

Electromagnetic Fields and Waves

Including Electric Circuits

Third Edition

Paul Lorrain

Université de Montréal and McGill University

Dale R. Corson

Cornell University

François Lorrain



W. H. Freeman and Company
New York

Cover image: Lines of *E* and of *H* in an optical waveguide. See page 668.
Back cover image: Differentiating circuit. See page 144.

Library of Congress Cataloging-in-Publication Data

Lorrain, Paul.

Electromagnetic fields and waves.

Includes index.

1. Electromagnetic fields. 2. Electromagnetic waves.

I. Corson, Dale R. II. Lorrain, François

III. Title

QC665.E4L67 1987 530.1'41 86-31803

ISBN 0-716-71823-5

ISBN 0-716-71869-3 (pbk.)

Copyright © 1988 by W. H. Freeman and Company

No part of this book may be reproduced by any mechanical,
photographic, or electronic process, or in the form of a
phonographic recording, nor may it be stored in a retrieval
system, transmitted, or otherwise copied for public or
private use without written permission from the publisher.

Printed in the United States of America

2 3 4 5 6 7 8 9 0 VB 6 5 4 3 2 1 0 8 9

CONTENTS

Preface	viii	
1	VECTOR OPERATORS	1
2	PHASORS	29
3	ELECTRIC FIELDS I	42
	<i>Coulomb's Law and Gauss's Law</i>	
4	ELECTRIC FIELDS II	64
	<i>The Equations of Poisson and Laplace. Charge Conservation. Conductors</i>	
5	ELECTRIC FIELDS III	84
	<i>Electric Multipoles</i>	
6	ELECTRIC FIELDS IV	101
	<i>Energy, Capacitance, and Forces</i>	
*7	ELECTRIC FIELDS V	119
	<i>Electric Circuits A: RC Circuits</i>	
*8	ELECTRIC FIELDS VI	149
	<i>Electric Circuits B: Circuit Theorems</i>	
9	ELECTRIC FIELDS VII	172
	<i>Dielectric Materials A: Bound Charges and the Electric Flux Density \mathbf{D}</i>	
10	ELECTRIC FIELDS VIII	192
	<i>Dielectric Materials B: Real Dielectrics, Continuity Conditions at an Interface, and Stored Energy</i>	
11	ELECTRIC FIELDS IX	211
	<i>Images. Laplace's Equation in Rectangular Coordinates</i>	

* Asterisks mean that the chapters can be skipped without losing continuity.

12	ELECTRIC FIELDS X	225
	<i>Laplace's Equation in Spherical Coordinates. Poisson's Equation for \mathbf{E}</i>	
*13	RELATIVITY I	238
	<i>The Lorentz Transformation and Space-time</i>	
*14	RELATIVITY II	251
	<i>The Lorentz Contraction and Time Dilation. The Transformation of a Velocity</i>	
*15	RELATIVITY III	265
	<i>Mass, Momentum, Force, and Energy</i>	
*16	RELATIVITY IV	284
	<i>The Field of a Moving Electric Charge</i>	
*17	RELATIVITY V	308
	<i>Maxwell's Equations. The Four-Potential A</i>	
18	MAGNETIC FIELDS I	326
	<i>The Magnetic Flux Density \mathbf{B} and the Vector Potential A</i>	
19	MAGNETIC FIELDS II	348
	<i>The Vector Potential A. Ampère's Circuital Law</i>	
20	MAGNETIC FIELDS III	360
	<i>Magnetic Materials A: The Magnetic Flux Density \mathbf{B} and the Magnetic Field Strength \mathbf{H}</i>	
*21	MAGNETIC FIELDS IV	374
	<i>Magnetic Materials B: Ferromagnetism and Magnetic Circuits</i>	
22	MAGNETIC FIELDS V	387
	<i>The Magnetic Forces on Charges and Currents</i>	
23	MAGNETIC FIELDS VI	412
	<i>The Faraday Induction Law</i>	
*24	MAGNETIC FIELDS VII	437
	<i>Electric Circuits C: Mutual and Self-Inductance</i>	
*25	MAGNETIC FIELDS VIII	452
	<i>Electric Circuits D: Inductive Circuits and Transformers</i>	
26	MAGNETIC FIELDS IX	471
	<i>Magnetic Energy and Macroscopic Magnetic Forces</i>	
27	MAXWELL'S EQUATIONS	492

28	PLANE ELECTROMAGNETIC WAVES I	514
	<i>Uniform Plane Waves in Free Space, Nonconductors, and Conductors</i>	
*29	PLANE ELECTROMAGNETIC WAVES II	536
	<i>Waves in Good Conductors and in Plasmas</i>	
30	PLANE ELECTROMAGNETIC WAVES III	554
	<i>Reflection and Refraction A: The Basic Laws and Fresnel's Equations</i>	
31	PLANE ELECTROMAGNETIC WAVES IV	578
	<i>Reflection and Refraction B: Nonuniform Plane Waves and Total Reflection</i>	
32	PLANE ELECTROMAGNETIC WAVES V	594
	<i>Reflection and Refraction C: Reflection and Refraction at the Surface of a Good Conductor</i>	
33	GUIDED WAVES I	610
	<i>General Principles. The Coaxial and Microstrip Lines</i>	
34	GUIDED WAVES II	626
	<i>The Hollow Rectangular Waveguide</i>	
*35	GUIDED WAVES III	646
	<i>The Planar Optical Waveguide A: The Eigenvalue Equation</i>	
*36	GUIDED WAVES IV	662
	<i>The Planar Optical Waveguide B: The Guided Wave</i>	
37	RADIATION I	675
	<i>The Potentials V and A and the Fields E and B</i>	
38	RADIATION II	692
	<i>The Electric Dipole Transmitting Antenna</i>	
39	RADIATION III	712
	<i>The Half-Wave Antenna, Antenna Arrays, and the Magnetic Dipole Antenna</i>	
	Appendix A SI PREFIXES AND THEIR SYMBOLS	730
	Appendix B CONVERSION TABLE	731
	Appendix C WAVES	732
	ANSWERS	739
	INDEX	745

CHAPTER 1

VECTOR OPERATORS

1.1	VECTOR ALGEBRA	2
1.1.1	INVARIANCE	3
1.2	THE GRADIENT ∇f	3
	<i>Example: THE ELEVATION OF A POINT ON THE SURFACE OF THE EARTH</i>	5
1.3	INVARIANCE OF THE OPERATOR ∇	6
1.4	FLUX	7
	<i>Example: FLUID FLOW</i>	7
1.5	THE DIVERGENCE $\nabla \cdot \mathbf{B}$	8
1.6	THE DIVERGENCE THEOREM	9
	<i>Examples</i>	10
1.7	THE LINE INTEGRAL $\int_a^b \mathbf{B} \cdot d\mathbf{l}$. CONSERVATIVE FIELDS	10
1.8	THE CURL $\nabla \times \mathbf{B}$	11
	<i>Example: FLUID STREAM</i>	13
1.9	STOKES'S THEOREM	13
	<i>Example: CONSERVATIVE FIELDS</i>	15
1.10	THE LAPLACIAN OPERATOR ∇^2	15
1.11	ORTHOGONAL CURVILINEAR COORDINATES	15
1.11.1	CYLINDRICAL COORDINATES	17
1.11.2	SPHERICAL COORDINATES	18
1.11.3	THE GRADIENT	20
1.11.4	THE DIVERGENCE	20
1.11.5	THE CURL	22
1.11.6	THE LAPLACIAN	23
1.12	SUMMARY	24
	PROBLEMS	26

This introductory chapter is meant to help those readers who are not yet proficient in the use of vector operators.

We shall frequently refer to the fields of electric charges and currents. For example, we shall consider the force between two electric charges to arise from an interaction between either one of the charges and the field of the other.

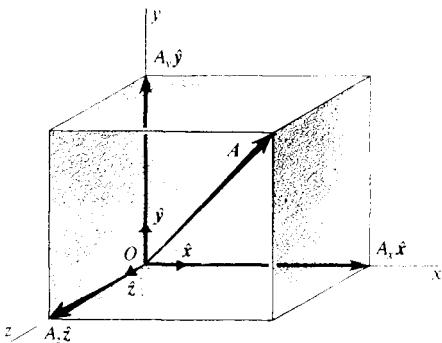


Fig. 1-1. A vector \mathbf{A} and its three component vectors $A_x\hat{x}$, $A_y\hat{y}$, $A_z\hat{z}$ which, when they are placed end to end, are equivalent to \mathbf{A} . The unit vectors \hat{x} , \hat{y} , \hat{z} point in the positive directions of the coordinate axes and are of unit magnitude.

Mathematically, a *field* is a function that describes a physical quantity at all points in space. In *scalar fields* this quantity is specified by a single number for each point. Temperature, density, and electric potential are examples of scalar quantities that can vary from one point to another in space. In *vector fields* the physical quantity is a vector, specified by both a number and a direction. Wind velocity and gravitational force are examples of such vector fields.

Vector quantities will be designated by ***boldface italic type***, and *unit vectors* will carry a circumflex: \hat{x} , \hat{y} , \hat{z} .

Scalar quantities will be designated by *lightface italic type*.

We shall follow the usual custom of using right-hand Cartesian coordinate systems as in Fig. 1-1: the positive z -direction is the direction of advance of a right-hand screw rotated in the sense that turns the positive x -axis into the positive y -axis through the 90° angle.

1.1 VECTOR ALGEBRA

Figure 1-1 shows a vector \mathbf{A} and its three *components* A_x , A_y , A_z . If we define two vectors

$$\mathbf{A} = A_x\hat{x} + A_y\hat{y} + A_z\hat{z}, \quad \mathbf{B} = B_x\hat{x} + B_y\hat{y} + B_z\hat{z}, \quad (1-1)$$

where \hat{x} , \hat{y} , \hat{z} are the *unit vectors* along the x -, y -, and z -axes, respectively, then

$$\mathbf{A} + \mathbf{B} = (A_x + B_x)\hat{x} + (A_y + B_y)\hat{y} + (A_z + B_z)\hat{z}, \quad (1-2)$$

$$\mathbf{A} - \mathbf{B} = (A_x - B_x)\hat{x} + (A_y - B_y)\hat{y} + (A_z - B_z)\hat{z}, \quad (1-3)$$

$$\mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y + A_z B_z = AB \cos \phi, \quad (1-4)$$

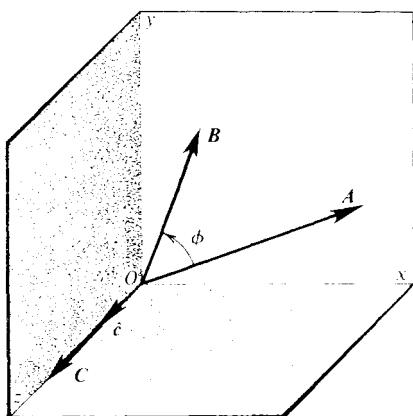


Fig. 1-2. Two vectors \mathbf{A} and \mathbf{B} and the unit vector $\hat{\mathbf{c}}$, normal to the plane containing \mathbf{A} and \mathbf{B} . The positive directions for ϕ and $\hat{\mathbf{c}}$ follow the right-hand screw rule. The vector product $\mathbf{A} \times \mathbf{B}$ is equal to $AB \sin \phi \hat{\mathbf{c}}$, and $\mathbf{B} \times \mathbf{A} = -\mathbf{A} \times \mathbf{B}$.

$$\mathbf{A} \times \mathbf{B} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} = AB \sin \phi \hat{\mathbf{c}} = \mathbf{C}. \quad (1-5)$$

as in Fig. 1-2, where

$$A = (A_x^2 + A_y^2 + A_z^2)^{1/2} \quad (1-6)$$

is the *magnitude* of \mathbf{A} , and similarly for \mathbf{B} .

The quantity $\mathbf{A} \cdot \mathbf{B}$, which is read “ \mathbf{A} dot \mathbf{B} ,” is the *scalar*, or *dot product* of \mathbf{A} and \mathbf{B} , while $\mathbf{A} \times \mathbf{B}$, read “ \mathbf{A} cross \mathbf{B} ,” is their *vector*, or *cross product*.

1.1.1 Invariance

The quantities \mathbf{A} , \mathbf{B} , and ϕ are independent of the choice of coordinate system. Such quantities are said to be *invariant*. A vector, say the gravitational force on a brick, is invariant, but its components are not; they depend on the coordinate system.

Both the dot and cross products are functions of only \mathbf{A} , \mathbf{B} , and ϕ and are thus also invariant.

The sum and the difference, $\mathbf{A} + \mathbf{B}$ and $\mathbf{A} - \mathbf{B}$, are themselves vectors and invariant.

1.2 THE GRADIENT ∇f

A *scalar point-function* is a scalar quantity, say temperature, that is a function of the coordinates. Consider a scalar point-function f that is

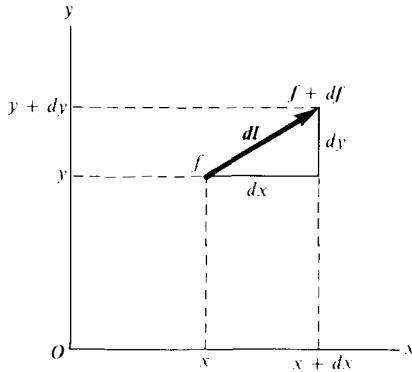


Fig. 1-3. A scalar-point function changes from f to $f + df$ over the distance dl .

continuous and differentiable. We wish to know how f changes over the infinitesimal distance dl in Fig. 1-3. The differential

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz \quad (1-7)$$

is the scalar product of the two vectors

$$dl = dx \hat{x} + dy \hat{y} + dz \hat{z} \quad (1-8)$$

and

$$\nabla f = \frac{\partial f}{\partial x} \hat{x} + \frac{\partial f}{\partial y} \hat{y} + \frac{\partial f}{\partial z} \hat{z}. \quad (1-9)$$

The second vector, whose components are the rates of change of f with distance along the coordinate axes, is called the *gradient* of f . The symbol

$$\nabla = \hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y} + \hat{z} \frac{\partial}{\partial z} \quad (1-10)$$

is read “del.”

Note the value of the magnitude of the gradient:

$$|\nabla f| = \left[\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 + \left(\frac{\partial f}{\partial z} \right)^2 \right]^{1/2}. \quad (1-11)$$

Thus

$$df = \nabla f \cdot dl = |\nabla f| |dl| \cos \theta, \quad (1-12)$$

where θ is the angle between the vectors ∇f and dl .

What direction should one choose for dl to maximize df ? That direction is the one for which $\cos \theta = 1$ or $\theta = 0$, that is, the direction of ∇f .

Therefore the gradient of a scalar function at a given point is a vector having the following properties:

- (1) Its components are the rates of change of the function along the directions of the coordinate axes.
- (2) Its magnitude is the maximum rate of change with distance.
- (3) Its direction is that of the maximum rate of change with distance.
- (4) It points toward larger values of the function.

The gradient is a vector point-function that derives from a scalar point-function.

Again, we have two definitions: ∇f is a vector whose magnitude and direction are those of the maximum space rate of change of f , and it is also the vector of Eq. 1-9. It is clear from the first definition that ∇f is invariant.

Example

THE ELEVATION OF A POINT ON THE SURFACE OF THE EARTH

As an example of the gradient, consider Fig. 1-4 in which E , the elevation above sea level, is a function of the x - and y -coordinates

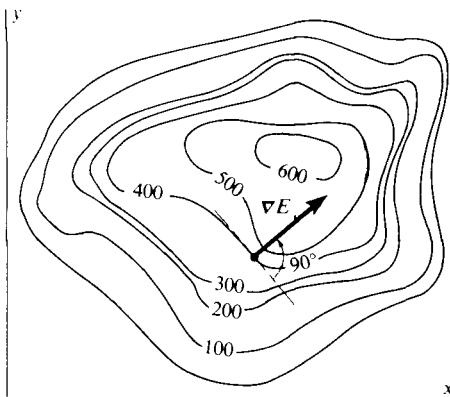


Fig. 1-4. Topographic map of a hill. The numbers shown give the elevation E in meters. The gradient of E is the slope of the hill at the point considered, and it points toward an increase in elevation. The arrow shows ∇E at one point where the elevation is 400 meters.

measured on a horizontal plane. Points at a given elevation define a *contour line*. The gradient of the elevation E at a given point then has the following properties:

- (1) It is perpendicular to the contour line at that point.
- (2) Its magnitude is equal to the maximum rate of change of elevation with displacement measured in a horizontal plane at that point.
- (3) It points toward an increase in elevation.

1.3 INVARIANCE OF THE OPERATOR ∇

We have just seen that ∇f is invariant. Is the operator ∇ itself also invariant? This requires careful consideration because the components of ∇ are not numbers, but operators.

Let S and S' be any two sets of Cartesian coordinates. Figure 1-5 shows two sets having a common origin, for simplicity. Then a given vector \mathbf{A} has the components A_x, A_y, A_z in S , and $A_{x'}, A_{y'}, A_{z'}$ in S' , with

$$A_{x'} = a_{xx}A_x + a_{xy}A_y + a_{xz}A_z, \quad (1-13)$$

$$A_{y'} = a_{yx}A_x + a_{yy}A_y + a_{yz}A_z, \quad (1-14)$$

$$A_{z'} = a_{zx}A_x + a_{zy}A_y + a_{zz}A_z. \quad (1-15)$$

The a coefficients depend only on the orientation of S' with respect to S .

If \mathbf{A} is ∇f , then its components are

$$A_x = \frac{\partial f}{\partial x}, \quad A_y = \frac{\partial f}{\partial y}, \quad A_z = \frac{\partial f}{\partial z}, \quad (1-16)$$

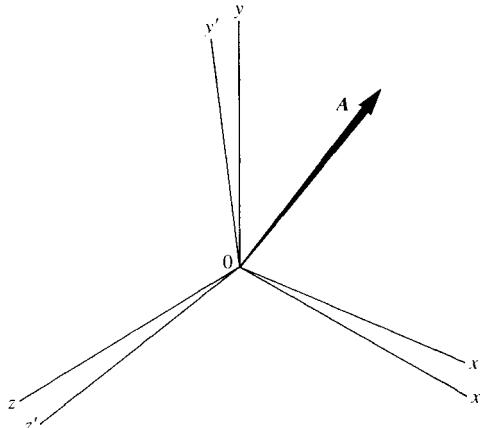


Fig. 1-5. A vector \mathbf{A} and two sets of coordinates S and S' having a common origin.

and

$$\frac{\partial f}{\partial x'} = a_{xx} \frac{\partial f}{\partial x} + a_{xy} \frac{\partial f}{\partial y} + a_{xz} \frac{\partial f}{\partial z}. \quad (1-17)$$

Since this is true for any differentiable f , we know that

$$\frac{\partial}{\partial x'} = a_{xx} \frac{\partial}{\partial x} + a_{xy} \frac{\partial}{\partial y} + a_{xz} \frac{\partial}{\partial z}, \quad (1-18)$$

and similarly for $\partial/\partial y'$ and $\partial/\partial z'$.

The components of ∇ in S' , namely $\partial/\partial x'$, $\partial/\partial y'$, and $\partial/\partial z'$, relate to those of ∇ in S , $\partial/\partial x$, $\partial/\partial y$, and $\partial/\partial z$, in the same way as the components of any vector \mathbf{A} in S' and in S . Therefore ∇ is invariant like any vector, and it transforms as a vector. We shall use this property of ∇ in the following sections.

1.4 FLUX

It is often necessary to calculate the flux of a vector quantity through a surface. By definition, the *flux* $d\Phi$ of \mathbf{B} through an infinitesimal surface $d\mathcal{A}$ is

$$d\Phi = \mathbf{B} \cdot d\mathcal{A}, \quad (1-19)$$

where the vector $d\mathcal{A}$ is normal to the surface. The flux $d\Phi$ is therefore the component of the vector normal to the surface, multiplied by $d\mathcal{A}$. For a surface of finite area \mathcal{A} ,

$$\Phi = \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A}. \quad (1-20)$$

If the surface is closed, the vector $d\mathcal{A}$ points *outward*, by convention.

Example

FLUID FLOW

Consider fluid flow, and let ρ be the density, \mathbf{v} the velocity, and $d\mathcal{A}$ an element of area situated in the fluid. The scalar product $\rho\mathbf{v} \cdot d\mathcal{A}$ is equal to the mass of fluid that crosses $d\mathcal{A}$ in 1 second, in the direction of the vector $d\mathcal{A}$. Then the flux of $\rho\mathbf{v}$ through a closed surface, or the integral of $\rho\mathbf{v} \cdot d\mathcal{A}$ over that surface, is equal to the net rate at which mass leaves the enclosed volume. In an incompressible fluid this flux would be equal to zero.

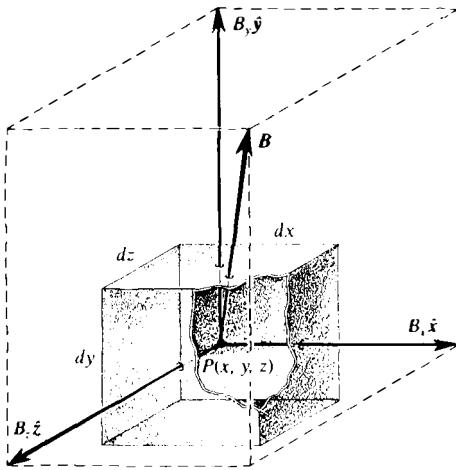


Fig. 1-6. Element of volume $dx dy dz$ and the vector \mathbf{B} at the point P .

1.5 THE DIVERGENCE $\nabla \cdot \mathbf{B}$

The outward flux of a vector through a closed surface can be calculated either from the above equation or as follows. Consider an infinitesimal volume $dx dy dz$ and a vector \mathbf{B} , as in Fig. 1-6, whose components B_x , B_y , B_z are functions of x , y , z . The value of B_x at the center of the right-hand face may be taken to be the average value over that face. Through the right-hand face of the volume element, the outgoing flux is

$$d\Phi_R = \left(B_x + \frac{\partial B_x}{\partial x} \frac{dx}{2} \right) dy dz, \quad (1-21)$$

since the normal component of \mathbf{B} at the right-hand face is the x -component of \mathbf{B} at that face. The volume being infinitesimal, we neglect higher-order derivatives of the components of \mathbf{B} .

At the left-hand face, the outgoing flux is

$$d\Phi_L = - \left(B_x - \frac{\partial B_x}{\partial x} \frac{dx}{2} \right) dy dz. \quad (1-22)$$

There is a minus sign before the parenthesis because $B_x \hat{x}$ points inward at this face and dA outward.

Thus the outward flux through the two faces is

$$d\Phi_L + d\Phi_R = \frac{\partial B_x}{\partial x} dx dy dz = \frac{\partial B_x}{\partial x} dv, \quad (1-23)$$

where dv is the volume of the infinitesimal element.

If we calculate the net flux through the other pairs of faces in the same manner, we find that the total outward flux for the element of volume $d\mathbf{v}$ is

$$d\Phi_{\text{tot}} = \left(\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} \right) d\mathbf{v}. \quad (1-24)$$

Suppose now that we have two adjoining infinitesimal volume elements and that we add the flux emerging through the bounding surface of the first volume to the flux emerging through the bounding surface of the second. At the common face, the fluxes are equal in magnitude but opposite in sign, and they cancel. The sum, then, of the flux from the first volume and that from the second is the flux emerging through the bounding surface of the combined volumes.

To extend this calculation to a finite volume, we sum the individual fluxes for each of the infinitesimal volume elements in the finite volume, and so the total outward flux is

$$\Phi_{\text{tot}} = \int_v \left(\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} \right) d\mathbf{v}. \quad (1-25)$$

At any given point in the volume, the quantity

$$\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z}$$

is thus the *outgoing flux per unit volume* and is invariant. We call this the *divergence of \mathbf{B}* at the point.

The divergence of a vector point-function is a scalar point-function.

According to the rule for the scalar product, we write the *divergence of \mathbf{B}* as

$$\nabla \cdot \mathbf{B} = \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z}. \quad (1-26)$$

The divergence is invariant also because both ∇ and the scalar product are invariant.

1.6 THE DIVERGENCE THEOREM

Now the total outward flux of a vector \mathbf{B} is equal to the surface integral of the normal outward component of \mathbf{B} . Thus, if we denote by \mathcal{A} the area of the surface bounding v , the total outward flux is

$$\Phi_{\text{tot}} = \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} = \int_v \left(\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} \right) dv = \int_v \nabla \cdot \mathbf{B} dv. \quad (1-27)$$

These relations apply to any continuously differentiable[†] vector field \mathbf{B} . Thus

$$\int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} = \int_v \nabla \cdot \mathbf{B} dv. \quad (1-28)$$

This is the *divergence theorem*, also called *Green's theorem*, or *Gauss's theorem*. Note that the first integral involves only the values of \mathbf{B} on the surface of area \mathcal{A} whereas the second involves the values of \mathbf{B} throughout the volume v .

Examples

In an incompressible fluid, let ρ be the mass density and \mathbf{v} the velocity at a given point. Then $\nabla \cdot (\rho \mathbf{v})$ is everywhere equal to zero, since the outward mass flux per unit volume is zero.
Within an explosion, $\nabla \cdot (\rho \mathbf{v})$ is positive.

1.7 THE LINE INTEGRAL $\int_a^b \mathbf{B} \cdot d\mathbf{l}$. CONSERVATIVE FIELDS

The integrals

$$\int_a^b \mathbf{B} \cdot d\mathbf{l}, \quad \int_a^b \mathbf{B} \times d\mathbf{l}, \quad \text{and} \quad \int_a^b f d\mathbf{l},$$

evaluated from the point a to the point b over some specified curve, are examples of *line integrals*.

In the first, which is especially important, the term under the integral sign is the product of an element of length $d\mathbf{l}$ on the curve, multiplied by the local value of \mathbf{B} according to the rule for the scalar product.

A vector field \mathbf{B} is *conservative* if the line integral of $\mathbf{B} \cdot d\mathbf{l}$ around any closed curve is zero:

$$\oint \mathbf{B} \cdot d\mathbf{l} = 0. \quad (1-29)$$

The circle on the integral sign indicates that the path of integration is closed.

[†] A function is *continuously differentiable* if its first derivatives are continuous. .

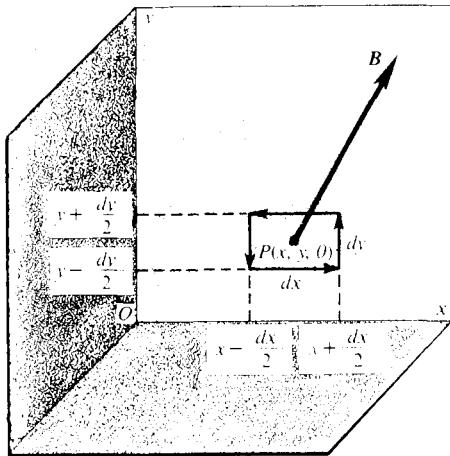


Fig. 1-7. Closed, rectangular path in the xy -plane, centered on the point $P(x, y, 0)$ where the vector \mathbf{B} has the value illustrated by the arrow. The integration around the path proceeds in the direction of the arrows, in accordance with the right-hand screw rule applied to the z -axis.

1.8 THE CURL $\nabla \times \mathbf{B}$

For any given field \mathbf{B} and for a closed path situated in the xy -plane,

$$\mathbf{B} \cdot d\mathbf{l} = B_x dx + B_y dy \quad (1-30)$$

and

$$\oint \mathbf{B} \cdot d\mathbf{l} = \oint B_x dx + \oint B_y dy. \quad (1-31)$$

Now consider the infinitesimal path in Fig. 1-7. There are two contributions to the first integral on the right-hand side of Eq. 1-31, one at $y - dy/2$ and one at $y + dy/2$:

$$\oint B_x dx = \left(B_x - \frac{\partial B_x}{\partial y} \frac{dy}{2} \right) dx - \left(B_x + \frac{\partial B_x}{\partial y} \frac{dy}{2} \right) dx. \quad (1-32)$$

There is a minus sign before the second term because the path element at $y + dy/2$ points in the negative x -direction. Therefore, for this infinitesimal path,

$$\oint B_x dx = - \frac{\partial B_x}{\partial y} dy dx. \quad (1-33)$$

Similarly,

$$\oint B_y dy = \frac{\partial B_y}{\partial x} dx \quad (1-34)$$

and

$$\oint \mathbf{B} \cdot d\mathbf{l} = \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} \right) dx dy \quad (1-35)$$

for the infinitesimal path of Fig. 1-7.

If we set

$$g_3 = \frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}, \quad (1-36)$$

then

$$\oint \mathbf{B} \cdot d\mathbf{l} = g_3 dA, \quad (1-37)$$

where $dA = dx dy$ is the area enclosed by the infinitesimal path. Note that this is correct only if the line integral runs in the positive direction in the xy -plane, that is, in the direction in which one would turn a right-hand screw to make it advance in the positive direction along the z -axis.

Consider now g_3 and the other two symmetric quantities as the components of a vector

$$\nabla \times \mathbf{B} = \left(\frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z} \right) \hat{x} + \left(\frac{\partial B_x}{\partial z} - \frac{\partial B_z}{\partial x} \right) \hat{y} + \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} \right) \hat{z}, \quad (1-38)$$

which may be written as

$$\nabla \times \mathbf{B} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ B_x & B_y & B_z \end{vmatrix} \quad (1-39)$$

This is the *curl* of \mathbf{B} . The quantity g_3 is its z -component.

If we choose a vector dA that points in the direction of advance of a right-hand screw turned in the direction chosen for the line integral, then

$$dA = dA \hat{z} \quad (1-40)$$

and

$$\oint \mathbf{B} \cdot d\mathbf{l} = (\nabla \times \mathbf{B}) \cdot dA. \quad (1-41)$$

This means that the line integral of $\mathbf{B} \cdot d\mathbf{l}$ around the edge of the area dA is equal to the scalar product of the curl of \mathbf{B} by this element of area, with the above sign convention.

We have arrived at this result for an element of area $dx dy$ in the xy -plane. Is this result general? Does it apply to any small area, whatever its orientation with respect to the coordinate axes? It does if it is invariant. We have already seen that the scalar product is invariant. Thus the above line integral is invariant. We have also seen that the operator ∇ and the vector product are invariant. Therefore $\nabla \times \mathbf{B}$ is invariant. This means that $\nabla \times \mathbf{B}$ is a vector whose value, defined by Eq. 1-41, is independent of the particular coordinate axes used, as long as they form a right-handed Cartesian system. Then Eq. 1-41 is indeed invariant; it does apply to any element of area $d\mathcal{A}$, and

$$(\nabla \times \mathbf{B})_n = \lim_{\mathcal{A} \rightarrow 0} \frac{1}{\mathcal{A}} \oint_C \mathbf{B} \cdot d\mathbf{l} \quad (1-42)$$

Thus the component of the curl of a vector normal to a small surface of area \mathcal{A} is equal to the line integral of the vector around the periphery C of the surface, divided by \mathcal{A} , when this area approaches zero.

In general, $\nabla \times \mathbf{B}$ is *not* normal to \mathbf{B} . See Prob. 1-7.

The curl of a gradient is identically equal to zero:

$$\nabla \times (\nabla f) = 0. \quad (1-43)$$

Example

FLUID STREAM

Near the bottom of a fluid stream the velocity \mathbf{v} is proportional to the distance from the bottom. Set the z -axis parallel to the direction of flow and the x -axis perpendicular to the stream bottom. Then

$$v_z = 0, \quad v_y = 0, \quad v_x = cx, \quad (1-44)$$

and the curl of the velocity vector is

$$\nabla \times \mathbf{v} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & 0 & cx \end{vmatrix} = -cy\hat{y}. \quad (1-45)$$

1.9 STOKES'S THEOREM

Equation 1-41 is true only for a path so small that $\nabla \times \mathbf{B}$ is nearly constant over the surface $d\mathcal{A}$ bounded by the path. What happens when the path is so large that this condition is not met? We divide the

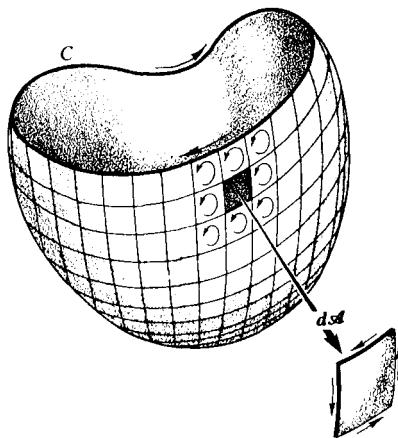


Fig. 1-8. An arbitrary surface of area \mathcal{A} bounded by the curve C . The sum of the line integrals around the curvilinear squares shown is equal to the line integral around C . The positive direction for the vector $d\mathcal{A}$ follows the right-hand screw rule.

surface—any finite surface[†] bounded by the path of integration in question—into elements of area $d\mathcal{A}_1$, $d\mathcal{A}_2$, and so forth, as in Fig. 1-8. For any one of these small areas,

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = (\nabla \times \mathbf{B}) \cdot d\mathcal{A}. \quad (1-46)$$

We add the left-hand sides of these equations for all the $d\mathcal{A}$'s and then all the right-hand sides. The sum of the left-hand sides is the line integral around the external boundary, since there are always two equal and opposite contributions to the sum along every common side between adjacent $d\mathcal{A}$'s. The sum of the right-hand sides is merely the integral of $(\nabla \times \mathbf{B}) \cdot d\mathcal{A}$ over the finite surface. Thus

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \int_{\mathcal{A}} (\nabla \times \mathbf{B}) \cdot d\mathcal{A}, \quad (1-47)$$

where \mathcal{A} is the area of any open surface bounded by the curve C .

This is *Stokes's theorem*. It relates the line integral over a given path to a surface integral over any finite surface bounded by that path. Figure 1-8 illustrates the sign convention.

[†] This must be an *orientable* surface, that is, a surface with two sides. Not all surfaces have two sides; a Möbius strip, for instance, has only one side. See J. E. Marsden and A. J. Tromba, *Vector Calculus*, Freeman, New York, 1976, p. 332.

Example

CONSERVATIVE FIELDS

Under what condition is a vector field conservative? In other words, under what condition is the line integral of $\mathbf{B} \cdot d\mathbf{l}$ around an arbitrary closed path equal to zero? From Stokes's theorem, the line integral of $\mathbf{B} \cdot d\mathbf{l}$ around an arbitrary closed path is zero if $\nabla \times \mathbf{B} = 0$ everywhere. This condition is met if

$$\mathbf{B} = \nabla f \quad (1-48)$$

or if

$$\nabla \times \mathbf{B} = 0. \quad (1-49)$$

A field \mathbf{B} that is the gradient of some scalar point-function f is therefore *conservative*.

1.10 THE LAPLACIAN OPERATOR ∇^2

The divergence of the gradient of f is the *Laplacian* of f :

$$\nabla \cdot \nabla f = \nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}, \quad (1-50)$$

where ∇^2 is the *Laplacian operator*.

The Laplacian is invariant because it is the result of two successive invariant operations.

We have defined the Laplacian of a scalar point-function f . It is also useful to define the Laplacian of a vector point-function \mathbf{B} :

$$\nabla^2 \mathbf{B} = \nabla^2 B_{xx} + \nabla^2 B_{yy} + \nabla^2 B_{zz}. \quad (1-51)$$

The Laplacian of a vector is also invariant. Equation 1-51 applies only to Cartesian coordinates. See Sec. 1.11.6.

1.11 ORTHOGONAL CURVILINEAR COORDINATES

It is frequently inconvenient, because of the symmetries that exist in certain fields, to use Cartesian coordinates. Of all the other possible coordinate systems, we shall restrict our discussion to cylindrical and spherical polar coordinates, the two most commonly used.

We could calculate the gradient, the divergence, and so on, directly in both cylindrical and spherical coordinates. However, it is easier and more general to introduce first the idea of orthogonal curvilinear coordinates.

Consider the equation

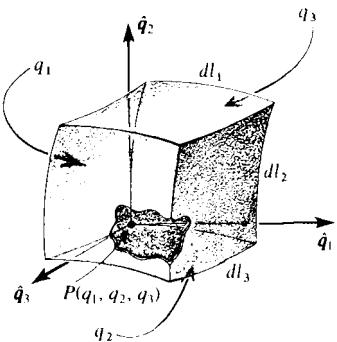


Fig. 1-9. Element of volume in curvilinear coordinates. The unit vectors $\hat{q}_1, \hat{q}_2, \hat{q}_3$ are, respectively, normal to the q_1, q_2, q_3 surfaces at the point P . These unit vectors are mutually perpendicular and oriented in such a way that $\hat{q}_1 \times \hat{q}_2 = \hat{q}_3$.

$$f(x, y, z) = q, \quad (1-52)$$

in which q is a constant. This equation defines a family of surfaces in space, each member characterized by a particular value of the parameter q . An obvious example is $x = q$, which defines surfaces parallel to the yz -plane in Cartesian coordinates.

Consider now three equations

$$f_1(x, y, z) = q_1, \quad f_2(x, y, z) = q_2, \quad f_3(x, y, z) = q_3 \quad (1-53)$$

defining three families of surfaces that are mutually orthogonal. The intersection of three of these surfaces, one of each family, then defines a point in space, and q_1, q_2, q_3 are the *orthogonal curvilinear coordinates* of that point, as in Fig. 1-9.

Call dl_1 an element of length normal to the surface q_1 . This is the distance between the surfaces q_1 and $q_1 + dq_1$ in the infinitesimal region considered. Then

$$dl_1 = h_1 dq_1, \quad (1-54)$$

where h_1 is, in general, a function of the coordinates q_1, q_2, q_3 . Similarly,

$$dl_2 = h_2 dq_2 \quad \text{and} \quad dl_3 = h_3 dq_3. \quad (1-55)$$

With Cartesian coordinates h_1, h_2, h_3 are all unity.

The unit vectors $\hat{q}_1, \hat{q}_2, \hat{q}_3$ are normal, respectively, to the q_1, q_2, q_3 surfaces and are oriented toward increasing values of these coordinates. We assign the subscripts 1, 2, 3 to the coordinates in order that $\hat{q}_1 \times \hat{q}_2 = \hat{q}_3$.

The orientations of the three unit vectors vary, in general, with q_1, q_2, q_3 . Only in Cartesian coordinates do the unit vectors point in fixed directions.

The volume element is

$$dv = dl_1 dl_2 dl_3 = h_1 h_2 h_3 (dq_1 dq_2 dq_3). \quad (1-56)$$

We can now find the q 's, the h 's, the elements of length, and the elements of volume for cylindrical and spherical coordinates.

1.11.1 Cylindrical Coordinates

In cylindrical coordinates, as in Fig. 1-10, $q_1 = \rho$, $q_2 = \phi$, $q_3 = z$.

At P there are three mutually orthogonal directions defined by the three unit vectors $\hat{\rho}$, $\hat{\phi}$, and \hat{z} . The unit vectors $\hat{\rho}$ and $\hat{\phi}$ do not maintain the same directions in space as the point P moves about. However, at any given point, the three unit vectors are mutually orthogonal.

The vector that defines the position of P is

$$\mathbf{r} = \rho \hat{\rho} + z \hat{z}. \quad (1-57)$$

Note that ϕ does not appear explicitly on the right-hand side; it is specified by the orientation of $\hat{\rho}$.

If the coordinates ϕ and z of the point P remain constant while ρ increases by $d\rho$, then P moves by $d\mathbf{r} = d\rho \hat{\rho}$. If ρ and z remain constant while ϕ increases by $d\phi$, then $d\mathbf{r} = \rho d\phi \hat{\phi}$. Finally, if ρ and ϕ are fixed while z increases by dz , then $d\mathbf{r} = dz \hat{z}$. For arbitrary increments $d\rho$, $d\phi$, dz , the distance element is thus

$$d\mathbf{r} = d\rho \hat{\rho} + \rho d\phi \hat{\phi} + dz \hat{z}. \quad (1-58)$$

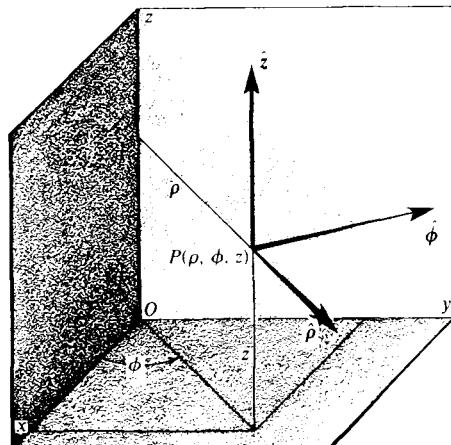


Fig. 1-10. Cylindrical coordinates.

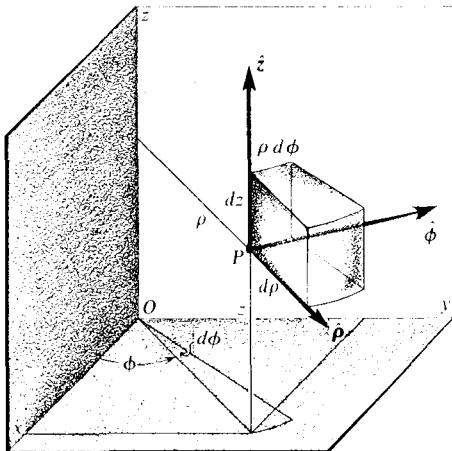


Fig. 1-11. Element of volume in cylindrical coordinates.

Figure 1-11 shows the volume element whose edges are the elements of length corresponding to infinitesimal increments in the coordinates at the point P of Fig. 1-10. The infinitesimal volume is

$$dV = \rho d\rho d\phi dz. \quad (1-59)$$

1.11.2 Spherical Coordinates

In spherical coordinates the position of a point P has the coordinates r, θ, ϕ as in Fig. 1-12. Again, the unit vectors $\hat{r}, \hat{\theta}, \hat{\phi}$ do not maintain the same orientations in space as P moves about.

The vector \mathbf{r} that defines the position of P is now simply $r\hat{r}$, the coordinates θ and ϕ being given by the orientation of \hat{r} . Also,

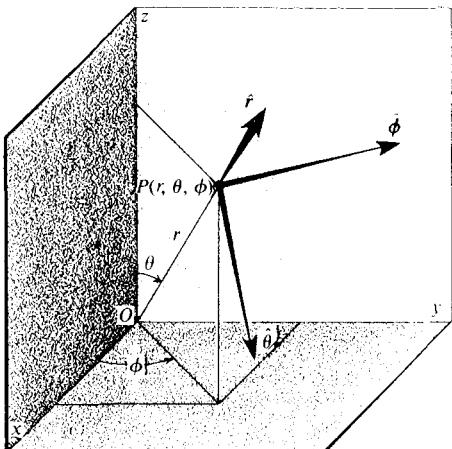


Fig. 1-12. Spherical coordinates.

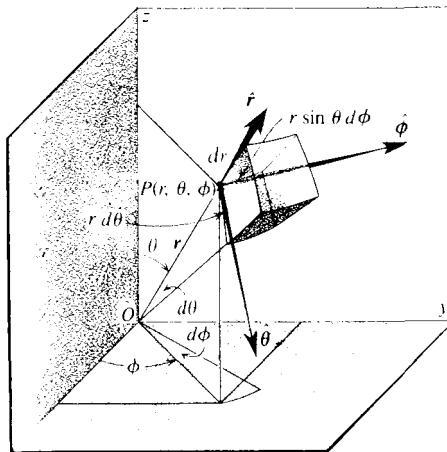


Fig. 1-13. Element of volume in spherical coordinates.

$$d\mathbf{r} = dr \hat{\mathbf{r}} + r d\theta \hat{\mathbf{\theta}} + r \sin \theta d\phi \hat{\mathbf{\phi}}. \quad (1-60)$$

The volume element, shown in Fig. 1-13, is

$$dV = r^2 \sin \theta dr d\theta d\phi. \quad (1-61)$$

Table 1-1 shows the correspondence between curvilinear, Cartesian, cylindrical, and spherical coordinates.

Note that the angle ϕ in both cylindrical and spherical coordinates is undefined for points on the z -axis.

With Cartesian coordinates, one uses the operator ∇ for the gradient of a scalar point-function and for the divergence and curl of a vector point-function. A single expression defines ∇ , and we obtain the gradient, the divergence, or the curl by performing the appropriate

Table 1-1

CURVILINEAR	CARTESIAN	CYLINDRICAL	SPHERICAL
q_1	x	ρ	r
q_2	y	φ	θ
q_3	z	z	φ
h_1	1	1	1
h_2	1	ρ	r
h_3	1	1	$r \sin \theta$
\hat{q}_1	\hat{x}	$\hat{\rho}$	\hat{r}
\hat{q}_2	\hat{y}	$\hat{\varphi}$	$\hat{\theta}$
\hat{q}_3	\hat{z}	\hat{z}	$\hat{\varphi}$

multiplication. This relatively simple situation is peculiar to the Cartesian coordinate system. With other coordinate systems, the divergence, gradient, and curl do not permit a single definition for ∇ but require more elaborate expressions that we shall now derive.

1.11.3 The Gradient

The gradient is the vector rate of change of a scalar function f :

$$\nabla f = \frac{\partial f}{\partial l_1} \hat{q}_1 + \frac{\partial f}{\partial l_2} \hat{q}_2 + \frac{\partial f}{\partial l_3} \hat{q}_3 \quad (1-62)$$

$$= \frac{1}{h_1} \frac{\partial f}{\partial q_1} \hat{q}_1 + \frac{1}{h_2} \frac{\partial f}{\partial q_2} \hat{q}_2 + \frac{1}{h_3} \frac{\partial f}{\partial q_3} \hat{q}_3. \quad (1-63)$$

For cylindrical coordinates, then,

$$\nabla f = \frac{\partial f}{\partial \rho} \hat{p} + \frac{1}{\rho} \frac{\partial f}{\partial \phi} \hat{\phi} + \frac{\partial f}{\partial z} \hat{z}. \quad (1-64)$$

With spherical coordinates,

$$\nabla f = \frac{\partial f}{\partial r} \hat{r} + \frac{1}{r} \frac{\partial f}{\partial \theta} \hat{\theta} + \frac{1}{r \sin \theta} \frac{\partial f}{\partial \phi} \hat{\phi}. \quad (1-65)$$

On the z -axis, ϕ is undefined and both ρ and $\sin \theta$ are zero, so these two expressions are meaningless.

1.11.4 The Divergence

To find the divergence, consider the volume element of Fig. 1-14. The quantity B_1 is the q_1 component of \mathbf{B} at the center, and h_1, h_2, h_3 are the h values at that point. Since the faces are mutually orthogonal, the outward flux through the left-hand face is

$$d\Phi_L = -B_{1L} h_{2L} h_{3L} dq_2 dq_3 \quad (1-66)$$

$$= -\left(B_1 - \frac{\partial B_1}{\partial q_1} \frac{dq_1}{2}\right) \left(h_2 - \frac{\partial h_2}{\partial q_1} \frac{dq_1}{2}\right) \left(h_3 - \frac{\partial h_3}{\partial q_1} \frac{dq_1}{2}\right) dq_2 dq_3. \quad (1-67)$$

Remember that h_2 and h_3 may be functions of q_1 , just as B_1 . We may neglect differentials of order higher than the third, and then

$$d\Phi_L = -B_1 h_2 h_3 dq_2 dq_3 + \frac{\partial}{\partial q_1} (B_1 h_2 h_3) \frac{dq_1}{2} dq_2 dq_3. \quad (1-68)$$

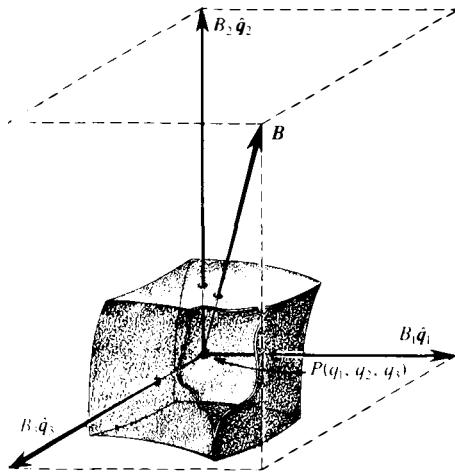


Fig. 1-14. Element of volume, centered on $P(q_1, q_2, q_3)$ where the vector \mathbf{B} has the value shown by the arrow.

By a similar argument,

$$d\Phi_R = B_1 h_2 h_3 \, dq_2 \, dq_3 + \frac{\partial}{\partial q_1} (B_1 h_2 h_3) \frac{dq_1}{2} \, dq_2 \, dq_3 \quad (1-69)$$

for the right-hand face. The net flux through this pair of faces is then

$$d\Phi_{LR} = \frac{\partial}{\partial q_1} (B_1 h_2 h_3) \, dq_1 \, dq_2 \, dq_3. \quad (1-70)$$

If we repeat the calculation for the other pairs of faces to find the net outward flux through the bounding surface and then divide by the volume of the element, we obtain the divergence:

$$\nabla \cdot \mathbf{B} = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1} (B_1 h_2 h_3) + \frac{\partial}{\partial q_2} (B_2 h_3 h_1) + \frac{\partial}{\partial q_3} (B_3 h_1 h_2) \right]. \quad (1-71)$$

In cylindrical coordinates,

$$\nabla \cdot \mathbf{B} = \frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho B_\rho) + \frac{1}{\rho} \frac{\partial B_\phi}{\partial \phi} + \frac{\partial B_z}{\partial z} \quad (1-72)$$

$$= \frac{B_\rho}{\rho} + \frac{\partial B_\rho}{\partial \rho} + \frac{1}{\rho} \frac{\partial B_\phi}{\partial \phi} + \frac{\partial B_z}{\partial z}. \quad (1-73)$$

In spherical coordinates,

$$\nabla \cdot \mathbf{B} = \frac{1}{r^2 \sin \theta} \left[\frac{\partial}{\partial r} (B_r r^2 \sin \theta) + \frac{\partial}{\partial \theta} (B_\theta r \sin \theta) + \frac{\partial}{\partial \phi} (B_\phi r) \right] \quad (1-74)$$

$$= \frac{2}{r} B_r + \frac{\partial B_r}{\partial r} + \frac{B_\theta}{r} \cot \theta + \frac{1}{r} \frac{\partial B_\theta}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial B_\phi}{\partial \phi}. \quad (1-75)$$

These divergences are also meaningless on the z -axis, where ρ and $\sin \theta$ are both zero.

1.11.5 The Curl

We apply the fundamental definition given in Eq. 1-42:

$$(\nabla \times \mathbf{B})_1 = \lim_{\mathcal{A} \rightarrow 0} \frac{1}{\mathcal{A}} \oint \mathbf{B} \cdot d\mathbf{l}, \quad (1-76)$$

where the path of integration C lies in the surface $q_1 = \text{constant}$ and where the direction of integration relates to the direction of the unit vector \hat{q}_1 by the right-hand screw rule. For the paths labeled a, b, c, d in Fig. 1-15, we have the following contributions to the line integral:

$$\begin{aligned} & -B_3 h_3 dq_3, \\ & \left(B_3 + \frac{\partial B_3}{\partial q_2} dq_2 \right) \left(h_3 + \frac{\partial h_3}{\partial q_2} dq_2 \right) dq_3, \\ & +B_2 h_2 dq_2, \\ & - \left(B_2 + \frac{\partial B_2}{\partial q_3} dq_3 \right) \left(h_2 + \frac{\partial h_2}{\partial q_3} dq_3 \right) dq_2. \end{aligned}$$

The sum of these four terms, divided by the element of area is equal to the 1-component of the curl of \mathbf{B} . Neglecting higher-order differentials,

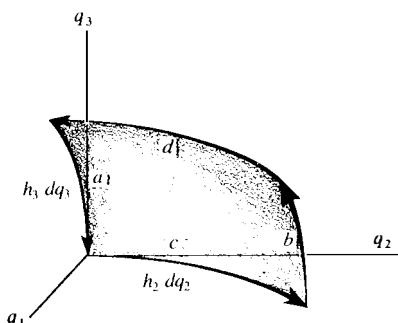


Fig. 1-15. Path of integration for component 1 of the curl in curvilinear coordinates.

$$(\nabla \times \mathbf{B})_1 = \frac{1}{h_2 h_3 dq_2 dq_3} \left[\frac{\partial}{\partial q_2} (B_3 h_3) dq_2 dq_3 - \frac{\partial}{\partial q_3} (B_2 h_2) dq_2 dq_3 \right] \quad (1-77)$$

$$= \frac{1}{h_2 h_3} \left[\frac{\partial}{\partial q_2} (B_3 h_3) - \frac{\partial}{\partial q_3} (B_2 h_2) \right]. \quad (1-78)$$

Corresponding expressions for the other components of the curl follow by rotating the subscripts. Finally,

$$\nabla \times \mathbf{B} = \frac{1}{h_1 h_2 h_3} \begin{vmatrix} h_1 \hat{q}_1 & h_2 \hat{q}_2 & h_3 \hat{q}_3 \\ \frac{\partial}{\partial q_1} & \frac{\partial}{\partial q_2} & \frac{\partial}{\partial q_3} \\ h_1 B_1 & h_2 B_2 & h_3 B_3 \end{vmatrix}. \quad (1-79)$$

For cylindrical coordinates,

$$\nabla \times \mathbf{B} = \frac{1}{\rho} \begin{vmatrix} \hat{\rho} & \rho \hat{\phi} & \hat{z} \\ \frac{\partial}{\partial \rho} & \frac{\partial}{\partial \phi} & \frac{\partial}{\partial z} \\ B_\rho & \rho B_\phi & B_z \end{vmatrix}, \quad (1-80)$$

and for spherical coordinates

$$\nabla \times \mathbf{B} = \frac{1}{r^2 \sin \theta} \begin{vmatrix} \hat{r} & r \hat{\theta} & r \sin \theta \hat{\phi} \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \phi} \\ B_r & r B_\theta & r \sin \theta B_\phi \end{vmatrix}. \quad (1-81)$$

These definitions are not valid on the z -axis.

11.6 The Laplacian

We calculate the Laplacian of a scalar function f in curvilinear coordinates by combining the expressions for the divergence and for the gradient:

$$\nabla^2 f = \nabla \cdot \nabla f \quad (1-82)$$

$$\begin{aligned} &= \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1} \left(\frac{h_2 h_3}{h_1} \frac{\partial f}{\partial q_1} \right) + \frac{\partial}{\partial q_2} \left(\frac{h_3 h_1}{h_2} \frac{\partial f}{\partial q_2} \right) \right. \\ &\quad \left. + \frac{\partial}{\partial q_3} \left(\frac{h_1 h_2}{h_3} \frac{\partial f}{\partial q_3} \right) \right]. \end{aligned} \quad (1-83)$$

For cylindrical coordinates,

$$\nabla^2 f = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial f}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2} \quad (1-84)$$

$$= \frac{1}{\rho} \frac{\partial f}{\partial \rho} + \frac{\partial^2 f}{\partial \rho^2} + \frac{1}{\rho^2} \frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2}, \quad (1-85)$$

except on the z -axis. For spherical coordinates,

$$\nabla^2 f = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2} \quad (1-86)$$

$$= \frac{2}{r} \frac{\partial f}{\partial r} + \frac{\partial^2 f}{\partial r^2} + \frac{\cot \theta}{r^2} \frac{\partial f}{\partial \theta} + \frac{1}{r^2} \frac{\partial^2 f}{\partial \theta^2} + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2}, \quad (1-87)$$

except, again, on the z -axis.

We have already seen in Sec. 1.10 that the Laplacian of a vector \mathbf{B} in *Cartesian* coordinates is itself a vector whose components are the Laplacians of B_x , B_y , B_z . Then

$$\nabla \times (\nabla \times \mathbf{B}) = \nabla(\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B} \quad (1-88)$$

is an identity in Cartesian coordinates.

With other coordinates, $\nabla^2 \mathbf{B}$ is, by definition, the vector whose components are those of $\nabla(\nabla \cdot \mathbf{B}) - \nabla \times (\nabla \times \mathbf{B})$, and *not* the sum of the Laplacians of B_1 , B_2 , B_3 :

$$\nabla^2 \mathbf{B} = \nabla(\nabla \cdot \mathbf{B}) - \nabla \times (\nabla \times \mathbf{B}). \quad (1-89)$$

1.12 SUMMARY

The *gradient* ∇f is a vector whose magnitude and direction are those of the maximum rate of increase of the scalar point-function f with distance at a point.

The *flux* Φ of a vector \mathbf{B} through a surface of area \mathcal{A} is the scalar

$$\Phi = \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A}. \quad (1-20)$$

If the surface is closed, the vector $d\mathcal{A}$ points outward, by convention.

The *divergence* of \mathbf{B}

$$\nabla \cdot \mathbf{B} = \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} \quad (1-26)$$

is the outward flux of \mathbf{B} per unit volume at a point.

The *divergence theorem* states that

$$\int_v \nabla \cdot \mathbf{B} dv = \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A}, \quad (1-28)$$

where \mathcal{A} is the area of the closed surface bounding the volume v .

The *line integral*

$$\int_a^b \mathbf{B} \cdot d\mathbf{l}$$

over a specified curve is the sum of the $\mathbf{B} \cdot d\mathbf{l}$ terms for each element $d\mathbf{l}$ of the curve between points a and b . For a closed curve C that bounds an oriented surface S (see footnote to Sec. 1.9), we have *Stokes's theorem*:

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \int_{\mathcal{A}} (\nabla \times \mathbf{B}) \cdot d\mathcal{A}, \quad (1-47)$$

where

$$\nabla \times \mathbf{B} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ B_x & B_y & B_z \end{vmatrix} \quad (1-39)$$

is the *curl* of the vector point-function \mathbf{B} . The above surface integral applies to any oriented surface of area \mathcal{A} bounded by the curve C .

The *Laplacian* is the divergence of the gradient:

$$\nabla \cdot \nabla f = \nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}. \quad (1-50)$$

The *Laplacian of a vector* in *Cartesian* coordinates is defined as

$$\nabla^2 \mathbf{B} = \nabla^2 B_x \hat{x} + \nabla^2 B_y \hat{y} + \nabla^2 B_z \hat{z}. \quad (1-51)$$

In *cylindrical coordinates* (Figs. 1-10 and 1-11),

$$\mathbf{r} = \rho \hat{\rho} + z \hat{z}, \quad (1-58)$$

$$d\mathbf{r} = d\rho \hat{\rho} + \rho d\phi \hat{\phi} + dz \hat{z}, \quad (1-58)$$

$$dv = \rho d\rho d\phi dz. \quad (1-59)$$

The formulas for the gradient, the divergence, the curl, and the Laplacian are given on the back of the front cover. These vector operators are meaningless on the z -axis, where $\rho = 0$.

In *spherical coordinates* (Figs. 1-12 and 1-13),

$$\mathbf{r} = r\hat{\mathbf{r}},$$

$$d\mathbf{r} = dr \hat{\mathbf{r}} + r d\theta \hat{\theta} + r \sin \theta d\phi \hat{\phi}, \quad (1-60)$$

$$dv = r^2 \sin \theta dr d\theta d\phi. \quad (1-61)$$

The vector operators for spherical coordinates also appear on the back of the front cover. They are meaningless on the z -axis, where $\sin \theta = 0$.

In other than Cartesian coordinates, one defines $\nabla^2 \mathbf{B}$ as follows:

$$\nabla^2 \mathbf{B} = \nabla(\nabla \cdot \mathbf{B}) - \nabla \times (\nabla \times \mathbf{B}). \quad (1-89)$$

We shall have occasion to use several other identities, given on the page facing the front cover.[†]

PROBLEMS

- 1-1. (1.1)[‡] Show that the angle between $\mathbf{A} = 2\hat{x} + 3\hat{y} + \hat{z}$ and $\mathbf{B} = \hat{x} - 6\hat{y} + \hat{z}$ is 130.6° .
- 1-2. (1.1) (a) Show that $(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C}$ is the volume of the parallelepiped whose edges are \mathbf{A} , \mathbf{B} , \mathbf{C} , when the vectors start from the same point.
(b) Show that $(\mathbf{A} \times \mathbf{C}) \cdot \mathbf{B} = -(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C}$. Observe how the sign changes when the cyclic order of the vectors changes.
- 1-3. (1.1) Let C be a plane closed curve. Prove that the area \mathcal{A} enclosed by C is given by

$$\mathcal{A} = \frac{1}{2} \oint_C \mathbf{r} \times d\mathbf{l},$$

where the vector \mathbf{r} goes from an *arbitrary* origin to the element $d\mathbf{l}$ on the curve and where the positive directions for \mathcal{A} and for $d\mathbf{l}$ obey the right-hand screw rule. You can prove this as follows.

[†] See Jean Van Bladel, *Electromagnetic Fields*, McGraw-Hill, New York, 1964, Appendices 1 and 2, for an extensive collection of vector identities and theorems.

[‡] Section number.

- (a) The origin is at O , in the plane, and inside C . Show that the equation is valid.
- (b) The origin is at O' , again in the plane, but outside C . Show that the equation is still valid.
- (c) The origin is at O'' , at some point outside the plane. Show that the integral is again valid.

- 1-4. (1.2) The vector \mathbf{r} points from $P'(x', y', z')$ to $P(x, y, z)$.
- (a) Show that if P is fixed and P' is allowed to move, then $\nabla'(1/r) = \hat{\mathbf{r}}/r^2$, where $\hat{\mathbf{r}}$ is the unit vector along \mathbf{r} .
- (b) Show that, similarly, if P' is fixed and P is allowed to move, then $\nabla(1/r) = -\hat{\mathbf{r}}/r^2$.
- 1-5. (1.6) (a) Show that $\nabla \cdot \mathbf{r} = 3$.
- (b) What is the flux of \mathbf{r} through a spherical surface of radius a ?
- 1-6. (1.6) Show that

$$\int_v \nabla f d\mathbf{v} = \int_s f d\mathcal{A},$$

where \mathcal{A} is the area of the closed surface bounding the volume v . You can prove this by multiplying both sides by \mathbf{c} , where \mathbf{c} is any vector independent of the coordinates. Then use Identity 3 (from inside the front cover) and the divergence theorem.

- 1-7. (1.8) Since $\mathbf{A} \times \mathbf{B}$ is normal to \mathbf{B} , it seems, offhand, that $\nabla \times \mathbf{B}$ must be normal to \mathbf{B} . That is *wrong*.
- As a counterexample, show that $(\nabla \times \mathbf{B}) \cdot \mathbf{B} = -1$ if $\mathbf{B} = y\hat{x} + \hat{z}$.

- 1-8. (1.11.1) (a) Check, by inspection of Fig. 1-10, that the unit vectors in Cartesian and cylindrical coordinates are related as follows:

$$\hat{\rho} = \cos \phi \hat{x} + \sin \phi \hat{y}, \quad \hat{\phi} = -\sin \phi \hat{x} + \cos \phi \hat{y}, \quad \hat{z} = \hat{z}.$$

- (b) Deduce from this set of equations that

$$\hat{x} = \cos \phi \hat{\rho} - \sin \phi \hat{\phi}, \quad \hat{y} = \sin \phi \hat{\rho} + \cos \phi \hat{\phi}, \quad \hat{z} = \hat{z}.$$

You can check this second set by inspection.

- 1-9. (1.11.2) (a) Check, by inspection of Fig. 1-12, that the unit vectors in Cartesian and spherical coordinates are related as follows:

$$\hat{r} = \sin \theta \cos \phi \hat{x} + \sin \theta \sin \phi \hat{y} + \cos \theta \hat{z},$$

$$\hat{\theta} = \cos \theta \cos \phi \hat{x} + \cos \theta \sin \phi \hat{y} - \sin \theta \hat{z}, \quad \hat{\phi} = -\sin \phi \hat{x} + \cos \phi \hat{y}.$$

(b) Show that

$$\begin{aligned}\hat{x} &= \sin \theta \cos \phi \hat{r} + \cos \theta \cos \phi \hat{\theta} - \sin \phi \hat{\phi}, \\ \hat{y} &= \sin \theta \sin \phi \hat{r} + \cos \theta \sin \phi \hat{\theta} + \cos \phi \hat{\phi}, \quad \hat{z} = \cos \theta \hat{r} - \sin \theta \hat{\theta}.\end{aligned}$$

- 1-10. (1.11.2) A vector \mathbf{F} has the same magnitude and direction at all points in space. Choose the z -axis parallel to \mathbf{F} . Then, in Cartesian and in cylindrical coordinates, $\mathbf{F} = F\hat{z}$.

Express \mathbf{F} in spherical coordinates.

- 1-11. (1.11.2) Show, by differentiating the appropriate expressions for \mathbf{r} , that the velocity $\dot{\mathbf{r}}$ in cylindrical coordinates is $\dot{\rho}\hat{\rho} + \rho\dot{\phi}\hat{\phi} + \dot{z}\hat{z}$, while in spherical coordinates it is $\dot{r}\hat{r} + r\dot{\theta}\hat{\theta} + r\sin\theta\dot{\phi}\hat{\phi}$.

- 1-12. (1.11.5) A force \mathbf{F} is of the form $(K/r^3)\hat{r}$ in spherical coordinates, where K is a constant. Is the field conservative?

- 1-13. (1.11.6) Show that, in cylindrical coordinates,

$$(a) \nabla\rho = \hat{\rho}, \quad (b) \nabla \cdot (\rho\hat{\rho}) = 2, \quad (c) \nabla \times (\rho\hat{\rho}) = 0,$$

$$(d) \nabla \times (z\hat{\rho}) = \rho\hat{\phi}, \quad (e) \nabla^2\rho = \frac{1}{\rho}.$$

- 1-14. In the coordinate systems that we have used until now, vectors and the operator ∇ all have three components. However, in relativity theory (Chaps. 13 to 17), it is often more convenient to consider only *two* components, one that is parallel to a given direction and one that is perpendicular. For example, one writes that $\mathbf{r} = \mathbf{r}_{||} + \mathbf{r}_{\perp}$.

If the chosen direction is the x -axis, then

$$\mathbf{r}_{||} = x\hat{x} \quad \text{and} \quad \mathbf{r}_{\perp} = y\hat{y} + z\hat{z}.$$

Also, $\nabla = \nabla_{||} + \nabla_{\perp}$, with

$$\nabla_{||} = \hat{x} \frac{\partial}{\partial x}, \quad \nabla_{\perp} = \hat{y} \frac{\partial}{\partial y} + \hat{z} \frac{\partial}{\partial z}.$$

Then

$$\nabla V = \nabla_{||} V + \nabla_{\perp} V.$$

Show that

$$\nabla \cdot \mathbf{A} = \nabla_{||} \cdot \mathbf{A}_{||} + \nabla_{\perp} \cdot \mathbf{A}_{\perp}, \quad \nabla \times \mathbf{A} = \nabla_{||} \times \mathbf{A}_{\perp} + \nabla_{\perp} \times \mathbf{A}_{||} + \nabla_{\perp} \times \mathbf{A}_{\perp}.$$

CHAPTER 2

PHASORS

1	COMPLEX NUMBERS	30
1.1	ADDITION AND SUBTRACTION OF COMPLEX NUMBERS	31
1.2	MULTIPLICATION AND DIVISION OF COMPLEX NUMBERS	31
	<i>Examples</i>	32
2	PHASORS	32
3	USING PHASORS	34
	<i>Example: SOLVING A SECOND-ORDER LINEAR DIFFERENTIAL EQUATION WITH PHASORS</i>	35
4	PRODUCTS OF PHASORS	36
5	QUOTIENTS OF PHASORS	37
6	ROTATING VECTORS	37
7	NOTATION	38
8	SUMMARY	39
	PROBLEMS	40

This short chapter discusses a second mathematical prerequisite for the study of electromagnetic fields, namely, phasors.

One uses phasors to represent quantities that are sine or cosine functions of the time, or of space coordinates, or of both.

The functions $\sin \omega t$ and $\cos \omega t$ play a major role in modern technology, mostly because of the relative ease with which they can be generated. They are also relatively easy to manipulate mathematically. All other periodic functions, square waves, for example, are much more difficult to generate and much more difficult to manipulate mathematically.

One often has to solve linear differential equations involving sine and cosine functions with constant coefficients. As we shall see, the use of phasors then has the immense advantage of transforming these differential equations into simple algebraic equations.

But first we must review complex numbers.

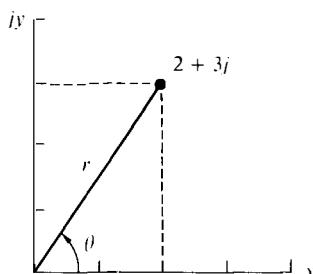


Fig. 2-1. The complex number $2 + 3j$ plotted in the complex plane.

2.1 COMPLEX NUMBERS

A *complex number* is of the form

$$z = a + jb, \quad (2-1)$$

where $j = (-1)^{1/2}$ and where a and b are real numbers, such as 2.5, 3, or -10. Complex numbers can be plotted in the *complex plane*, as in Fig. 2-1. The quantity a is said to be the *real part*, and bj the *imaginary part*, of the complex number.

One can express complex numbers in *Cartesian form*, as above, or in *polar form*, as follows. First,

$$z = a + jb = r \cos \theta + jr \sin \theta = r(\cos \theta + j \sin \theta), \quad (2-2)$$

where

$$r = (a^2 + b^2)^{1/2} \quad (2-3)$$

is the *modulus* of the complex number z , and the angle θ is its *argument*.

With the angle θ expressed in *radians*,

$$\cos \theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \dots, \quad (2-4)$$

$$\sin \theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \dots, \quad (2-5)$$

$$\exp j\theta = 1 + j\theta - \frac{j\theta^2}{2!} - \frac{j\theta^3}{3!} + \frac{\theta^4}{4!} + \frac{j\theta^5}{5!} - \dots, \quad (2-6)$$

and

$$\cos \theta + j \sin \theta = \exp j\theta. \quad (2-7)$$

Of course, $j \times j = -1$, $j \times j \times j = -j$, etc.

Thus, from Eq. 2-2,

$$z = a + jb = r \exp j\theta. \quad (2-8)$$

We have here a complex number expressed both in Cartesian form and in polar form.

If a is positive, then θ is in either the first or the fourth quadrant. If a is negative, θ is in either the second or third quadrant. Use the proper angle! For example, the argument of $-1 + j$ is $3\pi/4$, not $-\pi/4$.

