pi = [AJ_q(\chi\rho) + BN_q(\chi\rho)](C\cos\xi z + D\sin\xi z)e^{\mp iq\varphi}, \quad \chi^2 + \xi^2 = \beta^2. \tag{14-1}$$

This function is of the same general form as that given by (13-4); but $T(\rho)$ must be real and it must satisfy certain boundary conditions. The form (1) is more suitable for the latter purpose than (13-4). Similarly the field of electrically oriented circulating waves is obtained from a potential Ψ which is of the same form as (1).

Except when q is an integer, circulating waves can exist only within some φ-interval $(\varphi_0, \varphi_0 + \psi)$ where ψ is not greater than 2π. Of course, waves in a wave guide wound into a helical coil are approximately circulating waves in the above defined sense and then φ is unrestricted. The restrictions on the range of φ are physical and not mathematical. Mathematically we may consider fields in a *Riemann Space*, wound on itself; such a space is an idealization of the wave guide wound into a helix.

Some circulating waves have already been studied in section 8.23 in connection with wave propagation in bent wave guides. While the field in such wave guides can be expressed exactly in terms of Bessel functions regarded as functions of their order q, there are no tables available to enable us to apply the formal results to practical problems and we must resort to approximate methods.

10.15. *Relations between Plane, Cylindrical, and Spherical Waves*

From the mathematical point of view the difference between plane, cylindrical, and spherical waves is merely the difference in the coordinates used in the expressions for the wave functions. By transforming coordinates it should be possible to express a spherical wave function in terms of cylindrical wave functions or in terms of plane wave functions. The choice of the particular coordinate system and the particular type of wave function is dictated by the nature of the source and of the boundaries dividing the medium into homogeneous regions. Each current element in a homogeneous medium emits a spherical wave; hence if we know the complete current distribution it is natural to regard the field as the resultant of spherical waves emitted by the current elements. On the other hand, if an element is in a medium consisting of two homogeneous semi-infinite media, separated by a plane interface, we express the spherical wave emitted by the element in terms of plane waves " moving " toward the plane boundary and use the relatively simple formulae for the reflection and transmission coefficients of such waves in order to obtain the field of the element as modified by the change in the environment. Similarly if a plane wave is striking a spherical obstacle, we express it as a resultant of spherical waves incident on the obstacle and use the formulae for the reflection and transmission coefficients of spherical waves at a spherical boundary. In either case the reflection and transmission coefficients depend on the impedance ratio for a typical elementary wave.

At times a mathematical transformation of a wave function throws a new light on a physical phenomenon. In section 8.21 we have considered dominant waves in a rectangular wave guide. Above the cut-off these waves travel along the guide with a velocity greater than that of light. This high phase velocity and the filterlike characteristics of the guide are the properties of all transmission lines having a shunt inductance in parallel with the shunt capacitance. On the other hand if we express the sines and the cosines in (8.21–15) in terms of exponential functions we obtain

$$E_y = -\tfrac{1}{2}iE[e^{i[(\pi/a)x-\bar{\beta}z]} - e^{-i[(\pi/a)x+\bar{\beta}z]}], \quad \bar{\beta} = \sqrt{\beta^2 - \frac{\pi^2}{a^2}},$$

and corresponding expressions for the H-components. The amplitudes of the individual terms are independent of the coordinates; hence these terms represent uniform plane waves. The sum of the squares of the coefficients of x and z in the exponents equals β^2 and a real angle C can be found to satisfy $\bar{\beta} = \beta \cos C$, $\pi/a = \beta \sin C$. Comparing with the equations of section 4.10, we find that our equations represent two uniform

plane waves moving in directions making the following angles with the coordinate axes

$$A = \frac{\pi}{2} + C, \quad B = \frac{\pi}{2}, \quad C = C; \quad \text{and} \quad A = \frac{\pi}{2} - C, \quad B = \frac{\pi}{2}, \quad C = C.$$

For each wave the corresponding total magnetic vector is perpendicular to the electric vector and the ratio of the intensities is η. Either wave is obtained from the other by reflection at one of the faces of the wave guide (Fig. 10.10). This picture also helps to explain why the velocity in the direction of the guide is higher than the velocity of uniform plane waves in free space. The directions in which the uniform components are moving depend on the frequency. When the frequency is sufficiently high $\bar{\beta}$ is nearly equal to β and C is nearly zero; then the waves are moving almost in the direction of the guide. At the cut-off $\bar{\beta} = 0$ and $C = 90°$; in this case the waves cease to advance in the direction of the guide. Below the cut-off this picture fails altogether.

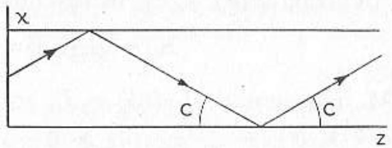


FIG. 10.10. Directions of uniform plane waves into which guided waves in rectangular tubes may be resolved.

If instead of dominant waves we consider waves whose electric intensity varies as $\sin n\pi x/a$, we obtain the following expression for the angle C

$$\sin C = \frac{n\pi}{\beta a} = \frac{n\lambda}{2a}.$$

If λ is small compared with $2a$, there will be several real values of C satisfying this equation, and these will be the " permissible " angles for uniform plane waves between two parallel perfectly conducting planes. Each permissible angle corresponds to a particular transmission mode.

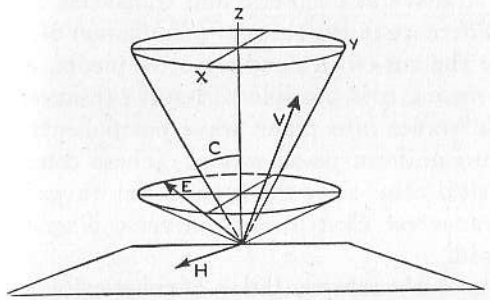


FIG. 10.11. Cone of directions of elementary uniform plane waves.

Even inside circular wave guides it is possible to express guided waves above the cut-off as bundles of plane waves repeatedly reflected from the cylindrical boundary. Consider for instance a bundle of uniform plane waves with H parallel to the xy-plane and let the wave normals of elementary components form a cone coaxial with the z-axis (Fig. 10.11). If, at the origin, the amplitude of the electric intensity of a typical elementary wave is $A(\hat{\varphi})\,d\hat{\varphi}$, then

$$E_z = \sin C \exp(-i\beta z \cos C) \int_0^{2\pi} A(\hat{\varphi}) \exp[-i\beta(x \cos\hat{\varphi} + y \sin\hat{\varphi}) \sin C]\, d\hat{\varphi}. \tag{15–1}$$

Since $x \cos\hat{\varphi} + y \sin\hat{\varphi} = \rho \cos(\varphi - \hat{\varphi})$, the above equation becomes

$$E_z = \sin C \exp(-i\beta z \cos C) \int_0^{2\pi} A(\hat{\varphi}) \exp[-i\beta\rho \sin C \cos(\varphi - \hat{\varphi})]\, d\hat{\varphi}.$$

Letting $A(\hat{\varphi}) = E_0$, using equation (3.7–6), and integrating we have

$$E_z = 2\pi E_0 \sin C J_0(\beta\rho \sin C) e^{-i\beta z \cos C}.$$

More generally, if $A(\hat{\varphi}) = E_0 \cos n\hat{\varphi}$, where n is an integer, then

$$E_z = (-i)^n 2\pi E_0 \sin C J_n(\beta\rho \sin C) \cos n\varphi\, e^{-i\beta z \cos C}. \tag{15–2}$$

Thus we have transverse magnetic waves in which

$$\chi = \beta \sin C, \quad \beta_z = \beta \cos C.$$

For waves inside a perfectly conducting cylinder of radius a, E_z must vanish on the boundary; thus

$$\beta a \sin C = k, \quad \sin C = \frac{k}{\beta a},$$

where k is a zero of $J_n(x)$.

Similarly it is possible to express the field component H_z of transverse electric waves as an integral of the type (1). Hence transverse electric waves above the cut-off may also be regarded as bundles of uniform plane waves. The difference between transverse magnetic and transverse electric waves becomes merely the difference in the state of polarization of the elementary components. Below the cut-off in nondissipative media, and at all frequencies in dissipative media, it is possible to resolve transverse electric and transverse magnetic waves into plane wave components of "exponential type" but not into uniform plane waves. These components, however, are of little practical value since the exponential waves are no simpler than the original transverse electric or transverse magnetic waves from which they are derived.

We shall now turn our attention to the representation of spherical waves in terms of cylindrical, conical, and plane waves. We may confine our attention to an electric current element of moment Il. Such an element may be regarded as an infinitely long current filament in which the current is zero everywhere except in a short interval. If the element is situated at the origin and directed along the z-axis, the current $I(z)$ in the filament is the impulse function (Fig. 10.12); the moment Il is represented by the area of the impulse. In section 2.9 this function has been expressed as the following contour integral

$$I(z) = \frac{I}{i\pi} \int_{(C)} \frac{\sinh \frac{\gamma l}{2}}{\gamma} e^{\gamma z}\, d\gamma, \tag{15–3}$$

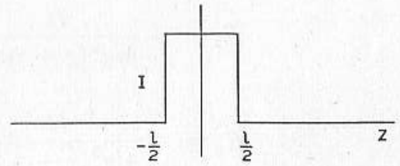


FIG. 10.12. Electric current element regarded as an infinitely long current filament.

where the contour (C) is the imaginary axis indented at the origin (see Fig. 2.19). Thus the current is now represented as the resultant of an infinite number of elementary progressive waves traveling along the z-axis.

Each elementary current wave produces a field of type (13–12) in which

$$\Pi = \hat{T}T = A(\gamma) K_0(\Gamma\rho) e^{\gamma z}\, d\gamma, \quad \Gamma = \sqrt{\sigma^2 - \gamma^2}.$$

Thus for each elementary wave we have

$$H_\varphi = \Gamma A(\gamma) K_1(\Gamma\rho) e^{\gamma z} d\gamma, \quad E_z = -\frac{\Gamma^2}{g + i\omega\epsilon} A(\gamma) K_0(\Gamma\rho) e^{\gamma z}\, d\gamma,$$

$$E_\rho = -\frac{\Gamma\gamma}{g + i\omega\epsilon} A(\gamma) K_1(\Gamma\rho) e^{\gamma z}\, d\gamma, \quad K_\rho^+ = \frac{\Gamma K_0(\Gamma\rho)}{(g + i\omega\epsilon) K_1(\Gamma\rho)}.$$

Assuming that a is the radius of the filament and using $2\pi a H_\varphi(a,z) = I(z)$, we obtain

$$A(\gamma) = \frac{I \sinh \frac{\gamma l}{2}}{2i\pi^2 \gamma \Gamma a K_1(\Gamma a)} \to \frac{I \sinh \frac{\gamma l}{2}}{2i\pi^2 \gamma}, \quad \text{as} \quad a \to 0,$$

and the field produced by $I(z)$ is

$$H_\varphi = \frac{I}{2i\pi^2} \int_{(C)} \frac{\Gamma \sinh \frac{\gamma l}{2}}{\gamma} K_1(\Gamma\rho) e^{\gamma z}\, d\gamma, \quad \Pi = \frac{I}{2i\pi^2} \int_{(C)} \frac{\sinh \frac{\gamma l}{2}}{\gamma} K_0(\Gamma\rho) e^{\gamma z}\, d\gamma,$$

$$E_z = -\frac{I}{2i\pi^2(g + i\omega\epsilon)} \int_{(C)} \frac{\Gamma^2 \sinh \frac{\gamma l}{2}}{\gamma} K_0(\Gamma\rho) e^{\gamma z}\, d\gamma, \tag{15–4}$$

$$E_\rho = -\frac{I}{2i\pi^2(g + i\omega\epsilon)} \int_{(C)} \Gamma \sinh \frac{\gamma l}{2} K_1(\Gamma\rho) e^{\gamma z}\, d\gamma.$$

If l is infinitely small, the above expressions become

$$\Pi = \frac{Il}{4i\pi^2}\int_{(C)} K_0(\Gamma\rho)e^{\gamma z}\,d\gamma, \quad \Gamma = \sqrt{\sigma^2-\gamma^2},$$

$$E_z = -\frac{Il}{4i\pi^2(g+i\omega\epsilon)}\int_{(C)} \Gamma^2 K_0(\Gamma\rho)e^{\gamma z}\,d\gamma,$$

$$E_\rho = -\frac{Il}{4i\pi^2(g+i\omega\epsilon)}\int_{(C)} \gamma\Gamma K_1(\Gamma\rho)e^{\gamma z}\,d\gamma, \qquad (15\text{–}5)$$

$$H_\varphi = \frac{Il}{4i\pi^2}\int_{(C)} \Gamma K_1(\Gamma\rho)e^{\gamma z}\,d\gamma.$$

On the other hand, using spherical coordinates, the field of the current element may be derived from $\Pi = A_z$ in (6.2–7). Comparing with (5), we have

$$\frac{e^{-\sigma r}}{r} = \frac{1}{i\pi}\int_{(C)} K_0(\rho\sqrt{\sigma^2-\gamma^2})e^{\gamma z}\,d\gamma, \quad \rho \neq 0. \qquad (15\text{–}6)$$

This identity could have been established mathematically and equations (4) would then follow.

The integrand of (6) has two branch points $\gamma = \sigma$, $\gamma = -\sigma$. If $g = 0$, these branch points are on the imaginary axis and the contour (C) should be indented as shown in Fig. 10.13. Since $\Gamma = \sqrt{\sigma^2-\gamma^2}$ is a double-valued function it should be remembered that the proper value, corresponding to the principal branch of the K_0-function, is in the first quadrant or on its boundaries. When γ is at infinity on the negative imaginary axis, Γ must be positive real so that the initial phases of $\sigma - \gamma$ and $\sigma + \gamma$ should be

$$\text{ph}(\sigma-\gamma) = \frac{\pi}{2}, \quad \text{ph}(\sigma+\gamma) = -\frac{\pi}{2}, \quad \gamma = -i\infty.$$

The change in ph Γ along the contour may easily be followed if the above differences are represented graphically as in Fig. 10.14. Thus at infinity on the positive imaginary axis we have

$$\text{ph}(\sigma-\gamma) = -\frac{\pi}{2}, \quad \text{ph}(\sigma+\gamma) = \frac{\pi}{2}, \quad \gamma = i\infty,$$

and Γ is again positive real.

In the case of nondissipative media Γ is either real or imaginary on (C) except on the two infinitely small indentations;* it is real outside the interval $\gamma = -i\beta$, $\gamma = i\beta$ and imaginary inside it. When Γ is real, the

* There is no singularity at the origin and therefore (C) need not be indented there.

equiphase surfaces of the elementary waves are normal to the z-axis and the waves are plane; when Γ is imaginary, the equiphase surfaces are inclined to the z-axis (and curved near it) and the waves are conical.

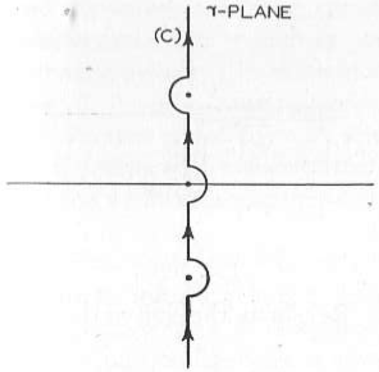


FIG. 10.13.

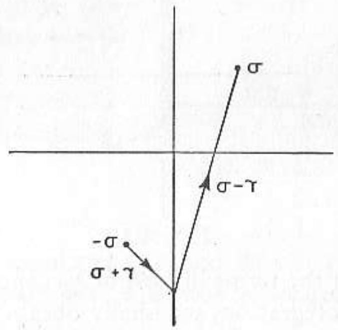


FIG. 10.14.

Separating the integral into two parts, one taken along the positive and the other along the negative imaginary axis, and reversing the sign of the variable of integration in the second part, we obtain

$$\frac{e^{-i\beta r}}{r} = \frac{2}{i\pi}\int_0^{i\infty} K_0(\rho\sqrt{-\beta^2-\gamma^2})\cosh\gamma z\,d\gamma = \frac{2}{\pi}\int_0^\infty K_0(\rho\sqrt{\xi^2-\beta^2})\cos\xi z\,d\xi. \qquad (15\text{–}7)$$

The second integral is taken along the positive real axis indented above $\xi = \beta$. The equiphase surfaces of elementary waves are cylinders coaxial with the z-axis.

The above expressions are valid either for positive or negative values of z. If $z > 0$, the exponential function in the integrand vanishes on the circle of infinite radius in the second quadrant and (C) can be deformed into the contour (C_1) shown in Fig. 10.15; on the negative real axis ph $\Gamma = \pi/2$. The exponential function vanishes also in the third quadrant and (C_1) can be deformed into (C_2) of Fig. 10.16. On the lower side of the negative real axis ph $\Gamma = -\pi/2$. Between $\gamma = 0$ and $\gamma = -i\beta$, ph $\Gamma = \pi/2$ on the right side of the contour and ph $\Gamma = -\pi/2$ on the left side. By (3.4–11) we have

$$\frac{2}{\pi}K_0(\rho\sqrt{-\beta^2-\gamma^2}) = -N_0(\rho\sqrt{\beta^2+\gamma^2}) - iJ_0(\rho\sqrt{\beta^2+\gamma^2}), \quad \text{ph}\,\Gamma = \frac{\pi}{2},$$
$$= -N_0(\rho\sqrt{\beta^2+\gamma^2}) + iJ_0(\rho\sqrt{\beta^2+\gamma^2}), \quad \text{ph}\,\Gamma = -\frac{\pi}{2}, \qquad (15\text{–}8)$$

where the arguments of the J_0 and N_0 functions are real on (C_2) excepting the infinitely small circle round $\gamma = -i\beta$. The integral round this circle is zero because the N_0-function becomes infinite as $\log|\gamma + i\beta|$ while the length of the circumference vanishes as $|\gamma + i\beta|$. On the two halves of (C_2), $d\gamma$ is in opposite directions; thus substituting from (8) in (6), we find

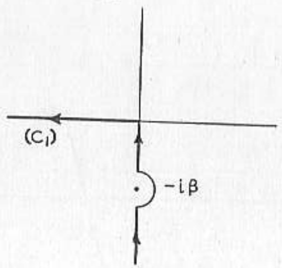


FIG. 10.15.

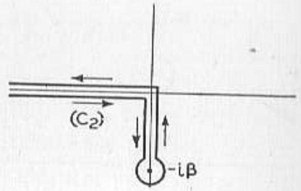


FIG. 10.16.

that the terms involving N_0 cancel out. Reversing the sign of the variable of integration, we finally obtain

$$\frac{e^{-i\beta r}}{r} = \int_0^\infty J_0(\rho\sqrt{\beta^2 + \gamma^2})e^{-\gamma z}\,d\gamma - \int_0^{i\beta} J_0(\rho\sqrt{\beta^2 + \gamma^2})e^{-\gamma z}\,d\gamma.$$

Introducing a new variable $\chi = \sqrt{\beta^2 + \gamma^2}$ and combining the two integrals in one, we have

$$\frac{e^{-i\beta r}}{r} = \int_{(C_3)} J_0(\chi\rho)e^{-z\sqrt{\chi^2-\beta^2}}\frac{\chi\,d\chi}{\sqrt{\chi^2 - \beta^2}}, \quad z > 0, \tag{15-9}$$

where (C_3) is the positive real axis indented above $\chi = \beta$ (Fig. 10.17). Unlike (6) this integral converges when $\rho = 0$; it converges also when $z = 0$ as long as $\rho \neq 0$. In this representation all elementary waves are plane waves traveling in the positive z-direction. When $\chi > \beta$, the propagation constant in the z-direction has all real values between zero and infinity; when $\chi < \beta$, the propagation constant is imaginary and the phase velocity assumes all values in the interval (v, ∞) where v is the characteristic velocity. This second group of waves can be expressed as a group of waves traveling with constant velocity v but in different directions; we need only express $J_0(\chi\rho)$ as an integral of type (1).

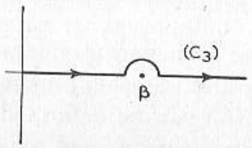


FIG. 10.17.

When $z < 0$, we have

$$\frac{e^{-i\beta r}}{r} = \int_{(C_3)} J_0(\chi\rho)e^{z\sqrt{\chi^2-\beta^2}}\frac{\chi\,d\chi}{\sqrt{\chi^2 - \beta^2}}$$
$$= \frac{1}{i\pi}\int_{(C_4)} K_0\left(\rho\sqrt{-\beta^2 - \gamma^2}\right)e^{\gamma z}\,d\gamma, \tag{15-10}$$

where the contour of integration (C_4) is shown in Fig. 10.18. On the real axis the elementary waves are plane and their propagation constants are real; on the imaginary axis above $\gamma = i\beta$, Γ is real and the waves are plane, traveling in the negative z-direction with velocities smaller than the characteristic velocity; in the region between $\gamma = 0$ and $\gamma = i\beta$, Γ and γ are imaginary, with the sum of their squares equal to $-\beta^2$, and the waves are conical, traveling with velocity v in directions making acute angles with the negative z-axis.

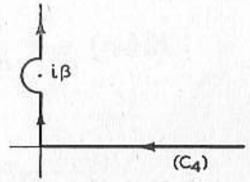


FIG. 10.18.

10.16. *Waves on an Infinitely Long Wire*

Consider an infinitely long cylindrical wire of radius a and assume that an electromotive force V is applied uniformly over a section of length s; then, assuming that the wire is along the z-axis,

$$E_z^i(a,z) = \frac{V}{s}, \quad -\frac{s}{2} < z < \frac{s}{2}; \tag{16-1}$$

outside the interval the applied intensity is zero. Representing this impulse function as a contour integral we have

$$E_z^i(a,z) = \frac{V}{i\pi s}\int_{(C)} \frac{\sinh\frac{\gamma s}{2}}{\gamma}e^{\gamma z}\,d\gamma. \tag{16-2}$$

Consequently the magnetic intensity at the surface of the wire is

$$H_\varphi(a,z) = \frac{V}{i\pi s}\int_{(C)} \frac{\sinh\frac{\gamma s}{2}\,e^{\gamma z}}{\gamma[K_\rho^+(a) + K_\rho^-(a)]}\,d\gamma. \tag{16-3}$$

In order to obtain the magnetic intensity at any distance ρ from the axis of the wire we multiply each element of the integrand in the above expression by the H-transfer ratio; thus

$$H_\varphi(\rho,z) = \frac{V}{i\pi s}\int_{(C)} \frac{\sinh\frac{\gamma s}{2}\,K_1(\Gamma\rho)e^{\gamma z}}{\gamma[K_\rho^+(a) + K_\rho^-(a)]K_1(\Gamma a)}\,d\gamma. \tag{16-4}$$

The corresponding E_z is obtained if we multiply the integrand by $-K_\rho^+(\rho)$; hence

$$E_z(\rho,z) = -\frac{V}{i\pi s(g + i\omega\epsilon)}\int_{(C)} \frac{\Gamma\sinh\frac{\gamma s}{2}\,K_0(\Gamma\rho)e^{\gamma z}}{\gamma[K_\rho^+(a) + K_\rho^-(a)]K_1(\Gamma a)}\,d\gamma. \tag{16-5}$$

In the special case when the wire is a perfect conductor and the medium is nondissipative these equations reduce to

$$E_z(\rho,z) = -\frac{V}{i\pi s}\int_{(C)} \frac{\sinh\frac{\gamma s}{2} K_0(\Gamma\rho)e^{\gamma z}}{\gamma K_0(\Gamma a)}\,d\gamma, \quad \Gamma = \sqrt{-\gamma^2-\beta^2},$$

$$H_\varphi(\rho,z) = \frac{\omega\epsilon V}{\pi s}\int_{(C)} \frac{\sinh\frac{\gamma s}{2} K_1(\Gamma\rho)e^{\gamma z}}{\gamma\Gamma K_0(\Gamma a)}\,d\gamma. \tag{16–6}$$

The current $I(z)$ in the wire is then

$$I(z) = \frac{2\omega\epsilon a V}{s}\int_{(C)} \frac{\sinh\frac{\gamma s}{2} K_1(\Gamma a)e^{\gamma z}}{\gamma\Gamma K_0(\Gamma a)}\,d\gamma. \tag{16–7}$$

If the input admittance Y_i is defined as the ratio of the current $I(s/2)$ through the upper terminal of the "generator" to the applied voltage, then

$$Y_i = \frac{2\omega\epsilon a}{s}\int_{(C)} \frac{e^{\gamma s/2}\sinh\frac{\gamma s}{2} K_1(\Gamma a)}{\gamma\Gamma K_0(\Gamma a)}\,d\gamma. \tag{16–8}$$

If s is small, the real part of the input admittance is nearly independent of s; but the reactive part of the admittance is positive and approaches infinity as s approaches zero. This is natural since the capacitance between two infinitely close cylindrical wires must be infinite. For thin wires the above integrals yield the results of section 8.13.

10.17. *Waves on Coaxial Conductors*

Consider now two perfectly conducting coaxial cylinders and assume that an electromotive force V_0, defined by equation (16–1), is applied to the inner cylinder. The longitudinal electric intensity is then the negative of the impressed intensity (16–2) and

$$E_z(a,z) = -\frac{V_0}{i\pi s}\int_{(C)} \frac{\sinh\frac{\gamma s}{2}}{\gamma}e^{\gamma z}\,d\gamma. \tag{17–1}$$

Between the coaxial cylinders we have

$$E_z(\rho,z) = -\frac{V_0}{i\pi s}\int_{(C)} \frac{\sinh\frac{\gamma s}{2}}{\gamma}\chi_{E_z}(a,\rho)e^{\gamma z}\,d\gamma, \tag{17–2}$$

$$H_\varphi(\rho,z) = \frac{V_0}{i\pi s}\int_{(C)} \frac{\sinh\frac{\gamma s}{2}\chi_{E_z}(a,\rho)}{\gamma Z_\rho^+(\rho)}e^{\gamma z}\,d\gamma, \tag{17–2}$$

$$E_\rho(\rho,z) = \frac{V_0}{i\pi s}\int_{(C)} K_z^+\frac{\sinh\frac{\gamma s}{2}\chi_{E_z}(a,\rho)}{\gamma Z_\rho^+(\rho)}e^{\gamma z}\,d\gamma.$$

For each elementary wave we may assume

$$E_z = [AJ_0(\chi\rho) + BN_0(\chi\rho)]e^{\gamma z}, \quad \chi^2-\gamma^2=\beta^2,$$

$$H_\varphi = \frac{i\omega\epsilon}{\chi}[AJ_1(\chi\rho) + BN_1(\chi\rho)]e^{\gamma z}, \tag{17–3}$$

$$E_\rho = -\frac{\gamma}{\chi}[AJ_1(\chi\rho) + BN_1(\chi\rho)]e^{\gamma z}.$$

The longitudinal impedance of a typical wave is then $K_z^+ = -\gamma/i\omega\epsilon$. Since $E_z(b) = 0$, we have $A = PN_0(\chi b)$, $B = -PJ_0(\chi b)$; hence

$$\chi_{E_z}(a,\rho) = \frac{N_0(\chi b)J_0(\chi\rho) - J_0(\chi b)N_0(\chi\rho)}{N_0(\chi b)J_0(\chi a) - J_0(\chi b)N_0(\chi a)},$$

$$Z_\rho^+(\rho) = -\frac{\chi[N_0(\chi b)J_0(\chi\rho) - J_0(\chi b)N_0(\chi\rho)]}{i\omega\epsilon[N_0(\chi b)J_1(\chi\rho) - J_0(\chi b)N_1(\chi\rho)]}. \tag{17–4}$$

Hence equations (2) assume the following form

$$E_z(\rho,z) = -\frac{V_0}{i\pi s}\int_{(C)} \frac{\sinh\frac{\gamma s}{2}S(\chi\rho,\chi b)}{\gamma S(\chi a,\chi b)}e^{\gamma z}\,d\gamma,$$

$$H_\varphi(\rho,z) = \frac{\omega\epsilon V_0}{\pi s}\int_{(C)} \frac{\sinh\frac{\gamma s}{2}U(\chi\rho,\chi b)}{\gamma\chi S(\chi a,\chi b)}e^{\gamma z}\,d\gamma, \tag{17–5}$$

$$E_\rho(\rho,z) = -\frac{V_0}{i\pi s}\int_{(C)} \frac{\sinh\frac{\gamma s}{2}U(\chi\rho,\chi b)}{\chi S(\chi a,\chi b)}e^{\gamma z}\,d\gamma,$$

$$S(\chi\rho,\chi b) = J_0(\chi\rho)N_0(\chi b) - N_0(\chi\rho)J_0(\chi b),$$

$$U(\chi\rho,\chi b) = N_1(\chi\rho)J_0(\chi b) - J_1(\chi\rho)N_0(\chi b).$$

The transverse voltage $V(z)$ is obtained by integrating E_ρ along a radius and the longitudinal current $I(z)$ in the inner cylinder by multiplying

$H_\varphi(a)$ by $2\pi a$; thus we have

$$V(z) = \frac{V_0}{i\pi s}\int_{(C)} \frac{\sinh \frac{\gamma s}{2}}{\gamma^2 + \beta^2} e^{\gamma z}\, d\gamma, \quad I(z) = \frac{2\omega\epsilon a V_0}{s}\int_{(C)} \frac{\sinh \frac{\gamma s}{2}\, U(\chi a, \chi b)}{\gamma\chi S(\chi a, \chi b)} e^{\gamma z}\, d\gamma. \tag{17-6}$$

The first of these integrals can be calculated at once by the method of residues. Thus if $z > s/2$, (C) can be closed in the left half of the plane. The point $\gamma = -i\beta$ is the only pole enclosed and the value of the integral is

$$V(z) = \tfrac{1}{2}V_0 \frac{\sin \frac{\beta s}{2}}{\frac{\beta s}{2}} e^{-i\beta z}, \quad z > \frac{s}{2}. \tag{17-7}$$

Thus the transverse voltage is given by the principal wave alone.

To evaluate $I(z)$ we note that the integrand is a single-valued function of γ and consequently we may again use the method of residues. The poles of the integrand are the roots of

$$S(\chi a, \chi b) = 0 \quad \text{or} \quad Z_\rho^+(a) = 0, \tag{17-8}$$

and thus represent the natural propagation constants. The contrast between this case and the case of an isolated wire is worth noting. In the latter case the radial impedance function is multiple-valued and has no poles. If χ_n is the nth root of (8), then $\gamma_n = \pm\Gamma_n$, $\Gamma_n = \sqrt{\chi_n^2 - \beta^2}$. When the frequency is such that $\beta < \chi_1$, all these poles are on the real axis. As the frequency increases, some poles move to the imaginary axis. In addition there are two poles $\gamma_0 = \pm i\beta$ which are always on the imaginary axis. Evaluating $I(z)$ for $z > s/2$, we have

$$I(z) = \sum_{n=0}^{\infty} I_n(z), \quad I_n(z) = \frac{4\pi i\omega\epsilon a V_0 \sinh \frac{\Gamma_n s}{2}\, U(\chi_n a, \chi_n b)}{s\Gamma_n \chi_n \frac{dS}{d\chi_n}} e^{-\Gamma_n z},$$

$$I_0(z) = \frac{V(z)}{K}, \quad K = \frac{\eta}{2\pi}\log\frac{b}{a}. \tag{17-9}$$

When $\beta < \chi_1$, all current waves except the dominant wave $I_0(z)$ are attenuated. The impedance to the dominant wave is a pure resistance; the additional current waves add a reactance in parallel with this resistance. If we define the input admittance as the ratio of $I(s/2)$ to the applied voltage V_0, then we have

$$Y_i = \sum_{n=0}^{\infty} Y_n, \quad Y_0 = \frac{1}{2K}\frac{\sin\frac{\beta s}{2}}{\frac{\beta s}{2}} e^{-i\beta s/2}, \quad Y_n = \frac{I_n\left(\frac{s}{2}\right)}{V_0}. \tag{17-10}$$

When s is vanishingly small, the resistive component of the input admittance is $1/2K$. In this case, however, the reactance in parallel with $2K$ is infinitely large; an infinitely small gap provides an infinitely large capacitance and effectively short-circuits the gap. Nevertheless the capacitance approaches infinity so slowly that for the "small gap" encountered in practice the capacitance is usually small.

If the applied electromotive force were not distributed uniformly round the circumference of the inner conductor, all possible natural waves might be generated in the region between the two coaxial conductors. A study of the roots of the equations in section 10.7 indicates that if the wavelength is greater than the circumference of the outer conductor, all natural waves, except the principal, are attenuated.

10.18. *Waves on Parallel Wires*

Consider now two thin perfectly conducting parallel wires, with their axes separated by distance l, and let the impressed voltage V_0 be distributed along wire 1 in accordance with equations (16–1) and (16–2). The longitudinal electric intensities due to the currents in each wire are respectively

$$E_z(\rho_1, z) = \int_{(C)} A(\gamma)K_0(\Gamma\rho_1)e^{\gamma z}\, d\gamma, \quad E_z(\rho_2, z) = \int_{(C)} B(\gamma)K_0(\Gamma\rho_2)e^{\gamma z}\, d\gamma, \tag{18-1}$$

where ρ_1 and ρ_2 are the distances from the axes of the wires. At the surface of each wire the boundary conditions are

$$A(\gamma)K_0(\Gamma a) + B(\gamma)K_0(\Gamma l) = -\frac{V_0 \sinh\frac{\gamma s}{2}}{i\pi s\gamma},$$
$$A(\gamma)K_0(\Gamma l) + B(\gamma)K_0(\Gamma a) = 0. \tag{18-2}$$

Taking one half of the sum and of the difference of these equations, we have

$$\tfrac{1}{2}[A(\gamma) + B(\gamma)] = -\frac{V_0 \sinh\frac{\gamma s}{2}}{2i\pi s\gamma[K_0(\Gamma a) + K_0(\Gamma l)]} = C(\gamma),$$
$$\tfrac{1}{2}[A(\gamma) - B(\gamma)] = \frac{V_0 \sinh\frac{\gamma s}{2}}{2i\pi s\gamma[K_0(\Gamma l) - K_0(\Gamma a)]} = D(\gamma). \tag{18-3}$$

Adding and subtracting we obtain

$$A(\gamma) = C(\gamma) + D(\gamma), \quad B(\gamma) = C(\gamma) - D(\gamma). \tag{18–4}$$

Hence the total longitudinal intensity may be expressed as $E_z = \bar{E}_z + \hat{E}_z$, where

$$\bar{E}_z = \int_{(C)} C(\gamma)[K_0(\Gamma\rho_1) + K_0(\Gamma\rho_2)]e^{\gamma z}\,d\gamma,$$

$$\hat{E}_z = \int_{(C)} D(\gamma)[K_0(\Gamma\rho_1) - K_0(\Gamma\rho_2)]e^{\gamma z}\,d\gamma. \tag{18–5}$$

Thus the total field consists of two parts, the one a symmetric function of ρ_1 and ρ_2 and the other an anti-symmetric. The first part alone would be produced if a voltage equal to $\frac{1}{2}V_0$ were applied in phase to each wire; at distances large compared with the interaxial separation between the wires the wave is similar to the wave on a single wire and approximately spherical. The second part alone would be produced if the voltages were applied in push-pull, that is, $\frac{1}{2}V_0$ applied to one wire and $-\frac{1}{2}V_0$ to the other. In what follows we shall consider only the second part.

The remaining field intensities are obtained from equations (13–12), that is, from

$$H_\varphi = \frac{i\omega\epsilon}{\Gamma^2}\frac{\partial E_z}{\partial \rho}, \quad E_\rho = -\frac{1}{\Gamma^2}\frac{\partial^2 E_z}{\partial\rho\,\partial z}, \quad \Gamma = \sqrt{-\gamma^2 - \beta^2}.$$

The transverse voltage $V(z)$ from the first wire to the second is found to be the same as in the case of a coaxial pair. The electric current in the first wire is

$$I(z) = -2\pi i\omega\epsilon a \int_{(C)} \frac{1}{\Gamma} D(\gamma)K_1(\Gamma a)e^{\gamma z}\,d\gamma.$$

We have seen that the integrand in the expression for the current in the inner wire of a coaxial pair is a single-valued function having an infinite number of poles. The present integrand is multiple-valued and thus resembles the integrand for the current in a single wire. On the other hand, it differs from the latter in that its contribution to the current from the part of the contour in the vicinity of one or other of the branch points is finite. This contribution is found to be equivalent to a pair of plane waves guided by the wires and moving in opposite directions from the source. Thus if $z > s/2$, we deform (C) into (C_2) of Fig. 10.16. The contribution from the infinitely small circle (C_0) round $\gamma = -i\beta$ is then

$$I_0(z) = \frac{V_0}{2K}\frac{\sin\frac{\beta s}{2}}{\frac{\beta s}{2}}e^{-i\beta z}, \quad K = \frac{\eta}{\pi}\log\frac{l}{a}.$$

The quantity K is seen to be the characteristic impedance of two parallel wires to a transverse electromagnetic plane wave moving parallel to the wires. As the length of the section over which the impressed voltage is distributed approaches zero, the second factor in the current formula approaches unity.