Note that

$$\exp j \frac{\pi}{2} = j, \quad \exp j\pi = -1, \quad \exp \left(-j \frac{3\pi}{2}\right) = -j, \quad \exp j2\pi = 1. \quad (2-9)$$

II

$$z = a + bj = r \exp j\theta, \quad (2-10)$$

then the *complex conjugate* of z is

$$z^* = a - bj = r \exp (-j\theta). \quad (2-11)$$

To obtain the complex conjugate of a complex expression, one changes the sign before j everywhere. For example, if

$$z = \frac{a + bj}{c + dj}, \quad \text{then} \quad z^* = \frac{a - bj}{c - dj}. \quad (2-12)$$

2.1.1 Addition and Subtraction of Complex Numbers

With complex numbers in Cartesian form, one simply adds or subtracts the real and imaginary parts separately:

$$(a + bj) + (c + dj) = (a + c) + (b + d)j. \quad (2-13)$$

If the numbers are in polar form, one first transforms them into Cartesian form.

2.1.2 Multiplication and Division of Complex Numbers

In Cartesian form, one proceeds as follows:

$$(a + bj)(c + dj) = (ac - bd) + (ad + bc)j, \quad (2-14)$$

$$\frac{a + bj}{c + dj} = \frac{(a + bj)(c - dj)}{(c + dj)(c - dj)} = \frac{(ac + bd) + (-ad + bc)j}{c^2 + d^2} \quad (2-15)$$

In polar form,

$$(r_1 \exp j\theta_1)(r_2 \exp j\theta_2) = r_1 r_2 \exp j(\theta_1 + \theta_2), \quad (2-16)$$

$$\frac{r_1 \exp j\theta_1}{r_2 \exp j\theta_2} = \frac{r_1}{r_2} \exp j(\theta_1 - \theta_2). \quad (2-17)$$

Remember to express the angles in *radians*.

Examples

$$(4 + 5j) + (2 - 3j) = 6 + 2j, \quad (2-18)$$

$$(4 + 5j)^2 = (16 - 25) + 40j = -9 + 40j, \quad (2-19)$$

$$\frac{1}{4 + 5j} = \frac{4 - 5j}{(4 + 5j)(4 - 5j)} = \frac{4 - 5j}{16 + 25} = \frac{4 - 5j}{41} = 0.098 - 0.122j, \quad (2-20)$$

$$\left(5 \exp j \frac{\pi}{3} \right) \left(2 \exp j \frac{\pi}{2} \right) = 10 \exp j \frac{5\pi}{6}, \quad (2-21)$$

$$\frac{5 \exp (j\pi/3)}{2 \exp (j\pi/2)} = 2.5 \exp \left(-j \frac{\pi}{6} \right). \quad (2-22)$$

2.2 PHASORS

Electric currents and voltages, electric fields, and magnetic fields are often sinusoidal functions of the time. For example, an *alternating current* is of the form

$$I = I_m \cos (\omega t + \alpha), \quad (2-23)$$

where I_m is the *maximum value* of the current, $\omega = 2\pi f$ is the *circular, or angular, frequency*, and f is the *frequency*. The quantity in parentheses is the *phase, or phase angle*, α being the phase at $t = 0$.

The point I of Fig. 2-2 rotates on a circle of radius I_m in the complex plane at an angular velocity ω . Then its projection on the real axis is

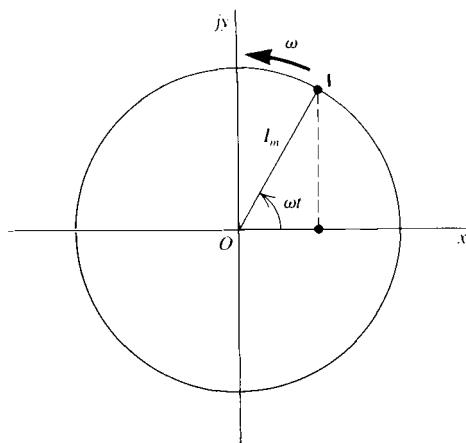


Fig. 2-2. Point I in the complex plane describes a circle of radius I_m about the origin O . It represents the phasor $I_m \exp j\omega t$. Its projection on the x -axis is $I_m \cos \omega t$.

$$I = I_m \cos \omega t. \quad (2-24)$$

We have set $\alpha = 0$ for simplicity. Therefore

$$I = \operatorname{Re}(I_m \exp j\omega t) = \operatorname{Re} I, \quad (2-25)$$

where the operator Re means “Real part of what follows.” The quantity in parentheses is the phasor I of Fig. 2-2:

$$I = I_m \exp j\omega t = I_m \cos \omega t + jI_m \sin \omega t. \quad (2-26)$$

So the phasor I is equal to the variable I plus the parasitic imaginary term $jI_m \sin \omega t$.

Then

$$\frac{dI}{dt} = j\omega I, \quad \frac{d^2I}{dt^2} = (j\omega)^2 I = -\omega^2 I, \quad \frac{d^nI}{dt^n} = (j\omega)^n I, \quad \text{etc.} \quad (2-27)$$

You can easily check that

$$\operatorname{Re}(j\omega I) = \frac{dI}{dt}, \quad \operatorname{Re}(-\omega^2 I) = \frac{d^2I}{dt^2}, \quad \text{etc.} \quad (2-28)$$

In other words,

$$\operatorname{Re} I = I, \quad (2-29)$$

$$\operatorname{Re} \frac{dI}{dt} = \operatorname{Re} j\omega I = \frac{dI}{dt}, \quad (2-30)$$

$$\operatorname{Re} \frac{d^2I}{dt^2} = \operatorname{Re} (j\omega)^2 I = \frac{d^2I}{dt^2}, \quad \text{etc.} \quad (2-31)$$

Therefore, if one replaces the variable I by the phasor I , then the operator d/dt becomes a factor $j\omega$, and a differential equation involving time derivatives becomes an algebraic equation!

A phasor can be multiplied by a complex number:

$$(a + bj)I = (a + bj)I_m \exp j\omega t = rI_m \exp j(\omega t + \theta), \quad (2-32)$$

where r is the modulus of $a + bj$ and θ is its argument.

One can also divide a phasor by a complex number:

$$\frac{I}{a + bj} = \frac{I_m}{r} \exp j(\omega t - \theta). \quad (2-33)$$

Instead of having a cosine function of the time t , one might have a cosine of a coordinate:

$$E = E_m \cos kx. \quad (2-34)$$

The corresponding phasor would then be

$$\mathbf{E} = E_m \exp jkx. \quad (2-35)$$

In a wave, one has a cosine function of both t and, say, z :

$$E = E_m \cos (\omega t - kz), \quad (2-36)$$

where k is the *wave number*. This wave travels in the positive direction of the z -axis. See Appendix C. In phasor form,

$$\mathbf{E} = E_m \exp j(\omega t - kz) \quad (2-37)$$

and

$$\frac{d\mathbf{E}}{dt} = j\omega \mathbf{E}, \quad \frac{d\mathbf{E}}{dz} = -jk \mathbf{E}. \quad (2-38)$$

Vector quantities can also be expressed in phasor form. For example, a force could be a cosine function of time:

$$\mathbf{F} = (F_m \cos \omega t) \hat{\mathbf{x}}. \quad (2-39)$$

Then

$$\mathbf{F} = (F_m \exp j\omega t) \hat{\mathbf{x}} \quad (2-40)$$

is both a vector and a phasor.

2.3 USING PHASORS

To use phasors, one first expresses the sine or cosine functions in the form $x_m \cos (\omega t + \theta)$, and then one uses the phasor

$$\mathbf{x} = x_m \exp j(\omega t + \theta) \equiv x_m \cos (\omega t + \theta) + jx_m \sin (\omega t + \theta). \quad (2-41)$$

One then performs the calculation with the phasors. The result almost invariably stays in phasor form. However, if one requires a real function, one simply rejects the imaginary part.

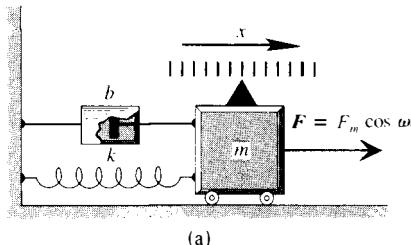


Fig. 2-3. Mass m subjected to a force $F = F_m \cos \omega t$, to a restoring force $-kx$ exerted by a spring, and to a damping force $-b dx/dt$ exerted by a dashpot. At rest, $x = 0$.

Example

SOLVING A SECOND-ORDER LINEAR DIFFERENTIAL EQUATION WITH PHASORS

One of the most common types of differential equation is the following:

$$m \frac{d^2x}{dt^2} + b \frac{dx}{dt} + kx = F_m \cos \omega t. \quad (2-42)$$

Here all the terms are real. This equation describes, for example, the motion of a mass m subjected to the applied force $F_m \cos \omega t$, to a restoring force $-kx$, and to a damping force $-bv$, as in Fig. 2-3: the product of the mass m by the acceleration d^2x/dt^2 is equal to the sum of the applied forces.

The steady-state solution is of the form $x_m \cos(\omega t + \theta)$, where x_m is the amplitude of the motion. It is a simple matter to solve this equation with phasors. We use the phasor

$$\mathbf{x} = x_m \exp j(\omega t + \theta), \quad (2-43)$$

whose real part is the displacement x . We also set

$$\mathbf{F} = F_m \exp j\omega t. \quad (2-44)$$

Substitution into the differential equation is trivial:

$$-m\omega^2 \mathbf{x} + b j\omega \mathbf{x} + k \mathbf{x} = \mathbf{F} \quad (2-45)$$

and

$$\mathbf{x} = \frac{\mathbf{F}}{k - m\omega^2 + jb\omega}. \quad (2-46)$$

Thus, expressing the denominator in polar form,

$$x_m = \frac{F_m}{[(k - m\omega^2)^2 + b^2\omega^2]^{1/2}}, \quad \theta = -\arctan \frac{b\omega}{k - m\omega^2}. \quad (2-47)$$

The actual displacement is the real part of the phasor \mathbf{x} , or

$$x = x_m \cos(\omega t + \theta). \quad (2-48)$$

2.4 PRODUCTS OF PHASORS

One often requires the *average* value of the product of two sinusoidal quantities. Now if one tries to multiply phasors, one runs into trouble. Consider a simple example. Suppose an alternating voltage $V = V_m \cos \omega t$ is applied across a resistance R . Then $I = (V_m \cos \omega t)/R$. The instantaneous power dissipated in the resistor is

$$P_{\text{inst}} = IV = \frac{V_m^2 \cos^2 \omega t}{R}, \quad (2-49)$$

and the average power is

$$P_{\text{av}} = \frac{V_m^2}{2R} = \frac{V_{\text{rms}}^2}{R}, \quad (2-50)$$

the average value of $\cos^2 \omega t$ being $\frac{1}{2}$. Here V_{rms} is the *root mean square* voltage, or the square root of the mean value of the square of V :

$$V_{\text{rms}} = \frac{V_m}{2^{1/2}} = 0.707 V_m \quad (2-51)$$

for a sine or a cosine function.

If one uses the phasors

$$\mathbf{V} = V_m \exp j\omega t \quad \text{and} \quad \mathbf{I} = \frac{V_m \exp j\omega t}{R}, \quad (2-52)$$

then

$$\mathbf{VI} = \frac{V_m^2 \exp 2j\omega t}{R}, \quad (2-53)$$

whose real part $(V_m^2/R) \cos 2\omega t$ is *neither* the instantaneous *nor* the average power. So phasors must *not* be multiplied in this way.

Suppose one has two sinusoidal quantities of the *same frequency*

$$A = A_m \cos \omega t \quad \text{and} \quad B = B_m \cos (\omega t + \theta). \quad (2-54)$$

Then the time-averaged value of their product is

$$\langle A_m(\cos \omega t)B_m \cos (\omega t + \theta) \rangle = \langle A_m B_m \cos \omega t (\cos \omega t \cos \theta - \sin \omega t \sin \theta) \rangle \quad (2-55)$$

$$= \langle A_m B_m (\cos^2 \omega t \cos \theta - \cos \omega t \sin \omega t \sin \theta) \rangle, \quad (2-56)$$

where the signs $\langle \dots \rangle$ mean “average value of.” Now the average value of $\cos^2 \omega t$ over one full cycle is $\frac{1}{2}$, as we saw above, while the average value of $\cos \omega t \sin \omega t$ is zero. Then

$$\langle A_m(\cos \omega t)B_m \cos (\omega t + \theta) \rangle = \frac{1}{2} A_m B_m \cos \theta = A_{\text{rms}} B_{\text{rms}} \cos \theta. \quad (2-57)$$

If one uses the phasors

$$\mathbf{A} = A_m \exp j\omega t \quad \text{and} \quad \mathbf{B} = B_m \exp j(\omega t + \theta), \quad (2-58)$$

then the average value of the product of their real parts is given correctly by

$$\frac{1}{2} \operatorname{Re} (\mathbf{AB}^*) = \frac{1}{2} \operatorname{Re} \{A_m(\exp j\omega t)B_m \exp [-j(\omega t + \theta)]\} \quad (2-59)$$

$$= \frac{1}{2} \operatorname{Re} [A_m B_m \exp (-j\theta)] = \frac{1}{2} A_m B_m \cos \theta \quad (2-60)$$

$$= A_{\text{rms}} B_{\text{rms}} \cos \theta, \quad (2-61)$$

as above.

2.5 QUOTIENTS OF PHASORS

Dividing one phasor by another of the *same frequency* yields a complex number:

$$\frac{r_1 \exp j(\omega t + \alpha)}{r_2 \exp j(\omega t + \beta)} = \frac{r_1}{r_2} \exp j(\alpha - \beta). \quad (2-62)$$

2.6 ROTATING VECTORS

If one expresses a rotating vector in phasor form, one runs into another kind of trouble. Suppose

$$\mathbf{E} = \hat{\mathbf{x}} E_m \cos \omega t + \hat{\mathbf{y}} E_m \sin \omega t. \quad (2-63)$$

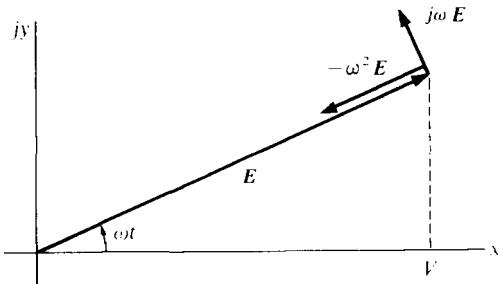


Fig. 2-4. The vector \mathbf{E} has a constant magnitude, but rotates counterclockwise at an angular velocity of ω radians/second. Its time derivative is $d\mathbf{E}/dt$, or $j\omega\mathbf{E}$, in the direction shown. Its second time derivative is $-\omega^2\mathbf{E}$.

Then the vector \mathbf{E} rotates in real space as in Prob. 2-10 and Fig. 2-4, and $d\mathbf{E}/dt$ is perpendicular to \mathbf{E} .

In phasor form,

$$\mathbf{E} = E_m (\exp j\omega t) \hat{x} + E_m \exp j\left(\omega t - \frac{\pi}{2}\right) \hat{y} \quad (2-64)$$

and

$$\frac{d\mathbf{E}}{dt} = j\omega \mathbf{E}. \quad (2-65)$$

This equation is correct. The trouble here is that it *appears* to say that $d\mathbf{E}/dt$ is collinear with \mathbf{E} , which is wrong, as one can see from the figure. The factor j rotates a phasor by $+\pi/2$.

2.7 NOTATION

We have used boldface sans-serif type for phasors, and the usual lightface italic type for the other variables. However, it is customary to use lightface italic type for phasors, as for any other variable, and to omit the operator Re . So, in practice, we write

$$x = x_m \exp j\omega t, \quad \frac{dx}{dt} = j\omega x, \quad \frac{d^2x}{dt^2} = -\omega^2 x, \quad \text{etc.}, \quad (2-66)$$

with the tacit understanding that the imaginary parts are parasitic.

2.8 SUMMARY

A *complex number* z is of the form $a + bj$, where $j = (-1)^{1/2}$ and a and b are real numbers. It is the custom to plot complex numbers in the complex plane as in Fig. 2-1, and thus

$$z = a + bj = r \exp(j\theta). \quad (2-8)$$

The *complex conjugate* of a complex number is its mirror image with respect to the real axis:

$$z^* = a - bj = r \exp(-j\theta). \quad (2-11)$$

Addition and *subtraction* of complex numbers are simpler with the Cartesian form:

$$(a + bj) + (c + dj) = (a + c) + j(b + d). \quad (2-13)$$

However, multiplication and division are simpler with the polar form:

$$(r_1 \exp j\theta_1)(r_2 \exp j\theta_2) = r_1 r_2 \exp j(\theta_1 + \theta_2), \quad (2-16)$$

$$\frac{r_1 \exp j\theta_1}{r_2 \exp j\theta_2} = \frac{r_1}{r_2} \exp j(\theta_1 - \theta_2). \quad (2-17)$$

If one has to deal with the time derivatives of a quantity of the form $I = I_m \cos \omega t$, it is usually advisable to substitute the *phasor*

$$\mathbf{I} = I_m \exp j\omega t = I_m \cos \omega t + jI_m \sin \omega t. \quad (2-26)$$

Then

$$\frac{d\mathbf{I}}{dt} = j\omega \mathbf{I}, \quad \frac{d^2\mathbf{I}}{dt^2} = -\omega^2 \mathbf{I}, \quad \frac{d^n\mathbf{I}}{dt^n} = (j\omega)^n \mathbf{I}, \quad (2-27)$$

with the understanding that only the real parts are meaningful.

The phasor

$$\mathbf{E} = E_m \exp j(\omega t - kz) \quad (2-37)$$

represents a *plane wave* traveling in the positive direction of the z -axis, where k is the *wave number*. The quantities \mathbf{E} and E_m can be vectors. Then \mathbf{E} is both a phasor and a vector.

One occasionally requires the time average of the product of two sinusoidal quantities of *same frequency* such as I and V in an alternating-current circuit. This is given by $\frac{1}{2} \operatorname{Re} \mathbf{IV}^*$, where \mathbf{I} and \mathbf{V} are phasors.

The ratio of two phasors, again of the *same frequency*, is a complex number.

PROBLEMS

2-1. (2.1) Complex numbers in polar form are often written as $r\angle\theta$, where r is the modulus and θ is the argument, expressed in radians.

Express $1 + 2j$ in this way.

2-2. (2.1) (a) Express the complex numbers $1 + 2j$, $-1 + 2j$, $-1 - 2j$, and $1 - 2j$ in polar form.

(b) Simplify the following expressions, leaving them in Cartesian coordinates: $(1 + 2j)(1 - 2j)$, $(1 + 2j)^2$, $1/(1 + 2j)^2$, $(1 + 2j)(1 - 2j)$.

2-3. (2.1) What happens to a complex number in the complex plane when it is (a) multiplied by j , (b) multiplied by j^2 , (c) divided by j ?

2-4. (2.2) Find the real parts of the phasors $(1 + 3j)\exp j(\omega t + 2)$ and $[\exp j(\omega t + 2)]/(1 + 3j)$.

2-5. (2.3) Solve the following differential equation by means of phasors:

$$2 \frac{d^2x}{dt^2} + 3 \frac{dx}{dt} + 4x = 5 \cos 6t.$$

2-6. (2.4) Find the rms values for the waveforms shown in Fig. 2-5.

2-7. (2.4) A certain electric circuit draws a current of 1.00 ampere rms when it is fed at 120 volts rms, 60 hertz. The current lags the voltage by $\pi/4$ radian.

(a) Express V and I in the form of phasors, and calculate the time-averaged power dissipation.

(b) Now calculate the power $V_{\text{rms}} I_{\text{rms}} \cos \theta$, where θ is $\pi/4$.

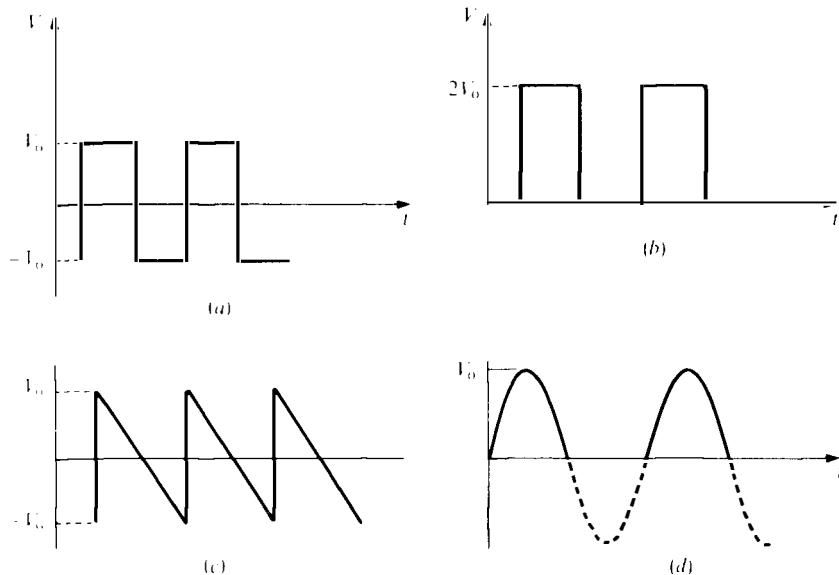
2-8. (2.5) Find the value of

$$\frac{5.14 \exp j(\omega t + 3)}{3.72 \exp j(\omega t + 5)}$$

in Cartesian form.

2-9. (2.5) Show that

$$\ln(a + bj) = \frac{1}{2} \ln(a^2 + b^2) + j \arctan \frac{b}{a}.$$

**Fig. 2-5.**

if a is positive. Then the point $a + bj$ lies in either the first or the fourth quadrant. If a is not positive, be careful to use the proper angle (Sec. 2.1).

- 2-10. (2.6) Show that the phasor $\mathbf{V} = V_m[\hat{\mathbf{x}} \exp j\omega t + \hat{\mathbf{y}} \exp j(\omega t - \pi/2)]$ represents a vector of constant magnitude V_m that rotates in the xy -plane at the angular velocity ω .

CHAPTER 3

ELECTRIC FIELDS I *Coulomb's Law and Gauss's Law*

3.1 COULOMB'S LAW	42
3.1.1 THE AMPERE	44
3.2 THE ELECTRIC FIELD STRENGTH \mathbf{E}	44
3.3 THE PRINCIPLE OF SUPERPOSITION	45
3.4 THE ELECTRIC POTENTIAL V AND THE CURL OF \mathbf{E}	46
3.4.1 THE ELECTRIC POTENTIAL V AT A POINT	48
3.5 THE ELECTRIC FIELD INSIDE AND OUTSIDE MACROSCOPIC BODIES	49
3.6 EQUIPOTENTIAL SURFACES AND LINES OF \mathbf{E}	50
3.7 GAUSS'S LAW	50
<i>Example: THE FIELD OF A UNIFORM SPHERICAL CHARGE DISTRIBUTION</i>	52
<i>Example: THE AVERAGE POTENTIAL OVER A SPHERICAL SURFACE. EARNSHAW'S THEOREM</i>	54
<i>Example: THE AVERAGE \mathbf{E} INSIDE A SPHERICAL VOLUME CONTAINING A POINT CHARGE Q</i>	56
3.8 SUMMARY	57
PROBLEMS	58

In Chaps. 3 to 12 we study the electric fields that result from accumulations of electric charges. The charges are usually stationary. When the charges do move, we assume that their velocities and accelerations are small. This ensures that the electric fields are nearly the same as if the charges were stationary. We also disregard magnetic fields.

This first chapter on electric fields concerns Coulomb's and Gauss's laws. Both are fundamental and well established.

3.1 COULOMB'S LAW

Experiments show that the force exerted by a stationary point charge Q_a on a stationary point charge Q_b situated a distance r away is given by

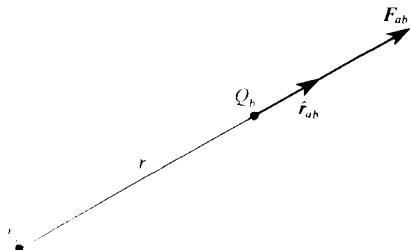


Fig. 3-1. Charges Q_a and Q_b separated by a distance r . Coulomb's law gives the force \mathbf{F}_{ab} exerted by Q_a on Q_b if Q_a is stationary.

$$\mathbf{F}_{ab} = \frac{Q_a Q_b}{4\pi\epsilon_0 r^2} \hat{\mathbf{r}}_{ab}, \quad (3-1)$$

where the unit vector $\hat{\mathbf{r}}_{ab}$ points from Q_a to Q_b , as in Fig. 3-1. This is *Coulomb's law*.† The force is repulsive if the two charges have the same sign, and attractive if they have different signs. The charges are measured in coulombs, the force in newtons, and the distance in meters. The constant ϵ_0 is the *permittivity of free space* and has the following value:

$$\epsilon_0 = 8.854187817 \times 10^{-12} \text{ farad/meter.} \quad (3-2)$$

Substituting the value of ϵ_0 , we find that

$$F_{ab} \approx 9 \times 10^9 \frac{Q_a Q_b}{r^2} \quad \text{newtons,} \quad (3-3)$$

where the factor of 9 is too large by about one part in a thousand.

We shall not be able to define the coulomb until Chap. 22. For the moment, we may take the value of ϵ_0 to be given, and use this law as a provisional definition of the unit of charge.

To what extent does Coulomb's law remain valid when Q_a and Q_b are not stationary?

(1) If Q_a is stationary and Q_b is not, then Coulomb's law applies to the force on Q_b , whatever the velocity of Q_b . This is an experimental fact. Indeed, the trajectories of charged particles in oscilloscopes, mass spectrographs, and ion accelerators are invariably calculated on that basis.

(2) If Q_a is not stationary, Coulomb's law is no longer strictly valid.

† The exponent of r is known to be equal to 2 within one part in 10^{16} .

Coulomb's law applies to a pair of charges situated in a vacuum. It also applies in dielectrics and conductors if F_{ab} is the direct force between Q_a and Q_b , irrespective of the forces arising from other charges within the medium.

With extended charges, "the distance between the charges" has no definite meaning. Moreover, the presence of Q_b can modify the charge distribution within Q_a , and vice versa, leading to a complicated variation of force with distance.

Electric forces in nature are enormous when compared to gravitational forces, for which

$$F = 6.67259 \times 10^{-11} \frac{m_a m_b}{r^2}. \quad (3-4)$$

For example, the gravitational force on a proton at the surface of the sun (mass = 2×10^{30} kilograms, radius = 7×10^8 meters) is equal to the electric force between one proton and one *microgram* of electrons, separated by a distance equal to the sun's radius. Or the electric repulsion between two protons (mass = 9.1×10^{-31} kilogram) is about 4.2×10^{42} times larger than their gravitational attraction.

There are two reasons why, fortunately, we are not normally conscious of the enormous electric forces. First, ordinary matter is truly neutral, or so it seems. Experiments have shown that no atom or molecule carries a charge greater than 10^{-20} times the electronic charge. Second, the mobility of some of the electrons in matter prevents the accumulation of any appreciable quantity of charge of either sign.

3.1.1 The Ampere

If charge flows through, say, a wire at the rate of 1 coulomb/second, then the current is 1 *ampere*. This is not an appropriate definition of the ampere, because it rests on the above provisional definition of the coulomb. See Sec. 22.3.1.

3.2 THE ELECTRIC FIELD STRENGTH \mathbf{E}

The force between two electric charges Q_a and Q_b results from the interaction of Q_b with the *field* of Q_a at the position of Q_b , or vice versa.

We thus define the *electric field strength* \mathbf{E} at a point as the force exerted on a unit test charge situated at that point. Thus, at a distance r from charge Q_a ,

$$\mathbf{E}_a = \frac{\mathbf{F}_{ab}}{Q_b} = \frac{Q_a}{4\pi\epsilon_0 r^2} \hat{\mathbf{r}} \quad \text{newtons/coulomb, or volts/meter,} \quad (3-5)$$

where 1 volt equals 1 joule/coulomb. The field of Q_a is the same, whether the test charge Q_b lies in the field or not, even if Q_b is larger than Q_a .

3.3 THE PRINCIPLE OF SUPERPOSITION

If there are several charges, each one imposes its own field, and the resultant \mathbf{E} is simply the vector sum of all the individual \mathbf{E} 's. This is the *principle of superposition*.

For a continuous distribution of charge, as in Fig. 3-2, the electric field strength at (x, y, z) is

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\rho \hat{\mathbf{r}}}{r^2} dV', \quad (3-6)$$

where ρ is the volume charge density at the source point (x', y', z') , as in the figure, $\hat{\mathbf{r}}$ is the unit vector pointing from the *source* point $P'(x', y', z')$ to the *field* point $P(x, y, z)$, r is the distance between these two points, and dV' is the element of volume $dx' dy' dz'$. If there exist surface distributions of charge, then we must add a similar integral, with

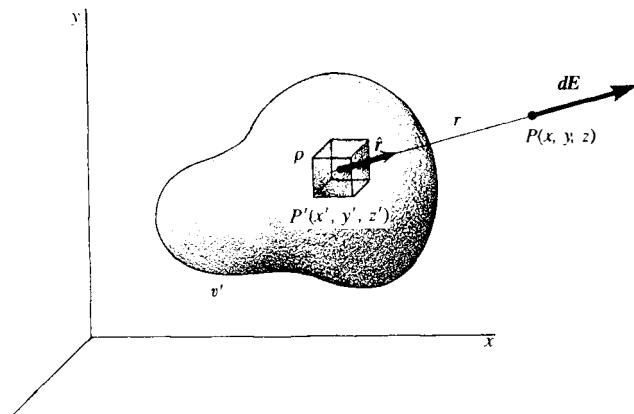


Fig. 3-2. Charge distribution of volume density ρ occupying a volume v' . The element of volume at $P'(x', y', z')$ has a field $d\mathbf{E}$ at $P(x, y, z)$.

ρ replaced by the surface charge density σ and v' by the area \mathcal{A}' of the charged surfaces.

3.4 THE ELECTRIC POTENTIAL V AND THE CURL OF \mathbf{E}

Consider a test charge Q' that can move about in an electric field. The energy \mathcal{E} required to move it at a constant velocity from a point A to a point B along a given path is

$$\mathcal{E} = - \int_A^B \mathbf{E} Q' \cdot d\mathbf{l}. \quad (3-7)$$

Because of the negative sign, \mathcal{E} is the work done *against* the field. We assume that Q' is so small that it does not disturb the charge distributions appreciably.

If the path is closed, the total work done on Q' is

$$\mathcal{E} = - \oint \mathbf{E} Q' \cdot d\mathbf{l}. \quad (3-8)$$

Let us evaluate this integral. We first consider the electric field of a single stationary point charge Q . Then

$$\oint \mathbf{E} Q' \cdot d\mathbf{l} = \frac{QQ'}{4\pi\epsilon_0} \oint \frac{\hat{\mathbf{r}} \cdot d\mathbf{l}}{r^2}. \quad (3-9)$$

Now the term under the integral on the right is simply dr/r^2 , or $-d(1/r)$. But the sum of the increments of $1/r$ over a closed path is zero, since r has the same value at the beginning and at the end. So the line integral is zero, and the net work done in moving Q' around any closed path in the field of Q , which is fixed, is zero.

If the electric field is that of some fixed charge distribution, then the line integrals corresponding to each individual charge of the distribution are all zero. Thus, for any distribution of fixed charges,

$$\oint \mathbf{E} \cdot d\mathbf{l} = 0. \quad (3-10)$$

An electrostatic field is therefore conservative (Example, Sec. 1.9). This important property follows from the fact that the Coulomb force is a central force: the force in the field of a point charge is radial.

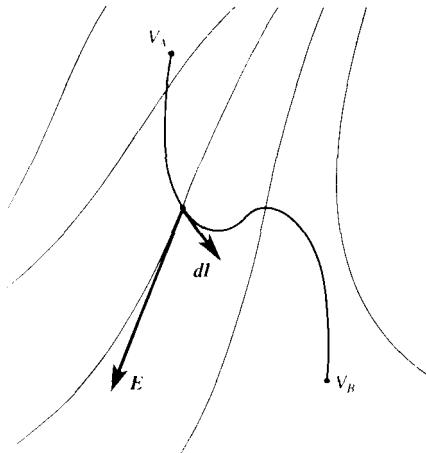


Fig. 3-3. The potential difference $V_A - V_B$ between two points is equal to the line integral of $\mathbf{E} \cdot d\mathbf{l}$ from A to B , where \mathbf{E} is the electric field strength and $d\mathbf{l}$ is an element of the path along which the integral runs. The light lines are lines of \mathbf{E} .

We can now show that the work done in moving a test charge at a constant velocity from a point A to a point B is independent of the path. Let m and n be any two paths leading from A to B . Then these two paths together form a closed curve, and the work done in going from A to B along m and then from B back to A along n is zero. Then the work done in going from A to B is the same along m as it is along n .

Now let us choose a datum point $R(x_0, y_0, z_0)$, and let us define a scalar function V of $P(x, y, z)$ such that

$$V_P = \int_P^R \mathbf{E} \cdot d\mathbf{l}. \quad (3-11)$$

This definition is unambiguous because the integral is the same for all paths leading from P to R . Then, for any pair of points A and B ,

$$-\int_A^B \nabla V \cdot d\mathbf{l} = V_A - V_B = \int_A^B \mathbf{E} \cdot d\mathbf{l}, \quad (3-12)$$

as in Fig. 3-3, and therefore

$$\mathbf{E} = -\nabla V. \quad (3-13)$$

The *electric potential* $V(x, y, z)$ describes the field completely. The negative sign makes \mathbf{E} point toward a *decrease* in V .

Note that V is not uniquely defined, because point R is arbitrary. In fact, one can add to V any quantity that is independent of the coordinates without affecting \mathbf{E} .

From Eq. 3-10 and from Stokes's theorem (Sec. 1.9),

$$\nabla \times \mathbf{E} = 0. \quad (3-14)$$

This is also obvious from the fact that

$$\nabla \times \mathbf{E} = -\nabla \times \nabla V = 0. \quad (3-15)$$

Remember that we are dealing here with *static* fields. If there were time-dependent currents, $\nabla \times \mathbf{E}$ would not necessarily be zero, and $-\nabla V$ would then describe only part of \mathbf{E} . We shall investigate these more complicated phenomena later.

3.4.1 The Electric Potential V at a Point

Equation 3-12 shows that \mathbf{E} concerns only *differences* between the potentials at two points. When one wishes to speak of the potential at a given point, one must arbitrarily define V in a given region of space to be zero. In the previous section, for instance, we made V equal to zero at point R . When the charges extend over only a finite region, it is usually convenient to choose the potential V at infinity to be zero. Then, at point P ,

$$V = \int_P^{\infty} \mathbf{E} \cdot d\mathbf{l}. \quad (3-16)$$

The energy \mathcal{E} required to bring a charge Q from a point where V is zero, by definition, to P is VQ . Thus V is \mathcal{E}/Q , and the unit of V is 1 joule/coulomb, or 1 volt.

If the field is that of a single point charge, then

$$V = \int_r^{\infty} \frac{Q}{4\pi\epsilon_0 r^2} dr = \frac{Q}{4\pi\epsilon_0 r}. \quad (3-17)$$

The sign of this V is the same as that of Q .

The principle of superposition applies to V as well as to \mathbf{E} , and for any charge distribution of density ρ ,

$$V = \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\rho dv'}{r}, \quad (3-18)$$

with r as in Fig. 3-2. The volume v' encloses all the charges. If there are surface charges, one adds a surface integral.

For some geometries, this integral diverges. Then one calculates \mathbf{E} directly, after which one integrates to find V .

3.5 THE ELECTRIC FIELD INSIDE AND OUTSIDE MACROSCOPIC BODIES

Macroscopic bodies consist of positively charged nuclei and negative electrons. This brings up three questions.

(1) Can one calculate the field *outside* an electrically charged body by assuming that the charge distribution inside the body is continuous? If so, then one can calculate the field by integrating over the charge distribution. Otherwise, one must find some other form of calculation.

It is, in fact, usually appropriate to treat the discrete charges carried by nuclei and by electrons within macroscopic bodies as though they were continuous. Even the largest nuclei have diameters that are only of the order of 10^{-14} meter. Nuclei and electrons are so small and so closely packed, compared to the dimensions of ordinary macroscopic objects, that one may assume a smoothly varying electric charge density measured in coulombs per cubic meter or per square meter.

(2) Now what about the electric field *inside* a charged body? Clearly the electric field strength in the immediate neighborhood of a nucleus or of an electron is enormous. Also at a given fixed point, this electric field changes erratically with time, since the charges are never perfectly stationary. It is not useful for our purposes to look at the electric field as closely as that. We shall be satisfied to calculate space- and time-averaged values of \mathbf{E} and V inside a charged body by assuming a continuous distribution of charge.

(3) Is it, then, really possible to define the electric field at a point P inside a continuously distributed charge? It appears at first sight that the dV contributed by the charge element $\rho dv'$ at P is infinite, since r is zero. In fact, it is not infinite.

Consider a spherical shell of thickness dr and radius r centered on P . The charge in this shell contributes at P a dV of $(4\pi r^2 dr \rho)/(4\pi\epsilon_0 r) = dr \rho$. Another shell of smaller radius contributes a smaller dV . The electric potential V therefore converges, and the integral is finite. A similar argument shows that \mathbf{E} also converges.

One can therefore calculate the electric fields of real charge distributions by the usual techniques of the integral calculus, both inside and outside the distributions.

3.6 EQUIPOTENTIAL SURFACES AND LINES OF \mathbf{E}

The set of all points in space that are at a given potential defines an *equipotential surface*. For example, the equipotential surfaces about a point charge are concentric spheres. Since $\mathbf{E} = -\nabla V$ (Eq. 3-13), \mathbf{E} is everywhere normal to the equipotential surfaces (Sec. 1.2).

If we join end-to-end infinitesimal vectors $d\mathbf{s}$ pointing in the direction of \mathbf{E} , we obtain a *line of \mathbf{E}* that is everywhere normal to the equipotential surfaces. We shall return to lines of \mathbf{E} in Sec. 6.5.1.

3.7 GAUSS'S LAW

Gauss's law relates the flux of \mathbf{E} through a closed surface to the total charge enclosed within that surface.

Consider Fig. 3-4, in which a finite volume v bounded by a surface \mathcal{A} encloses a charge Q . We can calculate the outward flux of \mathbf{E} through \mathcal{A} as follows. The flux of \mathbf{E} through the element of area $d\mathcal{A}$ is

$$\mathbf{E} \cdot d\mathcal{A} = \frac{Q}{4\pi\epsilon_0} \frac{\hat{r} \cdot d\mathcal{A}}{r^2}. \quad (3-19)$$

Now $\hat{r} \cdot d\mathcal{A}$ is the projection of $d\mathcal{A}$ on a plane normal to \hat{r} . Then

$$\mathbf{E} \cdot d\mathcal{A} = \frac{Q}{4\pi\epsilon_0} d\Omega, \quad (3-20)$$

where $d\Omega$ is the solid angle subtended by $d\mathcal{A}$ at the point P' .

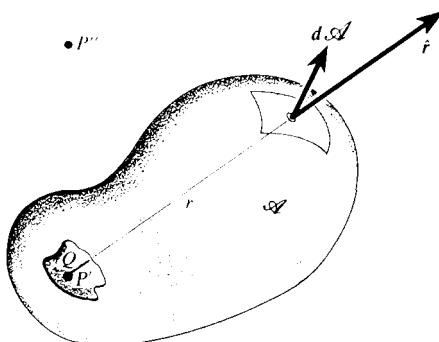


Fig. 3-4. A point charge Q located inside a volume v bounded by the surface of area \mathcal{A} . Gauss's law states that the surface integral of $\mathbf{E} \cdot d\mathcal{A}$ over \mathcal{A} is equal to Q/ϵ_0 . The vector $d\mathcal{A}$ points *outward*.

To find the outward flux of \mathbf{E} , we integrate over the area \mathcal{A} , or over a solid angle of 4π . Thus

$$\int_{\mathcal{A}} \mathbf{E} \cdot d\mathcal{A} = \frac{Q}{\epsilon_0}. \quad (3-21)$$

If Q is outside the surface at P'' , the integral is equal to zero. The solid angle subtended by any closed surface (or set of closed surfaces) is 4π at a point P' inside and zero at a point P'' outside.

If more than one charge resides within v , the fluxes add algebraically and the total flux of \mathbf{E} leaving v is equal to the total enclosed charge Q divided by ϵ_0 :

$$\int_{\mathcal{A}} \mathbf{E} \cdot d\mathcal{A} = \frac{Q}{\epsilon_0}. \quad (3-22)$$

This is *Gauss's law* in integral form.[†]

If the charge occupies a finite volume, then

$$\int_{\mathcal{A}} \mathbf{E} \cdot d\mathcal{A} = \frac{1}{\epsilon_0} \int_v \rho dv, \quad (3-23)$$

where \mathcal{A} is the area of the surface bounding the volume v , and ρ is the electric charge density. We assumed that there are no surface charges on the bounding surface.

If we apply the divergence theorem to the left-hand side, we have that

$$\int_v \nabla \cdot \mathbf{E} dv = \frac{1}{\epsilon_0} \int_v \rho dv. \quad (3-24)$$

Since this equation applies to any finite volume v , the integrands are equal and

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad (3-25)$$

at every point in space.

[†]We have followed the usual custom of starting out with Coulomb's law and then deducing from it Gauss's law. This procedure seems rational enough, but the latter law is, in fact, more general. Indeed, Gauss's law applies to moving charges, whatever be their velocity or acceleration, while Coulomb's law, as stated in Sec. 3.1, is valid only if Q_a is stationary.

This is *Gauss's law in differential form*. Observe that it relates the local charge density to the *derivatives* of E , and not to E itself.

When it is expressed in differential form, Gauss's law is a *local* law in that it relates the behavior of E in the infinitesimal neighborhood of a given point to the value of the charge density at that point. However, when it is expressed in integral form, as in Eq. 3-22, Gauss's law is *nonlocal*, because it concerns a finite region and not a specific point in space.

Many laws of nature, in particular the fundamental laws of electromagnetism, can be formulated in two such equivalent forms, one local and one nonlocal. With the local forms of physical laws, in the guise of differential equations, one views phenomena as the result of processes occurring in the immediate neighborhood of every point in space.

Example

THE FIELD OF A UNIFORM SPHERICAL CHARGE DISTRIBUTION

A spherical charge distribution has a radius R and a uniform density ρ , as in Fig. 3-5. Let us find E and V as functions of the distance r from the center of the sphere. By symmetry, both E and V are independent of the spherical coordinates θ and ϕ .

By symmetry, E is radial. It points outward if Q is positive.

(a) The electric field strength E

Outside the charge distribution, at P' where $r > R$, we imagine a sphere of radius r and surface area $4\pi r^2$. The enclosed charge is

$$Q = \frac{4}{3}\pi R^3 \rho. \quad (3-26)$$

From Gauss's law,

$$E_o = \frac{Q}{4\pi\epsilon_0 r^2} = \frac{R^3 \rho}{3\epsilon_0 r^2} \quad (3-27)$$

as if all the charge were situated at the center O .

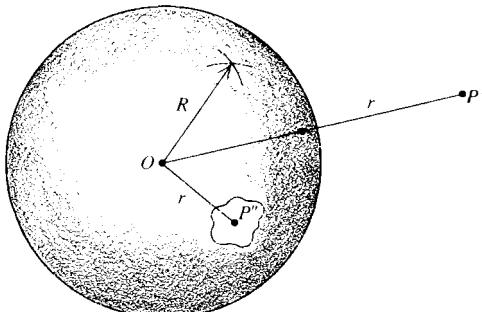


Fig. 3-5. Spherical charge distribution.

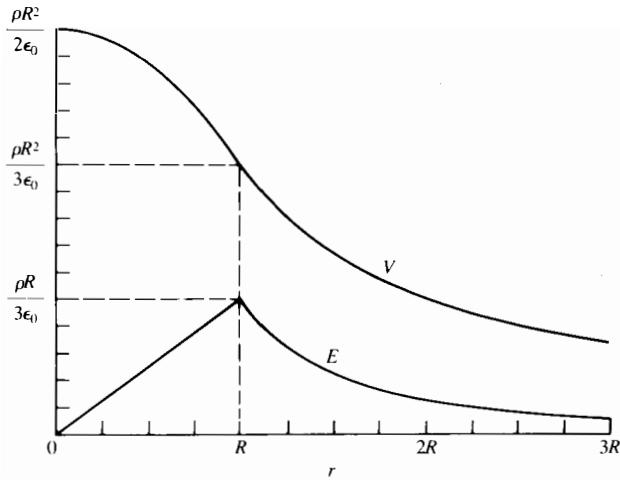


Fig. 3-6. The potential V and the electric field strength E as functions of the radial distance r from a spherical charge distribution of radius R and volume charge density ρ .

Inside the sphere, at P'' , the charge enclosed by the imaginary sphere is $Q(r/R)^3$. Using again Gauss's law, we find that

$$E_i = \frac{Q(r/R)^3}{4\pi\epsilon_0 r^2} = \frac{Qr}{4\pi\epsilon_0 R^3} = \frac{\rho r}{3\epsilon_0}. \quad (3-28)$$

Figure 3-6 shows E as a function of r .

Note that $E_i = E_o$ at $r = R$. This is in accordance with Gauss's law because a spherical shell of infinitesimal thickness just inside the surface of the sphere carries zero charge.

(b) The electric potential V

At P' ,

$$V_o = \frac{Q}{4\pi\epsilon_0 r} \quad (3-29)$$

because the field is the same as that of a point charge Q at O .

To find the potential at P'' , we use Eq. 3-16:

$$V_i = \int_r^\infty E dr = \int_r^R E_i dr + \int_R^\infty E_o dr. \quad (3-30)$$

The last integral is simply the potential at $r = R$ of a point charge Q situated at O , or $Q/(4\pi\epsilon_0 R)$. Thus

$$V_i = \int_r^R \frac{Qr dr}{4\pi\epsilon_0 R^3} + \frac{Q}{4\pi\epsilon_0 R} = \frac{Q}{4\pi\epsilon_0 R} \left(\frac{3}{2} - \frac{r^2}{2R^2} \right). \quad (3-31)$$

See Fig. 3-6.

Example

THE AVERAGE POTENTIAL OVER A SPHERICAL SURFACE. EARNSHAW'S THEOREM

As another illustration of the use of Gauss's law, we shall prove that the average potential $\langle V \rangle$ over any spherical surface of radius R and center O has the following two properties: if there are no charges inside, then $\langle V \rangle$ is equal to the value of V at the center O . If there is a net charge Q inside and no charge outside, then $\langle V \rangle$ is equal to $Q/(4\pi\epsilon_0 R)$.

(a) Let us first think of a spherical shell of radius R carrying a uniform surface charge density σ and a total charge $Q = 4\pi R^2 \sigma$, as in Fig. 3-7a. Then, from Gauss's law, at some point P *outside*, at a distance r from O ,

$$E = \frac{Q}{4\pi\epsilon_0 r^2} \quad \text{and} \quad V = \frac{Q}{4\pi\epsilon_0 r}. \quad (3-32)$$

But

$$V = \int \frac{\sigma dA}{4\pi\epsilon_0 r'} = \frac{Q}{4\pi R^2} \int \frac{dA}{4\pi\epsilon_0 r'}, \quad (3-33)$$

where A is the area of the shell.

Equating these two values of V gives a purely geometric relation concerning a sphere of radius R and a point P at a distance $r > R$ from its center:

$$\frac{Q}{4\pi\epsilon_0 r} = \frac{1}{4\pi R^2} \int_A \frac{Q dA}{4\pi\epsilon_0 r'}. \quad (3-34)$$

We now shift our attention to Fig. 3-7b. We now have an imaginary sphere of radius R and a charge Q situated outside at a distance r . The left-hand side of Eq. 3-34 is equal to the potential at the center of the imaginary sphere, and the right-hand side is the average potential $\langle V \rangle$ on the spherical surface. So we have demonstrated that, in Fig. 3-7b, the average V on the sphere is just the value of V at the center.

This result applies to any charge distribution situated outside the sphere, as in Fig. 3-7c, because of the principle of superposition (Sec. 3.3).

Now imagine for a moment that there is a potential maximum at some point O in a region where $\rho = 0$. Then the average potential over some sphere centered on O must be lower than the potential at O , which is contrary to the above result. *Thus there can never be a potential maximum in a charge-free region. For the same reason, there can never be a potential minimum either.*

This is *Earnshaw's theorem*.

It is occasionally desirable to create a potential well in space so as to trap either ions or electrons. Earnshaw's theorem shows that this is impossible with electrostatic fields.

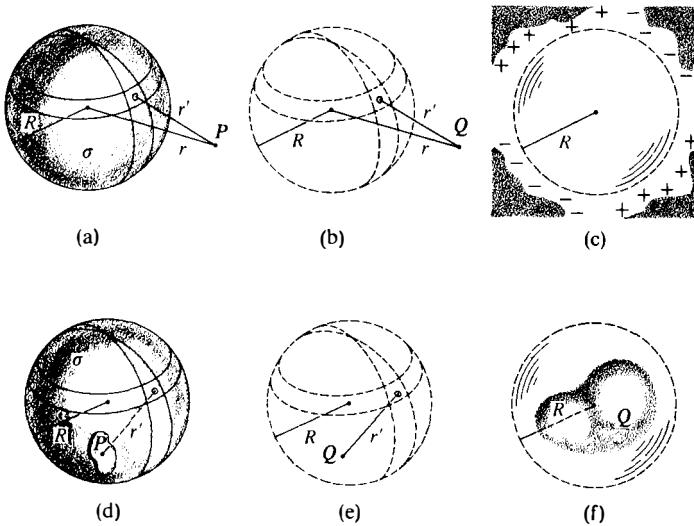


Fig. 3-7. (a) Spherical surface carrying a uniform charge distribution σ . (b) Point charge Q outside an imaginary spherical surface of radius R . (c) Imaginary spherical surface in a charge-free region. The average potential over the surface is equal to the potential at the center. (d) Spherical surface carrying a uniform charge distribution σ . Point P is situated inside. (e) Point charge Q inside an imaginary spherical surface of radius R . (f) Imaginary spherical surface enclosing a charge Q . The average potential over the surface is equal to $Q/(4\pi\epsilon_0 R)$.

(b) We can proceed in a similar fashion to find the average potential over a spherical surface when the charges are *inside*.

We start again with a charge Q spread uniformly over the surface, as in Fig. 3-7d. At any point P inside, E is zero for the following reason. By symmetry, E_θ and E_ϕ are zero. To find E_r , we apply Gauss's law to a concentric spherical surface having a radius smaller than R and thus enclosing zero charge. We find that $E_r = 0$. Then $E = 0$ inside, and the V at a point P inside is equal to the V at the surface, namely, $Q/(4\pi\epsilon_0 R)$. So, at P in Fig. 3-7d,

$$V = \frac{Q}{4\pi\epsilon_0 R} = \int_{\text{sphere}} \frac{\sigma dA}{4\pi\epsilon_0 r'} = \frac{1}{4\pi R^2} \int_{\text{sphere}} \frac{Q dA}{4\pi\epsilon_0 r'}. \quad (3-35)$$

The last term is just the average potential over the imaginary sphere of Fig. 3-7e. So the average potential over a spherical surface of radius R containing a point charge Q is $Q/(4\pi\epsilon_0 R)$, regardless of the position of Q inside! The same applies to a charge distribution Q of finite volume inside the sphere, as in Fig. 3-7f.

*Example*THE AVERAGE \mathbf{E} INSIDE A SPHERICAL VOLUME CONTAINING A POINT CHARGE Q

Figure 3-8 shows an imaginary spherical volume of radius R containing a point charge Q at \mathbf{r}' , and a point P at $\mathbf{r}' + \mathbf{r}$. We wish to find the average \mathbf{E} inside the sphere.

First, at the point P ,

$$\mathbf{E} = \frac{Q\hat{\mathbf{r}}}{4\pi\epsilon_0 r^2}. \quad (3-36)$$

Then, over the volume v of the sphere, the average \mathbf{E} is

$$\langle \mathbf{E} \rangle = \frac{1}{\frac{4}{3}\pi R^3} \int_v \frac{Q}{4\pi\epsilon_0 r'^2} \hat{\mathbf{r}} dv \quad (3-37)$$

$$= \frac{1}{4\pi\epsilon_0} \int_v \frac{Q}{\frac{4}{3}\pi R^3 r'^2} \hat{\mathbf{r}} dv. \quad (3-38)$$

Thanks to Gauss's law, we can find the value of this last integral without much effort. Suppose that, instead of having a single point charge Q , we had a uniform charge distribution of density $Q/(\frac{4}{3}\pi R^3)$. Then, from Coulomb's law, the \mathbf{E} at \mathbf{r}' would be given by the term on the right in Eq. 3-38, and this is $\rho\mathbf{r}'/(3\epsilon_0)$, from the example on page 52.

So, finally, the average \mathbf{E} inside a spherical volume containing a charge Q situated at \mathbf{r}' is

$$\langle \mathbf{E} \rangle = -\frac{Q\hat{\mathbf{r}}'}{4\pi\epsilon_0 R^3}. \quad (3-39)$$

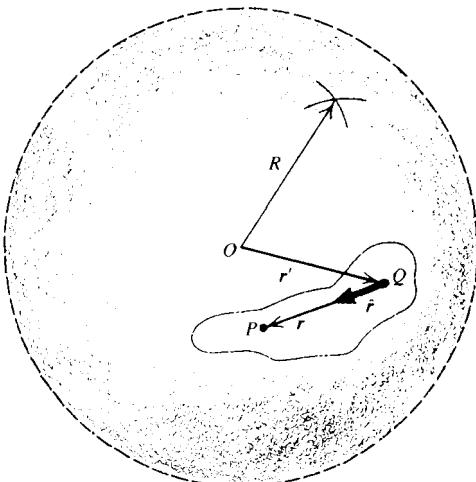


Fig. 3-8. Point charge Q inside a spherical volume of radius R .

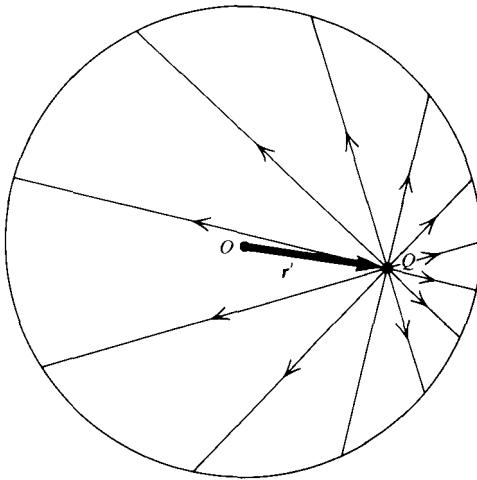


Fig. 3-9. Positive charge Q situated at r' inside an imaginary sphere. The average field of Q inside the sphere points in the direction opposite to r' .

| The reason for the negative sign should be clear from Fig. 3-9.

3.8 SUMMARY

The force exerted by a stationary point charge Q_a on a point charge Q_b , either stationary or in motion, is given by

$$\mathbf{F}_{ab} = \frac{Q_a Q_b}{4\pi\epsilon_0 r^2} \hat{\mathbf{r}}_{ab} \approx 9 \times 10^9 \frac{Q_a Q_b}{r^2} \hat{\mathbf{r}}_{ab}, \quad (3-1)$$

where $\epsilon_0 = 8.85 \times 10^{-12}$ farad/meter, r is the distance between the charges, and $\hat{\mathbf{r}}_{ab}$ is the unit vector pointing from Q_a to Q_b . This is Coulomb's law.

We consider the force to result from the interaction between Q_b and the electric field \mathbf{E}_a of Q_a at the position occupied by Q_b , and

$$\mathbf{E}_a = \frac{Q_a}{4\pi\epsilon_0 r^2} \hat{\mathbf{r}}, \quad (3-5)$$

where the unit vector $\hat{\mathbf{r}}$ points away from Q_a .

A current flowing at the rate of 1 coulomb/second has a magnitude of 1 ampere.

According to the principle of superposition, two or more \mathbf{E} 's acting at the same point add vectorially.

An electrostatic field is conservative:

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = 0, \quad (3-10)$$

where C is any closed curve. It follows that

$$\nabla \times \mathbf{E} = 0 \quad (3-14)$$

and that

$$\mathbf{E} = -\nabla V, \quad (3-13)$$

where

$$V = \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\rho dv'}{r} \quad (3-18)$$

is the *electric potential* at a point P . Here ρ is the volume charge density, r is the distance between $P(x, y, z)$ and the element of volume dv' at $P'(x', y', z')$, and v' encloses all the charges. This integral applies to finite charge distributions and it assumes that $V = 0$ at infinity.

Gauss's law follows from Coulomb's law. In integral form,

$$\int_{\mathcal{A}} \mathbf{E} \cdot d\mathbf{A} = \frac{Q}{\epsilon_0}, \quad (3-21)$$

where Q is the net charge contained inside the closed surface of area \mathcal{A} . In differential form,

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}. \quad (3-25)$$

PROBLEMS

3-1. (3.1)† Coulomb's law

The force of attraction between two charges of 1 coulomb and of opposite signs, separated by a distance of 1 meter, is about 9×10^9 newtons.

How large is a cube of lead that has a weight of 9×10^9 newtons? Lead has a density of 1.13×10^4 kilograms/meter³.

3-2. (3.4) Electrostatic seed-sorting device

It is possible to separate normal seeds from discolored ones and from foreign objects by means of a device that operates as follows. The seeds

† Section number.

drop one by one between a pair of photocells. If the color is not right, voltage is applied to a needle that deposits a charge on the seed. The seeds then fall between a pair of electrically charged plates that deflect the undesired ones into a separate bin. One such machine can sort peas at the rate of 100 per second, or about 2 metric tons per 24-hour day.

(a) If the seeds fall at the rate of 100 per second, over what distance must they fall if they must be spaced vertically by 20 millimeters when they pass between the photocells? Neglect air resistance.

(b) Assume that the seeds acquire a charge of 1.5×10^{-9} coulomb, that the deflecting plates are parallel and 50 millimeters apart, and that the potential difference between them is 25,000 volts. How long should the plates extend below the charging needle if the charged seeds must deflect by 40 millimeters on leaving the plates? Assume that the charging needle and the top of the deflecting plates are close to the photocell.

3-3. (3.4) Rutherford discovers the nucleus

In 1906, in the course of a historic experiment that demonstrated the small size of the atomic nucleus, Rutherford observed that an alpha particle ($Q_1 = 2 \times 1.6 \times 10^{-19}$ coulomb) having a kinetic energy of 7.68×10^6 electron volts ($7.68 \times 10^6 \times 1.6 \times 10^{-19}$ joule) rebounds backward in a head-on collision with a gold nucleus ($Q_2 = 79 \times 1.6 \times 10^{-19}$ coulomb).

(a) What is the distance of closest approach where the electrostatic potential energy is equal to the initial kinetic energy? Express your result in femtometers (10^{-15} meter).

(b) What is the maximum force of repulsion?

(c) What is the maximum acceleration in g 's? The mass of the alpha particle is about 4 times that of a proton, or $4 \times 1.7 \times 10^{-27}$ kilogram.

3-4. (3.4) Electrostatic ion thruster

Ion thrusters correct either the attitude or the trajectory of satellites.

The force exerted by a thruster is equal to $m'v$, where m' is the mass of propellant ejected per second and v is the exhaust velocity with respect to the thruster.

Figure 3-10 shows a schematic diagram of a thruster that ejects a beam of charged particles. The propellant enters at P and is ionized in S . Electrodes

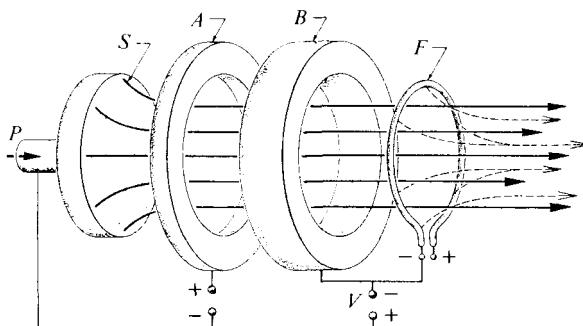


Fig. 3-10

A and *B* form a lens that accelerates the positive ions. A beam of positive ions exits on the right at a velocity determined by the accelerating voltage *V*. The ions of mass *m* carry charges *ne*, where *e* is the magnitude of the electronic charge. The current is *I*. Electrons emitted by the filament *F* neutralize the beam so as to prevent the satellite from charging up.

- Show that the thrust is given by $F = I[2Vm/(ne)]^{1/2}$.
- What is the value of *F* for a 0.1-ampere beam of protons when *V* = 50 kilovolts?
- If *P* is the power *IV* spent in accelerating the particles, show that

$$F = (2Pm')^{1/2} = \frac{2P}{v} = P\left(\frac{2m}{neV}\right)^{1/2}.$$

Thus, for given values of *P* and *m'*, the thrust is independent of the charge-to-mass ratio of the ions. Or, for a given *P*, *F* is inversely proportional to *v*. The last expression shows that, for a given power expenditure *P*, it is preferable to use heavy ions carrying a single charge (*n* = 1) and to use as low an accelerating voltage *V* as possible.

(d) If the electron source is turned off and if the beam current *I* is 1 ampere, how long will it take the body of the rocket to attain a voltage equal to the accelerating voltage, if *V* is 50 kilovolts? Assume that the rocket is spherical and that it has a radius of 1 meter. At that point the thruster ceases to operate because the ions follow the satellite.

3-5. (3.7) Possible and impossible fields

An electric field points everywhere in the *z*-direction.

- What can you conclude about the value of the partial derivatives of *E* with respect to *x*, *y*, *z* (i) if the space charge density *ρ* is zero and (ii) if *ρ* is not zero?
- Sketch lines of *E* for one possible and for one impossible field, on the assumption that $\nabla \times E = 0$.

3-6. (3.7) The conduction electron density at the surface of electrically charged copper

A copper atom has a diameter of about 0.3 nanometer.

- Calculate (i) the approximate number of atoms per square meter, (ii) the approximate charge density that would result if each atom gained one free electron, and (iii) the corresponding electric field strength.
- The maximum possible electric field strength in air is 3×10^6 volts/meter. How far apart are the excess electrons at that value of *E*?

3-7. (3.7) The earth's electric charge

The electric field strength in the atmosphere near the surface of the earth is about 100 volts/meter and points downward. The potential increases with increasing height, up to about 300,000 volts. This field is maintained by thunderstorms, which deposit negative charge on the earth at the average rate of about 10^3 amperes.

Calculate the electric charge carried by the earth.

3-8. (3.7) The coaxial line

Figure 33-4 shows a coaxial line.

Show that, at a distance ρ from the axis in the region between the two conductors, $E = \lambda/(2\pi\epsilon_0\rho)$, where λ is the charge per unit length on the inner conductor. The vector \mathbf{E} points outward if λ is positive.

3-9. (3.7) The force between a point charge and a line charge

A uniform linear distribution of charge of λ coulombs/meter is situated at a distance r from a point charge Q of opposite sign.

(a) Calculate the force of attraction.

(b) Show that the force is the same as if the linear distribution were replaced by a single charge $Q' = 2\lambda r$ situated at the foot of the perpendicular drawn from Q .

3-10. (3.7) Proton beam

A 1.00-microampere beam of protons is accelerated through a difference of potential of 10,000 volts.

(a) Calculate the charge density in the beam, once the protons have been accelerated, assuming that the current density is uniform over a diameter of 2.00 millimeters and is zero outside.

(b) Calculate the radial E both inside and outside the beam.

(c) Draw a graph of the radial E for values of r ranging from 0 to 10.0 millimeters.

(d) The beam is situated on the axis of a grounded cylindrical conducting tube with an inside radius of 10.0 millimeters. Draw a graph of V inside the tube.

(e) Calculate the electric charge density per unit length on the inside of the tube.

3-11. (3.7) The field of an atomic nucleus

The radial dependence of the electric charge density inside a certain atomic nucleus of radius a is roughly described by $\rho = \rho_0(1 - r^2/a^2)$, for $r \leq a$, where $\rho_0 = 5.0 \times 10^{25}$ coulombs/meter³ and $a = 3.4$ femtometers.

(a) What is the total charge Q ?

(b) Find E and V outside the nucleus. What are the values of E and V at the surface?

(c) Find E and V inside the nucleus. What is the value of V at the center?

(d) Show that E is maximum at $r/a = 0.745$.

(e) Draw graphs showing $E/(2\rho_0/15\epsilon_0)$ and $V/(2\rho_0/15\epsilon_0)$ as functions of r for $r/a = 0$ to 5.

3-12. (3.7) Van de Graaff accelerator

A Van de Graaff particle accelerator has a high-voltage electrode maintained under pressure in gaseous SF_6 in a metal tank. It is possible to maintain much higher voltages in this way than if the electrode were in air.

Assume that the electrode is spherical and that its radius is r_1 . Its voltage is V . The tank has a radius r_2 and is grounded. The electric field strength is highest at the surface of the electrode. You are required to find values of r_1 and r_2 that will minimize this E .

For a given value of r_1 , the optimum value of r_2 is infinite, which is absurd. Of course, cost, weight, and space limit r_2 . So you must optimize r_1 for a given r_2 , which is 483 millimeters in one specific case.

- (a) Show that E at the surface of the high-voltage electrode ($r = r_1$) has a minimum value of $2V/r_1$ when $r_1 = r_2/2$.
- (b) Explain qualitatively why there is an optimum radius r_1 .
- (c) Identifying an optimum condition is not sufficient. You must also evaluate how critical the condition is. So plot E/V at $r = r_1$ for $r_2 = 0.483$ meter and for values of r_1 ranging from 100 to 400 millimeters.
- (d) What range of values of r_1 is permissible if E can be 10% larger than $2V/r_1$?
- (e) Calculate $2V/r_1$ for $V = 5 \times 10^5$ volts and for the optimum r_1 .

3-13. (3.7) The equilibrium potential at the surface of a star

- (a) Calculate the escape energy for a particle of mass m and charge q situated at the surface of a star of mass M , charge Q , and radius R .
- (b) Calculate the equilibrium potential V of the star. Assume a sphere of fully ionized atomic hydrogen, with the electrons and protons at the same temperature and zero net current. The fraction of the electrons, or protons, that possess enough energy to escape is

$$\exp\left(-\frac{\text{escape kinetic energy}}{kT}\right),$$

where k is Boltzmann's constant, 1.37×10^{-28} joule per degree. At equilibrium, the electron and proton currents are equal. It is this phenomenon that causes the solar wind.

- (c) Show that $V \approx 10^3$ volts for the sun.

This phenomenon does not appear to have any appreciable astrophysical significance. Even giant galaxies have center-to-surface potential differences that are only of the order of 1000 volts, like the sun.

3-14. (3.7) Electrostatic precipitation

Electrostatic precipitation serves to eliminate dust particles from industrial gases, for example, to eliminate fly ash from the smoke of coal-fired electric power plants. A corona discharge ionizes the gas, and the ions charge the dust particles, which drift in the electric field to the electrodes, where they collect. Periodically, the electrodes are shaken, and the dust falls into a container.

In one type of precipitator, the anode is a grounded cylinder having a radius R of 150 millimeters, and the cathode is an axial wire maintained at a potential V of -50 kilovolts. The gas ionizes, and ions of both signs form in the corona discharge near the wire. The positive ions quickly reach the center wire, while the negative ions move out radially to the cylinder. The space charge is thus negative over most of the volume of the cylinder.

Under those conditions, experiments show that E is approximately uniform and equal to V/R for all values of r . If the dust particles are at least slightly conducting, they acquire a negative charge Q of $12\pi\epsilon_0 Ea^2$, where a is their radius. The charge is somewhat smaller if they are nonconducting.

- (a) Let I be the radial electric current per meter and \mathcal{M} the mobility (speed/ E) of the negative ions.

Show that, for any r , $I = 2\pi r \rho M E$, where ρ is the space charge density $\epsilon_0 E/r$.

(b) The drift velocity of the dust particles is given by Stokes's law: it is the force EQ divided by $6\pi\eta a$, where η is the viscosity of the gas.

Show that their drift velocity v is $2\epsilon_0 E^2 a / \eta$.

(c) Calculate I , ρ , v , and the time required for a dust particle to drift from the cathode to the anode when $M = 2 \times 10^{-4}$ meter²/(volt-second), $a = 5$ micrometers, and $\eta = 2 \times 10^{-5}$ kilogram/(meter-second).

This simplified theory neglects turbulence, which is important in practice.

3-15. (3.7) The expansion of the universe

In 1959 Lyttleton and Bondi suggested that the expansion of the universe could be explained on the basis of Newtonian mechanics if matter carries a net electric charge.

Imagine a spherical volume V of astronomical size containing un-ionized atomic hydrogen of uniform density N atoms per cubic meter, and assume that the proton charge e_p is equal to $(1 + y)e$, where e is the magnitude of the electron charge.

(a) Find E at the radius R .

(b) Show that, for $y > 10^{-18}$, the electric repulsion becomes greater than the gravitational attraction, so the gas expands.

(c) Show that the force of repulsion on an atom is then proportional to its distance R from the center and that, as a consequence, the radial velocity of an atom at R is proportional to R . Assume that the density is maintained constant by the continuous creation of matter in space.

(d) Show that the velocity v is R/T , where T is the time required for the radial distance R of a given atom to increase by a factor of e . This time T can be taken to be the age of the universe.

(e) In the Millikan oil-drop experiment, an electrically charged droplet of oil is suspended in the electric field between two plane horizontal electrodes. It is observed that the charge carried by the droplet changes by integral amounts within an accuracy of about 1 part in 10^5 .

Show that the Millikan oil-drop experiment leads us to believe that y is less than about 10^{-16} .

3-16. (3.7) The volume average of E over a spherical volume is equal to the value of E at the center. An alternative proof.

We know that the force exerted by a uniform spherical charge distribution on an outside charge is the same as if the spherical charge were concentrated at its center.

(a) Use this fact to show that the field of a point charge is such that its volume average over a sphere is equal to its value at the center.

(b) Show that the same applies to any electrostatic field in a charge-free region.

CHAPTER 4

ELECTRIC FIELDS II

The Equations of Poisson and of Laplace. Charge conservation. Conductors

4.1	THE EQUATIONS OF POISSON AND OF LAPLACE	65
	<i>Example: THE FIELD OF A UNIFORM SPHERICAL CHARGE DISTRIBUTION</i>	65
	<i>Example: THE VACUUM DIODE</i>	66
4.2	THE LAW OF CONSERVATION OF ELECTRIC CHARGE	69
4.3	CONDUCTION	69
4.3.1	RESISTANCE	70
	<i>Examples</i>	71
4.3.2	CONDUCTION IN A STEADY ELECTRIC FIELD	72
4.3.3	THE MOBILITY M OF CONDUCTION ELECTRONS	73
4.3.4	CONDUCTION IN AN ALTERNATING ELECTRIC FIELD	74
4.3.5	THE VOLUME CHARGE DENSITY ρ IN A CONDUCTOR	75
4.3.6	THE JOULE EFFECT	76
4.4	ISOLATED CONDUCTORS IN STATIC FIELDS	77
	<i>Example: HOLLOW CONDUCTOR ENCLOSING A CHARGED BODY</i>	77
4.5	SUMMARY	78
	PROBLEMS	79

Poisson's equation relates the local volume charge density ρ to the spatial rates of change of the potential V . This is again a fundamental relation. It follows immediately from Gauss's law. Laplace's equation is Poisson's, with ρ equal to zero. Both equations serve to calculate electric fields. We shall return to them in Chaps. 11 and 12.

Charge conservation is an experimental fact. Whatever the circumstances, the net electric charge carried by a closed system is constant. We shall frame that law in a simple mathematical form and then apply it a few sections later. We shall return to it on several occasions.

The major part of this chapter pertains to conductors. Ordinary electric conductors contain conduction electrons that drift in the direction opposite to the applied E . We shall find, among other things, that this drift velocity is surprisingly low and that the net volume charge density is normally zero.

4.1 THE EQUATIONS OF POISSON AND OF LAPLACE

Let us replace E by $-\nabla V$ in Eq. 3-25. Then

$$\nabla^2 V = -\frac{\rho}{\epsilon_0}. \quad (4-1)$$

This is *Poisson's equation*. It relates the space charge density ρ at a given point to the second space derivatives of V in the region of that point.

In a region where the charge density ρ is zero,

$$\nabla^2 V = 0, \quad (4-2)$$

which is *Laplace's equation*.

The general problem of finding V in the field of a given charge distribution amounts to finding a solution to either Laplace's or Poisson's equation that will satisfy the given boundary conditions.

Example

THE FIELD OF A UNIFORM SPHERICAL CHARGE DISTRIBUTION

Consider again a spherical charge distribution of uniform volume density ρ and radius R as in Fig. 3-5.

Outside the sphere, $\rho = 0$ and

$$\nabla^2 V_o = 0. \quad (4-3)$$

Now, by symmetry, V_o is independent of both θ and ϕ . Therefore, from Sec. 1.11.6,

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial V_o}{\partial r} \right) = 0, \quad \frac{\partial}{\partial r} \left(r^2 \frac{\partial V_o}{\partial r} \right) = 0, \quad (4-4)$$

$$\frac{\partial V_o}{\partial r} = \frac{A}{r^2}, \quad E_o = -\frac{A}{r^2}, \quad (4-5)$$

where A is a constant of integration. This is in agreement with Eq. 3-27 with $A = -Q/(4\pi\epsilon_0)$.

Inside the sphere,

$$\nabla^2 V_i = -\frac{\rho}{\epsilon_0}, \quad (4-6)$$

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial V_i}{\partial r} \right) = -\frac{\rho}{\epsilon_0}, \quad (4-7)$$

$$\frac{\partial}{\partial r} \left(r^2 \frac{\partial V_i}{\partial r} \right) = -\frac{\rho r^2}{\epsilon_0}, \quad (4-8)$$

$$r^2 \frac{\partial V_i}{\partial r} = -\rho \frac{r^3}{3\epsilon_0} + B, \quad (4-9)$$

$$E_i = \frac{\rho r}{3\epsilon_0} - \frac{B}{r^2}, \quad (4-10)$$

where B is another constant of integration.

It is intuitively obvious that E_i cannot become infinite at $r = 0$; so B is zero and

$$E_i = \frac{\rho r}{3\epsilon_0} = \frac{Qr}{4\pi\epsilon_0 R^3}, \quad (4-11)$$

as in the example in Sec. 3.7 on page 52.

Example

THE VACUUM DIODE

Let us find the potential distribution between the plates of a vacuum diode whose cathode and anode are plane, parallel, and separated by a distance s that is small compared to their linear extent, as in Fig. 4-1.

We assume that the cathode is at zero potential and the anode at a positive potential V_a . The hot cathode emits electrons that accelerate in the direction of the anode. We also assume that the

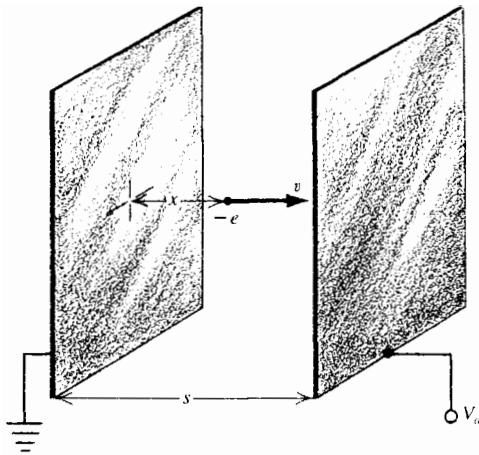


Fig. 4-1. Schematic diagram of a plane-parallel vacuum diode. The cathode, on the left, emits electrons. An electron of charge $-e$ drifts in the direction of the anode at a velocity v . We have shown widely separated electrodes for clarity.

electrons have zero initial velocity and that the current is not limited by the cathode temperature but can be increased at will by increasing V_a .

Since V depends only on x , by hypothesis, Poisson's equation reduces to

$$\frac{d^2V}{dx^2} = -\frac{\rho}{\epsilon_0}, \quad (4-12)$$

where the electron space charge density ρ is negative. Thus d^2V/dx^2 is positive, but we have not yet found ρ .

Now ρ is equal to the current density divided by the electron velocity v , so

$$\frac{d^2V}{dx^2} = \frac{J}{\epsilon_0 v}, \quad (4-13)$$

where J is the magnitude of the current density.

By conservation of energy,

$$\frac{mv^2}{2} = eV, \quad (4-14)$$

where m is the mass of an electron and $-e$ is its electric charge: $e = +1.6 \times 10^{-19}$ coulomb. Then

$$\frac{d^2V}{dx^2} = \frac{J}{\epsilon_0 (2eV/m)^{1/2}}. \quad (4-15)$$

To integrate, we multiply the left-hand side by $2(dV/dx) dx$ and the right-hand side by $2dV$. Then

$$\left(\frac{dV}{dx}\right)^2 = \frac{4J(mV/2e)^{1/2}}{\epsilon_0} + A, \quad (4-16)$$

where A is a constant of integration.

We now find the value of A . At the cathode, $V = 0$ and $A = (dV/dx)^2$. But dV/dx is zero at the cathode for the following reason. If one applies a voltage to the anode when the cathode is cold, dV/dx is positive and equal to V_a/s . If now one heats the cathode, it emits electrons, there is a negative space charge, and dV/dx decreases. As long as dV/dx is positive at the cathode, the emitted electrons accelerate toward the anode and cannot return to the cathode. The current is then limited by the thermionic emission and not by V_a . This is contrary to what we assumed at the beginning. However, if dV/dx was negative, electrons could never leave the cathode, and there would be zero space charge, which is

absurd. So dV/dx can be neither positive nor negative. So it is zero, and A is also zero. (However, see below.) Then

$$\frac{dV}{dx} = 2 \left(\frac{J}{\epsilon_0} \right)^{1/2} \left(\frac{m}{2e} \right)^{1/4} V^{1/4}, \quad (4-17)$$

$$V^{3/4} = 1.5 \left(\frac{J}{\epsilon_0} \right)^{1/2} \left(\frac{m}{2e} \right)^{1/4} x + B. \quad (4-18)$$

The constant of integration B is zero because V is zero at $x = 0$. So

$$V = \left(\frac{9J}{4\epsilon_0} \right)^{2/3} \left(\frac{m}{2e} \right)^{1/3} s^{4/3} \left(\frac{x}{s} \right)^{4/3}. \quad (4-19)$$

When $x = s$, $V = V_a$. Therefore

$$V = V_a \left(\frac{x}{s} \right)^{4/3}. \quad (4-20)$$

Also, disregarding the sign of E ,

$$E = \frac{4}{3} \frac{V_a}{s} \left(\frac{x}{s} \right)^{1/3}, \quad (4-21)$$

$$J = \frac{4\epsilon_0(2e/m)^{1/2}V_a^{3/2}}{9s^2} = 2.335 \times 10^{-6} \frac{V_a^{3/2}}{s^2} \quad \text{amperes/meter}^2, \quad (4-22)$$

$$\rho = \frac{4\epsilon_0 V_a}{9s^2(x/s)^{2/3}}. \quad (4-23)$$

Equation 4-22 is known as the *Child-Langmuir law*. This law is valid only for a plane-parallel diode with negligible edge effects

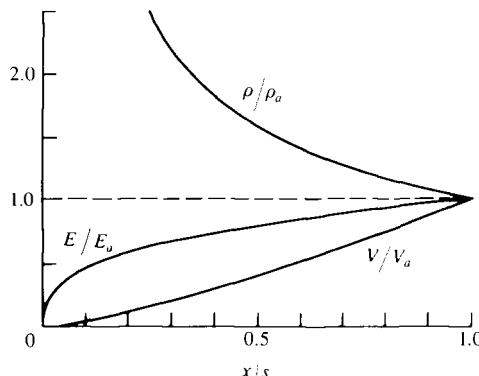


Fig. 4-2. The space charge density ρ , the electric field strength E , and the potential V as a functions of the distance x from the cathode of a plane-parallel diode. The subscript a refers to the value of the variable at the anode.

and for electrons emitted with zero velocity. However, for any geometry, J is proportional to $V_a^{3/2}$.

Figure 4-2 shows V , E , and ρ as functions of x/s .

In an actual diode, the emission velocity of the electrons is finite and there is a potential minimum immediately in front of the cathode. Only electrons with an initial velocity larger than a certain value get past this minimum.

4.2 THE LAW OF CONSERVATION OF ELECTRIC CHARGE

Consider a closed surface of area \mathcal{A} enclosing a volume v . The volume charge density inside is ρ . Charges flow in and out, and the current density at a given point on the surface is \mathbf{J} amperes/meter².

It is a well-established experimental fact that there is never any net creation of electric charge. Then any net outflow depletes the enclosed charge Q : at any given instant,

$$\int_{\mathcal{A}} \mathbf{J} \cdot d\mathcal{A} = -\frac{d}{dt} \int_v \rho dv = -\frac{dQ}{dt}, \quad (4-24)$$

where the vector $d\mathcal{A}$ points *outward*, according to the usual sign convention.

Applying now the divergence theorem on the left, we find that

$$\int_v \nabla \cdot \mathbf{J} dv = -\int_v \frac{\partial \rho}{\partial t} dv. \quad (4-25)$$

We have transferred the time derivative under the integral sign, but then we must use a partial derivative because ρ can be a function of x , y , z , as well as of t .

Now the volume v is of any shape or size. Therefore

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t}. \quad (4-26)$$

Equations 4-24 and 4-26 are, respectively, the integral and differential forms of the *law of conservation of electric charge*.

4.3 CONDUCTION

In good conductors such as copper or aluminum, each atom possesses one or two conduction electrons that are free to roam about in the material.

Table 4-1

Conductor	Conductivity σ , siemens/meter
Aluminium	3.54×10^7
Brass (65.8 Cu, 34.2 Zn)	1.59×10^7
Chromium	3.8×10^7
Copper	5.80×10^7
Gold	4.50×10^7
Graphite	7.1×10^4
Iron	1.0×10^7
Mumetal (75 Ni, 2 Cr, 5 Cu, 18 Fe)	0.16×10^7
Nickel	1.3×10^7
Seawater	~ 5
Silver	6.15×10^7
Tin	0.870×10^7
Zinc	1.86×10^7

Semiconductors may contain two types of mobile charges: conduction electrons and positive holes. A *hole* is a vacancy left by an electron liberated from the valence bond structure in the material. A hole behaves as a free particle of charge $+e$, and it moves through the semiconductor much as an air bubble rises through water.

In most good conductors and semiconductors, the current density \mathbf{J} is proportional to \mathbf{E} :

$$\mathbf{J} = \sigma \mathbf{E}, \quad (4-27)$$

where σ is the *electric conductivity* of the material expressed in siemens per meter, where 1 *siemens*[†] is 1 ampere/volt. This is *Ohm's law* in a more general form. As we shall see later, an electric conductivity can be complex. We shall find a still more general form of Ohm's law in Chap. 23.

Table 4-1 shows the conductivities of some common materials.

Ohm's law does not always apply. For example, in a certain type of ceramic semiconductor, \mathbf{J} is proportional to the fifth power of E . Also some conductors are not isotropic.

4.3.1 Resistance

If Ohm's law applies, the *resistance* between two electrodes fixed to a sample of material is

[†] After Ernst Werner von Siemens (1816–1892). The word therefore takes a terminal *s* in the singular: one siemens. The siemens was formerly called a "mho."

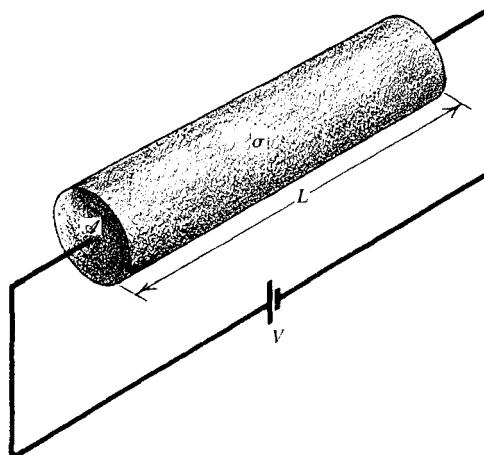
$$R = \frac{V}{I}, \quad (4-28)$$

where V is the potential difference between the two electrodes and I is the current.

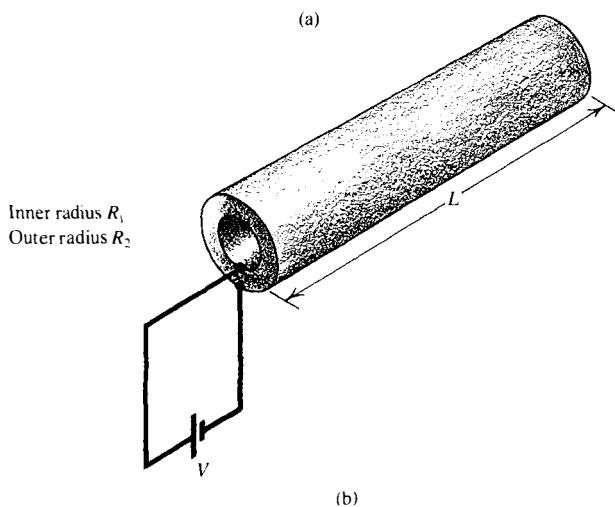
Examples

For a cylinder of cross section \mathcal{A} , length L , uniform conductivity σ , and with electrodes on the ends as in Fig. 4-3a,

$$I = \mathcal{A}J = \mathcal{A}\sigma E = \frac{\sigma \mathcal{A}V}{L}, \quad (4-29)$$



(a)



(b)

Fig. 4-3. (a) Cylinder of weakly conducting material with electrodes at both ends. (b) Tube of weakly conducting material with electrodes on the inner and outer surfaces.

$$R = \frac{L}{\sigma A}. \quad (4-30)$$

The tube of Fig. 4-3b has inner and outer radii R_1 and R_2 , respectively, a length L , and a uniform conductivity σ . There are copper electrodes on the inner and outer cylindrical surfaces. The resistance of a cylindrical element of thickness dr is $dr/\sigma 2\pi r L$. Then

$$R = \int_{R_1}^{R_2} \frac{dr}{\sigma 2\pi r L} = \frac{1}{2\pi\sigma L} \ln \frac{R_2}{R_1}. \quad (4-31)$$

4.3.2 Conduction in a Steady Electric Field

For simplicity, we assume that the charge carriers are conduction electrons.

The detailed motion of an individual conduction electron is exceedingly complex because, every now and then, it collides with an atom and rebounds. The atoms, of course, vibrate about their equilibrium positions, because of thermal agitation, and exchange energy with the conduction electrons.

However, on the average, each electron has a kinetic energy of $\frac{3}{2}kT$, where k is Boltzmann's constant and T is the temperature in kelvins. Thus, at room temperature, the velocity v_{th} associated with thermal agitation is given by

$$\frac{mv_{th}^2}{2} = \frac{3}{2}kT = \frac{3}{2}(1.38 \times 10^{-23} \times 300) \approx 6 \times 10^{-21} \text{ joule}, \quad (4-32)$$

and

$$v_{th} \approx \left(\frac{12 \times 10^{-21}}{9.1 \times 10^{-31}} \right)^{1/2} \approx 10^5 \text{ meters/second}. \quad (4-33)$$

Under the action of a steady electric field, the cloud of conduction electrons drifts at a constant velocity v_d such that

$$\mathbf{J} = \sigma \mathbf{E} = -N e \mathbf{v}_d, \quad (4-34)$$

where \mathbf{v}_d points in the direction opposite to \mathbf{J} and to \mathbf{E} , and N is the number of conduction electrons per cubic meter.

The drift velocity is low. In copper, $N = 8.5 \times 10^{28}$. If a current of 1 ampere flows through a wire having a cross section of 1 millimeter², $\mathbf{J} = 10^6$ and \mathbf{v}_d works out to about 10^{-4} meter/second, or about

300 millimeters/*hour*! Then the drift velocity is smaller than the thermal agitation velocity by *nine* orders of magnitude!

In Eq. 4-34 v_d is small, but Ne is very large. In copper,

$$Ne = 8.5 \times 10^{28} \times 1.6 \times 10^{-19} \approx 10^{10} \text{ coulombs/meter}^3. \quad (4-35)$$

The low drift velocity of conduction electrons is the source of many paradoxes. For example, a radio transmitting antenna is about 75 meters long and operates at about 1 megahertz. How can conduction electrons go from one end to the other and back in 1 microsecond? The answer is that they do not. They drift back and forth by a distance of the order of 1 atomic diameter, and that is enough to generate the required current.

4.3.3 The Mobility \mathcal{M} of Conduction Electrons

The mobility of conduction electrons

$$\mathcal{M} = \frac{|v_d|}{E} = \frac{\sigma}{Ne} \quad (4-36)$$

is, by definition, a positive quantity.[†] It is independent of E in linear conductors. Thus

$$\sigma = Ne\mathcal{M} \quad (4-37)$$

where, as usual, we have taken e to be the *magnitude* of the electronic charge.

If the driving electric field is constant, then the drift velocity is constant. This means that the time-averaged net force on a conduction electron is zero, or that the average braking force due to the collisions just cancels the $-eE$ force exerted by the field.

What is the magnitude of this braking force? It is

$$-(-eE) = eE = \frac{e |v_d|}{\mathcal{M}} = -\frac{e}{\mathcal{M}} v_d, \quad (4-38)$$

from the definition of the mobility \mathcal{M} . The braking force and v_d point in opposite directions.

This situation is analogous to that of a body falling through water; after a while, the viscous force exactly cancels the gravitational force, and the magnitudes of the two opposing forces are equal.

[†] Some authors assign to the mobility the sign of the charge carrier.

The quantities N , \mathcal{M} , and σ for good conductors (gc) and for semiconductors (sc) are related as follows:

$$N_{\text{gc}} \gg N_{\text{sc}}, \quad \sigma_{\text{gc}} \gg \sigma_{\text{sc}}, \quad \mathcal{M}_{\text{gc}} \ll \mathcal{M}_{\text{sc}}. \quad (4-39)$$

4.3.4 Conduction in an Alternating Electric Field

If we disregard thermal agitation, there are thus two forces acting on a conduction electron: the driving force eE and the braking force of Eq. 4-38.

In an alternating electric field these two forces are unequal, and the equation of motion is

$$m^* \frac{dv_d}{dt} = -eE_m \exp j\omega t - \frac{e}{\mathcal{M}} v_d, \quad (4-40)$$

where m^* is the *effective mass*. This quantity takes collisions into account. As a rule, m^* is smaller than the mass of an isolated electron.

In silicon $m^* = 0.97m$, but in gallium arsenide (GaAs) the effective mass is only $0.07m$. Electron velocities in solid-state devices are of the order of 10^5 meters/second in silicon and about 4 times larger in GaAs.

Replacing the time derivative by $j\omega$ and simplifying, we find that

$$v_d = -\frac{\mathcal{M}}{1 + j\omega m^* \mathcal{M}/e} E_m \exp j\omega t. \quad (4-41)$$

But $J = \sigma E = -Nev_d$. It follows that

$$\sigma = \frac{Ne\mathcal{M}}{1 + j\omega m^* \mathcal{M}/e}. \quad (4-42)$$

With $\omega = 0$ we revert to Eq. 4-37. This relation does *not* apply at frequencies of the order of 1 gigahertz ($f = 10^9$) or higher, where atomic phenomena become prominent.

For copper at ambient temperature,

$$\frac{\omega m^* \mathcal{M}}{e} \approx \frac{\omega m \sigma}{Ne^2} \quad (4-43)$$

$$\approx \frac{2\pi f \times 9.1 \times 10^{-31} \times 5.8 \times 10^7}{8.5 \times 10^{28} \times (1.6 \times 10^{-19})^2} \quad (4-44)$$

$$\approx 1.5 \times 10^{-13} f. \quad (4-45)$$

The imaginary term in Eq. 4-42 is negligible for $f \ll 7 \times 10^{12}$. Up to about 1 gigahertz the cloud of conduction electrons moves in phase with E , and $\sigma = NeM$.

Decreasing the temperature increases the mean free path of the carriers, which increases the mobility M . At very low temperatures the conductivity of pure metals is complex:

$$\sigma = -j \frac{Ne^2}{m^* \omega}. \quad (4-46)$$

4.3.5 The Volume Charge Density ρ in a Conductor

(1) Assume steady-state conditions and a homogeneous conductor. Then $\partial\rho/\partial t = 0$ and, from Sec. 4.2, $\nabla \cdot J = 0$. If J is the conduction current density in a homogeneous conductor that satisfies Ohm's law $J = \sigma E$, then

$$\nabla \cdot J = \nabla \cdot \sigma E = \sigma \nabla \cdot E = 0, \quad \nabla \cdot E = 0. \quad (4-47)$$

But the divergence of E is proportional to the volume charge density ρ , from Sec. 3.7. Thus, under steady-state conditions and in homogeneous conductors (σ independent of the coordinates), ρ is zero.

As a rule, the *surface* charge density on a conducting body carrying a current is not zero.

(2) Now suppose that one injects charge into a piece of copper by bombarding it with electrons. What happens to the charge density? In that case, from Sec. 4.2,

$$\nabla \cdot J = -\frac{\partial \rho}{\partial t}. \quad (4-48)$$

But, from Sec. 3.7,

$$\nabla \cdot J = \sigma \nabla \cdot E = \frac{\sigma \rho}{\epsilon_r \epsilon_0}, \quad (4-49)$$

where ϵ_r is the relative permittivity of the material (Sec. 9.9). Thus

$$\frac{\partial \rho}{\partial t} = -\frac{\sigma \rho}{\epsilon_r \epsilon_0}, \quad \rho = \rho_0 \exp\left(-\frac{\sigma t}{\epsilon_r \epsilon_0}\right), \quad (4-50)$$

and ρ decreases exponentially with time.

The relative permittivity ϵ_r of a good conductor is not measurable because conduction completely overshadows polarization. One may presume that ϵ_r is of the order of 3, as in common dielectrics.

The inverse of the coefficient of t in the above exponent is the *relaxation time*.

We have neglected the fact that σ is frequency-dependent and is thus itself a function of the relaxation time. Relaxation times in good conductors are, in fact, short; and ρ may be set equal to zero, in practice. For example, the relaxation time for copper at room temperature is about 4×10^{-14} second, instead of $\approx 10^{-19}$ second according to the above calculation.

(3) In a homogeneous conductor carrying an alternating current, ρ is zero because Eq. 4-47 applies.

(4) In a nonhomogeneous conductor carrying a current, ρ is not zero. For example, under steady-state conditions,

$$\nabla \cdot \mathbf{J} = \nabla \cdot (\sigma \mathbf{E}) = (\nabla \sigma) \cdot \mathbf{E} + \sigma \nabla \cdot \mathbf{E} = 0 \quad (4-51)$$

and

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_r \epsilon_0} = -\frac{(\nabla \sigma) \cdot \mathbf{E}}{\sigma}. \quad (4-52)$$

(5) If there are magnetic forces on the charge carriers, then $\mathbf{J} = \sigma \mathbf{E}$ does not apply and there can exist a volume charge density. See Sec. 22.4.1.

4.3.6 The Joule Effect

In the absence of an electric field, the cloud of conduction electrons remains in thermal equilibrium with the lattice of the host conductor. Upon application of an electric field, the electrons gain kinetic energy between collisions, and they share this extra energy with the lattice. The conductor thus heats up. This is the *Joule effect*.

What is the kinetic energy gained by the conduction electrons? Consider a cube of the conductor, with side a . Apply a voltage V between opposite faces. The current is I . Then the kinetic energy gained is VI , and the power dissipated as heat per cubic meter is

$$P' = \frac{VI}{a^3} = \left(\frac{V}{a}\right)\left(\frac{I}{a^2}\right) = EJ \quad (4-53)$$

$$= \sigma E^2 = \frac{J^2}{\sigma} \quad \text{watts/meter}^3. \quad (4-54)$$

If E and J are sinusoidal functions of the time,

$$P'_{\text{av}} = E_{\text{rms}} J_{\text{rms}} = \sigma E_{\text{rms}}^2 = \frac{J_{\text{rms}}^2}{\sigma}. \quad (4-55)$$

4.4 ISOLATED CONDUCTORS IN STATIC FIELDS

If one charges an isolated homogeneous conductor, the conduction electrons move about until they have reached their equilibrium positions and then, inside the conductor, there is zero E .

It follows that (1) all points inside the conductor are at the same potential; (2) the volume charge density is zero, from Eq. 4-47; (3) any net static charge resides on the surface of the conductor; (4) E is normal at the surface of the conductor, for otherwise charges would flow along the surface; (5) just outside the surface, $E = \sigma_{\text{ch}}/\epsilon_0$, where σ_{ch} is the surface charge density, from Gauss's law.

Note the paradox: one can express E at the surface of a conductor in terms of the *local* surface charge density alone, in spite of the fact that E depends on the magnitudes and positions of *all* the charges, whether they reside on the conductor or elsewhere.

What if the conductor is not homogeneous? For example, one might have a copper wire pressed onto a gold-plated terminal. Then conduction electrons drift across the interface and establish a *contact potential*, usually of a fraction of a volt. The magnitude and sign of the contact potential depend on the nature of the materials.

Example

HOLLOW CONDUCTOR ENCLOSING A CHARGED BODY

Figure 4-4 shows a cross section of a hollow conductor with a net electric charge Q within the cavity. The Gaussian surface lying within the conductor in a zero E encloses a zero net charge, because of Gauss's law. Then the surface charge on the *inside* surface of the conductor is $-Q$.

If the conductor carries a zero net charge, then the total charge on the *outside* surface is Q .

The surface charge density at a given point on the *outside* surface of the conductor is independent of the distribution of Q in the cavity. It is the same as if the conductor were solid and carried a net charge Q .

Inversely, the field inside the cavity is independent of the field outside the conductor. The conductor then acts as an *electrostatic shield*.

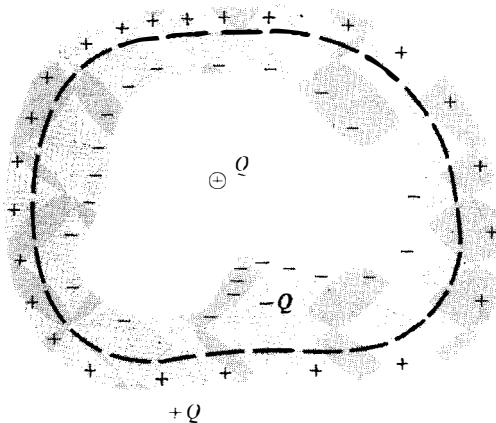


Fig. 4-4. Section through a hollow conductor enclosing a body that carries a net charge Q . The dashed line is a section through a Gaussian surface lying entirely in the conducting material, where $E = 0$.

4.5 SUMMARY

Poisson's equation follows from the differential form of Gauss's law and from the relation $\mathbf{E} = -\nabla V$:

$$\nabla^2 V = -\frac{\rho}{\epsilon_0}. \quad (4-1)$$

Setting the volume charge density ρ equal to zero yields *Laplace's equation*:

$$\nabla^2 V = 0. \quad (4-2)$$

The *law of conservation of electric charge* states that, whatever the circumstances, the net electric charge of a closed system is constant. Mathematically,

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t}, \quad (4-26)$$

where \mathbf{J} and ρ are, respectively, the current and charge densities.

Most conductors obey *Ohm's law*:

$$\mathbf{J} = \sigma \mathbf{E}, \quad (4-27)$$

where σ is the *conductivity*, expressed in siemens per meter.

In a time-independent electric field,

$$\mathbf{J} = \sigma \mathbf{E} = -Ne\mathbf{v}_d \quad (4-34)$$

where N is the number of charge carriers per cubic meter, $-e$ is the charge on one of them, and \mathbf{v}_d is their *drift velocity*. The *mobility* is defined by the equation

$$\mathcal{M} = \frac{|\mathbf{v}_d|}{E} = \frac{\sigma}{Ne} \quad (4-36)$$

and

$$\sigma = Ne\mathcal{M}. \quad (4-37)$$

In an alternating electric field the conductivity σ is complex:

$$\sigma = \frac{Ne\mathcal{M}}{1 + j\omega m^* \mathcal{M}/e}, \quad (4-42)$$

where m^* is the *effective mass* of a charge carrier.

If the current density is J and the electric field strength is E , then the time-averaged power dissipated per cubic meter in the form of heat is

$$P'_{av} = E_{rms} J_{rms} = \sigma E_{rms}^2 = \frac{J_{rms}^2}{\sigma} \quad \text{watts/meter}^3. \quad (4-55)$$

This is the *Joule effect*.

Under static conditions there exists a net electric charge solely at the surface of a conductor, and \mathbf{E} is zero inside. Just outside, \mathbf{E} is normal to the surface, and its magnitude is σ_{ch}/ϵ_0 , where σ_{ch} is the surface charge density.

PROBLEMS

4-1. (4.3) The conduction electron density in copper

The object of this problem is to illustrate the enormous magnitude of the electric charge densities in matter.

We take the example of the conduction electrons in copper. A copper atom contains 29 electrons, one of which is a conduction electron. Copper has an atomic weight of 64 and a density of 8.9×10^3 kilograms/meter³. Suppose that you have two copper spheres, each one having a volume of 1 centimeter³. The spheres are depleted of their conduction electrons and separated by a distance of 100 millimeters.

Calculate the force of repulsion. Show that this force is equal to about 0.5% of the force of attraction between the sun and the earth. See the table of physical constants on the page facing the back cover.

4-2. (4.3) The drift velocity of conduction electrons

Copper has an atomic weight of 64 and a density of 8.9×10^3 kilograms/meter³.

(a) Calculate the number of atoms per cubic meter and the approximate diameter of an atom.

(b) Calculate the charge λ carried by the conduction electrons in 1 meter of copper wire 1 millimeter in diameter. There is one conduction electron per atom.

(c) Calculate the drift velocity of the conduction electrons in meters per hour when the wire carries a current of 1 ampere.

4-3. (4.3) Refraction of lines of E at the interface between media of different conductivities

We shall see in Sec. 10.2.3 that the tangential component of E is continuous at the interface between two media.

Show that, at the boundary between two media of conductivities σ_1 and σ_2 , a line of E , or a line of J , is “refracted” in such a way that $\tan \theta_1/\sigma_1 = \tan \theta_2/\sigma_2$, where θ_1 and θ_2 are the angles formed by a line of E with the normal to the interface.

4-4. (4.3) The surface charge density at the interface between media of different conductivities.

A current of density J flows in the direction normal to the interface between two media of conductivities σ_{co1} and σ_{co2} . The current flows from medium 1 to medium 2.

Show that the surface charge density σ_{ch} is $\epsilon_r \epsilon_0 J (1/\sigma_{co2} - 1/\sigma_{co1})$. Assume that ϵ_r has the same value on both sides. If the current is not normal to the interface, then the above J is the normal component of the current density.

4-5. (4.3) Conduction in a nonhomogeneous medium

In a nonhomogeneous medium, the conductivity σ is a function of the coordinates.

Show that, under static conditions, or when $E = -\nabla V$, and if σ is nowhere equal to zero,

$$\nabla^2 V + \nabla V \cdot \nabla \tau = 0,$$

where $\tau = \ln \sigma$.

4-6. (4.3) Geophysical prospecting by the resistivity method

One can locate resistivity anomalies in the ground as in Fig. 4-5. The current I flowing between electrodes C_1 and C_2 establishes an electric field in the ground, and one measures the voltage V between a pair of electrodes P_1 and P_2 maintained at a fixed spacing b . With $b \ll a$, V/b is equal to E at the position x . Anomalies in ground conductivity show up in the curve of E as a function of x .

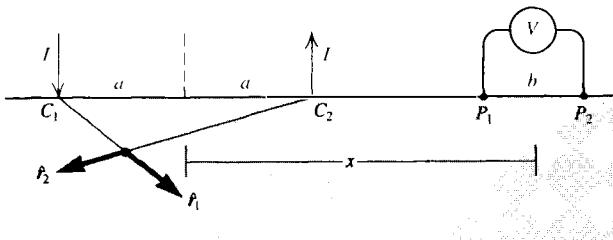


Fig. 4-5.

Show that if the substrate conductivity is uniform and equal to σ , then

$$\frac{V}{b} = -\frac{2axI}{\pi\sigma(x^2 - a^2)^2}.$$

The electrodes are of finite size. However, you can perform the calculation on the assumption that they are infinitely small, disregarding the fact that E and J would then be infinite at their surfaces.

You can use the principle of superposition as follows. The current in the ground is the sum of a radial distribution emanating from C_1 plus another radial distribution converging on C_2 . Thus, at a point r_1, r_2 ,

$$\mathbf{E} = \frac{I\hat{\mathbf{r}}_1}{2\pi\sigma r_1^2} - \frac{I\hat{\mathbf{r}}_2}{2\pi\sigma r_2^2}.$$

4-7. (4.3.1) The resistance of a spherical shell

A spherical shell of uniform conductivity has inner and outer radii R_1 and R_2 , respectively. It has copper electrodes plated on the inner and outer surfaces.

Show that the resistance is $(1/R_1 - 1/R_2)/4\pi\sigma$.

4-8. (4.3.1) Resistive film

A square film of Nichrome, an alloy of nickel and chromium, has copper electrodes deposited on two opposite edges.

Show that the resistance between the electrodes depends only on the thickness of the film and on its conductivity, as long as the film is square. This *surface resistance* is expressed in *ohms per square*.

4-9. (4.3.1) A theorem on the resistance of a plate

A rectangular plate $ABCD$ has a thickness s and a conductivity σ . With conducting electrodes on edges AB and CD , the resistance is R_1 . With electrodes on BC and DA , the resistance is R_2 .

Show that $R_1 R_2 = 1/(\sigma^2 s^2)$.

This equation also applies to any region bounded by equipotentials and lines of current flow. We shall use this theorem in Prob. 33.5. It was first proved by Isukada.[†]

[†] D. J. Epstein, Proc. IEEE, vol. 56, p. 198 (1968).

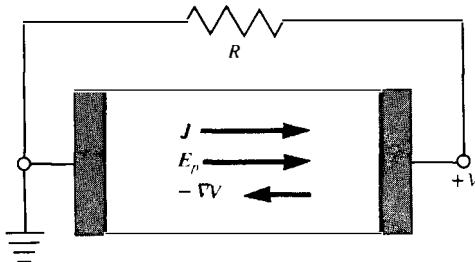


Fig. 4-6.

4-10. (4.3.1) E and J inside a battery

A battery feeds a resistance R as in Fig. 4-6. The battery acts as a pump, forcing conduction electrons toward the negative electrode. The battery is cylindrical, of length s and cross-sectional area \mathcal{A} , with electrodes at each end. Then $|\nabla V| = V/s$ and, inside the battery, $J = \sigma(E_p - |\nabla V|)$, where E_p is the “pumping field.”

Find the output voltage as a function of the current. Set $R = s/\sigma\mathcal{A}$ as the output resistance of the battery.

4-11. (4.3.4) Mobility and electron drift

A simple model for the drift of a conduction electron is the following. The electron describes a ballistic trajectory for a while, under the action of the ambient electric field, and then the electron suffers an impact. Its velocity just after the impact is unrelated to its velocity before the impact, and we set it equal to zero. The electron then starts out on another ballistic trajectory, and the process repeats itself. Let the mean time between the collisions be Δt and the effective mass (Sec. 4.3.4) be m^* .

Find the mobility in terms of Δt and m^* .

4-12. (4.3.4) Conduction by holes

We derived Eq. 4-40 on the assumption that the charge carriers are electrons. Suppose the carriers are holes. Then the charge changes sign, and both terms on the right are positive. If that is so, the complementary function, which one obtains on disregarding the forcing term $eE_m j\omega t$, is

$$v_d = v_{d0} \exp\left(\frac{e}{m^* \mathcal{M}} t\right).$$

Then v_d increases exponentially with time, which is absurd.

Show that, if the charge carriers are holes, then

$$m^* \frac{dv_d}{dt} = +eE_m \exp j\omega t - \frac{e}{\mathcal{M}} v_d,$$

and that

$$v_d = +\frac{\mathcal{M}}{1 + j\omega m^* \mathcal{M}/e} E_m \exp j\omega t, \quad \sigma = \frac{Ne\mathcal{M}}{1 + j\omega m^* \mathcal{M}/e}.$$

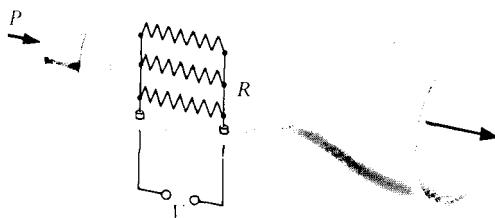


Fig. 4-7.

4-13. (4.3.6) The resistojet

Figure 4-7 shows the principle of operation of a *resistojet* used as a thruster for correcting the trajectory or the attitude of a satellite.

Assuming complete conversion of the electric energy to kinetic energy, calculate the thrust for a power input of 3 kilowatts and a flow of 0.6 gram of hydrogen per second.

CHAPTER 5

ELECTRIC FIELDS III *Electric Multipoles*

5.1 THE ELECTRIC DIPOLE	84	
5.2 THE LINEAR ELECTRIC QUADRUPOLE	87	
*5.3 ELECTRIC MULTipoLES	88	
*5.4 THE ELECTRIC FIELD OUTSIDE A CHARGE DISTRIBUTION, EXPANDED IN TERMS OF MULTipoLES	89	
*5.4.1 THE VALUE OF V . THE LEGENDRE POLYNOMIALS	90	
*5.4.2 THE MONOPOLE TERM	93	
*5.4.3 THE DIPOLE TERM	93	
*5.4.4 THE QUADRUPOLE TERM	94	
Example: THE FIELD OF A SET OF SIX POINT CHARGES SET SYMMETRICALLY ABOUT THE ORIGIN		95
5.5 SUMMARY	97	
PROBLEMS	98	

Electric multipoles are sets of point charges possessing certain symmetries. Their interest lies in the fact that real charged objects, such as antennas and atomic nuclei, possess electric fields that may be expressed as sums of multipole fields.

Aside from the monopole, which is a single point charge, the most useful type of multipole is the dipole, which consists of two charges of equal magnitudes and opposite signs, some distance apart. Most molecules act like small dipoles. Also many antennas radiate like oscillating electric dipoles.

5.1 THE ELECTRIC DIPOLE

The *electric dipole* is a common type of charge distribution. We return to it later in this chapter and in Chaps. 37 and 38.

The electric dipole consists of two charges, one positive and one negative, of the same magnitude, and separated by a distance s . We find V and \mathbf{E} at a point P situated at a distance $r \gg s$, as in Fig. 5-1. At P ,

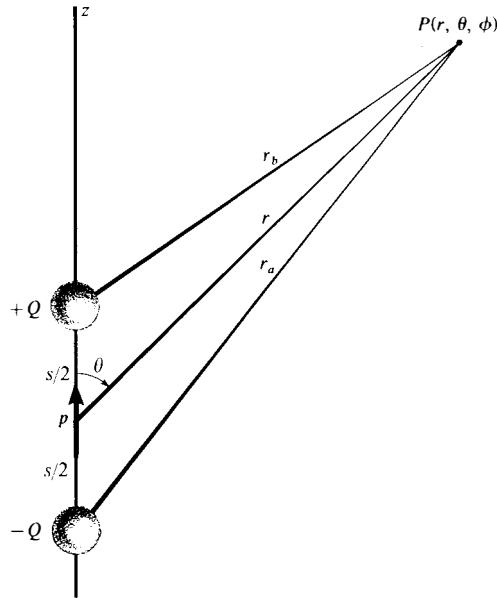


Fig. 5-1. Two charges $+Q$ and $-Q$ separated by a distance s and forming a dipole. The dipole moment is \mathbf{p} . We calculate the potential at point P by summing the potentials of the two charges.

$$V = \frac{Q}{4\pi\epsilon_0} \left(\frac{1}{r_b} - \frac{1}{r_a} \right), \quad (5-1)$$

where

$$r_a^2 = r^2 + \left(\frac{s}{2}\right)^2 + rs \cos \theta. \quad (5-2)$$

We now divide both sides by r^2 and take the inverse:

$$\frac{r}{r_a} = \left[1 + \left(\frac{s}{2r}\right)^2 + \frac{s}{r} \cos \theta \right]^{-1/2} \quad (5-3)$$

$$= 1 - \frac{1}{2} \left(\frac{s^2}{4r^2} + \frac{s}{r} \cos \theta \right) + \frac{3}{8} \left(\frac{s^2}{4r^2} + \frac{s}{r} \cos \theta \right)^2 - \dots \quad (5-4)$$

If we neglect terms of order $(s/r)^3$ and higher, then

$$\frac{r}{r_a} = 1 - \frac{s}{2r} \cos \theta + \frac{s^2}{4r^2} \frac{3 \cos^2 \theta - 1}{2}. \quad (5-5)$$

Similarly,

$$\frac{r}{r_b} = 1 + \frac{s}{2r} \cos \theta + \frac{s^2}{4r^2} \frac{3 \cos^2 \theta - 1}{2} \quad (5-6)$$

and

$$V = \frac{Qs}{4\pi\epsilon_0 r^2} \cos \theta \quad (r^3 \gg s^3). \quad (5-7)$$

Note that the potential in the field of a dipole falls off as $1/r^2$, whereas the potential of a single point charge varies only as $1/r$. This comes from the fact that the charges of a dipole appear close together for an observer some distance away, so their fields cancel more and more as the distance r increases.

The *dipole moment* $\mathbf{p} = Q\mathbf{s}$ is a vector that is directed from the negative to the positive charge. Then

$$V = \frac{\mathbf{p} \cdot \hat{\mathbf{r}}}{4\pi\epsilon_0 r^2}. \quad (5-8)$$

We can now find the electric field strength \mathbf{E} . In spherical coordinates,

$$E_r = -\frac{\partial V}{\partial r} = \frac{2p \cos \theta}{4\pi\epsilon_0 r^3}, \quad (5-9)$$

$$E_\theta = -\frac{1}{r} \frac{\partial V}{\partial \theta} = \frac{p \sin \theta}{4\pi\epsilon_0 r^3}, \quad (5-10)$$

$$E_\phi = -\frac{1}{r \sin \theta} \frac{\partial V}{\partial \phi} = 0, \quad (5-11)$$

$$\mathbf{E} = \frac{\mathbf{p}}{4\pi\epsilon_0 r^3} (2 \cos \theta \hat{\mathbf{r}} + \sin \theta \hat{\boldsymbol{\theta}}). \quad (5-12)$$

Thus E falls off as the *cube* of the distance r .

Figure 5-2 shows lines of \mathbf{E} and *equipotential lines* for an electric dipole. Rotating equipotential lines about the vertical axis generates *equipotential surfaces*.

More generally, the dipole moment of a charge distribution is

$$\mathbf{p} = \int_{v'} \mathbf{r}' \rho dv' = Q \frac{\int_{v'} \mathbf{r}' \rho dv'}{\int_{v'} \rho dv'} = Q \bar{\mathbf{r}}', \quad (5-13)$$

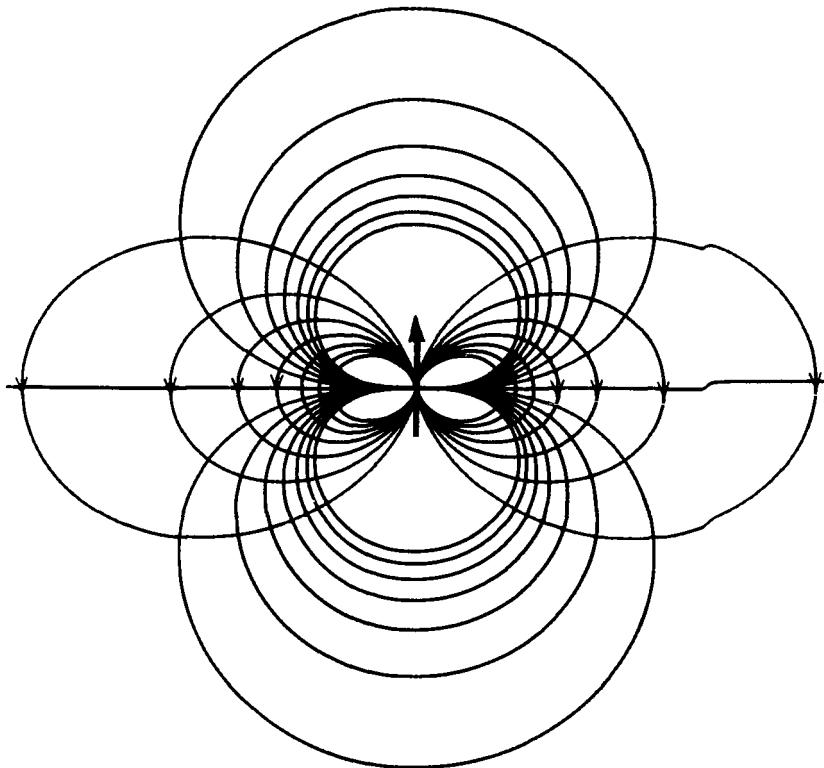


Fig. 5-2. Lines of \mathbf{E} , shown with arrows, and equipotential lines for the electric dipole of Fig. 5-1. In the central region the lines become too close together to be shown. The central arrow is the vector \mathbf{p} .

where Q is the net total charge occupying the volume v' , and $\bar{\mathbf{r}'}$ defines the position of the *center of charge*, by analogy with the center \mathbf{r} of mass in mechanics.

If $Q = 0$, then $\bar{\mathbf{r}'} \rightarrow \infty$ and $Q\bar{\mathbf{r}'}$ is indeterminate. However, the integral of $\mathbf{r}'\rho d\mathbf{v}'$ still provides the correct value of \mathbf{p} . If $Q = 0$, the dipole moment is independent of the choice of the origin (Prob. 5-1).

If $Q \neq 0$, the dipole moment of the distribution is zero when the origin is at the center of charge, for then $\bar{\mathbf{r}'} = 0$.

5.2 THE LINEAR ELECTRIC QUADRUPOLE

The *linear electric quadrupole* is a set of three charges, as in Fig. 5-3. The separation s is again small compared to the distance r to the point P . At P ,

$$V = \frac{1}{4\pi\epsilon_0} \left(\frac{Q}{r_a} - \frac{2Q}{r} + \frac{Q}{r_b} \right) = \frac{Q}{4\pi\epsilon_0 r} \left(\frac{r}{r_a} + \frac{r}{r_b} - 2 \right). \quad (5-14)$$

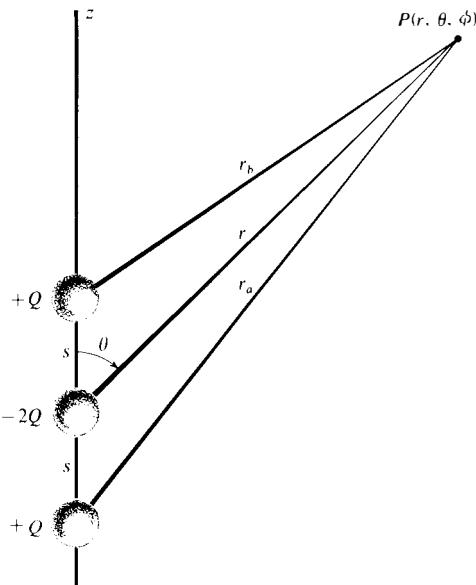


Fig. 5-3. Charges $+Q$, $-2Q$, $+Q$ forming a linear quadrupole.

We can expand the ratios r/r_a and r/r_b as previously, except that s now replaces $s/2$. Thus, if we neglect terms of order $(s/r)^3$ and higher,

$$\frac{r}{r_a} = 1 - \frac{s}{r} \cos \theta + \frac{s^2}{r^2} \frac{(3 \cos^2 \theta - 1)}{2}, \quad (5-15)$$

$$\frac{r}{r_b} = 1 + \frac{s}{r} \cos \theta + \frac{s^2}{r^2} \frac{(3 \cos^2 \theta - 1)}{2}, \quad (5-16)$$

and

$$V = \frac{2Qs^2}{4\pi\epsilon_0 r^3} \frac{(3 \cos^2 \theta - 1)}{2} \quad (r^3 \gg s^3). \quad (5-17)$$

The potential V of a linear electric quadrupole varies as $1/r^3$, whereas E , calculated as for the dipole, varies as $1/r^4$. The fields of the three charges cancel almost completely for $r \gg s$.

*5.3 ELECTRIC MULTipoles

It is possible to extend the concept of dipole and quadrupole to larger numbers of positive and negative charges. Such charge arrangements are

* Starred sections can be omitted without losing continuity. However, the Legendre polynomials of Sec. 5.4.1 are required for the starred Sec. 12.1.

known as *multipoles*. A single point charge is a *monopole*. A *dipole* is obtained by displacing a monopole through a small distance s_1 and replacing the original monopole by another of the same magnitude but of opposite sign. Likewise, a *quadrupole* is obtained by displacing a dipole by a small distance s_2 and then replacing the original dipole by one of equal magnitude but of opposite sign. For the linear quadrupole, $s_2 = s_1$.

The multipole concept can extend indefinitely. For example, the quadrupole can be displaced by a small distance s_3 , and the original quadrupole replaced by one in which the signs of all the charges have been changed. This gives an *octupole*. A 2^l -pole requires l displacements s_1, s_2, \dots, s_l .

We have seen that the dipole potential varies as $1/r^2$ and that the quadrupole potential varies as $1/r^3$. For the 2^l -pole, V varies as $1/r^{l+1}$ and E as $1/r^{l+2}$.

We have calculated the potential V in the field of a dipole and of a quadrupole in the following way. We found the sum of the potentials of the individual charges, expanded the sum as a power series, and then truncated. This approach is straightforward, and it has the advantage of showing exactly what approximations are involved. Problems 5-5 and 5-6 explore a different method that is more elegant, but that does not reveal the exact nature of the approximations.

*5.4 THE ELECTRIC FIELD OUTSIDE A CHARGE DISTRIBUTION, EXPANDED IN TERMS OF MULTipoles

A charge distribution of density $\rho(x', y', z')$ occupies a volume v' and extends to a maximum distance r'_{\max} from an arbitrary origin O , as in Fig. 5-4. We select O either within the volume or close to it.

We shall see that the potential V at a point P outside the charge distribution such that $r > r'_{\max}$ is the same as (1) the potential V_1 of a point charge, or monopole, equal to the net charge of the distribution, plus (2) the potential V_2 of a point dipole with a dipole moment equal to that of the charge distribution, plus (3) the potential V_3 of a point quadrupole with a quadrupole moment equal to that of the charge distribution, and so on, the monopole, dipole, quadrupole, etc., being all located at the arbitrary origin.

Similarly, the E at point P is the sum of the E 's of the above monopole, dipole, quadrupole, etc.

If $Q = 0$, then V_2 is independent of the choice of origin. More generally, V_l is independent of the choice of origin if all the multipole moments up to the 2^{l-1} -pole are zero.

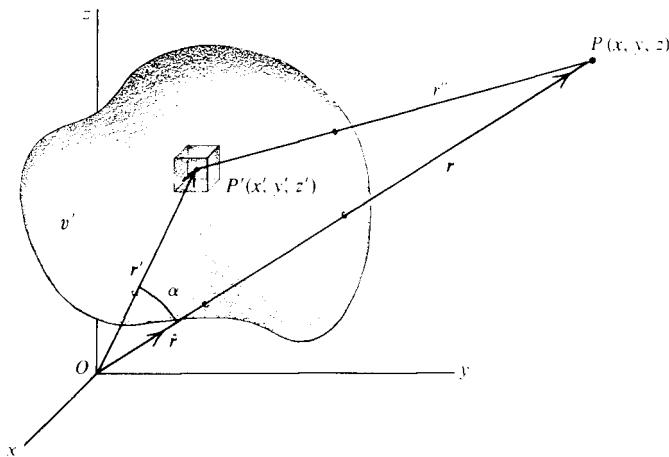


Fig. 5-4. Arbitrary charge distribution enclosed within a volume v' . The potential at P is the same as if one had, at the origin, a monopole, plus a dipole, plus a quadrupole, plus an octupole, etc. However, r must be larger than the maximum value of r' .

Since the V of a monopole decreases as $1/r$, that of a dipole as $1/r^2$, that of a quadrupole as $1/r^3$, etc., then a long distance away, where $r \gg r'_{\max}$, the field of any charge distribution is simple. It is that of a point charge at the origin and $V \approx V_1$. Closer in, $V \approx V_1 + V_2$. Still closer in, V_3 becomes discernible, then V_4 , etc., and the field becomes more and more complex.

*5.4.1 The Value of V . The Legendre Polynomials

We wish to find V at some point P such that $r > r'_{\max}$. From Fig. 5-4, this is

$$V = \int_{v'} \frac{\rho dv'}{4\pi\epsilon_0 r''}, \quad (5-18)$$

where

$$r'' = |\mathbf{r} - \mathbf{r}'| = [(x - x')^2 + (y - y')^2 + (z - z')^2]^{1/2}. \quad (5-19)$$

The point $P(x, y, z)$ is fixed. Thus r'' is a function of x', y', z' , and we can expand $1/r''$ as a Taylor series near the origin.

Let

$$w(u) = \frac{1}{|u\mathbf{r}' - \mathbf{r}|}. \quad (5-20)$$

The factor u is dimensionless. We require $w(1)$, which is $1/r''$:

$$w(1) = \frac{1}{r''} = w(0) + \left(\frac{\partial w}{\partial u} \right)_{u=0} + \frac{1}{2!} \left(\frac{\partial^2 w}{\partial u^2} \right)_{u=0} + \dots, \quad (5-21)$$

where $w(0) = 1/r$. To calculate the derivatives on the right-hand side, we require the partial derivative of $|u\mathbf{r}' - \mathbf{r}|$ with respect to u :

$$\frac{\partial}{\partial u} |u\mathbf{r}' - \mathbf{r}| = \frac{\partial}{\partial u} [(ux' - x)^2 + (uy' - y)^2 + (uz' - z)^2]^{1/2} \quad (5-22)$$

$$= \frac{1}{2|u\mathbf{r}' - \mathbf{r}|} (2ux'^2 - 2x'x + 2uy'^2 - 2y'y + 2uz'^2 - 2z'z) \quad (5-23)$$

$$= \frac{u\mathbf{r}'^2 - \mathbf{r}' \cdot \mathbf{r}}{|u\mathbf{r}' - \mathbf{r}|} = \frac{(u\mathbf{r}' - \mathbf{r}) \cdot \mathbf{r}'}{|u\mathbf{r}' - \mathbf{r}|}. \quad (5-24)$$

Thus

$$\frac{\partial w}{\partial u} = -\frac{1}{|u\mathbf{r}' - \mathbf{r}|^2} \frac{(u\mathbf{r}' - \mathbf{r}) \cdot \mathbf{r}'}{|u\mathbf{r}' - \mathbf{r}|} = -\frac{(u\mathbf{r}' - \mathbf{r}) \cdot \mathbf{r}'}{|u\mathbf{r}' - \mathbf{r}|^3} \quad (5-25)$$

and

$$\left(\frac{\partial w}{\partial u} \right)_{u=0} = \frac{\mathbf{r} \cdot \mathbf{r}'}{r^3} = \frac{\hat{\mathbf{r}} \cdot \mathbf{r}'}{r^2} = \frac{r' \cos \alpha}{r^2}, \quad (5-26)$$

where $\hat{\mathbf{r}}$ and α are as in Fig. 5-4. Also,

$$\frac{\partial^2 w}{\partial u^2} = \frac{3[(u\mathbf{r}' - \mathbf{r}) \cdot \mathbf{r}']^2}{|u\mathbf{r}' - \mathbf{r}|^5} - \frac{r'^2}{|u\mathbf{r}' - \mathbf{r}|^3} \quad (5-27)$$

and

$$\left(\frac{\partial^2 w}{\partial u^2} \right)_{u=0} = \frac{3(\mathbf{r}' \cdot \mathbf{r})^2}{r^5} - \frac{r'^2}{r^3} = \frac{r'^2(3 \cos^2 \alpha - 1)}{r^3}. \quad (5-28)$$

In general,

$$\left(\frac{\partial^n w}{\partial u^n} \right)_{u=0} = \frac{r'^n}{r^{n+1}} n! P_n(\cos \alpha), \quad (5-29)$$

where

$$P_n(\cos \alpha) = \frac{1}{2^n n!} \frac{d^n}{d(\cos \alpha)^n} (\cos^2 \alpha - 1)^n \quad (5-30)$$

Table 5-1 Legendre polynomials

<i>n</i>	$P_n(\cos \alpha)$
0	1
1	$\cos \alpha$
2	$\frac{3 \cos^2 \alpha - 1}{2}$
3	$\frac{5 \cos^3 \alpha - 3 \cos \alpha}{2}$
4	$\frac{35 \cos^4 \alpha - 30 \cos^2 \alpha + 3}{8}$
5	$\frac{63 \cos^5 \alpha - 70 \cos^3 \alpha + 15 \cos \alpha}{8}$

is a *Legendre polynomial*. Table 5-1 gives the first five, while Fig. 5-5 shows the first four as functions of the angle α . For any α ,

$$|P_n(\cos \alpha)| \leq 1. \quad (5-31)$$

Thus from Eq. 5-21,

$$\frac{1}{r''} = \frac{1}{r} + \frac{r' \cos \alpha}{r^2} + \frac{r'^2(3 \cos^2 \alpha - 1)}{2r^3} + \dots = \sum_{n=0}^{\infty} \frac{1}{r^{n+1}} r''^n P_n(\cos \alpha). \quad (5-32)$$

Since $|P_n(\cos \alpha)| \leq 1$ as above, this series converges for $r'_{\max} < r$.

Substituting in Eq. 5-18, we finally have that

$$\begin{aligned} V = & \frac{1}{4\pi\epsilon_0 r} \int_{v'} \rho dv' + \frac{1}{4\pi\epsilon_0 r^2} \int_{v'} r' \cos \alpha \rho dv' \\ & + \frac{1}{4\pi\epsilon_0 r^3} \int_{v'} r'^2 \frac{3 \cos^2 \alpha - 1}{2} \rho dv' \\ & + \frac{1}{4\pi\epsilon_0 r^4} \int_{v'} r'^3 \frac{5 \cos^3 \alpha - 3 \cos \alpha}{2} \rho dv' \\ & + \frac{1}{4\pi\epsilon_0 r^5} \int_{v'} r'^4 \frac{35 \cos^4 \alpha - 30 \cos^2 \alpha + 3}{8} \rho dv' + \dots \end{aligned} \quad (5-33)$$

$$= V_1 + V_2 + V_3 + V_4 + V_5 + \dots \quad (5-34)$$

Let us examine the first three terms in succession.

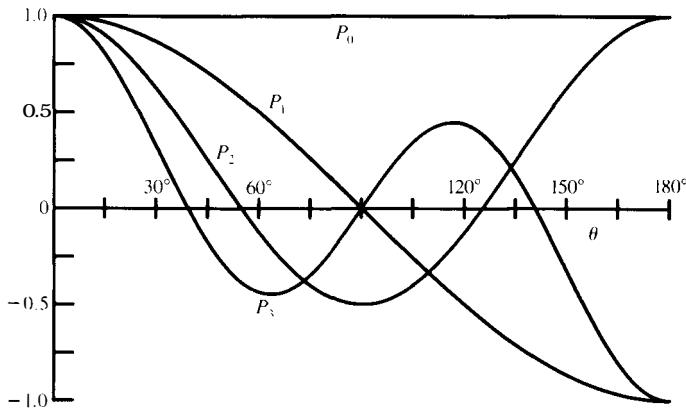


Fig. 5-5. The first four Legendre polynomials $P_n(\cos \alpha)$.

*5.4.2 The Monopole Term

The first term is the V that one would have at P if the whole charge was concentrated at the arbitrary origin:

$$V_1 = \frac{Q}{4\pi\epsilon_0 r}, \quad (5-35)$$

where Q is the net charge in the distribution. This is the *monopole* term. It is zero if the net charge is zero. Its value depends on the position chosen for the origin.

*5.4.3 The Dipole Term

The second term varies as $1/r^2$, like the electric potential of a dipole. From Eq. 5-33

$$V_2 = \frac{1}{4\pi\epsilon_0 r^2} \int_{v'} \mathbf{r}' \cos \alpha \rho dv' = \frac{\hat{\mathbf{r}}}{4\pi\epsilon_0 r^2} \cdot \int_{v'} \mathbf{r}' \rho dv', \quad (5-36)$$

where the integral on the right is the *dipole moment* of the charge distribution:

$$\mathbf{P} = \int_{v'} \mathbf{r}' \rho dv'. \quad (5-37)$$

Thus

$$V_2 = \frac{\mathbf{P} \cdot \hat{\mathbf{r}}}{4\pi\epsilon_0 r^2}, \quad (5-38)$$

as in Eq. 5-8.

*5.4.4 The Quadrupole Term

Now consider the term V_3 of Eq. 5-33. It involves a $1/r^3$ factor, like the V of the linear quadrupole of Sec. 5.2. If we calculate V_3 for the linear quadrupole with charges Q , $-2Q$, and Q at $z = s$, 0 , and $-s$, respectively, we find that it is equal to the V of Eq. 5-17. Then V_3 is the same as if one had a small quadrupole at the origin. Now

$$\begin{aligned} V_3 &= \frac{1}{4\pi\epsilon_0 r^3} \int_{v'} \frac{1}{2} r'^2 (3 \cos^2 \alpha - 1) \rho \, dv' \\ &= \frac{1}{4\pi\epsilon_0 r^3} \int_{v'} \frac{1}{2} [3(\hat{\mathbf{r}} \cdot \mathbf{r}')^2 - r'^2] \rho \, dv'. \end{aligned} \quad (5-39)$$

Set

$$\hat{\mathbf{r}} = l\hat{\mathbf{x}} + m\hat{\mathbf{y}} + n\hat{\mathbf{z}}, \quad (5-40)$$

where l, m, n are the *direction cosines* of $\hat{\mathbf{r}}$ and where

$$l^2 + m^2 + n^2 = 1. \quad (5-41)$$

Then expand and group terms. This yields

$$\begin{aligned} V_3 &= \frac{1}{4\pi\epsilon_0 r^3} \left(3mn \int_{v'} y' z' \rho \, dv' + 3nl \int_{v'} z' x' \rho \, dv' \right. \\ &\quad + 3lm \int_{v'} x' y' \rho \, dv' + \frac{3l^2 - 1}{2} \int_{v'} x'^2 \rho \, dv' \\ &\quad \left. + \frac{3m^2 - 1}{2} \int_{v'} y'^2 \rho \, dv' + \frac{3n^2 - 1}{2} \int_{v'} z'^2 \rho \, dv' \right). \end{aligned} \quad (5-42)$$

These integrals, like the integral of Eq. 5-37, depend solely on the distribution of electric charge within v' , and not on the coordinates x, y, z of the field point P . They specify the nine components of the quadrupole moment of the charge distribution:

$$p_{xx} = \int_{v'} x'^2 \rho \, dv' = Q \overline{x'^2}, \quad (5-43)$$

$$p_{yy} = \int_{v'} y'^2 \rho \, dv' = Q \overline{y'^2}, \quad (5-44)$$

$$p_{zz} = \int_{v'} z'^2 \rho \, dv' = Q \overline{z'^2}, \quad (5-45)$$

$$p_{yz} = p_{zy} = \int_{v'} y' z' \rho \, dv' = Q \overline{y' z'}, \quad (5-46)$$

$$p_{zx} = p_{xz} = \int_{v'} z' x' \rho \, dv' = Q \overline{z' x'}, \quad (5-47)$$

$$p_{xy} = p_{yx} = \int_{v'} x' y' \rho \, dv' = Q \overline{x' y'}. \quad (5-48)$$

The bar indicates, as usual, an average value. Thus

$$\begin{aligned} V_3 = & \frac{1}{4\pi\epsilon_0 r^3} \left(3mnp_{yz} + 3nlp_{zx} + 3lmp_{xy} \right. \\ & + \frac{3l^2 - 1}{2} p_{xx} \\ & \left. + \frac{3m^2 - 1}{2} p_{yy} + \frac{3n^2 - 1}{2} p_{zz} \right) \end{aligned} \quad (5-49)$$

If the charge distribution displays circular symmetry about the z -axis,

$$p_{yz} = p_{zx} = p_{xy} = 0, \quad p_{xx} = p_{yy}. \quad (5-50)$$

It is then convenient to define a single quantity

$$\mathcal{Q} = 2(p_{zz} - p_{xx}) \quad (5-51)$$

that is also called the *quadrupole moment* of the charge distribution. Then, remembering that $l^2 + m^2 + n^2 = 1$, we find that

$$V_3 = \frac{\mathcal{Q}}{16\pi\epsilon_0 r^3} (3n^2 - 1) = \frac{\mathcal{Q}}{16\pi\epsilon_0 r^3} (3 \cos^2 \theta - 1) \quad (5-52)$$

at the point r, θ, ϕ in spherical coordinates.

We can, of course, deduce the electric field strength from the relation $\mathbf{E} = -\nabla V$.

Example

THE FIELD OF A SET OF SIX POINT CHARGES SET SYMMETRICALLY ABOUT THE ORIGIN

Figure 5-6 shows six charges. A point P in space is at a distance $r > a$ from the origin. The potential at P is given by the series of Eq. 5-34. Let us calculate the first three terms. From Eq. 5-35

$$V_1 = \frac{12Q}{4\pi\epsilon_0 r}. \quad (5-53)$$

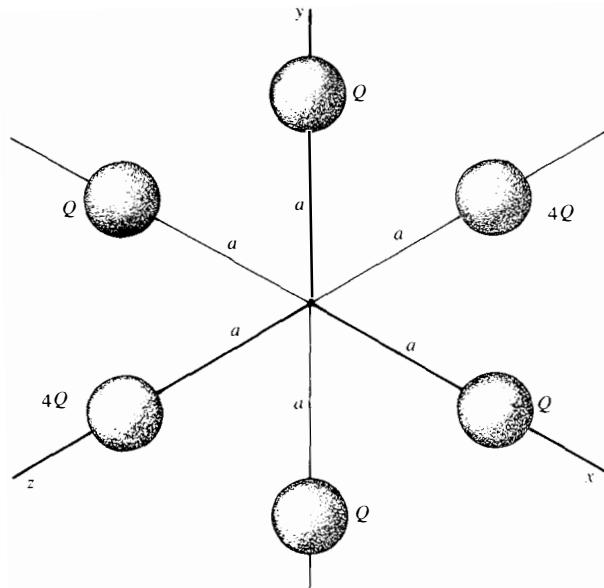


Fig. 5-6. Six charges arranged symmetrically about the origin O .

From Eqs. 5-37 and 5-38,

$$\mathbf{p} = \sum \mathbf{r}' Q = 0, \quad V_2 = 0. \quad (5-54)$$

From Eqs. 5-43 to 5-48,

$$p_{xx} = 2a^2 Q, \quad p_{yy} = 2a^2 Q, \quad p_{zz} = 8a^2 Q, \quad (5-55)$$

$$p_{yz} = p_{zx} = p_{xy} = 0. \quad (5-56)$$

Finally, from Eq. 5-49,

$$V_3 = \frac{1}{4\pi\epsilon_0 r^3} [(3l^2 - 1)a^2 Q + (3m^2 - 1)a^2 Q + (3n^2 - 1)4a^2 Q] \quad (5-57)$$

$$= \frac{a^2 Q}{4\pi\epsilon_0 r^3} (3l^2 + 3m^2 + 12n^2 - 6) \quad (5-58)$$

$$= \frac{a^2 Q}{4\pi\epsilon_0 r^3} (9n^2 - 3) = \frac{3a^2 Q}{4\pi\epsilon_0 r^3} (3n^2 - 1), \quad (5-59)$$

and

$$Q = 12a^2 Q. \quad (5-60)$$

It turns out that $V_4 = 0$, and that

$$V_5 = \frac{Qa^4}{16\pi\epsilon_0 r^5} [35(l^4 + m^4 + 4n^4) - 90n^2 - 12]. \quad (5-61)$$

At P ,

$$V = V_1 + V_3 + V_5 + \dots \quad (5-62)$$

$$\begin{aligned} &= \frac{6Q}{4\pi\epsilon_0 r} \left\{ 1 + \frac{3n^2 - 1}{2} \frac{a^2}{r^2} + \frac{1}{24} [35(l^4 + m^4 + 4n^4) \right. \\ &\quad \left. - 90n^2 - 12] \frac{a^4}{r^4} + \dots \right\}. \end{aligned} \quad (5-63)$$

Since the direction cosines l, m, n are each at most of the order of unity, with $l^2 + m^2 + n^2 = 1$, the coefficient of a^2/r^2 is also of the order of unity, and at $r \geq 10a$ the series reduces to its first term with an error of at most 1%. The terms V_3, V_5, \dots become progressively more prominent as r/a decreases.