The integral over the straight portions of the contour (C_2) may be transformed with the aid of (15–8) and (3.4–11) to obtain the following expressions

$$I_1(z) = \frac{4i\omega\epsilon V_0}{\pi s}\int_{-\infty}^{0} P(\gamma)e^{\gamma z}\,d\gamma, \quad I_2(z) = \frac{4i\omega\epsilon V_0}{\pi s}\int_{0}^{-i\beta} P(\gamma)e^{\gamma z}\,d\gamma,$$

$$P(\gamma) = \frac{\sinh\frac{\gamma s}{2}}{\gamma(\beta^2+\gamma^2)}\,\frac{1 + \frac{1}{2}\pi\hat{a}[N_1(\hat{a})J_0(\hat{l}) - J_1(\hat{a})N_0(\hat{l})]}{[N_0(\hat{a}) - N_0(\hat{l})]^2 + [J_0(\hat{a}) - J_0(\hat{l})]^2},$$

$$\hat{a} = a\sqrt{\beta^2+\gamma^2}, \quad \hat{l} = l\sqrt{\beta^2+\gamma^2}.$$

$I_1(z)$ is in quadrature with V_0 for all values of z and hence represents a stationary wave; $I_1(s/2)$ is positive imaginary and the corresponding part of the input admittance is positive and tends to infinity as s approaches zero. $I_2(z)$ represents a group of progressive current waves moving with velocities greater than the characteristic velocity; $I_2(z)$ approaches a limit as s approaches zero and $I_2(0)$ is in phase with the applied voltage V_0; consequently the corresponding admittance is a conductance. Thus the total input admittance of a parallel pair energized symmetrically in push-pull consists of three admittances in parallel as shown in Fig. 10.19. The first admittance is the admittance to the dominant transmission mode of the two halves of the parallel wire transmission line connected in series and it accounts for the power guided by the line. The second term is the capacitance representing the stationary field in the vicinity of the generator; in this local field the electric lines connect the two halves of each wire on the opposite sides of the generators. The third term is the radiation conductance; it accounts for power radiated in directions other than the direction parallel to the wires. In order to prove the last statement we should show that $I_2(z)$ approaches zero as z increases.

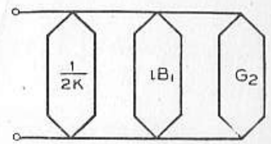


FIG. 10.19. Admittance seen by a generator in series with a parallel pair; K is the characteristic impedance to the principal wave, G_2 is the radiation conductance, and B_1 is the capacitive susceptance representing the local reactive field.

10.19. *Forced Waves in Metal Tubes*

Consider a perfectly conducting cylinder of radius a, coaxial with the z-axis. Let the source be an infinitely thin line source of finite length, parallel to the axis of the tube and passing through the point $(b,0)$ in the equatorial plane (Fig. 10.20). If the source consists of electric current elements, the longitudinal electric intensity will be of the following form

$$E_z^t = \int_{(C)} A(\gamma)K_0(i\chi\hat{\rho})e^{\gamma z}\,d\gamma, \quad \chi^2 = \gamma^2 + \beta^2, \tag{19–1}$$

where $\hat{\rho}$ is the distance from the axis of the source and χ is the radial phase constant. Thus in the case of an infinitely short current element in nondissipative dielectric, we have $A(\gamma) = -Il\chi^2/4\omega\epsilon\pi^2$. In the case of an infinitely thin antenna energized at the center we obtain from (9.25–14) and (15–6) (assuming the center at $z = 0$)

$$A(\gamma) = \frac{30I}{\pi}\,(2\cos\beta l - e^{\gamma l} - e^{-\gamma l}).$$

An antenna of finite radius may be regarded as consisting of infinitely thin filaments and the corresponding expressions for $A(\gamma)$ can easily be calculated.

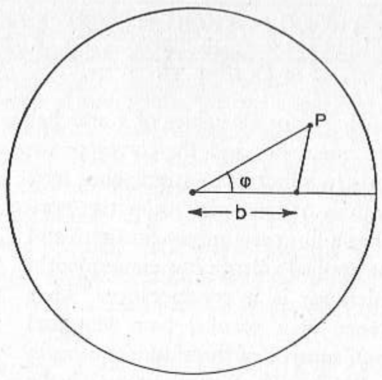


FIG. 10.20. Cross-section of a metal tube and a line source inside it.

In the region $\rho > b$ the K_0-function in (1) can be expressed in terms of the coordinates ρ and φ as in (8.17–2); thus we have

$$E_z^i = \sum_{n=0}^{\infty} i^n \epsilon_n \cos n\varphi \int_{(C)} A(\gamma) J_n(\chi b) K_n(i\chi\rho) e^{\gamma z}\, d\gamma, \tag{19–2}$$

where ϵ_n is the Neumann number defined by $\epsilon_0 = 1$, $\epsilon_n = 2$ if $n \neq 0$. This field is impressed on the cylinder and gives rise to a reflected field E_z^r; since the total field should vanish at $\rho = a$, we have

$$E_z^r = -\sum_{n=0}^{\infty} i^n \epsilon_n \cos n\varphi \int_{(C)} \frac{A(\gamma) J_n(\chi b) K_n(i\chi a) J_n(\chi\rho)}{J_n(\chi a)} e^{\gamma z}\, d\gamma.$$

The total longitudinal intensity is $E_z^i + E_z^r$; from this, using (1–13), we may obtain the stream function Π and thence the remaining field intensities. From (3.4–6) we have

$$K_n(ze^{i\pi}) = (-)^n K_n(z) - i\pi I_n(z),$$
$$K_n(ze^{-i\pi}) = (-)^n K_n(z) + i\pi I_n(z).$$

With the aid of these identities it can be shown that the total E_z is a single-valued function of γ; hence it can be evaluated in terms of the residues at the poles of the integrand. If z is greater than the z-coordinate of any current element producing the

wave, we can close (C) in the left half of the plane and thus obtain

$$E_z = -2\pi i \sum_{n,m}\sum \frac{i^n \epsilon_n A(-\Gamma_{n,m}) J_n\left(\frac{b}{a} k_{n,m}\right) K_n(ik_{n,m}) J_n\left(k_{n,m}\frac{\rho}{a}\right)}{J_n'(k_{n,m}) \left.\frac{d\chi a}{d\gamma}\right|_{\gamma=-\Gamma_{n,m}}} e^{-\Gamma_{n,m}z} \cos n\varphi$$

$$= -\frac{i\pi^2}{a} \sum_{n,m}\sum \frac{\epsilon_n k_{n,m} A(-\Gamma_{n,m}) J_n\left(\frac{b}{a} k_{n,m}\right) N_n(k_{n,m})}{\Gamma_{n,m} a J_n'(k_{n,m})} J_n\left(k_{n,m}\frac{\rho}{a}\right) e^{-\Gamma_{n,m}z} \cos n\varphi,$$

where $k_{n,m}$ is a typical zero of $J_n(\chi a)$ and $\Gamma_{n,m}$ is the corresponding natural propagation constant in the direction of the tube. From (3.7–13) we have $J_n'(k_{n,m}) = -2/\pi k_{n,m} N_n(k_{n,m})$, and

$$E_z = \frac{\pi^3}{2a} \sum_{n,m}\sum \frac{\epsilon_n k_{n,m}^2 A(-\Gamma_{n,m}) N_n^2(k_{n,m}) J_n\left(\frac{b}{a} k_{n,m}\right)}{\beta a \sqrt{1 - v_{n,m}^2}} J_n\left(k_{n,m}\frac{\rho}{a}\right) e^{-\Gamma_{n,m}z} \cos n\varphi.$$

Thus we have an expression for the amplitudes of various natural waves generated in the tube by a given distribution of current elements parallel to the axis of the tube. Analogous expressions can be obtained for a similar distribution of magnetic current elements and also for distributions of current elements perpendicular to the axis of the tube.

10.20. *Waves in Dielectric Wires*

First we shall consider natural waves in a nondissipative dielectric wire of radius a imbedded in a nondissipative medium. Except in the case of circular symmetry these waves are hybrid. The components of E and H parallel to the wire are given by the following expressions

$$E_z = AJ_n(\chi\rho)\cos n\varphi, \quad H_z = BJ_n(\chi\rho)\sin n\varphi,$$
$$E_z = CK_n(k\rho)\cos n\varphi, \quad H_z = DK_n(k\rho)\sin n\varphi, \tag{20–1}$$

where the first pair refers to the wire and the second to the medium surrounding it. The exponential factor $e^{-\Gamma z + i\omega t}$ is implied. The transverse phase constant χ in the wire, the transverse propagation constant k outside the wire, the longitudinal propagation constant Γ, the intrinsic phase constant β_1 of the wire, and the intrinsic phase constant β_2 of the external medium are connected by the following equations

$$\chi^2 = \Gamma^2 + \beta_1^2, \quad k^2 = -\Gamma^2 - \beta_2^2, \quad \beta_1 = \omega\sqrt{\mu_1\epsilon_1}, \quad \beta_2 = \omega\sqrt{\mu_2\epsilon_2}. \tag{20–2}$$

If the waves are to be plane, χ and k must be real.

From (1) we obtain the φ components of the field

$$E_\varphi = \left[A\frac{n\Gamma}{\chi^2\rho}J_n(\chi\rho) + B\frac{i\omega\mu_1}{\chi}J_n'(\chi\rho)\right]\sin n\varphi,$$
$$H_\varphi = -\left[A\frac{i\omega\epsilon_1}{\chi}J_n'(\chi\rho) + B\frac{n\Gamma}{\chi^2\rho}J_n(\chi\rho)\right]\cos n\varphi,$$
$$E_\varphi = -\left[C\frac{n\Gamma}{k^2\rho}K_n(k\rho) + D\frac{i\omega\mu_2}{k}K_n'(k\rho)\right]\sin n\varphi,$$
$$H_\varphi = \left[C\frac{i\omega\epsilon_2}{k}K_n'(k\rho) + D\frac{n\Gamma}{k^2\rho}K_n(k\rho)\right]\cos n\varphi, \tag{20-3}$$

where again the first pair refers to the wire and the second to the external medium. At $\rho = a$ the tangential intensities should be continuous and therefore

$$AJ_n(\chi a) = CK_n(ka), \quad BJ_n(\chi a) = DK_n(ka),$$
$$A\frac{n\Gamma}{\chi^2 a}J_n(\chi a) + B\frac{i\omega\mu_1}{\chi}J_n'(\chi a) = -C\frac{n\Gamma}{k^2 a}K_n(ka) - D\frac{i\omega\mu_2}{k}K_n'(ka), \tag{20-4}$$
$$A\frac{i\omega\epsilon_1}{\chi}J_n'(\chi a) + B\frac{n\Gamma}{\chi^2 a}J_n(\chi a) = -C\frac{i\omega\epsilon_2}{k}K_n'(ka) - D\frac{n\Gamma}{k^2 a}K_n(ka).$$

To solve these equations we write

$$A = SK_n(ka), \quad C = SJ_n(\chi a), \quad B = TK_n(ka), \quad D = TJ_n(\chi a), \tag{20-5}$$

and substitute in (4) to obtain

$$\frac{n\Gamma}{a}J_nK_n\left(\frac{1}{\chi^2}+\frac{1}{k^2}\right)S = -i\omega\left(\frac{\mu_2 J_nK_n'}{k} + \frac{\mu_1 K_nJ_n'}{\chi}\right)T,$$
$$-i\omega\left(\frac{\epsilon_2 J_nK_n'}{k} + \frac{\epsilon_1 K_nJ_n'}{\chi}\right)S = \frac{n\Gamma}{a}J_nK_n\left(\frac{1}{\chi^2}+\frac{1}{k^2}\right)T. \tag{20-6}$$

Eliminating S and T, and letting $\chi a = p$, $ka = q$, we have

$$\frac{\epsilon_1\mu_1 J_n'^2(p)}{p^2J_n^2(p)} + \frac{(\epsilon_1\mu_2+\mu_1\epsilon_2)J_n'(p)K_n'(q)}{pqJ_n(p)K_n(q)} + \frac{\mu_2\epsilon_2 K_n'^2(q)}{q^2K_n^2(q)}$$
$$= n^2\left(\frac{1}{p^2}+\frac{1}{q^2}\right)\left(\frac{\epsilon_1\mu_1}{p^2}+\frac{\epsilon_2\mu_2}{q^2}\right). \tag{20-7}$$

Since

$$p^2J_n'^2(p) = -p^2J_{n-1}(p)J_{n+1}(p) + n^2J_n^2(p),$$
$$q^2K_n'^2(q) = q^2K_{n-1}(q)K_{n+1}(q) + n^2K_n^2(q),$$

equation (7) becomes

$$-\frac{\mu_1\epsilon_1 J_{n-1}(p)J_{n+1}(p)}{p^2J_n^2(p)} + \frac{(\mu_1\epsilon_2+\mu_2\epsilon_1)J_n'(p)K_n'(q)}{pqJ_n(p)K_n(q)} + \frac{\mu_2\epsilon_2 K_{n-1}(q)K_{n+1}(q)}{q^2K_n^2(q)}$$
$$= n^2\frac{\mu_1\epsilon_1+\mu_2\epsilon_2}{p^2q^2}. \tag{20-8}$$

From this equation we obtain graphically or numerically pairs of values p, q in terms of which ω and Γ are given by

$$\omega = \frac{\sqrt{p^2+q^2}}{a\sqrt{\mu_1\epsilon_1-\mu_2\epsilon_2}}, \quad \Gamma = i\sqrt{\omega^2\mu_2\epsilon_2+\frac{q^2}{a^2}} = i\sqrt{\omega^2\mu_1\epsilon_1-\frac{p^2}{a^2}}. \tag{20-9}$$

Equation (8) has only real roots. When $q = 0$, we have

$$\omega = \frac{p_0}{a\sqrt{\mu_1\epsilon_1-\mu_2\epsilon_2}}, \quad \Gamma = i\beta_2. \tag{20-10}$$

The wave is traveling with the velocity characteristic of the medium external to the wire. The frequency given by the above expression is the lowest frequency for which there exist plane waves (with a particular field configuration) traveling in the direction of the wire; for this reason it may be called the cut-off frequency. When q becomes infinite, the frequency is also infinite and the wave is traveling with the velocity characteristic of the wire. For large values of q the field outside the wire varies exponentially with the distance from the surface; thus the field is confined to a relatively thin film of the external medium and is largely inside the wire.

In order to obtain the cut-off frequencies for various transmission modes we let q approach zero; thus for $n > 1$ we obtain

$$(\mu_1\epsilon_2+\mu_2\epsilon_1)\frac{pJ_{n-1}(p)}{J_n(p)} = n(\epsilon_1-\epsilon_2)(\mu_2-\mu_1) + \frac{\epsilon_2\mu_2}{n-1}p^2.$$

In the special case $\mu_1 = \mu_2$, this equation becomes

$$\frac{J_{n-1}(p)}{pJ_n(p)} = \frac{\epsilon_2}{(n-1)(\epsilon_1+\epsilon_2)}.$$

If the dielectric constant of the guide is very much higher than that of the surrounding medium, the first few roots of the above equation are in the vicinities of the zeros of $J_{n-1}(p)$. As q increases indefinitely equation (8) becomes

$$\frac{J_{n-1}(p)J_{n+1}(p)}{J_n^2(p)} = 0. \tag{20-11}$$

Thus in the limit the roots of (8) are exactly equal to the zeros of $J_{n-1}(p)$. In other words as q varies from 0 to ∞ p does not change much in any assigned transmission mode. From (11) it might appear that the limiting values of p could be the zeros of $J_{n+1}(p)$; but this is impossible because in the process of transition p would have to pass through the intermediate zero of $J_n(p)$ and no value of q is consistent with such zero.

The case $n = 1$ requires special examination; in this case equation (8) may be written

$$\frac{\mu_1\epsilon_1 J_0(p)J_2(p)}{p^2J_1^2(p)} + \frac{(\mu_1\epsilon_2+\mu_2\epsilon_1)J_1'(p)}{pJ_1(p)}\left[\frac{1}{q^2}+\frac{K_0(q)}{qK_1(q)}\right]$$
$$-\frac{\mu_2\epsilon_2 K_0(q)}{K_1^2(q)}\left[\frac{K_0(q)}{q^2}+\frac{2K_1(q)}{q^3}\right] = -\frac{\mu_1\epsilon_1+\mu_2\epsilon_2}{p^2q^2}.$$

Multiplying by q^2 and letting q approach zero, we have

$$\frac{(\mu_1\epsilon_2 + \mu_2\epsilon_1)J_1'(p)}{pJ_1(p)} = 2\mu_2\epsilon_2 K_0(q) - \frac{\mu_1\epsilon_1 + \mu_2\epsilon_2}{p^2}.$$

Hence as q approaches zero, p should approach a zero of $J_1(p)$. The smallest zero is $p = 0$; in this case the above equation gives approximately

$$p = \sqrt{\frac{(\mu_1 + \mu_2)(\epsilon_1 + \epsilon_2)}{2\mu_2\epsilon_2 K_0(q)}},$$

for sufficiently small values of q. Thus in the case $n = 1$, the lowest cut-off frequency is theoretically zero.

A study of forced waves in a dielectric wire could be made by assuming an electric current element or an antenna either in the wire itself or outside it. The resulting expressions for the field are analogous to the corresponding expressions for the interior of the metal tube, except that in the present case the integrands are multiple-valued functions. In consequence of this property the total field will consist of natural plane waves corresponding to the poles of the integrand and of a residual spherical field.

10.21. *Waves over a Dielectric Plate*

The principal reason for considering wave propagation over a dielectric plate is to contrast it with wave propagation at the interface of two semi-infinite media. The solution of the latter problem should give us information regarding wave propagation over the earth in any region so restricted that the curvature of the earth may be neglected. Thus let us assume an electric current element or an antenna perpendicular to the plate (Fig. 10.21). The field of an element at the origin is given by (15–5); that of an element at height h above the xy-plane is obtained if we multiply the integrands of (15–5) by $e^{-\gamma h}$; and the field for any linear distribution of elements may then be calculated by integrating the latter field with respect to h. Thus in general the field impressed on the upper surface of the plate, assumed to be the xy-plane, is

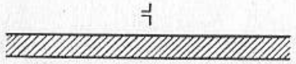


FIG. 10.21. Dielectric plate and antenna above it.

$$E_z^i = \int_{(C_5)} A(\gamma)K_0(i\chi\rho)e^{\gamma z}\,d\gamma, \quad \chi = \sqrt{\gamma^2 + \beta^2},$$

$$E_\rho^i = -i\int_{(C_5)} \frac{\gamma}{\chi} A(\gamma)K_1(i\chi\rho)e^{\gamma z}\,d\gamma, \qquad (21\text{–}1)$$

$$H_\varphi^i = -\omega\epsilon \int_{(C_5)} \frac{1}{\chi} A(\gamma)K_1(i\chi\rho)e^{\gamma z}\,d\gamma,$$

where the contour (C_5) is the one shown in Fig. 10.22 and is obtained from (C_4) of Fig. 10.18 in the same way as (C_2) of Fig. 10.16 was obtained from (C_1) of Fig. 10.15. This contour may be used only for values of z such that $z - h < 0$, where h is the height of a typical current element of the source, and it can be transformed into (C_3) of Fig. 10.17 which represents an elementary spherical wave function as a group of plane waves traveling downward with different propagation constants.

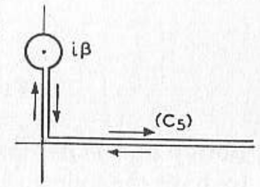


FIG. 10.22.

The wave impedance in the z-direction of each elementary wave is $K = \gamma/i\omega\epsilon$. Hence if the impedance looking into the plate is Z, the reflection coefficients for the various field intensities and the reflected field are

$$q_H = q_{E_z} = q, \quad q_{E_\rho} = -q, \quad q = \frac{K - Z}{K + Z} = \frac{\gamma - i\omega\epsilon Z}{\gamma + i\omega\epsilon Z}.$$

$$(21\text{–}2)$$

$$E_z^r = \int_{(C_5)} qA(\gamma)K_0(i\chi\rho)e^{-\gamma z}\,d\gamma, \quad E_\rho^r = i\int_{(C_5)} q\frac{\gamma}{\chi}A(\gamma)K_1(i\chi\rho)e^{-\gamma z}\,d\gamma,$$

$$H_\varphi^r = -\omega\epsilon\int_{(C_5)} \frac{q}{\chi}A(\gamma)K_1(i\chi\rho)e^{-\gamma z}\,d\gamma.$$

In the dielectric plate the propagation constant $\hat{\gamma}$ and the wave impedance in the z-direction are given by

$$\hat{\gamma} = \sqrt{\chi^2 - \hat{\beta}^2}, \quad \hat{K} = \frac{\hat{\gamma}}{i\omega\hat{\epsilon}}. \qquad (21\text{–}3)$$

Hence the impedance Z normal to the plate is

$$Z = \hat{K}\frac{K\cosh\hat{\gamma}l + \hat{K}\sinh\hat{\gamma}l}{\hat{K}\cosh\hat{\gamma}l + K\sinh\hat{\gamma}l}, \qquad (21\text{–}4)$$

where l is the thickness of the plate.

The natural waves in the dielectric plate and the poles of the integrands in (2) are the roots of

$$q = \infty, \quad K + Z = 0. \qquad (21\text{–}5)$$

Substituting from (4), we have

$$\frac{K\cosh\hat{\gamma}l + \hat{K}\sinh\hat{\gamma}l}{\hat{K}\cosh\hat{\gamma}l + K\sinh\hat{\gamma}l} = -\frac{K}{\hat{K}}, \quad \tanh\hat{\gamma}l = -\frac{2K\hat{K}}{K^2 + \hat{K}^2}. \qquad (21\text{–}6)$$

In order to solve these equations we make the following substitutions

$$\hat{\gamma}l = il\sqrt{\hat{\beta}^2 - \chi^2} = i\hat{p}, \quad \gamma l = l\sqrt{\chi^2 - \beta^2} = p,$$

Substituting in (6) and solving we have

$$\tan \hat{p} = \frac{2p\hat{p}}{\frac{\epsilon}{\hat{\epsilon}}\hat{p}^2 - \frac{\hat{\epsilon}}{\epsilon}p^2},$$

$$p_1 = \frac{\epsilon}{\hat{\epsilon}}\hat{p}\tan\frac{\hat{p}}{2}, \quad p_2 = -\frac{\epsilon}{\hat{\epsilon}}\hat{p}\cot\frac{\hat{p}}{2}. \qquad (21\text{–}7)$$

Both p and $\hat{p}$ are positive, and for each such pair of values satisfying (7), we have the following values of the frequency and the radial phase constant

$$\omega = \frac{\sqrt{\hat{p}^2 + p^2}}{l\sqrt{\hat{\mu}\hat{\epsilon} - \mu\epsilon}}, \quad \chi = \sqrt{\omega^2\mu\epsilon + \frac{p^2}{l^2}} = \sqrt{\omega^2\hat{\mu}\hat{\epsilon} - \frac{\hat{p}^2}{l^2}}.$$

Natural waves in dielectric plates are cylindrical and in nondissipative plates the amplitudes of their field intensities vary inversely as the square root of the distance. On the other hand, in free space, waves generated by sources in a finite region are spherical and the amplitudes of these waves vary inversely as the distance. Some energy is abstracted from the sources and guided in directions parallel to the plate. The plate acts analogously to parallel wires or dielectric wires which tend to guide waves by converting portions of spherical waves into plane waves. There is this difference between plane waves in free space and those in presence of wave guides; the former carry finite power per unit area (of an equiphase plane) and infinite total power, while the latter carry finite total power. Similarly, cylindrical waves in free space carry finite power per unit length (along the axis of the waves) and infinite total power, while in the presence of dielectric plates there may exist cylindrical waves carrying finite total power. Hence in free space no system of sources in a finite region can possibly generate either plane or cylindrical waves but they may do so in presence of "wave guides." Cylindrical guided waves conform to the physical idea of *surface waves* introduced by Zenneck, Sommerfeld, and other early writers on this subject.

In the case of a single interface between two semi-infinite media we have $Z = \hat{K}$ and the condition (5) for the existence of natural waves becomes

$$K + \hat{K} = 0. \qquad (21\text{–}8)$$

For transverse magnetic waves the characteristic wave impedance is either a resistance or a negative reactance; hence, this equation can have no roots, the integrands can have no poles, and there are no surface waves. This conclusion is contrary to that reached by early writers on the subject. Inadvertently the condition for matching impedances was substituted for the equation of natural waves. The impedance concept has only recently been introduced into field theory and no intuitive check on formal manipulations was previously available. Equation (8) for the poles was usually written in its explicit form

$$\frac{\sqrt{\chi^2 - \beta^2}}{\epsilon} + \frac{\sqrt{\chi^2 - \hat{\beta}^2}}{\hat{\epsilon}} = 0,$$

and then was rationalized and solved

$$\frac{\chi^2 - \beta^2}{\epsilon^2} = \frac{\chi^2 - \hat{\beta}^2}{\hat{\epsilon}^2}, \quad \chi = \omega\sqrt{\frac{\epsilon\hat{\epsilon}(\mu\hat{\epsilon} - \hat{\mu}\epsilon)}{\hat{\epsilon}^2 - \epsilon^2}}.$$

However this value of χ is a solution of $K - \hat{K} = 0$ and not of equation (8).

10.22. *Waves over a Plane Earth*

In the actual problem of wave propagation over the earth the conductivity of the earth plays an important role and the propagation constant $\hat{\gamma}$ in the direction normal to the ground surface is $\hat{\gamma} = \sqrt{\chi^2 + \hat{\sigma}^2}$ instead of (21–3). Assuming an electric current element of moment Il at height h above ground (Fig. 10.23) as our source, we find from (15–5) the following expression for $A(\gamma)$ in (21–1)

$$A(\gamma) = -\frac{Il\chi^2 e^{-\gamma h}}{4\pi^2\omega\epsilon}.$$

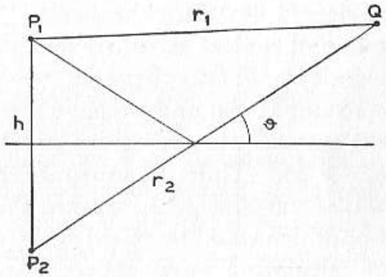


FIG. 10.23. Electric current element at P_1 over a plane earth and image element at P_2.

The impedance normal to the ground is

$$\hat{K} = \frac{\hat{\gamma}}{\hat{g} + i\omega\hat{\epsilon}} = \frac{\sqrt{\gamma^2 + \beta^2 + \hat{\sigma}^2}}{\hat{g} + i\omega\hat{\epsilon}}.$$

The intrinsic propagation constant $\hat{\sigma}$ of the ground is greater than that of the air above, especially in the frequency range in which the ground is a quasiconductor (see section 4.9). When $\hat{\sigma}$ is large, then we have approximately $\hat{K} = \hat{\eta}$ over a substantial part of the path of integration. In this case the ground behaves like an impedance sheet whose surface impedance is equal to the intrinsic impedance of the ground.

The incident and reflected components of E_z in the present case are

$$E_z^i = -\frac{Il}{4\pi^2\omega\epsilon}\int_{(C_4)} (\gamma^2 + \beta^2)K_0(i\rho\sqrt{\gamma^2 + \beta^2})e^{\gamma(z-h)}\,d\gamma,$$

$$E_z^r = -\frac{Il}{4\pi^2\omega\epsilon}\int_{(C_4)} q(\gamma)(\gamma^2 + \beta^2)K_0(i\rho\sqrt{\gamma^2 + \beta^2})e^{-\gamma(z+h)}\,d\gamma,$$

where (C_4) is shown in Fig. 10.18.

The contour (C_5) was used in the case of the dielectric plate to ensure that the incident field would be convergent over the entire surface of the plate (including $\rho = 0$); but there is nothing to prevent us from deforming this contour back to (C_4). In fact the following physical considerations show that it would have been permissible to use (C_4) to begin with. Consider E_z^i along the real axis of (C_4); this part of the integral represents cylindrical waves traveling parallel to the ground surface with velocities less than the characteristic velocity (since the radial phase constant $\sqrt{\gamma^2 + \beta^2}$ is greater than β). Along the imaginary axis from $\gamma = 0$ to $\gamma = i\beta$ we may write

$$\gamma = i\beta \cos \vartheta, \quad \sqrt{\gamma^2 + \beta^2} = \beta \sin \vartheta,$$

where ϑ lies in the interval $(0,\pi/2)$. The equation of the equiphase surfaces is then*

$$\rho \cos \vartheta + (z - h) \sin \vartheta = \text{constant}.$$

The corresponding wave normals make an angle ϑ with the ρ-lines and an angle $\pi/2 + \vartheta$ with the z-axis. Thus this part of the integral represents a group of conical waves traveling toward the ground with the characteristic velocity. Finally between $\gamma = i\beta$ and $\gamma = i\infty$ the radial propagation constant is real and we have a group of plane waves traveling normally to the ground with velocities smaller than the characteristic velocity (since $\gamma/i > \beta$). Thus the integral represents the spherical wave as a group of waves traveling either toward the ground or parallel to it. The only question which could be raised against using this group of waves for the purpose of calculating waves reflected from the ground would concern that wave group which is traveling parallel to the ground. It is reasonable to suppose, however, that the reflection coefficient which applies for waves at near grazing incidence would also apply in the limit to waves at grazing incidence. It was to remove the above mentioned objection that contour (C_5) was chosen. The original contour (C) is not permissible in the present case since the corresponding integral includes a group of waves traveling from the ground as well as toward it.

This care in the choice of the correct contour of integration for a particular problem is essential. Thus for the reflected wave in the present problem the reflection coefficient $q(\gamma)$ adds two branch points at $\gamma = \pm\sqrt{-\beta^2 - \hat{\sigma}^2}$ to the integrand in addition to those at $\gamma = \pm i\beta$ already present in the incident wave. The new branch points are in the second and fourth quadrants; and because of the branch point in the fourth quadrant contour (C) would give a different value for the reflected fields

* Except in the vicinity of the z-axis where the equiphase surfaces depart from the conical shape.

from the one given by (C_4) or (C_5). If the ground conductivity is permitted to approach zero one of the branch points of $q(\gamma)$ will approach the positive real axis from below and the contours (C_4) and (C_5) must be indented.

Let us now consider the incident and reflected fields in greater detail. In order to simplify the discussion we shall introduce the corresponding stream functions. In accordance with (15–5) these functions are obtained if we multiply the integrands of E_z^i and E_z^r by $i\omega\epsilon/(\gamma^2 + \beta^2)$; thus

$$\Pi^i = \frac{Il}{4i\pi^2} \int_{(C_4)} K_0(i\rho\sqrt{\gamma^2 + \beta^2}) e^{\gamma(z-h)}\, d\gamma,$$

$$\Pi^r = \frac{Il}{4i\pi^2} \int_{(C_4)} q(\gamma) K_0(i\rho\sqrt{\gamma^2 + \beta^2}) e^{-\gamma(z+h)}\, d\gamma.$$

If the reflection coefficient $q(\gamma)$ were independent of γ, the reflected field would be equal to the field of an image current element of moment qIl (Fig. 10.23)

$$\Pi^r = \frac{qIle^{-i\beta r_2}}{4\pi r_2},$$

where the reflection coefficient q then corresponds to plane waves whose wave normals make the angle ϑ with the ground plane.

Consider next the field along some particular line P_2Q whose equation is $(z + h) \cos \vartheta - \rho \sin \vartheta = 0$. For the distance r_2 measured along this line from P_2, we have $r_2 = (z + h) \sin \vartheta + \rho \cos \vartheta$. A pencil of elementary waves traveling in directions making an angle ϑ with the ground plane corresponds to the values of γ in the vicinity of $\gamma = \bar{\gamma} = i\beta \sin \vartheta$. Let us now expand the reflection coefficient in the following power series

$$q(\gamma) = q(\bar{\gamma}) + (\gamma - \bar{\gamma})q'(\bar{\gamma}) + \tfrac{1}{2}(\gamma - \bar{\gamma})^2 q''(\bar{\gamma}) + \cdots$$

and substitute in Π^r. Then we have

$$\Pi^r = \Pi_0^r + \Pi_1^r + \Pi_2^r + \cdots,$$

$$\Pi_0^r = \frac{q(i\beta \sin \vartheta)Ile^{-i\beta r_2}}{4\pi r_2},$$

$$\Pi_1^r = \frac{q'(\bar{\gamma})Il}{4i\pi^2} \int_{(C_4)} (\gamma - \bar{\gamma}) K_0(i\rho\sqrt{\gamma^2 + \beta^2}) e^{-\gamma(z+h)}\, d\gamma,$$

$$\Pi_2^r = \frac{q''(\bar{\gamma})Il}{8i\pi^2} \int_{(C_4)} (\gamma - \bar{\gamma})^2 K_0(i\rho\sqrt{\gamma^2 + \beta^2}) e^{-\gamma(z+h)}\, d\gamma.$$

The first term Π_0^r represents the reflected wave that would be obtained from the incident wave by assuming that the latter is reflected as if it were a uniform wave traveling in the direction making the angle ϑ with the ground plane. In order to compute the following terms we take

$$\hat{\Pi} = \frac{Il e^{-i\beta r_2}}{4\pi r_2} = \frac{Il}{4i\pi^2}\int_{(C_4)} K_0(i\rho\sqrt{\gamma^2+\beta^2})e^{-\gamma(z+h)}\,d\gamma,$$

and differentiate it with respect to z to obtain

$$\frac{\partial\hat{\Pi}}{\partial z} = -\frac{Il}{4i\pi^2}\int_{(C_4)} \gamma K_0(i\rho\sqrt{\gamma^2+\beta^2})e^{-\gamma(z+h)}\,d\gamma,$$

and

$$\Pi_1^r = -q'(\bar{\gamma})\left(\frac{\partial\hat{\Pi}}{\partial z} + \bar{\gamma}\hat{\Pi}\right), \quad \Pi_2^r = \tfrac{1}{2}q''(\bar{\gamma})\left(\frac{\partial}{\partial z}+\bar{\gamma}\right)^2\hat{\Pi},$$

$$\Pi_n^r = \frac{(-1)^n q^{(n)}(\bar{\gamma})}{n!}\left(\frac{\partial}{\partial z}+\bar{\gamma}\right)^n\hat{\Pi}.$$

Next we evaluate $\partial\hat{\Pi}/\partial z$ and substitute in Π_1^r

$$\frac{\partial\hat{\Pi}}{\partial z} = \frac{d\hat{\Pi}}{dr_2}\frac{\partial r_2}{\partial z} = \sin\vartheta\frac{d\hat{\Pi}}{dr_2} = -\frac{i\beta\sin\vartheta Il e^{-i\beta r_2}}{4\pi r_2}\left(1+\frac{1}{i\beta r_2}\right),$$

$$\Pi_1^r = \frac{q'(i\beta\sin\vartheta)\sin\vartheta Il e^{-i\beta r_2}}{4\pi r_2^2}.$$

The amplitude of this term of the reflected Π^r varies inversely as the square of the distance from the image source while the amplitude of the first term varies inversely as the first power of the distance.