5.5 SUMMARY

The *electric dipole* consists of a pair of charges of equal magnitudes Q but of opposite signs, separated by a distance \mathbf{s} . Its *dipole moment* \mathbf{p} is $Q\mathbf{s}$, a vector directed from the negative to the positive charge.

At a distance r from the dipole,

$$V = \frac{\mathbf{p} \cdot \hat{\mathbf{r}}}{4\pi\epsilon_0 r^2}, \quad (r^3 \gg s^3). \quad (5-8)$$

The *linear electric quadrupole* consists of four charges $Q, -2Q, Q, -Q$, as in Fig. 5-3. Then

$$V = \frac{2Qs^2}{4\pi\epsilon_0 r^3} \frac{(3\cos^2\theta - 1)}{2}, \quad (r^3 \gg s^3) \quad (5-17)$$

where $\cos\theta = \hat{\mathbf{z}} \cdot \hat{\mathbf{r}}$.

Some distance outside a charge distribution the potential can be written as a series

$$V = V_1 + V_2 + V_3 + V_4 + V_5 + \dots \quad (5-34)$$

Here V_1 is the potential at P due to a single charge, called a *monopole*, equal to the net charge of the distribution and situated at the position of the arbitrary origin. Similarly, V_2 is the potential at P due to a *dipole* whose dipole moment is equal to that of the distribution and also situated

at the origin, etc.:

$$V_1 = \frac{Q}{4\pi\epsilon_0 r}, \quad (5-35)$$

$$V_2 = \frac{\mathbf{p} \cdot \hat{\mathbf{r}}}{4\pi\epsilon_0 r^2}, \quad (5-38)$$

$$\begin{aligned} V_3 = \frac{1}{4\pi\epsilon_0 r^3} & \left(3mnp_{yz} + 3nlp_{zx} + 3lmp_{xy} \right. \\ & \left. + \frac{3l^2 - 1}{2} p_{xx} + \frac{3m^2 - 1}{2} p_{yy} + \frac{3n^2 - 1}{2} p_{zz} \right), \end{aligned} \quad (5-49)$$

where l, m, n are the direction cosines of the vector \mathbf{r} defining the position of P and where

$$p_{xx} = Q\overline{x'^2}, \quad p_{yy} = Q\overline{y'^2}, \quad p_{zz} = Q\overline{z'^2}, \quad (5-43), (5-45)$$

$$p_{yz} = p_{zy} = Q\overline{y'z'}, \quad p_{zx} = p_{xz} = Q\overline{z'x'}, \quad p_{xy} = p_{yx} = Q\overline{x'y'}. \quad (5-46), (5-48)$$

PROBLEMS

5-1. (5.1) The dipole moment of a charge distribution whose net charge is zero
Show that, if the net charge Q is zero, then the dipole moment of a charge distribution is independent of the choice of origin.

5-2. (5.1) The dipole moment of parallel line charges

Two line charges $+Q$ and $-Q$ extend, respectively, from $(-a, 0, c)$ to $(a, 0, c)$ and from $(-a, 0, -c)$ to $(a, 0, -c)$. Calculate their dipole moment.

5-3. (5-1) The dipole moment of a spherical shell of charge

Calculate the dipole moment of a spherical shell of radius R bearing a surface charge density $\sigma = \sigma_0 \cos \theta$.

5-4. (5.1) The dipole moment of a spherical shell of charge

(a) Calculate the dipole moment of a spherical shell of radius R whose surface charge density is $\sigma_0(1 + \cos \theta)$.

(b) What is the dipole moment if the center of the sphere is at $Z\hat{z}$?

(c) What is the dipole moment if the center of the sphere is at $X\hat{x} + Y\hat{y} + Z\hat{z}$?

5-5. (5.1) An alternate expression for the potential in the field of an electric dipole

We found that, in the field of an electric dipole,

$$V = \frac{Q}{4\pi\epsilon_0} \left(\frac{1}{r_b} - \frac{1}{r_a} \right).$$

Refer to Fig. 5-7. Show that, if the length of the dipole is small, then

$$V = \frac{Qs}{4\pi\epsilon_0} \left[\frac{d}{dz'} \left(\frac{1}{r'} \right) \right]_{z'=0},$$

where z' is the position of a point on the z -axis and $\mathbf{r}' = x\hat{x} + y\hat{y} + (z - z')\hat{z}$.

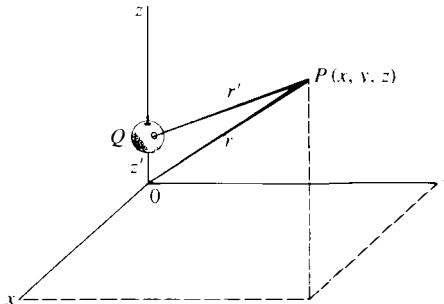


Fig. 5-7.

- 5-6. (5.2) An alternate expression for the potential in the field of an electric quadrupole

See Prob. 5-5 and refer to Fig. 5-8. Show that the potential in the field of a linear electric quadrupole is

$$V = \frac{ps}{4\pi\epsilon_0} \left[\frac{d}{dz'} \left(\frac{\cos\theta}{r'^2} \right) \right]_{z'=0}$$

- 5-7. (5.4) Multipolar expansion of the field of a single point charge

A single point charge Q is situated at $P'(0, 0, s)$ as in Fig. 5-7.

First expand its potential at point P in terms of multipoles. The vector \mathbf{r} that defines the position of P forms an angle θ with the z -axis, and $r \gg s$. The distance from Q to P is r' . Disregard terms of the order of $(s/r)^4$ and higher. Then write out the values of V_1 , V_2 , V_3 .

- 5-8. (5.4.4) The potential close to a dipole

Calculate V for a dipole exactly, and identify the quadrupole and octupole terms. The octupole term varies as $(s/r)^4$. You can therefore disregard terms in $(s/r)^5$, $(s/r)^6$, etc.

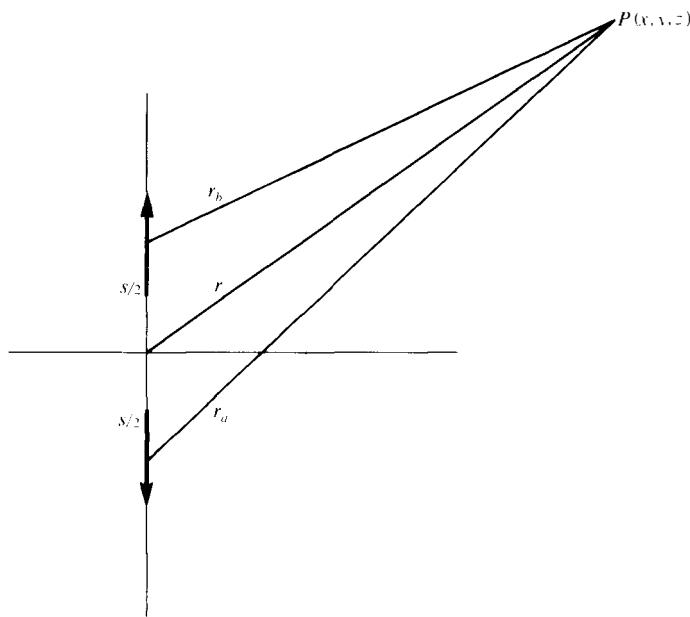
- 5-9. (5.4.4) The field of a charged cube

A cube of side $2a$ carries a uniform volume charge density ρ . The origin of coordinates is at the center. Calculate V_1 , V_2 , V_3 .

- 5-10. (5.4.4) The field of a line charge

A line charge Q extends from $z = -a/2$ to $z = a/2$.

(a) Calculate the monopole, dipole, and quadrupole terms in the expansion for V .

**Fig. 5-8.**

(b) For what value of the distance r to the center of the charge is the quadrupole term less than 1% of the monopole term, if $3n^2 - 1$ is of the order of unity?

- 5-11. (5.4.4) The field of a set of six equal point charges
In Fig. 5-6, let all the charges be Q . Calculate V_4 and V_5 .

CHAPTER 6

ELECTRIC FIELDS IV *Energy, Capacitance, and Forces*

6.1 THE POTENTIAL ENERGY \mathcal{E} OF A CHARGE DISTRIBUTION EXPRESSED IN TERMS OF CHARGES AND POTENTIALS	101
6.1.1 THE POTENTIAL ENERGY OF A SET OF POINT CHARGES	101
6.1.2 THE POTENTIAL ENERGY OF A CONTINUOUS CHARGE DISTRIBUTION	104
6.1.3 TRUE POINT AND LINE CHARGES	104
6.2 THE POTENTIAL ENERGY \mathcal{E} OF AN ELECTRIC CHARGE DISTRIBUTION EXPRESSED IN TERMS OF E	105
Example: THE POTENTIAL ENERGY \mathcal{E} OF A CHARGED CONDUCTING SPHERE	105
6.2.1 THE POTENTIAL ENERGY \mathcal{E} OF A CHARGE DISTRIBUTION EXPRESSED IN TERMS OF E : GENERAL PROOF	107
6.3 THE CAPACITANCE OF AN ISOLATED CONDUCTOR	108
Example: THE CAPACITANCE OF A CONDUCTING SPHERE	108
6.4 THE CAPACITANCE BETWEEN TWO CONDUCTORS	109
Example: THE PARALLEL-PLATE CAPACITOR	109
Example: CAPACITORS IN PARALLEL AND IN SERIES	110
6.5 ELECTRIC FORCES ON CONDUCTORS	110
6.5.1 ELECTRIC FORCES AND LINES OF E	112
6.6 CALCULATING ELECTRIC FORCES BY THE METHOD OF VIRTUAL WORK	113
Example: THE PARALLEL-PLATE CAPACITOR	114
6.7 SUMMARY	115
PROBLEMS	116

This chapter concerns the energy stored in an electric field and the resulting forces exerted on charged conductors. Capacitors are devices designed to store electric energy.

6.1 THE POTENTIAL ENERGY \mathcal{E} OF A CHARGE DISTRIBUTION EXPRESSED IN TERMS OF CHARGES AND POTENTIALS

6.1.1 The Potential Energy of a Set of Point Charges

Imagine a set of N point charges distributed in space as in Fig. 6-1. There are no other charges in the neighborhood. A given charge occupies a

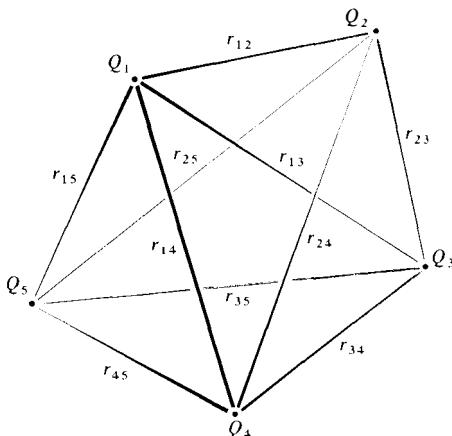


Fig. 6-1. Set of point charges Q_1, Q_2, Q_3, \dots separated by distances r_{12}, r_{13}, r_{14} , etc.

point where the potential due to the other charges is V . That particular charge therefore possesses a potential energy, which is either positive or negative. The system as a whole possesses a potential energy \mathcal{E} that we shall calculate.

Assume that the charges remain in equilibrium under the action of both the electric forces and restraining mechanical forces.

The potential energy of the system is equal to the work performed by the electric forces in the process of dispersing the charges out to infinity. After dispersal, the charges are infinitely remote from each other, and there is zero potential energy.

First, let Q_1 recede to infinity slowly, keeping the electric and the mechanical forces in equilibrium. There is zero acceleration and zero kinetic energy. The other charges remain fixed. The decrease in potential energy \mathcal{E}_1 is equal to Q_1 multiplied by the potential V_1 due to the other charges at the original position of Q_1 :

$$\mathcal{E}_1 = \frac{Q_1}{4\pi\epsilon_0} \left(\frac{Q_2}{r_{12}} + \frac{Q_3}{r_{13}} + \dots + \frac{Q_N}{r_{1N}} \right). \quad (6-1)$$

All the charges except Q_1 appear in the series between parentheses.

With Q_1 removed, let Q_2 recede to infinity, to some point infinitely distant from Q_1 . The decrease in potential energy is now

$$\mathcal{E}_2 = \frac{Q_2}{4\pi\epsilon_0} \left(\frac{Q_3}{r_{23}} + \frac{Q_4}{r_{24}} + \dots + \frac{Q_N}{r_{2N}} \right). \quad (6-2)$$

The series for \mathcal{E}_2 has $N - 2$ terms. We continue the process for all the remaining charges, until finally the N th charge can stay in position, since it lies in a zero field.

The total potential energy of the original charge distribution is then

$$\mathcal{E} = \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 + \cdots + \mathcal{E}_N \quad (6-3)$$

$$\begin{aligned} &= \frac{Q_1}{4\pi\epsilon_0} \left(0 + \frac{Q_2}{r_{12}} + \frac{Q_3}{r_{13}} + \frac{Q_4}{r_{14}} + \cdots + \frac{Q_N}{r_{1N}} \right) \\ &\quad + \frac{Q_2}{4\pi\epsilon_0} \left(0 + 0 + \frac{Q_3}{r_{23}} + \frac{Q_4}{r_{24}} + \cdots + \frac{Q_N}{r_{2N}} \right) \\ &\quad + \frac{Q_3}{4\pi\epsilon_0} \left(0 + 0 + 0 + \frac{Q_4}{r_{34}} + \cdots + \frac{Q_N}{r_{3N}} \right) + \cdots \\ &\quad + \frac{Q_N}{4\pi\epsilon_0} (0 + 0 + 0 + 0 + \cdots + 0). \end{aligned} \quad (6-4)$$

We now rewrite this array, adding, to the left of and below the diagonal line of zeros, terms that are equal to their counterparts to the right of and above the diagonal. Then every term of the series appears twice and

$$\begin{aligned} 2\mathcal{E} &= \frac{Q_1}{4\pi\epsilon_0} \left(0 + \frac{Q_2}{r_{12}} + \frac{Q_3}{r_{13}} + \frac{Q_4}{r_{14}} + \cdots + \frac{Q_N}{r_{1N}} \right) \\ &\quad + \frac{Q_2}{4\pi\epsilon_0} \left(\frac{Q_1}{r_{21}} + 0 + \frac{Q_3}{r_{23}} + \frac{Q_4}{r_{24}} + \cdots + \frac{Q_N}{r_{2N}} \right) \\ &\quad + \frac{Q_3}{4\pi\epsilon_0} \left(\frac{Q_1}{r_{31}} + \frac{Q_2}{r_{32}} + 0 + \frac{Q_4}{r_{34}} + \cdots + \frac{Q_N}{r_{3N}} \right) + \cdots \\ &\quad + \frac{Q_N}{4\pi\epsilon_0} \left(\frac{Q_1}{r_{N1}} + \frac{Q_2}{r_{N2}} + \frac{Q_3}{r_{N3}} + \frac{Q_4}{r_{N4}} + \cdots + 0 \right). \end{aligned} \quad (6-5)$$

On the right, the first line is $Q_1 V_1$, the second line is $Q_2 V_2$, and so forth, where V_i is the potential in the undisturbed system due to all the charges except Q_i at the point occupied by Q_i . Thus

$$2\mathcal{E} = Q_1 V_1 + Q_2 V_2 + Q_3 V_3 + \cdots + Q_N V_N, \quad (6-6)$$

and the potential energy of the initial charge configuration is

$$\mathcal{E} = \frac{1}{2} \sum_{i=1}^N Q_i V_i. \quad (6-7)$$

The reason for the factor of $\frac{1}{2}$ follows from the above calculation. Let all the charges be positive. Then the potential at the position of a given charge, just before it moves out to infinity, is less (except for Q_1) than

the potential at the same point in the original charge distribution. On the average, the potential just before removal is one-half the potential in the original charge distribution.

This energy \mathcal{E} , which does *not* include the energy required to assemble the individual charges themselves, can be positive, negative, or zero. For example, for two charges of the same sign, \mathcal{E} is positive. For charges of opposite signs, \mathcal{E} is negative. For a single charge, \mathcal{E} is zero.

But what is the energy required to simply modify a charge distribution, without dispersing it to infinity? This energy is clearly equal to the final potential energy minus the initial potential energy, whatever method one may choose to effect the change.

6.1.2 The Potential Energy of a Continuous Charge Distribution

For a continuous electric charge distribution, we replace Q_i by ρdv and the summation by an integration over any volume v that contains all the charge:

$$\mathcal{E} = \frac{1}{2} \int_v V \rho \, dv. \quad (6-8)$$

This integral is equal to the work performed by the electric forces in going from the given charge distribution to the situation where $\rho = 0$ everywhere, by dispersing all the charge to infinity, or by letting positive and negative charges coalesce, or by both processes combined.

At first sight, this is an obvious extension of the previous equation. It is not, because we have now included the energies required to assemble the individual macroscopic charges. In fact, as we shall see in Sec. 6.2, the above integral is always positive.

Observe that the potential V under the integral sign does not include the part that originates in the element of charge ρdv itself. We saw in Sec. 3.5 that the infinitesimal element of charge at a given point contributes nothing to V .

If there are surface charge densities σ , then their stored energy is

$$\mathcal{E} = \frac{1}{2} \int_{\mathcal{A}} \sigma V \, d\mathcal{A}, \quad (6-9)$$

where \mathcal{A} includes all the surfaces carrying charge.

6.1.3 True Point and Line Charges

Suppose we have a spherical charge Q of uniform volume density and radius R . Then, using the value of V_i that we calculated in Sec. 3.7.1,

$$\mathcal{E} = \frac{1}{2} \int_0^R \left(\frac{Q}{4\pi R^3/3} \right) \left[\frac{Q}{4\pi\epsilon_0 R} \left(\frac{3}{2} - \frac{r^2}{2R^2} \right) \right] 4\pi r^2 dr = \frac{3Q^2}{20\pi\epsilon_0 R}. \quad (6-10)$$

If R is zero, then \mathcal{E} is infinite, which is nonsense. Electrons are presumably true point charges, and dealing with this absurd result poses difficult problems whose solutions are well beyond the scope of this book. With a true line charge, \mathcal{E} is similarly infinite.

True point and line charges are therefore not allowed in the present context. Nonetheless, we follow the usual custom of speaking loosely of point and line charges when the radius is negligibly small. True surface charges cause no problems.

6.2 THE POTENTIAL ENERGY \mathcal{E} OF AN ELECTRIC CHARGE DISTRIBUTION EXPRESSED IN TERMS OF \mathbf{E}

We have expressed the potential energy \mathcal{E} of a charge distribution in terms of the charge density ρ and the potential V . Now both ρ and V are related to \mathbf{E} . So it should be possible to express \mathcal{E} solely in terms of \mathbf{E} . This is what we shall do here. We shall find that

$$\mathcal{E} = \int_v \frac{\epsilon_0 E^2}{2} dv, \quad (6-11)$$

where the volume v includes all the regions where \mathbf{E} exists. Thus we can calculate \mathcal{E} by assigning to each point in space an *electric energy density* of $\epsilon_0 E^2/2$.

Since the above \mathcal{E} is positive, that of Eq. 6-8 is also always positive. These two equivalent expressions for \mathcal{E} , one in terms of ρ and V and the other in terms of \mathbf{E} , are both important. They emphasize different, but complementary, aspects of electrical phenomena. With the first expression, \mathcal{E} is the potential energy of a system of *charges*; with the second, \mathcal{E} is the energy stored in a *field*.

We first apply the above formula to the field of a charged spherical conductor, and then we give a general proof.

Example

THE POTENTIAL ENERGY \mathcal{E} OF A CHARGED CONDUCTING SPHERE

We find the potential energy \mathcal{E} of a conducting sphere of radius R carrying a charge Q in three different ways.

First method

The whole charge Q is at the potential $Q/4\pi\epsilon_0 R$. Then

$$\mathcal{E} = \frac{1}{2} Q \frac{Q}{4\pi\epsilon_0 R} = \frac{Q^2}{8\pi\epsilon_0 R}. \quad (6-12)$$

Second method

Imagine that the radius of the charged spherical conductor increases slowly and eventually becomes infinite. The total mechanical work performed by the charges is equal to the initial potential energy.

During the expansion, the field outside remains unaffected (Gauss's law again!). An element of charge $\sigma dA = \epsilon_0 E dA$ is subjected to an electric field strength equal to $E/2$ (Sec. 6.5), and the work performed by this element, when the radius of the sphere increases by dR , is

$$d\mathcal{E} = (\epsilon_0 E dA) \left(\frac{E}{2}\right) dR = \frac{\epsilon_0 E^2}{2} dv, \quad (6-13)$$

where dv is the element of volume swept by the element of area dA . After the radius has expanded to infinity, the total work performed by the charges is

$$\mathcal{E} = \int_R^\infty \frac{\epsilon_0 E^2}{2} dv = \int_R^\infty \frac{\epsilon_0}{2} \left(\frac{Q}{4\pi\epsilon_0 r^2}\right)^2 4\pi r^2 dr = \frac{Q^2}{8\pi\epsilon_0 R} \quad (6-14)$$

as above.

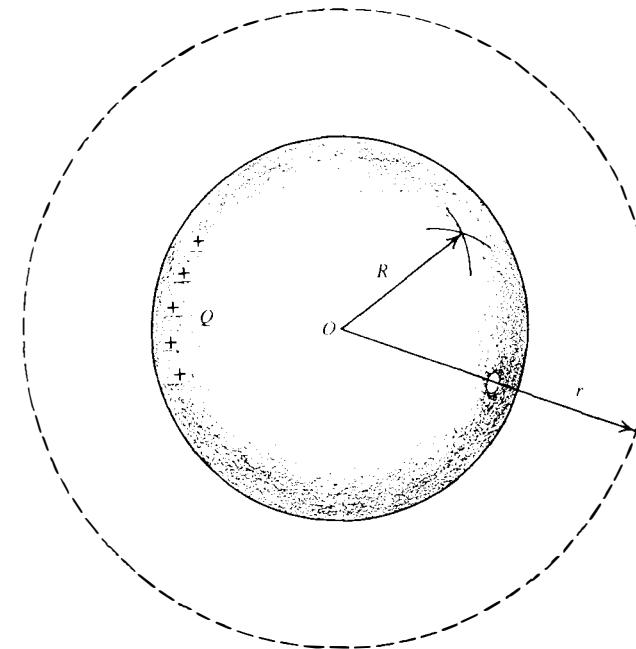


Fig. 6-2. Spherical conductor carrying a charge Q , and a concentric imaginary spherical surface of area $A(r)$.

Third method

Let $\mathcal{A}(r)$ be the area of any surface of radius r , concentric with the conducting sphere and outside it, as in Fig. 6-2. Then, from Gauss's law (Sec. 3.7),

$$\mathcal{E} = \frac{1}{2}QV = \frac{1}{2} \int_{\mathcal{A}(r)} \epsilon_0 E d\mathcal{A} \int_r^{\infty} E dr. \quad (6-15)$$

Since the two integrals are independent of each other,

$$\mathcal{E} = \frac{1}{2} \int_R^{\infty} \left(\int_{\mathcal{A}(r)} \epsilon_0 E d\mathcal{A} \right) E dr \quad (6-16)$$

$$= \frac{1}{2} \int_R^{\infty} (4\pi r^2 \epsilon_0 E) E dr = \int_R^{\infty} \frac{\epsilon_0 E^2}{2} 4\pi r^2 dr \quad (6-17)$$

$$= \int_R^{\infty} \frac{\epsilon_0 E^2}{2} dv = \frac{Q^2}{8\pi\epsilon_0 R}, \quad (6-18)$$

as previously.

6.2.1 The Potential Energy \mathcal{E} of a Charge Distribution Expressed in Terms of E : General Proof

We exclude unrealistic cases where V would be discontinuous, for this would require an infinite E . We also set $V = 0$ at infinity, which excludes charges of infinite extent. Then V has a finite maximum V_{\max} and a finite minimum V_{\min} , with $V_{\max} \geq 0$, $V_{\min} \leq 0$.

Let V_{\min} be negative and imagine a conductor that occupies all points where $V = V_{\min}$. The conductor expands out to the equipotential $V_{\min} + dV$. This does not affect the rest of the field. Eventually, the conductor reaches the equipotential $V = 0$. In the course of the expansion, any charge encountered accumulates on the surface of the conductor.

According to the second method above, the work performed by the charges is equal to the integral of $\epsilon_0 E^2/2$ over the volume swept out.

Now let V_{\max} be greater than zero, and imagine another conductor occupying the region where $V = V_{\max}$. It expands as above until it reaches the equipotential $V = 0$. Again the work performed is the integral of $\epsilon_0 E^2/2$ over the volume swept out.

If the two conductors meet, then, immediately before contact, the surface charge densities are equal in magnitude (same E), opposite in sign, and at the same potential. They cancel. The charge density is now zero everywhere, either because the charges are dispersed to infinity or because positive and negative charges have neutralized.

The initial stored energy is thus given correctly by the integral of $\epsilon_0 E^2/2$.

6.3 THE CAPACITANCE OF AN ISOLATED CONDUCTOR

Imagine a finite conductor situated a long distance from any other body and carrying a charge Q . If Q changes, the conductor's potential also changes. As we shall see, the ratio Q/V is a constant. The *capacitance* of the isolated conductor is

$$C = \frac{Q}{V}. \quad (6-19)$$

Thus the capacitance of an isolated conductor is equal to the extra charge required to increase its potential by 1 volt. The unit of capacitance is the *farad*, or coulomb per volt.

The energy stored in the field of an isolated conductor is

$$\mathcal{E} = \frac{QV}{2} = \frac{CV^2}{2} = \frac{Q^2}{2C}. \quad (6-20)$$

We now show that the capacitance C of an isolated conductor depends solely on its size and shape. The conductor is in air. The potential in the region surrounding the conductor is $V(x, y, z)$. It obeys Laplace's equation, and it is zero at infinity because the conductor is of finite size, by hypothesis. At the surface of the conductor the charge density σ is $\epsilon_0 E$ (Sec. 4.4), or ϵ_0 times the rate of change of $V(x, y, z)$ in the direction normal to the surface. An immediate consequence is that the value of the conductor potential determines the surface charge density $\sigma = -\epsilon_0 |\nabla V|$ on the conductor. Therefore the conductor potential also determines the total charge Q on the conductor.

Observe now that the equation $\nabla^2 V = 0$ is linear, so that any multiple of $V(x, y, z)$ is also a solution. If V increases everywhere by some factor a , this new V obeys Laplace's equation outside the conductor, is zero at infinity, and is equal to aV_c on the conductor. Furthermore, it is the only continuous function of x, y, z that satisfies these three conditions. We conclude that if the conductor's potential increases by the factor a , then V increases everywhere by the same factor a , and both σ and Q likewise increase by the same factor.

The charge Q on an isolated conductor is thus proportional to its voltage, and its capacitance $C = Q/V$ depends solely on the size and shape of the conductor.

Example

THE CAPACITANCE OF A CONDUCTING SPHERE

If an isolated conducting sphere of radius R carries a charge Q , the potential at its surface is $Q/4\pi\epsilon_0 R$ and

$$C = 4\pi\epsilon_0 R = 1.11 \times 10^{-10} R \text{ farad}$$

$$= 111R \text{ picofarads.} \quad (6-21)$$

6.4 THE CAPACITANCE BETWEEN TWO CONDUCTORS

We now have two uncharged isolated conductors. Transferring a charge Q from one to the other establishes a potential difference V between them. By definition, the *capacitance* between the conductors is Q/V . The capacitance depends solely on the geometry of the conductors and on their relative positions.

Pairs of conductors arranged specifically to possess capacitance are called *capacitors*.

Example THE PARALLEL-PLATE CAPACITOR

A parallel-plate capacitor (Fig. 6-3) consists of two conducting plates of area \mathcal{A} , separated by a distance s . The plates carry charges Q and $-Q$. We neglect edge effects. From Gauss's law,

$$E = \frac{\sigma}{\epsilon_0} = \frac{Q}{\mathcal{A}\epsilon_0}, \quad V = \frac{Qs}{\mathcal{A}\epsilon_0}. \quad (6-22)$$

Then

$$C = \frac{\epsilon_0 \mathcal{A}}{s}. \quad (6-23)$$

Also, the stored energy is

$$\mathcal{E} = \frac{QV}{2} = \frac{CV^2}{2} = \frac{Q^2}{2C} = \frac{Q^2 s}{2\epsilon_0 \mathcal{A}}, \quad (6-24)$$

or

$$\mathcal{E} = \int \frac{\epsilon_0 E^2}{2} dv = \frac{\epsilon_0}{2} \left(\frac{Q}{\epsilon_0 \mathcal{A}} \right)^2 \mathcal{A}s = \frac{Q^2 s}{2\epsilon_0 \mathcal{A}}. \quad (6-25)$$

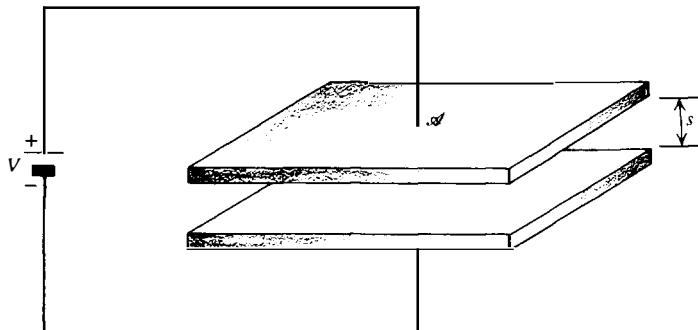


Fig. 6-3. Parallel-plate capacitor connected to a battery.

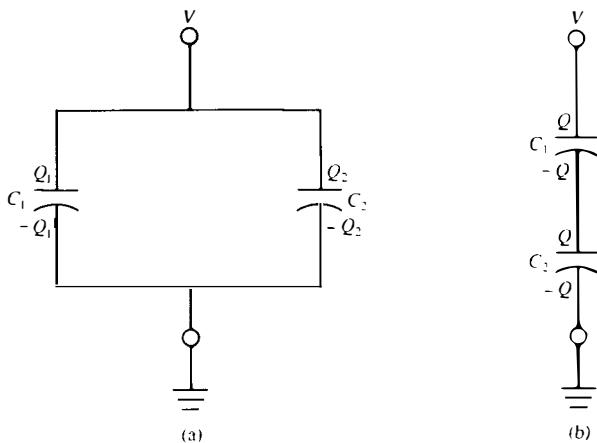


Fig. 6-4. (a) Two capacitors connected in parallel. (b) Two capacitors connected in series.

Example

CAPACITORS IN PARALLEL AND IN SERIES

Capacitors connected in parallel share the same voltage. Thus, for two capacitors, as in Fig. 6-4(a),

$$C = \frac{Q_1 + Q_2}{V} = \frac{Q_1}{V} + \frac{Q_2}{V} = C_1 + C_2. \quad (6-26)$$

Capacitors connected in series carry the same charges, as in Fig. 6-4(b). Then

$$V = \frac{Q}{C_1} + \frac{Q}{C_2} = \frac{Q}{C}, \quad (6-27)$$

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2}, \quad C = \frac{C_1 C_2}{C_1 + C_2}. \quad (6-28)$$

So capacitors connected in parallel add as resistors in series, and capacitors connected in series add as resistors in parallel.

6.5 ELECTRIC FORCES ON CONDUCTORS

An element of charge σdA on the surface of a conductor experiences the electric field of all the *other* charges and is therefore subjected to an electric force. Under static conditions this force is perpendicular to the surface, for otherwise there would be a tangential field and a tangential current. The force also acts on the conductor, to which σdA is bound by internal electric forces.

To calculate the magnitude of the electric force, consider a conductor carrying a surface charge density σ with an electric field strength E near

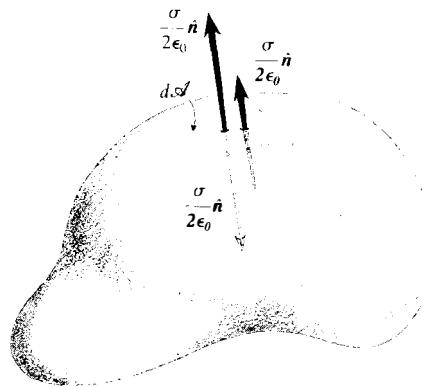


Fig. 6-5. The local surface charge density σ on a conductor gives two oppositely directed electric fields, as shown by the two arrows on the left. All the other charges together give the field shown on the right. The net result is a field strength of σ/ϵ_0 outside and zero inside. The unit vector \hat{n} points *outward*, as usual.

the surface. From Gauss's law, this E is σ/ϵ_0 . Now the force on $\sigma d\mathcal{A}$ is *not* $E\sigma d\mathcal{A}$, because the field that acts on $\sigma d\mathcal{A}$ is only that of the *other* charges in the system.

We can find the field of $\sigma d\mathcal{A}$ itself from Gauss's law. The flux of E emerging from $\sigma d\mathcal{A}$ is $\sigma d\mathcal{A}/\epsilon_0$, half of it inward and half outward, as in Fig. 6-5. Then $\sigma d\mathcal{A}$ provides exactly half the total E at a point outside, close to the surface, and cancels the field of the other charges, inside.

Therefore the E acting on $\sigma d\mathcal{A}$ is $\sigma/2\epsilon_0$, and the force on the element of area $d\mathcal{A}$ of the conductor is

$$dF = \frac{\sigma}{2\epsilon_0} \sigma d\mathcal{A} = \frac{\sigma^2}{2\epsilon_0} d\mathcal{A}. \quad (6-29)$$

The surface force density is

$$F' = \frac{dF}{d\mathcal{A}} = \frac{\sigma^2}{2\epsilon_0} = \frac{\epsilon_0 E^2}{2} \quad \text{newtons/meter}^2. \quad (6-30)$$

The force per unit area on a conductor is equal to the energy density in the field.

The net electrostatic force on a conductor of area \mathcal{A} is

$$\mathbf{F} = \frac{\epsilon_0}{2} \oint_{\mathcal{A}} E^2 d\mathcal{A}, \quad (6-31)$$

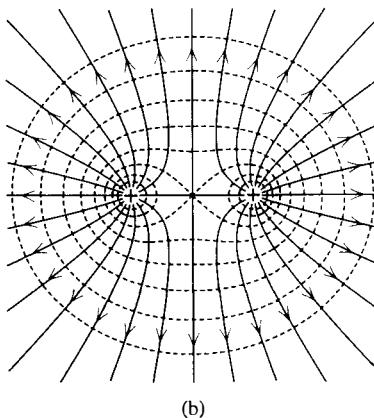
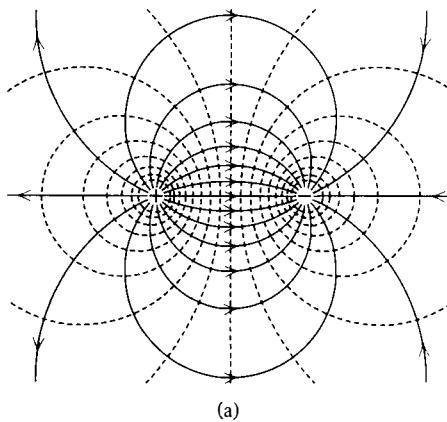
where the vector $d\mathcal{A}$ points outward. The local electric force tends to pull the conductor into the field. In other words, an electrostatic field exerts a negative pressure on a conductor. The net force on a conductor depends on the way σ and \mathbf{E} vary along its surface. In air, or in a vacuum, electric forces are usually negligible. However, they can be quite large in dielectrics.

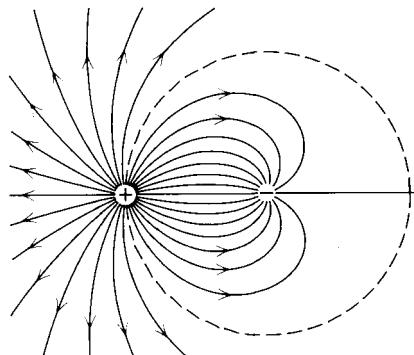
6.5.1 Electric Forces and Lines of \mathbf{E}

Figure 6-6 shows lines of \mathbf{E} for four pairs of line charges.

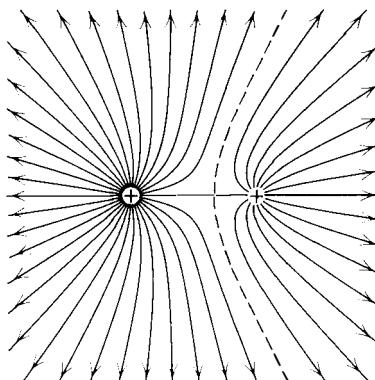
In Fig. 6-6(a) and (c), the force is attractive and the lines of \mathbf{E} are clearly “under tension.” Indeed, the tensile force per square meter is $\epsilon_0 E^2/2$, as we can infer from the previous section.

In Fig. 6-6(b) and (d), on the other hand, we can see lines of \mathbf{E} “repelling” each other laterally. The repulsive surface force density in the





(c)



(d)

Fig. 6-6. Lines of \mathbf{E} between pairs of line charges. It is useful to visualize attractive electric forces as being caused by a tension in the lines of \mathbf{E} , as in (a) and (c). Similarly, repulsive electric forces may be thought of as being caused by a lateral repulsion between lines of \mathbf{E} , as in (b) and (d). In (c) and (d) the charge on the left is twice as large as the other.

region where the lines of \mathbf{E} are parallel is also $\epsilon_0 E^2/2$. See Prob. 6-8.

Later we shall see that magnetic fields behave similarly.

6.6 CALCULATING ELECTRIC FORCES BY THE METHOD OF VIRTUAL WORK

We can also calculate electric forces by the *method of virtual work*. This method consists in postulating an infinitesimal displacement and then applying the principle of conservation of energy. We first define a system,

and then we calculate the energy fed into it in the course of the displacement. This energy is equal to the increase in the internal energy of the system.

The method of virtual work is a general and reliable method for calculating forces, but on two conditions: (1) one must be perfectly clear about exactly what system one is talking about, and (2) one must be particularly careful to use the proper signs.

Example

THE PARALLEL-PLATE CAPACITOR

A parallel-plate capacitor is connected to a battery supplying a fixed voltage V . See Fig. 6-7. We assume that the distance s between the plates increases by ds , and we apply the principle of virtual work. We are going to calculate energies related to the capacitor. Thus the net energy fed into the capacitor in the course of the displacement ds of the top plate will be equal to the increase in the electric energy stored in the capacitor.

Let \mathcal{E}_F be the mechanical energy fed *into* the capacitor by the force F , let \mathcal{E}_B be the electric energy fed *in* by the battery, and let \mathcal{E}_E be the increase in the electric energy of the capacitor. Then

$$\mathcal{E}_F + \mathcal{E}_B = \mathcal{E}_E, \quad (6-32)$$

with

$$\mathcal{E}_F = F ds, \quad \mathcal{E}_B = V dQ, \quad (6-33)$$

where dQ is the extra charge fed into the capacitor by the battery.

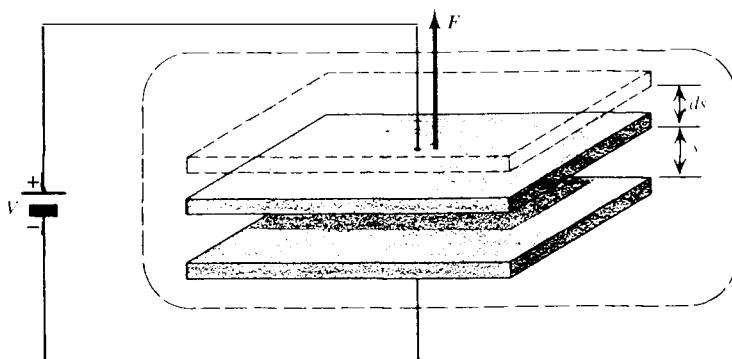


Fig. 6-7. Parallel-plate capacitor connected to a battery supplying a fixed voltage V . We can calculate the force of attraction between the plates by imagining an equal but opposite force F that pulls one of the plates away by an infinitesimal distance ds . The dashed curve reminds us that we apply the principle of the conservation of energy to the capacitor alone.

Here

$$\mathcal{E}_s = V(V dC) = V^2 d \frac{\epsilon_0 \mathcal{A}}{s} = -V^2 \epsilon_0 \mathcal{A} \frac{ds}{s^2} = -(\epsilon_0 E^2)(\mathcal{A} ds). \quad (6-34)$$

If ds is positive, dC is negative and energy flows from the capacitor into the battery.

Finally, since $E = V/s$,

$$\mathcal{E}_E = d \left(\frac{\epsilon_0 E^2}{2} \mathcal{A} s \right) = d \left(\frac{\epsilon_0 V^2 \mathcal{A} s}{2s^2} \right) = \epsilon_0 \frac{\mathcal{A} V^2}{2} \left(-\frac{ds}{s^2} \right) = -\frac{\epsilon_0 E^2}{2} \mathcal{A} ds. \quad (6-35)$$

Note the negative sign! The energy density $\epsilon_0 E^2 / 2V$ decreases faster than the volume $\mathcal{A}s$ increases.

Thus

$$F ds = \epsilon_0 E^2 \mathcal{A} ds = -\frac{\epsilon_0 E^2}{2} \mathcal{A} ds, \quad (6-36)$$

$$F = \frac{\epsilon_0 E^2}{2} \mathcal{A}, \quad (6-37)$$

and the force per unit area is equal to the energy per unit volume, as previously.

Half the energy supplied by the source becomes mechanical energy, and the other half becomes electric energy. This is a general rule.

6.7 SUMMARY

The *potential energy* of a charge distribution is given by either one of two integrals:

$$\mathcal{E} = \frac{1}{2} \int_v V \rho dv \quad (6-8)$$

$$\mathcal{E} = \frac{1}{2} \int_v \epsilon_0 E^2 dv. \quad (6-11)$$

In the first integral the volume v contains all the charges, while in the second it includes all the field. The assignment of an *electric energy density* $\epsilon_0 E^2 / 2$ to every point in space leads to the correct potential energy of a charge distribution.

If the potential of an isolated conductor is V when its charge is Q , then its *capacitance* C is Q/V farads. This quantity depends solely on the geometry of the conductor. The capacitance between two conductors is again Q/V , where Q is now the charge transferred from one to the other

and V is the potential between them. In both cases the stored energy is

$$\mathcal{E} = \frac{QV}{2} = \frac{CV^2}{2} = \frac{Q^2}{2C}. \quad (6-20)$$

The force per unit area exerted on a conductor situated in an electric field is equal to the electric energy density at the surface:

$$\mathbf{F}' = \frac{\epsilon_0 E^2}{2} = \frac{\sigma^2}{2\epsilon_0}, \quad (6-30)$$

where σ is the surface charge density.

It is useful to visualize electric forces as being caused by lines of E that are under tension and that repel each other laterally.

We can also calculate the electric force on a conductor by the principle of virtual work, which is simply the principle of conservation of energy applied to an infinitesimal disturbance of a system.

PROBLEMS

6-1. (6.1) The dipole and the quadrupole

Calculate the potential energies of an electric dipole and of an electric quadrupole.

6-2. (6.1) The potential energy of a sphere of charge

(a) Calculate the electric potential energy of a sphere of radius R carrying a total charge Q uniformly distributed throughout its volume.

(b) Calculate the gravitational potential energy of a sphere of radius R' and total mass M .

(c) Calculate the gravitational potential energy of the moon. See The Table of Physical Constants at the end of the book.

(d) Imagine that you can assemble a sphere of protons with a density equal to that of water. What would be the radius of this sphere if its electric potential energy were sufficient to blow up the moon?

(e) What is the voltage at the surface of the sphere of protons?

6-3. (6.2) The energy in the field of a sphere of charge

A sphere of radius R contains a charge Q , uniformly distributed throughout its volume. Calculate (a) the energy, (b) the energy stored in the field inside the sphere, and (c) the energy stored in the field outside the sphere. There is five times more energy outside than inside.

6-4. (6.4) The reciprocity theorem for electrostatics

Consider a set of n conductors of arbitrary sizes, shapes, and positions. Conductors 1, 2, 3, ... carry charges Q_1, Q_2, Q_3, \dots , and their voltages are V_1, V_2, V_3, \dots . Without disturbing the conductors, you change the charges to Q'_1, Q'_2, Q'_3 .

According to the *reciprocity theorem for electrostatics*,

$$Q_1 V'_1 + Q_2 V'_2 + Q_3 V'_3 + \dots = Q'_1 V_1 + Q'_2 V_2 + Q'_3 V_3 + \dots,$$

or

$$\sum QV' = \sum Q'V.$$

We shall find analogous reciprocity theorems in Chaps. 8 and 27.

You can prove this theorem by calculating the energy required to change the charges from Q to Q' and equating this energy to $\sum Q'V'/2 - \sum QV/2$. To do this, set the charge and voltage on conductor 1 equal to $(1-x)Q_1 + xQ'_1$ and $(1-x)V_1 + xV'_1$. Then you can go from one state to the other by letting x go from zero to unity.

6-5. (6.4) Cylindrical capacitor

- (a) Show that the capacitance per unit length of a cylindrical capacitor is $C' = 2\pi\epsilon_0/\ln(R_2/R_1)$, where R_1 and R_2 are the inner and outer radii.
- (b) Calculate the capacitance per meter when $R_2/R_1 = e = 2.718$.

6-6. (6.4) Spherical capacitor

- (a) Show that the capacitance of a spherical capacitor of inner and outer radii R_1 and R_2 is

$$C = \frac{4\pi\epsilon_0 R_1 R_2}{R_2 - R_1}.$$

- (b) Calculate the capacitance when $R_1 = 100$ millimeters and $R_2 = 200$ millimeters.

6-7. (6.4) Connecting charged capacitors in parallel

Two capacitors C_1 and C_2 are charged to voltages V_1 and V_2 , respectively, and then connected in parallel, positive terminal to positive terminal and negative to negative.

- (a) What is the final voltage?
- (b) What happens to the stored energy?

6-8. (6.5.1) Electric forces and lines of E

- (a) Use the method of Sec. 6.5.1 to calculate the force of attraction between two charges Q and $-Q$ separated by a distance $2D$.
- (b) Repeat the calculation for two charges of equal sign.

6-9. (6.5) High-voltage generator

Imagine the following mechanism for generating high voltages. One plate of a parallel-plate capacitor is fixed and connected to ground. The other plate is movable. When the plates are close together at a distance s , a contact closes and the movable plate charges to the voltage V . Then the contact opens, the movable plate moves out to a distance ns , and its voltage increases to nV , disregarding edge effects. At this point another contact closes, and the movable plate discharges to ground through a load resistance R .

- (a) Verify that there is conservation of energy.
- (b) Can you suggest a more convenient geometry for such a high-voltage generator?

6-10. (6.5) The surface force on a balloon carrying an electric charge

It is suggested that a balloon made of light conducting material could be kept approximately spherical by connecting it to a high-voltage supply. The balloon has a diameter of 100 millimeters, and the maximum breakdown field in air is 3 megavolts/meter.

- (a) What is the maximum permissible voltage?
- (b) What gas pressure, in atmospheres, inside the balloon would have the same effect?
- (c) How large could the surface mass density of the balloon be?

6-11. (6.6) Stored energy

Four charges $+Q, -Q, +Q, -Q$ occupy the corners of a square of side a , with the positive charges on one diagonal and the negative charges on the other.

- (a) Calculate the stored energy \mathcal{E} , and sketch a curve of \mathcal{E} as a function of a .
- (b) A mechanism constrains the charges to stay at the corners of a square but allows a to vary. What will happen?
- (c) Calculate the forces on the charges by the method of virtual work.
- (d) Compare with the values deduced from Coulomb's law.

6-12. (6.6) The forces on the plates of a parallel-plate capacitor

Show that the force of attraction between the plates of a parallel-plate capacitor that is not connected to a battery is $\epsilon_0 E^2 A/2$, as in Sec. 6.6.

6-13. (6.6) Half the battery energy becomes mechanical work, and the other half is stored in the electric field.

Rewrite Eq. 6-32 in the form $\mathcal{E}_B = \mathcal{E}_E - \mathcal{E}_F$, and show that, if ds is negative, one-half of the energy supplied by the battery becomes electric energy, while the other half performs mechanical work.

CHAPTER 7

*ELECTRIC FIELDS V

Electric Circuits A: RC Circuits

1	SYMBOLS AND DEFINITIONS	120
2	KIRCHHOFF'S LAWS FOR DIRECT CURRENTS	121
2.1	THE MESH METHOD FOR CALCULATING BRANCH CURRENTS	122
	<i>Example: SIMPLE TWO-MESH RESISTIVE CIRCUIT WITH A VOLTAGE SOURCE</i>	122
	<i>Example: SIMPLE TWO-MESH RESISTIVE CIRCUIT WITH A CURRENT SOURCE</i>	123
	<i>Example: SIMPLE TWO-MESH RC CIRCUIT WITH A VOLTAGE SOURCE. SOLVING AN ORDINARY DIFFERENTIAL EQUATION WITH CONSTANT COEFFICIENTS</i>	124
2.2	THE NODE METHOD FOR CALCULATING NODE VOLTAGES	126
	<i>Example: SIMPLE TWO-MESH RESISTIVE CIRCUIT WITH A VOLTAGE SOURCE</i>	127
3	ALTERNATING CURRENTS	127
3.1	THREE-WIRE SINGLE-PHASE ALTERNATING CURRENTS	128
3.2	THREE-PHASE ALTERNATING CURRENT	128
4	ALTERNATING CURRENTS IN RESISTORS	130
5	ALTERNATING CURRENTS IN CAPACITORS	132
6	THE IMPEDANCE Z AND THE ADMITTANCE Y	134
	<i>Example: SIMPLE RC CIRCUIT</i>	135
7	POWER IN AC CIRCUITS	137
8	KIRCHHOFF'S LAWS FOR ALTERNATING-CURRENT CIRCUITS	138
	<i>Example: SIMPLE RC CIRCUIT SOLVED BY BOTH THE MESH AND NODE METHODS</i>	138
9	SUMMARY	139
	PROBLEMS	141

This is the first of four chapters on electric circuits. For the moment, we limit ourselves to resistive and capacitive components. The next chapter deals with circuit theorems and is general, even though it appears to

* Chapters 7 and 8 are not required for what follows. However, Chap. 7 is a prerequisite for Chap. 8.

suffer from the same limitation. We have to wait until Chaps. 24 and 25 to study inductive circuits and transformers.

Unless specified otherwise, we assume that all circuit components are linear and time-independent.

If you are not familiar with phasors, you should read Chap. 2 before working on the material on alternating currents that starts in Sec. 7.3.

7.1 SYMBOLS AND DEFINITIONS

Figure 7-1 shows the circuit-related standard symbols that we use in this book. An ideal *voltage source* supplies a voltage that is independent of the current drawn. A good commercial voltage-stabilized power supply is close to ideal, up to a specified current, after which the voltage drops off, as in Fig. 7-2. Similarly, an ideal *current source* supplies a current that is independent of the output voltage. Commercial current-stabilized power supplies are nearly ideal, up to a specified voltage, beyond which the current drops, as in Fig. 7-3.

Figure 7-4 shows part of an electric circuit. Points such as *A*, *B*, *C* are called *nodes*; connections between nodes such as *AB*, or *BC*, are *branches*; and a closed circuit such as *ABC* is a *mesh*.

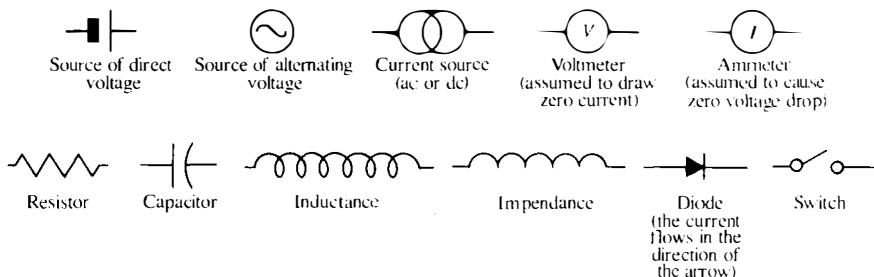


Fig. 7-1. Symbols for circuit components.

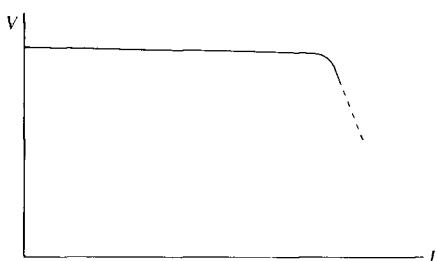


Fig. 7-2. Typical curve of output voltage as a function of output current for a commercial *voltage-stabilized power supply*.

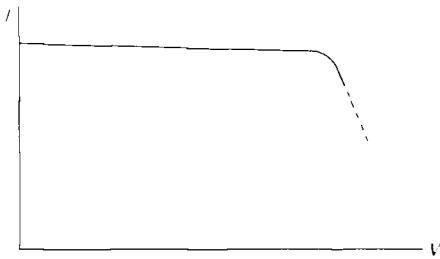


Fig. 7-3. Typical curve of output current as a function of output voltage for a commercial *current-stabilized* power supply.

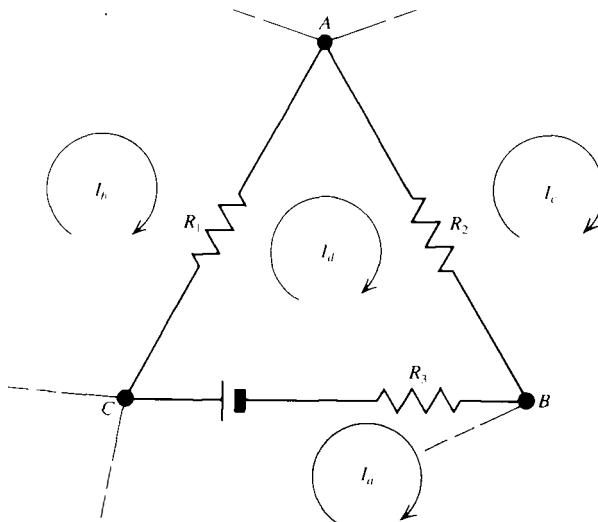


Fig. 7-4. Section of an electric circuit. This mesh has three branches and three nodes.

An *active circuit* comprises sources, while a *passive circuit* does not. A circuit component is said to be *linear* if the current passing through it is proportional to the applied voltage.

By convention, current flows from plus to minus, outside a source, in the direction opposite to the electron drift.

7.2 KIRCHHOFF'S LAWS FOR DIRECT CURRENTS

Kirchhoff's two laws are self-evident and appear trivial. They are, in fact, the fundamental laws of circuit theory, and they are completely general: they are valid for both linear and nonlinear, and for both passive and

active circuits. They remain valid even if the components are time-dependent or if they display hysteresis.[†]

They also apply, whatever be the time dependence of the voltages or currents or whatever be the initial conditions. All circuit theorems derive, in some way, from these two laws. Kirchhoff's laws serve to calculate branch currents and node voltages in circuits comprising known sources and known components.

We disregard alternating currents until Sec. 7.3.

The *Kirchhoff current law* (KCL) states that the algebraic sum of the currents entering a node is equal to zero. This stands to reason because, otherwise, charge would accumulate at the nodes. A node, by itself, has zero capacitance to ground. If a given connection *A* possesses an appreciable capacitance to ground, then that capacitance must appear on the circuit as a separate branch between *A* and ground.

The *Kirchhoff voltage law* (KVL) is equally obvious: the sum of the voltage drops round a mesh is equal to zero.

7.2.1 The Mesh Method for Calculating Branch Currents

With the *mesh method* one assigns *mesh currents* I_a, I_b, I_c, \dots , as in Fig. 7-4. This current distribution satisfies the KCL automatically. The meshes may be chosen arbitrarily, as long as no branch is left out. Then the KVL provides one equation for each mesh. Solving yields the mesh currents, the branch currents, and then the node voltages, as we shall see in the examples below.

If the circuit comprises only sources and resistors, then the equations are algebraic. However, if there are also capacitors, then the equations involve time derivatives.

Example

SIMPLE TWO-MESH RESISTIVE CIRCUIT WITH A VOLTAGE SOURCE

Refer to Fig. 7-5. Choosing mesh currents as in the figure, we apply the KVL to the two meshes, starting at *B* in both cases. For mesh *a*,

$$V - I_a R - (I_a - I_b)R = 0. \quad (7-1)$$

For mesh *b*,

$$-(I_b - I_a)R - 2I_b R = 0. \quad (7-2)$$

Solving, we find that

[†]Hysteresis is a property of certain devices or media whose parameters depend on their previous history. Ferroelectric media are hysteretic (Sec. 10.1.5).

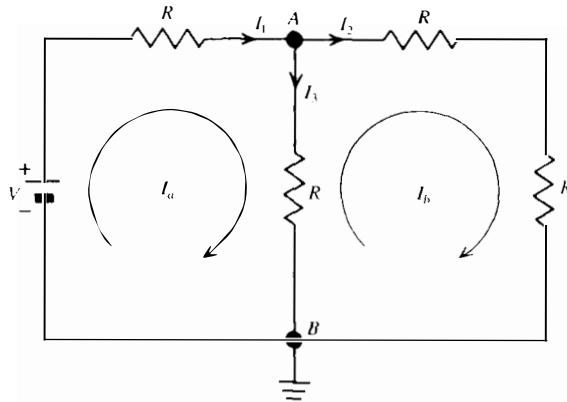


Fig. 7-5. Two-mesh resistive circuit fed by a source supplying a steady voltage V .

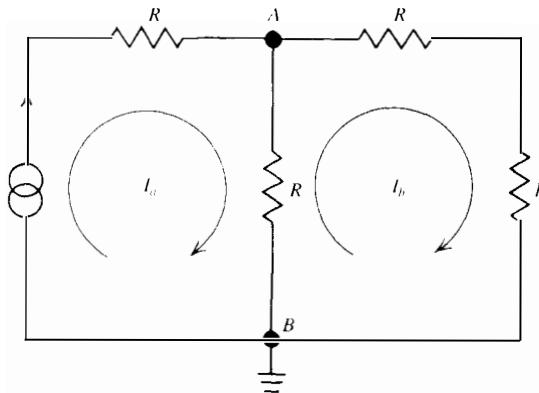


Fig. 7-6. Two-mesh resistive circuit fed by a source supplying a steady current I_a .

$$I_a = \frac{0.6V}{R}, \quad I_b = \frac{0.2V}{R}. \quad (7-3)$$

The downward current in branch AB is $0.4V/R$ amperes, and thus the voltage at A is $0.4V$ volts.

Example

SIMPLE TWO-MESH RESISTIVE CIRCUIT WITH A CURRENT SOURCE

We now have the circuit of Fig. 7-6, with a current source substituted for the voltage source of Fig. 7-5. Now the current I_a is known, but the voltage V across the source is unknown. Equations 7-1 and 7-2 still apply. Solving, we find that

$$V = \frac{5}{3} I_a R, \quad I_b = \frac{I_a}{3}. \quad (7-4)$$

Example

SIMPLE TWO-MESH RC CIRCUIT WITH A VOLTAGE SOURCE. SOLVING AN ORDINARY DIFFERENTIAL EQUATION WITH CONSTANT COEFFICIENTS

The circuit of Fig. 7-7 is similar to that of Fig. 7-5, except that we have added a switch and substituted a capacitor for one of the resistors. The switch closes at $t = 0$. The mesh equations are now as follows for $t \geq 0$:

$$V - I_a R - (I_a - I_b)R = 0, \quad (7-5)$$

$$-(I_b - I_a)R - I_b R - \frac{Q}{C} = 0. \quad (7-6)$$

Rewriting, with $I_b = dQ/dt$ if I_b and Q are chosen as in the figure,

$$2I_a R - R \frac{dQ}{dt} = V, \quad (7-7)$$

$$I_a R - 2R \frac{dQ}{dt} - \frac{Q}{C} = 0. \quad (7-8)$$

Eliminating now I_a yields

$$3R \frac{dQ}{dt} + 2\frac{Q}{C} = V. \quad (7-9)$$

This is an *ordinary differential equation with constant coefficients*. The term *ordinary* refers to the fact that the equation comprises ordinary, and not partial, derivatives. The equation is *nonhomogeneous* because its right-hand side is not a function of Q . It is *linear* because it comprises only the first powers of the variable Q and of its derivatives and no products such as Q^2 or $Q(dQ/dt)$.

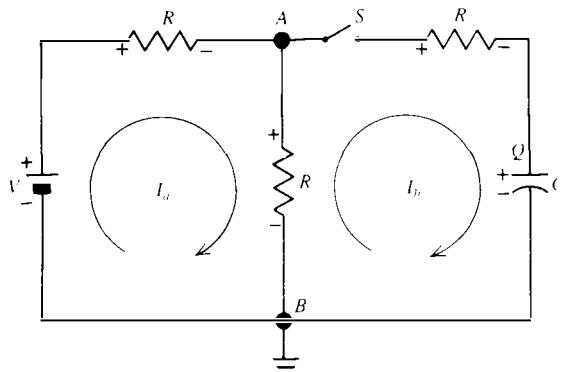


Fig. 7-7. *RC* circuit fed by a voltage source. Switch S remains open until $t = 0$.

As we shall see, there exist an infinite number of solutions $Q(t)$ that satisfy this equation. However, it is a simple matter to select the correct one, in a particular situation.

We first find one obvious solution that we call Q_p :

$$Q_p = \frac{CV}{2}. \quad (7-10)$$

You can easily check that this is a solution, by substitution. This is the *particular solution*. We found it by conserving only the last term on the left.

Now let Q be some other solution. By definition, it also satisfies Eq. 7-9. Then the function

$$Q_c = Q - Q_p \quad (7-11)$$

satisfies the homogeneous equation

$$3R \frac{dQ_c}{dt} + 2 \frac{Q_c}{C} = 0. \quad (7-12)$$

The function Q_c is the *complementary solution*.

In other words, the general solution Q is the sum of Q_p and Q_c :

$$Q = Q_p + Q_c. \quad (7-13)$$

Now the solution of Eq. 7-12 is simple. Rewriting, we get

$$\frac{dQ_c}{dt} = -\frac{2}{3RC} Q_c \quad (7-14)$$

and

$$Q_c = A \exp\left(-\frac{2t}{3RC}\right), \quad (7-15)$$

where A is a *constant of integration*.

Thus

$$Q = \frac{CV}{2} + A \exp\left(-\frac{2t}{3RC}\right). \quad (7-16)$$

Since A is any constant, there exists an infinite number of solutions. For given conditions, the value of A follows from, say, the value of Q at $t = 0$. In this example, $Q = 0$ at $t = 0$, and $A = -CV/2$. Thus

$$Q = \frac{CV}{2} \left[1 - \exp\left(-\frac{2t}{3RC}\right) \right]. \quad (7-17)$$

Figure 7-8(a) shows Q as a function of t for $CV = 2$ and $RC = 2$. For $t \gg 3RC/2$, $Q \approx 1$. Then $I_b = 0$ and the voltage across the capacitor is equal to the voltage between A and ground. Figure 7-8(b) shows $I_b = dQ/dt$ as a function of t .

We have here a circuit that can exist in two states, first with S open and then with S closed. With S open and until $t = 0$, $I_a = V/2R$, $I_b = 0$, $Q = 0$. With S closed and $t \gg 3RC/2$, the same

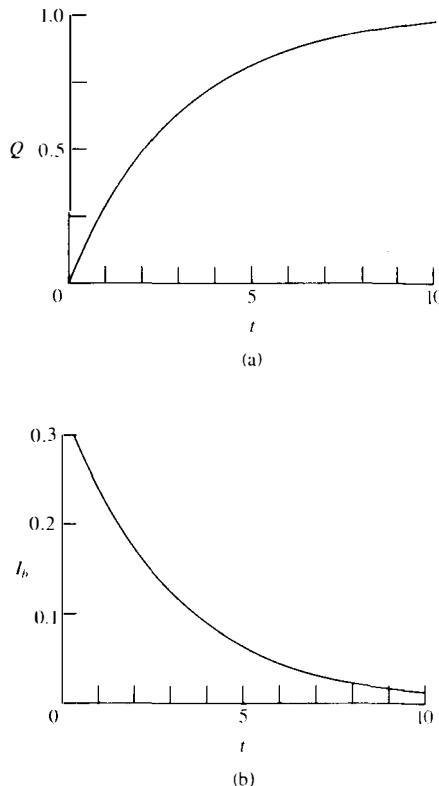


Fig. 7-8. (a) The charge Q on the capacitor of Fig. 7-7, as a function of the time, for $CV = 2$, $RC = 2$. (b) The current in mesh b as a function of the time.

values of I_a and I_b apply, but $Q = CV/2$. In the interval between $t = 0$ and $t \gg 3RC/2$, the circuit adapts itself to the new situation. Phenomena that occur during such periods of adaptation are called *transients*.

7.2.2 The Node Method for Calculating Node Voltages

The *node method* is an alternative to the mesh method. As a rule, one of the two methods requires a smaller number of simultaneous equations and is therefore preferable.

We assume that one of the nodes is at ground potential ($V = 0$) and that the other *node voltages* V_a , V_b , V_c , ... are unknown. Application of the KCL at each node in succession yields the node voltages and thus the branch currents.

We illustrate the node method with a single example.

Example

SIMPLE TWO-MESH RESISTIVE CIRCUIT WITH A VOLTAGE SOURCE

Refer again to Fig. 7-5. According to the KCL, $I_1 = I_2 + I_3$, with

$$I_1 = \frac{V - V_A}{R}, \quad I_2 = \frac{V_A}{2R}, \quad I_3 = \frac{V_A}{R}. \quad (7-18)$$

Therefore

$$\frac{V - V_A}{R} = \frac{V_A}{2R} + \frac{V_A}{R}. \quad (7-19)$$

Solving, we find that

$$V_A = 0.4V, \quad (7-20)$$

as in the first example in Sec. 7.2.1. The values of the branch currents I_1 , I_2 , I_3 follow immediately.

Observe that with the node method we had a single equation to solve, namely Eq. 7-19, versus two with the mesh method that we used for the same circuit in the first example in Sec. 7.2.1. With this particular circuit, the node method is simpler.

7.3 ALTERNATING CURRENTS

Most electric and magnetic devices operate with either fluctuating or alternating currents. There are many reasons for this, but the two major ones have to do with power technology and with the transmission and processing of information.

First, with alternating currents, the electric power supplied by a source at a given voltage can be made available at almost any other convenient voltage by means of transformers. This makes electric power adaptable to a broad variety of uses. (The power supplied by a *direct-current* source can also be changed from one voltage to another. This is done by first switching the current periodically, to obtain an alternating current, next feeding this to a transformer, and then rectifying and filtering the output. The operation is relatively costly.)

The second reason for using alternating, or fluctuating, currents is that they can transmit information. For example, a microphone transforms the information contained in a spoken word into a complex fluctuating current.

We assume that the voltages, currents, and charges are all cosine functions of the time, with appropriate phases. This is not always the case. The current through a microphone is not normally sinusoidal. Or one might have only the positive part of the cosine function or a cosine function whose amplitude is a function of the time. However, any

periodic function is the sum of an infinite series of sine and cosine terms called a Fourier series (Sec. 11.3, example).

7.3.1 Three-Wire Single-Phase Alternating Currents

In North America electric utilities supply electric power to individual houses and apartments at both 120 and 240 volts rms, 60 hertz, as in Fig. 7-9. This is the *three-wire single-phase alternating-current* system. Thus

$$V_A = 120 \times 1.414 \cos(2\pi \times 60t) = 170 \cos(2\pi \times 60t), \quad (7-21)$$

$$V_C = -120 \times 1.414 \cos(2\pi \times 60t) = -170 \cos(2\pi \times 60t), \quad (7-22)$$

$$V_A - V_C = 240 \times 1.414 \cos(2\pi \times 60t) = 340 \cos(2\pi \times 60t). \quad (7-23)$$

Low-power devices operate at 120 volts rms across *AD* or *CD*, while high-power devices such as water heaters operate at 240 volts rms across *A* and *C*.

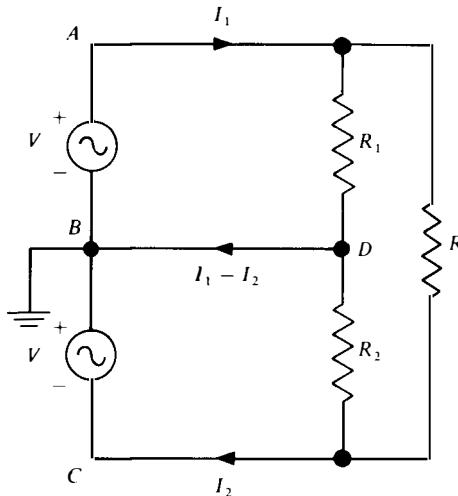


Fig. 7-9. Three-wire single-phase ac system. The plus and minus signs mean that the two sources have opposite phases. If $R_1 = R_2$, the connection *BD* is unnecessary.