The reflection coefficient is

$$q = \frac{K-\hat{K}}{K+\hat{K}} = \frac{(\hat{g}+i\omega\hat{\epsilon})\gamma - i\omega\epsilon\sqrt{\gamma^2+\beta^2+\hat{\sigma}^2}}{(\hat{g}+i\omega\hat{\epsilon})\gamma + i\omega\epsilon\sqrt{\gamma^2+\beta^2+\hat{\sigma}^2}}.$$

When $\hat{\sigma}$ is sufficiently large, then we have approximately

$$q = \frac{\gamma - i\omega\epsilon\hat{\eta}}{\gamma+i\omega\epsilon\hat{\eta}} = 1 - \frac{2i\omega\epsilon\hat{\eta}}{\gamma+i\omega\epsilon\hat{\eta}} = -1 + \frac{2\gamma}{\gamma+i\omega\epsilon\hat{\eta}}.$$

When ϑ is small, $q(\bar{\gamma})$ is nearly equal to -1. In the ground plane, where $r_1 = r_2$, we have then

$$\Pi^i + \Pi_0^r = \frac{2\eta\sin\vartheta}{\hat{\eta}}\frac{Il e^{-i\beta r_1}}{4\pi r_1} = \frac{\eta h}{\hat{\eta}}\frac{Il e^{-i\beta r_1}}{2\pi r_1^2}.$$

Since $q'(\gamma) \simeq 2/i\omega\epsilon\hat{\eta}$, we have

$$\Pi_1^r \simeq \frac{hIl e^{-i\beta r_1}}{2\pi i\omega\epsilon\hat{\eta}r_1^3}.$$

In this instance the total field is approximately proportional to the height of the source above ground and inversely proportional to the square of the distance from it. However, the total reflected field does not vanish

with h; Π_2^r contains a term independent of h and varies as the cube of the distance from the source.

The above expressions for the components of Π^r are asymptotic and, taking into consideration the physical picture to which they correspond, it may be expected that they are more suitable for numerical calculations when either the transmitter or the receiver or both are fairly high above ground than when they are both near the earth's surface. A practical rule to follow is to use the above expressions in that range of the variables for which the first and second terms are much larger than the succeeding terms.

10.23. *Wave Propagation between Concentric Spheres*

The curvature effect on wave propagation has already been considered in section 8.23 where several problems concerning waves in wave guides bent into toroids have been formulated exactly and then solved by approximate methods. We shall now consider another problem in which bending takes place in two perpendicular planes.

A study of wave propagation between two concentric spheres of nearly equal radii (Fig. 10.24) should give us an indication of the magnitude of the curvature effect on the propagation of cylindrical waves between parallel planes. We have seen that in the latter case the principal wave is a uniform cylindrical wave for which the electric lines are straight lines normal to the planes and the magnetic lines circles coaxial with the axis of the wave. The field of the corresponding wave between two concentric spheres should then be independent of φ; and the electric lines should be approximately radial. These lines cannot be exactly radial since the fundamental electromagnetic equations are inconsistent with the assumption of a field whose components other than E_r and H_φ vanish. These equations show that E_θ cannot vanish identically. Nevertheless, on the grounds of continuity, we expect that for the transmission mode under consideration E_θ should be small compared with E_r.

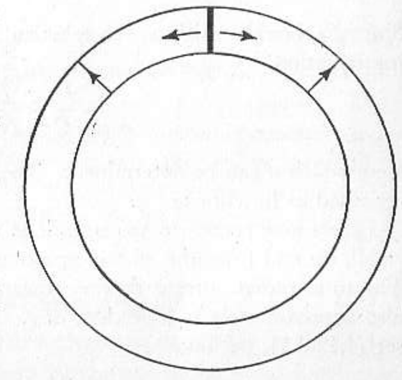


FIG. 10.24. Wave propagation between concentric spheres.

The appropriate solutions can be obtained from the general expressions in section 10. Thus if Π is independent of φ, we have

$$\Pi(r,\theta) = T(\theta)\hat{T}(r), \tag{23–1}$$

where $T(\theta)$ and $\hat{T}(r)$ are solutions of

$$\frac{d}{d\theta}\left(\sin\theta\frac{dT}{d\theta}\right) = -n(n+1)\sin\theta\, T, \tag{23–2}$$

$$\frac{d^2\hat{T}}{dr^2} = \left[-\beta^2 + \frac{n(n+1)}{r^2}\right]\hat{T}. \tag{23–3}$$

Equation (2) is the Legendre equation and its general solution is

$$T(\theta) = AP_n(-\cos\theta) + BP_n(\cos\theta). \tag{23-4}$$

In general when n is not an integer the first term becomes infinite at $\theta = 0$ and the second term is infinite at $\theta = \pi$. Hence if the source of the waves is along the axis, $\theta = 0$, we must have $B = 0$ and (4) becomes

$$T(\theta) = AP_n(-\cos\theta). \tag{23-5}$$

The general solution of (3) in terms of the functions best suited to waves in nondissipative media bounded by conducting spheres is

$$\hat{T}(r) = C\hat{J}_n(\beta r) + D\hat{N}_n(\beta r), \tag{23-6}$$

where $\hat{J}_n$ and $\hat{N}_n$ are defined in section 3.5. From Π we obtain the field as in section 10; thus

$$E_r = \frac{n(n+1)}{i\omega\epsilon r^2}\hat{T}(r)P_n(-\cos\theta), \quad rH_\varphi = -\hat{T}(r)\frac{d}{d\theta}P_n(-\cos\theta),$$
$$rE_\theta = \frac{1}{i\omega\epsilon}\frac{d\hat{T}}{dr}\frac{d}{d\theta}P_n(-\cos\theta). \tag{23-7}$$

Since E_θ should vanish on the spherical surfaces $r = a$ and $r = b$, we have the following equation

$$-\frac{D}{C} = \frac{\hat{J}_n'(\beta a)}{\hat{N}_n'(\beta a)} = \frac{\hat{J}_n'(\beta b)}{\hat{N}_n'(\beta b)} \tag{23-8}$$

from which n can be determined. The solution depends on properties of $\hat{J}_n$ and $\hat{N}_n$ regarded as functions of n.

Let us now return to the equations (4.12–11) for our field. Assuming that E_θ is small, we find from the second equation that rH_φ is approximately independent of r. The total radial current flowing toward the center, $I(\theta) = -2\pi r \sin\theta\, H_\varphi$, is then also approximately independent of r. Introducing $I(\theta)$ in the first equation of the set (4.12–11), we have

$$\frac{1}{r^2}\frac{\partial I}{\partial\theta} = -2\pi i\omega\epsilon \sin\theta\, E_r.$$

Integrating along the radius and introducing the transverse voltage $V(\theta)$ we obtain

$$\frac{\partial I}{\partial\theta} = -i\omega\bar{C}V, \quad \bar{C} = \frac{2\pi\epsilon ab\sin\theta}{b-a}, \quad V(\theta) = \int_a^b E_r\,dr, \tag{23-9}$$

where $\bar{C}$ is the capacity per radian. If $\bar{C}$ is integrated from $\theta = 0$ to $\theta = \pi$, the total capacity between the two spheres is found to agree with the exact value at zero frequency. The second transmission equation is obtained from the third equation of the set (4.12–11); thus introducing $I(\theta)$ and integrating from $r = a$ to $r = b$, we have

$$\frac{\partial V}{\partial\theta} = -i\omega\bar{L}I, \quad \bar{L} = \frac{\mu(b-a)}{2\pi\sin\theta}. \tag{23-10}$$

Replacing the angle θ by the distance $s = c\theta$ along the circumference of the circle of radius $c = \sqrt{ab}$ in the transmission equations, we have

$$\frac{\partial V}{\partial s} = -i\omega LI, \quad \frac{\partial I}{\partial s} = -i\omega CV,$$
$$L = \frac{\mu(b-a)}{2\pi c\sin\theta}, \quad C = \frac{2\pi\epsilon c\sin\theta}{b-a}. \tag{23-11}$$

If E_θ is eliminated from the second and third equations of the set (4.12–11), we obtain

$$\frac{\partial E_r}{\partial\theta} = i\omega\mu(rH_\varphi) - \frac{1}{i\omega\epsilon}\frac{\partial^2}{\partial r^2}(rH_\varphi). \tag{23-12}$$

For each transmission mode rH_φ is proportional to $\hat{T}$ and therefore satisfies equation (3). Hence (12) becomes

$$\frac{\partial E_r}{\partial\theta} = -\frac{n(n+1)}{i\omega\epsilon r}H_\varphi.$$

Introducing $I(\theta)$ into this equation and into the first equation of the set (4.12–11), we have

$$\frac{\partial}{\partial\theta}(r^2E_r) = -i\frac{n(n+1)}{2\pi\omega\epsilon\sin\theta}I, \quad \frac{\partial I}{\partial\theta} = -2\pi i\omega\epsilon\sin\theta(r^2E_r).$$

For any fixed value of r, E_r and I and hence V and I satisfy nonuniform transmission line equations. For the principal transmission mode as r approaches infinity the series reactance tends to become an inductance independent of the frequency; for finite values of r this inductance is slightly modified by the longitudinal displacement currents proportional to E_θ.

10.24. *Natural Oscillations in Cylindrical Cavity Resonators*

A cylindrical cavity resonator of arbitrary cross-section may be regarded as a wave guide. Assuming that the generators of the cylindrical boundary are parallel to the z-axis and that the flat faces are $z = 0$ and $z = l$, we may derive the fields of various natural oscillations from the following two wave functions

$$\Pi = T(x,y)\cos\frac{p\pi z}{l}, \quad p = 0,1,2,\cdots, \tag{24-1}$$

$$\Psi = T(x,y)\sin\frac{p\pi z}{l}, \quad p = 1,2,3,\cdots. \tag{24-2}$$

Since

$$\chi^2 + \frac{p^2\pi^2}{l^2} = \beta^2 = \frac{\omega^2}{v^2} = \frac{4\pi^2}{\lambda^2}, \tag{24-3}$$

where χ is the transverse phase constant, we have the following expressions for the natural frequencies and wavelengths

$$\omega = v\sqrt{\chi^2 + \frac{p^2\pi^2}{l^2}}, \quad \lambda = \frac{2\pi}{\sqrt{\chi^2 + \dfrac{p^2\pi^2}{l^2}}}. \tag{24–4}$$

For transverse magnetic waves in a cavity of rectangular cross-section, using $T(x,y)$ as given in section 5, we have

$$\Pi = \sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}\cos\frac{p\pi z}{l}. \tag{24–5}$$

Hence the natural wavelength of the $TM_{m,n,p}$-oscillation mode is

$$\lambda = \frac{2}{\sqrt{\dfrac{m^2}{a^2} + \dfrac{n^2}{b^2} + \dfrac{p^2}{l^2}}}. \tag{24–6}$$

Similarly the field of the $TE_{m,n,p}$-oscillation mode may be obtained from

$$\Psi = \cos\frac{m\pi x}{a}\cos\frac{n\pi y}{b}\sin\frac{p\pi z}{l}. \tag{24–7}$$

The natural wavelength is again given by equation (6). In the present case the designations TM and TE are arbitrary since the same parallelopipedal resonator can be regarded as a section of three different wave guides and the same oscillation mode may be obtained from transverse magnetic waves in one of these guides and from transverse electric waves in another.

In the case of cylindrical resonators of circular cross-section the natural wavelength is given by (4), where $\chi = k_{n,m}/a$ and $k_{n,m}$ is the mth nonvanishing zero of $J_n(x)$ or $J_n'(x)$ according as the oscillations correspond to waves of transverse magnetic or transverse electric type.

For each oscillation mode the energy stored in the resonator may be obtained either by calculating the magnetic energy at the instant when the magnetic field is maximum and the electric field zero or by calculating the electric energy at the instant when the electric field is maximum and the magnetic field zero. For oscillations of TM-type we choose the former method and find

$$W = \tfrac{1}{4}\chi^2\mu l \iint_{(S)} T^2\,dS, \quad \text{if } p \neq 0,$$

$$= \tfrac{1}{2}\chi^2\mu l \iint_{(S)} T^2\,dS, \quad \text{if } p = 0.$$

For oscillations of TE-type we choose the second method and obtain

$$W = \tfrac{1}{4}\chi^2\epsilon l \iint T^2\,dS.$$

For oscillations of TM-type the conduction current in the cylindrical part of the resonator is parallel to the axis and its density is $(\partial T/\partial n)\cos p\pi z/l$, where the derivative is taken in the direction of the outward normal. In the faces $z = 0$ and $z = l$ the conduction current density is respectively grad T and $(-)^{p+1}$ grad T. Hence the power absorbed by the walls of the resonator is

$$\mathscr{W} = \mathscr{W}_1 + \mathscr{W}_2, \quad \text{where} \quad \mathscr{W}_2 = \chi^2\mathcal{R}\iint_{(S)} T^2\,dS,$$

$$\mathscr{W}_1 = \tfrac{1}{4}\mathcal{R}l\int_{(s)}\left(\frac{\partial T}{\partial n}\right)^2 ds, \quad \text{if } p \neq 0,$$

$$= \tfrac{1}{2}\mathcal{R}l\int_{(s)}\left(\frac{\partial T}{\partial n}\right)^2 ds, \quad \text{if } p = 0.$$

In the case of oscillations of TE type the z component of the current density in the cylindrical wall is $(ip\pi/\eta\beta l)(\partial T/\partial s)\cos p\pi z/l$, where the derivative is taken in the counterclockwise direction. The transverse current density in the counterclockwise direction is $(\chi^2 T/i\omega\mu)\sin p\pi z/l$; and the current densities in the faces $z = 0$ and $z = l$ are respectively

$$\frac{ip\pi}{\eta\beta l}\,\text{grad}\,T \text{ and } \frac{(-)^{p+1}ip\pi}{\eta\beta l}\,\text{grad}\,T.$$

Hence the power absorbed by the walls is

$$\mathscr{W} = \mathscr{W}_1 + \mathscr{W}_2, \quad \text{where} \quad \mathscr{W}_2 = \frac{\mathcal{R}\chi^2p^2\pi^2}{\eta^2\beta^2l^2}\iint_{(S)} T^2\,dS,$$

$$\mathscr{W}_1 = \frac{\mathcal{R}p^2\pi^2}{4\eta^2\beta^2 l}\left[\int_{(s)}\left(\frac{\partial T}{\partial s}\right)^2 ds + \frac{\chi^4 l^2}{p^2\pi^2}\int_{(s)} T^2\,ds\right].$$

From the above expressions for the stored energy and the dissipated power we can express the Q of the resonator in a general form. Thus in the case of TM-oscillations we have

$$Q = \frac{\omega\mu\iint_{(S)} T^2\,dS}{\mathcal{R}\left[\chi^{-2}\int_{(s)}\left(\dfrac{\partial T}{\partial n}\right)^2 ds + 4l^{-1}\iint_{(S)} T^2\,dS\right]}, \quad \text{if } p \neq 0,$$

$$Q = \frac{\omega\mu \iint_{(S)} T^2\, dS}{\mathcal{R}\left[\chi^{-2}\int_{(s)}\left(\frac{\partial T}{\partial n}\right)^2 ds + 2l^{-1}\iint_{(S)} T^2\, dS\right]}, \quad \text{if } p = 0;$$

and in the case of TE-oscillations

$$Q = \frac{\omega\mu\chi^2\beta^2 \iint_{(S)} T^2\, dS}{\mathcal{R}\left[\chi^4\int_{(s)} T^2\, ds + p^2\pi^2 l^{-2}\int_{(s)}\left(\frac{\partial T}{\partial s}\right)^2 ds + 4\chi^2 p^2\pi^2 l^{-3}\iint_{(S)} T^2\, dS\right]}.$$

For a circular cylinder of radius a and length l the Q of the TM-oscillations becomes

$$Q = \frac{\omega\mu a}{2\mathcal{R}\left(1 + \frac{2a}{h}\right)}, \quad \text{if } p \neq 0; \quad Q = \frac{\omega\mu a}{2\mathcal{R}\left(1 + \frac{a}{h}\right)}, \quad \text{if } p = 0.$$

For $TE_{n,m,p}$-oscillations we have

$$Q = \frac{\omega\mu a(k_{n,m}^2 l^2 + p^2\pi^2 a^2)(k_{n,m}^2 - n^2)}{2\mathcal{R}[p^2 n^2\pi^2 a^2 + k_{n,m}^4 l^2 + 2p^2\pi^2 a^3 l^{-1}(k_{n,m}^2 - n^2)]}.$$

In the above calculations it has been assumed that there is no dielectric loss in the resonator. If there is such loss but no loss in the walls of the cavity, then instead of (3) we have

$$\chi^2 + \frac{p^2\pi^2}{l^2} = -\sigma^2 = -u\mu(g + u\epsilon),$$

where u is the oscillation constant. Solving for u, we obtain

$$u = -\frac{g}{2\epsilon} \pm i\sqrt{\frac{1}{\mu\epsilon}\left[\chi^2 + \frac{p^2\pi^2}{l^2} - \left(\tfrac{1}{2}g\sqrt{\frac{\mu}{\epsilon}}\right)^2\right]}.$$

By (5.11–16) we have

$$Q = \frac{\omega\epsilon}{g};$$

that is, the Q of the resonator is equal to the Q of the dielectric.

The Q of the resonator having both metal and dielectric losses may be obtained from the general formula

$$\frac{1}{Q} = \frac{1}{Q_1} + \frac{1}{Q_2}.$$

This formula follows immediately from the definition of Q in terms of stored energy and absorbed power provided the losses are not so large as to affect appreciably the natural frequency.

CHAPTER XI

ANTENNA THEORY

11.1. *Biconical Antenna*

The antenna theory developed in the following pages is based fundamentally on the conception that the antenna and the space surrounding it are two wave guides. Thus in the case of an antenna consisting of two equal coaxial cones (Fig. 11.1) in free space the two wave guides are:

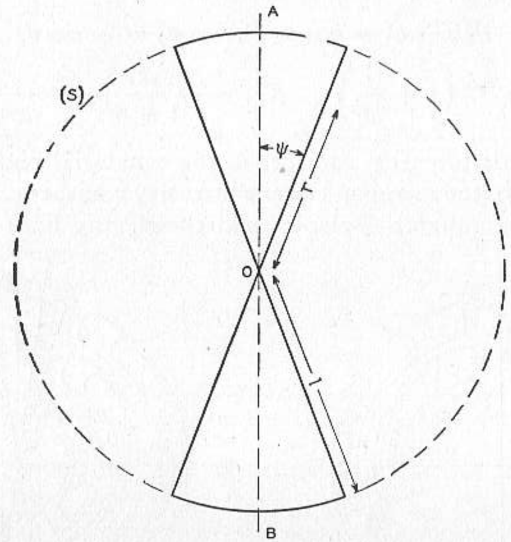


FIG. 11.1. The cross-section of a conical antenna of length l and of the "boundary sphere" S.

(1) the antenna region bounded by the conical surfaces and the boundary sphere (S) concentric with the center of the antenna and passing through the outer ends of the cones, (2) the space external to (S). There are infinitely many transmission modes appropriate to either wave guide. If the voltage is applied between the apices of the cones, then the waves possess circular symmetry; the magnetic lines are circles coaxial with the cones and the electric lines lie in axial planes; but the number of transmission modes is still infinite.

Circular magnetic waves in free space are described by the equations of section 10.10. Because of circular symmetry we use $m = 0$ in the expression (10.10–22) for T; since the field must be finite for all values of θ, n must be an integer; consequently for progressive waves traveling outward in region (2) we have $T(\theta) = P_n(\cos\theta)$ and the field of the nth "zonal wave" is

$$\Pi(r,\theta) = \hat{K}_n(i\beta r)P_n(\cos\theta), \quad rE_\theta = \eta\hat{K}_n'(i\beta r)P_n^1(\cos\theta),$$

$$rH_\varphi = -\hat{K}_n(i\beta r)\frac{d}{d\theta}P_n(\cos\theta) = -\hat{K}_n(i\beta r)P_n^1(\cos\theta), \qquad (1\text{–}1)$$

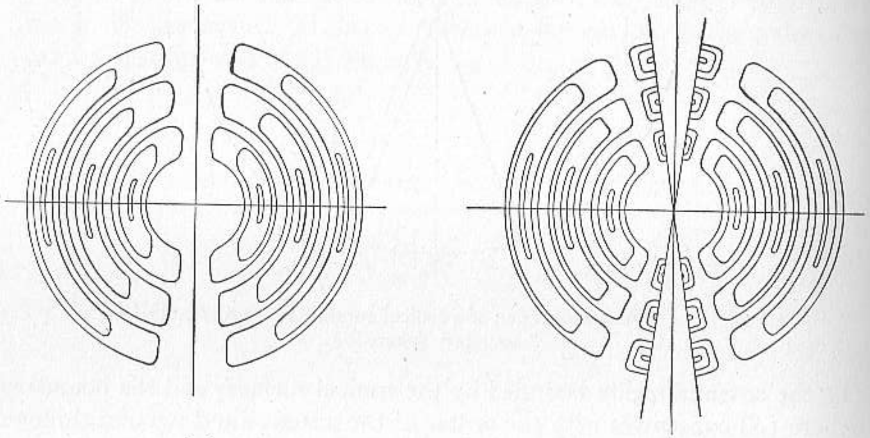
$$i\omega\varepsilon r^2E_r = n(n+1)\hat{K}_n(i\beta r)P_n(\cos\theta).$$

When $n = 0$ the field vanishes identically and the principal or dominant wave corresponds to $n = 1$. For this wave

$$P_1(\cos\theta) = \cos\theta,\ P_1^1(\cos\theta) = -\sin\theta,$$

$$\hat{K}_1(i\beta r) = e^{-i\beta r}\left(1 + \frac{1}{i\beta r}\right), \quad K_{r,1}^+ = \eta\frac{\beta^2r^2}{1+\beta^2r^2} + \frac{\eta}{i\beta r(1+\beta^2r^2)}.$$

The radial electric intensity vanishes in the equatorial plane and is maximum on the axis; the meridian electric intensity vanishes on the axis and is maximum in the equatorial plane; and the electric lines have the form



(a) (b)

FIG. 11.2. Electric lines for the first order transverse magnetic spherical waves: (a) lines in free space; (b) lines in the presence of two coaxial conical conductors.

shown in Fig. 11.2(a). For large values of r the radial impedance is substantially equal to the intrinsic impedance; for small values of r the radial impedance is largely reactive and is approximately equal to $1/i\omega\varepsilon r$. For small values of r the radial displacement current is forced to flow across

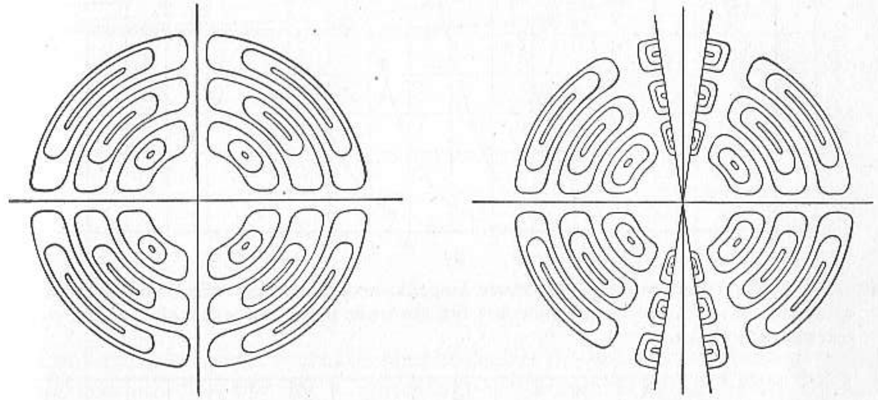
a small area, the series capacitance is small, and the input impedance is large.

For the zonal wave of the second order we have

$$P_2(\cos\theta) = \tfrac{1}{2}(3\cos^2\theta - 1), \quad P_2^1(\cos\theta) = -3\sin\theta\cos\theta,$$

$$\hat{K}_2(i\beta r) = e^{-i\beta r}\left(1 + \frac{3}{i\beta r} - \frac{3}{\beta^2r^2}\right), \quad K_{r,2}^+ = -\eta\frac{\hat{K}_2'(i\beta r)}{\hat{K}_2(i\beta r)}.$$

The radial electric intensity vanishes when $\cos\theta = \pm 1/\sqrt{3}$, while the meridian intensity vanishes on the axis and in the equatorial plane; the electric lines are shown approximately in Fig. 11.3(a). For large values



(a) (b)

FIG. 11.3. Electric lines for the second order transverse magnetic spherical waves: (a) lines in free space; (b) lines in the presence of two coaxial conical conductors.

of r the radial impedance is nearly equal to η and for small values it is nearly $2/i\omega\varepsilon r$.

As the order of the zonal wave increases the number of different sets of closed loops in the field also increases. The radial admittance of the nth zonal wave may be expressed in the following form

$$\eta M_n = \frac{1 - i(u_nu_n' + v_nv_n')}{2 - (u_n^2 + v_n^2) + (u_n'^2 + v_n'^2)},$$

$$u_n = 1 - \frac{(n-1)n(n+1)(n+2)}{2\cdot 4\beta^2r^2} + \frac{(n-3)(n-2)\cdots(n+4)}{2\cdot 4\cdot 6\cdot 8\beta^4r^4} - \cdots,$$

$$v_n = -\frac{n(n+1)}{2\beta r} + \frac{(n-2)(n-1)\cdots(n+3)}{2\cdot 4\cdot 6\beta^3r^3} - \frac{(n-4)(n-3)\cdots(n+5)}{2\cdot 4\cdots 10\beta^5r^5} + \cdots. \qquad (1\text{–}2)$$

Figures 11.4 and 11.5 show respectively the products ηG_n and ηB_n of the intrinsic impedance with the radial conductance and the radial susceptance. For small values of r the ratio of conductance to susceptance diminishes very rapidly with the increasing order of the wave; hence a given average

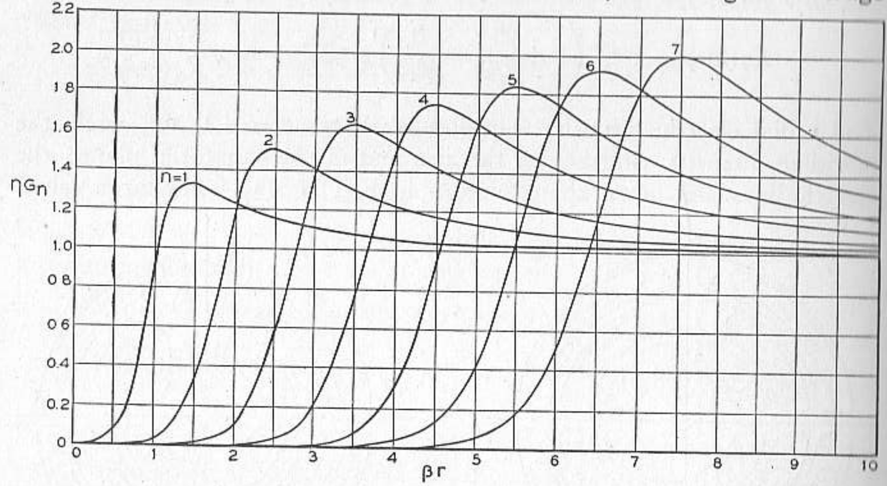


FIG. 11.4. The product of the intrinsic impedance η and the radial conductance G_n as a function of the phase distance βr from the wave origin for each of the first seven transmission modes.

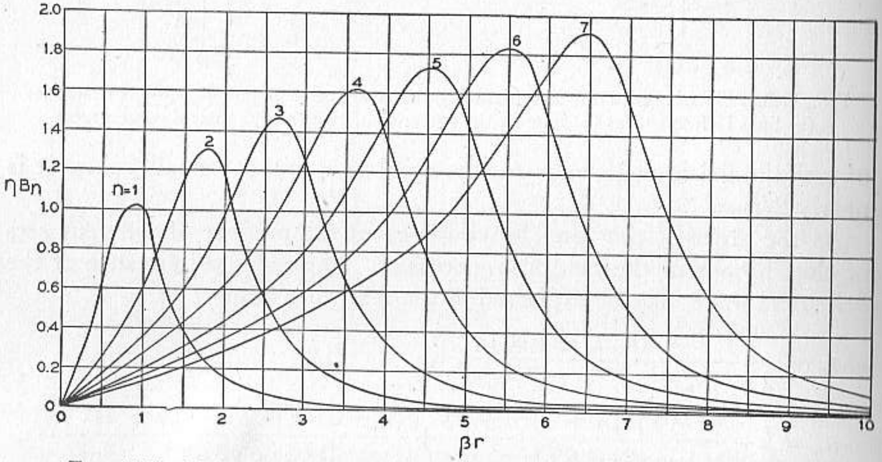


FIG. 11.5. The product of the intrinsic impedance and the radial susceptance.

radial power flow is associated with increasingly strong reactive fields in the vicinity of the center of the wave.

In the antenna region we have corresponding transmission modes except that the electric lines in the vicinity of the conical conductors form small

half-loops terminating on the conductors as shown in Figs. 11.2(b) and 11.3(b). As the cone angle becomes vanishingly small, these added half-loops become vanishingly small and the field configurations become nearly the same as in free space. In addition to these transmission modes there exists a mode in which the electric lines coincide with the meridians (Fig. 11.6); this is the principal mode in the antenna region and its theory has been developed in section 8.12.

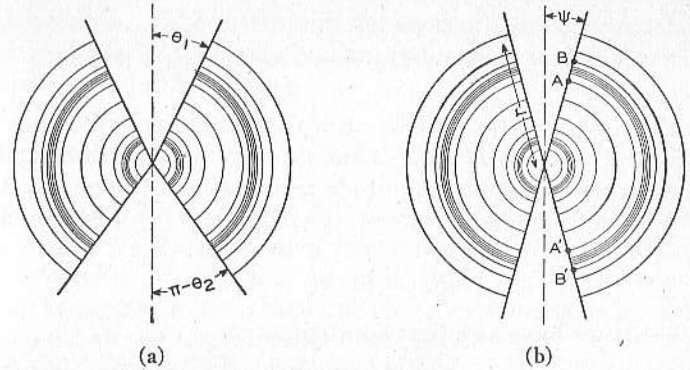


FIG. 11.6. Cross-sections of infinitely long conical conductors and electric lines of force for principal waves.

Let us now consider more closely the higher modes in the antenna region. Since the radii $\theta = 0$ and $\theta = \pi$ are excluded from this region by the conical conductors we have

$$T(\theta) = AP_n(-\cos\theta) + BP_n(\cos\theta).$$

This function is proportional to E_r and must vanish on the conical conductors, assuming that they are perfect conductors. Thus if the cone angle is ψ, then

$$\begin{aligned} AP_n(-\cos\psi) + BP_n(\cos\psi) &= 0, \\ AP_n(\cos\psi) + BP_n(-\cos\psi) &= 0. \end{aligned} \qquad (1\text{–}3)$$

These equations require $A^2 = B^2$, $A = \pm B$. Hence equations (3) are reduced to

$$P_n(-\cos\psi) = P_n(\cos\psi), \quad A = -B; \qquad (1\text{–}4)$$
$$P_n(-\cos\psi) = -P_n(\cos\psi), \quad A = B. \qquad (1\text{–}5)$$

In the case represented by equation (4) we have (setting $B = \frac{1}{2}$)

$$T(\theta) = \tfrac{1}{2}[P_n(\cos\theta) - P_n(-\cos\theta)],$$
$$rH_\varphi = -\hat{T}(\beta r)\frac{dT}{d\theta}, \quad rE_\theta = -i\eta\hat{T}'(\beta r)\frac{dT}{d\theta}, \qquad (1\text{–}6)$$
$$i\omega\epsilon r^2 E_r = n(n+1)\hat{T}(\beta r)T(\theta), \quad \hat{T}(\beta r) = A\hat{J}_n(\beta r) + B\hat{N}_n(\beta r).$$

Replacing θ by $\pi - \theta$ in the above expression for T, we see that

$$T(\pi - \theta) = -T(\theta), \quad \frac{dT(\pi - \theta)}{d(\pi - \theta)} = \frac{dT(\theta)}{d\theta}.$$

Hence, along the radii making angles θ and $\pi - \theta$ with the axis, the radial electric intensities are in opposite directions and the magnetic intensities are in the same direction. The current in the upper cone at distance r from the apex is

$$I(r,\psi) = 2\pi r \sin\psi\, H_\varphi = -2\pi\hat{T}(\beta r)\frac{dT(\psi)}{d\psi}\sin\psi; \tag{1–8}$$

similarly the current in the lower cone, if regarded as positive when flowing upward is $I(r,\pi - \psi) = I(r,\psi)$. Thus the currents at points equidistant from the center are equal in magnitude and flow in the same direction.

If $\frac{1}{2}\psi^2$ is small compared with unity, equation (4) becomes approximately

$$P_n(-\cos\psi) = 1. \tag{1–9}$$

From (3.6–10) we have as a first approximation

$$P_n(-\cos\psi) = \frac{2}{\pi}\sin n\pi\left[\log\frac{\psi}{2} + \psi(n) - \psi(0)\right] + \cos n\pi;$$

hence (9) becomes

$$\tan\frac{n\pi}{2} = -\frac{2}{\pi}\left[\log\frac{2}{\psi} + \psi(0) - \psi(n)\right] \simeq -\frac{2}{\pi}\log\frac{2}{\psi}.$$

Thus the approximate values of n are

$$n = 2m + 1 + \frac{1}{\log\frac{2}{\psi}}, \quad m = 0, 1, 2, \cdots.$$

Since for small cone angles the characteristic impedance of the cone to the principal wave is $K = 120\log 2/\psi$, we may write

$$n = 2m + 1 + \frac{120}{K}. \tag{1–10}$$

Thus as the characteristic impedance tends to infinity, the roots of (4) approach the odd integers. From equations (3.7–43) and (3.7–44) we find that

$$T_{2m+1+\Delta}(\theta) \to P_{2m+1}(\cos\theta) \text{ as } \Delta \to 0. \tag{1–11}$$

Thus all the transmission modes in the antenna region, except the principal, approach the corresponding modes in free space.

In the case represented by equation (5), we have

$$T(\theta) = \tfrac{1}{2}[P_n(\cos\theta) + P_n(-\cos\theta)],$$
$$T(\pi - \theta) = T(\theta), \quad \frac{dT(\pi - \theta)}{d(\pi - \theta)} = -\frac{dT(\theta)}{d\theta}. \tag{1–12}$$

Along the radii making angles θ and $\pi - \theta$ with the axis the radial electric intensities are in the same direction and the magnetic intensities are in opposite directions. The currents in the cones at points equidistant from the center are equal and flow in opposite directions. For small cone angles the approximate solution of (5) is

$$n = 2m + \frac{120}{K}. \tag{1–13}$$

In this case $T_{2m+\Delta}(\theta) \to P_{2m}(\cos\theta)$ as $\Delta \to 0$. The modes for values of $m > 0$ approach the corresponding free space transmission modes. The mode corresponding to $m = 0$ is the principal "anti-symmetric mode" in the antenna region; in free space there is no corresponding mode because $P_0(\cos\theta) = 1$ and the field vanishes identically.

In particular for small values of Δ we have

$$P_\Delta(-\cos\theta) \simeq 1 + 2\Delta\log\sin\frac{\theta}{2},$$

$$P_{1+\Delta}(-\cos\theta) \simeq -\left(1 + 2\Delta\log\sin\frac{\theta}{2}\right)\cos\theta.$$

Figure 11.7 illustrates the behavior of these functions. The corresponding T-functions are

$$T_\Delta(\theta) = 1 + \Delta\log\sin\frac{\theta}{2}, \quad T_{1+\Delta}(\theta) = \left(1 + \Delta\log\sin\frac{\theta}{2}\right)\cos\theta.$$

The behavior of these functions is illustrated in Figs. 11.8 and 11.9.

For $n > 0$ the function $\hat{N}_n(\beta r)$ becomes infinite as r^{-n} at $r = 0$, and its derivative becomes infinite as r^{-n-1}, hence B in (6) must be zero; then $\hat{T}_n(\beta r) = A\hat{J}_n(\beta r)$. In the vicinity of $r = 0$ the function $\hat{J}_n(\beta r)$ is proportional to r^n; thus the current associated with the higher order waves vanishes at the apices of the conical conductors $I_n(0) = 0$. The voltage $V_n(r)$ along a typical meridian vanishes at all distances

$$V_n(r) = \int_\psi^{\pi-\psi} rE_\theta\, d\theta = -i\eta\hat{T}'(\beta r)[T(\pi - \psi) - T(\psi)] = 0.$$

Hence *the input voltage and current depend only on the principal wave.* Consequently the *input impedance depends only on the principal wave.*

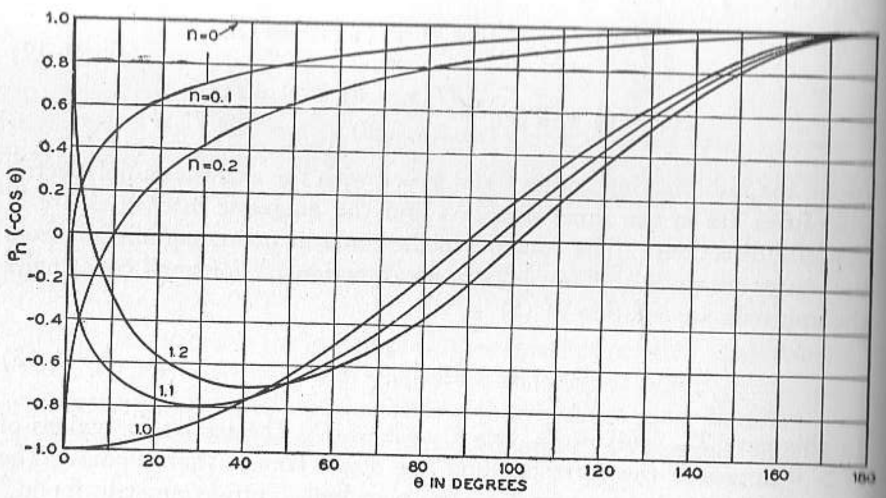


FIG. 11.7. Legendre functions of fractional order.