7.3.2 Three-Phase Alternating Current

Figure 7-10(a) shows three sources of alternating current feeding resistors R_1 , R_2 , R_3 . We have oriented the sources 120° apart on the figure so as to

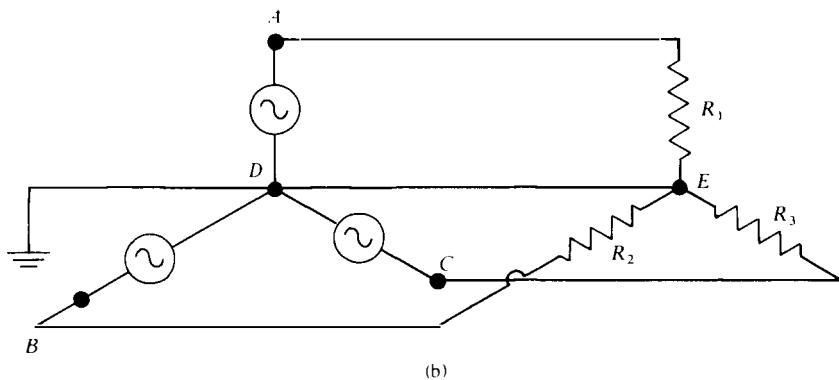
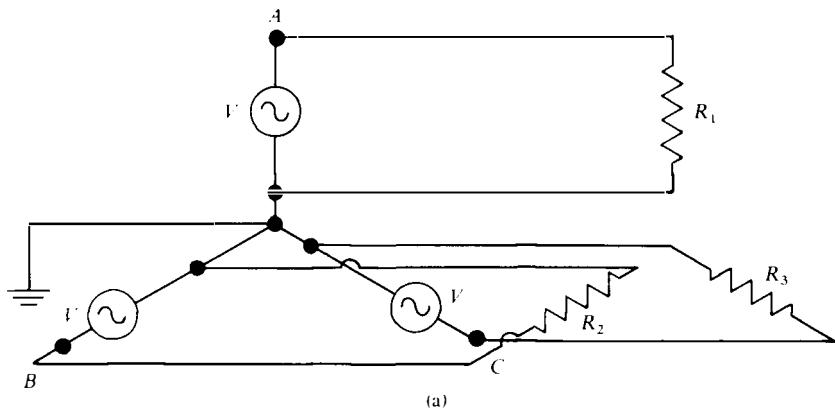


Fig. 7-10. (a) Three identical sources of alternating current, with phase differences of 120° , feeding three resistors R_1 , R_2 , R_3 . (b) Three-phase supply. If the resistances are equal, the connection DE is unnecessary.

reflect their relative phases. For example, if the phase at A is zero at $t = 0$, then

$$V_A = V_m \cos \omega t, \quad (7-24)$$

$$V_B = V_m \cos \left(\omega t + \frac{2\pi}{3} \right), \quad (7-25)$$

$$V_C = V_m \cos \left(\omega t + \frac{4\pi}{3} \right). \quad (7-26)$$

If the three resistances are equal, the sum of the currents flowing in the three grounded wires is

$$\frac{V_m}{R} \left[\cos \omega t + \cos \left(\omega t + \frac{2\pi}{3} \right) + \cos \left(\omega t + \frac{4\pi}{3} \right) \right] = 0. \quad (7-27)$$

Then the wire *DE* in Fig. 7-10(b) can be dispensed with. If the resistances are unequal, the currents do not completely cancel in *DE*. We now have four wires doing the work of six, with low I^2R losses in the wire *DE*.

A set of three sources, star-connected and phased as in Fig. 7-10(b), supplies *three-phase alternating current*. The main advantage of three-phase current is that it can generate the revolving magnetic fields of large electric motors. See Prob. 18-6.

Electric power stations usually generate three-phase alternating current. That is why ordinary high-voltage transmission lines have either three or six wires, plus one or two light wires. Three-wire single-phase power is obtained from a transformer whose primary is connected between one phase of a three-phase line and ground.

Except for a few problems at the end of this chapter, we shall be concerned henceforth solely with single-phase currents.

7.4 ALTERNATING CURRENTS IN RESISTORS

Figure 7-11(a) shows an alternating voltage source connected to a resistor. Let

$$V = V_m \cos \omega t, \quad (7-28)$$

where V_m is the *maximum*, or *peak*, voltage; $\omega = 2\pi f$ is the *circular frequency*, expressed in radians per second; and f is the *frequency* expressed in hertz. In the figure,

$$I = \frac{V_m \cos \omega t}{R} = I_m \cos \omega t, \quad I_m = \frac{V_m}{R}. \quad (7-29)$$

The current is in phase with the applied voltage, as in Fig. 7-11(b).

At any given instant the power dissipated in the resistor is

$$\begin{aligned} P_{\text{inst}} &= VI = (V_m \cos \omega t)(I_m \cos \omega t) = V_m I_m \cos^2 \omega t \\ &= \frac{V_m^2}{R} \cos^2 \omega t = I_m^2 R \cos^2 \omega t. \end{aligned} \quad (7-30)$$

Since the average value of $\cos^2 \omega t$ is $\frac{1}{2}$, the average power dissipation is

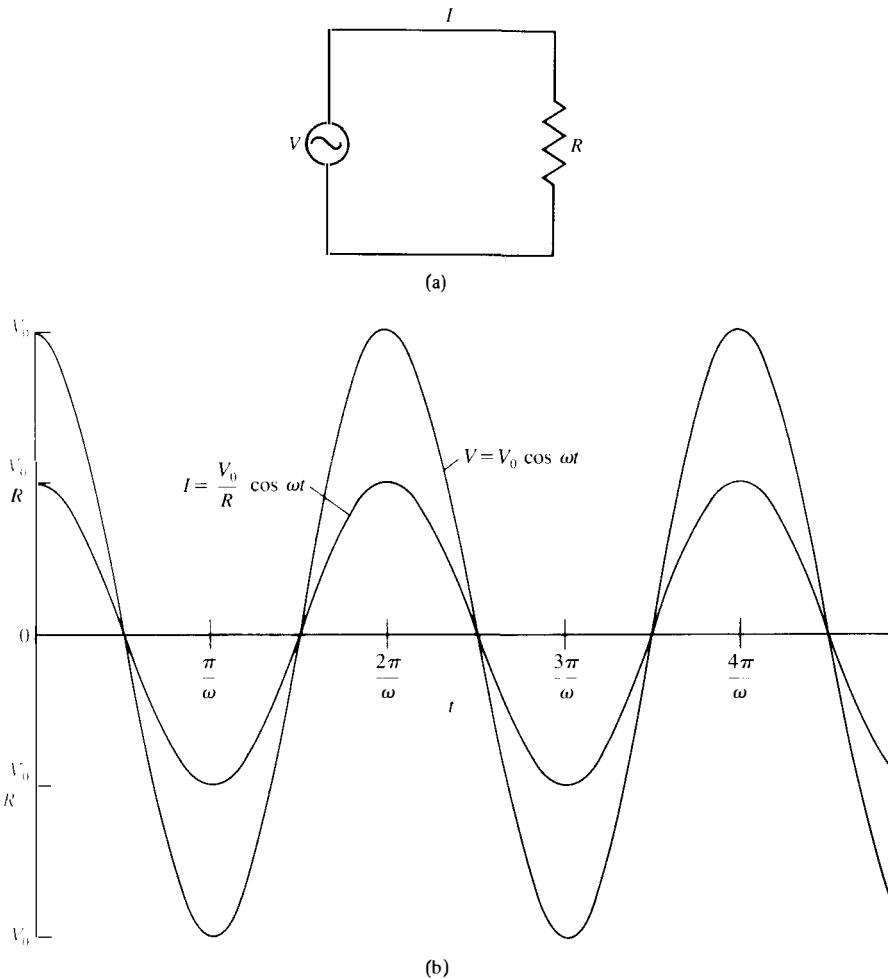


Fig. 7-11. (a) Resistor R connected to a source supplying a voltage V and a current I . (b) V and I plotted as functions of t for a resistor.

$$P_{av} = \frac{V_m I_m}{2} = \frac{V_m^2}{2R} = \frac{I_m^2 R}{2}, \quad (7-31)$$

$$= V_{rms} I_{rms} = \frac{V_{rms}^2}{R} = I_{rms}^2 R, \quad (7-32)$$

as in Fig. 7-12.

Recall that the subscript “rms” stands for root mean square, or the square root of the mean value of the square (Sec. 2.4). In practice, the subscripts “av” and “rms” prove to be unnecessary, and we simply write

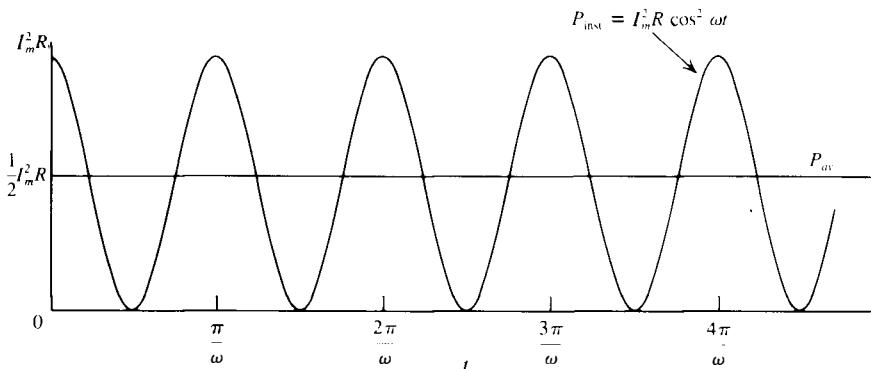


Fig. 7-12. Instantaneous and average power dissipated in a resistor.

$$I = \frac{V}{R}, \quad P = VI = \frac{V^2}{R} = I^2 R, \quad (7-33)$$

as with direct currents. We shall nonetheless use the subscript “rms” wherever it applies.

Unless stated explicitly otherwise, all voltmeters and ammeters show rms values.

If the voltage or current is an alternating but not a sinusoidal function of the time, then the ratio V_{rms}/V_m depends on the nature of the function. For example, if a voltage simply alternates between $+V_m$ and $-V_m$, the rms value is V_m .

7.5 ALTERNATING CURRENTS IN CAPACITORS

Figure 7-13(a) shows a source of alternating current connected to a capacitor. Then

$$Q = CV_m \cos \omega t, \quad (7-34)$$

$$I = \frac{dQ}{dt} = -\omega CV_m \sin \omega t = \omega CV_m \cos \left(\omega t + \frac{\pi}{2} \right). \quad (7-35)$$

The current *leads* the voltage by $\pi/2$ radians, as in Fig. 7-13(b).

In phasor notation (Chap. 2),

$$I = j\omega Q = j\omega CV, \quad (7-36)$$

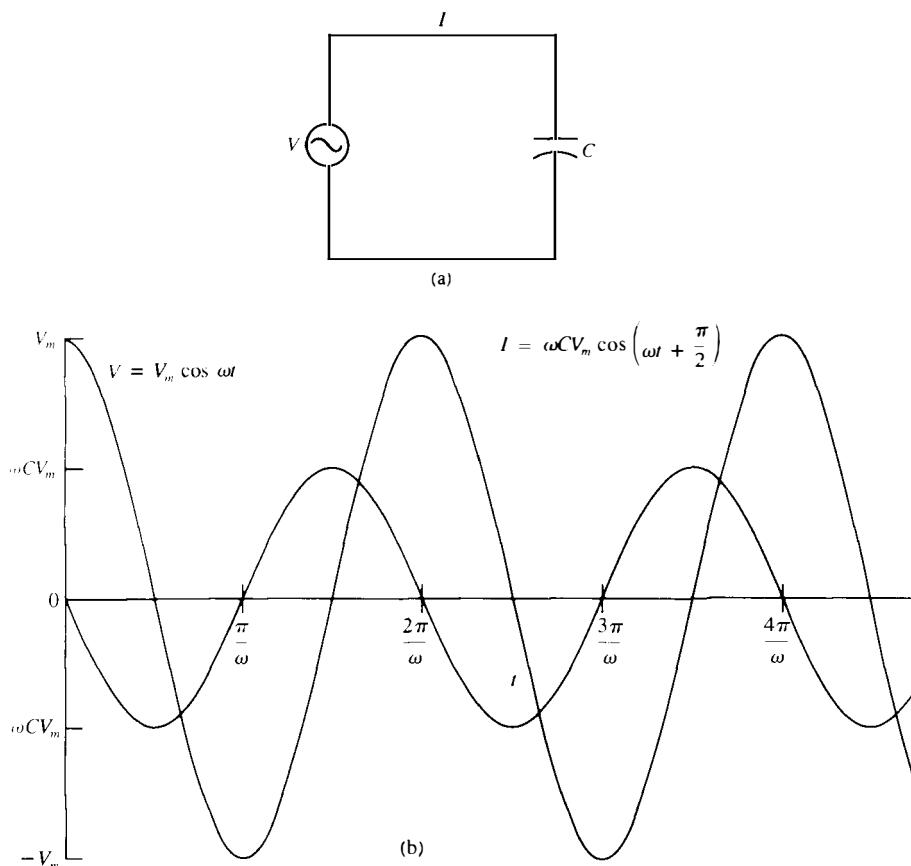


Fig. 7-13. (a) Capacitor C connected to a source supplying a voltage V . The circled V and I represent a voltmeter and an ammeter. (b) V and I as functions of the time for a capacitor. The current leads the voltage by 90° and $\omega C = 0.5$, here.

where I, Q, V are all phasors. At a given instant the energy stored in the capacitor is

$$\mathcal{E}_{\text{inst}} = \frac{CV^2}{2} = \frac{CV_m^2 \cos^2 \omega t}{2} \quad (7-37)$$

and the average stored energy is

$$\mathcal{E}_{\text{av}} = \frac{CV_m^2/2}{2} = \frac{CV_{\text{rms}}^2}{2}. \quad (7-38)$$

7.6 THE IMPEDANCE Z AND THE ADMITTANCE Y

By analogy with Ohm's law for dc circuits, we have *Ohm's law for ac circuits*:

$$I = \frac{V}{Z}, \quad (7-39)$$

where both I and V are phasors and Z is the *impedance* of the circuit. The impedance of a resistance is R .

For a capacitor, from Eq. 7-36,

$$Z = \frac{1}{j\omega C}. \quad (7-40)$$

Impedances in series and in parallel operate like resistances in series and in parallel.

Impedances are in general complex and thus of the form

$$Z = R + jX, \quad (7-41)$$

where X is the *reactance*. Given a two-terminal passive and linear circuit that comprises only resistors and capacitors, however complex, the impedance between its terminals is a complex number $R + jX$.

This reactance X is negative for the following reason. The reactance X of a single capacitance C is negative:

$$jX = \frac{1}{j\omega C}, \quad X = -\frac{1}{\omega C}. \quad (7-42)$$

Now the current in a capacitor leads the applied voltage, and intuitively, you can guess that, whatever the arrangement of resistors and capacitors, the current at the input terminals will lead the voltage. Thus the reactance for the complete circuit will be negative, or capacitive.

Thus the circuit of Fig. 7-14(a) is equivalent to that of Fig. 7-14(b) with

$$R_s = R, \quad C_s = -\frac{1}{\omega X}. \quad (7-43)$$

The inverse of an impedance is an *admittance*:

$$Y = \frac{1}{Z} = \frac{1}{R + jX} = G + jB, \quad (7-44)$$

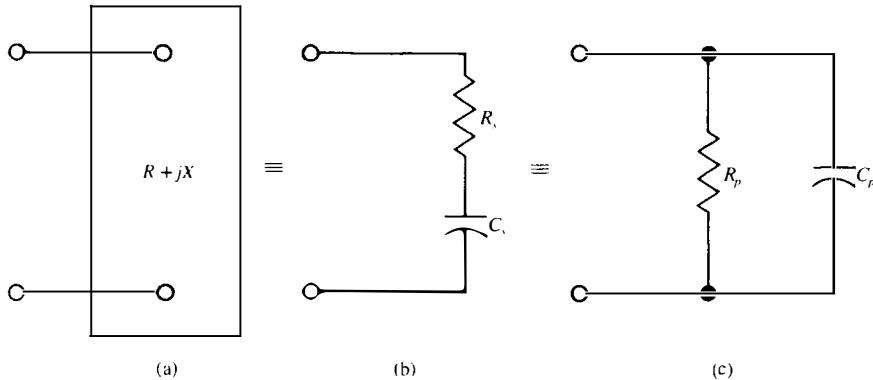


Fig. 7-14. (a) Any passive and linear circuit comprising only resistors and capacitors. The impedance at the input terminals is $R + jX$. (b) Equivalent series circuit. (c) Equivalent parallel circuit.

where G is the *conductance* and B is the *susceptance*. The admittance Y of a capacitor is $j\omega C$, and its susceptance B is ωC .

Any two-terminal passive and linear RC circuit is equivalent to a conductance in parallel with a susceptance as in Fig. 7-14(c). You can easily show that

$$G = \frac{R}{R^2 + X^2} = \frac{R}{Z^2}, \quad B = -\frac{X}{R^2 + X^2} = -\frac{X}{Z^2}, \quad (7-45)$$

$$R = \frac{G}{G^2 + B^2} = \frac{G}{Y^2}, \quad X = -\frac{B}{G^2 + B^2} = -\frac{B}{Y^2}. \quad (7-46)$$

The quantity G is positive, like R , while B is either positive or negative, like X .

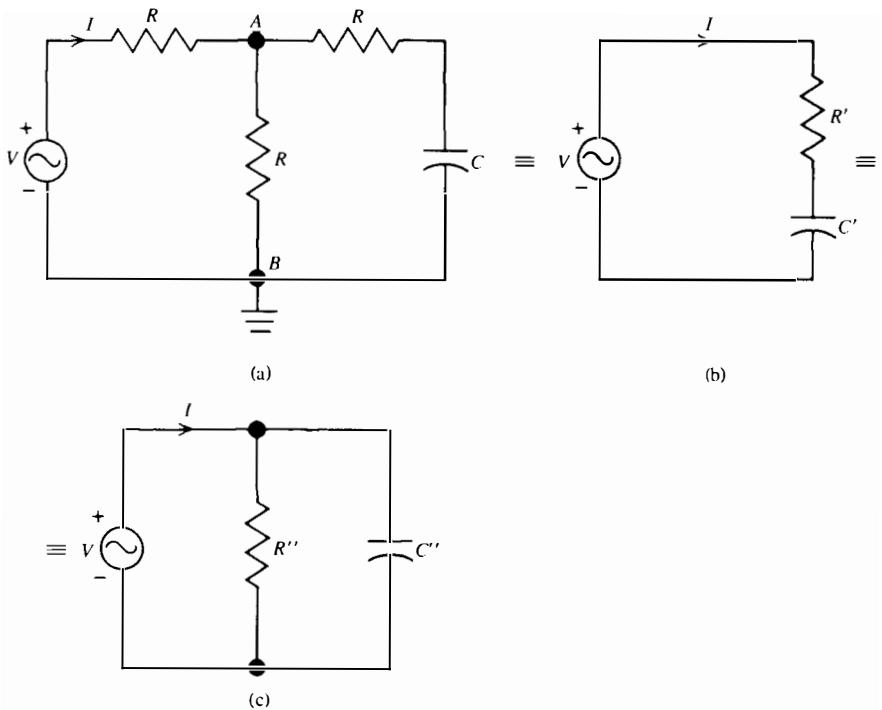
If a given circuit comprises only resistors and capacitors, then X is negative and B is positive. Also,

$$R_p = \frac{1}{G} = \frac{R^2 + X^2}{R}, \quad (7-47)$$

$$V = \omega C_p, \quad C_p = \frac{B}{\omega} = -\frac{X}{\omega(R^2 + X^2)}. \quad (7-48)$$

Example | SIMPLE RC CIRCUIT

In the circuit of Fig. 7-15(a) the source feeds an impedance



F g. 7-15. (a) Simple RC circuit. (b) Equivalent series circuit. (c) Equivalent parallel circuit.

$$Z = R + \frac{R[R + 1/(j\omega C)]}{2R + 1/(j\omega C)} = R \frac{2 + 3Rj\omega C}{1 + 2Rj\omega C} \quad (7-49)$$

$$= R \frac{(2 + 3Rj\omega C)(1 - 2Rj\omega C)}{1 + 4R^2\omega^2C^2} = R \frac{(2 + 6R^2\omega^2C^2) - Rj\omega C}{1 + 4R^2\omega^2C^2}. \quad (7-50)$$

In effect, this means that the circuit of Fig. 7-15(a) has the same impedance as that of Fig. 7-15(b), where the real part of Z ,

$$R' = 2R \frac{1 + 3R^2\omega^2C^2}{1 + 4R^2\omega^2C^2}, \quad (7-51)$$

is frequency-dependent. The value of C is given by the imaginary part of Z :

$$\frac{1}{j\omega C'} = \frac{-R^2 j\omega C}{1 + 4R^2 \omega^2 C^2}, \quad (7-52)$$

$$C' = \frac{1 + 4R^2\omega^2 C^2}{R^2\omega^2 C} = \frac{1}{R^2\omega^2 C} + 4C. \quad (7-53)$$

The value of C' is also frequency-dependent. Indeed, the

components of such equivalent circuits are almost invariably frequency-dependent.

The circuit of Fig. 7-15(a) is also equivalent to that of Fig. 7-15(c), with G' and B' given by Eq. 7-45:

$$G' = \frac{R'}{R'^2 + X'^2} = \frac{R'}{Z'^2}, \quad B' = -\frac{X'}{R'^2 + X'^2} = -\frac{X'}{Z'^2}, \quad (7-54)$$

where $jX' = 1/(j\omega C')$ and $X' = -1/(\omega C')$. After simplification, we find that

$$Z^2 = \frac{4 + 9R^2\omega^2C^2}{1 + 4R^2\omega^2C^2} R^2. \quad (7-55)$$

Then

$$G' = \frac{2}{R} \frac{1 + 3R^2\omega^2C^2}{4 + 9R^2\omega^2C^2}, \quad B' = \frac{\omega C}{4 + 9R^2\omega^2C^2}. \quad (7-56)$$

In Fig. 7-15(c), the resistor has a conductance G' and hence a resistance

$$R'' = \frac{1}{G'}. \quad (7-57)$$

As to the capacitance,

$$jB' = j\omega C'', \quad C'' = \frac{B'}{\omega}. \quad (7-58)$$

Let $R = 1000$ ohms, $\omega = 500$ radians/second, $C = 1$ microfarad. Then, from Eq. 7-50,

$$Z = 1000 \frac{(2 + 6 \times 10^6 \times 500^2 \times 10^{-12}) - j10^3 \times 500 \times 10^{-6}}{1 + 4 \times 10^6 \times 500^2 \times 10^{-12}} \quad (7-59)$$

$$= 1750 - 250j \quad (7-60)$$

$$\begin{aligned} &= 1768 \exp \left[j \arctan \left(-\frac{250}{1750} \right) \right] \\ &= 1768 \exp (-0.142j) \quad \text{ohms.} \end{aligned} \quad (7-61)$$

Angles are expressed in radians. If V is 10 volts rms,

$$I = \frac{10}{1768} \exp (0.142j) = 5.66 \exp (0.142j) \quad \text{milliamperes rms.} \quad (7-62)$$

The current leads the voltage by 0.142 radian, or by 8.13 degrees. The instantaneous value of the current is

$$\begin{aligned} I &= 2^{1/2} \times 5.66 \cos (500t + 0.142) \\ &= 8.00 \cos (500t + 0.142) \quad \text{milliamperes.} \end{aligned} \quad (7-63)$$

7.7 POWER IN AC CIRCUITS

Suppose a resistance R is in series with a capacitor C . One applies a voltage $V_m \exp j\omega t$ to the combination. Then

$$I = \frac{V_m \exp j\omega t}{Z} = I_m \exp j(\omega t - \theta), \quad (7-64)$$

where

$$Z = R + \frac{1}{j\omega C} = \left(R^2 + \frac{1}{\omega^2 C^2} \right)^{1/2} \exp j\theta, \quad \theta = \arctan \left(-\frac{1}{R\omega C} \right), \quad (7-65)$$

$$I_m = \frac{V_m}{[R^2 + 1/(\omega^2 C^2)]^{1/2}}. \quad (7-66)$$

The average power dissipation is

$$P_{av} = \frac{1}{2} V_m I_m \cos \theta = V_{rms} I_{rms} \cos \theta, \quad (7-67)$$

from Sec. 2.4. The term $\cos \theta = R/|Z|$ is the *power factor* of the impedance Z .

According to Sec. 2.4, we can also write

$$P_{av} = \frac{1}{2} \operatorname{Re} VI^*, \quad (7-68)$$

where the operator Re means “real part of” and where I^* is the complex conjugate of I (Sec. 2.1). One occasionally writes that

$$\frac{1}{2} VI^* = P_{av} + jQ_{av}, \quad (7-69)$$

where the term on the left is the *complex power*, P_{av} is the *time-averaged power*, and Q_{av} is the *time-averaged reactive power*.

7.8 KIRCHHOFF'S LAWS FOR ALTERNATING-CURRENT CIRCUITS

Both the Kirchhoff current law and the Kirchhoff voltage law apply to alternating-current circuits if one uses phasors and impedances.

Example

SIMPLE RC CIRCUIT SOLVED BY BOTH THE MESH AND NODE METHODS

(a) Referring to Fig. 7-16, we use mesh currents and apply the KVL as if we had direct currents:

$$V - I_a R - (I_a - I_b)R = 0, \quad (7-70)$$

$$-(I_b - I_a)R - I_b R - \frac{I_b}{j\omega C} = 0. \quad (7-71)$$

Solving yields

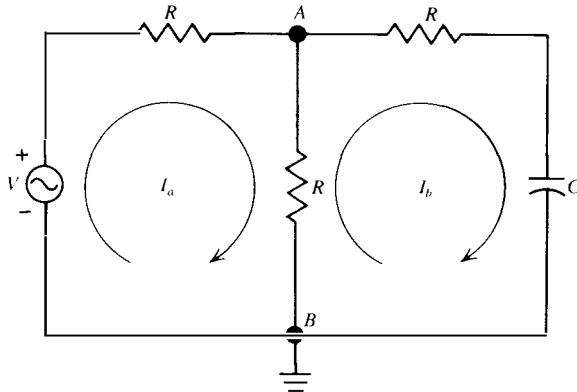


Fig. 7-16. Simple RC circuit.

$$I_a = V \left[\frac{1}{2R} + \frac{1}{6R + 4/(j\omega C)} \right], \quad I_b = \frac{V}{3R + 2/(j\omega C)}. \quad (7-72)$$

At A ,

$$V_A = V - I_a R = V \frac{1 + Rj\omega C}{2 + 3Rj\omega C}. \quad (7-73)$$

The ratio V/I_a should be equal to the impedance Z that we calculated in the example in Sec. 7.6. Let us check:

$$\frac{V}{I_a} = \frac{2R[6R + 4/(j\omega C)]}{6R + 4/(j\omega C) + 2R} = \frac{2R(6Rj\omega C + 4)}{8Rj\omega C + 4} = R \frac{2 + 3Rj\omega C}{1 + 2Rj\omega C}, \quad (7-74)$$

as previously.

(b) With the node method, we apply the KCL at node A , where we have, as with a dc circuit,

$$I_1 = I_2 + I_3, \quad I_1 = I_a, \quad I_2 = I_b, \quad I_3 = I_a - I_b, \quad (7-75)$$

with

$$I_1 = \frac{V - V_A}{R}, \quad I_2 = \frac{V_A}{R + 1/(j\omega C)}, \quad I_3 = \frac{V_A}{R}. \quad (7-76)$$

Thus

$$\frac{V - V_A}{R} = \frac{V_A}{R + 1/(j\omega C)} + \frac{V_A}{R}, \quad (7-77)$$

$$V_A = \frac{V}{2 + Rj\omega C/(1 + Rj\omega C)} = \frac{1 + Rj\omega C}{2 + 3Rj\omega C} V. \quad (7-78)$$

7.9 SUMMARY

Kirchhoff's current law (KCL) states that the algebraic sum of the currents entering a node is equal to zero. According to *Kirchhoff's*

voltage law (KVL), the sum of the voltage drops around a mesh is equal to zero.

These two laws serve to calculate voltages and currents at various points in a circuit, given the sources and the impedances. With the *mesh method* we assign mesh currents I_a, I_b, I_c, \dots to each mesh. We then apply the KVL to each mesh in succession to find the mesh currents. With the *node method* we assign node voltages V_a, V_b, V_c, \dots to each node. One of the nodes is set at $V = 0$. Then we apply the KCL to each node in succession to obtain the node voltages.

If the circuit is purely resistive, then either method gives a set of simultaneous algebraic equations. If the circuit comprises both resistors and capacitors, we obtain a set of simultaneous differential equations.

The *alternating current in a true resistor* is in phase with the applied voltage. The time-average dissipated power is

$$P_{av} = \frac{I_m^2 R}{2} = I_{rms}^2 R, \quad (7-31), (7-32)$$

where I_m is the maximum value of I and I_{rms} is the root mean square value of I . In practice, the subscripts "av" and "rms" are unnecessary.

The *alternating current in a capacitor* leads the applied voltage by 90° :

$$I = j\omega CV. \quad (7-36)$$

The time-averaged stored energy is $CV_{rms}^2/2$.

More generally,

$$I = \frac{V}{Z} \quad (7-39)$$

where Z is the *impedance* of the device across which the voltage is V . For a resistor, $Z = R$. For a capacitor, $Z = 1/(j\omega C)$. Impedances are usually complex:

$$Z = R + jX, \quad (7-41)$$

where X is the *reactance* of the circuit. The inverse of an impedance is an *admittance*:

$$Y = \frac{1}{Z}. \quad (7-44)$$

The *time-averaged power* dissipated in a circuit is

$$P_{av} = V_{rms} I_{rms} \cos \theta, \quad (7-67)$$

where $\cos \theta$ is the *power factor* $R/|Z|$ of Z . Also

$$P_{av} = \frac{1}{2} \operatorname{Re} VI^*. \quad (7-68)$$

Kirchhoff's current and voltage laws apply to alternating-current circuits.

PROBLEMS

7-1. (7.2) The potentiometer

Figure 7-17 shows a potentiometer circuit. Show that, when $I = 0$,

$$V_o = \frac{R_2}{R_1 + R_2} V_i.$$

This is a common type of circuit. It serves to measure a voltage, in this case V_b , without drawing current. In curve plotters the current I , after amplification, actuates a motor that displaces the pen and simultaneously moves the tap in the direction that decreases I . The resistances R_1 and R_2 act as a *potential divider*.

7-2. (7.2) The operational amplifier

Figure 7-18(a) shows a common type of amplifier.

The triangular figure is an *operational amplifier* whose gain is $-A$. Such amplifiers have gains of the order of 10^4 to 10^9 and draw a negligible amount of current at their input terminals. The accuracy of the gain of this circuit is limited only by the stability of the ratio R_2/R_1 . The drift in R_2/R_1 due to aging, temperature changes, and so forth is normally smaller than the drift in A by orders of magnitude.

As a first approximation, (a) the operational amplifier draws zero current, and the same current I flows in R_1 and in R_2 , (b) A is infinite, and the potential at the junction between R_1 and R_2 is therefore zero. Then $I \approx V_i/R_1 \approx -V_o/R_2$, and the gain is about $-R_2/R_1$.

(a) Find a more accurate expression for the gain V_o/V_i . You can take into account the fact that the gain is not infinite by setting $V_o = -AV_{iA}$ in

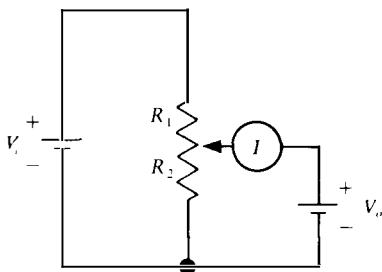


Fig. 7-17.

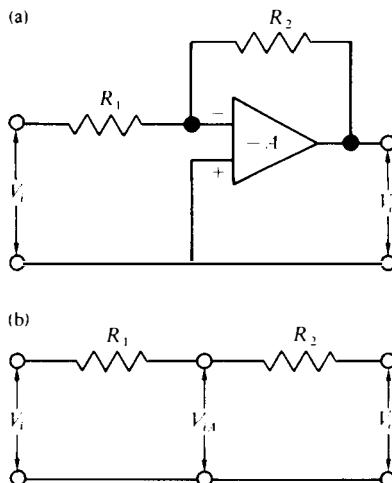


Fig. 7-18.

Fig. 7-18(b). The approximation $-R_2/R_1$ is valid only if $A \gg 1$ and if $A \gg R_2/R_1$.

(b) What is the minimum value of A if $R_1 = 1000$ ohms, $R_2 = 2000$ ohms, and the circuit gain must be equal to 2 within 0.1%?

7-3. (7.2) The $R/2R$ ladder network

Figure 7-19 shows a so-called *$R/2R$ ladder network* that serves for both digital-to-analog (D/A) and analog-to-digital (A/D) conversion. See Prob. 7-4.

Show that the resistance between A and ground, called the *input resistance*, is $2R$, whatever be the number of stages.

7-4. (7.2) Digital-to-analog conversion

The $R/2R$ ladder network of Prob. 7-3 can convert a binary† number to an analog voltage. The usual convention is that of *positive logic*, in which

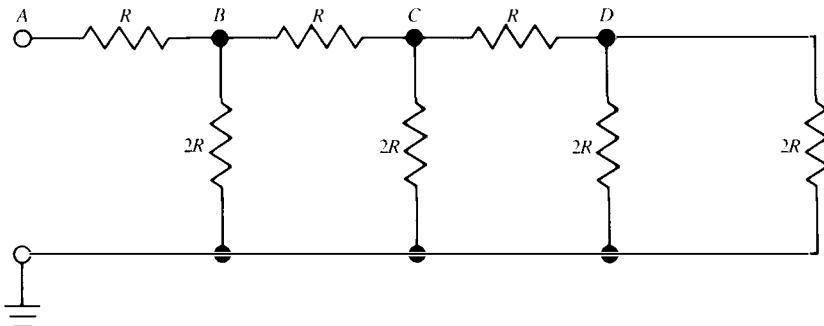


Fig. 7-19.

† A binary number uses only the digits 0 and 1: one = 01, two = 10, three = 11, four = 100, five = 101, etc.

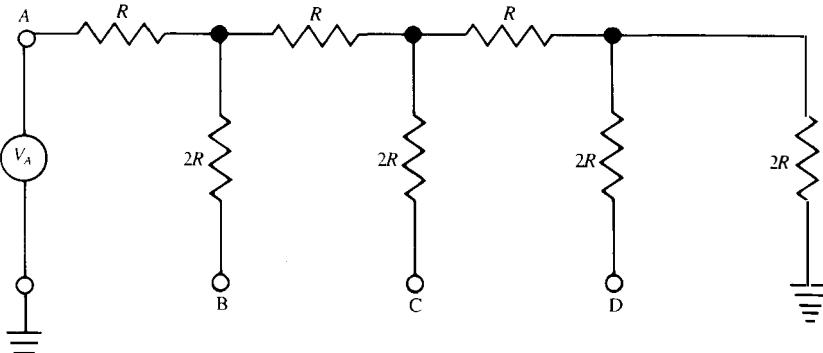


Fig. 7-20.

$V = 0$ means 0 and $V = V'$ means 1, where V' is a precisely regulated positive voltage.

Refer to Fig. 7-20. Say $V' = 1$ volt. The voltmeter connected between A and ground draws essentially zero current. The binary number enters at the terminals B, C, D, \dots and the analog number appears at A . For example, if the number is 4, or 100, $V_B = 1, V_C = 0, V_D = 0$.

Use the result of Prob. 7-3 to find the value of V_A when

- $V_B = 1, V_C = 0$, and $V_D = 0$;
- $V_B = 0, V_C = 1$, and $V_D = 0$;
- $V_B = 0, V_C = 0$, and $V_D = 1$.

7-5. (7.2) Charging a capacitor through a resistor

A source charges a capacitor C through a resistor R to a voltage V . Calculate the energy supplied by the source, that dissipated by the resistor, and that stored in the capacitor, after an infinite time.

You should find that the resistor dissipates half the energy and that the capacitor stores the other half.

7-6. (7.2) RC differentiating circuit

Figure 7-21 shows an RC differentiating circuit. The load resistance connected at V_o is large compared to R .

(a) Show that, if the voltage drop across R is negligible compared to that across C , then

$$V_o \approx RC \frac{dV_i}{dt}.$$

- (b) The input is a square wave. Sketch $V_o(t)$.

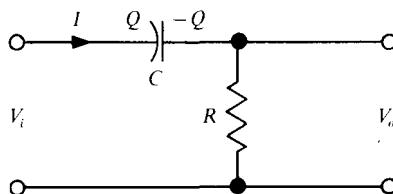


Fig. 7-21.

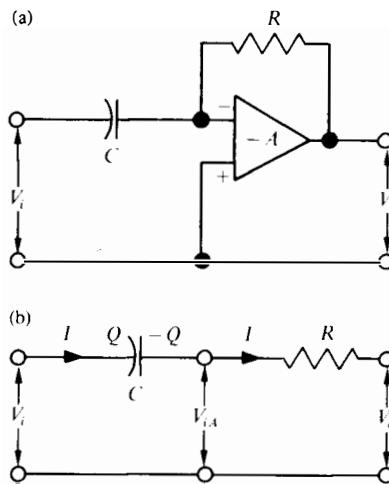


Fig. 7-22.

7-7. (7.2) Differentiating circuit using an operational amplifier

The circuit of Fig. 7-21(a) is simple and inexpensive, but $V_o \ll V_i$. Figure 7-22(a) shows a much superior, but more complex, *differentiating circuit*. The triangle represents an operational amplifier as in Prob. 7-2. Figure 7-22(b) shows the equivalent circuit. Show that

$$V_o = -RC \frac{dV_i}{dt},$$

as long as $A \gg 1$ and $|V_o|/RC \gg |dV_o/dt|/A$. Note that RC can be much larger than unity, so that V_o need not be much smaller than V_i .

7-8. (7.2) *RC* integrating circuit

Figure 7-23 shows an *RC* integrating circuit. The current through the load connected at V_o is negligible compared to dQ/dt in C .

(a) Show that as long as the voltage across C is small compared to that across R ,

$$V_o = \frac{1}{RC} \int_0^t V_i dt.$$

As in Prob. 7-6, $V_o \ll V_i$. We assume that $V_o = 0$ at $t = 0$.

(b) Sketch a curve of the output voltage as a function of time if V_i is a square wave.

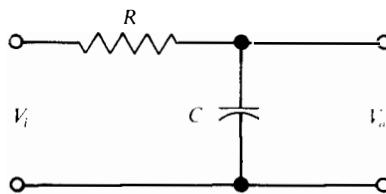


Fig. 7-23.

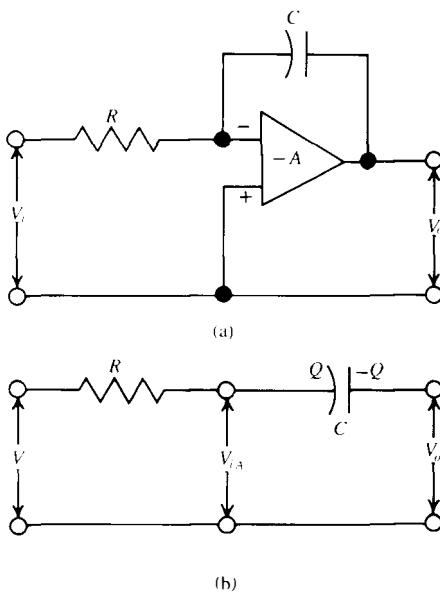


Fig. 7-24.

7-9. (7.2) Integrating circuit using an operational amplifier

The integrating circuit shown in Fig. 7-24(a) performs integrations without the limitation $V_o \ll V_i$ that applies to the circuit of Fig. 7-23. The triangle represents an operational amplifier, as in Prob. 7-2. Show that

$$V_o = -\frac{1}{RC} \int_0^t V_i dt,$$

if $A \gg 1$ and if $|V_o|/(RC) \ll A |dV_o/dt|$. Use the equivalent circuit of Fig. 7-24(b), and set

$$\frac{d}{dt} V_o = \frac{d}{dt} \left(V_{iA} - \frac{Q}{C} \right).$$

7-10. (7.2) Motion transducer

Figure 7-25 shows how the differentiating circuit of Prob. 7-7 can serve to measure a displacement.

Show that, with an alternating voltage at the input, the output voltage is proportional to the spacing s of the parallel-plate capacitor, neglecting edge effects.

7-11. (7.7) Compensated potential divider

The potential divider of Prob. 7-1 is not useful as such at high frequencies for the following reason. There are stray capacitances due to the wiring in parallel with R_1 and R_2 . If the frequency is high enough, these stray capacitances carry an appreciable current and V_o/V_i is a function of the frequency.

Show that, with the circuit of Fig. 7-26, the relation of Prob. 7-1 applies if $R_1 C_1 = R_2 C_2$. If the added capacitances are large compared to the stray capacitances, the potential divider is said to be *compensated*.

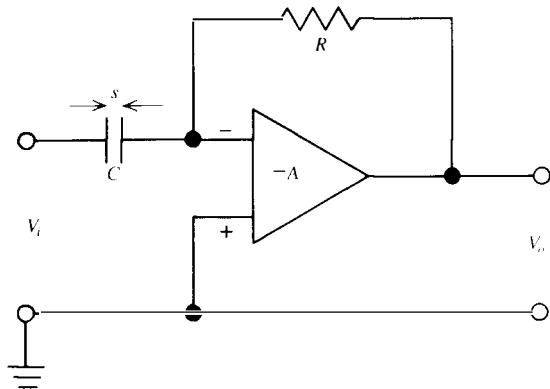


Fig. 7-25.

7-12. (7.7) Impedance bridges: The Wien bridge

There exist a wide variety of *impedance bridge* circuits for measuring the impedance of components. These bridges have now been largely superseded by *network analyzers* that can perform many sophisticated measurements on complex circuits. Figure 7-27(a) shows such an impedance bridge. If $Z_1/Z_2 = Z_3/Z_4$, then $V = 0$. The impedances must satisfy two independent equations to satisfy both the real and the imaginary parts of this equation.

One common type is the Wien bridge shown in Fig. 7-27(b). As a rule, one sets $R_1 = R_2/2$, $R_3 = R_4$, $C_3 = C_4$. Find the condition for balance.

The Wien bridge is used in tuned amplifiers and in oscillators as well as for measuring or monitoring a frequency. To measure a frequency, one changes R_3 and R_4 simultaneously until $R_3\omega C_3$ is equal to unity.

7-13. (7.7) Phase shifter

It is often necessary to shift the phase of a signal. Figure 7-28 shows a simple circuit for doing this without affecting the amplitude of the signal. The resistances are adjustable, but equal. Use the polarities shown. They mean that V_i is the voltage of the top terminal with respect to the bottom one, and V_o is the voltage of the right-hand terminal with respect to the left-hand one.

(a) Show that $V_o/V_i = \exp \{2j \arctan [1/(R\omega C)]\}$.

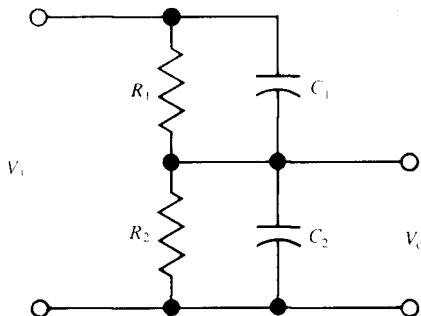


Fig. 7-26.

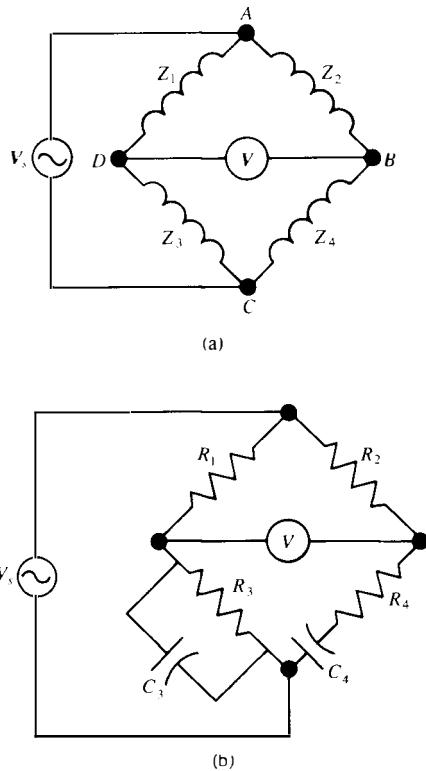


Fig. 7-27.

(b) Draw a graph of the phase of V_o with respect to V_i in the range $R\omega C = 0.1$ to 10. Use a logarithmic scale for $R\omega C$.

7-14. (7.8) Direct-current high-voltage transmission lines

From many points of view, alternating current is much preferable to direct current for power distribution. However, line losses are lower with direct current.

On a high-voltage overhead transmission line, the maximum operating voltage depends on several factors, such as corona losses (current losses

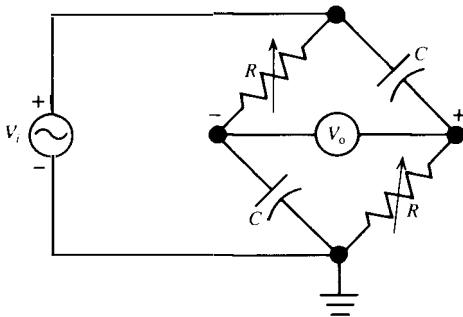


Fig. 7-28.

through ionization of the air), the size of the insulators, etc. So, for a given line, the instantaneous voltage between one conductor and ground must never exceed a certain value, say V_0 . Otherwise, the downtime and the cost of maintenance become excessive. The cost of a line increases rapidly with its voltage rating.

The current in the line can be made nearly as large as one likes without damaging it, since the conductors are well cooled by the ambient air. However, the power loss increases as the square of the current. So, the lower the current, the better.

With direct current, two conductors operate at $+V_0$ and $-V_0$ with respect to ground. The power delivered to the load is $2V_0I_{dc}$. With single-phase (SP) alternating current, there are two wires at $+V_0 \cos \omega t$ and $-V_0 \cos \omega t$.

(a) Show that for the same power at the load, the rms current I_{SP} is $2^{1/2}I_{dc}$. The I^2R losses in the line with single-phase alternating current are twice as large as those with direct current.

(b) With three-phase (TP) alternating current, we have three wires at $V_0 \cos \omega t$, $V_0 \cos(\omega t + 2\pi/3)$, and $V_0 \cos(\omega t + 4\pi/3)$. We assume that the three load resistances connected between these wires and ground are equal. Then the current in the ground wire is zero. Show that for the same total power delivered to the three load resistances, the rms currents I_{TP} are $(\frac{2}{3})2^{1/2}I_{dc} \approx I_{dc}$.

With three-phase alternating current, the rms currents are thus about the same as with direct current, but there are three current-carrying wires instead of two, so that the losses are 50% larger than with direct current. A three-wire line is also more expensive than a two-wire one.

CHAPTER 8

*ELECTRIC FIELDS VI *Electric Circuits B: Circuit Theorems*

8.1	THE PRINCIPLE OF SUPERPOSITION	149
8.2	THE SUBSTITUTION THEOREM	150
8.3	THÉVENIN'S THEOREM	150
	<i>Example: THE ZINC-CARBON BATTERY</i>	151
8.4	NORTON'S THEOREM	152
8.5	MILLMAN'S THEOREM	153
	<i>Example</i>	154
8.6	TELLEGREN'S THEOREM	154
8.6.1	A MORE GENERAL FORM OF TELLEGREN'S THEOREM	156
8.7	THE RECIPROCITY THEOREMS	157
	<i>Example</i>	158
	<i>Example: SEARCHING FOR ANOMALIES IN GROUND CONDUCTIVITY</i>	160
8.8	THE POWER-TRANSFER THEOREM	161
8.9	DELTA-STAR TRANSFORMATIONS	163
	<i>Example</i>	165
8.10	SUMMARY	166
	PROBLEMS	167

Generally speaking, circuit theorems serve to simplify calculations. There exists a seemingly infinite number of circuit theorems, but the eight that follow are probably the most useful. The delta-star transformations of Sec. 8.9 are even more useful.

All the material in this chapter is general and is *not* restricted to *RC* circuits. We, however, assume that all components are linear and time-*independent*, as in Chap. 7. A *linear* component is one that follows Ohm's law, $I = V/Z$.

8.1 THE PRINCIPLE OF SUPERPOSITION

The principle of *superposition* of Sec. 3.3 applies to linear circuits: each source acts independently of all the others. In other words, the current

* Chapters 7 and 8 are not required for what follows. However, Chap. 7 is a prerequisite for Chap. 8.

that flows through one branch is equal to the sum of the currents ascribable to each individual source.

8.2 THE SUBSTITUTION THEOREM

The *substitution theorem* is an obvious consequence of the Kirchhoff voltage law (KVL): if the voltage drop across a passive component in a circuit is V , then one can replace that component by an ideal voltage source supplying the same voltage, without disturbing any of the branch currents. Clearly, this does not disturb the voltage drops around the meshes. Figure 8-1 shows a trivial example.

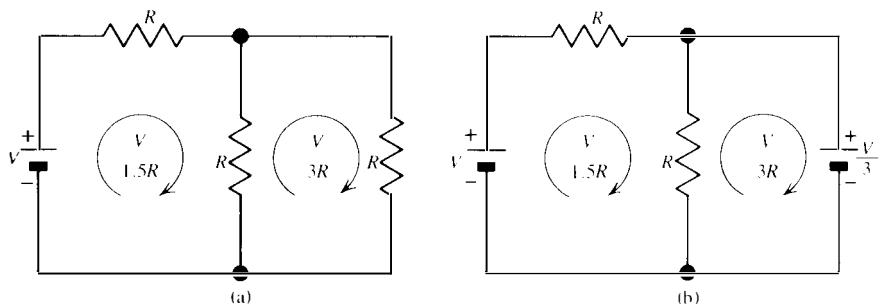


Fig. 8-1. The substitution theorem. (a) The voltage drop across the resistance on the right is $V/3$. (b) Replacing that resistance by a battery $V/3$ of the correct polarity leaves the mesh currents unchanged.

8.3 THÉVENIN'S THEOREM

Thévenin's theorem states that any active, linear, two-terminal circuit is equivalent to an ideal voltage source in series with an impedance Z_o , called the *output impedance* of the circuit, as in Fig. 8-2. As a rule, this impedance is resistive, and the source is said to have an *output resistance* R_o . A good voltage source has a low output resistance.

Thévenin's theorem applies to simple sources such as flashlight batteries, but it also applies to complex circuits such as audio amplifiers, power plant generators, etc. When applied to simple sources, Thévenin's theorem is an experimental fact, at least for a limited range of output currents. Granted that the theorem applies to simple sources, then it also applies to complex circuits (Prob. 8-3).

Thévenin's theorem is useful for calculating branch currents. We write

$$I_b = \frac{V_o}{Z_{Th}}, \quad (8-1)$$

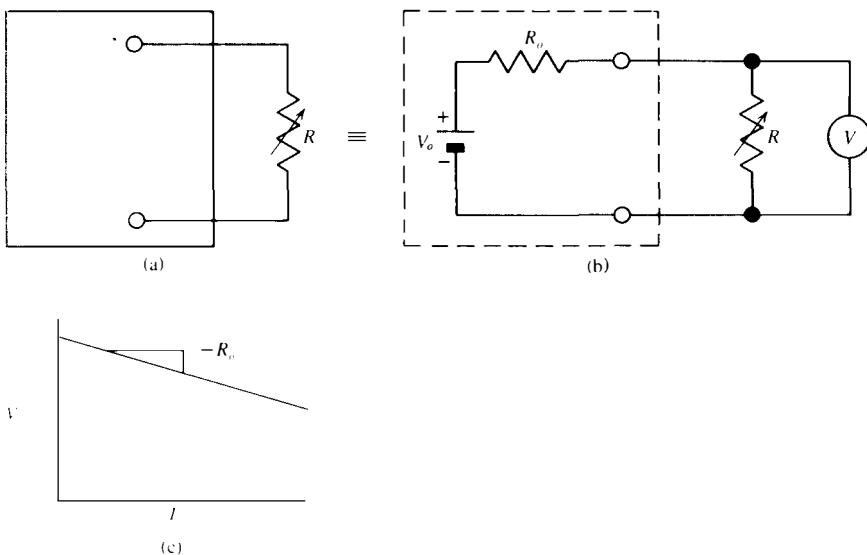


Fig. 8-2. Thévenin's theorem. (a) An active circuit A feeds current to a resistance R . (b) Circuit A is equivalent to an ideal voltage source V_o in series with an output impedance Z_o . As a rule, Z_o is a resistance R_o . (c) One can measure R_o by plotting $V = V_o - R_o I$ as a function of I . Decreasing R increases I and decreases V . The slope of the curve is $-R_o$.

where V_o is the voltage that would appear between the nodes at the ends of a branch b if the branch were removed, and Z_{Th} is the impedance calculated at the cut, including the impedance of the branch, with all the sources replaced by their output impedances. In other words,

$$I_b = \frac{V_o}{Z_o + Z_b}, \quad (8-2)$$

where Z_o is the output impedance of the circuit feeding the branch.

Example THE ZINC-CARBON BATTERY

A zinc-carbon battery, like the ones used in flashlights, performs as if it comprised an ideal voltage source set at $V_o = 1.5$ to 1.6 volts, depending on the actual materials used, in series with a resistance R_o . For a fresh size-D cell, $R_o = 0.27$ ohm. The value of R_o increases gradually with use and with time and eventually grows by many orders of magnitude. With time, V_o falls to zero.

A high-resistance voltmeter connected across the terminals measures V_o . To measure R_o one can either proceed as in Fig. 8-2, or measure both V_o and the “flash” (peak short-circuit) current.

8.4 NORTON'S THEOREM

Norton's theorem is the *dual* of Thévenin's. It states that any active, linear, two-terminal circuit is equivalent to an ideal *current source* I_o in *parallel* with an output *admittance* Y_o , which is usually a conductance. See Fig. 8-3. A good current source has a low Y_o .

It is easy to show that, for a given device,

$$Z_o Y_o = 1, \quad (8-3)$$

at least when Z_o and Y_o are real. From Fig. 8-2, R_o is equal to minus the slope of the V versus I curve. But, from Fig. 8-3, G_o is minus the slope of the I versus V curve. So $G_o = 1/R_o$ and $R_o G_o = 1$.

Norton's theorem is useful for calculating *branch voltages*. We again consider the complete circuit minus the branch as a source, and the branch as a load. Thus

$$V_b = \frac{I_o}{Y_o + Y_b}, \quad (8-4)$$

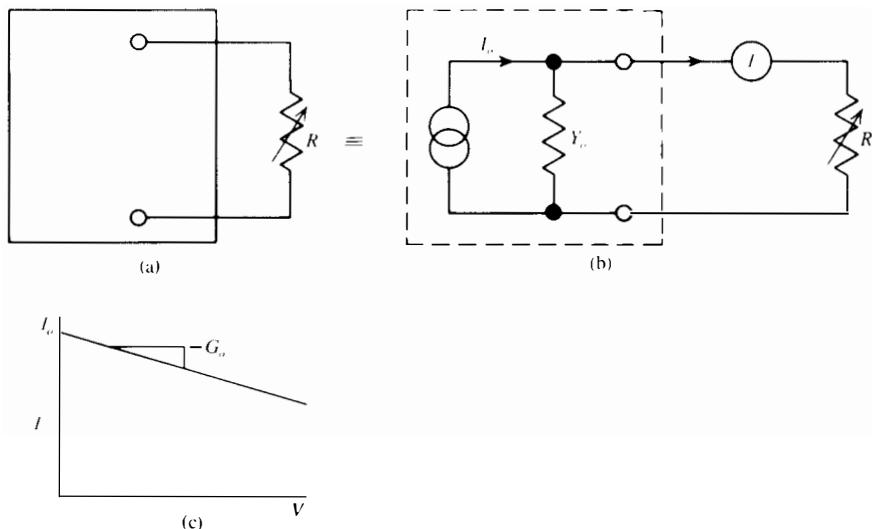


Fig. 8-3. Norton's theorem. (a) An active circuit A feeds a resistance R . (b) Circuit A is equivalent to an ideal current source in parallel with an output admittance Y_o . As a rule, Y_o is real and is thus a conductance G_o . (c) One can measure G_o by plotting $I = I_o - G_o V$ as a function of V . Increasing R increases V and decreases I . The slope of the curve is $-G_o$.

where I_o and Y_o are the parameters of the Norton source and Y_b is the admittance of the branch.

To calculate I_o and Y_o , we again remove the branch. Then I_o is the current through a zero-resistance ammeter connected to its terminals, and Y_o is the admittance calculated at those terminals, with all the sources replaced by their output admittances.

8.5 MILLMAN'S THEOREM

Millman's theorem serves to calculate the potential at a specific node in a circuit, and hence the incoming branch currents. It is a simple application of Kirchhoff's current law (KCL). It is best explained by an example.

In Fig. 8-4, $I_1 + I_2 + I_3 = 0$. Then

$$(V_1 - V)Y_1 + (V_2 - V)Y_2 + (V_3 - V)Y_3 = 0, \quad (8-5)$$

$$V = \frac{V_1 Y_1 + V_2 Y_2 + V_3 Y_3}{Y_1 + Y_2 + Y_3}. \quad (8-6)$$

This is *Millman's theorem*. Of course,

$$I_1 = (V_1 - V)Y_1, \quad \text{etc.} \quad (8-7)$$

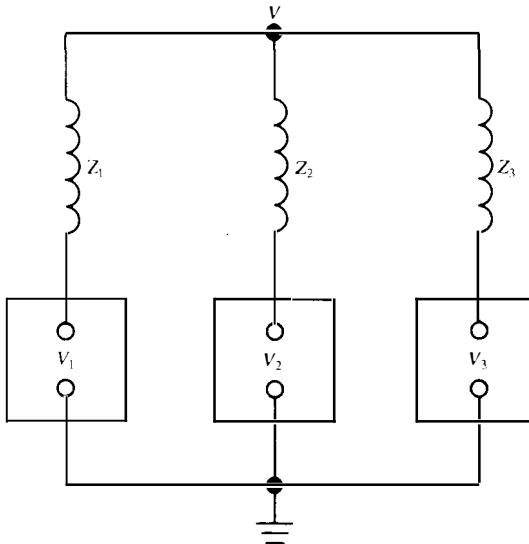


Fig. 8-4. Millman's theorem relates the voltage V at a node to the Y 's and V 's connected to it. The squares represent unspecified circuits that may be either passive or active.

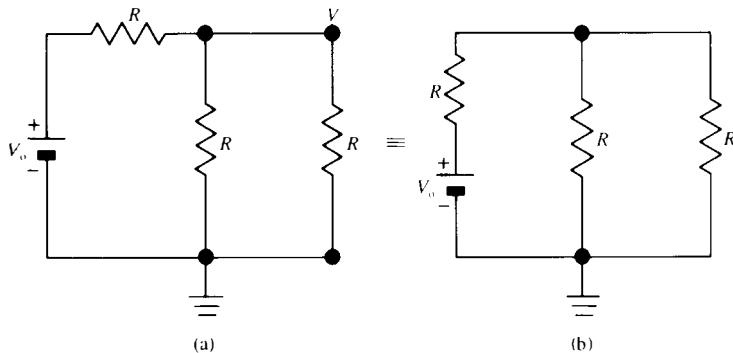


Fig. 8-5. (a) Simple circuit for testing Millman's theorem. (b) The same circuit, redrawn to resemble that of Fig. 8-4.

There can be any number of nodes connected to the specific node in question.

Example In Fig. 8-5, V is clearly equal to $V_o/3$. Applying Millman's theorem, we find that

$$V = \frac{V_o/R}{3(1/R)} = \frac{V_o}{3}. \quad (8-8)$$

8.6 TELLEGGEN'S THEOREM

Tellegen's theorem, in its original form, is like Kirchhoff's laws: it is self-evident and appears trivial. Again, this is a false impression; Tellegen's theorem is related to at least 150 other circuit theorems!†

Suppose we have only *direct* voltages and currents in a circuit. Set V_b equal to the voltage across branch b and I_b equal to the current flowing through it. Also, choose the signs of V_b and I_b in such a way that $V_b I_b$ is the power flowing *into* that branch at that instant, as in Fig. 8-6. Some branches comprise sources, others not. Then, according to *Tellegen's theorem*,

$$\sum V_b I_b = 0, \quad (8-9)$$

† Paul Penfield, Jr., Robert Spence, and Simon Duinker, *Tellegen's Theorem and Electrical Networks*, Research Monograph 58, M.I.T. Press, Cambridge, Mass., 1970.

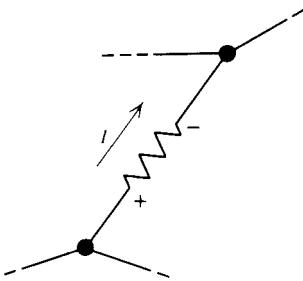


Fig. 8-6. The sign convention for Tellegen's theorem.

where the summation runs over all the branches of the circuit at that instant. This is simply a statement of the law of conservation of energy. You will be able to show in Prob. 8-7 that Tellegen's theorem follows from both the KCL and the KVL.

Now suppose that we have *alternating* currents in a circuit comprising only sources and linear passive components. Then the complex power (Sec. 7.7) flowing into branch b is $V_b I_b^*/2$, and the sum of the complex powers is zero,

$$\sum V_b I_b^* = 0, \quad (8-10)$$

for the following reason. Substitute for each starred branch current the difference between the neighboring starred mesh currents. Then, rearranging the sum, we have the first starred mesh current multiplied by the voltage drop around it, plus the second starred mesh current multiplied by the voltage drop around it, etc. Each term of this sum being zero, according to the KVL, the above equation is correct.

So the sum of the complex powers in a circuit is zero.

As a rule, a circuit comprises pairs of terminals called *input ports* that serve to feed either power or a signal. For example, in the Wheatstone bridge of Fig. 8-7, one connects a battery to the input port. As a rule, circuits also comprise one or more *output ports*, as exemplified again by the Wheatstone bridge. Tellegen's theorem, as stated above, applies to the complete circuit, including sources and load.

It is useful to rewrite Tellegen's theorem in a form that shows explicitly the power at the input and output ports. With the sign convention of Fig. 8-8 for the ports, the theorem becomes

$$\sum' V_b I_b - \sum V_p I_p = 0, \quad (8-11)$$

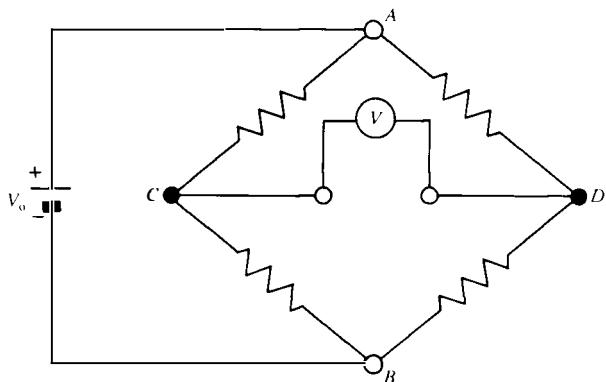


Fig. 8-7. Wheatstone bridge. The input port is AB , and the output port is CD .

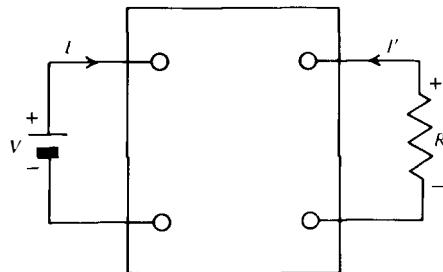


Fig. 8-8. The sign convention for Tellegen's theorem, as stated in Eq. 8-11.

where the primed summation over the branches excludes the ports, and where the power with subscripts p is the power flowing *out* of the devices connected to the ports. Thus

$$\sum' V_b I_b = \sum V_p I_p. \quad (8-12)$$

8.6.1 A More General Form of Tellegen's Theorem

We again limit ourselves to direct currents. Suppose that we have two circuits that are unrelated, except that they share the same *geometry*. The branch voltages and currents are V_b and I_b for one circuit, and V'_b and I'_b for the other. What is the value of $\sum V_b I'_b$? This is admittedly a weird expression, but we shall use it in the next section.

First we substitute for I'_b the difference between the neighboring mesh currents in the primed circuit. Rearranging the sum, we have the first mesh current of the primed circuit, multiplied by the voltage drops

around the first mesh of the unprimed circuit, the second . . . etc. Each term of this new series is zero, because of the KVL, and

$$\sum V_b I'_b = 0. \quad (8-13)$$

By symmetry,

$$\sum V'_b I_b = 0. \quad (8-14)$$

8.7 THE RECIPROCITY THEOREMS

The reciprocity theorems are most useful. We use the above form of Tellegen's theorem to prove one of them.

Figure 8-9(a) and (b) shows a two-port circuit P connected to external elements x in two different ways. Circuit P is passive and linear. The x 's can be either sources or passive linear elements.

We now rewrite Eqs. 8-13 and 8-14 in the form of Eq. 8-12, with the power flowing *out* of the devices connected to the ports on the left-hand side of the equations and the power flowing *into* P on the right-hand side:

$$V_1 I'_1 + V_2 I'_2 = \sum_P V_b I'_b, \quad (8-15)$$

$$V'_1 I_1 + V'_2 I_2 = \sum_P V'_b I_b. \quad (8-16)$$

We can easily show that the sums on the right are equal. Say the resistance of branch b of circuit A is R_b . Then

$$V_b I'_b = (I_b R_b) I'_b = I_b (R_b I'_b) = I_b V'_b. \quad (8-17)$$

It follows that, in Fig. 8-9,

$$V_1 I'_1 + V_2 I'_2 = V'_1 I_1 + V'_2 I_2. \quad (8-18)$$

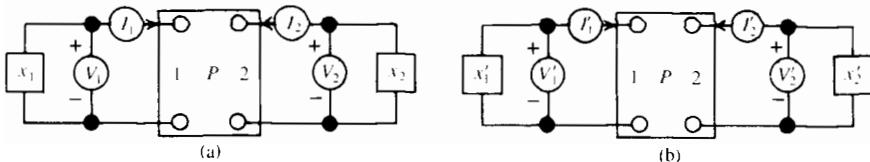


Fig. 8-9. Identical passive and linear circuits P connected to external elements in two different ways. As usual, a circled V represents an infinite-resistance voltmeter, and a circled I a zero-resistance ammeter. The signs and the arrows show which directions are taken to be positive.

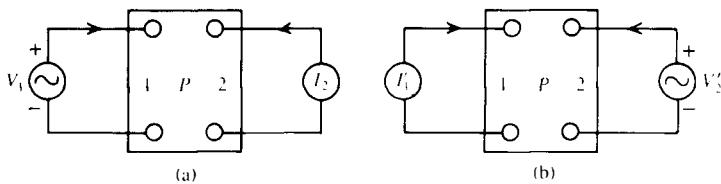


Fig. 8-10. (a) Circuit P with an ideal voltage source on the left and an ammeter on the right. (b) The same circuit P , with a voltage source on the right and an ammeter on the left. The most commonly used reciprocity theorem states that the ratios V/I are equal.

Now connect P as in Fig. 8-10(a), which makes $V_2 = 0$. Reconnect P as in Fig. 8-10(b) so as to make $V'_1 = 0$. Then, from Eq. 8-18,

$$V_1 I'_1 = V'_1 I_2. \quad (8-19)$$

Setting $V_1 = V'_1$ makes $I'_1 = I_2$.

In other words, *interchanging the voltage source and the ammeter does not affect the ratio V/I* . We have assumed steady direct currents and linear elements, but the result is general.

This is an important result. It means that given any linear and passive circuit with one input port and one output port, the short-circuit output current, for a given input voltage, is unaffected if the roles of the input and output ports are interchanged. In other words, the ratio V/I remains the same. The output impedance of the source need not be zero, but the impedance of the ammeter must be zero. This is the most commonly used *reciprocity theorem*.

This result is unexpected, because circuit P need not be symmetric at all. The power supplied by the source is *not* the same in the two configurations.

Figure 8-11 illustrates the four reciprocity theorems.

Example

Let us check the reciprocity theorem illustrated in Fig. 8-10(a) and (b) by applying it to the circuit of Fig. 8-12(a). We first apply V volts at port 1 and connect an ammeter at port 2, as in Fig. 8-12(b). We require the current I_b when port 2 is short-circuited by the ammeter. Let us calculate I_b in three different ways: by the mesh method of Sec. 7.2.1, by the node method of Sec. 7.2.2, and by Millman's theorem of Sec. 8.5.

With the *mesh method*, we apply the KVL to meshes a and b in succession:

$$V - I_a R_1 - (I_a - I_b) R_3 = 0, \quad -(I_b - I_a) R_3 - I_b R_2 = 0. \quad (8-20)$$

Solving,

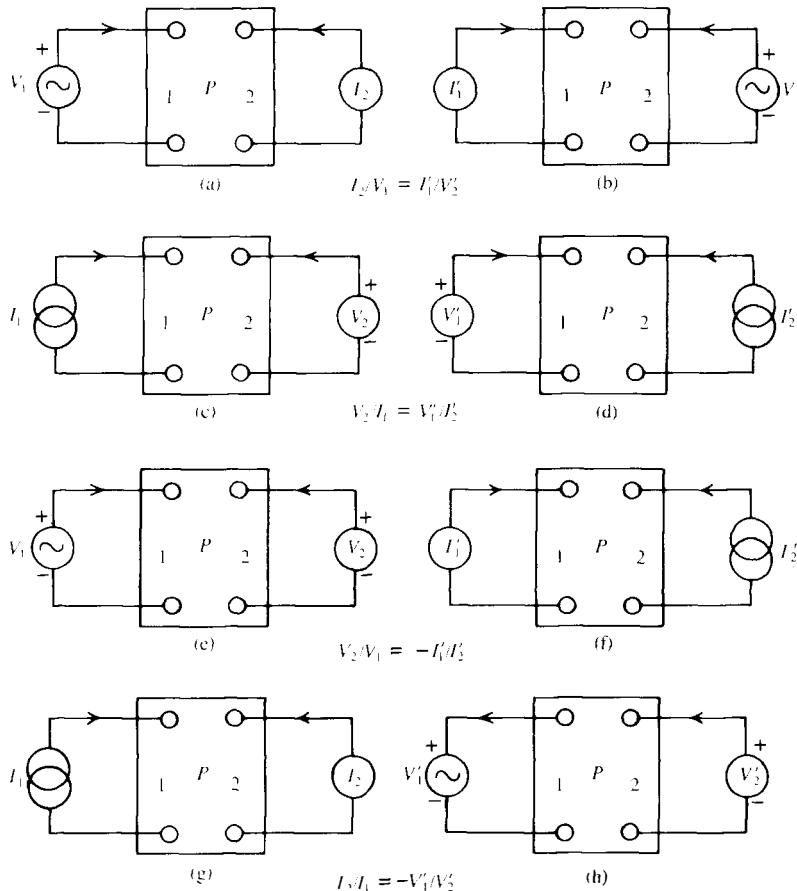


Fig. 8-11. The four reciprocity theorems. In all cases P is any passive and linear circuit. (a), (b) Interchanging the positions of the ideal voltage source and of the ammeter does not alter the ratio I/V . (c), (d) Interchanging the ideal current source and the voltmeter does not alter the ratio V/I . (e), (f) and (g), (h) Two other reciprocity theorems.

$$I_b = \frac{R_3}{R_2 R_3 + R_3 R_1 + R_1 R_2} V. \quad (8-21)$$

With the *node method*, we set the sum of the currents flowing into node A equal to zero:

$$\frac{V - V_A}{R_1} = \frac{V_A}{R_3} + \frac{V_A}{R_2}. \quad (8-22)$$

Solving, we find the same value for I_b :

$$V_A = \frac{R_2 R_3}{R_2 R_3 + R_3 R_1 + R_1 R_2} V, \quad (8-23)$$

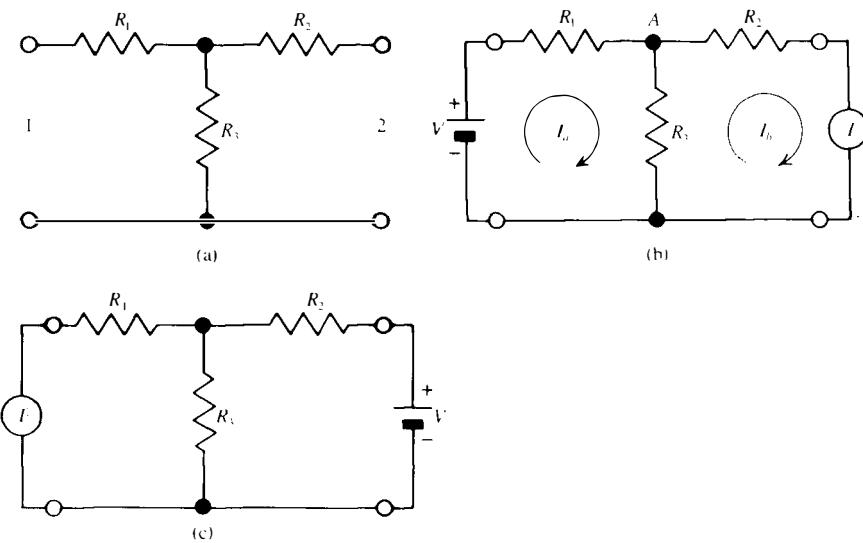


Fig. 8-12. (a) Simple circuit for testing the reciprocity theorem of Fig. 8-10(a) and (b). (b) We have added a voltage source on the left and an ammeter on the right. (c) We have now interchanged the source and the ammeter.

$$I_b = \frac{V_A}{R_2} = \frac{R_1 V}{R_2 R_3 + R_3 R_1 + R_1 R_2}. \quad (8-24)$$

We can apply *Millman's theorem* by referring again to Fig. 8-12(b):

$$V_A = \frac{V/R_1}{1/R_1 + 1/R_2 + 1/R_3}. \quad (8-25)$$

This yields the same value for \$V_A\$ and the same for \$I_b\$.

Now, to check the reciprocity theorem, we interchange the voltage source and the ammeter of Fig. 8-12(b) to obtain Fig. 8-12(c). We require \$I_a\$. That is easy! Circuit (d) is identical to circuit (b), except for the fact that \$R_1\$ and \$R_2\$ are interchanged. So we can write down the \$I_a\$ of circuit (d) by interchanging \$R_1\$ and \$R_2\$ in the above expression for the \$I_b\$ of circuit (b). The two currents are the same.

Problem 6-4 concerns a reciprocity theorem for electrostatic fields.

Example

SEARCHING FOR ANOMALIES IN GROUND CONDUCTIVITY

Geophysicists use many methods for locating potential ore bodies. Figure 4-7 shows one method. The ground acts as circuit \$A\$, port 1 is the pair of electrodes \$C_a C_b\$, and port 2 is the pair \$P_a P_b\$.

According to the reciprocity theorem illustrated in Fig. 8-11(c) and (d), the ratio V/I would be unaffected if, instead, the current was injected at P_aP_b and the voltage measured at C_aC_b .

8.8 THE POWER-TRANSFER THEOREM

The *power-transfer theorem* states that the condition for maximum power transfer to a load is that

$$Z_L = Z_o^*, \quad (8-26)$$

where Z_L is the load impedance and Z_o^* is the complex conjugate of the source output impedance.

The proof is simple. According to Thévenin's theorem (Sec. 8.3), a real source is equivalent to an ideal voltage source in series with an impedance Z_o , as in Fig. 8-13. Let

$$Z_o = R_o + jX_o, \quad Z_L = R_L + jX_L. \quad (8-27)$$

Then the power dissipated in the load is

$$P_L = I^2 R_L = \frac{V^2}{(R_o + R_L)^2 + (X_o + X_L)^2} R_L. \quad (8-28)$$

To maximize P_L , set $X_L = -X_o$. Then

$$P_L = \frac{R_L}{(R_o + R_L)^2} V^2. \quad (8-29)$$

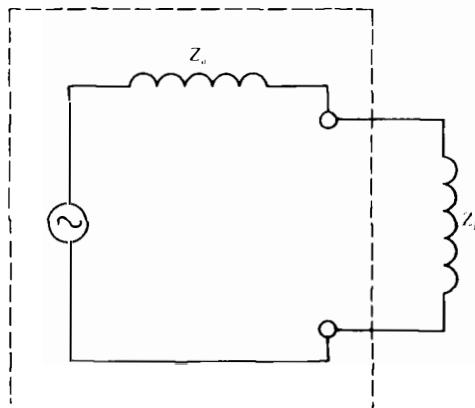


Fig. 8-13. Source, whose output impedance is Z_o , connected to a load impedance Z_L .

For a given source, with a given value of R_o , we set $dP_L/dR_L = 0$ and find that $R_L = R_o$. So P_L is maximum when

$$Z_L = R_L + jX_L = R_o - jX_o = Z_o^*, \quad (8-30)$$

as above.

Under those conditions, the power expended in the source is

$$P_o = I^2 R_o = I^2 R_L = \frac{V_m^2}{4R_L}. \quad (8-31)$$

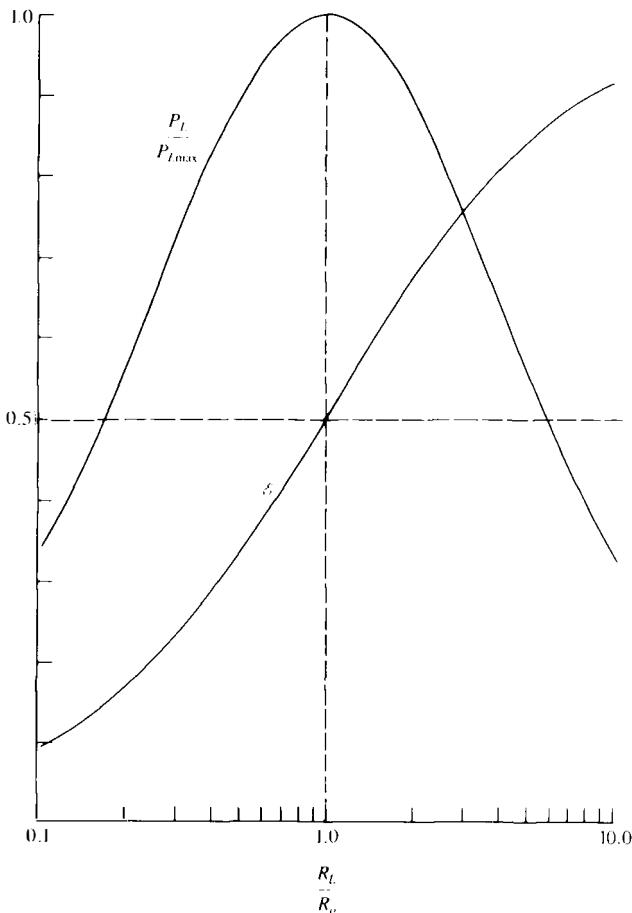


Fig. 8-14. The ratio $P_L/P_{L,\max}$ and the efficiency \mathcal{E} as functions of the ratio R_L/R_o for the circuit of Fig. 8-14, with $X_L = -X_o$. The power P_L dissipated in the load is maximum when the load resistance R_L is equal to the output resistance R_o of the source. But then only half the power goes to the load; the other half is lost in the source.

So the condition for maximum power transfer from a source to a load requires that one-half of the total power be dissipated in the source. In other words, the efficiency at maximum power transfer is 50%.

Figure 8-14 shows that the condition for maximum power transfer is not critical.

8.9 DELTA-STAR TRANSFORMATIONS

Figure 8-15(a) shows three nodes A , B , C forming part of some unspecified circuit. The nodes are *delta-connected* in (a) and *star-connected* in (b). As we shall see, we can substitute the star for the equivalent delta, or inversely, without disturbing the rest of the circuit. Indeed, if we had two boxes, one containing the delta and the other the equivalent star, with only the terminals A , B , C showing, there would be no way of telling which box was which.

It is useful to be able to transform a delta into an equivalent star, or inversely, on paper. This often simplifies the calculation of mesh currents and node voltages.

We could find Z_A , Z_B , Z_C in terms of Z_a , Z_b , Z_c , and, inversely, by assuming the same mesh currents I_A , I_B , I_C in the two circuits, and then making $V_A - V_B$, $V_B - V_C$, $V_C - V_A$ in the delta equal to the corresponding voltages in the star. See Prob. 8-11. Here is another approach that is somewhat less convincing, but shorter.

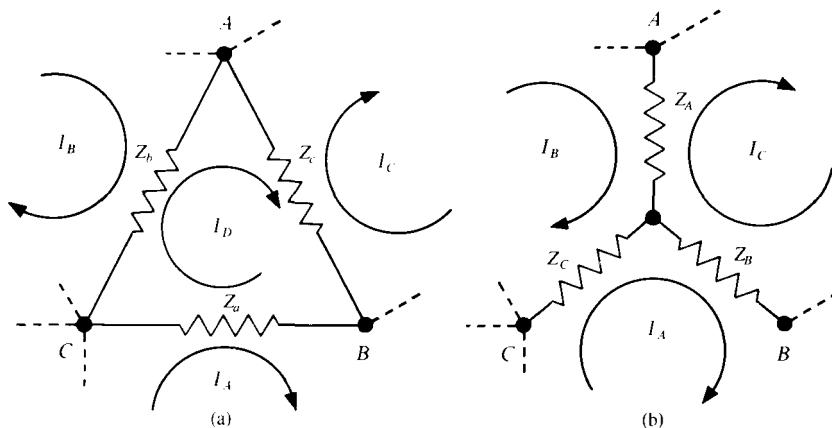


Fig. 8-15. The three nodes A , B , C are part of a larger circuit. (a) The three nodes are delta-connected. (b) The nodes are star-connected. Under certain conditions the delta is completely equivalent to the star, but only at a given frequency. If the six Z 's are all resistances, then the equivalence is valid at all frequencies.

Suppose we have a delta and the equivalent star. The impedance between A and B is the same in both. Then

$$Z_{AB} = \frac{Z_c(Z_a + Z_b)}{Z_a + Z_b + Z_c} = Z_A + Z_B. \quad (8-32)$$

Similarly,

$$Z_{BC} = \frac{Z_a(Z_b + Z_c)}{Z_a + Z_b + Z_c} = Z_B + Z_C, \quad (8-33)$$

$$Z_{CA} = \frac{Z_b(Z_c + Z_a)}{Z_a + Z_b + Z_c} = Z_C + Z_A. \quad (8-34)$$

Observe that Z_A is associated with Z_bZ_c , Z_B with Z_cZ_a , and Z_C with Z_aZ_b . Therefore

$$Z_A = \frac{Z_bZ_c}{Z_a + Z_b + Z_c}, \quad (8-35)$$

$$Z_B = \frac{Z_cZ_a}{Z_a + Z_b + Z_c}, \quad (8-36)$$

$$Z_C = \frac{Z_aZ_b}{Z_a + Z_b + Z_c}. \quad (8-37)$$

To find the inverse relationship, we use the admittance between node A and nodes B and C short-circuited together. This gives

$$Y_b + Y_c = \frac{Y_A(Y_B + Y_C)}{Y_A + Y_B + Y_C}. \quad (8-38)$$

Similarly,

$$Y_c + Y_a = \frac{Y_B(Y_C + Y_A)}{Y_A + Y_B + Y_C}, \quad (8-39)$$

$$Y_a + Y_b = \frac{Y_C(Y_A + Y_B)}{Y_A + Y_B + Y_C}. \quad (8-40)$$

Again by inspection,

$$Y_a = \frac{Y_B Y_C}{Y_A + Y_B + Y_C}, \quad (8-41)$$

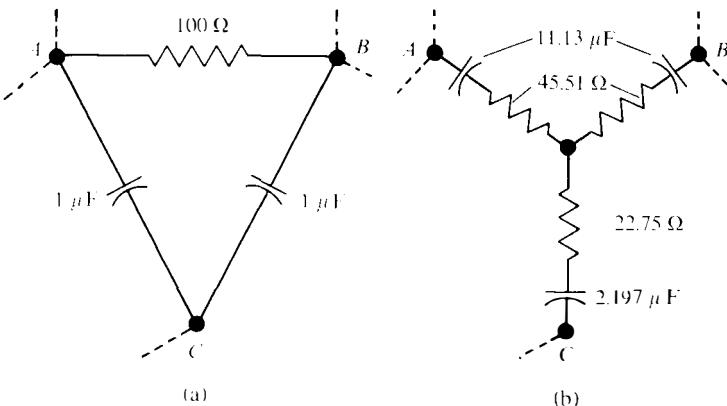


Fig. 8-16. (a) Delta-connected circuit. (b) Its equivalent star at 1 kilohertz. Note the negative resistance.

$$Y_b = \frac{Y_C Y_A}{Y_A + Y_B + Y_C}, \quad (8-42)$$

$$Y_c = \frac{Y_A Y_B}{Y_A + Y_B + Y_C}. \quad (8-43)$$

If a star or a delta comprises only resistances, then the equivalent circuit is also purely resistive. However, if the original circuit comprises capacitors (or inductors, or both), then the values of the resistances and capacitances (and inductances) in the equivalent circuit are themselves frequency-dependent.

For example, refer to Fig. 8-16(a) and call the components R and C . Then the resistance in the lower branch of the equivalent star on the right turns out to be $-R/(4 + R^2\omega^2C^2)$. This means to say that, for given components, the equivalence is valid at only one frequency. Moreover, the resistances of equivalent circuits can be negative, as in this instance. So a real circuit can be equivalent, on paper, to an impossible one.

Example

The delta and the star circuits of Fig. 8-16 are equivalent, as you can check by setting

$$Z_a = Z_b = \frac{1}{2\pi \times 10^3 \times 10^{-6}j} \quad \text{ohms}, \quad Z_c = 100 \text{ ohms}, \quad (8-44)$$

$$Z_A = Z_B = 45.51 + \frac{1}{2\pi \times 10^3 \times 11.13 \times 10^{-6}j} \quad \text{ohms}, \quad (8-45)$$

$$Z_c = -22.75 + \frac{1}{2\pi \times 10^3 \times 2.197 \times 10^{-6}j} \quad \text{ohms}. \quad (8-46)$$

8.10 SUMMARY

The *superposition theorem* states that the current flowing in a given branch of a circuit is equal to the sum of the currents ascribable to the various sources.

One can replace a passive element in a circuit by an ideal voltage source of the same polarity without disturbing the currents. This is the *substitution theorem*.

According to the *Thévenin's theorem*, any active, linear, two-terminal circuit is equivalent to an ideal voltage source in series with an impedance, called the *output impedance* of the circuit. This impedance is usually a resistance.

Norton's theorem is the dual of Thévenin's: any active, linear, two-terminal circuit is equivalent to an ideal current source in parallel with an admittance, which is usually a conductance.

Millman's theorem gives the voltage V at a node in terms of the voltages on the neighboring nodes and of the incoming currents.

Tellegen's theorem states that, for any circuit comprising sources and linear passive components,

$$\sum V_b I_b^* = 0, \quad (8-10)$$

where the summation is taken over all the branches. It follows that, for any pair of resistive circuits C and C' sharing the same geometry,

$$\sum V_b I'_b = 0, \quad \sum V'_b I_b = 0. \quad (8-13), (8-14)$$

The most common *reciprocity theorem* states that, given any passive and linear two-port circuit, with a voltage source connected to one port and an ammeter connected to the other, the ratio I/V is unaltered when the positions of the source and of the ammeter are interchanged. Figure 8-11 illustrates the four reciprocity theorems.

Finally, the *power-transfer theorem* states that maximum power transfer from a source to a load occurs when the load impedance is the complex conjugate of the source output impedance.

The delta-star transformations. Given three nodes A , B , C , either *delta-connected* as in Fig. 8-16(a) or *star-connected* as in Fig. 8-15(b), the two circuits are entirely equivalent if

$$Z_A = \frac{Z_b Z_c}{Z_a + Z_b + Z_c}, \quad (8-35)$$

$$Z_B = \frac{Z_c Z_a}{Z_a + Z_b + Z_c}, \quad (8-36)$$

$$Z_C = \frac{Z_a Z_b}{Z_a + Z_b + Z_c}, \quad (8-37)$$

or if

$$Y_a = \frac{Y_B Y_C}{Y_A + Y_B + Y_C}, \quad (8-41)$$

$$Y_b = \frac{Y_C Y_A}{Y_A + Y_B + Y_C}, \quad (8-42)$$

$$Y_c = \frac{Y_A Y_B}{Y_A + Y_B + Y_C}. \quad (8-43)$$

PROBLEMS

- 8-1. (8.3) The output resistance of an automobile battery.

The voltage at the terminals of a certain defective automobile battery drops from 12.5 to 11.5 volts when the headlights are turned on. What is the approximate value of the output resistance?

- 8-2. (8.3) The output resistance of a bridge circuit.

Show that the output resistance of the bridge circuit of Fig. 8-17, as seen at the voltmeter V , is R . Assume that the source has a zero output resistance.

- 8-3. (8.3) Calculating a branch current by Thévenin's theorem

(a) Calculate the current that flows to the right through the resistance R_o in Fig. 8-18. Use mesh currents and the KVL. The output resistance of the battery is negligible.

(b) Calculate the same current by Thévenin's theorem, considering the complete circuit minus R_o as the source.

- 8-4. (8.4) Norton's theorem

(a) Show that $V_o = I_o / R_o$.

(b) Show that Eq. 8-4 is equivalent to Eq. 8-2.

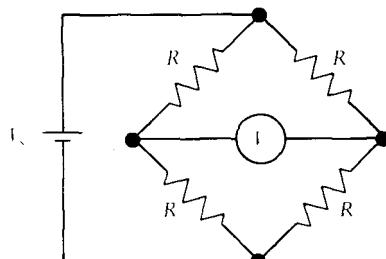


Fig. 8-17.

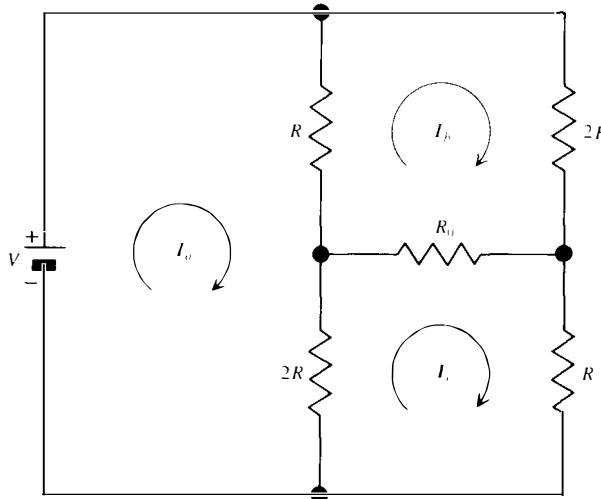


Fig. 8-18.

8-5. (8.4) Norton's theorem.

Calculate the voltage across the resistance R_o of Fig. 8-18 by Norton's theorem. The output conductance of the source is infinite.

8-6. (8.5) Millman's theorem

In the circuit of Fig. 8-19 the switch S closes at $t = 0$. Find the voltage V_C across C as a function of the time by means of Millman's theorem. Set $R_1 = 1$ ohm, $R_2 = 2$ ohms, $R_3 = 3$ ohms, $C = 1$ microfarad, $V = 100$ volts.

8-7. (8.6) Tellegen's theorem

- (a) Show that Tellegen's theorem is a consequence of the KCL.
- (b) Show that Tellegen's theorem is also a consequence of the KVL.

8-8. (8.7) The reciprocity theorems

Check the last three reciprocity theorems of Fig. 8-11 by applying them to the simple circuit of Fig. 8-12(a).

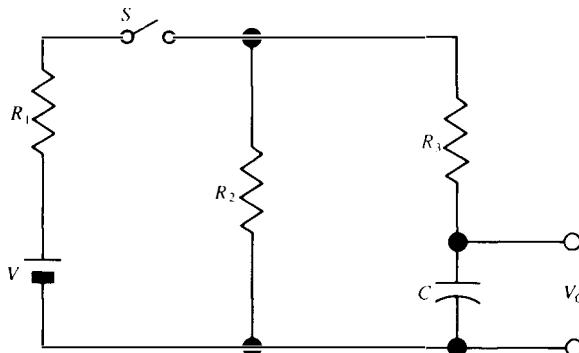


Fig. 8-19.

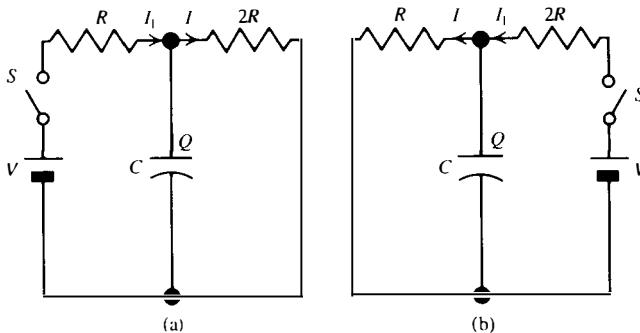


Fig. 8-20.

8-9. (8.7) The reciprocity theorem applies to transients

Show that the ratio I/V for the circuit of Fig. 8-20(a) is the same as for the circuit of Fig. 8-20(b). Find $Q(t)$ first.

8-10. (8.8) Audio amplifier

An audio power amplifier has an output resistance of 8 ohms and feeds a resistive load of 6.4 ohms. Calculate the efficiency.

8-11. (8.9) Delta-star transformation.

Find the equations for either the delta-star or the star-delta transformation by assuming mesh currents as in Fig. 8-15 and making the voltages $V_A - V_B$, $V_B - V_C$, $V_C - V_A$ in part (a) the same as those in part (b).

Find an equation of the form $(\dots)I_A + (\dots)I_B + (\dots)I_C = 0$. Since this equation must be valid whatever the values of the mesh currents, the parentheses must all be identically equal to zero. This will give you one of the equations of one set; the other two equations follow by symmetry.

8-12. (8.9) Delta-star transformation

Find the resistance of the circuit shown in Fig. 8-21.

8-13. Van der Pauw's theorem

It is difficult to measure the conductivity of small samples of semiconductor. First, they are brittle and thus difficult to machine. Second, contacts to the material are resistive. With Van der Pauw's theorem, however, it is possible to measure the conductivity of a sample in the form of a thin plate of arbitrary shape with four contacts around the periphery without

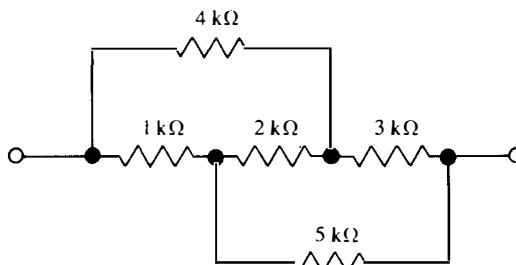
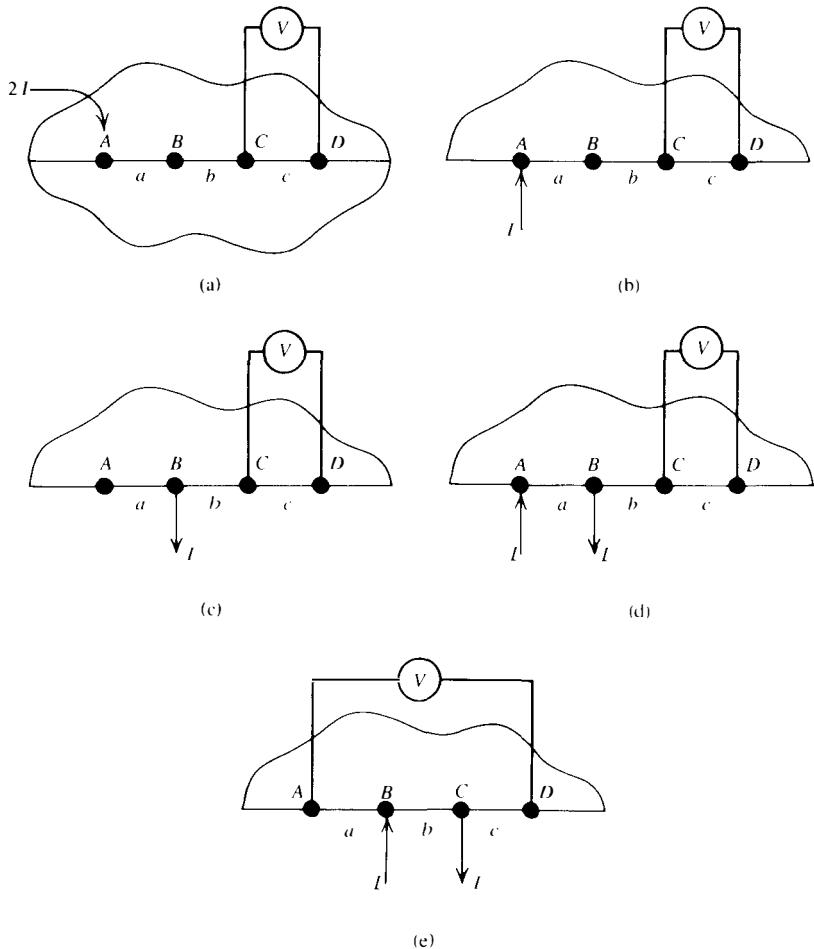


Fig. 8-21.

**Fig. 8-22.**

interference from the contact resistances. We deduce this theorem for the case of a semi-infinite plate.

(a) Imagine an infinite thin plate of thickness s and conductivity σ . A current $2I$ flows into point A in Fig. 8-22(a).

Show that, in the plate, $E = I/\pi\sigma rs$, where r is the radial distance to A.

(b) Now consider three points B, C, D as in Fig. 8-22(a), on a line going through A. Show that

$$V_C - V_D = \frac{I}{\pi\sigma s} \ln \frac{a+b+c}{a+b}.$$

If we cut the plate along the line and remove the lower half, the above

equation applies to the upper half, if I is now the current at A flowing into the upper half. So we now have a semi-infinite plate as in Fig. 8-22(b).

(c) Now suppose that a current I comes out of B as in Fig. 8-22(c). Calculate $V_C - V_D$ again; then superpose cases (b) and (c) to obtain Fig. 8-22(d). Show that, for case (d),

$$\frac{V'_C - V'_D}{I} = \frac{1}{\pi\sigma s} \ln \frac{(a+b+c)b}{(a+b)(b+c)}.$$

This ratio has the dimensions of a resistance; call it $R_{AB,CD}$. The contact resistances at A and B are unimportant because only the *current* I between A and B is significant. The contact resistances at C and D are also unimportant if their sum is much smaller than the resistance of the voltmeter that measures $V_C - V_D$.

(d) Show that, with currents as in Fig. 8-22(e),

$$R_{BC,DA} = \frac{1}{\pi\sigma s} \ln \frac{(a+b)(b+c)}{ac}.$$

(e) You can now derive Van der Pauw's theorem:

$$\exp(\pi\sigma s R_{AB,CD}) + \exp(\pi\sigma s R_{BC,DA}) = 1.$$

The only unknown is σ . This result, in fact, applies to a lamella of any shape, with contacts A, B, C, D around the periphery.

CHAPTER 9

ELECTRIC FIELDS VII

Dielectric Materials A: Bound Charges and the Electric Flux Density D

9.1	THE THREE BASIC POLARIZATION PROCESSES	173
9.2	THE ELECTRIC POLARIZATION P	173
9.3	FREE AND BOUND CHARGES	174
9.3.1	THE BOUND SURFACE CHARGE DENSITY σ_b	174
9.3.2	THE BOUND VOLUME CHARGE DENSITY ρ_b	175
9.3.3	THE POLARIZATION CURRENT DENSITY J_p	176
9.4	THE ELECTRIC FIELD OF A POLARIZED DIELECTRIC	176
	<i>Example: THE FIELD OF AN INFINITE-SHEET ELECTRET</i>	177
9.5	GAUSS'S LAW	178
9.6	POISSON'S AND LAPLACE'S EQUATIONS FOR V IN DIELECTRICS	178
9.7	THE ELECTRIC FLUX DENSITY D , THE DIVERGENCE OF D	179
	<i>Example: THE INFINITE-SHEET ELECTRET</i>	180
	<i>Example: THE BAR ELECTRET</i>	180
9.8	THE ELECTRIC SUSCEPTIBILITY χ_e	181
9.9	THE RELATIVE PERMITTIVITY ϵ_r	181
	<i>Example: THE FREE AND BOUND VOLUME CHARGE DENSITIES</i>	182
	<i>Example: THE FREE AND BOUND SURFACE CHARGE DENSITIES AT THE INTERFACE BETWEEN A DIELECTRIC AND A CONDUCTOR</i>	183
	<i>Example: DIELECTRIC-INSULATED PARALLEL-PLATE CAPACITOR</i>	184
	<i>Example: THE FIELD OF FREE CHARGES EMBEDDED IN A DIELECTRIC</i>	186
9.10	THE DISPLACEMENT CURRENT DENSITY $\partial D / \partial t$	186
9.11	SUMMARY	187
	PROBLEMS	189

Dielectrics differ from conductors in that they possess no carriers of free charge that can drift about under the control of an externally applied electric field. In a true dielectric the charges are all bound to their atoms or molecules, and they can be forced to move by only minute distances, positive charges going one way and negative charges the other way. A dielectric in which this displacement has taken place is said to be *polarized*.

Some molecules possess a permanent dipole moment and are said to be *polar*. Other molecules are *nonpolar*.

A given substance can be a dielectric under normal circumstances and become a conductor under appropriate conditions. For example, *photoconductors* are normally nonconducting, but become conducting when exposed to light.

In this first chapter on dielectric media we are concerned with the basic concepts of polarization, susceptibility, and relative permittivity, mostly in static fields.

9.1 THE THREE BASIC POLARIZATION PROCESSES

- (1) Under the action of an applied electric field, the center of charge† of the electron cloud in a molecule moves slightly with respect to the center of charge of the nuclei. This is *electronic polarization*. The displacement is minute, even on the atomic scale, typically 10^{-8} times the diameter of an atom.
- (2) Polar molecules align themselves and become further polarized in an applied electric field. This is *orientational polarization*. However, collisions arising from thermal agitation partly disrupt the alignment.
- (3) The third basic process is *atomic polarization*, in which *ions* of different signs in a solid such as NaCl move in different directions when subjected to an electric field. The ferroelectric dielectrics of Sec. 10.1.4 exhibit atomic polarization.

A polarized dielectric possesses its own field, which adds to that of the other charges. The two fields can be comparable in magnitude.

9.2 THE ELECTRIC POLARIZATION \mathbf{P}

If, in the neighborhood of a given point, the average vector dipole moment per molecule in a given direction is \mathbf{p} , and if there are N molecules per cubic meter, then

$$\mathbf{P} = N\mathbf{p} \quad (9-1)$$

is the *electric polarization* at that point. So \mathbf{P} is the *dipole moment per unit volume* at a given point. Recall from Sec. 5.4 that the dipole moment

† The center of charge is analogous to the center of mass in mechanics.

of a charge distribution is independent of the choice of origin when the net charge of the distribution is zero, as it is for a normal, neutral molecule.

9.3 FREE AND BOUND CHARGES

Polarization causes charges to accumulate, either within the dielectric or at its surface. We refer to such charges as *bound*. Other charges are said to be *free*. Examples of free charges are the conduction electrons in good conductors, the charge carriers in semiconductors, and electrons injected into a dielectric by means of a high-energy electron beam.[†]

9.3.1 The Bound Surface Charge Density σ_b

Imagine an element of area dA inside a dielectric as in Fig. 9-1. Say the dielectric is nonpolar. When the dielectric is polarized, the center of positive charge $+Q$ of a molecule lies at a distance s from the center of negative charge $-Q$. This s is the same for all the molecules over an infinitesimal region.

Upon application of an electric field, n_+ positive charges cross the element of area by moving in the direction of s , and n_- negative charges cross it by moving in the opposite direction. The net charge that crosses dA in the direction of s is therefore

$$dQ = n_+ Q - n_- (-Q) = (n_+ + n_-)Q. \quad (9-2)$$

Now $n_+ + n_-$ is simply the number of molecules within the imaginary parallelepiped of Fig. 9-1, whose volume is $s \cdot dA$. Then

$$dQ = NQs \cdot dA = N\mathbf{p} \cdot dA = \mathbf{P} \cdot dA, \quad (9-3)$$

where Qs is the dipole moment \mathbf{p} of a single molecule.

If dA lies on the surface of a dielectric, then dQ accumulates there and the *bound surface charge density* is

[†] The concepts of free and bound charges, and even the concept of polarization, are not, however, as clear-cut as one might wish. The distinction between free and bound charges rests on the assumption that the dielectric consists of electrically neutral molecules. If there are no well-defined molecules as in a crystal of NaCl, for example, the values of the free and bound surface charge densities are arbitrary. We confine most of our discussion here to dielectrics composed of molecules.

See Edward Purcell, *Electricity and Magnetism, Berkeley Physics Course*, vol. 2, McGraw-Hill, New York, 1965, p. 344.

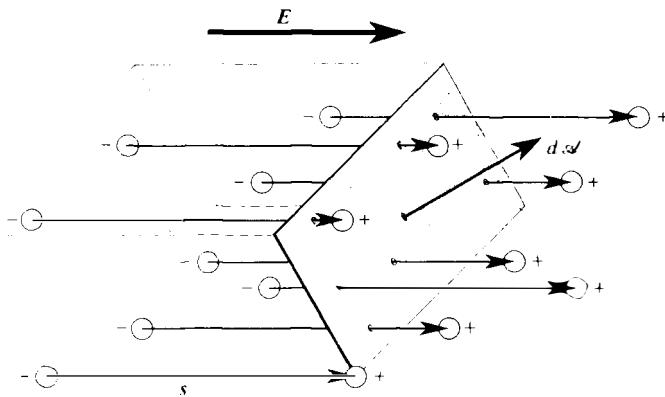


Fig. 9-1. Element of area $d\mathcal{A}$ situated inside a dielectric. The vector $d\mathcal{A}$ is normal to the element. Here E is the externally applied field plus the field of the dipoles within the dielectric. Under the action of this E , the molecules either stretch, or rotate and stretch, and a net charge $NQs \cdot d\mathcal{A}$ crosses the element of area.

$$\sigma_b = \frac{dQ}{d\mathcal{A}} = \mathbf{P} \cdot \hat{\mathbf{n}}, \quad (9-4)$$

where $\hat{\mathbf{n}}$ is the unit vector normal to the surface and pointing *outward*. Thus σ_b is equal in magnitude to the normal component of \mathbf{P} , pointing outward.

9.3.2 The Bound Volume Charge Density ρ_b

We now demonstrate that inside a dielectric the bound volume charge density ρ_b is equal to $-\nabla \cdot \mathbf{P}$. The net bound charge that flows out of a volume v across an element $d\mathcal{A}$ of its surface is $\mathbf{P} \cdot d\mathcal{A}$, as we found above. The net bound charge that flows out of the closed surface of area \mathcal{A} delimiting a volume v entirely situated within the dielectric (so as to exclude surface charges) is thus

$$Q_{\text{out}} = \int_{\mathcal{A}} \mathbf{P} \cdot d\mathcal{A}, \quad (9-5)$$

and the net charge that remains within v must be $-Q_{\text{out}}$. If ρ_b is the volume density of the charge remaining within v , then

$$\int_v \rho_b dv = -Q_{\text{out}} = -\int_{\mathcal{A}} \mathbf{P} \cdot d\mathcal{A} = -\int_v \nabla \cdot \mathbf{P} dv. \quad (9-6)$$

We have used the divergence theorem to transform the second integral into the third. Since this equation applies to any volume chosen as above, the integrands are equal at every point in the dielectric and the *bound volume charge density* is

$$\rho_b = -\nabla \cdot \mathbf{P}. \quad (9-7)$$

9.3.3 The Polarization Current Density \mathbf{J}_b

The motion of bound charges under the action of a time-dependent electric field generates a *polarization current*. Consider a small surface $d\mathcal{A}$ situated inside a dielectric as in Fig. 9-1, but normal to \mathbf{P} . As the polarization increases from zero to \mathbf{P} , a net charge $dQ_b = P d\mathcal{A}$ crosses $d\mathcal{A}$ in the direction of \mathbf{P} . More generally, if \mathbf{P} increases by $d\mathbf{P}$ in a time interval dt and if $d\mathcal{A}$ is normal to $d\mathbf{P}$, then a current

$$I = \frac{dQ_b}{dt} = \frac{|\mathbf{d}\mathbf{P}| d\mathcal{A}}{dt} \quad (9-8)$$

flows through $d\mathcal{A}$ in the direction of \mathbf{P} .

Thus if, at a given point in space, \mathbf{P} is a function of time, the motion of bound charge results in a *polarization current density*

$$\mathbf{J}_b = \frac{\partial \mathbf{P}}{\partial t}. \quad (9-9)$$

9.4 THE ELECTRIC FIELD OF A POLARIZED DIELECTRIC

We have seen that polarization causes charges to accumulate, either at the surface of a dielectric or inside, with charge densities σ_b and ρ_b , respectively. Now Coulomb's law applies to *any* net charge density, regardless of any matter that may be present. The potential V ascribable to a polarized dielectric that occupies a volume v' bounded by a surface of area \mathcal{A}' is therefore the same as if the bound charges were located in a vacuum:

$$V = \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\rho_b dv'}{r} + \frac{1}{4\pi\epsilon_0} \int_{\mathcal{A}'} \frac{\sigma_b d\mathcal{A}'}{r}, \quad (9-10)$$

where r is the distance between the element of bound charge at $P'(x', y', z')$ and the point $P(x, y, z)$ where one calculates V .

If there are also free charges present, then one adds similar integrals for the free charges. Therefore

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\rho \hat{r}}{r^2} dv' + \frac{1}{4\pi\epsilon_0} \int_{\mathcal{A}'} \frac{\sigma \hat{r}}{r^2} d\mathcal{A}', \quad (9-11)$$

where \hat{r} points from P' to P and where ρ and σ are total charge densities, free plus bound.

It does *not* follow that one can calculate \mathbf{E} from ρ and σ because both charge distributions depend on \mathbf{E} .

Example | THE FIELD OF AN INFINITE-SHEET ELECTRET

An *electret* is the electrical equivalent of a permanent magnet. In most dielectrics the polarization vanishes immediately upon removal of the electric field, but some dielectrics retain their polarization for long periods. For example, certain polymers have extrapolated lifetimes of several thousand years at room temperature.

As a rule, electrets have the form of sheets, with the polarization normal to the surface. One way of *poling* an electret is to place the material in an electric field of about 10^8 volts/meter at about 100°C . A bound surface charge builds up as the molecules orient themselves, and the sample is then cooled down to room temperature without removing the electric field.

One commonly used material is polyvinylidene fluoride (PVF_2). This is a polymer composed of a chain of CH_2-CF_2 units. Its remanent polarization is typically 50 to 70 millicoulombs/meter². This material is used in various types of transducers, microphones for example, because it also has the property of going into a metastable polarized state when stretched.

For a sheet electret, neglecting edge effects, the polarization P is uniform and equal to σ_b as in Fig. 9-2. Then ρ_b is zero, from Eq. 9-7. Inside the sheet,

$$E = \frac{\sigma_b}{\epsilon_0}. \quad (9-12)$$

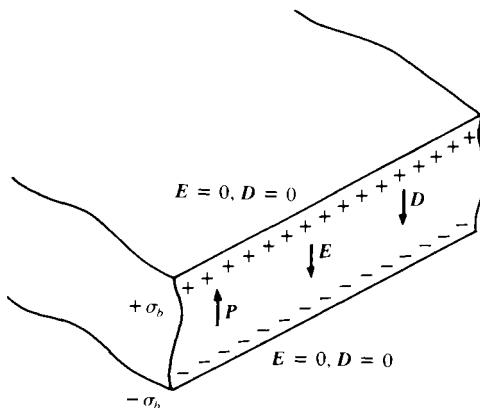


Fig. 9-2. Section through a sheet electret. Vectors \mathbf{E} and \mathbf{P} point in opposite directions.

9.5 GAUSS'S LAW

Say a given volume v contains various dielectrics, some of which may be partly inside and partly outside. The total free and bound charge within v is $Q_f + Q_b = Q$. There are no surface charges on the surface of v . Then *Gauss's law* relates the outward flux of \mathbf{E} through the surface of area \mathcal{A} to the net enclosed charge Q :

$$\boxed{\int_{\mathcal{A}} \mathbf{E} \cdot d\mathcal{A} = \frac{Q}{\epsilon_0}.} \quad (9-13)$$

This is Gauss's law in its more general form.

If the volume v lies entirely inside a dielectric, there are no surface charges and

$$\int_{\mathcal{A}} \mathbf{E} \cdot d\mathcal{A} = \frac{1}{\epsilon_0} \int_v (\rho_f + \rho_b) dv = \frac{1}{\epsilon_0} \int_v \rho dv, \quad (9-14)$$

where $\rho = \rho_f + \rho_b$ is the *total charge density*. Applying the divergence theorem to the surface integral of \mathbf{E} gives the volume integral of $\nabla \cdot \mathbf{E}$. Equating the integrands then yields

$$\boxed{\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0},} \quad (9-15)$$

which is again Gauss's law, expressed in differential form.

This is one of *Maxwell's four fundamental equations of electromagnetism*.

9.6 POISSON'S AND LAPLACE'S EQUATIONS FOR V IN DIELECTRICS

Since $\mathbf{E} = -\nabla V$, from Sec. 3.4, it follows that

$$\nabla^2 V = -\rho/\epsilon_0. \quad (9-16)$$

This is *Poisson's equation* for V in dielectrics.

Laplace's equation applies to regions where the total electric charge density is zero:

$$\nabla^2 V = 0, \quad (9-17)$$

as in Sec. 4.1.

9.7 THE ELECTRIC FLUX DENSITY \mathbf{D} AND ITS DIVERGENCE

According to Eq. 9-15,

$$\nabla \cdot \mathbf{E} = \frac{\rho_f + \rho_b}{\epsilon_0}. \quad (9-18)$$

But we found in Sec. 9.3.2 that $\rho_b = -\nabla \cdot \mathbf{P}$. Therefore

$$\nabla \cdot (\epsilon_0 \mathbf{E} + \mathbf{P}) = \rho_f. \quad (9-19)$$

We conclude that the vector

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \quad (9-20)$$

is such that its divergence is equal to the free volume charge density. This quantity is called the *electric flux density*. Thus

$$\nabla \cdot \mathbf{D} = \rho_f. \quad (9-21)$$

It follows that, for any volume v that lies entirely inside a dielectric and that encloses a free charge Q_f with no surface charges,

$$\int_v \nabla \cdot \mathbf{D} dv = \int_A \mathbf{D} \cdot d\mathcal{A} = \int_v \rho_f dv = Q_f. \quad (9-22)$$

This concurs with the relation $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$:

$$\int_A \mathbf{D} \cdot d\mathcal{A} = \epsilon_0 \int_A \mathbf{E} \cdot d\mathcal{A} + \int_A \mathbf{P} \cdot d\mathcal{A} = Q + \int_v \nabla \cdot \mathbf{P} dv \quad (9-23)$$

$$= Q - \int_v \rho_b dv = Q - Q_b = Q_f \quad (9-24)$$

The fact that the divergence of \mathbf{D} depends solely on the free charge density does *not* mean that \mathbf{D} itself depends only on ρ_f . To find \mathbf{D} , we must integrate Eq. 9-21 subject to whatever boundary conditions apply.

In using the divergences of \mathbf{E} , \mathbf{P} , and \mathbf{D} , we assume implicitly the existence of the space derivatives. If we have to deal with the interface between two media, where these derivatives do not exist, then we must use the integral form of Gauss's law, which is therefore more general.

Example

THE INFINITE-SHEET ELECTRET

Inside the electret of the example in Sec. 9.4, $\mathbf{E} = (\sigma_b/\epsilon_0)\hat{\mathbf{n}}$ and points downward. Also, \mathbf{P} points upward and its magnitude is σ_b . Therefore \mathbf{E} and \mathbf{P}/ϵ_0 are equal in magnitude, but point in opposite directions, as in Fig. 9-2.

Example

THE BAR ELECTRET

Assume that, inside the bar electret of Fig. 9-3, the polarization \mathbf{P} is uniform. (This is an interesting but unrealistic example because \mathbf{P} is in fact nonuniform.) Then $\nabla \cdot \mathbf{P}$ is zero and, from Eq. 9-7, ρ_b

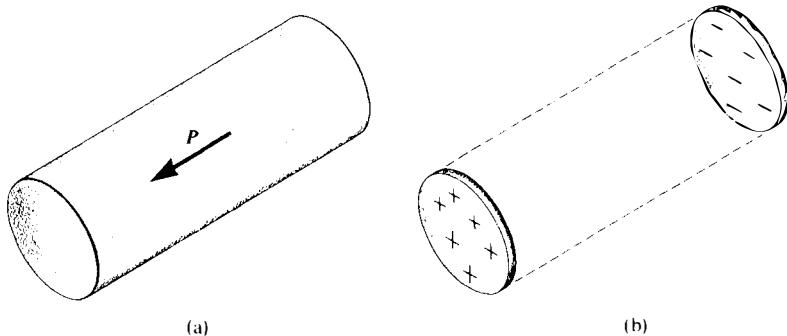


Fig. 9-3. (a) Idealized bar electret polarized uniformly parallel to its axis. (b) The \mathbf{E} field of the bar electret is the same as that of a pair of circular plates carrying uniform surface charge densities P of opposite polarities.

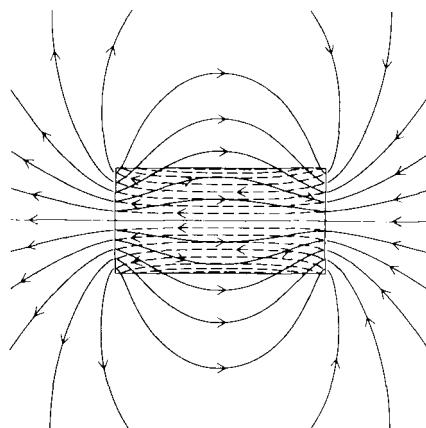


Fig. 9-4. Lines of \mathbf{E} (solid) for the idealized bar electret of Fig. 9-3. Lines of \mathbf{D} are shown dashed inside the electret; outside they follow the lines of \mathbf{E} . The polarization is uniform and points to the left.

is zero. There are bound charges solely on the end faces. Then the E -field, both inside and outside the bar electret, is that of a pair of circular sheets of charge of uniform charge densities $+P$ and $-P$, respectively, as in the figure. (The surface charge density on a conducting plate is larger near the edges than near the center.)

Outside, the lines of D are identical to the lines of E because $D = \epsilon_0 E$.

Inside, $D = \epsilon_0 E + P$, with E and P pointing in roughly opposite directions. The divergence of D is everywhere zero. See Fig. 9-4.

9.8 THE ELECTRIC SUSCEPTIBILITY χ_e

In most dielectrics P is proportional to E and points in the same direction. (This is *not* true of electrets, as we just saw!) Such dielectrics are *linear* and *isotropic*. So, in linear and isotropic dielectrics,

$$P = \epsilon_0 \chi_e E, \quad (9-25)$$

where χ_e is the *electric susceptibility* of the medium.

The dielectric is *homogeneous* if its electric susceptibility is independent of the coordinates.

9.9 THE RELATIVE PERMITTIVITY ϵ_r

In linear and isotropic dielectrics, Eq. 9-25 applies and

$$D = \epsilon_0 E + P = \epsilon_0 (1 + \chi_e) E = \epsilon_0 \epsilon_r E = \epsilon E, \quad (9-26)$$

where

$$\epsilon_r = 1 + \chi_e = \frac{\epsilon}{\epsilon_0} \quad (9-27)$$

is the *relative permittivity*. This quantity is dimensionless and larger than unity. The quantity ϵ is the *permittivity*. Its dimensions are the same as those of ϵ_0 , namely, farads per meter. Thus, for linear and isotropic dielectrics,

$$P = \epsilon_0 \chi_e E = \epsilon_0 (\epsilon_r - 1) E. \quad (9-28)$$

Table 9-1 shows the value of ϵ_r for some common dielectrics at three widely spaced frequencies.

The relative permittivities of gases at normal temperature and pressure are only slightly larger than unity. For example, $\epsilon_r = 1.000536$ for air at normal temperature and pressure.

Table 9-1 Relative permittivities of various materials near 25°C

TYPE	FREQUENCY		
	100	10 ⁶	10 ¹⁰
Barium titanate	1250	1140	100
Benzene	2.28	2.28	2.28
Birch (yellow)		2.7	1.95
Butyl rubber	2.43	2.40	2.38
Carbon tetrachloride	2.17	2.17	2.17
Fused silica	3.78	3.78	3.78
Glass (soda borosilicate)	5.0	4.84	4.82
Ice		4.15	3.20
Lucite	3.20	2.63	2.57
Neoprene	6.70	6.26	4.0
Polyethylene	2.26	2.26	2.26
Polystyrene	2.56	2.56	2.54
Sodium chloride		5.90	5.90
Soil (dry loam)		2.59	2.55
Steatite	6.55	6.53	6.51
Styrofoam	1.03	1.03	1.03
Teflon	2.1	2.1	2.08
Water	81	78.2	34
Wheat (red, winter)		4.3	2.6

The relative permittivities of good conductors are unknown because their conduction currents are so much larger than their polarization currents. One may assume that their permittivities are of the order of three to five, as for ordinary dielectrics.

Example

THE FREE AND BOUND VOLUME CHARGE DENSITIES

The presence of a free volume charge density ρ_f polarizes a dielectric and gives rise to a ρ_b . If the dielectric is homogeneous, isotropic, and linear, then

$$\mathbf{P} = \mathbf{D} - \epsilon_0 \mathbf{E} = \left(1 - \frac{1}{\epsilon_r}\right) \mathbf{D}, \quad (9-29)$$

$$\nabla \cdot \mathbf{P} = \left(1 - \frac{1}{\epsilon_r}\right) \nabla \cdot \mathbf{D}, \quad (9-30)$$

$$\rho_b = -\left(1 - \frac{1}{\epsilon_r}\right) \rho_f. \quad (9-31)$$

If ρ_f is zero, then ρ_b is zero and $\nabla \cdot \mathbf{P}$ is also zero. If ρ_f is not zero,

which is a rare occurrence, the total charge density ρ has the same sign as ρ_f , but it is smaller:

$$\rho = \rho_f + \rho_b = \frac{\rho_f}{\epsilon_r}. \quad (9-32)$$

Example

THE FREE AND BOUND SURFACE CHARGE DENSITIES AT THE INTERFACE BETWEEN A DIELECTRIC AND A CONDUCTOR

At the interface between a conductor and a linear and isotropic dielectric, the conductor carries a σ_f and the dielectric a σ_b . If the field is static, $E = 0$ and $D = 0$ inside the conductor. Since the tangential component of E is continuous across the interface, as we shall see in Sec. 10.2.3, both D and E are normal to the interface in the dielectric, close to the conductor.

Now imagine a thin Gaussian volume G straddling an area \mathcal{A} of the interface as in Fig. 9-5. Gauss's law, as stated in Sec. 9.5, tells us that the outward flux of D in Fig. 9-5 is equal to the enclosed free charge $\sigma_f \mathcal{A}$. Take the upward direction as positive. Then, since $P = \sigma_b$ in this example, from Sec. 9.3.1,

$$\sigma_f = D = \epsilon_r \epsilon_0 E = \epsilon_0 E + P = \epsilon_0 E - \sigma_b \quad (9-33)$$

and

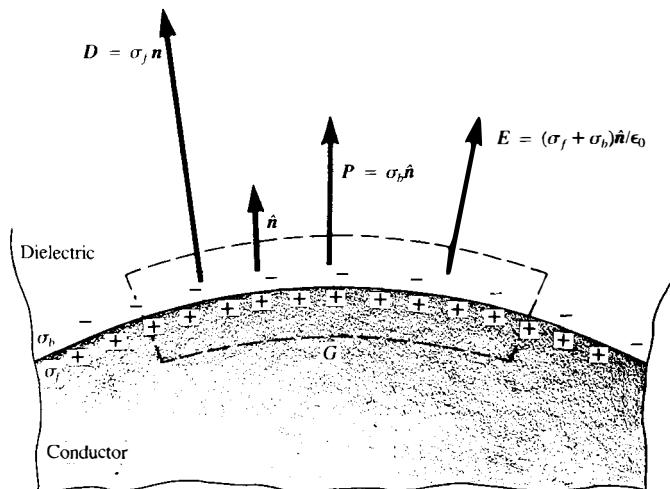


Fig. 9-5. At the interface between a dielectric and a conductor there exists both a free surface charge density σ_f on the conductor and a bound surface charge density σ_b on the dielectric. If σ_f is positive, as it is here, then σ_b is negative and the vectors D , P , E point in the directions shown.

$$E = \frac{\sigma_f + \sigma_b}{\epsilon_0}. \quad (9-34)$$

Observe that Gauss's law for E , as stated in Eq. 9-13, leads to the same value of E .

Since $D = \epsilon_r \epsilon_0 E$,

$$\sigma_f = D = \epsilon_r (\sigma_f + \sigma_b), \quad \sigma_b = -\left(1 - \frac{1}{\epsilon_r}\right) \sigma_f, \quad (9-35)$$

$$\sigma = \sigma_f + \sigma_b = \frac{\sigma_f}{\epsilon_r}. \quad (9-36)$$

Compare with Eqs. 9-31 and 9-32.

Example

DIELECTRIC-INSULATED PARALLEL-PLATE CAPACITOR

Figure 9-6 shows a cross-section of a parallel-plate capacitor. We have shown air spaces on either side of the dielectric sheet so as to render our discussion more instructive, but as a rule the plates are in contact with the dielectric.

We assume static charges and fields. Also, almost invariably, $\rho_f = 0$ in the dielectric and thus $\rho_b = 0$ and $\nabla \cdot \mathbf{P} = 0$, from the first example in Sec. 9.9 and from Sec. 9.3.2. Neglecting edge effects, the surface charge densities σ_f and σ_b are uniform. Similarly, the vectors \mathbf{P} , \mathbf{E} , and \mathbf{D} are uniform within the dielectric and uniform within the air spaces.

The Gaussian volume G straddles the air-conductor interface. Within the conductor, $\mathbf{E} = 0$ and $\mathbf{D} = 0$. Therefore, in the air space,

$$E_a = \frac{\sigma_f}{\epsilon_0}, \quad D_a = \sigma_f. \quad (9-37)$$

Similarly, for the Gaussian volume G' ,

$$E_d = \frac{\sigma_f + \sigma_b}{\epsilon_0}, \quad D_d = \sigma_f. \quad (9-38)$$

But

$$E_d = \frac{D_d}{\epsilon_r \epsilon_0} = \frac{\sigma_f}{\epsilon_r \epsilon_0}. \quad (9-39)$$

It follows that

$$\sigma_b = -\left(1 - \frac{1}{\epsilon_r}\right) \sigma_f. \quad (9-40)$$

Of course, $|\sigma_b| = P$, from Sec. 9.3.1.

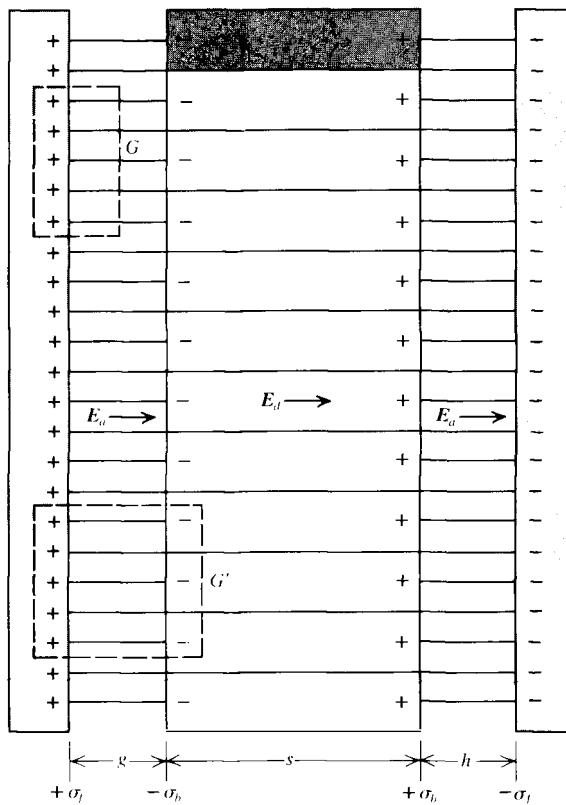


Fig. 9-6. Pictorial representation of the manner in which \mathbf{E} varies inside a parallel-plate capacitor containing a sheet of dielectric with $\epsilon_r = 2$. Of course, individual lines of \mathbf{E} do not exist. Here $|\sigma_b| = |\sigma_f|/2$.

The voltage V across the capacitor is thus

$$V = \frac{\sigma_f}{\epsilon_0} (g + h) + \frac{\sigma_f}{\epsilon_r \epsilon_0} s = \frac{Q}{\epsilon_0 \mathcal{A}} \left(g + h + \frac{s}{\epsilon_r} \right), \quad (9-41)$$

where $\sigma_f = Q/\mathcal{A}$, Q being the magnitude of the charge and \mathcal{A} the area of one plate.

The capacitance is

$$C = \frac{Q}{V} = \frac{\sigma_f \mathcal{A}}{V} = \frac{\epsilon_0 \mathcal{A}}{g + h + s/\epsilon_r}. \quad (9-42)$$

Without air spaces on either side of the dielectric, $g + h = 0$ and

$$V = \frac{Q}{\epsilon_r \epsilon_0 \mathcal{A}} s, \quad (9-43)$$

$$C = \frac{\epsilon_r \epsilon_0 \mathcal{A}}{s}. \quad (9-44)$$

This is ϵ_r times larger than the capacitance that we found in Sec. 6.4.1 for an air-insulated parallel-plate capacitor.

The measurement of a capacitance, with and without a dielectric, provides a convenient way of measuring a relative permittivity.

Example

THE FIELD OF FREE CHARGES EMBEDDED IN A DIELECTRIC

Imagine a small conducting sphere carrying a free charge Q_f and embedded in an infinite, homogeneous, isotropic, linear, and stationary (HILS) dielectric. The surface of the dielectric in contact with the charged sphere carries a bound charge of the opposite sign. At each point on the surface of the sphere Eq. 9-36 applies, and the net charge is Q_f/ϵ_r . Thus, at any point inside the dielectric at a distance r from the charged conducting sphere,

$$E = \frac{Q_f/\epsilon_r}{4\pi\epsilon_0 r^2} = \frac{Q_f}{4\pi\epsilon_r \epsilon_0 r^2}. \quad (9-45)$$

Similarly, if a given distribution of free charges ρ_f is located inside a dielectric as above, the net charge density is ρ_f/ϵ_r , and the E at any point is E_0/ϵ_r , where E_0 is the field that would obtain if the free charge distribution were situated in a vacuum.

If the dielectric is not infinite, then we must take into account the field of the bound charges at the surface.

9.10 THE DISPLACEMENT CURRENT DENSITY $\partial\mathbf{D}/\partial t$

Figure 9-7 shows a parallel-plate capacitor connected to a source of alternating voltage. The plates have an area \mathcal{A} and are spaced by a distance s . We assume that edge effects are negligible. Then

$$I = \frac{V}{Z} = \frac{Es}{1/j\omega C} = j\omega Es \frac{\epsilon_r \epsilon_0 \mathcal{A}}{s} \quad (9-46)$$

$$= \mathcal{A}j\omega\epsilon_r\epsilon_0 E = \mathcal{A}j\omega D = \mathcal{A} \frac{dD}{dt} = \frac{dQ}{dt}. \quad (9-47)$$

The current density in the capacitor is therefore $j\omega D$ or, more generally, $\partial\mathbf{D}/\partial t$. This quantity is called the *displacement current density* because \mathbf{D} was formerly called the *electric displacement*.

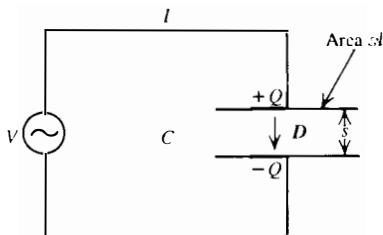


Fig. 9-7. Parallel-plate capacitor connected to a source of alternating voltage. The displacement current density in the dielectric is $\partial\mathbf{D}/\partial t$.

The displacement current density consists of two parts:

$$\frac{\partial \mathbf{D}}{\partial t} = \frac{\partial}{\partial t} (\epsilon_0 \mathbf{E} + \mathbf{P}) = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \frac{\partial \mathbf{P}}{\partial t}. \quad (9-48)$$

The second term on the right is the polarization current density of Sec. 9.3.3, which results from the motion of bound charges. The first term, $\epsilon_0 \partial \mathbf{E} / \partial t$, is nameless. It can exist even in a vacuum.

We shall have many occasions to refer to the displacement current density.

9.11 SUMMARY

In a *polar dielectric*, each molecule possesses a permanent dipole moment. Upon application of an electric field, the molecules orient themselves more or less, and their dipole moments increase. The molecules of a *nonpolar dielectric* acquire a dipole moment when subjected to an electric field.

The dipole moment per cubic meter at a given point is called the *electric polarization* \mathbf{P} .

Polarization causes bound charges to accumulate, with *surface* and *volume charge densities*

$$\sigma_b = \mathbf{P} \cdot \hat{\mathbf{n}}, \quad \rho_b = -\nabla \cdot \mathbf{P}, \quad (9-4), (9-7)$$

where $\hat{\mathbf{n}}$ is a unit vector that is normal to the surface and points *outward*.

In a fluctuating electric field the fluctuating polarization gives a *polarization current density*

$$\mathbf{J}_p = \frac{\partial \mathbf{P}}{\partial t}. \quad (9-9)$$

One can calculate \mathbf{E} both inside and outside dielectrics by treating both free and bound charges as if they were situated in a vacuum.

In dielectrics, *Gauss's law* states that, for a closed surface of area \mathcal{A} ,

$$\int_{\mathcal{A}} \mathbf{E} \cdot d\mathbf{A} = \frac{Q}{\epsilon_0}, \quad (9-13)$$

where Q is the total enclosed charge, free plus bound. The differential form of Gauss's law is

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}. \quad (9-15)$$

As a consequence, Poisson's equation for V is

$$\nabla^2 V = -\frac{\rho}{\epsilon_0}. \quad (9-16)$$

By definition, the *electric flux density* is

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \quad (9-20)$$

and

$$\nabla \cdot \mathbf{D} = \rho_f. \quad (9-21)$$

This is still another form of *Gauss's law*.

In linear and isotropic dielectrics,

$$\mathbf{P} = \epsilon_0 \chi_e \mathbf{E}, \quad (9-25)$$

where χ_e is a constant, independent of \mathbf{E} , called the *electric susceptibility*. Then

$$\mathbf{D} = \epsilon_0(1 + \chi_e) \mathbf{E} = \epsilon_r \epsilon_0 \mathbf{E} = \epsilon \mathbf{E}, \quad (9-26)$$

where ϵ_r is the *relative permittivity* and ϵ is the *permittivity* of the material. Also,

$$\mathbf{P} = \epsilon_0(\epsilon_r - 1) \mathbf{E}. \quad (9-28)$$

In a homogeneous, isotropic, and linear dielectric,

$$\rho_b = -\left(1 - \frac{1}{\epsilon_r}\right) \rho_f. \quad (9-31)$$

At the interface between a dielectric and a conductor, under static conditions,

$$\sigma_f = D = \epsilon_r \epsilon_0 E, \quad \sigma_b = -\left(1 - \frac{1}{\epsilon_r}\right) \sigma_f, \quad (9-33), \quad (9-35)$$

with the sign conventions of Fig. 9-5.

The *displacement current density* is $\partial \mathbf{D} / \partial t$.

PROBLEMS

9-1. (9.2) The dipole moment of a carbon atom

A sample of diamond has a density of 3.5×10^3 kilograms/meter³ and a polarization of 10^{-7} coulomb/meter².

(a) Calculate the average dipole moment per atom.

(b) What is the average separation between the centers of positive and negative charge? The carbon nucleus has a charge $+6e$ and is surrounded by six electrons. The diameter of an atom is of the order of 10^{-10} meter.

9-2. (9.3) Nonhomogeneous dielectric

Show that, in a nonhomogeneous dielectric, if $\rho_f = 0$, then $\rho_b = -(\epsilon_0 / \epsilon_r) \mathbf{E} \cdot \nabla \epsilon_r$.

9-3. (9.4) The field of a polarized dielectric sphere

The polarization \mathbf{P} of a given dielectric sphere is uniform. Here is an elegant way of calculating its field, both inside and outside.

Take the sum of the fields of two spherical charges of densities ρ and $-\rho$, displaced by a distance Δz , one with respect to the other, Δz being the distance between the centers of charge of the individual dipoles. Then $\mathbf{P} = \rho \Delta z \hat{z}$ if the vector \mathbf{P} points in the positive direction of the z -axis.

(a) Show that, inside, $\mathbf{E} = -\mathbf{P}/3\epsilon_0$ and is therefore uniform.

(b) Show that, outside, V and \mathbf{E} are those of a dipole of moment \mathbf{PV} situated at the center of the sphere of volume \mathcal{V} .

9-4. (9.6) Poisson's equation in nonhomogeneous media

Show that, in a nonhomogeneous medium, $\nabla^2 V + (\nabla V \cdot \nabla \epsilon) / \epsilon = -\rho_f / \epsilon$.

9-5. (9.7) The surface charge densities on a dielectric

The surface of a certain sample of dielectric carries a free charge of density of σ_f .

Show that, close to the surface, $E_{\text{air}} = \sigma_f / [(\epsilon_r + 1)\epsilon_0]$ and $E_{\text{dielectric}} = -\sigma_f / [(\epsilon_r + 1)\epsilon_0]$. We have chosen the outward direction as positive. Note that $D_{\text{air}} - D_{\text{dielectric}} = \sigma_f$, as expected. Also, in the dielectric, $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$.

9-6. (9.7) Electrically charged Lucite block

If high-energy electrons bombard a block of insulating material such as Lucite, the electrons penetrate the material and remain trapped inside. If one then gives the block a sharp knock with a conducting object, say with a center punch, the electrons escape, leaving a beautiful tree-like design where the plastic has broken down.

In one particular instance, a 0.1-microampere beam bombarded an area of 25 centimeter² of Lucite ($\epsilon_r = 3.2$) for 1 second, and essentially all the electrons were trapped about 6 millimeters below the surface in a region about 2 millimeters thick. The block was 12 millimeters thick.

In the following calculations, neglect edge effects and assume a uniform density for the trapped electrons. Assume also that both faces of the Lucite are in contact with grounded conducting plates.

- (a) What is the bound charge density in the charged region?
- (b) What is the bound charge density at the surface of the Lucite?
- (c) Sketch graphs of D , E , V as functions of position inside the dielectric.
- (d) Show that the potential at the center of the sheet of charge is about 4 kilovolts.
- (e) What is the energy stored in the block? Could the block explode?

9-7. (9.9) Capacitance with and without a dielectric

A conducting body A , of arbitrary shape, is grounded. Another conducting body B , of arbitrary shape and position, is maintained at a potential V . In air, the capacitance is C_0 . Show that the capacitance is $\epsilon_r C_0$ when the bodies are submerged in a large body of dielectric ϵ_r .

9-8. (9.9) The coaxial line

Figure 33-4 shows a section of coaxial line. A dielectric ϵ_r fills the space between the two conductors. From Probs. 6-5 and 9-7, the capacitance per unit length C' is $2\pi\epsilon_r\epsilon_0/\ln(R_2/R_1)$. The outer conductor is grounded, and the inner conductor is at the potential V .

- (a) Calculate the charge per unit length λ on the inner conductor.
 - (b) Show that the bound charge per unit length on the inner surface of the dielectric is $-\lambda(1 - 1/\epsilon_r)$.
- Thus the net charge per unit length at the radius R_1 is λ/ϵ_r . The bound charge per unit length on the outer surface of the dielectric is similarly $+\lambda(1 - 1/\epsilon_r)$, and the net charge per unit length at R_2 is $-\lambda/\epsilon_r$.
- (c) Show that the volume density of bound charge is zero.
 - (d) Draw graphs of D , E , and V as functions of the radius r from $r = R_1$ to $r = R_2$, for $V = 100$ volts, $R_1 = 1.00$ millimeter, $R_2 = 10.0$ millimeters, $\epsilon_r = 3.00$.

9-9. (9.9) A parallel-plate capacitor with a nonuniform dielectric

The dielectric of a parallel-plate capacitor has a permittivity that varies as $\epsilon_{r0} + ax$, where x is the distance from one plate. The area of a plate is \mathcal{A} , and their spacing is s .

- (a) Find the capacitance.
- (b) Show that, if ϵ_r varies from ϵ_{r0} to $2\epsilon_{r0}$, then C is 1.44 times larger than if a were zero.
- (c) Find P from the values of D and E for that case.
- (d) Deduce the value of ρ_b .
- (e) Now calculate ρ_b from the relation given in Prob. 9-2.
- (f) Draw curves of E , ρ_b , and P as functions of x for $\epsilon_{r0} = 3.00$, $a = \epsilon_{r0}/s$, $s = 1.00$ millimeter when the applied voltage is 1.00 volt.