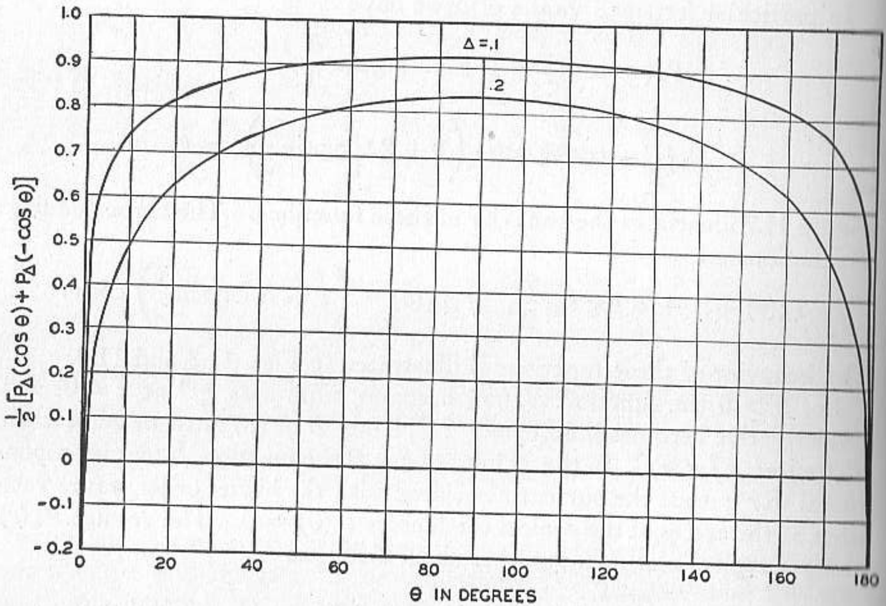


FIG. 11.8. Legendre functions of fractional order.

The total transverse voltage $V(r)$ and the total longitudinal current $I(r)$ in the upper cone may thus be written in the following form

$$V(r) = V_0(r), \quad I(r) = I_0(r) + \bar{I}(r), \quad \bar{I}(0) = 0, \tag{1-14}$$

where the principal voltage and current are

$$\begin{aligned} V_0(r) &= V_0(l) \cos \beta(l-r) + iKI_0(l) \sin \beta(l-r), \\ I_0(r) &= \frac{iV_0(l)}{K} \sin \beta(l-r) + I_0(l) \cos \beta(l-r), \end{aligned} \tag{1-15}$$

and l is the length of the cone. The total " complementary " current wave $\bar{I}(r)$ consists of an infinite number of current waves associated with higher order transmission modes

$$\bar{I}(r) = I_1(r) + I_3(r) + I_5(r) + \cdots. \tag{1-16}$$

As implied by this equation only the odd order waves, for which $T(\theta)$ is given by equation (11), appear when the cones are equal and when they are

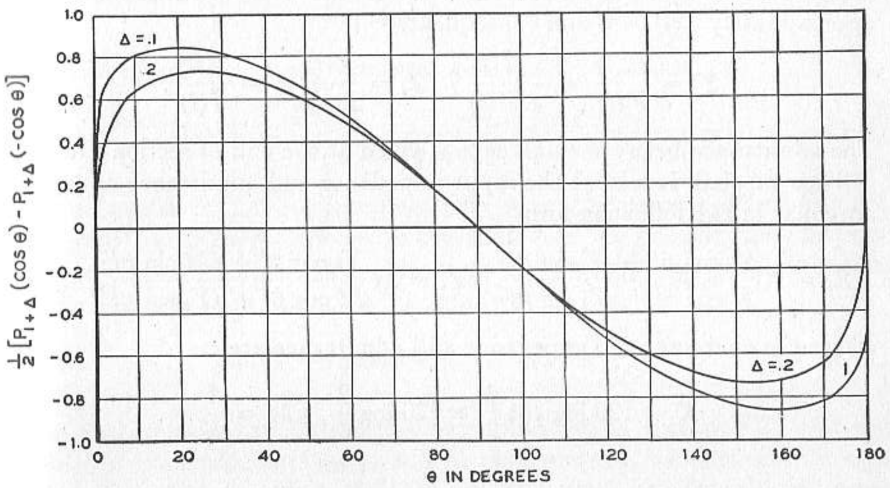


FIG. 11.9. Legendre functions of fractional order.

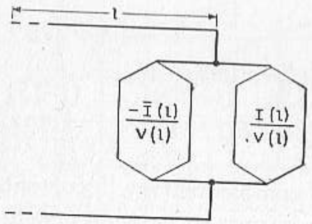
energized at the center. The currents associated with waves of even order flow in opposite directions and can only be produced by impressed voltages acting in opposite directions.

11.2. *General Considerations Concerning the Input Impedance and Admittance of a Conical Antenna*

The input impedance of a conical antenna is

$$Z_i = \frac{V(0)}{I(0)} = \frac{V_0(0)}{I_0(0)} = K\frac{V_0(l)\cos\beta l + iKI_0(l)\sin\beta l}{KI_0(l)\cos\beta l + iV_0(l)\sin\beta l}. \tag{2-1}$$

Thus in so far as the input impedance, the total voltage distribution, and the principal current distribution are concerned the effect of the complementary current waves is equivalent to a terminal impedance given by



$$Z_t = \frac{V_0(l)}{I_0(l)} = \frac{V(l)}{I(l) - \bar{I}(l)}. \qquad (2\text{–}2)$$

Taking the reciprocal, we have

$$Y_t = \frac{I(l)}{V(l)} - \frac{\bar{I}(l)}{V(l)}. \qquad (2\text{–}3)$$

FIG. 11.10. The terminal or radiation admittance of an antenna.

Thus the terminal admittance consists of two admittances in parallel and the conditions at $r = l$ may be represented diagrammatically as in Fig. 11.10. Since $I(l)$ is the current flowing into one spherical cap of the conical antenna and out of the other we may interpret $I(l)/V(l)$ as the admittance between the two caps. When the caps are small we have approximately $I(l) = 0$ and consequently

$$Y_t = \frac{I_0(l)}{V(l)} = -\frac{\bar{I}(l)}{V(l)}, \quad Z_t = \frac{V(l)}{I_0(l)} = -\frac{V(l)}{\bar{I}(l)}. \qquad (2\text{–}4)$$

The admittance between small caps is given at the end of section 11.6.

Thus we shall represent the input impedance and admittance of conical antennas in the following form

$$Z_i = K\frac{Z_t \cos \beta l + iK \sin \beta l}{K \cos \beta l + iZ_t \sin \beta l}, \quad Y_i = M\frac{Y_t \cos \beta l + iM \sin \beta l}{M \cos \beta l + iY_t \sin \beta l}, \qquad (2\text{–}5)$$

where the characteristic impedance and admittance are

$$K = 120 \log \cot \frac{\psi}{2} \simeq 120 \log \frac{2}{\psi}, \quad M = \frac{1}{K}.$$

11.3. *Current Distribution in the Antenna and the Terminal Impedance*

In section 1 we have seen that the radial electric intensity in the antenna region can be expressed as follows

$$2\pi i\omega\varepsilon r^2 E_r = \sum_n a_n \frac{\hat{J}_n(\beta r)}{\hat{J}_n(\beta l)} T_n(\theta), \qquad (3\text{–}1)$$

where $T_n(\theta)$ is defined by (1–6) and the summation is extended over the set (1–10). The complementary current in the antenna is then

$$\bar{I}(r) = -\sum_n \frac{a_n \hat{J}_n(\beta r)}{n(n+1)\hat{J}_n(\beta l)} \sin \psi \frac{d}{d\psi} T_n(\psi). \qquad (3\text{–}2)$$

When ψ is small, we have approximately

$$\frac{dT_n(\psi)}{d\psi} = \frac{\Delta}{\psi} = \frac{120}{K\psi}, \quad \bar{I}(r) = -\frac{120}{K}\sum_n \frac{a_n \hat{J}_n(\beta r)}{n(n+1)\hat{J}_n(\beta l)}. \qquad (3\text{–}3)$$

In free space the radial electric intensity may be written in the form

$$2\pi i\omega\varepsilon r^2 E_r = \sum_{k=1}^{\infty} b_k \frac{\hat{K}_k(i\beta r)}{\hat{K}_k(i\beta l)} P_k(\cos\theta). \qquad (3\text{–}4)$$

At $r = l$, E_r should be continuous, hence

$$\sum_n a_n T_n(\theta) = \sum_{k=1}^{\infty} b_k P_k(\cos\theta).$$

As K increases indefinitely and $\psi \to 0$, n approaches $2m+1$ and $T_n(\theta)$ approaches $P_{2m+1}(\cos\theta)$; thus the limiting value of a_n is

$$\lim a_n = \lim a_{2m+1+\Delta} = b_{2m+1} \text{ as } \Delta \to 0.$$

Hence for high impedance antennas we have approximately

$$\bar{I}(r) = -\frac{120}{K}\sum_{m=0}^{\infty} \frac{b_{2m+1}\hat{J}_{2m+1}(\beta r)}{2(m+1)(2m+1)\hat{J}_{2m+1}(\beta l)}. \qquad (3\text{–}5)$$

We have seen that for infinitely thin wires the current distribution is sinusoidal, with current nodes at the ends. In the present case the conclusion follows also from (5). Hence as $K \to \infty$, the current distribution in the antenna approaches the following value

$$I_0(r) = I_0 \sin \beta(l - r), \quad I_0 = \frac{iV_0(l)}{K}. \qquad (3\text{–}6)$$

The field of this distribution has been obtained in section 9.25 and the values of the coefficients b_{2m+1} will be determined if we expand $r^2 E_r$ as a series of zonal harmonics. For this purpo[illegible] ·h the following expression

$$C = 2\pi r \sin \theta H_\varphi = \tfrac{1}{2} i I_0 (e^{-i\beta r_1} + e^{-i\beta r_2} - 2e^{-i\beta r} \cos \beta l), \qquad (3\text{–}7)$$

obtained from (9.25–14). At great distances from the antenna this becomes

$$C = iI_0[\cos(\beta l \cos\theta) - \cos \beta l]e^{-i\beta r}.$$

Then from (10.10–6) we have

$$2\pi i\omega\varepsilon r^2 E_r = \frac{1}{\sin\theta}\frac{\partial C}{\partial\theta} = i\beta l I_0 \sin(\beta l \cos\theta) e^{-i\beta r}.$$

The expansion of sin ($\beta l \cos\theta$) in terms of zonal harmonics is known to be

$$\sin(\beta l \cos\theta) = \frac{1}{\beta l}\sum_{m=0}^{\infty}(-)^m(4m+3)\hat{J}_{2m+1}(\beta l)P_{2m+1}(\cos\theta).$$

Thus we have expressed the distant field in terms of zonal harmonics. On the other hand at great distances from the antenna equation (4) becomes

$$2\pi i\omega\epsilon r^2 E_r = \sum_{m=0}^{\infty}\frac{(-)^{m+1}b_{2m+1}P_{2m+1}(\cos\theta)e^{-i\beta r}}{\hat{J}_{2m+1}(\beta l) - i\hat{N}_{2m+1}(\beta l)}.$$

Comparing the alternative expressions for E_r we have

$$b_{2m+1} = -iI_0(4m+3)\hat{J}_{2m+1}(\beta l)[\hat{J}_{2m+1}(\beta l) - i\hat{N}_{2m+1}(\beta l)],$$

$$\bar{I}(r) = \frac{60iI_0}{K}\sum_{m=0}^{\infty}\frac{4m+3}{(m+1)(2m+1)}[\hat{J}_{2m+1}(\beta l) - i\hat{N}_{2m+1}(\beta l)]\hat{J}_{2m+1}(\beta r). \tag{3-8}$$

From (2–4), (6) and (8) we now obtain the terminal admittance

$$Y_t = \frac{Z_a(L)}{K^2}, \quad Z_a(L) = R_a(L) + iX_a(L),$$

$$R_a(L) = 60\sum_{m=0}^{\infty}\frac{4m+3}{(m+1)(2m+1)}\hat{J}^2_{2m+1}(L), \tag{3-9}$$

$$X_a(L) = -60\sum_{m=0}^{\infty}\frac{4m+3}{(m+1)(2m+1)}\hat{J}_{2m+1}(L)\hat{N}_{2m+1}(L),$$

where L is the " phase length " of each cone defined by $L = 2\pi l/\lambda$.

Thus for cones of small angles the terminal impedance and its inverse are

$$Z_t = \frac{K^2}{R_a(L) + iX_a(L)}, \quad \frac{K^2}{Z_t} = R_a(L) + iX_a(L). \tag{3-10}$$

In computing the input impedance of the antenna we may replace the terminal impedance Z_t by its inverse inserted in series with the transmission line representing the antenna, one quarter wavelength from its end.

11.4. *Calculation of the Inverse of the Terminal Impedance*

The series (3–9) for the inverse terminal reactance converges slowly and is not practical for numerical calculations unless L is small. A simple expression for $Z_a(L)$ can be obtained as follows. The input impedance is

$$Z_i = K\frac{Z_a \sin L - iK\cos L}{K\sin L - iZ_a\cos L}. \tag{4-1}$$

As K increases indefinitely we have

$$Z_i \to \frac{Z_a}{\sin^2 L} - iK\cot L. \tag{4-2}$$

Since the input current approaches the value $I_0 \sin\beta l$, the power input becomes

$$\Psi = \tfrac{1}{2}Z_i I_0^2 \sin^2 L = \tfrac{1}{2}[Z_a - \tfrac{1}{2}iK\sin 2L]I_0^2. \tag{4-3}$$

On the other hand the power flow for a sinusoidal distribution may be calculated by the method described in sections 9.26 and 9.27. The integral of the product of the current and the tangential electric intensity should be calculated on the surface of an infinitely thin cone. In the limit the real part of this integral is independent of the shape of the longitudinal cross-section of the antenna and $R_a = R$ as given by equation (9.27–2); but the reactive part is still a function of the shape of the longitudinal cross-section and must be calculated for the cone. In the limit the tangential component of the electric intensity is the sum of E_z and $E_\rho \sin\psi$, where E_z and E_ρ are given by (9.25–14).

In this way the following expressions are obtained

$$R_a(L) = 60(C + \log 2L - \text{Ci}\,2L) + 30(C + \log L - 2\,\text{Ci}\,2L + \text{Ci}\,4L)\cos 2L + 30(\text{Si}\,4L - 2\,\text{Si}\,2L)\sin 2L, \tag{4-4}$$

$$X_a(L) = 60\,\text{Si}\,2L + 30(\text{Ci}\,4L - \log L - C)\sin 2L - 30\,\text{Si}\,4L\cos 2L.$$

The inverse of the terminal impedance is shown in Fig. 11.11.

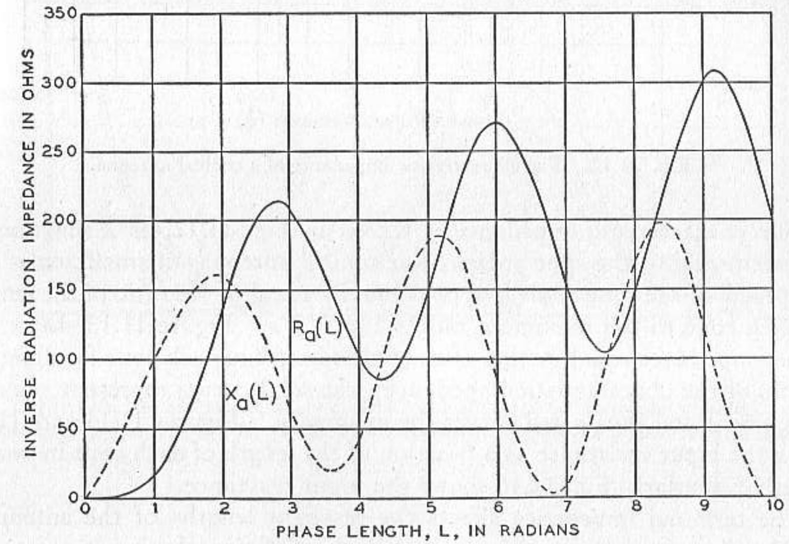


FIG. 11.11. Curves for the resistive and reactive components of the inverse terminal or radiation impedance of an antenna.

11.5. *The Input Impedance and Admittance of a Conical Antenna*

Separating the real and imaginary parts of Z_i and its reciprocal, we have

$$Z_i = \frac{R_a - i\left[\frac{1}{2}K \sin 2L + X_a \cos 2L - \frac{R_a^2 + X_a^2}{2K} \sin 2L\right]}{\sin^2 L + \frac{X_a}{K} \sin 2L + \frac{R_a^2 + X_a^2}{K^2} \cos^2 L}, \tag{5-1}$$

$$Y_i = \frac{R_a + i\left[\frac{1}{2}K \sin 2L + X_a \cos 2L - \frac{R_a^2 + X_a^2}{2K} \sin 2L\right]}{K^2\left[\cos^2 L - \frac{X_a}{K} \sin 2L + \frac{R_a^2 + X_a^2}{K^2} \sin^2 L\right]}. \tag{5-2}$$

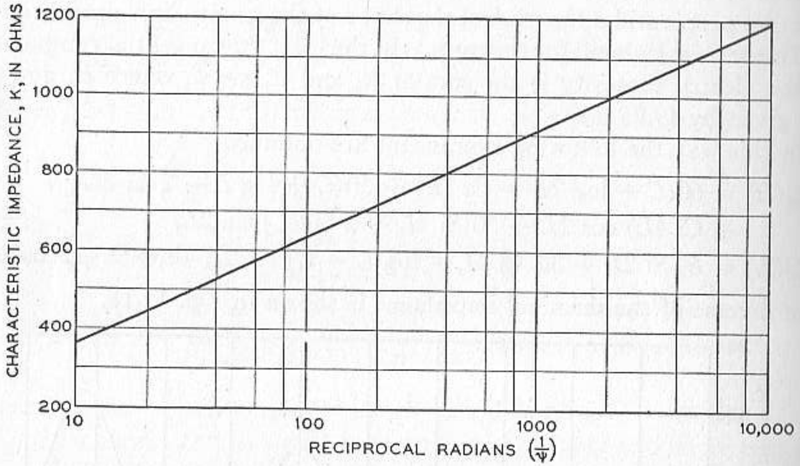


FIG. 11.12. The characteristic impedance of a conical antenna.

The characteristic impedance is shown in Fig. 11.12, as a function of the reciprocal of the cone angle. For conical antennas of small angles the reciprocal of the cone angle is approximately equal to the ratio of the length of each cone to the maximum radius $1/\psi = l/a$. Figure 11.13 shows the input impedance as a function of the phase length of each cone for different values of the characteristic impedance; the solid curves represent the real component and the dotted curves the imaginary. Figures 11.14 and 11.15 show the input resistance as a function of the length of each cone in wavelengths; similarly Fig. 11.16 shows the input reactance.

The terminal impedance affects the resonant lengths of the antennas. As the characteristic impedance approaches infinity the input reactance

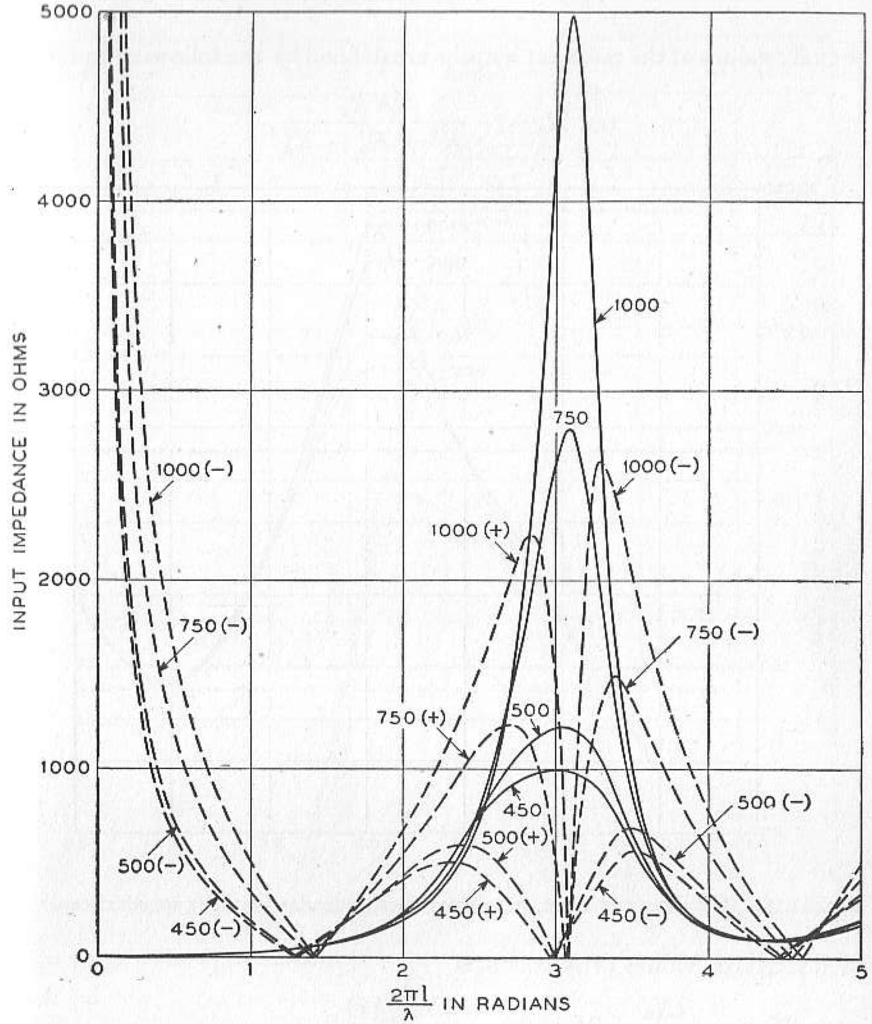


FIG. 11.13. The input impedance of hollow conical antennas (without spherical caps) as a function of $2\pi l/\lambda$ and K. Solid curves represent the real component and the dotted curves the imaginary.

vanishes for values of L approaching those given by

$$\sin 2L = 0, \quad 2L = k\pi, \quad 2l = \frac{k\lambda}{2}, \quad k = 1,2,3,\cdots;$$

the exact values of the resonant lengths are defined by the following equation

$$\tan 2L = -\frac{2KX_a}{K^2 - R_a^2 - X_a^2},$$

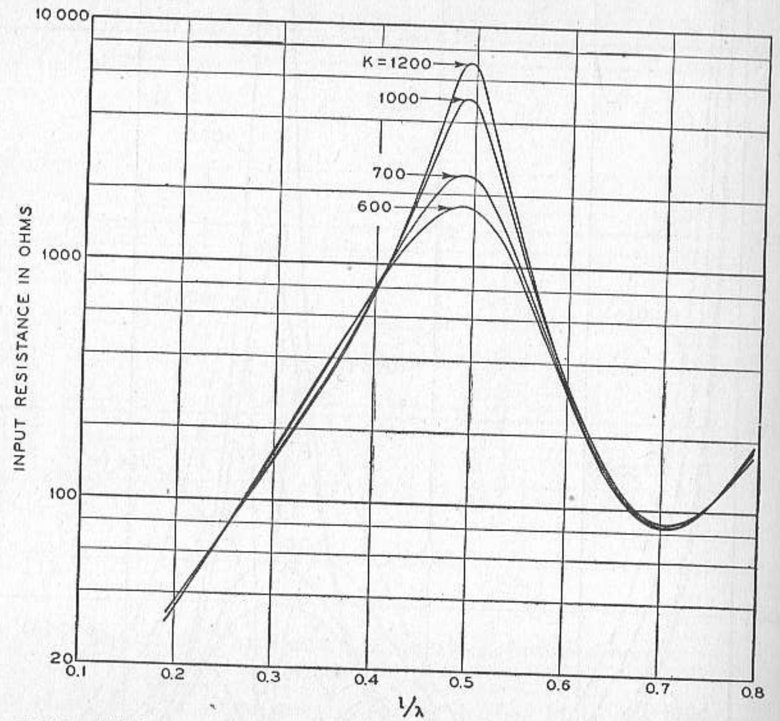


FIG. 11.14. The input resistance of hollow conical antennas (without spherical caps).

which for large values of K becomes

$$\tan 2L = -\frac{2X_a}{K}, \quad 2L = k\pi - \frac{2X_a(\frac{1}{2}k\pi)}{K},$$

$$\frac{4l}{k\lambda} = 1 - \frac{2X_a(\frac{1}{2}k\pi)}{\pi kK} = 1 - \frac{120 \text{ Si } k\pi + 60(-)^{k+1} \text{ Si } 2k\pi}{\pi kK}. \quad (5\text{–}3)$$

Figure 11.17 shows the deviation of the resonant length of the antenna from $k\lambda/2$.

The input impedance of the antenna when the length of the cone is equal

to an odd multiple of $\lambda/4$ is of course the inverse terminal impedance $Z_i = Z_a$. When $l = \lambda/4$, this impedance is

$$Z_i = 73.129 + i153.66.$$

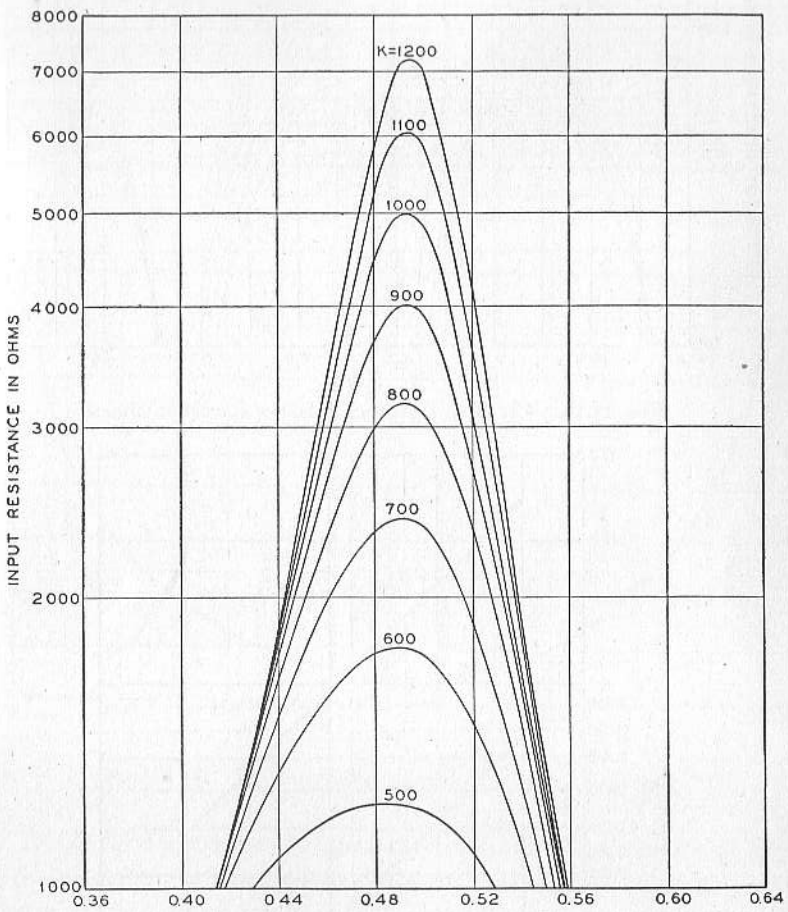


FIG. 11.15. The input resistance of hollow conical antennas in the neighborhood of the second resonance.

When l is a multiple of $\lambda/2$, then the input impedance is equal to the terminal impedance. If Z_a were independent of l, the above two impedances would be represented by inverse points in the impedance plane. In fact the whole impedance diagram would be a circle inverse to itself with respect to a circle of radius K. As it is the impedance diagram is not quite a circle but surprisingly close to it considering the variation in Z_a. Figure 11.18 shows such a diagram for $K = 750$.

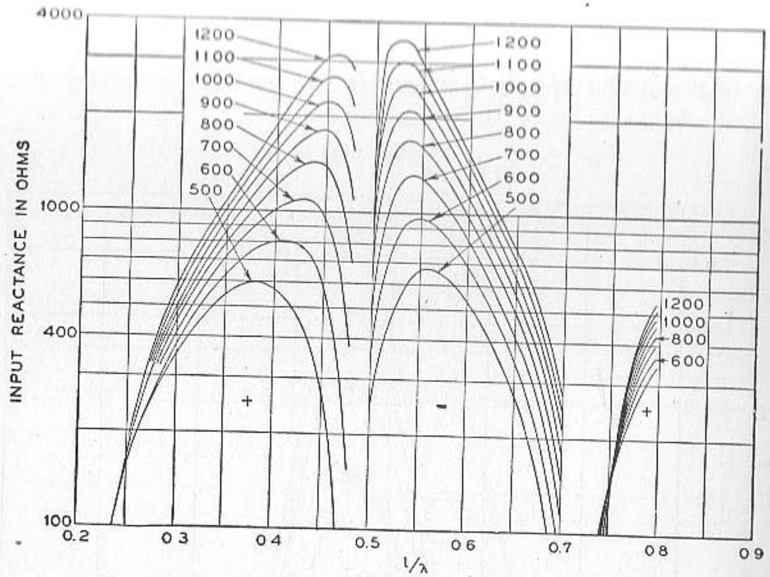


Fig. 11.16. The input reactance of hollow conical antennas.

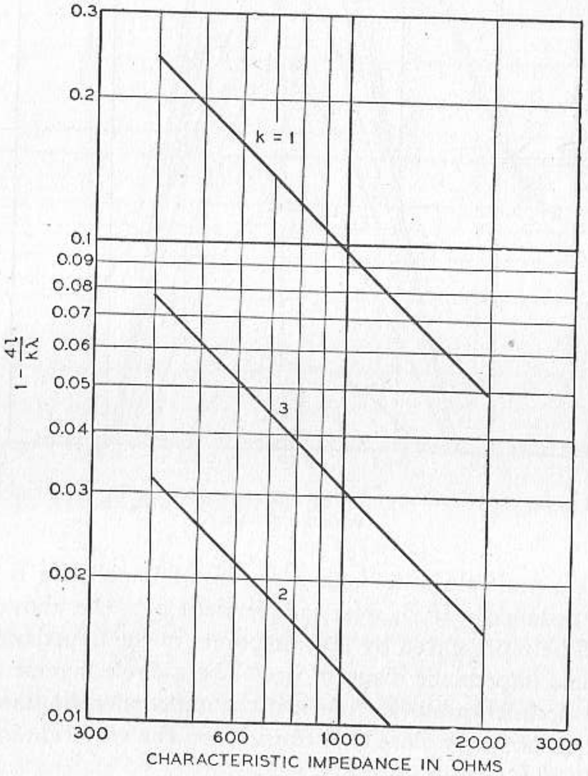


Fig. 11.17. Deviation of the resonant length of conical antennas (without spherical caps) from $2l = k\lambda/2$.

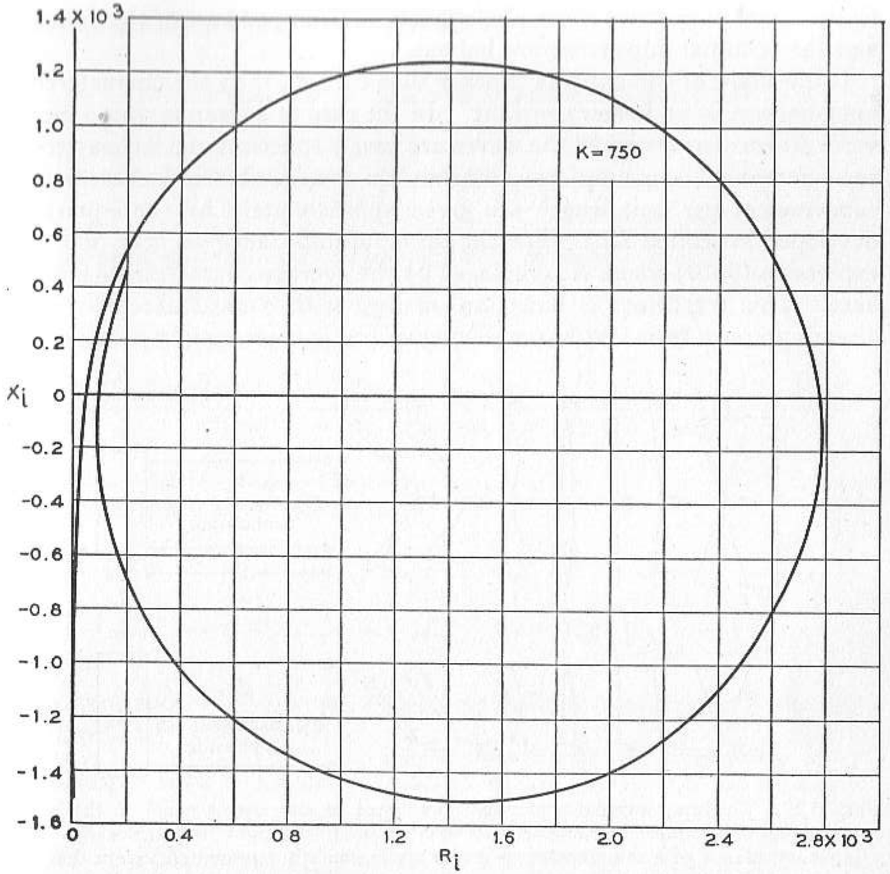


Fig. 11.18. The impedance diagram for a conical antenna.

11.6. *The Input Impedance of Antennas of Arbitrary Shape and End Effects*

In the preceding sections we have shown that a biconical antenna may be regarded as a wave guide with infinitely many transmission modes. Only the principal mode is generated if two infinitely long cones are energized at their common apex; but if the cones are of finite length there is a discontinuity at the boundary sphere which separates the antenna region from the surrounding space and reflection takes place. Then higher transmission modes are generated, and the effect of these modes on the amplitudes of the total voltage wave, the principal current wave, and the input impedance can be represented *exactly* by an appropriate terminal impedance. This property we represent diagrammatically as in Fig. 11.19. We have also determined an approximate value of the terminal impedance for cones of small angles. When a single cone is placed normally to a perfectly conducting plane sheet, the characteristic impedance of the antenna and the terminal impedance are halved.

If the shape of the antenna is other than conical, then the characteristic impedance K is no longer constant. In the case of antennas whose transverse dimensions are small the waves are nearly spherical and we may treat such antennas as nonuniform transmission lines whose inductances and capacitances per unit length are given approximately by the equations developed in section 8.13. For the terminal impedance we may take the expression (3–10) where K is replaced by the average characteristic impedance. This treatment is based on analogy with conical antennas; the direct approach based on Maxwell's equations is theoretically possible but

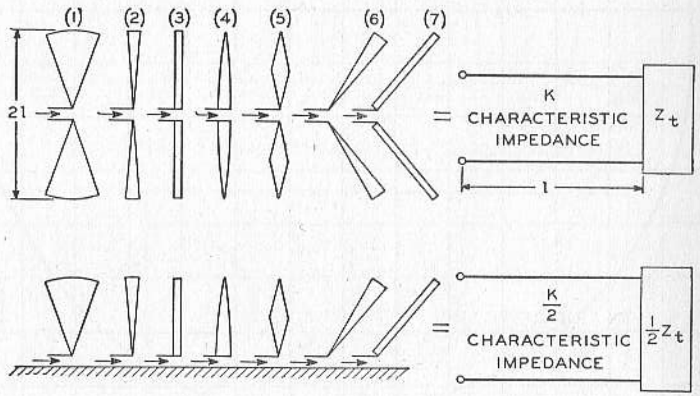


FIG. 11.19. The input impedance of a conical antenna of any size is equal to the input impedance of a uniform transmission line with a certain "output" impedance Z_t. The input impedance of a thin antenna of any shape is similarly represented, except that the characteristic impedance is variable.

at present is not practicable because no convenient method is available for computing the appropriate wave functions. In order to obtain such functions exactly it is necessary to use coordinate systems appropriate to the boundaries *and* to the physical phenomena. Spheroidal coordinates, for instance, fit the boundary of a spheroidal antenna but not the problem of the transmission of waves on it; and spheroidal wave functions are much more suitable for defining waves traveling away from the spheroid than waves traveling along it.