9-10. (9.9) The resistance and the capacitance between two electrodes

When the space between the plates of a parallel-plate capacitor is filled with a dielectric ϵ , it has a capacitance of C farads. If the dielectric is replaced by a material whose resistivity $\rho = 1/\sigma$ is much smaller than that of the electrodes, the resistance between the electrodes is R ohms.

(a) Show that $RC = \rho\epsilon$, neglecting edge effects.

(b) Show that this result also applies to cylindrical and spherical capacitors.

(c) Show that this result applies to any pair of electrodes submerged in a medium whose resistivity ρ is much larger than that of the electrodes.

You should be able to show that the field is unaffected by the conductivity, with the above restriction.

One important application of this fact is the electrolytic plotting tank, which is used for plotting electric fields in two, and in some cases three, dimensions.

If the medium occupies only the region between the electrodes, then RC is again equal to $\rho\epsilon$, except that C does not include the fringing field and is therefore smaller than the true capacitance.

(d) The capacitance per unit length between two parallel wires of diameter d and separated by a distance D is $C' = \pi\epsilon/\cosh^{-1}(D/d)$. Find the conductance between parallel wires 10 millimeters in diameter separated by 100 millimeters and submerged in sea water ($\sigma = 5$).

9-11. (9.9) The current between two electrodes in a conducting medium.

See Prob. 9-10. The electrodes of part (c) are initially charged, and discharge through the medium. Show that $J_f + \partial D/\partial t = 0$.

CHAPTER 10

ELECTRIC FIELDS VIII

Dielectric Materials B: Real Dielectrics, Continuity Conditions at an Interface, and Stored Energy

*10.1 REAL DIELECTRICS	193
*10.1.1 LOSSY DIELECTRICS	193
*10.1.2 THE FREQUENCY AND TEMPERATURE DEPENDENCE OF ϵ_r	195
*10.1.3 ANISOTROPY	195
*10.1.4 FERROELECTRICITY	196
*10.1.5 HYSTERESIS	197
10.2 THE CONTINUITY CONDITIONS AT AN INTERFACE	197
10.2.1 THE POTENTIAL V	197
10.2.2 THE NORMAL COMPONENT OF D	197
10.2.3 THE TANGENTIAL COMPONENT OF E	198
10.2.4 BENDING OF LINES OF E AT AN INTERFACE	199
10.3 THE POTENTIAL ENERGY OF A CHARGE DISTRIBUTION IN THE PRESENCE OF DIELECTRICS	200
10.3.1 THE POTENTIAL ENERGY EXPRESSED IN TERMS OF ρ_f AND V	200
<i>Example: THE DIELECTRIC-INSULATED PARALLEL-PLATE CAPACITOR</i>	201
10.3.2 THE ENERGY DENSITY EXPRESSED IN TERMS OF E AND D	202
<i>Example: THE DIELECTRIC-INSULATED PARALLEL-PLATE CAPACITOR</i>	203
10.3.3 THE ENERGY DENSITY ASSOCIATED WITH POLARIZATION	203
10.4 ELECTRIC FORCES IN THE PRESENCE OF DIELECTRICS	204
10.5 SUMMARY	204
PROBLEMS	205

In this second and last chapter on dielectric materials, we first describe briefly some characteristics of real dielectrics: finite conductivity, frequency and temperature dependence of ϵ_r , anisotropy, ferroelectricity, and hysteresis. Then we study the continuity conditions at the interface between two media. We shall require these conditions on many occa-

sions. Finally, we return to the energy stored in an electric field, this time in the presence of dielectrics.

*10.1 REAL DIELECTRICS

Some dielectrics, such as polyethylene, are close to ideal: they are homogeneous, linear, and isotropic; their conductivity is close to zero; and their relative permittivity is practically independent of frequency. However, most dielectrics exhibit a more complex behavior.

*10.1.1 Lossy Dielectrics

Dielectrics that are somewhat conducting are said to be *lossy*. For example, most natural substances, such as wood or wheat, show a slight conductivity that is associated with the presence of water.

With alternating electric fields, it is convenient to express the conductivity in terms of a complex permittivity as follows. Consider a parallel-plate capacitor containing such a material, as in Fig. 10-1. The current is the same as if one had an ideal nonconducting dielectric in the capacitor, with a resistance

$$R = \frac{s}{\sigma A} \quad (10-1)$$

connected in parallel, σ being the conductivity of the dielectric (Sec. 4.3.1).

Then, from Sec. 9.9.3,

$$I = Vj\omega C + \frac{V}{R} = Vj\omega \frac{\epsilon_r \epsilon_0 A}{s} + V \frac{\sigma A}{s}. \quad (10-2)$$

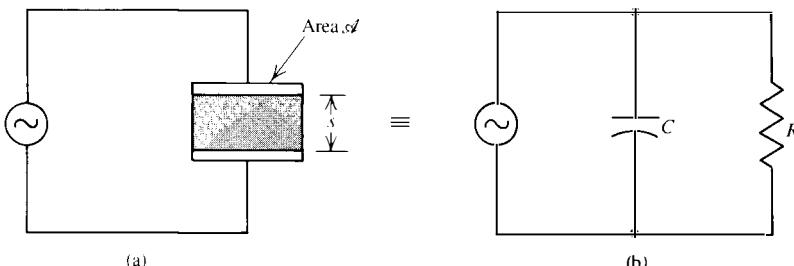


Fig. 10-1. A parallel-plate capacitor containing a lossy dielectric as in (a) is equivalent to an ideal lossless capacitor, in parallel with a resistance R , as in (b).

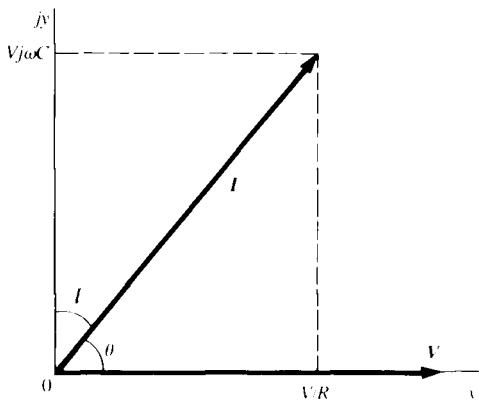


Fig. 10-2. The phasor I for the circuit of Fig. 10-1(b).

It is convenient to define ϵ_r differently and to rewrite this as follows:

$$I = \frac{Vj\omega(\epsilon'_r - j\epsilon''_r)\epsilon_0 A}{s} = \frac{Vj\omega\epsilon_r\epsilon_0 A}{s}. \quad (10-3)$$

We have now extended the definition of the relative permittivity to include conduction:

$$\epsilon_r = \epsilon'_r - j\epsilon''_r, \quad \epsilon''_r = \frac{\sigma}{\epsilon_0 \omega}. \quad (10-4)$$

This new ϵ_r is the *complex relative permittivity* whose real part is the ϵ_r of Eq. 10-2. Note the negative sign before the imaginary part. The imaginary part ϵ''_r pertains to the dissipation of energy as heat, and the real part ϵ'_r to the storage of energy in the electric field.

It is the custom to set

$$\epsilon_r = |\epsilon_r| \exp(-jl), \quad (10-5)$$

where l is the *loss angle*, and to call the ratio

$$\frac{\epsilon''_r}{\epsilon'_r} = \tan l \quad (10-6)$$

the *loss tangent* of the dielectric.

Unless specified otherwise, the term *relative permittivity*, as in Table 9-1, refers to the real part of ϵ_r .

Figure 10-2 shows the phasors \mathbf{V} and \mathbf{I} , the loss angle ℓ , and the complementary angle θ , known as the *power-factor angle*.

*10.1.2 The Frequency and Temperature Dependence of ϵ_r

Generally speaking, ϵ'_r decreases with frequency, as in Table 9-1, whereas ϵ''_r increases with frequency. This is only a rough rule of thumb because atoms and molecules exhibit resonances, especially above 1 gigahertz. The frequencies that are used for *dielectric heating* range from a few megahertz to a few gigahertz. Domestic microwave ovens usually operate at 2.45 gigahertz. They act on the water molecules in food.

A given dielectric can polarize through more than one of the processes described in Sec. 9.1, and the relative importance of a given process can vary with frequency. For example, water has a relative permittivity of 81 in an electrostatic field and 1.8 at optical frequencies. The large static value results from the orientation of the permanent dipole moments, but the rotational inertia of the molecules prevents any significant response at optical frequencies. Similarly, the relative permittivity of sodium chloride is 5.9 in an electrostatic field and 2.3 at optical frequencies. The larger static value comes from ionic motion, which again is impossible at high frequencies.

As a rule, the permittivity of a substance increases by a large factor upon melting. For example, the ϵ_r of nitrobenzene increases from 3 to 35 at the melting point near 279 kelvins. In the solid, the permanent dipoles are frozen in the lattice and cannot rotate under the influence of an applied field.

Observe that since ϵ_r is a function of frequency, it is strictly definable only for a pure sine wave.

*10.1.3 Anisotropy

Crystalline solids commonly possess different dielectric properties in different crystal directions, because the ions can move more easily in some directions than in others. As a result, their susceptibility depends on direction, and \mathbf{P} is not necessarily in the same direction as \mathbf{E} .

Then the x -component of \mathbf{P} is of the form

$$P_x = \epsilon_0(\chi_{exx}E_x + \chi_{exy}E_y + \chi_{exz}E_z), \quad (10-7)$$

and similarly for the other two components. All three components of \mathbf{P} depend on all three components of \mathbf{E} . The susceptibility thus has nine components and is a *tensor*. Actually, there are only six independent

components, and if one chooses the coordinate axes properly with respect to the crystal, these six components reduce to three.

If the various components of the susceptibility are *not* functions of E , then \mathbf{P} is a linear function of the components of E and the dielectric is linear.

In anisotropic dielectrics, it is still true that

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}, \quad \nabla \cdot \mathbf{D} = \nabla \cdot (\epsilon_0 \mathbf{E} + \mathbf{P}) = \rho_f, \quad (10-8)$$

but the three vectors are usually not parallel.

*10.1.4 Ferroelectricity

Ferroelectric dielectrics have the peculiar property of exhibiting spontaneous polarization over microscopic crystalline regions called *domains*. The word *ferroelectric* originates from the fact that their behavior is in several respects similar to that of *ferromagnetic* substances (Sec. 21.1). When placed in an electric field, domains that happen to be correctly polarized grow at the expense of neighboring domains and eventually coalesce.

If the temperature of a ferroelectric substance increases slowly, the spontaneous polarization varies, usually in some complex fashion, and

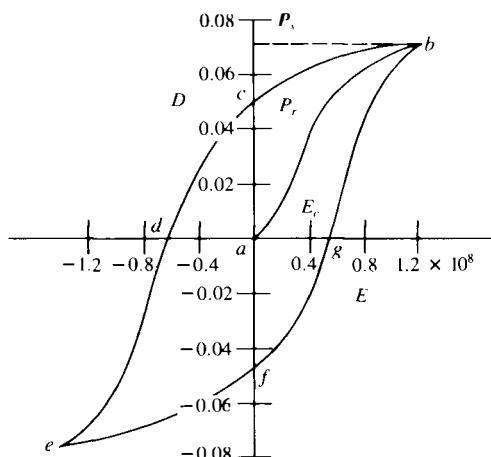


Fig. 10-3. Hysteresis curve for a typical ferroelectric material. Starting with an unpolarized sample and $E = 0$ at point a , an increase in E increases D to point b . Then decreasing E to zero leaves a residual D at c , etc. The value of D thus depends not only on the value of E , but also on the previous history of the specimen.

eventually disappears above the *Curie temperature*. This temperature is a characteristic of the material.

The main advantage of the ferroelectric dielectrics in common use in *ceramic capacitors* is that they have large relative permittivities, ranging up to 10,000. As a rule, these dielectrics are compounds of titanium.

Some ferroelectric dielectrics possess definite temperature coefficients that are useful in certain applications.

*10.1.5 Hysteresis

Ferroelectric materials also exhibit a property called *hysteresis*, illustrated in Fig. 10-3. Then the ratio D/E depends on the previous history of the material. Such materials are thus said to be *nonlinear* because D is not proportional to E .

The relative permittivity $\epsilon_r = D/\epsilon_0 E$ of a given ferroelectric sample thus has no definite value and can even be negative. When one quotes relative permittivities for ferroelectric materials, as we have done above and in Table 9-1, one refers to the order of magnitude of $D/\epsilon_0 E$, some distance from the origin $E = 0$, $D = 0$, in the first and third quadrants.

10.2 THE CONTINUITY CONDITIONS AT AN INTERFACE

10.2.1 The Potential V

The potential V is continuous across the boundary between two media. Otherwise, a discontinuity would imply an infinitely large E , which is physically impossible.

10.2.2 The Normal Component of \mathbf{D}

Consider a short imaginary cylinder spanning the interface, and of cross section \mathcal{A} as in Fig. 10-4. The top and bottom faces of the cylinder are parallel to the boundary and close to it. The interface carries a free surface charge density σ_f .

According to Gauss's law (Sec. 9.5), the net flux of \mathbf{D} coming out of the cylinder is equal to the enclosed free charge. Now the only flux of \mathbf{D} is that through the top and bottom faces because the height of the cylinder is small. If now the area \mathcal{A} is not too large, \mathbf{D} is approximately uniform over it, and then

$$(D_{2n} - D_{1n})\mathcal{A} = \sigma_f \mathcal{A}, \quad (\mathbf{D}_2 - \mathbf{D}_1) \cdot \hat{\mathbf{n}} = \sigma_f, \quad (10-9)$$

where $\hat{\mathbf{n}}$ is the unit vector normal to the interface and pointing from medium 2 to medium 1.

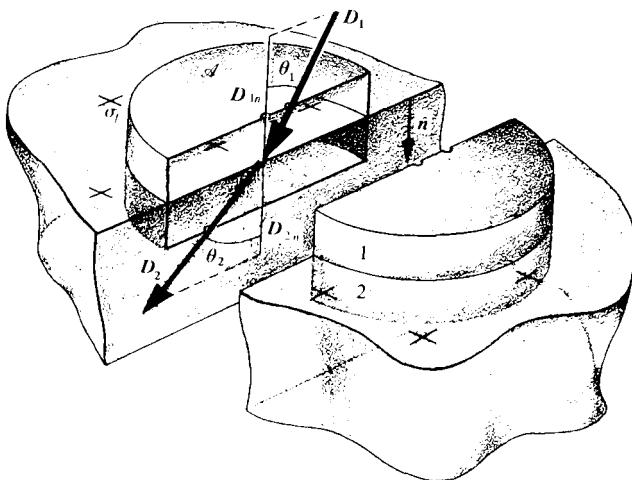


Fig. 10-4. Imaginary cylinder straddling the interface between media 1 and 2 and delimiting an area A' . The difference $D_{2n} - D_{1n}$ between the *normal* components of \mathbf{D} is equal to the free surface charge density σ_f .

As a rule, the boundary between two dielectrics does not carry *free* charges, and then the normal component of \mathbf{D} is continuous across the interface. Thus the normal component of \mathbf{E} is discontinuous.

On the other hand, if one medium is a conductor and the other a dielectric, and if \mathbf{D} is not a function of the time, then $\mathbf{D} = 0$ in the conductor and $D_n = \sigma_f$ in the dielectric. If \mathbf{D} is a function of the time, Eq. 10.9 still applies, but \mathbf{D} is not zero in the conductor.

10.2.3 The Tangential Component of \mathbf{E}

Consider now the path shown in Fig. 10-5, with two sides of length L parallel to the boundary and close to it. The other two sides are infinitesimal. If L is short, \mathbf{E} does not vary significantly over that distance, and integrating over the path yields

$$\oint \mathbf{E} \cdot d\mathbf{l} = E_{1t}L - E_{2t}L. \quad (10-10)$$

Now, from Sec. 3.4 this line integral is zero, and thus

$$E_{1t} = E_{2t}\dagger, \quad \text{or} \quad (\mathbf{E}_1 - \mathbf{E}_2) \times \hat{\mathbf{n}} = 0, \quad (10-11)$$

[†] The line integral of $\mathbf{E} \cdot d\mathbf{l}$ is zero in electrostatic fields. More generally, it is equal to minus the time derivative of the magnetic flux linking the path of integration, as we shall see in Sec. 23.4. However, since the enclosed area is zero, the flux is zero, and this equation is general.

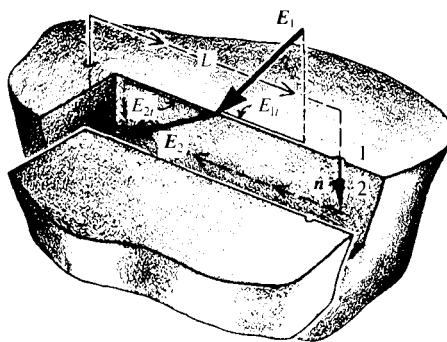


Fig. 10-5. Closed path of integration spanning the interface between media 1 and 2. The tangential components of \mathbf{E} are equal: $E_{1t} = E_{2t}$.

with $\hat{\mathbf{n}}$ defined as above. The tangential component of \mathbf{E} is continuous across any interface.

10.2.4 Bending of Lines of \mathbf{E} at an Interface

The \mathbf{E} and \mathbf{D} vectors change direction at the interface between two different linear and isotropic dielectrics. In Fig. 10-6, if there is zero free surface charge density, the continuity of the normal component of \mathbf{D} requires that

$$D_1 \cos \theta_1 = D_2 \cos \theta_2 \quad (10-12)$$

or that

$$\epsilon_r \epsilon_0 E_1 \cos \theta_1 = \epsilon_r \epsilon_0 E_2 \cos \theta_2. \quad (10-13)$$

Also, from the continuity of the tangential component of \mathbf{E} ,

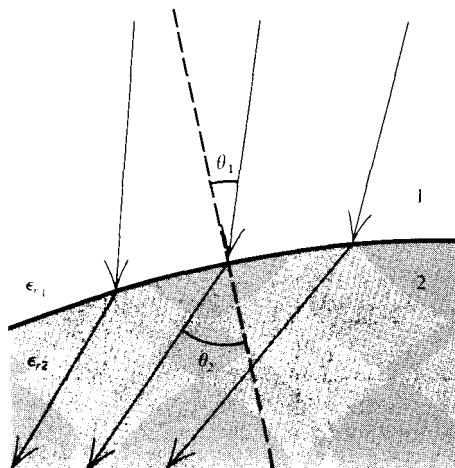


Fig. 10-6. Lines of \mathbf{E} crossing the interface between media 1 and 2. The permittivity of medium 2 is larger than that of medium 1.

$$E_1 \sin \theta_1 = E_2 \sin \theta_2. \quad (10-14)$$

Dividing the third equation by the second gives

$$\frac{\tan \theta_1}{\epsilon_{r1}} = \frac{\tan \theta_2}{\epsilon_{r2}}. \quad (10-15)$$

The larger angle from the normal is in the medium with the larger permittivity. The lines of \mathbf{D} "prefer" the medium with the higher permittivity. This rule does not apply to electrets. See Prob. 10-10.

Since the *normal* component of \mathbf{E} is discontinuous, there exists a bound surface charge density. From Gauss's law,

$$\sigma_b = \epsilon_0 E_2 \cos \theta_2 - \epsilon_0 E_1 \cos \theta_1 = \epsilon_0 E_2 \cos \theta_2 \left(1 - \frac{\epsilon_{r2}}{\epsilon_{r1}}\right). \quad (10-16)$$

For example, if medium 1 is air and medium 2 is a dielectric, then the bound surface charge density is negative if \mathbf{E} points into the surface of the dielectric.

10.3 THE POTENTIAL ENERGY OF A CHARGE DISTRIBUTION IN THE PRESENCE OF DIELECTRICS

In Chap. 6 we calculated the energy stored in an electric field when the charges reside either on conductors or in free space.

The addition of dielectrics to the field constitutes a major complication, but the problem becomes tractable with the following assumptions.

Our first two assumptions are realistic. (1) We assume that both the free and bound charge distributions are of finite extent. This makes the potential V equal to zero at infinity. (2) We assume that the dielectrics are linear (\mathbf{D} is proportional to \mathbf{E}), but not necessarily homogeneous or isotropic.

We also assume that the dielectrics are fixed in position and rigid. In other words, the dielectrics do not distort under the electric forces. This assumption is *not* realistic, but we can do no better in the present context.

10.3.1 The Potential Energy Expressed in Terms of ρ_f and V

At any point,

$$V = \frac{1}{4\pi\epsilon_0} \int_v \frac{\rho_f + \rho_b}{r} dv + \frac{1}{4\pi\epsilon_0} \int_{\mathcal{A}} \frac{\sigma_f + \sigma_b}{r} d\mathcal{A}, \quad (10-17)$$

as in Sec. 9.4. Since all the dielectrics in the field are linear, by hypothesis, ρ_b is proportional to ρ_f and σ_b to σ_f (Sec. 9.9, examples).

We start with a zero charge density everywhere and gradually pull in charges from infinity. We disregard surface charges for the moment. Let the final potential and the final volume free charge density at a point be, respectively, V and ρ_f . At a given moment, the free volume charge density is $\alpha\rho_f$, where α increases gradually from zero to unity.

Since

$$\nabla \cdot \mathbf{D} = \nabla \cdot (\epsilon_r \epsilon_0 \mathbf{E}) = -\epsilon_0 \nabla \cdot (\epsilon_r \nabla V) = \rho_f, \quad (10-18)$$

V is proportional to ρ_f . Thus the potential is αV when the charge density is $\alpha\rho_f$.

Suppose that α increases to $\alpha + d\alpha$. Then, at a given point,

$$d\rho_f = \rho_f d\alpha, \quad dV = V d\alpha. \quad (10-19)$$

The energy required to pull in the extra charge is

$$d\mathcal{E} = \int_v (\rho_f d\alpha) V \left(\alpha + \frac{d\alpha}{2} \right) dv, \quad (10-20)$$

where the volume v includes all the charges and where $V(\alpha + d\alpha/2)$ is the average potential during the operation. Disregarding the $(d\alpha)^2$ term, we find that

$$d\mathcal{E} = \int_v (\rho_f V dv) \alpha d\alpha, \quad (10-21)$$

$$\mathcal{E} = \int_0^1 \alpha d\alpha \int_v \rho_f V dv = \frac{1}{2} \int_v \rho_f V dv. \quad (10-22)$$

The potential energy stored in the field is thus equal to one-half of the volume integral of $\rho_f V$, exactly as in Sec. 6.1.2. However, V now depends on the nature, shape, size, and position of the dielectrics situated in the field.

Example | THE DIELECTRIC-INSULATED PARALLEL-PLATE CAPACITOR

A sheet of dielectric separates the plates of a parallel-plate capacitor, one of which is grounded. Then the energy required to charge the other plate to the potential V is $QV/2$, where $Q = CV$.

Thus

$$\mathcal{E} = \frac{QV}{2} = \frac{CV^2}{2}, \quad (10-23)$$

as in the first example in Sec. 6.4

For a given charge Q , the presence of the dielectric decreases V by a factor of ϵ_r , and hence decreases \mathcal{E} by the same factor. However, C is proportional to ϵ_r , and, for a given potential V , the stored energy is proportional to ϵ_r .

10.3.2 The Energy Density Expressed in Terms of E and D

Our thought experiment here will be similar to that of the starred Sec. 6.2.1.

Assume that there is one and only one region where V is minimum, with $V_{\min} \leq 0$. Similarly, there is one and only one region where $V = V_{\max} \geq 0$. The dielectrics are linear.

Now imagine a conducting surface of area \mathcal{A} along the equipotential $V = V_{\min}$. There is zero field inside. The conductor expands slowly, from one equipotential to the next, sweeping through the dielectrics and picking up the free charges that stand in its way.

The conductor carries a free surface charge density σ . At the surface, E is normal and its magnitude is equal to σ/ϵ_0 , as in Fig. 10-7. If the medium is not isotropic, D is not parallel to E . However,

$$\sigma d\mathcal{A} = D \cdot d\mathcal{A}, \quad (10-24)$$

from Gauss's law, D being zero inside the conductor.

Say the element of area $d\mathcal{A}$ moves out by a distance dl as in the figure, dl being normal to the surface. Then the work done by the field *on* the charge $\sigma d\mathcal{A}$ is

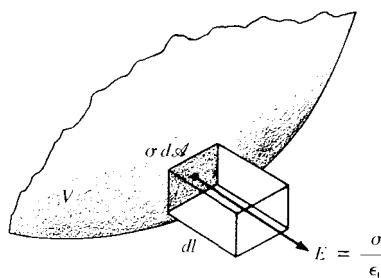


Fig. 10-7. Thought experiment for calculating the energy stored in an electric field in the presence of dielectrics. The shaded volume of area \mathcal{A} is a conductor that expands slowly, from one equipotential surface to the next. The surface charge density is σ . The element of area $d\mathcal{A}$ moves out by a distance dl .

$$d\mathcal{E} = (\sigma d\mathcal{A}) \left(\frac{E}{2} \right) dl, \quad (10-25)$$

the field acting on $\sigma d\mathcal{A}$ being $E/2$, from Sec. 6.5. So

$$d\mathcal{E} = \mathbf{D} \cdot d\mathcal{A} \frac{E}{2} dl = \frac{\mathbf{D} \cdot \mathbf{E}}{2} d\mathcal{A} dl = \frac{\mathbf{D} \cdot \mathbf{E}}{2} dv. \quad (10-26)$$

The conducting surface continues to expand until it reaches the equipotential $V = V_{\max}$. In the end, it has swept through all space, the free charges have been either removed to infinity or canceled by the addition of opposite charges on the conducting surface, and \mathbf{E} is zero everywhere.

Thus the total work done by the field is

$$\mathcal{E} = \int_v \frac{1}{2} \mathbf{E} \cdot \mathbf{D} dv, \quad (10-27)$$

and the energy density is

$$\mathcal{E}' = \frac{1}{2} \mathbf{E} \cdot \mathbf{D}. \quad (10-28)$$

If the dielectrics are linear and isotropic, $\mathbf{D} = \epsilon_r \epsilon_0 \mathbf{E}$ and

$$\mathcal{E} = \int_v \frac{\epsilon_r \epsilon_0 E^2}{2} dv. \quad (10-29)$$

The energy density is then $\epsilon_r \epsilon_0 E^2 / 2$. This integral extends over all the volume occupied by the field.

Example | THE DIELECTRIC-INSULATED PARALLEL-PLATE CAPACITOR

Returning to the capacitor of the example in Sec. 10.3.1, we find that, for a given E , its stored energy is proportional to ϵ_r , as previously. As usual, we have neglected edge effects.

10.3.3 The Energy Density Associated with Polarization

Since

$$\mathcal{E}' = \frac{1}{2} \mathbf{E} \cdot \mathbf{D} = \frac{1}{2} \mathbf{E} \cdot (\epsilon_0 \mathbf{E} + \mathbf{P}) = \frac{1}{2} \epsilon_0 E^2 + \frac{1}{2} \mathbf{E} \cdot \mathbf{P}, \quad (10-30)$$

the energy density associated with the polarization alone is $\mathbf{E} \cdot \mathbf{P}/2$.

In isotropic dielectrics,

$$\mathcal{E}' = \frac{1}{2}\epsilon_r\epsilon_0 E^2 = \frac{1}{2}(1 + \chi_e)\epsilon_0 E^2. \quad (10-31)$$

In most common dielectrics, the electric susceptibility χ_e ranges about from 1 to 4, and most of \mathcal{E} is polarization energy.

10.4 ELECTRIC FORCES IN THE PRESENCE OF DIELECTRICS

Electric forces between conductors in the presence of dielectrics are best calculated by the method of virtual work, as in Section 6.6.

When the conductors are immersed in a dielectric, the forces are smaller than those in air by a factor of ϵ_r , if the charges are the same. They are larger by the same factor if the electric fields, and hence the voltages, are the same.

The case of solid dielectrics is best illustrated by Probs. 10-13 to 10-15.

10.5 SUMMARY

Although some dielectrics are close to ideal, most exhibit a more complex behavior. If the conductivity σ of a dielectric is not negligible, then its *complex relative permittivity* is

$$\epsilon_r = \epsilon'_r - j\epsilon''_r, \quad \epsilon''_r = \frac{\sigma}{\epsilon_0\omega}. \quad (10-4)$$

Generally speaking, ϵ'_r decreases with frequency and ϵ''_r increases. Temperature effects can be large.

Some dielectrics are *nonisotropic*. Then each component of \mathbf{P} depends on the three components of \mathbf{E} , and the susceptibility is a tensor.

Ferroelectric dielectrics exhibit large permittivities and *hysteresis*: the value of \mathbf{D} depends on the previous values of \mathbf{E} .

At the *interface* between any two media, both \mathbf{V} and the tangential component of \mathbf{E} are continuous. The normal component of \mathbf{D} can be discontinuous, however, the discontinuity being equal to the free surface charge density on the interface.

Lines of \mathbf{E} bend at an interface, the larger angle from the normal being in the dielectric with the larger permittivity.

The *potential energy* stored in a charge distribution, when dielectrics are present, is given either by

$$\mathcal{E} = \frac{1}{2} \int_v \rho_f V dv, \quad (10-22)$$

where v is any volume containing all the free charges of the system, or by

$$\mathcal{E} = \int_v \frac{1}{2} \mathbf{E} \cdot \mathbf{D} dv, \quad (10-27)$$

where v now includes all the regions where the \mathbf{E} of the charge distribution exists. For linear and isotropic dielectrics,

$$\mathcal{E} = \int_v \frac{1}{2} \epsilon_r \epsilon_0 E^2 dv. \quad (10-29)$$

One may therefore ascribe to an electric field an *energy density* $\mathbf{E} \cdot \mathbf{D}/2$ or, for linear and isotropic dielectrics, $\epsilon_r \epsilon_0 E^2/2$. The energy density associated with the polarization itself in a dielectric is $\mathbf{E} \cdot \mathbf{P}/2$ or $\chi_c \epsilon_0 E^2/2$ if the dielectric is isotropic and linear.

PROBLEMS

10-1. (10.1.1) Parallel-plate capacitor with a conducting dielectric

Note: In this problem we call a surface charge density σ_{ch} and a conductivity σ_{co} .

The dielectric of a parallel-plate capacitor has a relative permittivity ϵ_r and a conductivity σ_{co} . The conductivity of the dielectric is much less than that of the plates, which makes \mathbf{E} uniform between the plates.

(a) Assume surface charge densities σ_{ch} on the plates, and use Gauss's law to relate E to σ_{ch} . Then show that the impedance of the capacitor is the same as that of a resistor and a capacitor in parallel.

(b) Show that when the capacitor is disconnected, the charge on the capacitor decreases by a factor of e in ϵ_r/σ_{co} seconds. This is the *relaxation time* of the capacitor.

10-2. (10.1.1) Parallel-plate capacitor with two conducting dielectrics

The dielectric of a parallel-plate capacitor is made up of two parts, as in Fig. 10-8.

(a) Find the impedance Z .

(b) Call a conductivity σ_{co} and a surface charge density σ_{ch} . Calculate the surface charge density on the interface.

10-3. (10.1.1) Parallel-plate capacitor with a nonuniform dielectric

A parallel-plate capacitor has plates of area \mathcal{A} separated by a distance s .

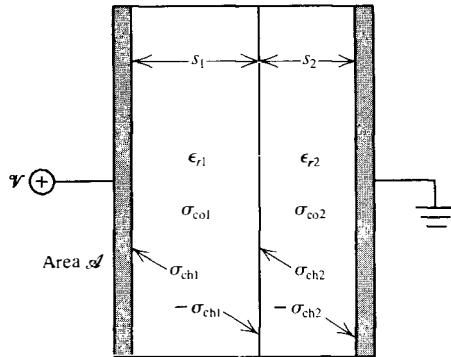


Fig. 10-8.

Its dielectric has a conductivity $\sigma = a + bx$, where x is the distance to one plate, and a uniform relative permittivity ϵ_r .

- Calculate the resistance R of the capacitor.
- Show that with a steady voltage V applied to the electrodes, there is a uniform volume density of free charge.
- Sketch lines of E for $b > 0$. The field is not uniform.
- With an alternating voltage across the electrodes,

$$I = \mathcal{A} \left(\sigma E + \frac{\partial D}{\partial t} \right) = \mathcal{A} J_t.$$

Show that $\nabla \cdot J_t = \partial J_t / \partial x = 0$. Then J_t is independent of x .

- Show that $E = J_t / (a + bx + j\omega\epsilon_r\epsilon_0)$.
- Now find the impedance $Z = V/I$. The real part of Z is not the R that we found above. However, if $\omega = 0$, we revert to R , as expected. See Prob. 10-1.
- Find ρ_f .

10-4. (10.1.4) Ceramic capacitor

One particular ceramic capacitor is cylindrical, with three electrodes as in Fig. 10-9. The ceramic disks each have a diameter of 21 millimeters and a thickness of 0.5 millimeter. The nominal capacitance is 0.05 microfarad within a range of -20% to $+80\%$. What is the approximate value of ϵ_r ?

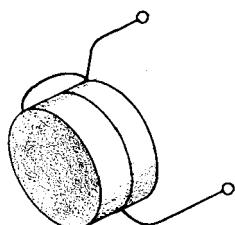
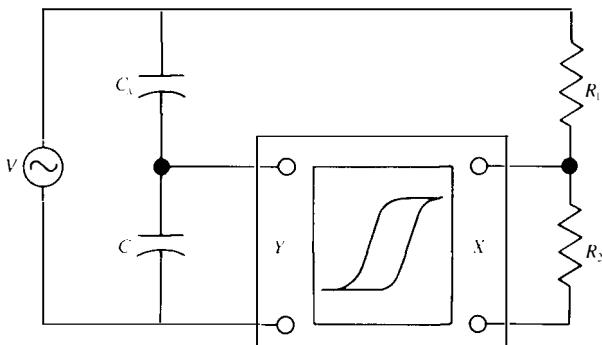


Fig. 10-9.

**Fig. 10-10.**

10.5. (10.1.4) The hysteresis curve of a ferroelectric material

Figure 10-10 shows a schematic diagram of an instrument that has been used to plot the hysteresis curves of ferroelectric materials. An oscillator applies an alternating voltage V to two capacitors in series, the parallel-plate capacitor C_x containing the material and a normal capacitor $C \gg C_x$. The voltage across C goes to the Y input of an XY recorder, while the voltage across $R_2 \ll R_1$ goes to the X input. The ferroelectric sample is a few millimeters thick, and $V \approx 10$ kilovolts, $f \approx 10^{-2}$ hertz.

Explain why the voltage across C is proportional to the D in the sample contained in C_x , while that across R_2 is proportional to E . The recorder draws essentially zero current at its X and Y terminals.

10.6. (10.1.5) The stored energy density in a nonlinear dielectric

A parallel-plate capacitor whose dielectric is nonlinear is connected to a power supply. The voltage V increases slightly by dV , and an extra charge dQ flows into the capacitor. Show that the density of stored energy increases by $E dD$.

10.7. (10.1.5) Hysteresis in ferroelectric materials

Show that the area of the hysteresis loop for a ferroelectric material is equal to the energy dissipated per cubic meter and per cycle.

10.8. (10.2) The boundary conditions at the interface between two conductors

Show that, at the interface between two conductors,

$$(\mathbf{E}_1 - \mathbf{E}_2) \times \hat{\mathbf{n}} = 0, \quad (\mathbf{J}_1 - \mathbf{J}_2) \times \hat{\mathbf{n}} = \frac{d\sigma_{ch}}{dt},$$

where $\hat{\mathbf{n}}$ is a unit vector that is normal to the interface and points away from conductor 1, and where σ_{ch} is the surface charge density on the interface.

10.9. (10.3.1) Capacitive energy storage for a small vehicle

Investigate the possibility of propelling a small vehicle with an electric motor fed by a charged capacitor. Consider only the problem of energy storage.

(a) Show that the maximum energy density in the dielectric of a parallel-plate capacitor is $\epsilon a^2/2$, where a is the *dielectric strength* of the insulator, or the maximum E before breakdown. A good dielectric to use

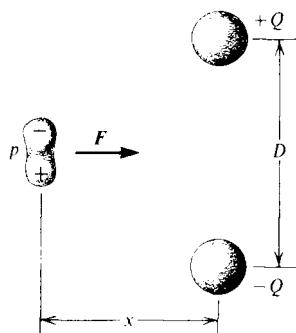


Fig. 10-11.

would be Mylar, which has a dielectric strength of 1.6×10^8 volts/meter when in the form of thin sheets and a relative permittivity of 3.2. The energy density would be about 10 times less with electrolytic capacitors.

(b) Calculate the energy density and the approximate size and mass of the capacitor that you would need to operate a 1-kilowatt motor for 1 hour.

10-10. (10.2.4) The bar electret

Figure 9-4 shows the E and D fields of a bar electret.

(a) Show that the lines of E do not bend at the cylindrical surface, but that the lines of D do bend.

(b) Show that the inverse is true at the end faces.

10-11. (10.3.2) The potential energy of a dipole situated in an electric field

(a) A dipole of *fixed* dipole moment p orients itself in a uniform electric field E . Show that its potential energy is $-pE$, assuming that the potential energy is zero when the dipole axis is perpendicular to the field.

(b) A nonpolar molecule acquires a dipole moment in a uniform electric field that gradually increases from zero to E . Show that its potential energy is $+pE/2$.

10-12. (10.3.2) An accelerator for neutral molecules

There exist accelerators for *neutral* molecules that operate as follows. Figure 10-11 shows a pair of spheres that carry charges $+Q$ and $-Q$, and a molecule of dipole moment p . The molecule accelerates toward the spheres until it reaches their midpoint. At that instant the spheres are discharged, and the molecule continues on its way at a constant velocity. Such accelerators serve to study the processes that occur during molecular collisions.

(a) Find the kinetic energy acquired by the molecule if the distance x in the figure is initially much larger than the distance D between the pair of electrodes. Consider the spheres as point charges, and apply the principle of conservation of energy. Assume that p is constant, and refer to Prob. 10-11. The dipole moment in fact increases as the molecule approaches, so that we have underestimated the energy.

(b) In one particular accelerator the electrode voltages are ± 40 kilovolts, their radius is 0.25 millimeter, and $D = 1.00$ millimeter. Calculate the approximate value of the kinetic energy of a molecule in electronvolts for

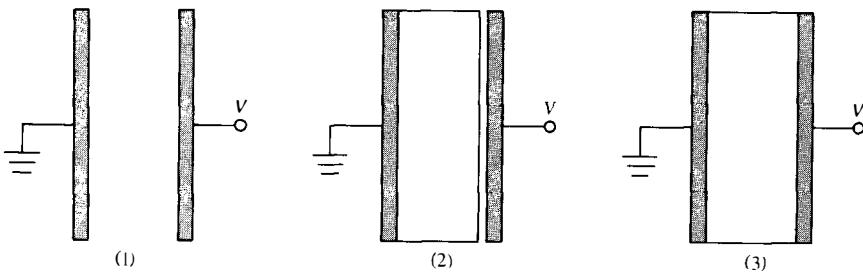


Fig. 10-12.

$p = 2 \times 10^{-29}$ coulomb-meter and for 700 stages. That accelerator has a length of 10 meters.

10-13. (10.4) Electrostatic clamps

Electrostatic clamps are used for holding workpieces while they are being machined, for holding silicon wafers during electron beam microfabrication, etc. They comprise an insulated conducting plate maintained at a potential of several thousand volts and covered with a thin insulating sheet. The workpiece or the wafer rests on the sheet and is grounded. It is advisable to apply a film of oil to the sheet to prevent sparking.

One particular type operates at 3000 volts and has holding power of 2 atmospheres (2×10^5 pascals). If the insulator is Mylar ($\epsilon_r = 3.2$), what is its thickness?

10-14. (10.4) Self-clamping capacitor

A certain capacitor consists of two polished circular aluminum plates, 237 millimeters in diameter, separated by a sheet of plastic 0.762 millimeter thick, with a relative permittivity of 3.0. Thin films of air subsist between the electrodes and the plastic. This reduces the capacitance below the rated value, and there is no way of clamping the plates mechanically with sufficient force.

(a) Someone suggests that the electric force alone might be sufficient to clamp the plates, at the operating voltage of 60 kilovolts. What is your opinion?

(b) Show that, if the complete capacitor is submerged in an oil with $\epsilon_r \approx 3$, the force is 3 times less. See the next problem.

10-15. (10.4) The clamping force, with and without an air film

In Prob. 10-14 we found that the electric clamping force on a capacitor is larger by a factor of ϵ_r , when there are air films between the electrodes and the dielectric, for a given applied voltage, or a given E . This is paradoxical. For a given E , the energy density is ϵ_r times larger in a dielectric than in air. Then the force should be ϵ_r times *larger* when there are no air films.

You can explain this paradox by considering the three capacitors of Fig. 10-12. In (2) the air film is much thinner than the dielectric. The electrode spacing is s in all three capacitors. For each case find the surface charge densities σ on the electrodes as well as D , E , and \mathcal{E} in the air and in the dielectric.

10-16. (10.4) Electric forces, with and without a dielectric

Electric forces on conductors immersed in liquid dielectrics are larger than in air by the factor ϵ , if the voltages are the same. They are smaller than in air by the same factor ϵ , if the charges are the same. Can you justify these general statements?

This is not in contradiction with Probs. 10-14 and 10-15, where we had a solid dielectric with thin films of air or oil next to the electrodes.

CHAPTER 11

ELECTRIC FIELDS IX

Images. Laplace's Equation in Rectangular Coordinates

11.1	THE UNIQUENESS THEOREM FOR ELECTROSTATIC FIELDS	211
*11.2	IMAGES	212
	<i>Example: POINT CHARGE NEAR A GROUNDED CONDUCTING PLATE</i>	212
	<i>Example: POINT CHARGE NEAR A DIELECTRIC</i>	213
*11.3	SOLVING LAPLACE'S EQUATION IN RECTANGULAR COORDINATES	215
	<i>Example: THE FIELD BETWEEN TWO GROUNDED SEMI-INFINITE PARALLEL ELECTRODES TERMINATED BY A PLANE ELECTRODE MAINTAINED AT A FIXED POTENTIAL; FOURIER SERIES</i>	218
11.4	SUMMARY	221
	PROBLEMS	223

Up to this point we have limited our discussion of electric fields to general considerations and to simple charge distributions. We now address the calculation of more complex fields, both in this chapter and in the next one. That will complete our study of purely electric fields.

11.1 THE UNIQUENESS THEOREM FOR ELECTROSTATIC FIELDS

According to the *uniqueness theorem* for electrostatic fields, there cannot exist more than one potential $V(x, y, z)$ that satisfies both Poisson's equation and a given set of boundary conditions. This theorem is important because it leaves us free to use any method, even intuition, to find V . If we can, somehow, discover a function $V(x, y, z)$ that meets these two requirements, then it is the only possible potential function.

Several proofs of this theorem are known, each one based on its own particular set of assumptions. We simply assume that the uniqueness theorem always applies.

*11.2 IMAGES

If an electric charge distribution lies in a uniform dielectric that is in contact with a conducting body, then the method of images often provides the simplest route for calculating the electric field. The method is best explained by examples such as the two given below, but the principle is the following.

Call the charge distribution Q , the dielectric D , and the conductor C . One replaces C , on paper, by more dielectric D' and by a second charge distribution Q' such that the original boundary conditions are not disturbed. Then the field in D is left undisturbed, according to the uniqueness theorem. The charge distribution Q' is said to be the *image* of Q . Of course, the dielectric D can be simply air or a vacuum.

If the charge distribution Q lies in a uniform dielectric D_1 that is in contact with a second dielectric D_2 , then we can calculate the fields in both D_1 and D_2 in a similar fashion, as we shall see in the second example below.

Example

POINT CHARGE NEAR A GROUNDED CONDUCTING PLATE

Figure 11-1(a) shows a point charge Q situated at a distance D from a large grounded conducting plate.[†]

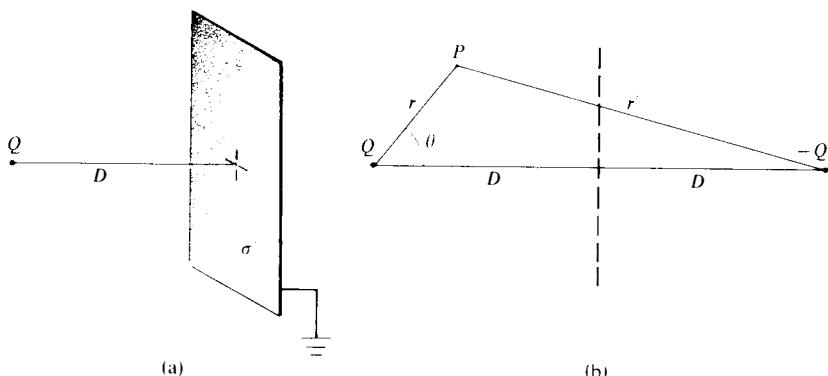


Fig. 11-1. (a) Point charge Q near a large, grounded, conducting sheet. (b) We have replaced the conducting sheet by the image charge $-Q$. The field to the left of the dashed line is unaffected.

[†] The earth carries a negative charge, and near the surface the electric field strength is about 100 volts/meter. However, the field of Q is independent of the field of the earth, according to the principle of superposition.

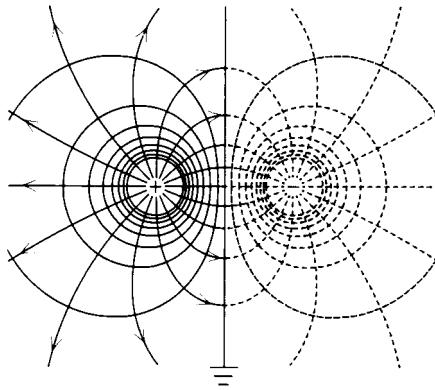


Fig. 11-2. Lines of \mathbf{E} (arrows) and equipotentials for the field of Fig. 11-1(a). Equipotentials near the charge get too close together to be shown. Rotating the figure about the horizontal axis of symmetry generates equipotential surfaces.

Clearly, if we remove the plate and add an image charge $-Q$ at a further distance D as in Fig. 11-1(b), then every point in the plane formerly occupied by the plate will be equidistant from the two charges and will thus be at zero potential. So the field in the region to the left of the position formerly occupied by the plate is the same in both figures.

This is remarkable. The image provides a trivial solution to a problem that would, otherwise, be rather difficult.

Figure 11-2 shows the lines of \mathbf{E} and the equipotentials.

Example

POINT CHARGE NEAR A DIELECTRIC

Our point charge Q now lies in air close to a large block of dielectric as in Fig. 11-3. We wish to know \mathbf{E} on both sides of the interface.

We could calculate \mathbf{E} without using images in the following way. The field of Q polarizes the dielectric, and a bound surface charge density σ_b appears on the interface. Then the \mathbf{E} at any point on either side of the interface is the same as if one had the point charge Q and the sheet of bound charges of density σ_b situated in a vacuum. As we shall see, it is simple enough to find σ_b . But then calculating \mathbf{E} would be rather awkward. The problem is much easier to solve with images.

We first calculate the value of the normal component of \mathbf{E} next to the interface and on both sides of it. It is this normal component, E_n , that we shall use as a boundary condition. In the process, we shall find σ_b .

The part of E_n that is ascribable to Q is $[Q/(4\pi\epsilon_0 r^2)](D/r)$. This field is the same on the two sides of the interface, as in Fig. 11-3.

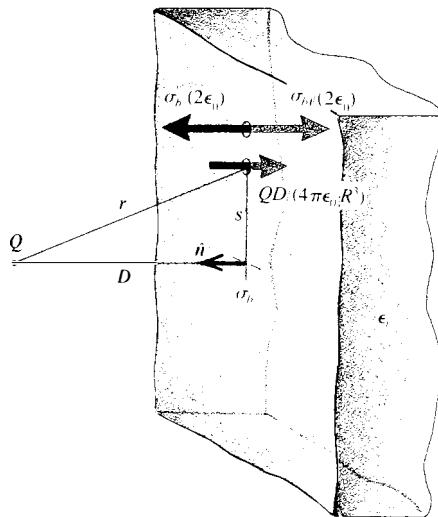


Fig. 11-3. Point charge Q near a large block of dielectric. The arrows for the fields are oriented on the assumption that both Q and σ_b are positive. If Q is positive, σ_b is, in fact, negative.

From Gauss's law, σ_b provides equal normal field strengths $\sigma_b/2\epsilon_0$ on either side of the interface, oriented as in the figure. We shall find that the sign of σ_b is opposite that of Q .

Therefore, just inside the dielectric, the normal component of \mathbf{E} pointing to the right is

$$E_{ni} = \frac{QD}{4\pi\epsilon_0 r^3} + \frac{\sigma_b}{2\epsilon_0}. \quad (11-1)$$

Now, from Eqs. 9-4 and 9-28,

$$\sigma_b = \mathbf{P} \cdot \hat{\mathbf{n}} = P_n = -\epsilon_0(\epsilon_r - 1)E_{ni}, \quad (11-2)$$

where the unit vector $\hat{\mathbf{n}}$ is normal to the interface and points left. Therefore

$$\sigma_b = -\frac{(\epsilon_r - 1)QD}{2\pi(\epsilon_r + 1)r^3}. \quad (11-3)$$

As expected, Q and σ_b have opposite signs.

If we take the right-hand direction as positive, the normal components of \mathbf{E} outside and inside the dielectric, respectively, are as follows:

$$E_{no} = \left(1 + \frac{\epsilon_r - 1}{\epsilon_r + 1}\right) \frac{QD}{4\pi\epsilon_0 r^3} \quad (11-4)$$

$$= \frac{2\epsilon_r}{\epsilon_r + 1} \frac{QD}{4\pi\epsilon_0 r^3}, \quad (11-5)$$

$$E_{ni} = \left(1 - \frac{\epsilon_r - 1}{\epsilon_r + 1}\right) \frac{QD}{4\pi\epsilon_0 r^3} \quad (11-6)$$

$$= \frac{2}{\epsilon_r + 1} \frac{QD}{4\pi\epsilon_0 r^3}. \quad (11-7)$$

Since $\epsilon_r E_{ni} = E_{no}$, the normal component of \mathbf{D} is continuous across the interface. This was to be expected because the free surface charge density is zero.

Thanks to the uniqueness theorem and to images, our problem is nearly solved.

The Field on the Left of the Interface

(a) The boundary condition is stated in Eqs. 11-4 and 11-5. From Eq. 11-4, we can remove the dielectric and replace it by an image charge Q' , as in Fig. 11-4(a), without disturbing E_{no} . Then, from the uniqueness theorem, the field everywhere on the left of the interface is simply the field of Q plus that of Q' .

(b) Now, according to Eq. 11-5, E_{no} is also unaffected if one removes the dielectric and replaces Q by a charge

$$Q'' = \frac{2\epsilon_r}{\epsilon_r + 1} Q. \quad (11-8)$$

But then we have modified the charge distribution, and hence the field, in the region we are interested in. So the correct equivalent field is that of Fig. 11-4(a).

The Field on the Right of the Interface

(a) From Eq. 11-6, E_{ni} is the same as if one had, instead of Q and the dielectric, the charge Q and a charge $-Q'$ at a distance D to the right of the interface. This is of no interest because this changes the field on the right.

(b) From Eq. 11-7, the field on the right of the interface is the same as if there were no dielectric and a charge Q'' were substituted for Q , as in Fig. 11-4(b).

(c) Figure 11-4(c) shows another possibility: the dielectric extends on both sides of the interface, and Q''' replaces Q .

Figure 11-5 shows lines of \mathbf{D} and equipotentials for this field.

*11.3 SOLVING LAPLACE'S EQUATION IN RECTANGULAR COORDINATES

Solutions of Laplace's equation $\nabla^2 V = 0$ are known as *harmonic functions*. These functions possess a number of general properties, of which we shall use the following one. If one can find solutions V_1, V_2, V_3, \dots of

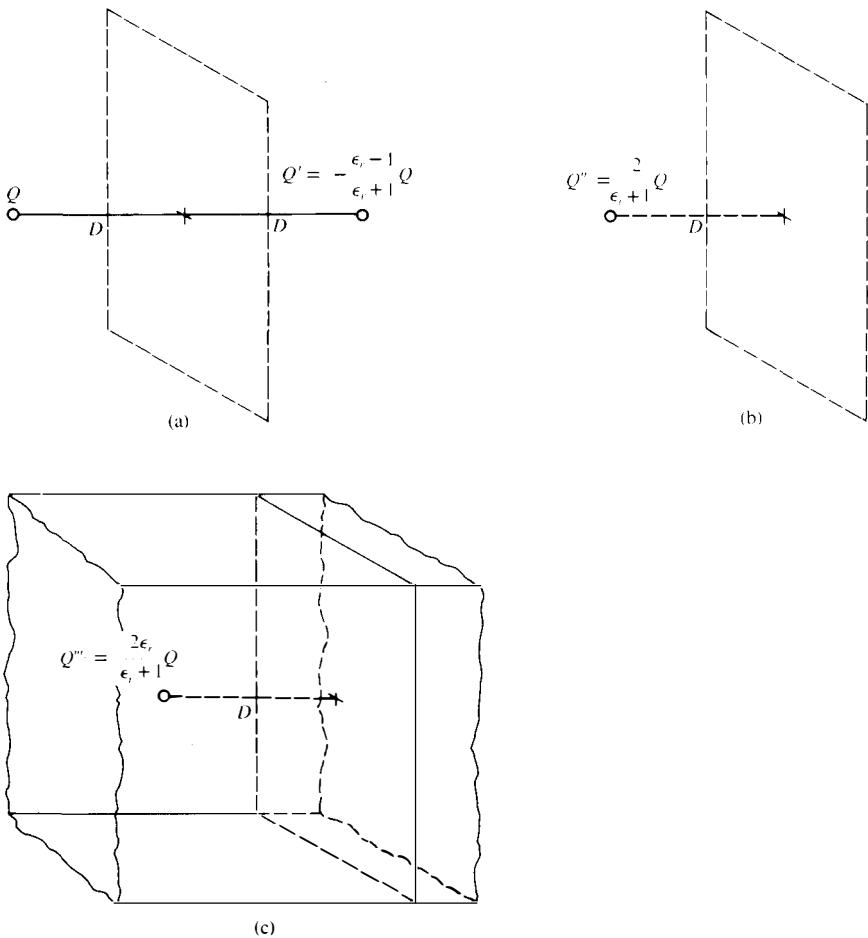


Fig. 11-4. (a) The field on the *left* of the interface of Fig. 11-3 is the same as that of this pair of charges. (b), (c) The field on the *right* of the interface is the same as that of either of these two arrangements.

Laplace's equation, then any linear combination $A_1 V_1 + A_2 V_2 + A_3 V_3 + \dots$, where the A 's are constants, is also a solution. This becomes obvious upon substituting the sum into the original equation.

As a rule, one can solve Laplace's equation by *separating the variables*. For example, in Cartesian coordinates, one can seek solutions of the form

$$V = X(x)Y(y)Z(z), \quad (11-9)$$

where X is a function of x only, Y a function of y only, and Z a function

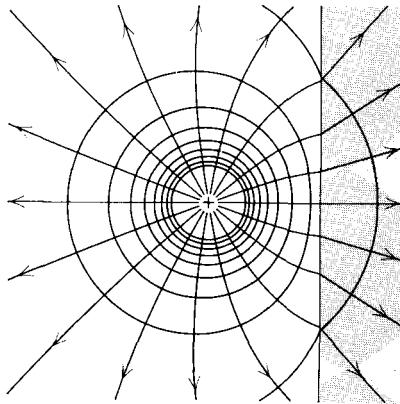


Fig. 11-5. Lines of \mathbf{D} (arrows) and equipotentials for a point charge situated near a dielectric. Rotating the figure about the horizontal axis of symmetry generates equipotential surfaces. The figure does not show equipotentials near the point charge.

of z only. Substituting into Laplace's equation yields

$$YZ \frac{d^2X}{dx^2} + ZX \frac{d^2Y}{dy^2} + XY \frac{d^2Z}{dz^2} = 0. \quad (11-10)$$

Dividing by XYZ gives

$$\frac{1}{X} \frac{d^2X}{dx^2} + \frac{1}{Y} \frac{d^2Y}{dy^2} + \frac{1}{Z} \frac{d^2Z}{dz^2} = 0, \quad (11-11)$$

where the first term depends solely on x , and similarly for the other two terms.

Since these three terms add to zero at any point in the field, each one is equal to a constant and

$$\frac{d^2X}{dx^2} = C_1 X, \quad \frac{d^2Y}{dy^2} = C_2 Y, \quad \frac{d^2Z}{dz^2} = C_3 Z, \quad (11-12)$$

with

$$C_1 + C_2 + C_3 = 0. \quad (11-13)$$

We have now separated the variables. Solving the three equations separately yields X , Y , Z , and thus V .

Observe how astute this is: we have transformed a *partial* differential equation in all three variables x, y, z into three simple *ordinary* differential equations in x , and y , and z .

Sometimes, as in the example below, one has to sum an infinite number of such solutions, each multiplied by a suitable coefficient, to fit a given boundary condition.

*Example

THE FIELD BETWEEN TWO GROUNDED SEMI-INFINITE PARALLEL ELECTRODES TERMINATED BY A PLANE ELECTRODE MAINTAINED AT A FIXED POTENTIAL; FOURIER SERIES

Figure 11-6 shows the electrodes. We wish to find $V(x, y)$ between the plates. By hypothesis, the field is independent of z and $C_3 = 0$. Then

$$\frac{d^2X}{dx^2} = k^2 X, \quad \frac{d^2Y}{dy^2} = -k^2 Y. \quad (11-14)$$

We have substituted k^2 for C_1 and $-k^2$ for C_2 to avoid square roots in the solution. As we shall see immediately, C_2 must be negative.

We solve the Y equation by setting

$$Y = A \sin ky + B \cos ky, \quad (11-15)$$

where A and B are constants. This solution is easily verified by substitution.

Now V must satisfy the following boundary conditions: (1) $V = 0$ at $x = 0, y = b$; (2) $V = V_0$ at $x = 0$; (3) $V \rightarrow 0$ as $x \rightarrow \infty$.

Since $V = 0$ at $y = 0$, then $B = 0$.

Also, the condition $V = 0$ at $y = b$ requires that

$$kb = n\pi, \quad k = \frac{n\pi}{b} \quad (n = 1, 2, 3, \dots). \quad (11-16)$$

So

$$Y = A \sin \frac{n\pi y}{b} \quad (n = 1, 2, 3, \dots). \quad (11-17)$$

We ignore the value $n = 0$ because it corresponds to zero field.

The differential equation for X is now

$$\frac{d^2X}{dx^2} = \left(\frac{n\pi}{b}\right)^2 X, \quad (11-18)$$

and

$$X = G \exp \frac{n\pi x}{b} + H \exp \left(-\frac{n\pi x}{b}\right), \quad (11-19)$$

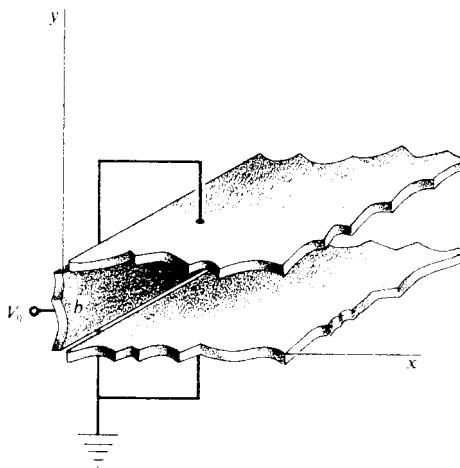


Fig. 11-6. Grounded, plane, parallel, and semi-infinite electrodes terminated by a plane electrode at the potential V_0 .

where G and H are constants of integration. We can again verify this solution by substitution. Since V tends to zero as x tends to infinity, $G = 0$. Thus

$$V = XY = C \sin \frac{n\pi y}{b} \exp \left(-\frac{n\pi x}{b} \right), \quad (11-20)$$

where C is another constant.

But this is not right! Although this V satisfies conditions (1) and (3), it clearly does not satisfy condition (2). We can also satisfy condition (2) by adding an infinite number of such solutions:

$$V = \sum_{n=1}^{\infty} C_n \sin \frac{n\pi y}{b} \exp \left(-\frac{n\pi x}{b} \right). \quad (11-21)$$

Then we use condition (2) to evaluate the coefficients C_n by setting $x = 0$. So

$$V_0 = \sum_{n=1}^{\infty} C_n \sin \frac{n\pi y}{b} \quad (11-22)$$

for all y between 0 and b .

An infinite series of the form

$$\sum_{n=0}^{\infty} \left(C_n \sin \frac{n\pi y}{b} + D_n \cos \frac{n\pi y}{b} \right),$$

where C_n and D_n are constants, is a *Fourier series* and forms a

complete set: given a reasonably well-behaved[†] function $V(y)$ defined in the interval $y = 0$ to $y = b$, there exists a Fourier series that is equal to $V(y)$ in this interval.[‡]

We can find the values of the C_n coefficients by an ingenious technique devised by Fourier. First, we multiply both sides of Eq. 11-22 by $\sin(p\pi y/b)$, where p is an integer, and then we integrate from $y = 0$ to $y = b$:

$$\int_0^b V_0 \sin \frac{p\pi y}{b} dy = \sum_{n=1}^{\infty} \int_0^b C_n \sin \frac{n\pi y}{b} \sin \frac{p\pi y}{b} dy. \quad (11-23)$$

On the left-hand side,

$$\int_0^b V_0 \sin \frac{p\pi y}{b} dy = \begin{cases} \frac{2bV_0}{p\pi} & \text{if } p \text{ is odd} \\ 0 & \text{if } p \text{ is even.} \end{cases} \quad (11-24)$$

The terms of a Fourier series are thus said to be *orthogonal*. On the right,

$$\int_0^b C_n \sin \frac{n\pi y}{b} \sin \frac{p\pi y}{b} dy = \begin{cases} 0 & \text{if } p \neq n \\ \frac{C_n b}{2} & \text{if } p = n. \end{cases} \quad (11-25)$$

It follows that

$$C_n = \begin{cases} \frac{4V_0}{n\pi} & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even.} \end{cases} \quad (11-26)$$

Finally,

$$V = \frac{4V_0}{\pi} \sum_{1,3,5,\dots}^{\infty} \frac{1}{n} \sin \frac{n\pi y}{b} \exp \left(-\frac{n\pi x}{b} \right). \quad (11-27)$$

The successive terms in the series become progressively less important, because of the factor $1/n$, but mostly because of the exponential function. Figure 11-7 shows the degree of approximation achieved with 1, 3, 10, and 100 terms of the series.

Note how fast the exponential function decreases with x . For $n = 1$ and $x = b$ it is down to $\exp(-\pi)$, or 4%. Roughly speaking, the field of the charged plate does not penetrate in the x direction beyond a distance equal to the spacing b between the plates.

Notice also how fast the exponential decreases with n . For $n = 1$ and $x = b$ it is 0.04, but for $n = 2$ and $x = b$ it is 0.002. This means

[†] By a "reasonably well-behaved" function, we mean one that is continuous, or at least piecewise continuous, and that does not become infinite at any point. Most boundary conditions encountered in practice possess these characteristics.

[‡] Except possibly at the endpoints $y = 0$ and $y = b$, but this discrepancy between V and the Fourier series is usually unimportant in practice.

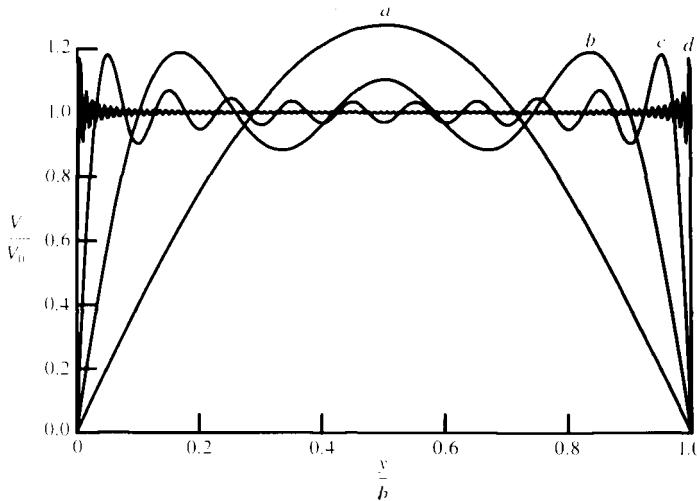


Fig. 11-7. The condition $V = V_0$ at $x = 0$, as satisfied by a Fourier series by taking (a) only the first term, (b) the first 3 terms, (c) the first 10 terms, and (d) the first 100 terms. The fit improves as the number of terms increases, but the spikes at $y = 0$ and $y = b$, where V is discontinuous, remain. This is the *Gibbs phenomenon*.

that $V(y/b)$ rapidly becomes a sine curve, with increasing x .

Figure 11-8 shows the equipotentials.

11.4 SUMMARY

The *uniqueness theorem* states that, for a given set of boundary conditions, there is only one possible field. This theorem is of enormous practical importance. If, somehow, one can find a $V(x, y, z)$ that satisfies both Laplace's (or Poisson's) equation and the given boundary conditions, then it is the correct potential function.

The *method of images* can often simplify the calculation of electric fields that involve interfaces between different media. It consists in setting up, on paper, a different field on the other side of the interface, with fictitious, or image, charges.

Solutions of Laplace's equation $\nabla^2 V = 0$ are termed *harmonic functions*. It is often possible to reduce Laplace's equation to a set of three independent, ordinary differential equations, one for each coordinate. Then V is of the form

$$V(x, y, z) = X(x)Y(y)Z(z), \quad (11-9)$$

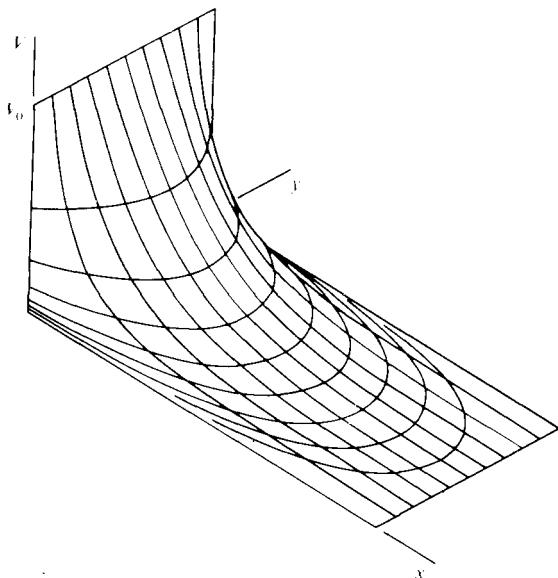


Fig. 11-8. Three-dimensional plot of the potential V for the field of Fig. 11-5. The U-shaped curves are equipotentials; the others show the intersections of the potential surface with planes parallel to the xV -plane.

where X is a function of x alone, etc. This operation is called *separating the variables*.

For certain fields one can fit the boundary conditions only by summing an infinite number of harmonic functions. We used a *Fourier series*

$$\sum_{n=1}^{\infty} \left(C_n \sin \frac{n\pi y}{b} + D_n \cos \frac{n\pi y}{b} \right),$$

where C_n and D_n are constants.

PROBLEMS

11-1. (11.1) The uniqueness theorem

According to this theorem, the Poisson equation $\nabla^2 V = -\rho_i/\epsilon_0$ can have only one solution if the potential V is defined at the boundaries of the field.

Show that two solutions can differ at most by a constant if the normal component of ∇V is defined at the boundaries.

11-2. (11.2) Point charge above a conducting plate

A point charge Q lies at a distance D above a grounded conducting plate.

(a) Calculate the surface charge density induced on the plate as a function of the radius r from the foot of the perpendicular drawn from the charge.

(b) Show that the total induced charge is $-Q$.

11-3. (11.2) Line charge near a conducting plate

A line charge of λ coulombs/meter is parallel to a flat conducting plate, at a distance a , as in Fig. 11-9. Find \mathbf{E} at a point (x, y) . The surface charge density on the plate is given by $\epsilon_0 E$ at $(x, 0)$.

11-4. (11.2) The field of a charge inside a hollow conducting sphere

A hollow conducting sphere of radius a contains a point charge Q at the radius b as in Fig. 11-10.

(a) Show that the field inside the sphere is the same as if there was no sphere and, instead, a charge $Q' = -(a/b)Q$ at $D = (a/b)a$. You can prove this by showing that the V of Q plus Q' is uniform over the surface of the sphere.

(b) Calculate the force of attraction.

(c) Calculate the surface charge density on the inside surface of the conducting sphere.

11-5. (11.2) The field of a point charge near a block of dielectric

Show that the potential at the surface of the dielectric, opposite Q in Fig. 11-3, is the same, whether one calculates the field in air, as in Fig. 11-3(a), or in the dielectric, as in Fig. 11-3(c). Set $\epsilon_r = 3$.

11-6. (11.3) Solutions of Laplace's equation can be of the form $X(x) + Y(y) + Z(z)$

Show that there exist solutions of Laplace's equation that are of the form $X(x) + Y(y) + Z(z)$.

11-7. (11.3) The function $1/r$ and its derivatives are solutions of Laplace's equation

(a) Show that $\nabla^2(1/r) = 0$.

(b) Use this fact to show that

$$\frac{\partial}{\partial x} \frac{1}{r}, \quad \frac{\partial^2}{\partial x^2} \frac{1}{r}, \quad \frac{\partial^2}{\partial x \partial y} \frac{1}{r}$$

are also solutions of Laplace's equation.

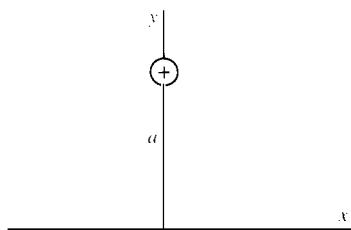
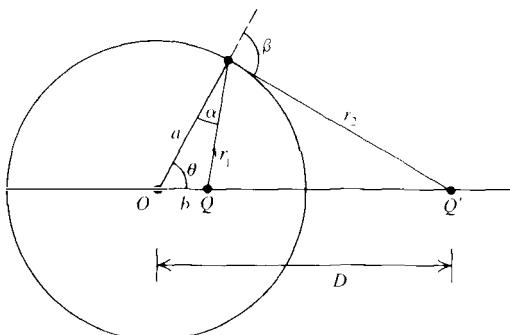


Fig. 11-9.