From the theory of nonuniform transmission lines we obtain the following approximate expression for the input impedance

$$Z_i = K_a \frac{R_a \sin L + i[(X_a - N)\sin L - (K_a - M)\cos L]}{[(K_a + M)\sin L + (X_a + N)\cos L] - iR_a \cos L}; \qquad (6\text{–}1)$$

where M and N are the functions defined by

$$M(L) = \beta \int_0^l [K_a - K(r,\rho)] \sin 2\beta r \, dr,$$
$$N(L) = \beta \int_0^l [K_a - K(r,\rho)] \cos 2\beta r \, dr. \qquad (6\text{–}2)$$

Separating the real and imaginary parts, we have

$$R_i = \frac{K_a R_a (K_a + N \sin 2L - M \cos 2L)}{R_a^2 \cos^2 L + [(K_a + M)\sin L + (X_a + N)\cos L]^2},$$

$$X_i = \frac{K_a[\frac{1}{2}(R_a^2 + X_a^2 + M^2 - N^2 - K_a^2)\sin 2L + (MN - K_a X_a)\cos 2L + (MX_a - K_a N)]}{R_a^2 \cos^2 L + [(K_a + M)\sin L + (X_a + N)\cos L]^2},$$

where K_a is the average impedance.

For cylindrical antennas we find

$$M(L) = 60(\log 2L - \text{Ci}\, 2L + C - 1 + \cos 2L),$$
$$N(L) = 60(\text{Si}\, 2L - \sin 2L), \quad K_a = 120\left(\log \frac{2l}{a} - 1\right).$$

In free space, for antennas whose longitudinal cross-section is rhombic, we have

$$M(L) = 60(C + \log 2L - \text{Ci}\, 2L)(1 + \cos 2L) - 60\, \text{Si}\, 2L \sin 2L,$$
$$N(L) = 60\, \text{Si}\, 2L(1 - \cos 2L) - 60(C + \log 2L - \text{Ci}\, 2L)\sin 2L,$$
$$K_a = 120 \log \frac{2l}{a},$$

where a is the maximum radius. For vertical antennas of triangular shape with base of radius a, above a perfectly conducting ground, the above values are halved.

For spheroidal antennas the corresponding formulas are

$$M(L) = R_a(L) - 60(1 - \cos 2L) \log 2,$$
$$N(L) = X_a(L) - 60 \log 2 \sin 2L, \quad K_a = 120 \log \frac{l}{a},$$

where a is the maximum radius.

In the case of antennas of rhombic cross-section above a perfectly conducting ground (cross-section (5) in Fig. 11.19), the first half of the antenna is uniform and the second half nonuniform. The input impedance of the second half is

$$Z_i = K_a \frac{R_a \sin \frac{L}{2} + i\left[(X_a - N)\sin\frac{L}{2} - (K_a - M)\cos\frac{L}{2}\right]}{\left[(K_a + M)\sin\frac{L}{2} + (X_a + N)\cos\frac{L}{2}\right] - iR_a \cos\frac{L}{2}},$$

$$M(L) = 60 \log 4 + 60(C + \log \tfrac{1}{2}L - \text{Ci } 2L) \cos L - 60 \text{ Si } 2L \sin L,$$

$$N(L) = 60(\text{Si } 2L - 2 \text{ Si } L) \cos L - 60(C + \log \tfrac{1}{2}L + \text{Ci } 2L - 2 \text{ Ci } L) \sin L,$$

$$K_a = 120 \log \frac{4l}{a}.$$

It should be noted that in the above equation for Z_i the quantities R_a, X_a, M, and N are functions of L not of $L/2$. Using this impedance as the terminal impedance of a uniform transmission line of phase length $L/2$ and characteristic impedance equal to that of the first half of the antenna, we obtain the input impedance of the entire antenna.

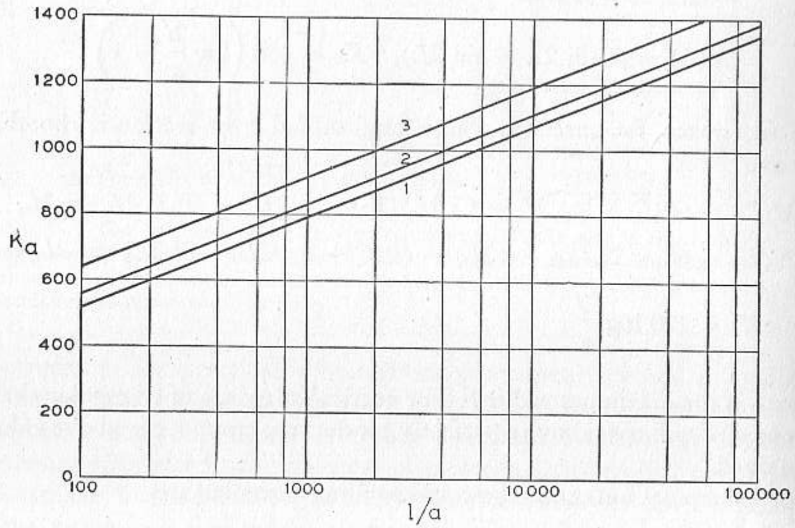


FIG. 11.20. The average characteristic impedance: (1) cylindrical antenna, (2) spheroidal antenna, (3) antenna of rhombic cross-section.

In Fig. 11.20 the average characteristic impedances are shown for several antenna shapes: curve (1) is for a cylindrical antenna, curve (2) for a spheroidal antenna, and curve (3) for an antenna of rhombic cross-section. In Figs. 11.21 and 11.22, the input resistance and reactance of cylindrical antennas are shown as functions of l/λ for different values of K_a. For

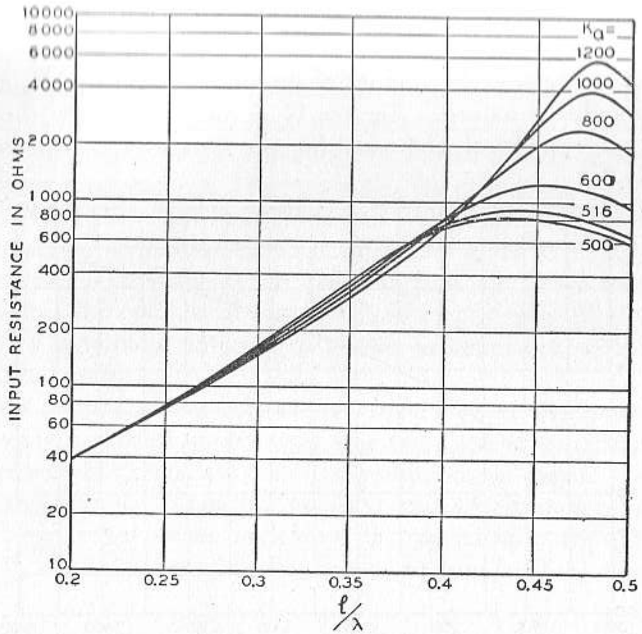


FIG. 11.21. The input resistance of hollow cylindrical antennas in free space. For vertical antennas over a perfectly conducting ground divide the ordinates and K_a by 2.

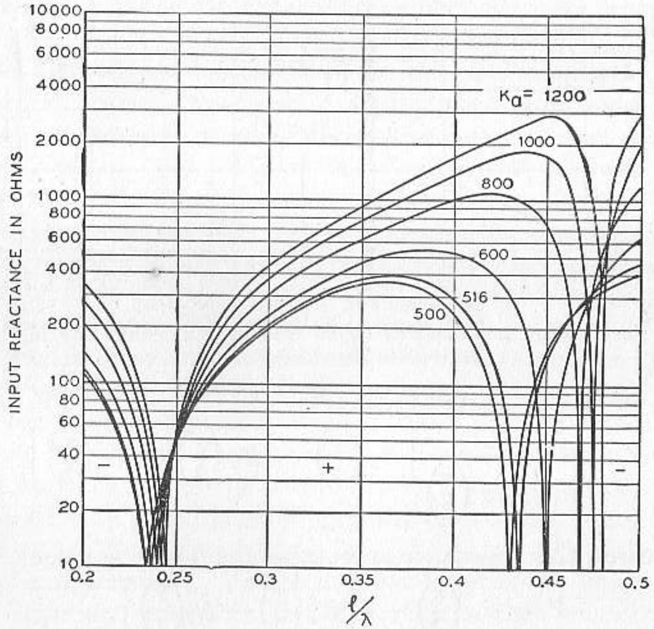


FIG. 11.22. The input reactance of hollow cylindrical antennas in free space. For vertical antennas over a perfectly conducting ground divide the ordinates and K_a by 2.

antennas above a perfectly conducting ground the values of K_a and of the ordinates should be halved. Figures 11.23 and 11.24 show the resonant impedance of cylindrical antennas for the first and second resonances respectively.*

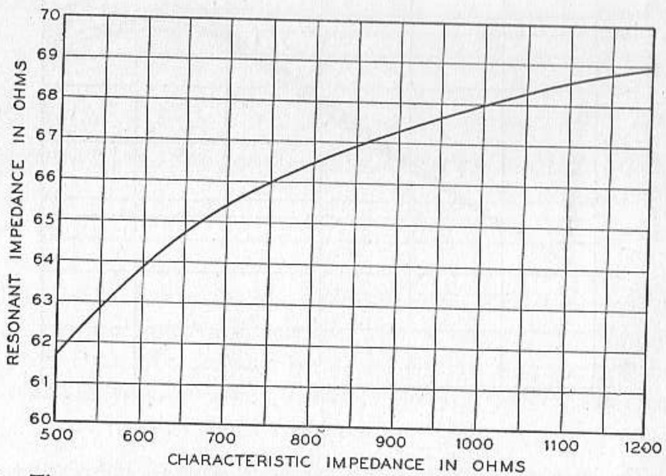


FIG. 11.23. The resonant impedance of hollow cylindrical antennas as a function of K when l is in the vicinity of $\lambda/4$.

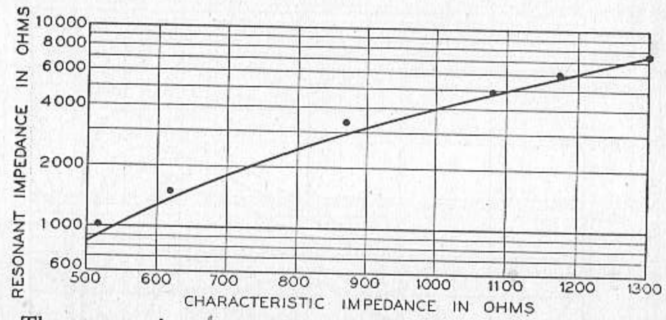


FIG. 11.24. The resonant impedance of hollow cylindrical antennas as a function of K when l is in the vicinity of $\lambda/2$.

When $l = \lambda/4$, the input impedance is

$$Z_i = \frac{K_a}{K_a + M\left(\frac{\pi}{2}\right)}\left[R_a\left(\frac{\pi}{2}\right) + iX_a\left(\frac{\pi}{2}\right) - iN\left(\frac{\pi}{2}\right)\right].$$

As K_a increases, this impedance approaches the following value

$$Z_i = R_a\left(\frac{\pi}{2}\right) + i\left[X_a\left(\frac{\pi}{2}\right) - N\left(\frac{\pi}{2}\right)\right].$$

* The points in Fig. 11.24 are experimental.

The limiting value of the input resistance is nearly 73.13 ohms. For any finite value of K_a, the input resistance depends also on $M(\pi/2)$, that is, on the shape of the longitudinal cross-section as well as on the mean size of the transverse cross-section. For example, for cylindrical and spheroidal antennas $M(\pi/2)$ is equal respectively to -21 and -10; hence the input resistances are somewhat higher than 73.13. Even when K_a is infinite, the input reactance depends on the shape of the longitudinal cross-section. Thus for cylindrical antennas the reactance is 30 Si $2\pi = 42.5$ ohms and for spheroidal antennas it is zero as compared with 154 ohms for conical antennas.

The effect of capacitance between spherical caps at the ends of the antennas may be included as follows. For two caps of small radius a the admittance between them is substantially equal to $i\omega C$ where C is the electrostatic capacitance. Thus the value of this admittance is $ia/30\lambda$ and it should be included, when necessary, in parallel with the terminal admittance Y_t of (2–4). Introducing this correction in (3–9) we find

$$Y_t = \frac{R_a(L)}{K^2} + i\left[\frac{X_a(L)}{K^2} + \omega C\right],$$

so that the effect of the cap capacitance is obtained if we replace X_a by $X_a + \omega CK^2$.

Consider for instance formula (5–3) for the resonant lengths of antennas. The cap capacitance causes a shortening of the resonant length which may be expressed as $\omega CK/L$. For conical antennas this may be written $K/30\pi \exp(-K/120)$, while for cylindrical antennas its value is $K/30\pi \exp(-K/120 - 1)$. This effect will be negligible when K is sufficiently large, but for $K = 600$ we find that the cap capacitance decreases the resonant length by 4.3 per cent for conical antennas and by 1.6 per cent for cylindrical antennas. If K is increased to 1200 these percentages are reduced to 0.058 and 0.021 respectively.

The explicit expressions for the M and N functions given in this section have been calculated from (6–2) on the assumption that the distance between terminals is negligibly small. Under practical conditions one may assume that M and N are independent of this distance so long as l is measured from the end of the antenna to its nearest input terminal. In more accurate work the effect of finite separation can be obtained by recomputing M and N from equations (6–2). The distance r should be taken from the input terminals. As the distance between the input terminals becomes larger and larger the field distortion in their vicinity becomes substantial and will effectively introduce an impedance across these terminals in parallel with that of the rest of the antenna.

11.7. *Current Distribution in Antennas*

The current distribution in a conical antenna is given by equations (1–14), (1–15), (1–16) and (3–8). In an antenna with nonuniform characteristic impedance equations (1–15) should be modified in accordance with the theory of nonuniform transmission lines. Thus the current distribution depends on the distribution of inductance and capacitance in the antenna as well as on the complementary current waves due to reflection at the boundary sphere.

A typical complementary current wave is given by

$$I_{2m+1}(r) = \frac{60(4m+3)I_0}{(m+1)(2m+1)}[\hat{N}_{2m+1}(\beta l) + i\hat{J}_{2m+1}(\beta l)]\hat{J}_{2m+1}(\beta r).$$

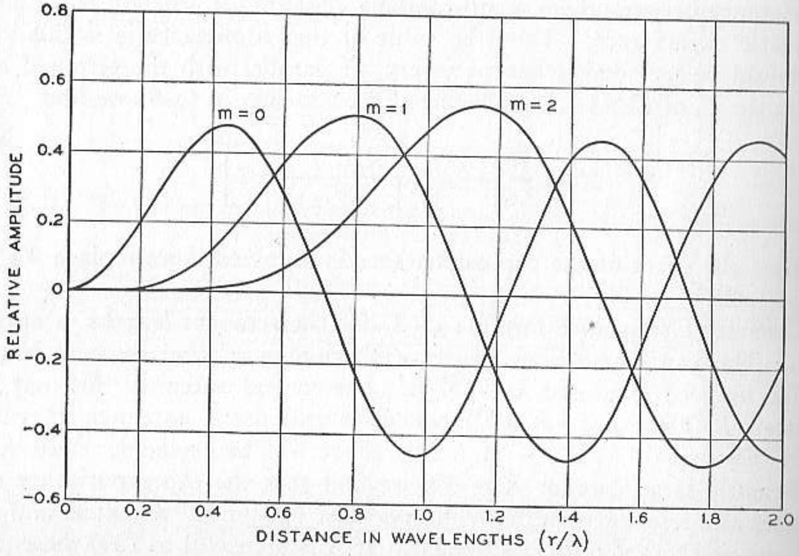


FIG. 11.25. The amplitudes of the first, third, and fifth ($m = 0,1,2$) secondary waves as functions of the distance from the center of the antenna.

Figure 11.25 shows the variation of the amplitudes of the first, third and fifth ($m = 0,1,2$) secondary waves as functions of the distance from the center of the antenna. In the vicinity of $r = 0$ the amplitude of the $(2m+1)$th current wave varies nearly as r^{2m+1}; thus the secondary current waves affect the current distribution mainly near the ends of the antenna. The maximum amplitudes of the secondary current waves depend on the characteristic impedance and on the length of the antenna. Figure 11.26 shows the secondary current waves when $K = 1000$ and $l = \lambda/2$. The solid curves show the components in phase with the principal current

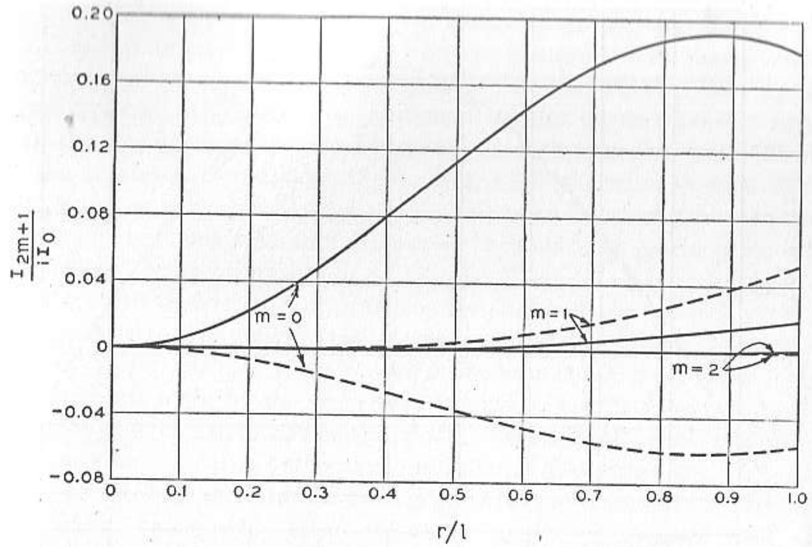


FIG. 11.26. Secondary current waves for $K = 1000$ and $l = \lambda/2$. The solid curves show the components in phase with the dominant current $I_0 \sin \beta(l - r)$, and the dotted curves show the quadrature components.

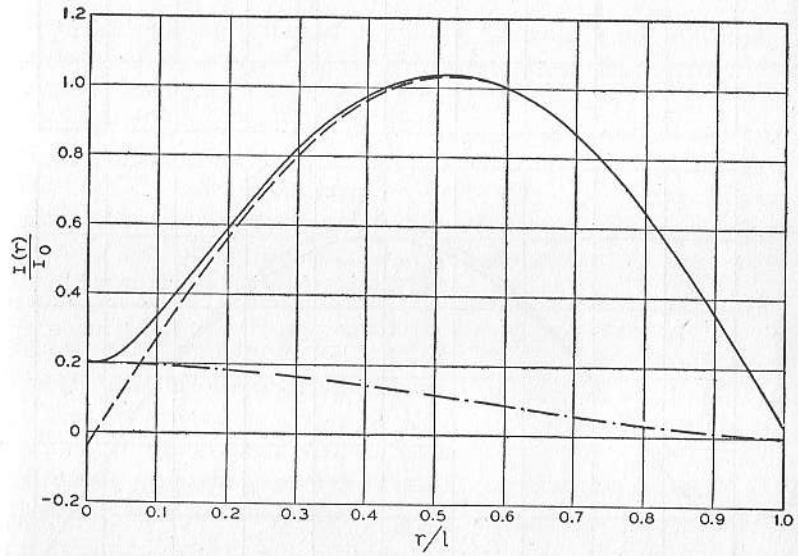


FIG. 11.27. The total current in the antenna of length $2l = \lambda$; $K = 1000$. The solid curve represents the amplitude of the total current; the dash curve represents the amplitude of the component in phase with I_0 and the dash-dot curve is the amplitude of the quadrature component.

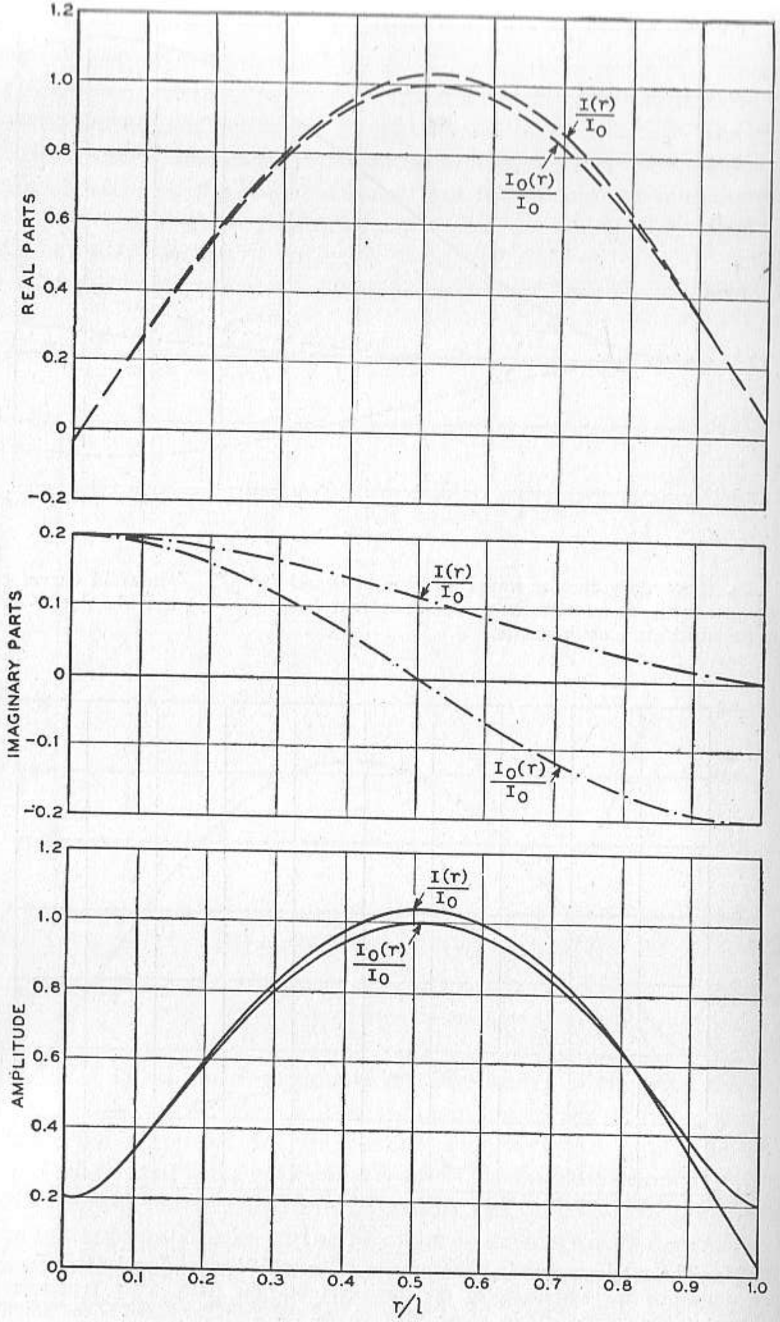


FIG. 11.28. Curves for the total current and the principal current.

component $I_0 \sin \beta(l - r)$; only the first of these is important, except in the immediate vicinity of $r = l$. The dotted curves show the quadrature components of which only the first two need be considered. In Fig. 11.27 the solid curve, the dash curve, and the dash-dot curve represent respectively the amplitude of the total current, the component in phase with I_0, and the quadrature component.* In Fig. 11.28 the total current is compared with the principal current. The difference between the real parts is quite small but the difference between the imaginary parts is relatively large except at the center.

The current distribution determines the shape of the radiation pattern. The quadrature components radiate independently. Since the radiation intensity is proportional to the square of the moment, the radiation pattern is affected but little by the current in quadrature with I_0 except in those directions in which the radiation is small, where a small absolute difference may contribute a large percentage deviation. Rewriting the expressions (1–15) for the voltage and the principal current in terms of the amplitude I_0 as defined by equation (3–6), we have

$$V(r) = -iKI_0 \cos \beta(l - r) + (R_a + iX_a)I_0 \sin \beta(l - r),$$

$$I_0(r) = I_0 \sin \beta(l - r) + \frac{X_a - iR_a}{K} I_0 \cos \beta(l - r).$$

The real part of the principal current is practically unaltered except in the vicinity of the current nodes. These equations are in agreement with the general theory concerning the current distribution on thin wires.

The imperfect conductivity of the ground has a much more marked effect on the radiation pattern than the deviations in the current distribution from the limiting distribution $I_0 \sin \beta(l - r)$. Thus the radiation intensity of a quarter-wave antenna in the ground plane is zero for an imperfectly conducting ground; an effect in comparison with which the effect of finite K is altogether negligible.

11.8. *Inclined Wires and Wires Energized Unsymmetrically*

The principal waves on inclined wires [see Figs. 11.19(6) and 11.19(7)] have been considered in section 8.14. The functions R_a and X_a may be calculated by the method outlined in section 4; the M and N functions are obtained from the theory of nonuniform transmission lines; in terms of these functions the impedance is given by equation (6–1).

* In this figure the current does not quite vanish at the end of the antenna because only the first two complementary waves were included; the higher order waves reduce the end current to zero but do not affect the current appreciably at any distance from the end.

The calculation of the complementary current waves is more complicated. If the angles ψ_1 and ψ_2 of the conical wires are small compared with the angle ϑ between their axes, the proximity effect is small and the current distribution in each wire is substantially uniform round its axis. Then the radial electric intensity is proportional to the following function

$$T(\theta_1,\theta_2) = AP_n(-\cos\theta_1) + BP_n(-\cos\theta_2),$$

where θ_1 and θ_2 are the angles made by a typical radius with the axes of the wires. This function must vanish on the surfaces of the wires and we have approximately

$$\begin{aligned} AP_n(-\cos\psi_1) + BP_n(-\cos\vartheta) &= 0, \\ AP_n(-\cos\vartheta) + BP_n(-\cos\psi_2) &= 0. \end{aligned} \tag{8–1}$$

Hence n must be a root of

$$P_n(-\cos\psi_1)P_n(-\cos\psi_2) = [P_n(-\cos\vartheta)]^2.$$

For equal cones this equation becomes $P_n(-\cos\psi) = \pm P_n(-\cos\vartheta)$. Once n is found, the ratio A/B is obtained from (1) and the T-function is determined except for a constant factor. The field is then computed from the T-function.

Other antenna problems arise if the generator is not at the center. For example, in the case of a biconical antenna the generator may be at some point A (Fig. 11.29). In general the current distribution in the antenna is not symmetric about the center and waves corresponding to all values of n given by (1–10) and (1–13) are generated. There are two principal waves, the symmetric corresponding to $n = 0$ and the antisymmetric corresponding to $n = 120/K$. The symmetric transmission modes correspond to the case in which one-half of the total electromotive force V is applied at A and the other half at A' (Fig. 11.29); the antisymmetric modes correspond to the case in which $\frac{1}{2}V$ is applied at A and $-\frac{1}{2}V$ at A'. In region (1) only $\hat{J}_n$-functions are permissible (when $n > 0$) and in region (2) $\hat{N}_n$-functions should be included.

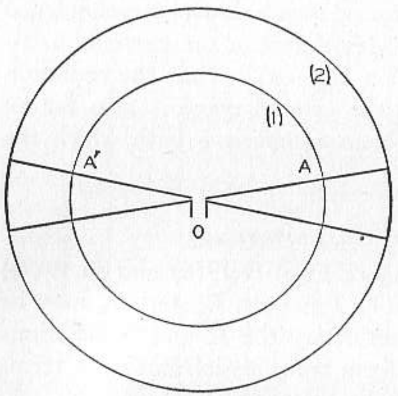


FIG. 11.29. The cross-section of a conical antenna and boundary spheres.

Another method of approach is based on considering two cones of unequal length with a generator at the common apex O (Fig. 11.30). The field in region (1) is similar to that in region (1) of the preceding problem. The field in region (2) consists of waves corresponding to a different set of values for n since in this region the field must be finite for $\theta = \pi$ and consequently $T(\theta) = AP_n(-\cos\theta)$. The values of n are then obtained from the equation $P_n(-\cos\psi) = 0$. If ψ is small, the roots are approximately $n = m + 60/K, m = 0,1,2,\cdots$.

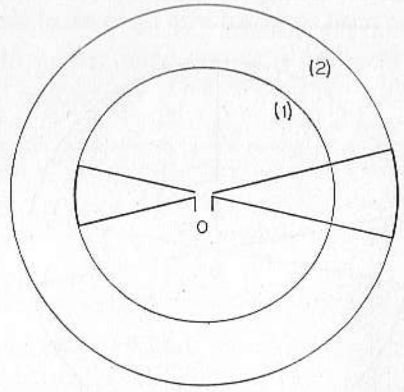


FIG. 11.30. The cross-section of a conical antenna and boundary spheres.

Transmission of waves on a wire energized near one of its ends may also be studied by considering waves on a cone surmounted by a small sphere (Fig. 11.31). The electromotive force may be applied either between the cone and the spherical surface round the circumference AB or between the apices at O, one apex belonging to the original cone and the other to a small cone leading to the surface of the sphere. Approximate solutions of these problems can be found without much difficulty and extended to wires of other shapes. These solutions will supplement the conclusions drawn from the solution of the principal problem of two equal cones of equal length.

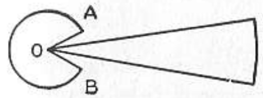


FIG. 11.31. The cross-section of a conical antenna and a small sphere at its apex.

11.9. *Spherical Antennas*

Equation (2–3) is the general expression for the terminal admittance Y_t of a biconical antenna consisting of two equal and oppositely directed cones. When the cone angle ψ and the length l are sufficiently small, the total current $I(l)$ is so nearly equal to zero that the terminal admittance is determined substantially by the complementary current. But as the cone angle increases the total current makes an increasingly important contribution to the admittance. When the cone angle ψ is nearly 90 degrees, the biconical antenna becomes a pair of nearly hemispherical conductors fed by a cone transmission line from the center (Fig. 11.32). In this case the complementary current $\bar{I}(l)$ becomes small and the terminal admittance is determined largely by the total current $I(l)$. That the complementary current becomes relatively small can be seen if we observe that in the present case the cone line is approximately a disc line with variable separation between the "parallel planes." When the separation is small compared with the wavelength, all transmission modes except the principal are attenuated and the energy associated with these modes is concentrated in the region where the conical feeders join the sphere. The depth of this region (in the direction toward the center) is comparable to the separation between the edges of the two hemispheres. The capacitance representing this local storage of energy will be small compared with the external capacitance between the hemispheres.

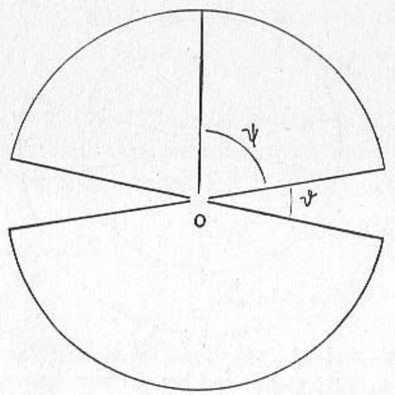


FIG. 11.32. The cross-section of a spherical antenna or a double cone of large angle ψ surmounted by large spherical caps.

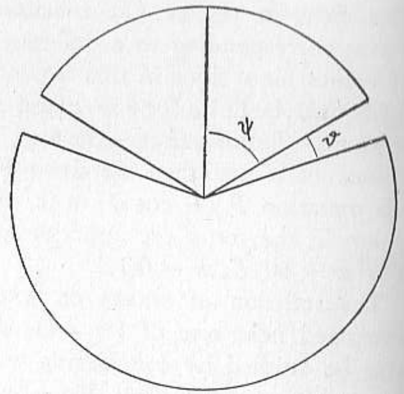


FIG. 11.33. A spherical antenna fed along an arbitrary circle of parallel.

Thus we may assume that the approximate voltage distribution over the aperture of the cone transmission line in Fig. 11.32 or in the more general case shown in Fig. 11.33 is governed by the principal wave. Let V be the total transverse voltage at $r = l = a$, where a is the radius of the sphere; then $E_\theta(a,\theta) = 0$ except in the interval $\psi < \theta < \psi + \vartheta$, where

$$E_\theta(a,\theta) = \frac{\eta V}{2\pi K a \sin\theta} = \frac{60V}{Ka\sin\theta}. \tag{9–1}$$

In accordance with equations (1–1) the meridian electric intensity for $r > a$ may be represented as follows

$$rE_\theta(r,\theta) = V\sum_{n=1}^{\infty} \frac{A_n \hat{K}_n'(i\beta r)}{\hat{K}_n'(i\beta a)} P_n^1(\cos\theta). \tag{9–2}$$

In the theory of spherical harmonics it has been established that an arbitrary function $f(\theta)$ (subject to certain limitations) can be expanded in the form

$$f(\theta) = \sum_{n=1}^{\infty} a_n P_n^1(\cos\theta), \quad a_n = \frac{2n+1}{2n(n+1)}\int_0^{\pi} f(\theta)P_n^1(\cos\theta)\sin\theta\, d\theta.$$

Hence the coefficient A_n in the expansion for $aE_\theta(a,\theta)$ is

$$A_n = \frac{60(2n+1)}{2n(n+1)K}\int_\psi^{\psi+\vartheta} \frac{d}{d\theta} P_n(\cos\theta)\, d\theta$$

$$= \frac{30(2n+1)[P_n(\cos\overline{\psi+\vartheta}) - P_n(\cos\psi)]}{n(n+1)K}. \tag{9–3}$$

From (2) and (1–1) we obtain the remaining intensities

$$rH_\varphi(r,\theta) = -V\sum_{n=1}^{\infty} A_n \frac{\hat{K}_n(i\beta r)}{\eta \hat{K}_n'(i\beta a)} P_n^1(\cos\theta),$$

$$i\omega\epsilon r^2 E_r = V\sum_{n=1}^{\infty} A_n \frac{n(n+1)\hat{K}_n(i\beta r)}{\eta \hat{K}_n'(i\beta a)} P_n(\cos\theta).$$

The conduction current in the sphere, flowing in the direction of *decreasing* coordinate θ, is

$$I(\theta) = \sum_{n=1}^{\infty} I_n(\theta), \quad I_n(\theta) = 2\pi A_n M_n V \sin\theta\, P_n^1(\cos\theta), \tag{9–4}$$

where M_n is the radial admittance of the nth zonal wave at the surface of the sphere

$$M_n = -\frac{\hat{K}_n(i\beta a)}{\eta \hat{K}_n'(i\beta a)} = \frac{\hat{J}_n(\beta a) - i\hat{N}_n(\beta a)}{\eta[\hat{N}_n'(\beta a) + i\hat{J}_n'(\beta a)]}.$$

The conjugate complex power flow across the aperture of the cone is

$$\Psi^* = \tfrac{1}{2}\int_\psi^{\psi+\vartheta} aE_\theta^*(a,\theta)I(\theta)\, d\theta$$

$$= \frac{1800\pi V^2}{K^2}\sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} M_n[P_n(\cos\overline{\psi+\vartheta}) - P_n(\cos\psi)]^2.$$

From this we obtain the average admittance at the aperture of the cone line $Y_{av} = 2\Psi^*/V^2$. This value may be taken as the approximate value of the terminal admittance $Y_t \simeq Y_{av}$. The exact value would be obtained if we used the exact expression for $E_\theta(a,\theta)$ instead of the approximation (1).

Let us compare this expression with the ratio of the total current $I(\psi)$ flowing across the edge of the upper hemisphere to the transverse voltage

$$\frac{I(\psi)}{V} = \frac{60\pi}{K}\sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} M_n \sin\psi\, P_n^1(\cos\psi)[P_n(\cos\overline{\psi+\vartheta}) - P_n(\cos\psi)]. \tag{9–5}$$

When $\psi + \vartheta/2 = \pi/2$ this ratio represents one component of the terminal admittance as defined by (2–3). For other values of ψ we should take the average value* $[I(\psi) + I(\psi+\vartheta)]/2V$. The characteristic impedance of the cone line is

$$K = 60\log\left(\cot\frac{\psi}{2}\tan\frac{\psi+\vartheta}{2}\right) = 60\log\frac{\sin\left(\psi+\frac{\vartheta}{2}\right) + \sin\frac{\vartheta}{2}}{\sin\left(\psi+\frac{\vartheta}{2}\right) - \sin\frac{\vartheta}{2}}.$$

* The inequality of $I(\psi)$ and $I(\psi+\vartheta)$ is due to secondary waves and affects the input impedance only indirectly through these waves. For the principal current $I_0(\psi) = I_0(\psi+\vartheta)$.

When ϑ is small compared with $\psi + \vartheta/2$, then we have approximately

$$K = \frac{120 \sin \dfrac{\vartheta}{2}}{\sin\left(\psi + \dfrac{\vartheta}{2}\right)} = \frac{60\vartheta}{\sin\left(\psi + \dfrac{\vartheta}{2}\right)}.$$

In this case we also have, except when n is large,

$$P_n(\cos\overline{\psi + \vartheta}) - P_n(\cos\psi) = \vartheta P_n^1\left[\cos\left(\psi + \frac{\vartheta}{2}\right)\right]. \tag{9-6}$$

Substituting these approximations in $Y_t = Y_{av}$ and in (5), we obtain

$$Y_t = \pi \sin^2\left(\psi + \frac{\vartheta}{2}\right) \sum \frac{2n+1}{n(n+1)} M_n \left\{P_n^1\left[\cos\left(\psi + \frac{\vartheta}{2}\right)\right]\right\}^2,$$

$$\frac{I(\psi)}{V} = \pi \sin\psi \sin\left(\psi + \frac{\vartheta}{2}\right) \sum \frac{2n+1}{n(n+1)} M_n P_n^1(\cos\psi) P_n^1\left[\cos\left(\psi + \frac{\vartheta}{2}\right)\right]. \tag{9-7}$$

The two expressions tend to become equal as ϑ approaches zero. For large values of n the original terms in the expansions for Y_t and $I(\psi)/V$ must be retained since (6) will no longer be a good approximation. While the individual terms are small, the approximation (6) will lead to a divergent series if used for unrestricted values of n.

Thus we have the following approximate expression for the terminal admittance

$$Y_t = \frac{3600\pi}{K^2} \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} M_n [P_n(\cos\overline{\psi + \vartheta}) - P_n(\cos\psi)]^2. \tag{9-8}$$

When computing this expression the first few terms may be replaced by the corresponding terms in the series (7) for Y_t. The terminal admittance tends to infinity as ϑ approaches zero; this is as it should be since the capacitance between the two hemispheres will increase indefinitely as the distance between the edges diminishes. On the other hand, Fig. 11.4 shows that when n increases, G_n rapidly approaches zero and only a few terms are needed to compute G_t quite accurately.

If $\psi = (\pi - \vartheta)/2$, only the odd terms in (8) remain; thus

$$Y_t = \frac{7200\pi}{K^2} \sum_{m=0}^{\infty} \frac{4m+3}{(m+1)(2m+1)} M_{2m+1} \left[P_{2m+1}\left(\cos\frac{\pi - \vartheta}{2}\right)\right]^2. \tag{9-9}$$

For small values of ϑ the real part may conveniently be computed from (7); thus

$$G_t = \frac{\pi}{2} \sum_{m=0}^{\infty} \frac{4m+3}{(m+1)(2m+1)} [P_{2m+1}^1(0)]^2 G_{2m+1}(\beta a). \tag{9-10}$$

The values of the associated Legendre functions and the A's are

$$P_{2m+1}^1(0) = (-)^{m+1} \frac{2(m+\frac{1}{2})!}{m!(-\frac{1}{2})!} = (-)^{m+1} \frac{1 \cdot 3 \cdot 5 \cdots (2m+1)}{2 \cdot 4 \cdot 6 \cdots 2m},$$

$$A_{2m+1} = -\frac{30(4m+3)P_{2m+1}\left(\cos\dfrac{\pi - \vartheta}{2}\right)}{(m+1)(2m+1)K} \simeq \frac{(4m+3)P_{2m+1}^1(0)}{4(m+1)(2m+1)},$$

$$A_1 = -\tfrac{3}{4}, \quad A_3 = \tfrac{7}{16}, \quad A_5 = -\tfrac{11}{32}, \cdots A_{2m+1} \to \frac{(-)^{m+1}}{\sqrt{m\pi}}\left(1 - \frac{3}{8m}\right).$$

Hence the first two terms of $I(\theta)$ are

$$I_1(\theta) = \frac{3\pi}{2} M_1 V \sin^2\theta, \quad I_3(\theta) = -\frac{21\pi}{4} M_3 V \sin^2\theta(1 - \tfrac{5}{4}\sin^2\theta).$$

The amplitude of the first current wave is maximum at the edges of the hemispheres and vanishes at their poles; the next wave has an additional maximum and an additional node. When the circumference of the sphere equals the wavelength, then $\beta a = 1$ and

$$M_1 = \frac{1+i}{120\pi}, \quad M_3 = \frac{0.00047 + i0.359}{120\pi};$$

$$I_1(\theta) = \frac{V(1+i)}{80} \sin^2\theta, \quad I_3(\theta) \simeq -i0.0157 V \sin^2\theta(1 - \tfrac{5}{4}\sin^2\theta).$$

The current wave in phase with V is given almost entirely by the real part of $I_1(\theta)$; on the other hand the reactive component of I_3 may be larger in magnitude than that of I_1.

We shall now prove that as ϑ approaches zero, B_t increases indefinitely as $\log 1/\vartheta$. If ϑ is very small and λ not too large, the principal term in the capacitance between two spheres must be equal to the principal term in the capacitance between two halves of a 180-degree wedge transmission line (section 8.7). Thus, if $s = a\vartheta$ is the distance between the edges of the hemispheres, $B_t \to 2i\omega\varepsilon a \log \lambda/s$ as $s \to 0$. The following direct method of proving this asymptotic property is also useful for computing B_t from (8). For large values of n we may use the asymptotic formulae

$$P_n(\cos\theta) \sim \sqrt{\frac{2}{n\pi \sin\theta}} \sin\left[(n + \tfrac{1}{2})\theta + \frac{\pi}{4}\right], \quad M_n \sim \frac{i\omega\varepsilon a}{n}.$$

Assuming that $\psi + \vartheta/2 = \pi/2$, substituting in the imaginary part of (8), rejecting the first term, expanding the coefficient in a power series, and retaining only the principal term, we obtain

$$B_t \simeq \frac{i\omega\varepsilon a}{\vartheta^2} \sum_{m=1}^{\infty} \frac{1 - \cos 2m\vartheta}{m^3} \simeq 2i\omega\varepsilon a\left(\log\frac{1}{2\vartheta} + 1.5\right).$$

As m increases the difference between the corresponding terms in the exact series for B_t and in the above approximate series approaches zero as $1/m^2$; hence the series formed by subtracting one of these series from the other converges more rapidly than the original series.

An approximate formula for the conductance of a large spherical antenna can also be obtained by regarding the sphere as a 180-degree wedge transmission line; thus

$G_t \simeq \beta a/120 = \pi a/60\lambda$. This approximation is fairly good even for moderate values of βa.

The outstanding feature of the spherical antenna which is shared by other "broad surface" radiators is the comparatively slow variation of its impedance with frequency as contrasted with the rapid variation of the impedance of thin wires.

11.10. *The Reciprocity Theorem*

The Reciprocity Theorem may be briefly stated as follows: *the positions of an impedanceless generator and ammeter may be interchanged without affecting the ammeter reading.* One type of proof is very similar to the proof of the Reciprocity Theorem for electric networks (see section 5.1). In this section we shall prove the theorem first for transmission lines and then for antennas.

Consider a transmission line of length l and let $E_1(x)$ and $E_2(x)$ be two distributions of applied series electromotive forces; then

$$\frac{dV_1}{dx} = -ZI_1 + E_1(x), \quad \frac{dI_1}{dx} = -YV_1,$$
$$\frac{dV_2}{dx} = -ZI_2 + E_2(x), \quad \frac{dI_2}{dx} = -YV_2.$$

Multiplying the first equation by I_2, the last by V_1 and adding, we obtain

$$\frac{d}{dx}(V_1I_2) = -ZI_1I_2 - YV_1V_2 + E_1I_2.$$

Integrating from $x = 0$ to $x = l$, and rearranging the terms, we have

$$\int_0^l E_1I_2\,dx = \int_0^l ZI_1I_2\,dx + \int_0^l YV_1V_2\,dx + V_1I_2\Big|_0^l.$$

The first two terms on the right are obviously symmetric in the subscripts; the last is also symmetric since

$$V_1I_2\Big|_0^l = Z(l)I_1(l)I_2(l) - Z(0)I_1(0)I_2(0),$$

where $Z(0)$ and $Z(l)$ are the terminal impedances. Therefore

$$\int_0^l E_1I_2\,dx = \int_0^l E_2I_1\,dx.$$

This is the general Reciprocity Theorem for transmission lines.

In the special case when $E_1(x)$ and $E_2(x)$ are concentrated in infinitely short intervals at $x = \xi_1$ and $x = \xi_2$, we have

$$I_2(\xi_1)\int_{\xi_1-0}^{\xi_1+0} E_1\,dx = I_1(\xi_2)\int_{\xi_2-0}^{\xi_2+0} E_2(x)\,dx,$$

or

$$I_2(\xi_1)V_1^i(\xi_1) = I_1(\xi_2)V_2^i(\xi_2).$$

If the impressed electromotive forces are equal, then $I_2(\xi_1) = I_1(\xi_2)$.

In three dimensions we write

$$\operatorname{curl} E_1 = -M_1 - i\omega\mu H_1, \quad \operatorname{curl} H_1 = (g + i\omega\epsilon)E_1,$$
$$\operatorname{curl} E_2 = -M_2 - i\omega\mu H_2, \quad \operatorname{curl} H_2 = (g + i\omega\epsilon)E_2.$$

Multiplying scalarly the first equation by H_2, the last by E_1, and subtracting,* we have

$$E_1 \cdot \operatorname{curl} H_2 - H_2 \cdot \operatorname{curl} E_1 = M_1 \cdot H_2 + i\omega H_1 \cdot H_2 + (g + i\omega\epsilon)E_1 \cdot E_2.$$

The left side equals div $(H_2 \times E_1)$; multiplying by dv, integrating over a volume (v), replacing the volume integral of the divergence by the surface integral of the normal component, and rearranging the terms, we have

$$\iiint M_1 \cdot H_2\,dv = -\iiint i\omega\mu H_1 \cdot H_2\,dv$$
$$-\iiint (g + i\omega\epsilon)E_1 \cdot E_2\,dv + \iint (H_2 \times E_1)_n\,dS,$$

where (S) is the boundary of (v). The first two terms on the right are symmetric in the subscripts. The last term is also symmetric since we are free to choose (v) as an infinite sphere, in which case $E_\theta = \eta H_\varphi$, $E_\varphi = -\eta H_\theta$, and consequently

$$\iint (H_2 \times E_1)_n\,dS = -\eta \iint (H_{\theta,1}H_{\theta,2} + H_{\varphi,1}H_{\varphi,2})\,dS,$$

and (10–1)

$$\iiint M_1 \cdot H_2\,dv = \iiint M_2 \cdot H_1\,dv.$$

This equation holds even if g, μ, and ϵ are functions of position.

Consider now a special case of two antennas (Fig. 11.34) energized at points A and B. Two voltages $V_1(A)$ and $V_2(B)$ may be applied by means of two infinitely small magnetic current loops round points A and B, carrying the following magnetic currents

$$K_1(A) = -V_1(A), \quad K_2(B) = -V_2(B).$$

Under these conditions (1) becomes

$$V_1(A)I_2(A) = V_2(B)I_1(B).$$

* Subtracting rather than adding since the electromagnetic equations are antisymmetric.

The above is the most frequently used reciprocity theorem. Another theorem is: *the positions of a generator and a voltmeter, both of infinite impedance, may be interchanged without affecting the voltmeter reading.* The more general form of this theorem for transmission lines is

$$\int_0^l J_1 V_2\, dx = \int_0^l J_2 V_1\, dx,$$

where $J_1(x)$ and $J_2(x)$ are two current distributions applied in shunt with the line. Similarly in the three dimensional case we have

$$\iiint J_1 \cdot E_2\, dv = \iiint J_2 \cdot E_1\, dv.$$

One corollary of the Reciprocity Theorem is: the directivity patterns of an antenna are the same whether the antenna is used as a transmitter or as a receiver. This follows at once from the definition of such patterns. The directive pattern of a transmitter is explored at an "infinite" distance by a tuned doublet so oriented that maximum power is received by the doublet. The directive pattern of a receiver is found by obtaining its response to plane waves arriving from different directions, the electric vectors of these waves being so oriented that maximum power is received by the antenna.

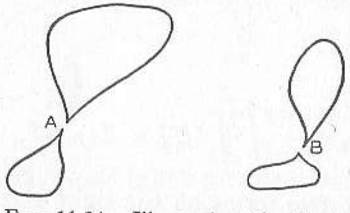


FIG. 11.34. Illustrating the Reciprocity Theorem.

11.11. *Receiving Antennas*

In the case of a receiving antenna the electromotive force is impressed at all points of the antenna and the complete theory of a receiving antenna depends on the solution of the most general transmitter problem. On the other hand, the solution for a transmitting antenna energized at the center is sufficient, in view of the Reciprocity Theorem, to enable us to find the performance of the same antenna when the electromotive force is removed and the internal impedance of the generator is used as a load.

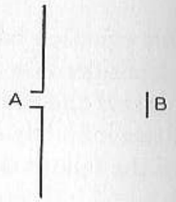


FIG. 11.35. Illustrating the Reciprocity Theorem.

Thus consider an antenna and a tuned current element at a large distance from it (Fig. 11.35), and suppose first that the antenna is being used as a transmitter. Let p be the moment of the current distribution in the antenna when an electromotive force V is acting at A, $p = \int I(z)\, dz$. The voltage V impressed on the element and hence the current in it are

$$V_1(B) = \frac{i\eta p l e^{-i\beta r}}{2\lambda r}, \quad I_1(B) = \frac{i\eta p l e^{-i\beta r}}{2\lambda r R},$$

where R is the radiation resistance of the element. Now let the current element act as transmitter with the antenna as receiver, then the current in the element and consequently the electric intensity impressed on the antenna are

$$I_2(B) = \frac{V}{R}, \quad E = \frac{i\eta V l e^{-i\beta r}}{2\lambda r R}.$$

By the Reciprocity Theorem the current at A in response to this field is equal to $I_1(B)$; expressing this in terms of E, we have

$$I_2(A) = I_1(B) = \frac{pE}{V} = \frac{E\int I(z)\, dz}{V} = \frac{Eh}{Z_A + Z_L},$$

where h is the effective height of the antenna, and Z_A, Z_L are respectively the antenna and load impedances. Thus the current in the antenna when used as a receiver has been expressed in terms of the current distribution in the antenna when used as a transmitter. In applying the Reciprocity Theorem care should be taken not to change the impedances.

CHAPTER XII

The Impedance Concept

12.1. *In Retrospect*

Most concepts grow with their use. Originally the word " number " meant what we now call " integer "; then in successive steps the meaning has been broadened to include rational fractions, irrational fractions, negative numbers, and finally complex numbers. The impedance concept is no exception and it has grown considerably in the half-century of its existence. The term " impedance " was proposed in 1886 by Oliver Heaviside for the voltage/current amplitude ratio in a circuit comprised of a resistor and an inductor.* In 1889 a further impetus to its use in the same sense was given by Oliver Lodge.† Three years later F. Bedell and A. Crehore proposed the term " impediment " for a similar ratio in a circuit including a capacitor as well as a resistor and an inductor. However, there was no real need for a separate term and the word " impedance " was soon used in a broader sense. It was not very long before it was used for any voltage/current ratio, expressed as a complex number with its absolute value equal to the amplitude ratio and its phase to the phase difference between the voltage and current. It is outside the scope of the present book to follow the historical development of the impedance concept and the above remarks have been made only to give a probable date of its inception. What concerns us primarily is the actual meaning of the concept and its uses.

First let us consider stationary fields and the simpler concept resistance.

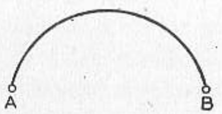


Fig. 12.1. A wire between two terminals.

The resistance of a conductor AB (Fig. 12.1) is defined as V/I, where V is the electromotive force along AB and I is the current in the conductor. The current in the conductor is the same across all cross-sections and the voltage between the terminals AB is independent of the path along which the electric intensity is integrated. Thus our definition of the resistance requires no further qualification.

The situation becomes quite different when an alternating voltage is applied between the terminals A and B. The voltage depends on the path joining the terminals and the current varies along the conductor.

* The Electrician, July 23, 1886, p. 212.
† Electrical Review, May 3, 1889.

If the distance between the terminals is small (Fig. 12.2), the voltage is nearly the same for paths going " more or less directly " from A to B and the current flowing out of A, let us say, is nearly equal to the current flowing into B. The input impedance is then defined as the voltage divided by the input current. In practice, the distance between the terminals is " small " if the voltage is nearly independent of the path from one terminal to the other. In theory we might be tempted to assume that the terminals are infinitely close; but unfortunately the capacitance, and hence the admittance, between terminals of finite size tends to infinity as the distance between them approaches zero. This is a purely theoretical situation; in practice the distance is merely small and the capacitance is also usually small. A way out of this difficulty is found by assuming that the terminals are tapered off to mere points (Fig. 12.3). If such conical terminals are coaxial, then the admittance between them, ignoring the rest of the circuit, is

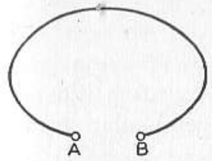


Fig. 12.2. A wire whose terminals are close together.

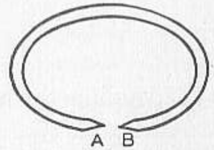


Fig. 12.3. A wire with pointed terminals which can be brought infinitely close together without making the capacitance between the terminals infinite.

$$Y = i\frac{2\pi l}{\lambda K} = i\frac{\pi l}{60\lambda \log \dfrac{1+\cos\psi}{\sin\psi}} = i\frac{\pi\sqrt{a^2+h^2}}{60\lambda \log \dfrac{h+\sqrt{a^2+h^2}}{a}}, \qquad (1\text{–}1)$$

where h and a are respectively the height and the maximum radius of each cone. If $h = a$, we have

$$Y = i\frac{\pi\sqrt{2}a}{60\lambda \log(1+\sqrt{2})} = i0.0840\frac{a}{\lambda}\text{ mhos.} \qquad (1\text{–}2)$$

This admittance is very small unless the waves are very short; even if it is doubled or trebled by increasing h, it will have a negligible effect on the impedance of the circuit unless the latter is very large.

If a is kept constant while h approaches zero, we have

$$Y \to \frac{i\pi a^2}{60\lambda h} = i0.0524\frac{a^2}{\lambda h}; \qquad (1\text{–}3)$$

this admittance is equal to the admittance of two parallel circular discs of radius a, separated by distance h. As h approaches zero its value increases indefinitely but, except for very short waves, h has to be very small indeed before the impedance of the entire circuit is noticeably affected.

The conception of impedance does not depend on the existence of terminals. Thus in dealing with the transmission of waves on perfectly conducting parallel wires (Fig. 12.4), the series impedance per unit length of the transmission line is defined as the following ratio

$$Z = \frac{V_{BC} - V_{AD}}{Il}, \quad l \doteq AB. \tag{1–4}$$

In defining the transverse voltage V_{BC} the path of integration must be restricted to the plane perpendicular to the wires; it does not have to lead directly from one wire to the other and may deviate considerably from a straight path as long as it remains in the transverse plane; but it should not leave the plane. The transverse voltage can be measured directly only when the separation between the wires is small and the impedance of the leads connecting B and C with the terminals of the voltmeter does not appreciably affect the measurement.

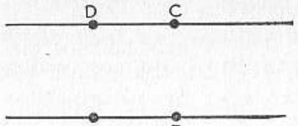


FIG. 12.4. Illustrating the impedance per unit length of a transmission line.

The impedance per unit length of the parallel pair of wires as defined above does not have quite the same meaning as the resistance per unit length of a wire, or the impedance per unit length of a long coil. In the latter cases there exist actual voltages between two points on the wire or between two turns of the coil, and the impedance is similar to the input impedance of a circuit. But in the transmission line consisting of perfectly conducting wires, the voltage between A and B along the wire is zero; so is the voltage between C and D. Yet the line as a whole acts as if it had an impedance between a "terminal" A, D and another terminal B, C.

The next step in extending the impedance concept was to include ratios of electric and magnetic intensities. This extension has served to unify many diverse problems concerning the reflection of electromagnetic waves so that each problem has become a special case of a general problem. Consider for example a metal tube (or a coaxial pair) filled with two homogeneous dielectrics and let the boundary between the dielectrics be perpendicular to the axis of the tube (Fig. 12.5). For each transmission mode the field distribution in transverse planes is independent of the constants of the dielectric and the reflection and transmission coefficients depend solely on the ratio K_z''/K_z' of the wave impedances in the direction of the tube. The same general formula applies when plane waves are incident normally or obliquely to the plane interface between two media, provided either E or H is parallel to the

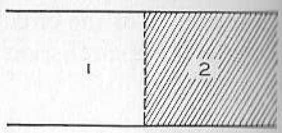


FIG. 12.5. Reflection in wave guides at a simple discontinuity.

interface; it applies when cylindrical waves are incident normally on cylindrical bodies and when spherical waves are incident normally on spherical bodies.

The impedance normal to a surface separating two media is defined as the ratio of the tangential electric intensity to the tangential magnetic intensity. The impedance normal to a thin film with a sheet of infinite impedance just behind it is equal to the surface impedance of the film itself, that is, the ratio of the electric intensity to the current density. In general the only practical means of approximating in effect a sheet of infinite impedance behind the film is to place a good conductor a quarter wavelength (or an odd number of quarter wavelengths) behind the film. The wavelength in question should be that in the direction normal to the film.

Any wave guide may be terminated in a resistance film, normal to its axis, and designed to absorb all the energy carried by a wave traveling in one particular transmission mode; we need only make the surface resistance R_s of the film equal to the wave impedance of the incoming wave

$$R_s = K_z. \tag{1–5}$$

The surface resistance of a thin film of thickness t is substantially equal to its d–c resistance $R_s = 1/gt$. Since the conductivity of very thin films is not necessarily equal to the conductivity of the material of the film in bulk, R_s should be measured.

For the dominant mode in air-filled coaxial pairs we have $K_z = 120\pi$ and the impedances are matched when $R_s = 120\pi$. The d–c resistance of an annular film is

$$R = \frac{R_s}{2\pi} \log \frac{b}{a}, \tag{1–6}$$

where a and b are the radii of the boundaries of the film. If the radii of the coaxial pair are equal to a and b, then the matching condition becomes $R = K$, where K is the integrated characteristic impedance of the coaxial pair.

In the case of circular wave guides the matching resistance films are circular discs. To determine the surface resistance of these discs the d–c resistance between a pair of concentric circular electrodes in contact with the discs may be measured. Since from (6) we have

$$R_s = \frac{2\pi R}{\log \dfrac{b}{a}},$$

where a and b are the radii of the electrodes, the matching condition (5) may be expressed in terms of this measured resistance

$$R = \frac{K_z}{2\pi}\log\frac{b}{a}.$$

For transverse magnetic and transverse electric waves this becomes respectively

$$R = \frac{\eta\sqrt{1-\nu^2}}{2\pi}\log\frac{b}{a}, \quad R = \frac{\eta}{2\pi\sqrt{1-\nu^2}}\log\frac{b}{a}.$$

In the case of rectangular guides the matching films are rectangular. If the dimensions of the rectangle are a and b and the d–c resistance R is measured between the sides having the length a, then $R = (b/a)R_s$ and the matching condition (5) becomes $R = (b/a)K_z$.

We have seen that a sufficiently thick conductor can be regarded as an impedance sheet whose surface impedance is equal to the intrinsic impedance $\hat{\eta}$ of the conductor. If the thickness t of the conductor is small compared with the radius of curvature, then the surface impedance is $\hat{\eta}$ coth σt. When an actual conductor is replaced by an impedance sheet, consisting of a sheet of finite impedance over a sheet of infinite impedance, many problems are simplified. No energy can flow across an infinite impedance and the media on either side of the impedance sheet may be treated as electromagnetically independent. Also it becomes unnecessary to consider the field in the conductor itself.

12.2. *Wave Propagation between Two Impedance Sheets*

Consider two parallel impedance sheets CD and $C'D'$ (Fig. 12.6) and let the surface impedance of each be Z. Then the field will be symmetric

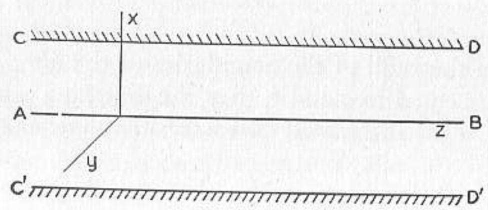


FIG. 12.6. Parallel impedance sheets.

with respect to the plane AB midway between the sheets. If we confine our attention to fields in which the H-lines are parallel to the y-axis, the non-vanishing field intensities will be E_x, E_z, H_y. From the assumed symmetry about the central plane we conclude that the electric lines are either perpendicular or parallel to this plane. In the former case a perfect conductor and in the latter a sheet of infinite impedance may be inserted at the plane, without disturbing the field distribution.

When the yz-plane is a perfect conductor, then for progressive waves in the positive z-direction we have

$$H_y = J\cosh\gamma x\, e^{-\Gamma z}, \qquad \Gamma^2 + \gamma^2 = -\beta^2,$$
$$E_x = J\frac{\Gamma}{i\omega\epsilon}\cosh\gamma x\, e^{-\Gamma z}, \quad E_z = J\frac{\gamma}{i\omega\epsilon}\sinh\gamma x\, e^{-\Gamma z}, \tag{2–1}$$

where Γ and γ are respectively the longitudinal and transverse propagation constants and J is the conduction current density in AB. The condition for natural waves is then

$$\frac{\gamma}{i\omega\epsilon}\tanh\gamma b = -Z, \tag{2–2}$$

where b is the distance between AB and CD. Thus we have the following expressions for the propagation constants

$$\gamma = \frac{w}{b}, \quad \Gamma = \sqrt{-\frac{w^2}{b^2} - \beta^2}, \tag{2–3}$$

where w is a root of the following equation

$$w\tanh w = -i\omega\epsilon bZ = -\frac{i\beta bZ}{\eta}. \tag{2–4}$$

When $Z = 0$, we have $w = n\pi i$. If $n = 0$, $\gamma = 0$, E_z vanishes identically and we have the principal wave. If $|\beta bZ|$ is small compared with η, then one root of (4) is small compared with unity and we have approximately

$$w^2 = -\frac{i\beta bZ}{\eta}, \quad \Gamma = \sqrt{-\beta^2 + \frac{i\beta Z}{\eta b}} = \sqrt{(i\omega\mu b + Z)\frac{i\omega\epsilon}{b}}. \tag{2–5}$$

We have already seen that this formula can be obtained directly from the integral equations of electromagnetic induction by assuming that the longitudinal displacement current is zero. Actually the longitudinal electric intensity varies almost linearly with x and the magnetic intensity is approximately constant

$$E_z = -ZJ\frac{x}{b}e^{-\Gamma z}, \quad H_y = J\left(1 - \frac{i\beta bZ}{2\eta}\frac{x^2}{b^2}\right)e^{-\Gamma z}.$$

Then the current density in the sheet CD is $-J + (i\beta bZ/2\eta)J$; thus this density is slightly different from the current density in the sheet AB.

At the other extreme when the right-hand side of (4) is infinite we have $w = i(n + \frac{1}{2})\pi$, and when the right-hand side is large but finite, then

$$w = i(n + \tfrac{1}{2})\pi \left(1 + \frac{i\eta}{\beta b Z}\right).$$

For the principal wave $n = 0$ and $w = i\pi/2 - \pi\eta/2\beta bZ$. The longitudinal electric intensity is nearly sinusoidal, rising from zero at the plane AB to a maximum at the plane CD. The current in the sheet CD nearly vanishes and the dielectric between the planes serves as the "return circuit."

In solving (4) for intermediate values of the right-hand side we shall consider only the case in which Z is a pure resistance

$$Z = R, \quad k = \frac{\beta bR}{\eta} = \frac{bR}{60\lambda},$$

the numerical coefficient in the last equation corresponding to free space between the planes. Since w is complex, we write $w = u + iv$. Substituting in (4) and separating the real and imaginary parts, we have

$$u \sinh u \cos v - v \cosh u \sin v = k \sinh u \sin v,$$

$$v \sinh u \cos v + u \cosh u \sin v = -k \cosh u \cos v.$$

Eliminating k, we obtain

$$u \sinh 2u = v \sin 2v. \qquad (2\text{–}6)$$

Since u is essentially real, the left side of this equation is positive and the values of v are confined to the following intervals

$$0 \leq v \leq \frac{\pi}{2}, \quad \pi \leq v \leq \frac{3\pi}{2}, \cdots n\pi \leq v \leq (n + \tfrac{1}{2})\pi.$$

Thus the functional relationship between the transverse attenuation constant and the transverse phase constant is independent of R; hence this is also true for the longitudinal attenuation and phase constants. At each end of the admissible v-intervals $u = 0$. From (5) we conclude that in the present case u and v are of opposite sign; and since v has been assumed positive, u must be negative. The left side of (6) is a monotonic increasing function of u and consequently u is minimum when the right-hand side is maximum, that is, when $\tan 2v = -2v$. The smallest root is 1.015^- and the corresponding maximum value of $v \sin 2v$ is 0.91. Thus in the entire first interval the absolute value of u is less than unity and a rough approximation for u is

$$u = -\sqrt{\tfrac{1}{2}v \sin 2v}. \qquad (2\text{–}7)$$

This approximation is at its worst in the vicinity of the minimum value of u; thus from equation (7) this value of u is found to be -0.67 while the exact value is near -0.60. Figure 12.7 shows the first branch of the curve

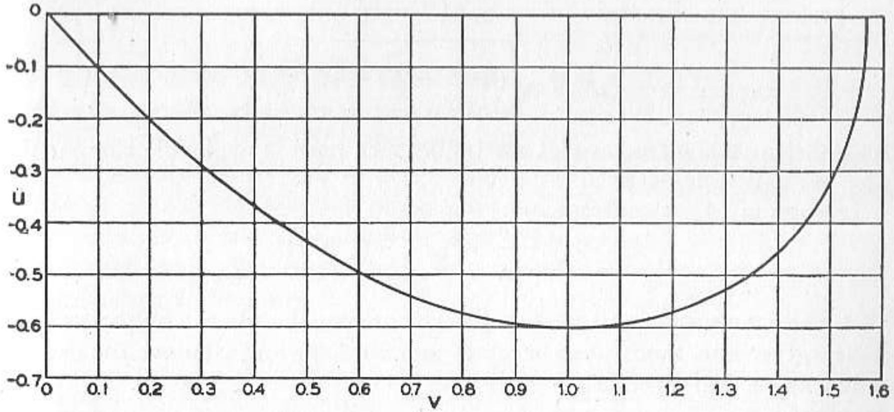


FIG. 12.7. Transverse attenuation constant vs. transverse phase constant.

represented by equation (6). For each pair of values of u and v, we obtain the longitudinal propagation constant from

$$\Gamma = \frac{1}{b}\sqrt{-\beta^2 b^2 - u^2 + v^2 - 2iuv}. \qquad (2\text{–}8)$$

Taking the absolute value of (4), we have

$$k = \frac{\beta bR}{\eta} = \sqrt{\frac{(u^2 + v^2)(\cosh 2u - \cos 2v)}{\cosh 2u + \cos 2v}}.$$

This equation enables us to plot k as a function of either u or v; Fig. 12.8 represents k as a function of v for the principal wave. It should be noted

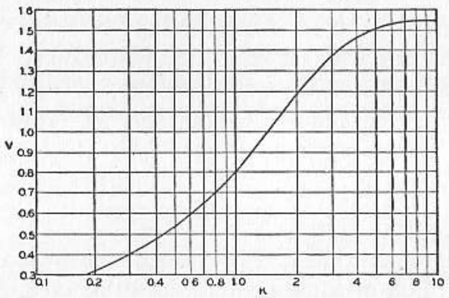


FIG. 12.8. Transverse phase constant vs. impedance ratio.

that while k may vary from zero to infinity, the ranges of u and v are limited for each transmission mode.

Consider now the principal wave as the frequency increases from zero to infinity while R remains constant. In the lower frequency range k is small and the propagation constant is given substantially by (5) which in the present case becomes

$$\Gamma = \sqrt{(i\omega\mu b + R)\frac{i\omega\epsilon}{b}}.$$

At sufficiently low frequencies the inductance term is negligible compared with the resistance and

$$\Gamma = \sqrt{\frac{i\omega\epsilon R}{b}} = \sqrt{\frac{\omega\epsilon R}{2b}}\,(1 + i).$$

Thus the attenuation and phase constants are equal. As the frequency increases, the inductance term becomes predominant and we have the usual transmission line formula

$$\Gamma = i\omega\sqrt{\mu\epsilon} + \frac{R}{2\eta b}.$$

As the frequency becomes so high that k is no longer small, Γ should be obtained from the exact formula (8). In this formula u and v are finite while βb increases indefinitely; hence the attenuation constant approaches zero and the phase constant remains approximately equal to the intrinsic phase constant. Over the entire frequency range we have roughly

$$\Gamma = \frac{1}{b}\sqrt{-\beta^2 b^2 + v^2\left(1 - \frac{\sin 2v}{2v}\right) + iv\sqrt{2v \sin 2v}}.$$

If the frequency is kept constant but R is made to increase from zero to infinity, the behavior of Γ will depend on the magnitude of βb in comparison with unity. When βb is large, the longitudinal phase constant is approximately equal to β in the entire range while the attenuation constant varies from zero to a maximum and then back to zero. If βb is not very large, but larger than $\pi/2$, the phase constant varies from β at $R = 0$ to $\sqrt{\beta^2 - \pi^2/4b^2}$ at $R = \infty$. Finally if $\beta b < \pi/2$, then the phase constant diminishes from β at $R = 0$ to zero at $R = \infty$; at the same time the attenuation constant rises from zero to $\sqrt{\pi^2/4b^2 - \beta^2}$.

It should be remembered that R is the impedance normal to the sheet CD and, if CD is a thin film of some conducting substance, R is equal to the surface resistance of the film itself only when the impedance normal to the sheet on the other side of it is infinite. This condition is obtained, for example, for all values of R when the resistance film is placed half-way between two perfectly conducting planes (Fig. 12.9) carrying equal currents in the same direction. At the plane of the film the magnetic intensity vanishes and we have in effect a surface of infinite impedance. The above equations will apply to this case if $R = 2R_{\text{film}}$.

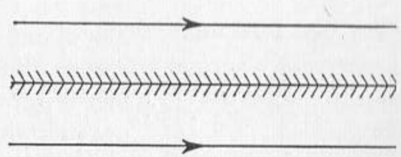


FIG. 12.9. Resistance sheet between perfectly conducting planes.

The above discussion has been based directly on the solutions of the differential equations (4.12–18) appropriate to the actual problem. For each transmission mode these equations can be converted into the conventional form of transmission line equations in which the x-coordinate is suppressed. Eliminating E_z we have

$$\frac{\partial E_x}{\partial z} = -\left(i\omega\mu - \frac{w^2}{i\omega\epsilon b^2}\right)H_y, \quad \frac{\partial H_y}{\partial z} = -i\omega\epsilon E_x.$$

Integrating each of these equations with respect to x from $x = 0$ to $x = b$, we have

$$\frac{dV}{dz} = -\left(\frac{i\omega\mu b p}{a} + \frac{(v^2 - u^2)p}{i\omega\epsilon ab} - \frac{2uvp}{\omega\epsilon ab}\right)I, \quad \frac{dI}{dz} = -\frac{i\omega\epsilon a}{pb}V,$$

$$V = \int_0^b E_x\,dx, \quad p = \frac{a}{bI}\int_0^b H_y\,dx.$$

I is the conduction current in the plane AB between $y = 0$ and $y = a$; the constant p is unity when H_y is uniform. When k is small, p is nearly unity, the second term in the expression for the distributed series impedance is negligible, the third term becomes R/a, and the equations reduce to the engineering form based on neglecting the longitudinal displacement currents. When k is comparable to unity, the expressions for the distributed series impedance and shunt admittance become complicated; as k becomes large, these expressions are again simplified.

If the AB-plane is a surface of infinite impedance, the equation for the transverse propagation constant becomes

$$\frac{\gamma}{i\omega\epsilon}\coth \gamma b = -Z.$$

This equation has no solution in the vicinity of $\gamma = 0$ and it defines transmission modes similar to the higher transmission modes in the preceding case

12.3. *On Impedance and Reflection of Waves at Certain Irregularities in Wave Guides*

Let us consider a wave guide of rectangular cross-section (Fig. 12.10) bounded by two planes of zero impedance $x = 0$, $x = a$, and two planes of infinite impedance $y = 0$, $y = b$. Such a wave guide is an idealization of a pair of parallel conducting strips; the infinite impedance sheets serve to eliminate the edge effect, and thus simplify the mathematical problem. The transverse electric intensity in the plane $z = 0$ is, in general, an arbitrary function of x and y. This function determines completely the field in the wave guide. We shall take a simple case as an example and assume that $E_x = 0$, $E_y(x,y,0) = f(y)$. In this case $H_y = H_z = 0$. Then for an infinitely long guide the field intensities are of the following form

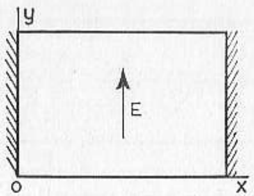


FIG. 12.10. Rectangular wave guide; two faces are of zero impedance and the other two of infinite impedance.

$$E_y = \sum_{n=0}^{\infty} E_n \cos\frac{n\pi y}{b} e^{-\Gamma_n z}, \quad \Gamma_n = \sqrt{\frac{n^2\pi^2}{b^2} - \beta^2},$$

$$H_x = -\sum_{n=0}^{\infty} M_n E_n \cos\frac{n\pi y}{b} e^{-\Gamma_n z}, \quad M_n = \frac{i\omega\epsilon}{\Gamma_n}, \tag{3-1}$$

$$E_z = -\frac{\pi}{i\omega\epsilon b}\sum_{n=1}^{\infty} nM_nE_n \sin\frac{n\pi y}{b} e^{-\Gamma_n z}.$$

In the plane $z = 0$ we have

$$E_y = \sum_{n=0}^{\infty} E_n \cos\frac{n\pi y}{b}, \quad H_x = -\sum_{n=0}^{\infty} M_nE_n \cos\frac{n\pi y}{b}. \tag{3-2}$$

Consequently the conjugate complex power flow across $z = 0$ is

$$\Psi^* = \frac{ab}{2}\left(M_0E_0E_0^* + \tfrac{1}{2}\sum_{n=1}^{\infty} M_nE_nE_n^*\right). \tag{3-3}$$

The coefficients E_n are obtained by expanding $f(y)$ in a cosine series of the form (2); thus

$$E_0 = \frac{1}{b}\int_0^b f(y)\,dy = \frac{V}{b}, \quad E_n = \frac{2}{b}\int_0^b f(y)\cos\frac{n\pi y}{b}\,dy, \tag{3-4}$$

where V is the transverse voltage between the conducting strips. Thus the power flow across the plane $z = 0$ may be expressed in terms of the transverse voltage and the form of its distribution over the plane. The input impedance may now be defined so that

$$\Psi^* = \tfrac{1}{2}Y_iVV^*. \tag{3-5}$$

Substituting from (4) in (3) and using (5), we obtain

$$Y_i = \frac{a}{\eta b} + i\frac{a}{2\eta b}\sum_{n=1}^{\infty}\frac{\beta}{\Gamma_n}\frac{E_nE_n^*}{E_0E_0^*}. \tag{3-6}$$

The first term is the characteristic admittance of the guide to the principal wave. If $\lambda > 2b$, the second term is imaginary. As λ increases indefinitely we have

$$\Gamma_n \to \frac{n\pi}{b}, \quad \frac{\beta}{\Gamma_n} \to \frac{2b}{n\lambda} \to 0,$$

and the reactive part of the input admittance approaches zero regardless of the form of the voltage distribution. Thus for "low frequencies" the input impedance of the guide is nearly equal to its characteristic impedance to the principal wave.

Let us consider a specific numerical example in which the electric intensity E_y in the plane $z = 0$ is zero except in the interval $d < y < d + s$,

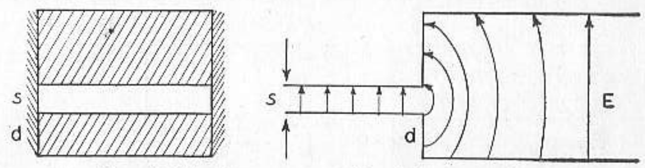


FIG. 12.11. Two wave guides joined together.

where it is V/s. This distribution approximates that at the mouth of a wave guide of height s, joining another wave guide of larger height b (Fig. 12.11). In this case we have

$$\frac{E_n}{E_0} = \frac{2b}{n\pi s}\left[\sin\frac{n\pi(d+s)}{b} - \sin\frac{n\pi d}{b}\right] = \frac{4b}{n\pi s}\sin\frac{n\pi s}{2b}\cos\frac{n\pi(2d+s)}{2b}.$$

When $d + s/2 = b/2$, we obtain

$$E_{2m+1} = 0, \quad \frac{E_{2m}}{E_0} = (-)^m\frac{2b}{m\pi s}\sin\frac{m\pi s}{b}.$$

Substituting in (6), we have $Y_i = G_i + iB_i$, $G_i = M = a/\eta b$, and

$$\frac{B_i}{M} = \frac{b^2}{\pi^2 s^2}\sum_{m=1}^{\infty}\frac{\beta\left(1 - \cos\dfrac{2m\pi s}{b}\right)}{\Gamma_{2m}m^2}$$

As the wavelength increases, this ratio becomes approximately

$$\frac{B_i}{M} = \frac{b^3}{\pi^2 s^2\lambda}\sum_{m=1}^{\infty}\frac{1 - \cos\dfrac{2m\pi s}{b}}{m^3}$$

By (3.7–53) we have

$$\frac{B_i}{M} = \frac{2b}{\lambda}\left[\log\frac{b}{2\pi s} + 1.5 + \frac{1}{144}\left(\frac{2\pi s}{b}\right)^2 + \cdots\right] = \frac{i\omega C_i}{M}.$$

$$C_i = \frac{\epsilon a}{\pi}\left[\log\frac{b}{s} - 0.338 + \frac{1}{144}\left(\frac{2\pi s}{b}\right)^2 + \cdots\right].$$

When $f = 1000$ and $b = 0.02$, $2b/\lambda = 1.3 \times 10^{-7}$. If $s = 0.01b$, then $B_i/M = 5.6 \times 10^{-7}$. Even if f is raised to 10^6, the ratio of the susceptance to the characteristic admittance is still small. But when the frequency becomes so high that λ is comparable to b, then the susceptance may become appreciable.

Consider now a wave guide of uniform rectangular cross-section with a metal diaphragm or iris across it (Fig. 12.12). Again we assume that two faces of the guide are of zero impedance and the other two of infinite impedance. Let a wave in the dominant mode impinge on the iris. The re-

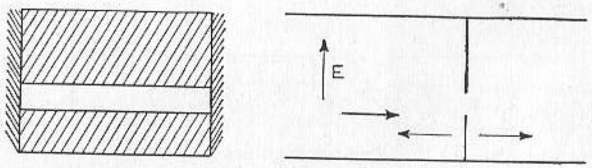


FIG. 12.12. An iris in a rectangular wave guide.

flected and transmitted waves will consist of higher order waves as well as the dominant one. Let the electric intensity of the incident wave in the plane of the iris (the xy plane), be E_0^i. The electric and magnetic intensities of the transmitted field are of the form (2). Since the total electric intensity tangential to the iris is continuous across the plane of the latter, the intensity of the reflected wave in this plane is

$$E_y^r = (E_0 - E_0^i) + \sum_{n=1}^{\infty} E_n \cos\frac{n\pi y}{b}.$$

From this we obtain the magnetic intensity of the reflected field

$$H_x^r = M_0(E_0 - E_0^i) + \sum_{n=1}^{\infty} M_n E_n \cos\frac{n\pi y}{b}.$$

The magnetic intensity is continuous over the aperture of the iris; there we have $H_x^i + H_x^r = H_x$, so that

$$-M_0E_0^i + M_0(E_0 - E_0^i) + \sum_{n=1}^{\infty} M_n E_n \cos\frac{n\pi y}{b} = -\sum_{n=0}^{\infty} M_n E_n \cos\frac{n\pi y}{b}.$$

Transposing the terms, we obtain

$$M_0E_0^i = \sum_{n=0}^{\infty} M_n E_n \cos\frac{n\pi y}{b}. \qquad (3\text{–}7)$$

Multiplying by

$$\sum_{n=0}^{\infty} E_n^* \cos\frac{n\pi y}{b} \qquad (3\text{–}8)$$

and integrating over the iris,* we have

$$M_0E_0^iE_0^*ab = M_0E_0E_0^*ab + \tfrac{1}{2}ab\sum_{n=1}^{\infty} M_nE_nE_n^*,$$

$$\frac{E_0^i}{E_0} = \frac{1}{p} = 1 + \frac{1}{2}\sum_{n=1}^{\infty}\frac{M_n}{M_0}\frac{E_nE_n^*}{E_0E_0^*}.$$

This equation represents the reciprocal of the transmission coefficient across the iris for the dominant wave.

Now the admittance (6) of either half of the wave guide as seen from the iris may be represented in the following form

$$Y_i = M + \hat{Y}, \quad \frac{\hat{Y}}{M} = \frac{1}{2}\sum_{n=1}^{\infty}\frac{M_n}{M_0}\frac{E_nE_n^*}{E_0E_0^*}.$$

Consequently we have $1/p = 1 + \hat{Y}/M$. Comparing with (7.13–10) we find that the transmission coefficient for the dominant wave is the same as if the iris acted as an admittance $Y_s = 2\hat{Y}$ in shunt with a transmission line whose characteristic admittance is M. The impedances of the two faces of the iris are thus in parallel, as is evident by inspection.

That the iris should act as an admittance in shunt with the guide could have been assumed to begin with, since the voltage of the dominant wave is continuous at the iris and the current is discontinuous. When $\lambda > 2b$, then $\hat{Y}$ and Y_s are pure reactances. But when $\lambda < 2b$, Y_s has in general a real component. While there is no loss of power at the iris, some power is carried beyond the iris and reflected back in other modes than the dominant; this represents an effective loss of power in the dominant wave.

The theory of transverse irises in wave guides in which all four faces are perfect conductors is similar to the above. Thus if the edges of the iris are parallel to the E-vector of the dominant wave, then in the range $a < \lambda < 2a$, the iris acts as an inductive reactance in shunt with the guide. If the E-vector is perpendicular to the edges, the effective reactance is negative and the iris acts as a capacitor. Similarly a transverse wire par-

* This is permissible even though (7) is true only over the aperture since (8) vanishes outside the aperture.

allel to the E-vector acts as a shunt inductance. These phenomena are not peculiar to high frequencies. When a conventional inductance coil or a capacitor is inserted in shunt with a transmission line consisting of parallel wires and operated at low frequencies, the field of the dominant wave is disturbed and the local field in and around the inserted structure abstracts some energy from the dominant wave during one half cycle and returns it during the other half; the dominant wave therefore suffers reflection.

12.4. *The Impedance Seen by a Transverse Wire in a Rectangular Wave Guide*

The impedance seen by a transverse wire in a rectangular wave guide (Fig. 12.13) is of interest because it approximates the impedance seen by a coaxial pair (Fig. 12.14) when the shorter side b is small compared with a quarter wavelength so that the current distribution in the wire is substantially uniform. There are at least three methods available for the calcula-

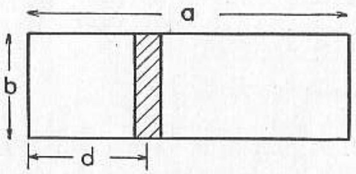


FIG. 12.13.

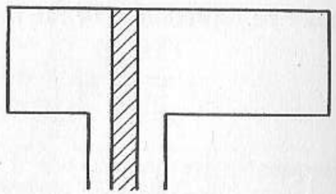


FIG. 12.14.

tion of this impedance. We can express the free space field of the current I in the wire by a contour integral of the type given in problem 10.9 and add to it another field, expressed by a similar integral, so as to satisfy the boundary conditions at the surface of the guide. The contour integrals are then evaluated. Another method consists in considering the total field in the guide as due to superposition of the free space fields of the current in the wire and its images in the walls of the guide. A third method is based on the following considerations. If the radius of the wire is small, then the field at the surface of the wire is nearly equal to that produced by an infinitely thin current filament on the axis of the wire. The latter filament can be regarded as the limit of a current strip; and the current strip represents a known discontinuity in H, which can be expanded in a series of the form $\sum H_n \sin n\pi x/a$, appropriate to TE-waves. Then the complete field is determined and the impedance is obtained as the ratio $-bE_y/I$, where E_y is the intensity on the surface of the wire.

Thus if the guide extends to infinity in both directions, then by the second method we obtain for the real and imaginary parts of the impedance seen

from a wire of radius r respectively

$$R = \tfrac{1}{4}\eta\beta b[J_0(\beta r) + 2\sum_{n=0}^{\infty} J_0(2n\beta a) - \sum_{n=0}^{\infty} J_0(2n\beta a + 2\beta d) - \sum_{n=0}^{\infty} J_0(2n\beta a + \beta a - 2\beta d)];$$

$$X = -\tfrac{1}{4}\eta\beta b[N_0(\beta r) + 2\sum_{n=0}^{\infty} N_0(2n\beta a) - \sum_{n=0}^{\infty} N_0(2n\beta a + 2\beta d) - \sum_{n=0}^{\infty} N_0(2n\beta a + \beta a - 2\beta d)].$$

Thus in the case of thin wires the resistance component is nearly independent of the radius and the reactance is a constant depending on a and d plus a logarithmic function of r. In fact we have

$$X(r_2) - X(r_1) = \frac{\eta b}{\lambda} \log \frac{r_1}{r_2}.$$

A simple expression for R can be obtained either by the first or by the third method. Thus in the frequency range between the absolute cut-off and the next higher it will be found that

$$R = \frac{\eta b}{a}\left(1 - \frac{\lambda^2}{4a^2}\right)^{-1/2} \sin^2 \frac{\pi d}{a}.$$

As the frequency increases and passes the cut-off frequencies for the successive TE-waves other terms will be added and we shall have in general $R = R_1 + R_2 + R_3 + \cdots$, where

$$R_n = \frac{\eta b}{a}\left(1 - \frac{n^2\lambda^2}{4a^2}\right)^{-1/2} \sin^2 \frac{n\pi d}{a}$$

is the resistance corresponding to the nth transmission mode.

The above impedance is that looking into the wave guide extending to infinity in *both* directions. The two semi-infinite halves of the guide are in parallel and consequently the impedance looking into either half is $2Z$. If now $a < \lambda < 2a$, so that only the dominant wave carries power to any distance along the guide, and if we have a conducting piston at distance l from the axis of the wire, then the resistance component of the impedance seen by the wire becomes

$$\frac{\eta b}{a}\left(1 - \cos\frac{4\pi l}{\bar{\lambda}}\right)\left(1 - \frac{\lambda^2}{4a^2}\right)^{-1/2} \sin^2 \frac{\pi d}{a},$$

where $\bar{\lambda}$ is the wavelength along the guide. The additional factor comes in because of reflection of the dominant wave by the piston. On the other hand the reactance represents the local field round the wire and is substantially unaffected by the piston unless it is quite close to the wire.

The methods outlined above can be used equally well for calculating the mutual impedance between two parallel wires in the guide and hence for the solution of any problem involving a system of parallel wires. The above results can be generalized by assuming a sinusoidal distribution of current in the wire.

Problems

1.1. Prove that if ψ is the angle between two directions (θ_1,φ_1) and (θ_2,φ_2), then

$$\cos\psi = \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2\cos(\varphi_1-\varphi_2).$$

1.2. Prove the following relations between vector components in cartesian, cylindrical, and spherical coordinates:

$$\begin{aligned}A_r &= A_x\sin\theta\cos\varphi + A_y\sin\theta\sin\varphi + A_z\cos\theta,\\ A_\theta &= A_x\cos\theta\cos\varphi + A_y\cos\theta\sin\varphi - A_z\sin\theta,\\ A_\varphi &= -A_x\sin\varphi + A_y\cos\varphi;\end{aligned}$$

$$\begin{aligned}A_x &= A_r\sin\theta\cos\varphi + A_\theta\cos\theta\cos\varphi - A_\varphi\sin\varphi,\\ A_y &= A_r\sin\theta\sin\varphi + A_\theta\cos\theta\sin\varphi + A_\varphi\cos\varphi,\\ A_z &= A_r\cos\theta - A_\theta\sin\theta;\end{aligned}$$

$$A_\rho = A_x\cos\varphi + A_y\sin\varphi,\quad A_\varphi = -A_x\sin\varphi + A_y\cos\varphi;$$
$$A_x = A_\rho\cos\varphi - A_\varphi\sin\varphi,\quad A_y = A_\rho\sin\varphi + A_\varphi\cos\varphi;$$

$$A_r = A_\rho\sin\theta + A_z\cos\theta,\quad A_\theta = A_\rho\cos\theta - A_z\sin\theta;$$
$$A_\rho = A_r\sin\theta + A_\theta\cos\theta,\quad A_z = A_r\cos\theta - A_\theta\sin\theta.$$

1.3. Orthogonal curvilinear cylindrical coordinates (u,v,z) may be defined by means of functions of a complex variable

$$u + iv = F(x+iy),\quad x+iy = f(u+iv),\quad z = z.$$

Prove that

$$ds^2 = |f'(u+iv)|^2(du^2+dv^2) + dz^2.$$

1.4. Bicylindrical coordinates (u,ϑ,z) may be defined as follows:

$$u + i\vartheta = \log\frac{a+(x+iy)}{a-(x+iy)},\quad x+iy = a\tanh\frac{u+i\vartheta}{2},\quad z = z.$$

Show that

$$u = \tfrac{1}{2}\log\frac{(a+x)^2+y^2}{(a-x)^2+y^2},\quad \vartheta = \tan^{-1}\frac{y}{a+x} + \tan^{-1}\frac{y}{a-x},$$

$$x = a\frac{\sinh u}{\cosh u+\cos\vartheta},\quad y = a\frac{\sin\vartheta}{\cosh u+\cos\vartheta},$$

$$\rho^2 = x^2+y^2 = a^2\frac{\cosh u-\cos\vartheta}{\cosh u+\cos\vartheta},$$

$$ds^2 = \frac{a^2(du^2+d\vartheta^2)}{(\cosh u+\cos\vartheta)^2} + dz^2.$$

1.5. Elliptic coordinates (u,ϑ,z) may be defined by the following equations:

$$x + iy = l \cosh (u + i\vartheta), \quad z = z,$$

or

$$x = l \cosh u \cos \vartheta, \quad y = l \sinh u \sin \vartheta.$$

Show that

$$ds^2 = \tfrac{1}{2}l^2(\cosh 2u - \cos 2\vartheta)(du^2 + d\vartheta^2) + dz^2,$$
$$\rho^2 = \tfrac{1}{2}l^2(\cosh 2u + \cos 2\vartheta).$$

The u-surfaces are elliptic cylinders

$$\frac{x^2}{l^2 \cosh^2 u} + \frac{y^2}{l^2 \sinh^2 u} = 1$$

and the ϑ-surfaces, confocal hyperbolic cylinders

$$\frac{x^2}{l^2 \cos^2 \vartheta} - \frac{y^2}{l^2 \sin^2 \vartheta} = 1.$$

If ρ_1 and ρ_2 are the distances from the focal lines, then

$$\rho_1 = l(\cosh u - \cos \vartheta), \quad \rho_2 = l(\cosh u + \cos \vartheta),$$
$$\rho_1\rho_2 = \tfrac{1}{2}l^2(\cosh 2u - \cos 2\vartheta).$$

1.6. Prolate spheroidal coordinates (u,ϑ,φ) may be defined as follows:

$$z + i\rho = l \cosh (u + i\vartheta), \quad \varphi = \varphi,$$
$$z = l \cosh u \cos \vartheta, \quad \rho = l \sinh u \sin \vartheta.$$

Show that

$$ds^2 = l^2(\sinh^2 u + \sin^2 \vartheta)(du^2 + d\vartheta^2) + l^2 \sinh^2 u \sin^2 \vartheta \, d\varphi^2.$$

1.7. Oblate spheroidal coordinates (u,ϑ,φ) may be defined as follows:

$$\rho + iz = l \cosh (u + i\vartheta), \quad \varphi = \varphi,$$
$$\rho = l \cosh u \cos \vartheta, \quad z = l \sinh u \sin \vartheta.$$

Show that

$$ds^2 = l^2(\cosh^2 u - \cos^2 \vartheta)(du^2 + d\vartheta^2) + l^2 \cosh^2 u \cos^2 \vartheta \, d\varphi^2.$$

2.1. Prove that $(\cos \varphi \pm i \sin \varphi)^n = \cos n\varphi \pm i \sin n\varphi$.

2.2. Prove that

$$\cos n\varphi = \cos^n \varphi - \frac{n(n-1)}{1 \cdot 2} \cos^{n-2} \varphi \sin^2 \varphi + \cdots,$$

$$\sin n\varphi = n \sin \varphi \cos^{n-1} \varphi - \frac{n(n-1)(n-2)}{1 \cdot 2 \cdot 3} \sin^3 \varphi \cos^{n-3} \varphi + \cdots.$$

2.3. Show that

$$\cos^n \varphi = \frac{1}{2^{n-1}}\left[\cos n\varphi + n \cos (n-2)\varphi + \frac{n(n-1)}{1 \cdot 2} \cos (n-4)\varphi + \cdots\right],$$

where the last term is $\dfrac{n(n-1)\cdots\left(\frac{n+3}{2}\right)}{1 \cdot 2 \cdots \frac{n-1}{2}} \cos \varphi$ if n is odd and

$$\frac{1}{2}\,\frac{n(n-1)\cdots\left(\frac{n}{2}+1\right)}{1 \cdot 2 \cdots \frac{n}{2}}$$ if n is even. Similarly if n is even, then

$$\sin^n \varphi = \frac{(-)^{n/2}}{2^{n-1}}\left[\cos n\varphi - n \cos (n-2)\varphi + \frac{n(n-1)}{1 \cdot 2} \cos (n-4)\varphi - \cdots\right];$$

and if n is odd, then

$$\sin^n \varphi = \frac{(-)^{(n-1)/2}}{2^{n-1}}\left[\sin n\varphi - n \sin (n-2)\varphi + \frac{n(n-1)}{1 \cdot 2} \sin (n-4)\varphi - \cdots\right].$$

2.4. The even part of e^x is called the *hyperbolic cosine* of x and is designated by $\cosh x$; the odd part is the *hyperbolic sine*, $\sinh x$. Obtain formulae for hyperbolic functions analogous to those in problems 2.1, 2.2, 2.3.

2.5. Let (ρ,φ) and (ρ_1,φ_1) be the polar coordinates of the same point with respect to two systems having the same polar axis. Let $(l,0)$ be the origin of the second system with respect to the first. Using $\rho_1 e^{i\varphi_1} = \rho e^{i\varphi} - l$, show that

$$\log \rho_1 = \log \rho - \sum_{n=1}^{\infty} \frac{l^n}{n\rho^n} \cos n\varphi, \quad \rho > l,$$

$$\rho_1^n \cos n\varphi_1 = \sum_{m=0}^{n} (-)^{n-m} \frac{n!}{m!\,(n-m)!} l^{n-m}\rho^m \cos m\varphi.$$

If $\rho < l$, then ρ and l are interchanged in the first equation; the second equation remains unaltered.

2.6. Obtain the following identities

$$\frac{1}{2\pi i}\int_{(C)} \frac{e^{pt}}{p^2}\,dp = 0, \quad t < 0, \qquad \frac{1}{2\pi i}\int_{(C)} \frac{e^{pt}}{p^{n+1}}\,dp = 0, \quad t < 0,$$
$$= t, \quad t > 0; \qquad = \frac{t^n}{n!}, \quad t > 0.$$

2.7. Let $F(t)$ be a function which vanishes for $t < 0$, starts rising linearly at $t = 0$ and reaches unity in τ seconds; subsequently it remains constant. Show that

$$F(t) = \frac{1}{2\pi i\tau}\int_{(C)} \frac{1 - e^{-p\tau}}{p^2} e^{pt}\,dp.$$

Note that this integral is the difference of two integrals, each representing a linear function of t; one function starts from $t = 0$ and the other from $t = \tau$.

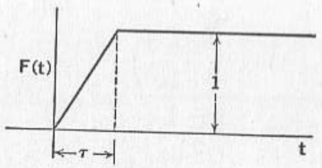


FIG. 1

2.8. Show that for the function defined by Fig. 2

$$F(t) = \frac{1}{2\pi i\tau}\int_{(C)} \frac{1 - e^{-p\tau} - e^{-p(T+\tau)} + e^{-p(T+2\tau)}}{p^2} e^{pt}\,dp.$$

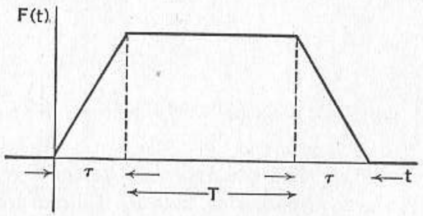


FIG. 2

2.9. Show that the spectrum of a sinusoid of unit amplitude and finite duration

$$\begin{aligned} F(t) &= 0, \quad t < 0; \\ &= \cos(\omega t + \varphi), \quad 0 < t < T; \\ &= 0, \quad t > T; \end{aligned}$$

is

$$S(p) = \frac{1}{4\pi i}\left[\frac{1 - e^{-(p-i\omega)T}}{p - i\omega} e^{i\varphi} + \frac{1 - e^{-(p+i\omega)T}}{p + i\omega} e^{-i\varphi}\right].$$

Note that

$$S(p) = S_1(p,\varphi) - e^{-pT} S_1(p,\varphi + \omega T),$$

where

$$S_1(p,\varphi) = \frac{1}{4\pi i}\left(\frac{e^{i\varphi}}{p - i\omega} + \frac{e^{-i\varphi}}{p + i\omega}\right).$$

The spectrum S_1 is independent of the duration T of the sinusoid; it is the spectrum of the sinusoid starting at $t = 0$ and continuing indefinitely. The second term in $S(p)$ is a similar sinusoid, starting at $t = T$ with just the right phase to cancel the first sinusoid ever after. Thus, for the semi-infinite sinusoid beginning at $t = 0$,

we have

$$F(t) = \frac{1}{4\pi i}\int_{(C)}\left(\frac{e^{i\varphi}}{p - i\omega} + \frac{e^{-i\varphi}}{p + i\omega}\right) e^{pt}\,dp = \frac{1}{2\pi i}\int_{(C)} \frac{p\cos\varphi - \omega\sin\varphi}{p^2 + \omega^2} e^{pt}\,dp.$$

2.10. For a circuit consisting of a resistor and an inductor in series $Z(p) = R + pL$. If a steady electromotive force V is impressed at $t = 0$ and is discontinued at $t = \tau$, then

$$I(t) = \frac{V}{R}(1 - e^{-(R/L)t}), \quad 0 < t < \tau,$$

and

$$I(t) = \frac{V}{R}[e^{-(R/L)(t-\tau)} - e^{-(R/L)t}], \quad t > \tau.$$

If $R\tau/L$ is large, then for $t > \tau$

$$I(t) = \frac{V}{R} e^{-(R/L)(t-\tau)}.$$

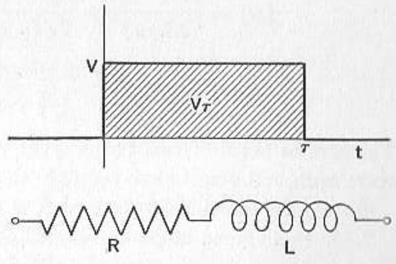


FIG. 3

2.11. For a circuit consisting of a resistor and a capacitor in series $Z(p) = R + 1/pC$. Show that for $0 < t < \tau$,

$$I(t) = \frac{V}{R}\exp\left(-\frac{t}{RC}\right);$$

and for $t > \tau$,

$$I(t) = \frac{V}{R}\exp\left(-\frac{t}{RC}\right) - \frac{V}{R}\exp\left(-\frac{t-\tau}{RC}\right).$$

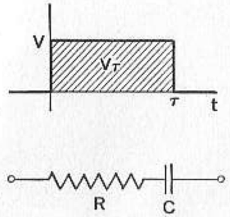


FIG. 4

If τ/RC is large, the first term in the second equation may be ignored.

From these equations obtain the electric charge in the capacitor

$$\begin{aligned} q(t) &= \int_0^t I(t)dt = CV\left[1 - \exp\left(-\frac{t}{RC}\right)\right], \quad 0 < t < \tau, \\ &= CV\left[\exp\left(-\frac{t-\tau}{RC}\right) - \exp\left(-\frac{t}{RC}\right)\right], \quad t > \tau. \end{aligned}$$

The electric charge could be obtained directly by the contour integral method if, instead of the impedance $Z(p)$, we used the impediment $Z(p) = pR + 1/C$.

2.12. For a circuit consisting of a resistor, a capacitor, and an inductor in series $Z(p) = R + pL + 1/pC$. Show that for $0 < t < \tau$

$$I(t) = I_1(t) = \frac{Ve^{p_1 t}}{R + 2p_1 L} + \frac{Ve^{p_2 t}}{R + 2p_2 L},$$

where p_1 and p_2 are the zeros of $Z(p)$, and for $t > \tau$

$$I(t) = I_1(t) - I_1(t - \tau).$$

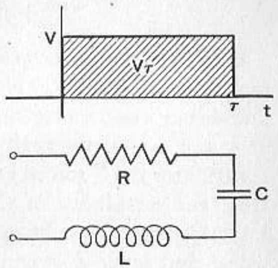


FIG. 5

2.13. An electromotive force

$$V(t) = 0, \quad t < 0;$$
$$= V\cos(\omega t + \varphi), \quad t > 0;$$

is impressed on a circuit consisting of R, L, and C in series. Show that

$$I(t) = \frac{Ve^{i(\omega t+\varphi)}}{2Z(i\omega)} + \frac{Ve^{-i(\omega t+\varphi)}}{2Z(-i\omega)} + \frac{Vp_1(p_1\cos\varphi - \omega\sin\varphi)e^{p_1 t}}{(p_1^2+\omega^2)(R+2p_1L)} + \frac{Vp_2(p_2\cos\varphi - \omega\sin\varphi)e^{p_2 t}}{(p_2^2+\omega^2)(R+2p_2L)}.$$

The sum of the first two terms is the real part of $Ve^{i(\omega t+\varphi)}/Z(i\omega)$; this is the steady state term and could have been obtained directly.

Obtain the solution for the case $R = 0$ and $\omega =$ the natural frequency of the circuit.

2.14. In a transmission line described by (2.10–3) the impressed electromotive force $E(x)$ is a progressive wave $Ee^{-i\beta_1 x}$ over a section $0 < x < l$ and is zero elsewhere. Show that

$$E(x) = \frac{E}{2\pi i}\int_{(C)} \frac{1 - e^{-(\gamma+i\beta_1)l}}{\gamma + i\beta_1} e^{\gamma x}\, d\gamma;$$

$$I(x) = \frac{1 - e^{-(\Gamma+i\beta_1)l}}{2K(\Gamma+i\beta_1)} Ee^{\Gamma x}, \quad x < 0;$$
$$= \frac{1 - e^{(\Gamma-i\beta_1)l}}{2K(-\Gamma+i\beta_1)} Ee^{-\Gamma x}, \quad x > l;$$
$$= \frac{YEe^{-i\beta_1 x}}{\Gamma^2+\beta_1^2} + \frac{Ee^{-\Gamma x}}{2K(-\Gamma+i\beta_1)} - \frac{Ee^{-i\beta_1 l}e^{\Gamma(x-l)}}{2K(\Gamma+i\beta_1)}, \quad 0 < x < l;$$

$$V(x) = -\frac{1 - e^{-(\Gamma+i\beta_1)l}}{2(\Gamma+i\beta_1)} Ee^{\Gamma x}, \quad x < 0;$$
$$= \frac{1 - e^{(\Gamma-i\beta_1)l}}{2(-\Gamma+i\beta_1)} Ee^{-\Gamma x}, \quad x > l;$$
$$= \frac{i\beta_1 Ee^{-i\beta_1 x}}{\Gamma^2+\beta_1^2} + \frac{Ee^{-\Gamma x}}{2(-\Gamma+i\beta_1)} + \frac{Ee^{-i\beta_1 l}e^{\Gamma(x-l)}}{2(\Gamma+i\beta_1)}, \quad 0 < x < l.$$

Prove these equations first by evaluating the contour integrals and second by the method of section 7.9. Note that $\gamma = -i\beta_1$ is not a pole of the complete integrands and hence contributes nothing either when $x < 0$ or when $x > l$; in the region $0 < x < l$, the integrand must be split into two terms, and $\gamma = -i\beta_1$ is a pole and contributes the "forced term." In the second method the interval $(0,l)$ again differs from the remainder of the line in that when x is an interior point the definite integrals must be split at this point because the integrands are different on its two sides; but when x is an exterior point such splitting is unnecessary.

Consider also the special case $\beta_1 = 0$ when the applied electromotive force is uniform, equal to E in the interval $(0,l)$ and to zero outside this interval.

2.15. Consider the function shown in Fig. 6. Show that

$$F(x) = \frac{1}{2\pi i}\int_{(C)} \frac{e^{\gamma x}}{\gamma}(1 - e^{-\gamma a} + e^{-2\gamma a} - e^{-3\gamma a} + \cdots)\, d\gamma$$
$$= \frac{1}{2\pi i}\int_{(C)} \frac{e^{\gamma x}}{\gamma(1+e^{-\gamma a})}\, d\gamma.$$

The poles of the integrand are given by

$$\gamma = 0 \quad \text{and} \quad 1 + e^{-\gamma a} = 0;$$

that is,

$$\gamma_0 = 0, \quad \gamma_m = \frac{(2m+1)\pi i}{a}, \quad m = 0, \pm 1, \pm 2, \cdots.$$

Hence, $F(x) = 0$ if $x < 0$; if $x > 0$, then

$$F(x) = \frac{1}{2} + \frac{1}{a}\sum_{-\infty}^{\infty}\frac{1}{\gamma_m}e^{\gamma_m x} = \frac{1}{2} + \frac{2}{\pi}\sum_{m=0}^{\infty}\frac{1}{2m+1}\sin\frac{(2m+1)\pi x}{a}.$$

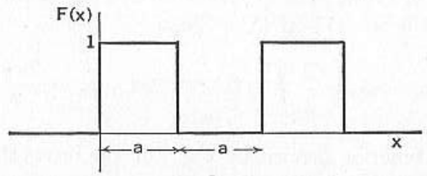


FIG. 6

2.16. Consider a function $f(x)$ which vanishes identically outside the interval $(0,l)$, so that

$$f(x) = \int_{(C)} S(\gamma)e^{\gamma x}\, d\gamma, \quad S(\gamma) = \frac{1}{2\pi i}\int_0^l f(x)e^{-\gamma x}\, dx. \tag{1}$$

Consider a new function $F(x)$ which vanishes identically for $x < 0$ and is periodic of period l for $x > 0$, being equal to $f(x)$ in the interval $(0,l)$. Then,

$$F(x) = \int_{(C)} S(\gamma)(1 + e^{-\gamma l} + e^{-2\gamma l} + \cdots)e^{\gamma x}\, d\gamma$$
$$= \int_{(C)} \frac{S(\gamma)}{1 - e^{-\gamma l}} e^{\gamma x}\, d\gamma. \tag{2}$$

The poles of the integrand are

$$\gamma_n = \frac{2n\pi i}{l}, \quad n = 0, \pm 1, \pm 2, \cdots;$$

hence, for $x > 0$,

$$F(x) = \frac{2\pi i}{l}\sum_{-\infty}^{\infty} S(\gamma_n)e^{\gamma_n x}.$$

If $f(x)$ is a real function, let

$$S(\gamma_n) = \frac{(a_n - ib_n)l}{4\pi i}, \tag{3}$$

so that

$$a_n = \frac{2}{l}\int_0^l f(x)\cos\frac{2n\pi x}{l}\,dx,$$

$$b_n = \frac{2}{l}\int_0^l f(x)\sin\frac{2n\pi x}{l}\,dx. \tag{4}$$

Then

$$F(x) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty}\left(a_n\cos\frac{2n\pi x}{l} + b_n\sin\frac{2n\pi x}{l}\right). \tag{5}$$

Evidently, the "Fourier series" (5) defines a periodic function in the interval $(-\infty,\infty)$ and not only in $(0,\infty)$. Equation (2), on the other hand, defines a function which vanishes for $x < 0$ and is given by (2) for $x > 0$.

The coefficients a_n, b_n of the Fourier series can be obtained either from the conventional formulae (4) or from (1) and (3). Thus

$$a_n - ib_n = \frac{2}{l}\int_0^l f(x)e^{-\gamma_n x}\,dx, \quad \gamma_n = \frac{2n\pi i}{l}.$$

2.17. Consider the function defined by Fig. 7 in the interval $(0,l)$. Prove that in this interval

$$f(x) = \frac{As}{l} + \sum_1^{\infty}\left(a_n\cos\frac{2n\pi x}{l} + b_n\sin\frac{2n\pi x}{l}\right),$$

where

$$a_n = \frac{A}{n\pi}\left[\sin\frac{2n\pi(d+s)}{l} - \sin\frac{2n\pi d}{l}\right] = \frac{2A}{n\pi}\sin\frac{n\pi s}{l}\cos\frac{2n\pi}{l}\left(d+\frac{s}{2}\right),$$

$$b_n = \frac{A}{n\pi}\left[\cos\frac{2n\pi d}{l} - \cos\frac{2n\pi(d+s)}{l}\right] = \frac{2A}{n\pi}\sin\frac{n\pi s}{l}\sin\frac{2n\pi}{l}\left(d+\frac{s}{2}\right).$$

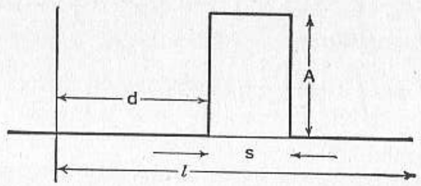


FIG. 7

Prove also that in this interval

$$f(x) = \frac{As}{l} + \frac{4A}{\pi}\sum_1^{\infty}\frac{1}{n}\sin\frac{n\pi s}{2l}\cos\frac{n\pi}{l}\left(d+\frac{s}{2}\right)\cos\frac{n\pi x}{l}.$$

Likewise, prove that in the same interval

$$f(x) = \frac{4A}{\pi}\sum_1^{\infty}\frac{1}{n}\sin\frac{n\pi s}{2l}\sin\frac{n\pi}{l}\left(d+\frac{s}{2}\right)\sin\frac{n\pi x}{l}.$$

4.1. Show that the external inductance L, the capacitance C, and the conductance G per unit length of two coaxial cylinders whose radii are a and b, $b > a$, are

$$L = \frac{\mu}{2\pi}\log\frac{b}{a}, \quad C = \frac{2\pi\epsilon}{\log\frac{b}{a}}, \quad G = \frac{2\pi g}{\log\frac{b}{a}}.$$

4.2. Show that for two parallel wires whose axes are separated by distance l, large compared with the radii, we have approximately

$$L = \frac{\mu}{\pi}\log\frac{l}{\sqrt{ab}}, \quad C = \frac{\pi\epsilon}{\log\frac{l}{\sqrt{ab}}}, \quad G = \frac{\pi g}{\log\frac{l}{\sqrt{ab}}}.$$

4.3. Show that if the radii a and b of two metal spheres are small compared with the distance l between their centers, then the capacitance is approximately

$$C = \frac{4\pi\epsilon}{\frac{1}{a}+\frac{1}{b}-\frac{2}{l}}.$$

4.4. Consider two perfectly conducting concentric spheres and a homogeneous conducting medium between them. Along some radius imagine a filament, insulated from the rest of the medium, and let an impressed electromotive force sustain a steady current I from the outer sphere to the inner. Calculate the field between the spheres, the impressed electromotive force, and the conductance between the spheres. What happens if the medium between the spheres is a perfect dielectric?

4.5. Show that the total force exerted by an electric particle q_1 moving with velocity $\bar{v}_1$ on a particle q_2 moving with velocity $\bar{v}_2$ is

$$F = \frac{q_1q_2\bar{r}_{12}}{4\pi\epsilon r_{12}^3} + \frac{\mu q_1q_2(\bar{r}_{12}\times\bar{v}_1)\times\bar{v}_2}{4\pi r_{12}^3}.$$

Show then that the force between two electric current elements of moments $I_1\bar{l}_1$ and $I_2\bar{l}_2$ is

$$F = \frac{\mu I_1I_2(\bar{r}_{12}\times\bar{l}_1)\times\bar{l}_2}{4\pi r_{12}^3}.$$

(An electric current element is a short filament carrying uniform current.)

5.1. Discuss free and forced oscillations in two inductively coupled simple series circuits. Consider in particular the case of two high Q circuits, tuned to the same frequency and show that in this case there exists a critical coefficient of coupling $\hat{k} = L_{12}/\sqrt{L_{11}L_{22}} = 1/\sqrt{Q_1Q_2}$ such that for $k > \hat{k}$ the current in the secondary circuit passes through two maxima and one minimum while for $k \le \hat{k}$ there is only one maximum. Show that if $k = \bar{k}$, the relative band width (in the same sense as for simple tuned circuits) is $\sqrt{2/Q_1Q_2}$ if Q_1 and Q_2 are of the same order of magnitude, but it is approximately $2/Q_1$ if $Q_1 \gg Q_2 \gg 1$.

5.2. Thévenin's Theorem: At a pair of accessible terminals any linear network containing one or more impedanceless generators acts as a generator whose electromotive force equals the voltage appearing across the terminals when no load impedance is connected and whose internal impedance is the impedance measured between the terminals if all the generators are short-circuited. Prove it.

6.1. Consider the wave produced by an electric current element situated at the origin along the z-axis. Let I be the radial current flowing outward through the hemisphere $\theta \leq \pi/2$ and V be the transverse voltage along a meridian (or along any path in the surface $r =$ const.) from the radius $\theta = 0$ to the radius $\theta = \pi$. Show that V and I satisfy the following transmission equations

$$\frac{dV}{dr} = -\left(\frac{i\omega\mu}{\pi} + \frac{2}{i\omega\epsilon\pi r^2}\right)I, \quad \frac{dI}{dr} = -i\omega\epsilon\pi V.$$

6.2. Consider the wave produced by a small electric current loop situated in the xy-plane at the origin. Let K be the radial magnetic current flowing outward through the hemisphere $\theta \leq \pi/2$ and U the magnetomotive force along a meridian from the radius $\theta = 0$ to the radius $\theta = \pi$. Show that K and U satisfy the following transmission equations

$$\frac{dK}{dr} = -i\omega\mu\pi U, \quad \frac{dU}{dr} = -\left(\frac{i\omega\epsilon}{\pi} + \frac{2}{i\omega\mu\pi r^2}\right)K.$$

6.3. Consider two equally and oppositely charged conductors. The regions subtended by the conductors and bounded by electric lines are called *tubes of flow*; the regions bounded by equipotential surfaces are *equipotential layers*. Show that the tubes of flow are in parallel with each other and that the equipotential layers are in series. Hence show that the capacitance of the two conductors is

$$C = \epsilon \iint \frac{dS}{\int \frac{ds}{F(s)}},$$

where $F(s,u,v)\, dS$ is the area of the normal cross-section of a typical elementary tube; s is the distance along the lines of flow, and u, v are the coordinates of a point on one of the conductors.

The corresponding formula for the conductance is obtained if ϵ is replaced by g.

6.4. Assume a conducting cylinder of radius a, placed in a uniform electric field normal to the axis of the cylinder. Find the charge distribution on the cylinder.

6.5. Assume a conducting cylinder of radius a and a uniformly charged filament parallel to the cylinder at distance l from the axis. Find the field when the charge per unit length of the cylinder is equal and opposite to that on the filament. Consider two cases. (1) the filament is outside the cylinder, (2) the filament is inside the cylinder.

6.6. Assume a circular cylinder of radius a whose permeability is μ placed in a medium with permeability μ_0 normally to a uniform magnetic field. Find the field inside the cylinder and the reflected field outside the cylinder.

6.7. Solve 6.4 if the cylinder is replaced by a sphere.

6.8. Solve 6.6 if the cylinder is replaced by (1) a sphere, (2) a spherical shell.

6.9. Solve problem 6.6 if the cylinder is replaced by a cylindrical shell.

7.1. Show that the admittance seen by the generator in Fig. 7.4 is

$$Y = \frac{1}{D}[K \cosh \Gamma\xi + Z_1 \sinh \Gamma\xi][K \cosh \Gamma(l - \xi) + Z_2 \sinh \Gamma(l - \xi)]$$

and that the impedance seen by the generator in Fig. 7.5 is

$$Z = \frac{K^2}{D}[K \sinh \Gamma\xi + Z_1 \cosh \Gamma\xi][K \sinh \Gamma(l - \xi) + Z_2 \cosh \Gamma(l - \xi)].$$

7.2. Consider a transmission line of length l, terminated at both ends into its characteristic impedance and let the impressed series voltage per unit length be $Ee^{-\gamma x}$, where x is the distance from one end. Find the transverse voltage and longitudinal current

$$V(x) = \tfrac{1}{2}E\left[\frac{2\gamma}{\Gamma^2 - \gamma^2}e^{-\gamma x} - \frac{1}{\Gamma - \gamma}e^{-\Gamma x} + \frac{e^{-\gamma l}}{\Gamma + \gamma}e^{-\Gamma(l-x)}\right],$$

$$I(x) = \frac{1}{2K}E\left[\frac{2\Gamma}{\Gamma^2 - \gamma^2}e^{-\gamma x} - \frac{1}{\Gamma - \gamma}e^{-\Gamma x} - \frac{e^{-\gamma l}}{\Gamma + \gamma}e^{-\Gamma(l-x)}\right].$$

7.3. Let Z_1 be connected in shunt with a line at distance l from the input terminals and let the line be terminated in Z_2 at distance l_2 beyond Z_1. Find the input impedance by two methods: (1) using (7.6–2) or (7.6–6), (2) using (7.11–11).

7.4. An attenuator is a device which, when inserted in a transmission line, absorbs power without introducing reflections. Design a symmetric T-type attenuator and calculate the attenuation ratio.

7.5. Design an attenuator, using series resistors only.

7.6. Discuss resonance in a non-dissipative transmission line shorted at one end and terminated into a capacitor at the other end. Show that if the terminal capacitance C_1 is small compared with the total d–c capacitance Cl of the line, then the longest resonant wavelength is $4(l + l_1)$, where $C_1 = Cl_1$.

7.7. Treat the problem of section 7.8 by another method. Starting with the impedances Z_L and Z_R looking respectively to the left and to the right from the generator, determine (in the case shown in Fig. 7.4) $V(\xi + 0)$ and $V(\xi - 0)$ in terms of $I(\xi)$; then use (7.4–10) to obtain the voltage and current distribution. Treat similarly the case shown in Fig. 7.5.

8.1. Consider n parallel thin wires and let E_m be the electric intensity impressed uniformly on the mth wire. Show that the currents in the wires may be obtained from

$$\sum_k Z_{mk}I_k = E_m, \quad m = 1, 2, \cdots n,$$

where $Z_{m,m}$ is the sum of the internal and external impedances of the mth wire and $Z_{mk} = (1/2\pi)i\omega\mu K_0(\sigma l_{mk})$, where l_{mk} is the interaxial distance between the wires.

Discuss the special case of two equal wires energized in parallel ($E_1 = E_2$) and in push-pull ($E_1 = -E_2 = \frac{1}{2}E$). Show that if the interaxial separation in the latter case is small, then $E = [2Z_i + (i\omega\mu/\pi) \log l/a]I$, where Z_i is the internal impedance per unit length of each wire.

8.2. Prove that in the case of three equal, perfectly conducting, equispaced, coplanar wires, energized in parallel ($E_1 = E_2 = E_3$) the approximate ratio of the current in the middle wire to that in either of the other wires is $1 - \log 2/\log l/a$, where l is the distance between adjacent wires and a is the radius.

8.3. Prove that if the interaxial distances between three equal, parallel wires are small and if $E_2 = E_3 = 0$, then $I_2 + I_3 \simeq -I_1$ and a larger fraction of the total current flows in the wire nearest to the first wire.

8.4. Show that the inductance of a solenoid of radius a, coaxial with a perfectly conducting shield of radius b, is $L = \dfrac{\mu N^2 S}{l}\left(1 - \dfrac{a^2}{b^2}\right)$, where l is the length, S the area of the cross-section of the solenoid, and N is the total number of turns. Show that the circulating current in the shield is opposite to that in the solenoid and that the current ratio is $a^2/(b^2 - a^2)$.

8.5. Obtain the exact expressions for the internal impedance of a conducting cylindrical shell: (1) with an external return, (2) with an internal return. Obtain the transfer impedance.

8.6. Taking into consideration only the principal wave discussed in section 8.14, obtain the expressions for the input impedance and the current in a large circular loop and in a rhombus fed at one of its vertices.

8.7. Consider an infinitely long electric current filament, carrying current I, and a conducting cylinder of radius a whose generators are parallel to the filament. Find the current distribution and the power dissipated in the cylinder on the assumption that the distance l between the filament and the axis of the cylinder is small. In particular consider the high frequency case when the current is near the surface of the cylinder.

8.8. Discuss the dominant transverse magnetic wave in a rectangular wave guide. For this wave magnetic lines form a single set of loops (Fig. 6.20) and the longitudinal electric intensity has only one maximum. Obtain the expressions for the field, the longitudinal current, and the transverse voltage between the axis and the walls of the guide. Show that the cut-off wavelength is $\lambda_c = 2ab/\sqrt{a^2 + b^2}$ and that the attenuation constant $\alpha = 2\mathcal{R}(a^3 + b^3)/[\eta ab(a^2 + b^2)\sqrt{1 - \nu^2}]$, where $\nu = \lambda/\lambda_c$. Obtain the integrated impedances

$$K_{V,I} = \frac{abK_z}{4(a^2 + b^2)}, \quad K_{W,I} = \frac{\pi^2 abK_z}{64(a^2 + b^2)}, \quad K_{W,V} = \frac{4abK_z}{\pi^2(a^2 + b^2)},$$

where the wave impedance at a typical point is $K_z = \eta\sqrt{1 - \nu^2}$.

8.9. In problem 8.8 the longitudinal propagation constant vanishes when $\lambda = \lambda_c$ and then the electric lines become parallel to the guide. Assuming two conducting planes normal to these lines, we obtain a parallelopipedal cavity with free oscillations in it. Show that the energy content is $W = \epsilon abV^2/8c$, where c is the dimension of the cavity parallel to E and V is the maximum voltage amplitude. Show that the

total power loss in the cavity and the Q are

$$W = \frac{\mathcal{R}V^2}{4\eta^2}\left[\frac{ab}{c^2} + \frac{\lambda^2}{2c}\left(\frac{a}{b^2} + \frac{b}{a^2}\right)\right], \quad Q = \frac{\pi\eta c}{\mathcal{R}\lambda[1 + \frac{1}{2}c\lambda^2(a^{-3} + b^{-3})]}.$$

If $b = a$, then $Q = \pi\eta c/\mathcal{R}\sqrt{2}(a + 2c)$.

8.10. Discuss circular magnetic waves (that is waves with circular magnetic lines) in a circular tube. Find the field and show that the cut-off frequencies are $\lambda_{c,m} = 2\pi a/k_m$ where k_m is a typical zero of $J_0(x)$. For the lowest cut-off $\lambda_c = 1.31d$, $d = 0.76\lambda_c$, where $d = 2a$. Show that the attenuation constant is $\alpha = \mathcal{R}/\eta a\sqrt{1 - \nu^2}$ and that

$$K_{W,I} = \frac{K_z}{4\pi}, \quad K_{V,I} = \frac{K_z}{2\pi kJ_1(k)}, \quad K_z = \eta\sqrt{1 - \nu^2}.$$

For the lowest mode $K_{V,I} \simeq 48\sqrt{1 - \nu^2}$ and $K_{W,I} = 30\sqrt{1 - \nu^2}$ if the tube is filled with air.

8.11. Show that the cut-off frequencies of circular electric waves are given by $\lambda_{c,m} = 2\pi a/k_m$, where k_m is a non-vanishing zero of $J_1(x)$. For the lowest mode $\lambda_c = 0.820d$. Obtain the attenuation constant $\alpha = \mathcal{R}\nu^2/\eta a\sqrt{1 - \nu^2}$.

8.12. Discuss circular magnetic waves in perfect dielectric "wires."

9.1. Obtain the radiation intensity and the power radiated by a uniform current loop of any radius a:

$$\Phi = 15\pi(\beta a)^2 I^2 J_1^2(\beta a \sin\theta),$$

$$W = 30\pi^2\beta a I^2 \int_0^{2\beta a} J_2(t)\, dt.$$

9.2. Obtain the radiation intensity and the power radiated by the condenser antenna (by a pair of parallel circular plates, energized from the axis of the condenser so formed):

$$\Phi = \frac{(\beta a)^2}{960\pi} V^2 J_1^2(\beta a \sin\theta),$$

$$W = \frac{\beta a}{480} V^2 \int_0^{2\beta a} J_2(t)\, dt.$$

9.3. Calculate the power radiated from an open end of a coaxial pair. $W = \dfrac{\pi^2}{360}\left(\dfrac{S}{\lambda^2 \log b/a}\right)^2 V^2$, where V is the voltage across the open end.

9.4. Obtain the radiation intensity of a rhombic antenna in free space. Consider the case in which progressive waves are established in the antenna and assume that the amplitude is unaffected by radiation.

10.1. Prove the *orthogonality* of the T-functions corresponding to regions enclosed by perfectly conducting cylindrical surfaces; that is, show that $\iint_{(S)} T_1 T_2\, dS = 0$, where (S) is the cross-section of the cylindrical guide and T_1, T_2 are functions corresponding to the same boundary condition ($T = 0$ or $\partial T/\partial n = 0$) but to different values of χ.

10.2. Consider a uniform plane wave impinging on a conducting wire of small radius a in such a way that H is normal to the wire. Let the angle ϑ between the wire and the direction of the plane wave be arbitrary. Find the scattered wave and the current in the wire.

10.3. Study the normal incidence of a uniform plane wave on a homogeneous cylinder with arbitrary electromagnetic properties and of arbitrary radius a. Consider both cases: (1) E is parallel to the cylinder, (2) H is parallel to the cylinder.

10.4. Consider the problem of reflection of uniform plane waves from a perfectly conducting sphere. Find the reflected field and study some special cases.

10.5. Consider a perfectly conducting sphere and a current element in the direction of some radius. Find the field.

10.6. Solve the preceding problem for a small electric current loop coaxial with some radius.

10.7. Obtain the field of a typical transverse current element in a metal tube of circular cross-section.

10.8. Consider a cylindrical cavity of radius a and height h and inside it a uniform electric current filament parallel to the axis of the cavity. Find the impedance seen by the filament.

10.9. Prove that $K_0(\sigma\rho) = \dfrac{1}{2i}\displaystyle\int_{(C)} \frac{e^{\gamma y \mp \Gamma x}}{\Gamma}\, d\gamma,\quad x \gtrless 0$, where $\Gamma = \sqrt{\sigma^2 - \gamma^2}$.

11.1. Consider two conducting wires of length l normal to a conducting disc large enough to ensure almost complete reflection. Show that at the principal resonance such a structure behaves as a simple (that is, single transmission mode) line of length l, short-circuited at one end and terminated into an admittance $G + i\omega C$, where $G = 60\pi^2 s^2/K^2\lambda^2$ and $C = 2\epsilon a + 60\,(s - a)/K^2 v_0$, s being the distance between the axes of the wires and K the characteristic impedance. Show that the resonant wavelength is approximately $\lambda = 4l + 4v_0CK = 4l + (8/\pi)\, a \log(s/a) + 240\,(s - a)/K$.

12.1. Find an approximate expression for the reactance of the iris shown in Fig. 12.11 to the dominant wave in the frequency range between the absolute cut-off and the next higher when all faces of the wave guide are conductors. Consider the case in which the E-lines are parallel to the edges of the iris and then the case in which they are normal. Show that in the first case the iris possesses an inductive reactance and in the second capacitive.

Questions and Exercises

1. What is the capacitance of a sphere of radius 1 cm. in free space? 1.1 $\mu\mu$f.

2. What is the magnetic intensity inside a conducting wire of radius a carrying a uniform current I? $I\rho/2\pi a^2$.

3. What is the internal magnetic energy per unit length of a wire carrying a uniformly distributed current? $\frac{1}{2}L_i I^2$, where $L_i = \mu/8\pi$.

4. What is the internal inductance of a copper wire? An iron wire whose relative permeability is 100? 0.05 μh, 5 μh per meter.

5. What is the capacitance between parallel plates in air, one millimeter apart, if the radius of each plate is 10 cm? 278 $\mu\mu$f (except for the edge effect).

6. What is the capacitance in Ex. 5 if mica is inserted between the metal plates?

7. What is the capacitance per meter of a coaxial pair, with air between the cylinders, for the diameter ratio e? What is the inductance? 55.6 $\mu\mu$f, 0.2 μh.

8. Estimate the external inductance of a circular loop of radius b, made of wire of radius a. $\mu b \log b/a$.

9. What is the approximate resistance of a 40 watt electric bulb?

10. What is the motional electromotive force developed in a rectangular wire loop rotating in a uniform magnetic field with the frequency ω radians per second? Assume that initially the plane of the loop is normal to the field and that S is the area of the loop. $BS\omega \sin \omega t$.

11. What is the order of magnitude of the Q of coils employed in radio communication? The Q of a coil is defined as the ratio $\omega L/R$ where R is the resistance of the coil.

12. The *power factor* of a capacitor is defined as the reciprocal of its $Q = \omega C/G$, where G is the conductance of the capacitor. What is the order of magnitude of the power factor of capacitors employed in radio communication?

13. What is the resonant frequency and the characteristic impedance of a circuit in which $L = 10$ mh, $C = 100\ \mu\mu$f? $f = (1/2\pi)10^6 = 159$ kilocycles per second, $K = 10{,}000$ ohms.

14. Let the Q of the circuit in Ex. 13 be 200. What is the series impedance (1) at resonance, (2) at twice the resonant frequency, (3) at half the resonant frequency? 50, $50 + i15{,}000$, $50 - i15{,}000$ ohms.

15. In Ex. 14 what is the shunt impedance (the impedance measured across the coil or the capacitor? (1) at resonance, (2) at twice the resonant frequency, (3) at half the resonant frequency? 2×10^6, $22 - i6667$, $22 + i6667$ ohms.

16. In the case of natural oscillations in the above circuit how long would it take for the amplitude to decrease by 1 neper? 400 microseconds.

17. In the case of natural oscillations what is the rate of decay in nepers (1) per second, (2) per radian, (3) per cycle? $\omega/2Q$, $1/2Q$, π/Q.

18. In a series resonant circuit, what is the approximate ratio of the current at resonance to that at twice (or half) the resonant frequency? $1.5Q$.

19. In a parallel resonant circuit, what is the approximate ratio of the voltage across the capacitor at resonance to that at twice the resonant frequency?

20. What is the radiation resistance of a wire one meter long, energized at the center, when $\lambda_0 = 10$ m? 1.97 ohms.

21. What is the electric intensity at distance 100 km from a current element radiating 10 watts in free space? 300 microvolts per meter.

22. In Ex. 21 assume that the current element is at the ground surface (assumed to be a perfect conductor) and normal to it. What is the intensity? 424 microvolts per meter.

23. What is the electric intensity if the current element is replaced by a small current loop?

24. Estimate the capacitance between the outside surfaces of a capacitor formed by two parallel circular discs of radius a, distance h apart (the external capacitance of the discs). Assume $a \gg h$. Roughly $\epsilon a \log a/h$.

25. What is the reflection coefficient in the case of two to one impedance mismatch? $\pm\frac{1}{3}$.

26. What is the amplitude of the reflection coefficient if the line is terminated into $K(1+i)$? 0.45.

27. What is the resonant impedance of a dissipative quarter wave section of a transmission line short-circuited at the far end? $4K/\alpha\lambda$.

28. Express the answer in Ex. 27 in terms of the series resistance R per unit length, assuming that there are no shunt losses. $8K^2/R\lambda$.

29. Express the shunt conductance by an equivalent series resistance. GK^2.

30. Express the quarter wave resonant impedance in Ex. 27 in terms of the Q. $(4/\pi)KQ$.

31. What is the Brewster angle when the ground Q is unity? When Q is small? When Q is large?

$$\cos\vartheta_0 = \sqrt{\frac{\hat{\mu}_r}{\hat{\epsilon}_r\sqrt{2}}},\quad \sqrt{\frac{\hat{\mu}_r Q}{\epsilon_r}},\quad \sqrt{\frac{\hat{\mu}_r}{\hat{\epsilon}_r}}.$$

32. What is the Q of an air-filled cylindrical resonator of radius 20 cm and height 5 cm, assuming copper walls?

33. Show that for an air-filled cylindrical cavity $Q = kh\sqrt{\lambda_0}/(a+h)$, where $k = 13\sqrt{\hat{g}}$.

34. Calculate the longitudinal electric intensity in a coaxial pair having air as the dielectric. $E_z = \frac{60I}{K}\left(Z_a \log\frac{b}{\rho} + Z_b \log\frac{a}{\rho}\right)$ where Z_a and Z_b are the internal impedances of the cylinders.

35. How does the total longitudinal displacement current between coaxial cylinders compare with the conduction current in the inner cylinder?

36. If the diameter of the outer cylinder is fixed, what is the diameter ratio for which the attenuation is minimum? $\log\frac{b}{a} = 1 + \frac{a}{b}$, $\frac{b}{a} = 3.59$.

37. What are the conditions for maximum Q in a coaxial section when the diameter of the outer cylinder is small compared with the length of the section?

38. What is the approximate Q of a single circular turn of wire if the radius of the wire is a and that of the loop b? $Q = (\beta a \log b/a)\eta/\mathcal{R}$.

39. What is the approximate Q of a doublet antenna assumed so short that radiation losses can be neglected? $Q = \frac{3a\lambda}{2\pi l^2}\left(\log\frac{2l}{a} - 1\right)\frac{\eta}{\mathcal{R}}$.

40. What is the expression for the maximum received power W_r in terms of the power W_t radiated by the transmitter? $W_r = (1/16\pi^2)g_1g_2(\lambda/r)^2W_t$. If the directivities are measured with respect to short doublets, then $W_r = 0.0142\bar{g}_1\bar{g}_2(\lambda/r)^2W_t$. In terms of the effective areas of the receiver and transmitter, $W_R/W_T = S_RS_T/\lambda^2r^2$.

BIBLIOGRAPHICAL NOTE

A complete bibliography on the subject of wave theory would make a book of sizable proportions. Those who are interested in following the subject into periodical literature will find it advantageous to start with the following references which have been chosen mainly to call attention to the names of recent contributors in this field.

Barrow, W. L., "Transmission of Electromagnetic Waves in Hollow Tubes of Metal," *Proc. I.R.E.*, October, 1936.

Brillouin, L., "Theoretical Study of Dielectric Cables," *Electrical Communication*, April, 1938.

Carson, J. R., "The Guided and Radiated Energy in Wire Transmission," *A.I.E.E. Jl.*, October, 1924.

Carson, J. R., Mead, S. P., and Schelkunoff, S. A., "Hyper-frequency Wave Guides — Mathematical Theory," *Bell System Tech. Jour.*, April, 1936.

Carter, P. S., "Circuit Relation in Radiating Systems and Applications to Antenna Problems," *Proc. I.R.E.*, June, 1932.

Chu, L. J., and Stratton, J. A., "Forced Oscillations of a Prolate Spheroid," *Jour. App. Phys.*, March, 1941.

Condon, E. U., "Forced Oscillations in Cavity Resonators," *Jour. App. Phys.*, February, 1941.

Foster, R. M., "A Reactance Theorem," *Bell System Tech. Jour.*, April, 1924; "Directive Diagrams of Antenna Arrays," *Bell System Tech. Jour.*, April, 1926.

Fry, T. C., "Plane Waves of Light — III," *Jour. Opt. Soc. Amer. and Rev. Sci. Instr.*, June, 1932.

Gray, M. C., "Horizontally Polarized Electromagnetic Waves over a Spherical Earth," *Phil. Mag.*, April, 1939.

Hahn, W. C., "New Method of the Calculation of Cavity Resonators," *Jour. App. Phys.*, January, 1941.

Hallen, E., "Theoretical Investigations into the Transmitting and Receiving Qualities of Antennas," *Nova Acta (Upsala)*, November, 1938.

Hansen, W. W. and Richtmeyer, R. D., "On Resonators Suitable for Klystron Oscillators," *Jour. App. Phys.*, March, 1939.

King, L. V., "On the Radiation Field of a Perfectly Conducting Base Insulated Cylindrical Antenna over a perfectly Conducting Plane Earth, and the Calculation of Radiation Resistance and Reactance," *Phil. Trans.*, Ser. A., November 2, 1937.

Page, L. and Adams, N. I., J., "The Electrical Oscillations of a Prolate Spheroid," *Phys. Rev.*, May, 1938.

Ramo, S., "Space Charge an. Field Waves in an Electron Beam," *Phys. Rev.*, August 1, 1939.

SOUTHWORTH, G. C., "Certain Factors Affecting the Gain of Directive Antennas," *Proc. I.R.E.*, September, 1930.

SCHELKUNOFF, S. A., "The Electromagnetic Theory of Coaxial Transmission Lines and Cylindrical Shields," *Bell Syst. Tech. Jour.*, October, 1934.

VAN DER POL, B., and BREMMER H., "The Propagation of Radio Waves over a Finitely Conducting Spherical Earth — III," *Phil. Mag.*, June, 1938.

SYMBOLS USED IN TEXT NOT INCLUDED IN TABLE I, PAGE 61

Symbol	Meaning	PAGE
f:	frequency in cycles per second	21
i:	imaginary unit	14
k:	impedance ratio	212
p:	oscillation constant	23
p (with subscript):	transmission coefficient	211
q (with subscript):	reflection coefficient	210
v:	wave velocity	23
A:	magnetic vector potential	128
B:	susceptance = $\mathrm{im}(Y)$	27
C:	Euler's constant	48
F:	electric vector potential	128
G:	conductance = $\mathrm{re}(Y)$	27
Q:	of a medium	83
Q:	of a circuit	115
R:	resistance = $\mathrm{re}(Z)$	27
*W:	power	
X:	reactance = $\mathrm{im}(Z)$	27
α:	attenuation constant = $\mathrm{re}(\Gamma)$	23
β:	phase constant = $\mathrm{im}(\Gamma)$	23
η:	intrinsic impedance	81
λ:	wavelength	23
ν:	frequency ratio	317
ξ:	growth constant = $\mathrm{re}(p,$	22
σ:	intrinsic propagation constant	81
χ (with subscript):	transfer ratio	206
ω:	angular velocity = $\mathrm{im}(p)$	21
Γ:	propagation constant	23
Δ:	Laplacian	12
Φ:	radiation intensity	333
Ψ:	complex power	31
Ψ:	stream function	174
Ω:	solid angle	161
*$\mathfrak{E}$:	work or energy	
$\mathfrak{R}$:	intrinsic resistance	82

* No page references are given for W and $\mathfrak{E}$ since their exact significance varies somewhat from one formula to another. In each case the symbol is clearly defined in the context.

INDEX

If the guard plates are removed, the electric lines near the edges of our parallel strips will bulge out as shown in Fig. 4.1 and the magnetic lines will bend round to enclose one conductor or the other. Subsequent analysis will show that a wave of this modified type may exist at all frequencies and that the shape and distribution of the electric and magnetic lines are independent of the frequency. For such a wave the edge effect is small if b is small compared to a since the energy is distributed largely between the plates.

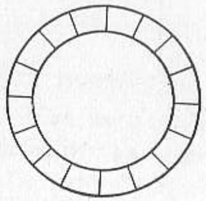


FIG. 8.3. Coaxial cylinders of nearly equal radii.

If b is small compared to a, the strips can be bent into cylinders to form coaxial conductors with nearly equal radii (Fig. 8.3). Electric lines will run along radii and magnetic lines will be coaxial circles between the conductors. There will be a slight "curvature effect" instead of the edge effect. The curvature effect is comparatively small; thus if the radii of the conductors are a and b $(b > a)$, then by the parallel plane formula (using the average circumference for the approximate length of the magnetic lines), we have*

$$K = \frac{\eta(b-a)}{\pi(b+a)} = \frac{120(b-a)}{b+a}.$$

If $b = 2a$, this gives $K = 40$ ohms; the exact value is 41.6.

Since the voltages along various parts of a given radius are added while the magnetomotive force is the same for all magnetic lines, the characteristic impedance of a coaxial pair is the sum of the characteristic impedances of coaxial shells into which the space between the conductors might be subdivided. Thus if $b - a$ is divided into n equal parts, the exact value of K may be expressed in the following form

$$K = \lim 120(b-a)\left[\frac{1}{(2n-1)a+b} + \frac{1}{(2n-3)a+3b} + \cdots \frac{1}{a+(2n-1)b}\right]$$
$$\text{as } n \to \infty.$$

Taking again $b = 2a$ and choosing $n = 2$, we obtain $K = 41.1$; this value differs from the exact value by about 1 per cent.

Let us now return to waves in an unlimited medium. With transverse dimensions fading out of the picture, we fix our attention on the field intensities E and H, rather than on their integrated values, and define the ratio E/H as the *wave impedance* in the direction of wave propagation. A uniform plane wave can be generated by a plane current sheet of uniform

* When a numerical value is ascribed to the intrinsic impedance, free space is usually assumed.

density. Consider such a sheet in the xy-plane and let its density be J_x. Since the electric intensity is continuous at the sheet while the magnetic intensity is discontinuous, we have

$$E_x(+0) = E_x(-0), \quad H_y(+0) - H_y(-0) = -J_x.$$

The current sheet acts as a shunt generator and sends out plane waves in both directions

$$E_x^+(z) = -\tfrac{1}{2}\eta J_x e^{-\sigma z}, \quad H_y^+(z) = -\tfrac{1}{2}J_x e^{-\sigma z}, \quad z > 0,$$
$$E_x^-(z) = -\tfrac{1}{2}\eta J_x e^{\sigma z}, \quad H_y^-(z) = \tfrac{1}{2}J_x e^{\sigma z}, \quad z < 0.$$

The complex power (per unit area) contributed to the field by the impressed forces is

$$\Psi = -\tfrac{1}{2}E_x(0)J_x^* = \tfrac{1}{4}\eta J_x J_x^*.$$

If the medium is nondissipative, then the power carried by each wave per unit area in an equiphase plane is

$$\Psi^+ = \tfrac{1}{2}E_x^+(z)[H_y^+(z)]^* = \tfrac{1}{8}\eta J_x J_x^*, \quad \Psi^- = -\tfrac{1}{2}E_x^-(z)[H_y^-(z)]^* = \tfrac{1}{8}\eta J_x J_x^*.$$

The sum is equal to the power contributed to the field.

The total power carried by a uniform plane wave in an unlimited medium is infinite and the wave cannot possibly be started by an ordinary generator. The principal reason for considering such waves at all is their simplicity, combined with the fact that at great distances from any antenna and in a sufficiently limited region the wave is nearly plane.

If the medium is nondissipative it is possible to send all the energy in one direction only. Consider two parallel equal current sheets (1) and (2), a quarter wavelength apart, and let the currents be in quadrature. If the current in the left-hand sheet (2) is 90 degrees ahead, then the right-hand wave generated by it will be in phase with the right-hand wave generated by the sheet (1); the two waves will reinforce each other. The left-hand wave from (1) will be 180 degrees out-of-phase with the left-hand wave from (2); the two waves will destroy each other to the left of the plane (2). The electric intensity of the wave produced by the sheet (1) will directly oppose the electric intensity of the second sheet and reduce the total intensity at that sheet to zero; hence the second sheet contributes no power and may be taken to be a perfect conductor. The electric intensities of the two waves reinforce each other at the sheet (1). Assuming that this sheet is in the plane $z = 0$, we have therefore

$$E_x^+(z) = -\eta J_x e^{-i\beta z}, \quad H_y^+(z) = -J_x e^{-i\beta z}, \quad z > 0.$$

The power emitted by the sheet is twice that which would be emitted by an isolated sheet.