**Fig. 11-10.**

11-8. (11.3) The field of Fig. 11-6

Plot $\exp(-n\pi x/b)$ as a function of x/b from $x/b = 0$ to 1 and for $n = 1, 2$, and 3.

11-9. (11.3) The field of Fig. 11-6

Plot V/V_0 as a function of y/b for the field of Fig. 11-8 for $x = 0$, $x = 0.5b$, and $x = b$, up to $n = 100$. You can truncate the series when $\exp(-n\pi x/b)$ is less than 0.01.

CHAPTER 12

*ELECTRIC FIELDS X *Laplace's Equation in Spherical Coordinates.* *Poisson's Equation for E*

12.1	SOLVING LAPLACE'S EQUATION IN SPHERICAL COORDINATES.	
	LEGENDRE'S EQUATION	225
	<i>Example: UNCHARGED CONDUCTING SPHERE IN A PREVIOUSLY UNIFORM ELECTRIC FIELD</i>	228
	<i>Example: DIELECTRIC SPHERE IN A PREVIOUSLY UNIFORM ELECTRIC FIELD</i>	231
12.2	POISSON'S EQUATION FOR E IN ELECTROSTATIC FIELDS	233
	<i>Example: THE FIELD OF A SPHERE OF CHARGE WHOSE DENSITY IS A FUNCTION OF THE RADIUS</i>	234
12.3	SUMMARY	234
	PROBLEMS	235

In this last chapter on electric fields we shall see how to solve Laplace's equation in spherical coordinates. This will require the Legendre polynomials that we discussed in the starred Sec. 5.4.1. Then we shall deduce and discuss briefly Poisson's equation for E , which leads to an integral for E that is strangely different from the one that follows from Coulomb's law.

12.1 SOLVING LAPLACE'S EQUATION IN SPHERICAL COORDINATES. LEGENDRE'S EQUATION

Some electric fields are best treated in spherical polar coordinates. Solutions of Laplace's equation expressed in these coordinates are known as *spherical harmonic functions*.

We limit ourselves to fields possessing axial symmetry and therefore to fields that are independent of the azimuthal angle ϕ . Then Laplace's equation takes the following form:

$$\frac{\partial}{\partial r} \left(r^2 \frac{\partial V}{\partial r} \right) + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial V}{\partial \theta} \right) = 0. \quad (12-1)$$

To separate the variables as in Chap. 11, we set

$$V = R(r)\Theta(\theta), \quad (12-2)$$

where R is a function of r only and Θ a function of θ only. Then, by substitution,

$$\Theta \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) + \frac{R}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Theta}{\partial \theta} \right) = 0. \quad (12-3)$$

Dividing by $R\Theta$, we get

$$\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{1}{\Theta \sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) = 0. \quad (12-4)$$

We have now written total instead of partial derivatives because R and Θ are each functions of a single variable.

The second term is independent of r . Then the first term is also independent of r and is equal to a constant:

$$\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) = k. \quad (12-5)$$

Then

$$\frac{1}{\Theta \sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) = -k, \quad (12-6)$$

since the sum of the two constants must equal zero.

Let us examine the R equation first. Multiplying both sides by R and differentiating the term enclosed in parentheses, we have that

$$r^2 \frac{d^2 R}{dr^2} + 2r \frac{dR}{dr} - kR = 0. \quad (12-7)$$

The solution of this equation is of the form

$$R = A'r^n + B'r^{-(n+1)}, \quad (12-8)$$

with

$$n(n+1) = k, \quad (12-9)$$

as you can check by substitution.

Rewriting the Θ equation yields

$$\frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) + n(n+1)\Theta \sin \theta = 0. \quad (12-10)$$

Now set

$$\mu = \cos \theta, \quad (12-11)$$

remembering that, for any function $f(\mu)$,

$$\frac{df}{d\theta} = \frac{df}{d\mu} \frac{d\mu}{d\theta} = -\sin \theta \frac{df}{d\mu} = -(1 - \mu^2)^{1/2} \frac{df}{d\mu}. \quad (12-12)$$

Then the Θ equation becomes *Legendre's equation*:

$$\frac{d}{d\mu} \left[(1 - \mu^2) \frac{d\Theta}{d\mu} \right] + n(n+1)\Theta = 0. \quad (12-13)$$

When n is an integer, its solutions are the Legendre polynomials of Sec. 5.4.1:

$$\Theta = P_n(\cos \theta). \quad (12-14)$$

We shall use the following property of the Legendre equation. Since

$$n(n+1) = n'(n'+1), \quad \text{if } n' = -(n+1), \quad (12-15)$$

then

$$P_{-(n+1)}(\cos \theta) = P_n(\cos \theta). \quad (12-16)$$

It follows that, for every solution of Laplace's equation of the form

$$V = Ar^n P_n(\cos \theta), \quad (12-17)$$

there exists another one of the form

$$V = Br^{-(n+1)} P_{-(n+1)}(\cos \theta) = Br^{-(n+1)} P_n(\cos \theta). \quad (12-18)$$

Observe that this result is in agreement with Eq. 12-8.

So the general solution of Laplace's equation for fields possessing axial symmetry is

$$V = \sum_{n=0}^{\infty} (A_n r^n + B_n r^{-(n+1)}) P_n(\cos \theta). \quad (12-19)$$

Table 12-1 Solutions of Laplace's equation in spherical polar coordinates for fields possessing axial symmetry

n	$r^n P_n(\cos \theta)$	$r^{-(n+1)} P_n \cos \theta$
0	1	r^{-1}
1	$r \cos \theta$	$r^{-2} \cos \theta$
2	$r^2 \frac{(3 \cos^2 \theta - 1)}{2}$	$r^{-3} \frac{(3 \cos^2 \theta - 1)}{2}$
3	$r^3 \frac{(5 \cos^3 \theta - 3 \cos \theta)}{2}$	$r^{-4} \frac{(5 \cos^3 \theta - 3 \cos \theta)}{2}$
4	$r^4 \frac{(35 \cos^4 \theta - 30 \cos^2 \theta + 3)}{8}$	$r^{-5} \frac{(35 \cos^4 \theta - 30 \cos^2 \theta + 3)}{8}$
5	$r^5 \frac{(63 \cos^5 \theta - 70 \cos^3 \theta + 15 \cos \theta)}{8}$	$r^{-6} \frac{(63 \cos^5 \theta - 70 \cos^3 \theta + 15 \cos \theta)}{8}$

Table 12-1 shows the first six terms.

This series is analogous to a Fourier series (example, Sec. 11.3) in the following ways. First, the expressions under the summation sign form a complete set: the series can satisfy any reasonably well-behaved boundary condition exhibiting axial symmetry. Second,

$$\int_{-1}^{+1} P_m(\cos \theta) P_n(\cos \theta) d(\cos \theta) = \begin{cases} 0 & \text{if } m \neq n \\ \frac{2}{2n+1} & \text{if } m = n. \end{cases} \quad (12-20)$$

Legendre polynomials are thus orthogonal. Third, we can use this orthogonality to calculate the values of the A_n and B_n coefficients.

Example

UNCHARGED CONDUCTING SPHERE IN A PREVIOUSLY UNIFORM ELECTRIC FIELD

An insulated and uncharged conducting sphere is situated in a previously uniform electric field \mathbf{E}_0 . This applied field originates in remotely situated charge distributions that are unaffected by the presence of the sphere.

At any point in space, either inside or outside the sphere, $\mathbf{E} = \mathbf{E}_0 + \mathbf{E}_i$, where \mathbf{E}_i is the field of the charges induced on the sphere. The induced charges arrange themselves so as to render the net field inside equal to zero.

We calculate the field outside the sphere by solving Laplace's equation in two different ways.

(a) We first use spherical polar coordinates, with the origin at the center of the sphere and the polar axis along \mathbf{E}_0 . The boundary conditions are then as follows:

$$(1) \quad V_{r=a} = 0, \quad (2) \quad V_{r \rightarrow \infty} = -E_0 z = -E_0 r \cos \theta. \quad (12-21)$$

From boundary condition 1 and from Eq. 12-19,

$$0 = \sum_{n=0}^{\infty} A_n a^n P_n(\cos \theta) + \sum_{n=0}^{\infty} B_n a^{-(n+1)} P_n(\cos \theta). \quad (12-22)$$

We now evaluate the A and B coefficients in a manner analogous to that of the example in Sec. 11.3. We multiply both sides of the equation by $P_m(\cos \theta)$ and integrate from $\cos \theta = -1$ to $\cos \theta = +1$:

$$\begin{aligned} 0 &= \sum_{n=0}^{\infty} \int_{-1}^{+1} A_n a^n P_n(\cos \theta) P_m(\cos \theta) d(\cos \theta) \\ &\quad + \sum_{n=0}^{\infty} \int_{-1}^{+1} B_n a^{-(n+1)} P_n(\cos \theta) P_m(\cos \theta) d(\cos \theta). \end{aligned} \quad (12-23)$$

According to Eq. 12-20, the only nonvanishing terms are those for which $m = n$. Then each summation reduces to a single term:

$$0 = A_n a^n \int_{-1}^{+1} P_n^2(\cos \theta) d(\cos \theta) + B_n a^{-(n+1)} \int_{-1}^{+1} P_n^2(\cos \theta) d(\cos \theta) \quad (12-24)$$

$$= \frac{A_n a^n + B_n a^{-(n+1)}}{n + \frac{1}{2}}. \quad (12-25)$$

So

$$B_n = -A_n a^{2n+1}. \quad (12-26)$$

Substituting into Eq. 12-19 yields

$$V = \sum_{n=0}^{\infty} A_n (r^n - a^{2n+1} r^{-(n+1)}) P_n(\cos \theta). \quad (12-27)$$

Now boundary condition (2) concerns the value of V at infinity where all inverse powers of r tend to zero. Thus, at $r \rightarrow \infty$,

$$V = -E_0 r \cos \theta = -E_0 r P_1(\cos \theta) = \sum_{n=0}^{\infty} A_n r^n P_n(\cos \theta) \quad (12-28)$$

for all θ . By inspection, the only nonzero term on the right is that for which $n = 1$. See Table 12-1. Then

$$A_1 = -E_0 \quad (12-29)$$

and all the other A 's are zero. Also, from Eq. 12-26, all the B 's are zero except B_1 :

$$B_1 = -A_1 a^3 = E_0 a^3. \quad (12-30)$$

Finally, at any point outside the sphere,

$$V = -E_0 r \cos \theta + \frac{E_0 a^3 \cos \theta}{r^2} = -\left(1 - \frac{a^3}{r^3}\right) E_0 r \cos \theta, \quad (12-31)$$

$$E_r = -\frac{\partial V}{\partial r} = \left(1 + \frac{2a^3}{r^3}\right) E_0 \cos \theta, \quad (12-32)$$

$$E_\theta = -\frac{1}{r} \frac{\partial V}{\partial \theta} = -\left(1 - \frac{a^3}{r^3}\right) E_0 \sin \theta. \quad (12-33)$$

The surface charge density on the conducting sphere is

$$\sigma = \epsilon_0 E_r (r=a) = 3\epsilon_0 E_0 \cos \theta. \quad (12-34)$$

In the above expression for V , observe that the first term comes from E_0 , while the second is the V of a point dipole (Sec. 5.1) situated at the origin and oriented along the z -axis, of moment $4\pi\epsilon_0 a^3 E_0$.

(b) We can also find V by a much less formal method, as follows.

We require the term $-E_0 r \cos \theta$ so as to fit the condition at infinity. No other function with a positive power of r is permissible. This one term, however, is inadequate to fit the

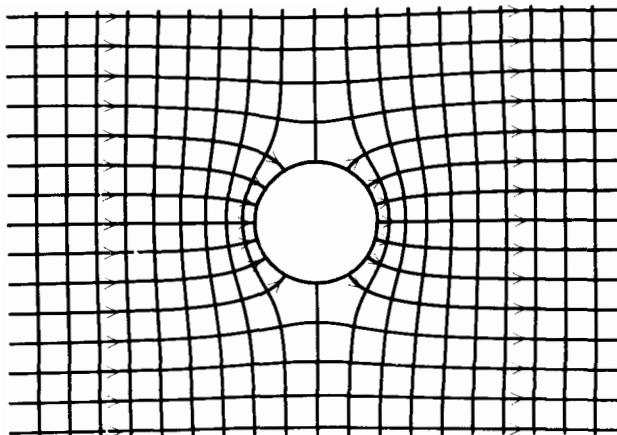


Fig. 12-1. Lines of \mathbf{E} (arrows) and equipotentials for an uncharged *conducting* sphere situated in a previously uniform electric field. The lines of \mathbf{E} are normal to the surface, and there is zero field inside. Observe that the field is hardly disturbed at distances larger than one radius from the surface of the sphere. The origin of the spherical polar coordinates used for the calculation is at the center of the sphere.

condition at $r = a$, where V must be independent of θ . We must therefore add another function that also includes the $\cos \theta$ factor so that the coefficient of $\cos \theta$ will be zero at $r = a$. Then, from Eq. 12-19,

$$V = -E_0 r \cos \theta + \frac{B \cos \theta}{r^2}. \quad (12-35)$$

We finally set $B = E_0 a^3$ to make $V = 0$ at $r = a$. Our solution satisfies both Laplace's equation and the boundary conditions; it is therefore the correct solution, according to the uniqueness theorem of Sec. 11.1. See Fig. 12-1.

Example

DIELECTRIC SPHERE IN A PREVIOUSLY UNIFORM ELECTRIC FIELD

We now have a previously uniform field E_0 , as in the preceding section, and the following boundary conditions, where a is the radius of the sphere:

- (1) V is continuous at $r = a$.
- (2) The normal component of \mathbf{D} is continuous at $r = a$.
- (3) $V_{r \rightarrow \infty} = -E_0 r \cos \theta$.

Instead of going through a formal solution, as we did in part (a) of the previous example, we proceed as in (b) and devise a combination of spherical harmonics that satisfies all three conditions.

There now exists a field inside the sphere. So we require two solutions, one that is valid inside and one that is valid outside. The field *outside* must satisfy boundary condition (3). So we require a term $-E_0 r \cos \theta$ and no other positive power of r . All negative powers of r qualify. *Inside* the sphere, there must be no negative powers of r , because V cannot become infinite at $r = 0$.

Let V_o and V_i be the potentials outside and inside the sphere, respectively. Then

$$V_o = -E_0 r \cos \theta + \sum_{n=0}^{\infty} B_n r^{-(n+1)} P_n(\cos \theta), \quad (12-36)$$

$$V_i = \sum_{n=0}^{\infty} C_n r^n P_n(\cos \theta). \quad (12-37)$$

Boundary conditions (1) and (2) require that, at $r = a$,

$$V_o = V_i, \quad -\frac{\partial V_o}{\partial r} = -\epsilon_r \frac{\partial V_i}{\partial r}. \quad (12-38)$$

Then we have the following two equations:

$$-E_0 a \cos \theta + \frac{B_0}{a} + \frac{B_1 \cos \theta}{a^2} + \frac{B_2 P_2(\cos \theta)}{a^3} + \dots$$

$$= C_0 + C_1 a \cos \theta + C_2 a^2 P_2(\cos \theta) + \dots \quad (12-39)$$

and

$$\begin{aligned} E_0 \cos \theta + \frac{B_0}{a^2} + \frac{2B_1 \cos \theta}{a^3} + \frac{3B_2 P_2(\cos \theta)}{a^4} + \dots \\ = -\epsilon_r [C_1 \cos \theta + 2C_2 a P_2(\cos \theta) + \dots]. \end{aligned} \quad (12-40)$$

These two equations apply for all values of θ . It follows that the coefficient of a given function of θ on the left must be equal to the coefficient of the same function on the right. From the first equation,

$$\frac{B_0}{a} = C_0, \quad -E_0 a + \frac{B_1}{a^2} = C_1 a, \quad \frac{B_2}{a^3} = C_2 a^2, \quad \dots \quad (12-41)$$

From the second,

$$\frac{B_0}{a^2} = 0, \quad E_0 + \frac{2B_1}{a^3} = -\epsilon_r C_1, \quad \frac{3B_2}{a^4} = -2\epsilon_r C_2 a, \quad \dots \quad (12-42)$$

Combining these equations gives

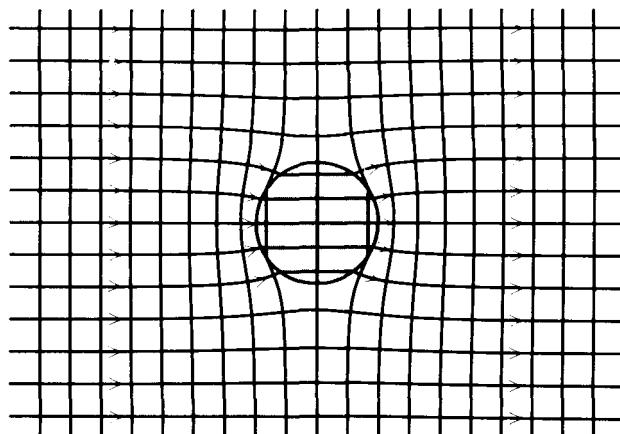


Fig. 12-2. Lines of \mathbf{D} (arrows) and equipotentials for a dielectric sphere ($\epsilon_r = 3$) situated in a previously uniform electric field. The lines of \mathbf{D} crowd into the sphere, and \mathbf{D} is larger inside than outside. The equipotentials spread out inside, so \mathbf{E} is smaller inside than outside. There exists a bound surface charge density, and \mathbf{E} is discontinuous at the surface. As for the conducting sphere (Fig. 12-1), the field is hardly disturbed at distances larger than one radius from the surface. The field inside is uniform.

$$B_0 = C_0 = 0, \quad B_1 = \frac{\epsilon_r - 1}{\epsilon_r + 2} E_0 a^3, \quad (12-43)$$

$$C_1 = -\frac{3}{\epsilon_r + 2} E_0, \quad B_n = C_n = 0 \quad (n > 1). \quad (12-44)$$

Finally,

$$V_o = -\left(1 - \frac{\epsilon_r - 1}{\epsilon_r + 2} \frac{a^3}{r^3}\right) E_0 r \cos \theta, \quad (12-45)$$

$$V_i = -\frac{3}{\epsilon_r + 2} E_0 r \cos \theta. \quad (12-46)$$

Of course, $r \cos \theta = z$.

Notice that, if $\epsilon_r \gg 1$, then V_o is approximately the same as the V of the conducting sphere of the previous section.

Also, the \mathbf{E} inside the sphere is uniform:

$$\mathbf{E}_i = -\nabla V_i = \frac{3}{\epsilon_r + 2} E_0 \hat{z}. \quad (12-47)$$

See Fig. 12-2.

12.2 POISSON'S EQUATION FOR \mathbf{E} IN ELECTROSTATIC FIELDS

We discussed Poisson's equation for V in Sec. 4.1. There also exists a Poisson equation for \mathbf{E} . From Sec. 1.11.6,

$$\nabla \times (\nabla \times \mathbf{E}) = -\nabla^2 \mathbf{E} + \nabla(\nabla \cdot \mathbf{E}). \quad (12-48)$$

Now, in electrostatic fields, the curl of \mathbf{E} is zero, so that

$$\nabla^2 \mathbf{E} = \nabla(\nabla \cdot \mathbf{E}) = \frac{\nabla \rho}{\epsilon_0}, \quad (12-49)$$

from Sec. 9.5, where ρ is the total charge density, free plus bound.

We can solve this equation by first separating it into its three components, the x -component being

$$\nabla^2 E_x = \frac{\partial}{\partial x} \frac{\rho}{\epsilon_0}. \quad (12-50)$$

This scalar equation is similar to the Poisson equation for V (Eq. 4-1), whose solution is the V of Eq. 3-18. Thus

$$E_x = -\frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\partial\rho/\partial x'}{r} dv', \quad (12-51)$$

and similarly for the other two components. As usual, the primes refer to the charge distribution, and r is the distance between the field point $P(x, y, z)$ and the source point $P'(x', y', z')$. The solution of Laplace's equation for \mathbf{E} is thus

$$\mathbf{E} = -\frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\nabla' \rho}{r} dv'. \quad (12-52)$$

This equation relates \mathbf{E} to the *gradient* of the total charge density. Note the *negative sign* and the *first* power of r in the denominator. This equation is valid whatever the nature of the media that are present in the field, as long as the gradient is definable.

The more usual expression for the \mathbf{E} of a volume distribution of charge is a consequence of Coulomb's law and it is the one that we found in Sec. 3.3:

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\rho \hat{r}}{r^2} dv'. \quad (12-53)$$

Although the two integrals for \mathbf{E} are equal, the integrands are obviously unequal. One integral applies only where $\nabla\rho$ is not zero, while the other extends over the complete charge distribution.

Example | THE FIELD OF A SPHERE OF CHARGE WHOSE DENSITY IS A FUNCTION OF THE RADIUS

Imagine a sphere of charge of uniform density, except near the periphery, where the density gradually falls to zero. With Eq. 12-52, it is only the region near the surface that contributes to the integral, since the gradient of the charge density is zero everywhere else. Nonetheless, this equation leads to the same result as Eq. 12-53. See Prob. 12-6.

12.3 SUMMARY

In spherical coordinates, if V is independent of the ϕ coordinate, we set

$$V(r, \theta) = R(r)\Theta(\theta). \quad (12-2)$$

Then Laplace's equation becomes a pair of ordinary differential equations

$$r^2 \frac{d^2 R}{dr^2} + 2r \frac{dR}{dr} - n(n+1)R = 0 \quad (12-7), (12-9)$$

and Legendre's equation

$$\frac{d}{d\mu} \left[(1-\mu^2) \frac{d\Theta}{d\theta} \right] + n(n+1)\Theta = 0, \quad (12-13)$$

where $\mu = \cos \theta$. The first equation defines a function of the type

$$R = Ar^n + \frac{B}{r^{n+1}}, \quad (12-8)$$

while the Legendre polynomials of Sec. 5.4.1 are solutions of Legendre's equation.

The *general solution of Laplace's equation in spherical coordinates, for axial symmetry* is thus

$$V = \sum_{n=0}^{\infty} (A_n r^n + B_n r^{-(n+1)}) P_n(\cos \theta). \quad (12-19)$$

Poisson's equation for \mathbf{E} is

$$\nabla^2 \mathbf{E} = \frac{\nabla \rho}{\epsilon_0}, \quad (12-49)$$

and its solution is

$$\mathbf{E} = -\frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\nabla' \rho}{r} dv'. \quad (12-52)$$

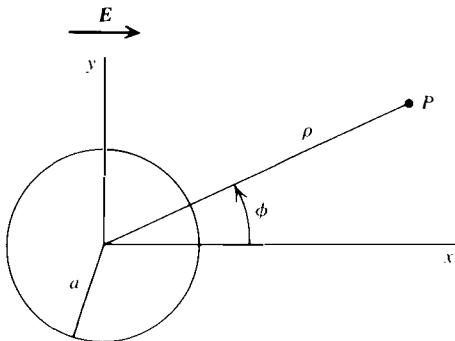
For electrostatic fields, we can calculate \mathbf{E} either as above or through Coulomb's law:

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\rho \hat{r}}{r^2} dv'. \quad (12-53)$$

PROBLEMS

12-1. (12.1) Grounded cylindrical conductor in a uniform \mathbf{E}

A grounded, infinite, circular cylindrical conductor of radius a lies in a previously uniform electric field, with its axis perpendicular to \mathbf{E}_0 , as in Fig. 12-3. Show that $V = -E_0(1 - a^2/\rho^2)\rho \cos \phi$.

**Fig. 12-3.**

- 12-2. (12.2) The volume integral of $\nabla' \times (\mathbf{E}/r)$ is zero

Show that $\int_{v'} \nabla' \times (\mathbf{E}/r) dv' = 0$ for any finite charge distribution. Use identity 19 from the front endpaper.

- 12-3. (12.2) Equality of the two integrals for \mathbf{E}

Show that the two integrals for \mathbf{E} ,

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\rho}{r^2} \hat{\mathbf{r}} dv' \quad \text{and} \quad \mathbf{E} = -\frac{1}{4\pi\epsilon_0} \int_{v'} \nabla' \frac{\rho}{r} dv'$$

are equal, at least if v' is finite. Use identities 15 and 18 from the front endpaper.

- 12-4. (12.2) The volume integral of $\nabla'(\rho/r)$ is zero

We have shown that, for static fields,

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\rho}{r^2} \hat{\mathbf{r}} dv' = -\frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\nabla' \rho}{r} dv'.$$

Show that, as a consequence,

$$\int_{v'} \nabla' \frac{\rho}{r} dv' = 0.$$

- 12-5. (12.2) The Dirac positron

Dirac proposed at one time that a positron could be considered as a hole in an infinite sea of negative electrons. Assume that positrons and electrons have finite dimensions.

Deduce the field of a positron from Coulomb's law on these assumptions. You can find this field without having to integrate!

- 12-6. (12.2) The electric field of an atomic nucleus

Figure 12-4 shows the density as a function of the radius for a given sphere of charge. This corresponds roughly to the charge distribution inside an atomic nucleus.

(a) Use Gauss's law to find E at a distance $r > \beta$ from the center of the sphere.

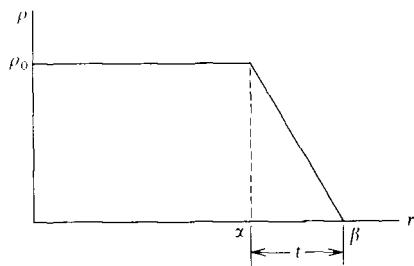


Fig. 12-4.

- (b) Show that the expression for E given in Eq. 12-52 leads to the same result. Both calculations remain valid when $\beta - \alpha$ tends to zero.

CHAPTER 13

*RELATIVITY I *The Lorentz Transformation and Space-Time*

13.1	REFERENCE FRAMES AND OBSERVERS	239
13.2	THE GALILEAN TRANSFORMATION	240
13.3	THE PRINCIPLE OF RELATIVITY	240
13.4	THE LORENTZ TRANSFORMATION	241
	<i>Example: THE INVARIANCE OF $x^2 + y^2 + z^2 - c^2t^2$ UNDER A LORENTZ TRANSFORMATION</i>	243
	<i>Example: THE PHOTON</i>	244
13.5	SPACE-TIME DIAGRAMS AND WORLD LINES	244
13.5.1	THE MINKOWSKI DIAGRAM	245
13.5.2	CAUSALITY AND MAXIMUM SIGNAL VELOCITY	247
13.6	SUMMARY	247
	PROBLEMS	248

We have now studied electric fields at quite some length. At this point either you can skip to Chap. 18 and go on to magnetic fields, or you can study Chaps. 13 to 17, which deal with relativity.

There are two reasons for this digression. First, relativity reveals the fundamental aspects of electromagnetism. Second, there are many phenomena, mostly related to high velocities, that are baffling without relativity. The daggered problems, from Chap. 18 on, are examples of these.

The longer path is more interesting, as always, but it may not be the better one. The five chapters on relativity do not replace the more conventional approach that comes afterward. Selecting one path or the other is a matter of time and personal taste.

You can therefore go directly to Chap. 18 without losing continuity.

Chapters 13, 14, and 15 set forth the fundamentals of relativity, with little reference to electromagnetism. Then, in Chaps. 16 and 17, we first

relate electric fields to magnetic fields and then establish relations that we shall later rediscover, starting with Chap. 18, without referring to relativity.[†]

13.1 REFERENCE FRAMES AND OBSERVERS

Consider the two rigid reference frames S and S' of Fig. 13-1, where S' moves at some constant velocity \mathcal{V} with respect to S .

Both frames are inertial, by assumption. We define an *inertial frame* as one that does not accelerate and that does not rotate. In other words, an inertial frame is one with respect to which there are no inertial forces.[‡]

Each reference frame carries an *observer*, which is either a human being equipped with instruments, or some device that can take readings or photographs, either automatically or under remote control. Observer o is situated in S , and o' in S' .

The *special theory of relativity* concerns measurements made by these two observers. As a rule, we refer to the two specific frames of Fig. 13-2.

We are *not* concerned with the *general theory of relativity*, which deals with accelerated frames and gravitation.

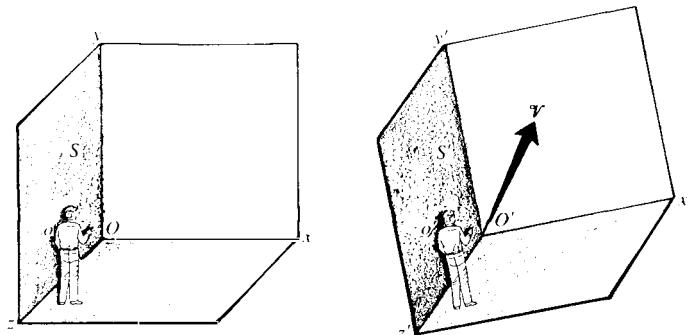


Fig. 13-1. Reference frames S and S' , with S' moving at some arbitrary constant velocity \mathcal{V} with respect to S , *without* rotation. Observers o and o' perform various experiments, each one in his or her own frame.

[†]For a more detailed introduction to relativity, see Edwin F. Taylor and John A. Wheeler, *Spacetime Physics*, W. H. Freeman, New York, 1966.

[‡]The forces that one feels while stationary inside an accelerating vehicle are *inertial forces*. The inertial force on a person of mass m is $-ma$, where a is the acceleration of the vehicle.

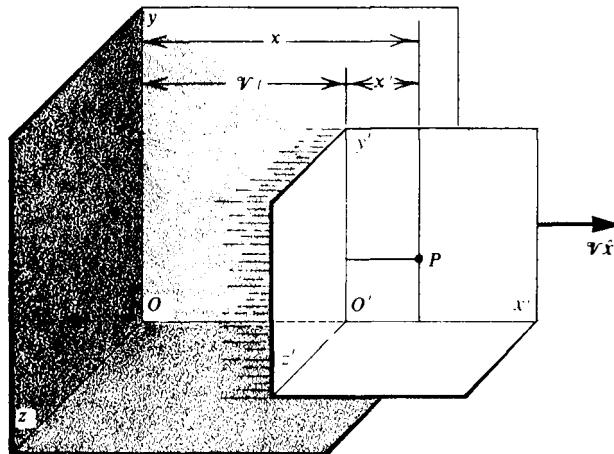


Fig. 13-2. The reference frame S' moves at a velocity $\mathcal{V}\hat{x}$ with respect to S .

13.2 THE GALILEAN TRANSFORMATION

According to *classical* physics, observers in reference frames S and S' of Fig. 13-2 may select the same time scale. If they set $t = 0$ at the instant when their frames overlap, then the coordinates of a given point in space, with respect to the two frames, satisfy the equations

$$x = x' + \mathcal{V}t, \quad y = y', \quad z = z', \quad t = t'. \quad (13-1)$$

This is the *Galilean transformation*.

In particular, suppose observer o' finds that the velocity of a certain light pulse relative to S' is c , in the positive direction of the common x -axis. Then, according to the Galilean transformation and according to elementary mechanics, observer o in reference frame S should find that the velocity of that particular light pulse is $c + \mathcal{V}$.

The Galilean transformation proves to be in error for high velocities. Experiments show that the velocity of the light pulse is the *same* for the two observers. This fact goes against common sense, of course. The Galilean transformation also proves to be incompatible with electromagnetism because, under that transformation, Maxwell's equations are not invariant.

13.3 THE PRINCIPLE OF RELATIVITY

The *principle of relativity* states that it is physically impossible to detect whether an inertial frame is at rest or in motion from observations made entirely within that frame.

This principle was formulated by Galileo in 1638. It is quite clear in itself, but it is so fundamental that we state it in another way, to emphasize its meaning. The principle means that *any* experiment leads to precisely the same result, whether it is performed in S or in S' . The result is the same, whether the experiment takes place in a stationary or in a moving vehicle, as long as the vehicle moves in a straight line at a constant velocity.

For example, observer o' in reference frame S' can measure the period of a pendulum suspended at some point in S' , or the collision of billiard balls on a table at rest in S' , etc. In all instances the phenomenon observed by o' is precisely the same as if she performed her experiment in reference frame S .

This principle is firmly established. In particular, it is an experimental fact that the speed of light is the same for all observers traveling at a constant velocity.⁺

To illustrate, consider the law

$$\mathbf{F} = \frac{d}{dt}(m\mathbf{v}), \quad (13-2)$$

as observed in reference frame S . According to the principle of relativity, there exists an identical law

$$\mathbf{F}' = \frac{d}{dt'}(m'\mathbf{v}') \quad (13-3)$$

that applies in S' . Physical *laws* are thus said to be *invariant*.

It follows that there exist mathematical relations that *transform* unprimed quantities to primed quantities, and inversely. We shall discover several such *transformations* in these five chapters on relativity.

With the *Galilean* transformation, $F = F'$, $m = m'$, $t = t'$, $v = v' + \mathcal{V}$, and the principle of relativity applies. However, nature is not that simple; these four equations are, in fact approximations. Relativistic transformations show that $F \neq F'$, $m \neq m'$, $t \neq t'$, and $v \neq v' + \mathcal{V}$!

13.4 THE LORENTZ TRANSFORMATION

The *Lorentz transformation* relates the space and time coordinates of reference frame S to those of S' , and inversely. For the frames of Fig. 13-2,

⁺ See Wolfgang K. H. Panofsky and Melba Phillips, *Classical Electricity and Magnetism*, Addison-Wesley, Reading, Mass., 1962, chap. 15.

$$x = \frac{x' + \mathcal{V}t'}{(1 - \mathcal{V}^2/c^2)^{1/2}}, \quad x' = \frac{x - \mathcal{V}t}{(1 - \mathcal{V}^2/c^2)^{1/2}}, \quad (13-4)$$

$$y = y', \quad y' = y, \quad (13-5)$$

$$z = z', \quad z' = z, \quad (13-6)$$

$$t = \frac{t' + (\mathcal{V}/c^2)x'}{(1 - \mathcal{V}^2/c^2)^{1/2}}, \quad t' = \frac{t - (\mathcal{V}/c^2)x}{(1 - \mathcal{V}^2/c^2)^{1/2}}, \quad (13-7)$$

where c is the speed of light in a vacuum.

This set of eight equations forms the basis of special relativity.[†] We spend the rest of this chapter, and the next four, discussing some of its strange consequences. For the moment, we note six fairly obvious features.

- (1) The right-hand column is identical to the left-hand column, except that the primed and unprimed quantities are interchanged, and that $-\mathcal{V}$ replaces \mathcal{V} . The reason is simply that it is immaterial whether S' moves at the velocity $\mathcal{V}\hat{x}$ with respect to S , or whether S moves at the velocity $-\mathcal{V}\hat{x}$ with respect to S' .
- (2) There are only four independent equations because the right-hand column follows from the left-hand one, and inversely. You can easily check this.
- (3) The origins O and O' coincide at $t = t' = 0$ because, if $x = y = z = t = 0$, then $x' = y' = z' = t' = 0$, and inversely.
- (4) The Lorentz transformation reduces to the Galilean transformation (Eq. 13-1) if we set the speed of light c equal to infinity.
- (5) The relative velocity \mathcal{V} of the two frames cannot exceed c , for otherwise either x and t or x' and t' become imaginary.
- (6) The fact that $t \neq t'$ means that, if observer o measures a time interval T between two events, then o' will, in general, measure a different time interval T' between the same pair of events. In particular, if two events are simultaneous for o , then they are not necessarily simultaneous for o' .

Points 1 and 2 illustrate two general rules.

[†]The Lorentz transformation follows from the principle of relativity applied to the velocity of light. We omit the proof, for lack of space. See, for example, E. Taylor and J. Wheeler, *Spacetime Physics*, W. H. Freeman, New York, 1966, p. 43.

- (1) The relation between quantities in one inertial reference frame and the corresponding quantities in another inertial frame can always be expressed by either one of two equations, or sets of equations, that are equivalent.
- (2) One obtains the inverse relation by adding primes to unprimed quantities, deleting primes on primed quantities, and changing the sign of \mathcal{V} .

It is the custom to set

$$\beta = \frac{\mathcal{V}}{c} \quad \text{and} \quad \gamma = \frac{1}{(1 - \beta^2)^{1/2}}. \quad (13-8)$$

Note that $\beta \leq 1$ and $\gamma \geq 1$. Then

$$x = \gamma(x' + \mathcal{V}t'), \quad x' = \gamma(x - \mathcal{V}t), \quad (13-9)$$

$$y = y', \quad y' = y, \quad (13-10)$$

$$z = z', \quad z' = z, \quad (13-11)$$

$$t = \gamma\left(t' + \frac{\mathcal{V}x'}{c^2}\right), \quad t' = \gamma\left(t - \frac{\mathcal{V}x}{c^2}\right). \quad (13-12)$$

It is often more convenient to express the transformation in vector form, as in Table 13-1, where the subscripts \parallel and \perp refer, respectively, to the components that are either parallel or perpendicular to the motion of S' with respect to S .

Table 13-1 Lorentz transformation

$\mathbf{r} = \gamma(\mathbf{r}'_{\parallel} + \mathcal{V}\mathbf{r}'_{\perp}) + \mathbf{r}'_{\perp}$	$\mathbf{r}' = \gamma(\mathbf{r}_{\parallel} - \mathcal{V}) + \mathbf{r}_{\perp}$
$t = \gamma\left(t' + \frac{\mathcal{V}r'_{\parallel}}{c^2}\right)$	$t' = \gamma\left(t - \frac{\mathcal{V}r_{\parallel}}{c^2}\right)$

Example THE INVARIANCE OF $x^2 + y^2 + z^2 - c^2t^2$ UNDER A LORENTZ TRANSFORMATION

A quantity is said to be *invariant* if its numerical value is the same in all inertial frames. For example, under a Galilean transformation, the distance r_{ab} between two points a and b is the same whether one performs the measurement in S or in S' .

Note that the word *invariant* is *not* synonymous with the word *constant*. Indeed, if point a is fixed in S , and b is fixed in S' , then r_{ab} is a function of the time but is still invariant under a Galilean transformation.

With the *Lorentz* transformation, r_{ab} is *not* invariant, but $r^2 - c^2t^2$ is. You can easily check that

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

or $r^2 - c^2t^2 = r'^2 - ct'^2. \quad (13-13)$

We shall use the invariance of $r^2 - c^2t^2$ below. We shall also find several other analogous invariant quantities.

Example

THE PHOTON

Imagine a flash of light emitted at O or O' at the instant when the two origins coincide. According to the principle of relativity, the light propagates in all directions at the *same* velocity c in both reference frames. Therefore the position of a photon satisfies the equation

$$\frac{x^2 + y^2 + z^2}{t^2} = \frac{x'^2 + y'^2 + z'^2}{t'^2} = c^2, \quad (13-14)$$

in agreement with the previous equations.

13.5 SPACE-TIME DIAGRAMS AND WORLD LINES

A *space-time diagram* is a graph of ct as a function of x for a given event, as in Fig. 13-3. Such diagrams help us visualize some of the implications of the Lorentz transformation for the reference frames of Fig. 13-2.

The *world line* of an object or event is its curve of ct as a function of x . Say an electron is stationary at $x = x_0$, $ct = 0$. Its world line is the straight line a of Fig. 13-3. If the electron moves in some way along the x -axis, then its world line is a curve such as b . However, b is not any arbitrary curve because the speed of any object is always less than c . This makes the magnitude of the slope everywhere larger than unity, and the electron has no access to the shaded region.

If the electron moves in a plane normal to the x -axis, then $y = y'$, $z = z'$, from the Lorentz transformation.

Now say a flash of light occurs at $x = 0$, $ct = 0$. Photons travel along the lines $ct = \pm x$. If we draw a y -axis perpendicular to the x - and ct -axes, the photons travel along a cone whose axis is the ct -axis. In four-dimensional space-time, they follow a *light cone*.

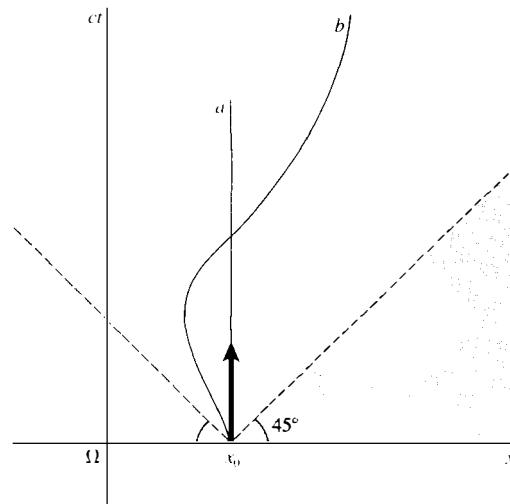


Fig. 13-3. (a) World line of a stationary electron at $(x_0, 0, 0)$. (b) World line of an electron whose velocity dx/dt is first negative and then positive along the x -axis. The shaded area is inaccessible to this electron because $|dx/dt|$ cannot exceed c .

13.5.1 The Minkowski Diagram

We can picture the Lorentz transformation by superposing the (x, ct) and the (x', ct') planes in a manner devised by Minkowski as in Fig. 13-4. First,

$$x' = \gamma(x - \mathcal{V}t). \quad (13-15)$$

Setting $x' = 0$ defines a line

$$x = \mathcal{V}t = \left(\frac{\mathcal{V}}{c}\right)ct \quad (13-16)$$

in the (x, ct) plane. This line, along which $x' = 0$, is the ct' -axis, as in the figure. Also,

$$t' = \gamma\left(t - \frac{\mathcal{V}x}{c^2}\right). \quad (13-17)$$

Setting $t' = 0$ defines a line

$$ct = \left(\frac{\mathcal{V}}{c}\right)x \quad (13-18)$$

in the (x, ct) plane. This line, along which $ct' = 0$, is the x' -axis. Lines

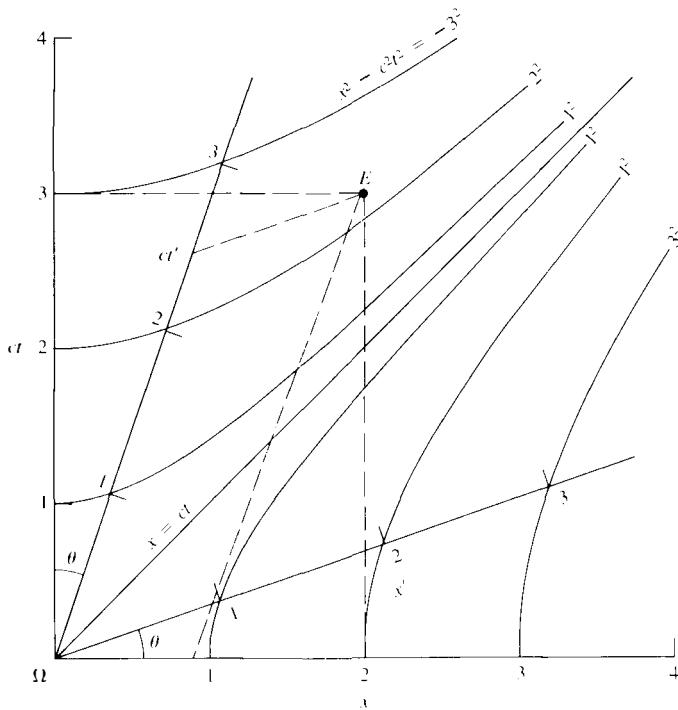


Fig. 13-4. The (x, ct) plane of reference frame S and, superimposed, x' - and ct' -axes for a reference frame S' that travels at a velocity $0.364c\hat{x}$ with respect to S . Thus $\theta = 20.0^\circ$ and $\gamma = 1.074$. The origins $x = 0, t = 0$ and $x' = 0, t' = 0$ coincide at Ω . Event E occurs at $x = 2, ct = 3$, or at $x' = 0.975, ct' = 2.44$. The curves are parabolas $x^2 - c^2t^2 = \pm a^2$. This is a *Minkowski diagram*.

$x' = \text{constant}$ are parallel to the ct' -axis, while lines $ct' = \text{constant}$ are parallel to the x' -axis.

Notice the following points.

- (1) *The scales are not the same.* For example, a length corresponding to 3 meters is longer on the x' -axis than on the x -axis.
- (2) The angle between the ct - and ct' -axes is the same as that between the x - and x' -axes. This reveals a certain symmetry between space and time. However, there are three space dimensions and only one time dimension, so the symmetry is only partial.
- (3) The lines $x = \pm ct$ coincide with the lines $x' = \pm ct'$ because a velocity c along the x -axis is invariant.
- (4) Points on the x' -axis correspond to events that are simultaneous for observer o' in reference frame S' since $t' = 0$ all along this line. However,

those events clearly occur at different values of t and are therefore not simultaneous in S . Innumerable paradoxes result from this fact.

If two events F and G occur at the same t and at the same x (but not necessarily the same y and the same z) in S , then they are also simultaneous in S' , and inversely.

13.5.2 Causality and Maximum Signal Velocity

The *order* in which two events F and G occur can be different in different frames because

$$t'_G - t'_F = \gamma \left[(t_G - t_F) - \left(\frac{\gamma}{c^2} \right) (x_G - x_F) \right], \quad (13-19)$$

and the signs of $t'_G - t'_F$ and $t_G - t_F$ can be different. This is disturbing because, according to the *principle of causality*, a cause necessarily precedes its effect.

For example, imagine that observer o throws a ball in the direction of the x -axis. After a flight of a few seconds, the ball breaks a windowpane. The Lorentz transformation surely cannot mean that these events could occur backward, for certain observers.

Say event F occurs at origins O and O' at the instant when they coincide. Event F is the cause of event G , which occurs at x_G at a later time t_G in S and at t'_G in S' . Event G cannot occur before event F in S' . Then t'_G must not be negative. Now

$$t'_G = \gamma \left[t_G - \left(\frac{\gamma}{c^2} \right) x_G \right] = \gamma \left[t_G - \left(\frac{\gamma}{c^2} \right) v t_G \right] \quad (13-20)$$

$$= \gamma t_G \left(1 - \frac{\gamma v}{c^2} \right), \quad (13-21)$$

where v is the speed at which a “signal” propagates from F to G . For the above example, v is the horizontal velocity of the ball. Therefore $\gamma v \leq c^2$. Since $\gamma \leq c$ from Sec. 13.4, $v \leq c$. A “signal” cannot therefore propagate at a speed that is larger than c .

13.6 SUMMARY

An *inertial reference frame* does not accelerate and does not rotate.

The *special theory of relativity* concerns observations and measurements on a given phenomenon made by observers situated in two inertial

frames, one of which moves at a constant velocity with respect to the other. As a rule, the two frames referred to are those of Fig. 13-2.

The *principle of relativity* states that it is physically impossible to detect whether an *inertial* reference frame is at rest or in uniform motion from observations made entirely within that frame.

Thus *physical laws are invariant*: the law describing a given phenomenon is mathematically the same, in whatever inertial frame the phenomenon occurs.

The *Lorentz transformation* relates x, y, z, t in S to x', y', z', t' in S' (Fig. 13-2):

$$x = \gamma(x' + \mathcal{V}t'), \quad x' = \gamma(x - \mathcal{V}t), \quad (13-9)$$

$$y = y', \quad y' = y, \quad (13-10)$$

$$z = z', \quad z' = z, \quad (13-11)$$

$$t = \gamma\left(t' + \frac{\mathcal{V}x'}{c^2}\right), \quad t' = \gamma\left(t - \frac{\mathcal{V}x}{c^2}\right), \quad (13-12)$$

where $\gamma = (1 - \mathcal{V}^2/c^2)^{-1/2}$.

Table 13-1 shows these equations in vector form. The subscripts \parallel and \perp refer, respectively, to components that are parallel or perpendicular to the velocity of S' with respect to S .

A quantity is said to be *invariant* if its numerical value is the same in all inertial reference frames.

A *space-time diagram* shows ct as a function of x for a given object or event. The *world line* of an object or event is its curve of ct as a function of x . One can visualize the Lorentz transformation by superposing the space-time diagrams for S and for S' .

A signal cannot propagate at a speed larger than c .

PROBLEMS[†]

13-1. (13.4) The Lorentz transformation

For what value of β is the value of γ equal to 1.01?

13-2. (13.4) The Lorentz transformation

Calculate γ , β , and v for a conduction electron whose energy is 10 electronvolts. The rest energy of an electron is 5.11×10^5 electronvolts.

[†] Several of the problems on relativity are adapted, with permission, from Edwin F. Taylor and John A. Wheeler, *Spacetime Physics*, W. H. Freeman, New York, 1966.

13-3. (13.4) Signaling problems with a fast train

Three persons A , O' , and B , ride on a train moving at a velocity \mathcal{V} , with A in front, O' in the middle, and B in the rear. A fourth person, O , stands beside the rails. At the moment O' passes O , light signals from A and B reach both O and O' . Persons O and O' are asked who emitted her light signal first. What do they answer?

13-4. (13.4) Transformation of an angle

A straight line passing through the origin O' of S' forms an angle α' with the x -axis. (a) Find a relation between α and α' .

(b) What is the value of α when \mathcal{V} tends to c ?

13-5. (13.4) Things that move faster than a photon

The Lorentz transformation implies that the relative velocity \mathcal{V} of two frames of reference cannot exceed the speed of light c . We have also shown that a signal cannot exceed the speed of light. Discuss the following cases.

(a) A long, straight rod forms a small angle θ with another rod, which is horizontal and stationary. The first rod moves downward at a velocity v .

What is the speed of the point of intersection of the lower edge of the moving rod with the fixed rod? Can this speed be greater than c ? Can the point of intersection be used to transmit a signal?

(b) The upper rod is initially at rest with the point of intersection at the origin. The rod is struck a downward blow at the origin with a hammer.

Can the motion of the point of intersection be used to transmit a signal at a speed greater than the speed of light?

(c) A powerful laser rotates rapidly about an axis perpendicular to its length.

Can the azimuthal speed of the beam exceed the speed of light? Can the beam transmit a signal between two points at a speed greater than c ?

(d) The manufacturers of some oscilloscopes claim writing speeds in excess of the speed of light. Is this possible?

13-6. (13.5) Space-time diagrams

Show that

$$x' + ct' = \left(\frac{1 - \beta}{1 + \beta} \right)^{1/2} (x + ct).$$

Substituting $-c$ for c gives

$$x' - ct' = \left(\frac{1 + \beta}{1 - \beta} \right)^{1/2} (x - ct).$$

13-7. (13.5) c is the ultimate speed

Imagine a series of reference frames S , S' , S'' , S''' , etc., with S' moving at a velocity $\mathcal{V}\hat{x}$ with respect to S , S'' moving at the same velocity with respect to S' , etc. According to the Galilean transformation, a particle at rest in S'' moves at a velocity $n\mathcal{V}\hat{x}$ with respect to S , where $n\mathcal{V}$ is arbitrarily large.

Show that, according to relativity, the velocity of that particle with respect to S is always less than c .

13-8. (13.5) Three-dimensional space-time x , y , ct

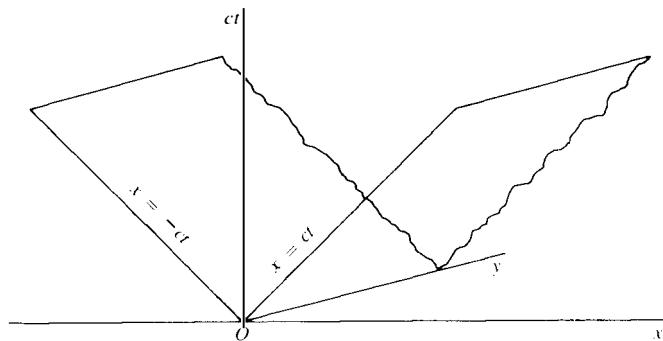
**Fig. 13-5.**

Figure 13-5 shows the xy -plane, the ct -axis, and the planes $x = ct$ and $x = -ct$.

- (a) Show that a Lorentz transformation contracts this space by the factor $[(1 + \beta)/(1 - \beta)]^{1/2}$ in the direction of the plane $x = ct$, and dilates it by the same factor in the direction of $x = -ct$.
- (b) Show that a point x_0, y_0, ct_0 transforms to another point on the same light cone.

13-9. (13.5.1) The Minkowski diagram

- (a) Draw a Minkowski diagram similar to that of Fig. 13-4 with $\gamma = -0.346c$ and an event E that occurs at $x = 2, ct = 2$.
- (b) What are the values of coordinates x' and ct' ?

CHAPTER 14

*RELATIVITY II

The Lorentz Contraction and Time Dilation. The Transformation of a Velocity

14.1	TRANSFORMATION OF A LENGTH. THE LORENTZ CONTRACTION	251
14.1.1	TRANSFORMATION OF AN ELEMENT OF AREA dA	253
	Example: THE APPARENT SHAPE OF A RAPIDLY MOVING OBJECT	254
14.2	TRANSFORMATION OF A TIME INTERVAL. TIME DILATION	254
	Example: THE TIME READ ON A RAPIDLY MOVING CLOCK	255
	Example: THE RELATIVISTIC DOPPLER EFFECT FOR ELECTROMAGNETIC WAVES	257
14.3	THE INCREMENT OF PROPER TIME FOR AN ACCELERATED PARTICLE	258
14.4	TRANSFORMATION OF A VELOCITY	259
	Example: TRANSFORMING THE SPEED OF LIGHT	260
14.5	SUMMARY	260
	PROBLEMS	261

This second chapter on relativity concerns a few immediate consequences of the Lorentz transformation. The Lorentz contraction makes an object appear shortened in the direction of its motion, while time dilation makes a time interval on a moving object appear longer for a stationary observer. However, a moving clock can also appear to run fast, as we shall see. We end this chapter with the transformation of a velocity.

14.1 TRANSFORMATION OF A LENGTH. THE LORENTZ CONTRACTION

Imagine that observer o' on reference frame S' fixes a ruler of length l_0 on the x' -axis of S' so that its extremities are at $x' = 0$ and $x' = l_0$. The quantity l_0 is the length of the ruler, as measured in its own reference frame, and is called its *proper length*. According to o , the ruler sweeps the shaded area in Fig. 14-1.

What is the length of the *same* ruler for observer o on S ? That observer performs his measurement by noting the positions of the two extremities

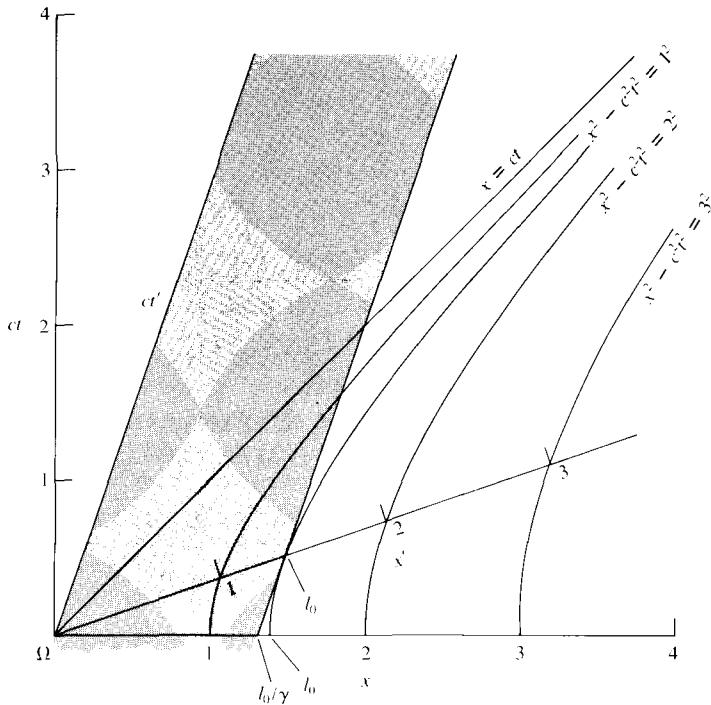


Fig. 14-1. A ruler of proper length l_0 , fixed in reference frame S' , sweeps the shaded region as time goes by. In reference frame S its apparent length is l_0/γ . Note that the scale on the x' -axis is *different* from that on the x -axis. The curves are parabolas $x^2 - c^2t^2 = a$.

of the ruler at the same time, say $t = 0$. These are events Ω and $(l_0/\gamma, 0)$ in Fig. 14-1. Clearly, the length l_0/γ is shorter than l_0 . How much shorter? At the right-hand end $x = l$ and $t = 0$. From the Lorentz transformation (Sec. 13.4),

$$l_0 = x' = \gamma(x - vt) = \gamma l, \quad l = \frac{l_0}{\gamma}. \quad (14-1)$$

Thus a ruler moving in the direction of its length at a velocity v relative to an observer appears to be shortened by the factor $1/\gamma$. Remember that $\gamma \geq 1$, from Sec. 13.4. This *Lorentz contraction* is independent of the sign of v .

Of course, the Lorentz contraction applies if the ruler is anywhere else along the x' -axis.

If the ruler lies on the x -axis in S , then o' finds it shortened by the same factor $1/\gamma$, as in Fig. 14-2.

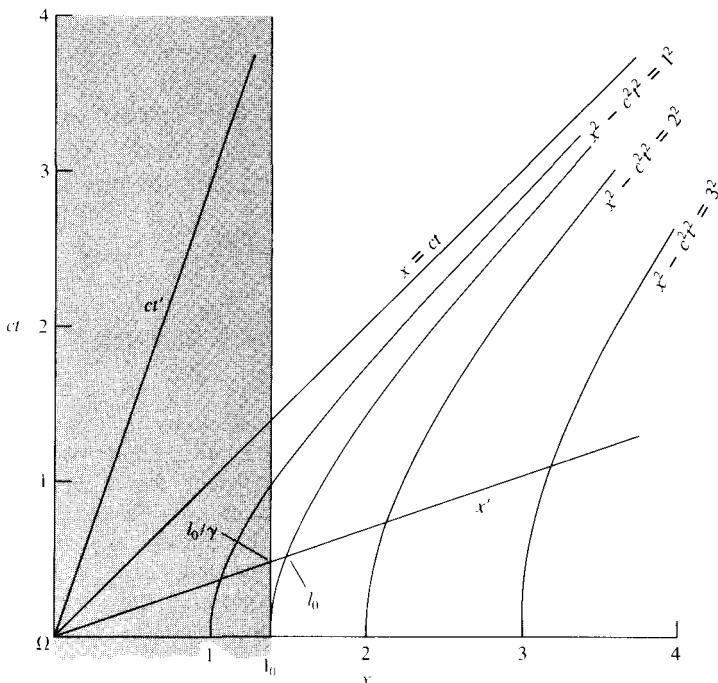


Fig. 14-2. The ruler of proper length l_0 is now fixed in reference frame S . It sweeps the shaded region and, for an observer on frame S' , its length is l_0/γ .

Thus o says that the meters of o' are too short, and o' maintains that those of o are too short. This is not absurd because the comparisons are quite complex; they involve eight separate measurements.

If the ruler moves relative to o in a direction perpendicular to its length, then you can show that there is no Lorentz contraction: its length, measured by o , is equal to its proper length l_0 . A length l therefore transforms as follows:

$$l = \frac{l_{0\parallel}}{\gamma} + l_{0\perp}, \quad (14-2)$$

where $l_{0\parallel}$ is the component of l_0 that is parallel to the motion and $l_{0\perp}$ is the orthogonal component.

14.1.1 Transformation of an Element of Area dA

Say the element of area is a rigid parallelogram of sides dl_{10} and dl_{20} of arbitrary magnitudes and orientations in its own reference frame. Then

$$d\mathcal{A}_0 = dl_{10} \times dl_{20}. \quad (14-3)$$

This vector is normal to the element of area.

Now if S is an inertial reference frame with respect to which the element of area moves at a velocity \mathcal{V} , it turns out (Prob. 14-3) that, in S , the element of area suffers a Lorentz contraction in the direction of \mathcal{V} :

$$d\mathcal{A} = dl_1 \times dl_2 = d\mathcal{A}_{0\parallel} + \frac{d\mathcal{A}_{0\perp}}{\gamma}. \quad (14-4)$$

Example

THE APPARENT SHAPE OF A RAPIDLY MOVING OBJECT

Suppose one looked through a telescope at a far-away cube moving at a velocity $\mathcal{V} \approx c$ perpendicular to the line of sight. Then the face normal to the line of sight would appear to be foreshortened in the direction of motion by the factor $1/\gamma$. Also, one would see the trailing face for the following reason. At a given instant, the eye senses the photons that arrive at that instant. Photons originating from distant parts of the object have left earlier than the others, and the object has moved in the meantime. The net effect is that the cube would appear to be rotated through an angle $\arctan(\mathcal{V}/c)$. If the cube were not far away, then it would appear distorted in peculiar ways, depending on its distance and velocity.

This effect, which has never been observed, was discovered by James Terrell in 1959, 54 years after the publication of Einstein's first paper on relativity.[†]

14.2 TRANSFORMATION OF A TIME INTERVAL. TIME DILATION

Observer o' measures the duration of a certain phenomenon that occurs at $x' = 0$ in S' . The phenomenon starts at $t' = 0$, namely at Ω in Fig. 14-3, and it ends at $t' = T_0$, or at Q' . The time T_0 measured in the reference frame of the phenomenon is the *proper time*. Observer o' has a single clock located at $x' = 0$.

Observer o measures the same time interval with two identical and synchronized clocks, one at $x = 0$, where the phenomenon starts, and the other at the position of O' at the end of the time interval.

[†] See V. F. Weisskopf, *Physics Today*, September 1960, p. 24.

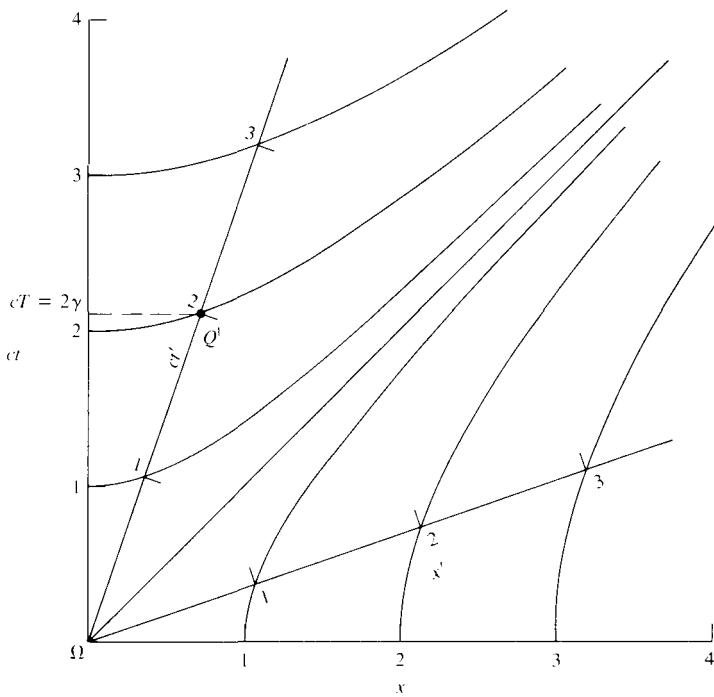


Fig. 14-3. The proper time interval between events Ω and Q' that occur at $x' = 0$ in reference frame S' is T_0 , with $cT_0 = 2$. For an observer on S , this time interval is γ times longer. The scale on the ct' -axis is *not* the same as on the ct -axis.

From Fig. 14-3, the interval T , as measured in S , is longer than the proper time interval T_0 . Indeed, according to the Lorentz transformation, the event Q' for which $x' = 0$ and $t' = T_0$ occurs at $t = \gamma T_0$ for o in S . So

$$T = \gamma T_0, \quad (14-5)$$

For o , the time interval is longer than T_0 . This is the phenomenon called *time dilation*. In other words, a moving clock appears to run slow by the factor γ if one measures its rate as above.[†]

Example | THE TIME READ ON A RAPIDLY MOVING CLOCK

We have seen that, if o uses two identical clocks at two different x 's on S to measure a given time interval that o' measures at a

[†] See the film entitled “Time Dilation: An Experiment on Mu-Mesons” by F. Friedman, D. Frisch, and J. Smith, produced by the Educational Development Center, Newton, Mass.

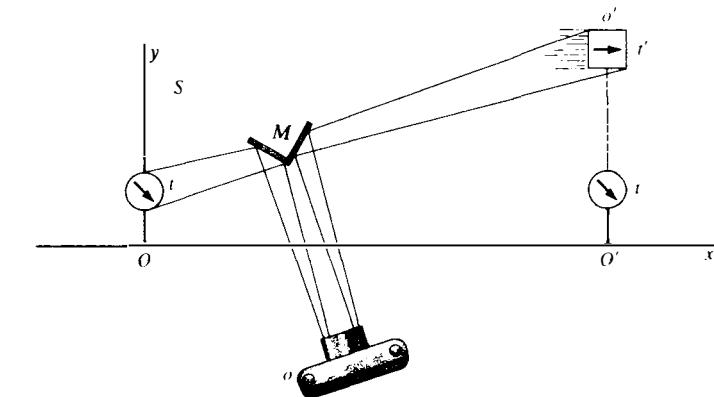


Fig. 14-4. Observer o uses a double mirror M to photograph simultaneously both her or his own clock and the moving clock. The relation $t = \gamma t'$ holds, but the light from the moving clock does not reach the camera until a later time $t'' > t$.

fixed point on S' with a single clock, then o finds a time interval that is longer than that of o' by the factor γ .

What if o uses a single clock and *looks* at the moving clock of o' as in Fig. 14-4? Let the primed clock be at O' , and let both V and t be positive. Then O' is to the right of O and moves *away*.

Imagine that o has a set of identical and synchronized clocks all along his x -axis. As the primed clock goes by each one, the relation $t = \gamma t'$ holds. But o stands at the origin O of S , and the light from the primed clock at O' takes some time to reach O . Suppose observer o reads a time t' on the moving clock. What time is it on his own clock? Call this time t'' . Then t'' is the above t , plus the time required for light to travel the distance Vt :

$$t'' = t + \frac{Vt}{c} = \left(1 + \frac{V}{c}\right)t = (1 + \beta)t = (1 + \beta)\gamma t', \quad (14-6)$$

with $\beta = V/c$, as usual. Thus, when o reads t' on a clock that is moving away, her own time is

$$t'' = \frac{1 + \beta}{(1 - \beta)^{1/2}} t' = \left(\frac{1 + \beta}{1 - \beta}\right)^{1/2} t'. \quad (14-7)$$

With V and β both positive, $t'' > t'$.

Therefore, if one *looks* at a clock that is moving *away*, the moving clock appears to run even slower than with the measurements of Sec. 14.2.

What if the moving clock is *approaching*? Then the origin O' is to the left of O . With V positive, both t and t' are negative and

$$t'' = t + \frac{|Vt|}{c} = t - \frac{Vt}{c} = (1 - \beta)t = (1 - \beta)\gamma t' \quad (14-8)$$

is also negative, since $\beta \leq 1$. Thus, when o reads a time t' on an approaching clock, his own time is

$$t'' = \left(\frac{1 - \beta}{1 + \beta} \right)^{1/2} t'. \quad (14-9)$$

Does the approaching clock appear to run fast, or slow, with respect to the stationary clock? Say observer o reads a time $t' = -1$ second on the approaching clock, and the above square root equals 0.9. Then $t'' = -0.9$. When the moving clock reaches 0, both t' and t'' will be zero. So, in the interval, the moving clock will have advanced by 1 second, and the fixed clock by 0.9 second. The approaching clock appears to run *fast*.

Example

THE RELATIVISTIC DOPPLER EFFECT FOR ELECTROMAGNETIC WAVES

The *Doppler effect* is the frequency shift observed when a source of waves moves with respect to a detector. This phenomenon is well known in the field of acoustics.

Imagine a source of electromagnetic waves of proper frequency f_0 situated at O' and a detector at O . What is the apparent frequency at O ?

This problem is identical to the clock problem that we just solved, because the source beats periods $1/f_0$, instead of seconds. Therefore, with a receding source, the period measured by a fixed observer at O is

$$T = \left(\frac{1 + \beta}{1 - \beta} \right)^{1/2} T_0, \quad (14-10)$$

and

$$f = \frac{1}{T} = \left(\frac{1 - \beta}{1 + \beta} \right)^{1/2} \frac{1}{T_0} = \left(\frac{1 - \beta}{1 + \beta} \right)^{1/2} f_0 < f_0. \quad (14-11)$$

For an approaching source,

$$f = \left(\frac{1 + \beta}{1 - \beta} \right)^{1/2} f_0 > f_0. \quad (14-12)$$

Note that it is only the *relative* velocity that counts. See Fig. 14-5.

If, at a given instant, the relative velocity forms a right angle with the line joining the source to the detector, there is still a Doppler effect because of time dilation. This is a purely relativistic effect. The frequency measured at the receiver is either larger or smaller than f_0 , depending on the reference frame in which the angle is 90° (Prob. 14-7):

$$f = \gamma f_0 \quad (\text{source}), \quad \text{or} \quad f = \frac{f_0}{\gamma} \quad (\text{detector}). \quad (14-13)$$

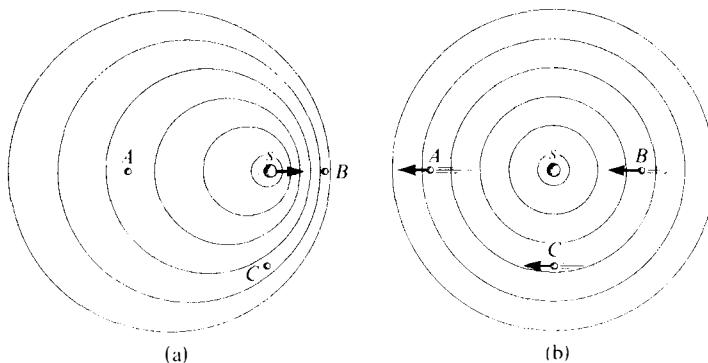


Fig. 14-5. The Doppler effect. (a) Source S moves to the right with respect to stationary receivers A , B , C . (b) The source is stationary, and the receivers all move to the left. The frequency shifts are the same in the two figures. The measured frequency is lower than the proper frequency at A , and higher at B . At C , the measured frequency is either lower or higher than the proper frequency, as in Prob. 14-7.

14.3 THE INCREMENT OF PROPER TIME FOR AN ACCELERATED PARTICLE

Consider a particle moving at some arbitrary time-dependent velocity. Say the particle moves in a straight line along the x -axis. Then its world line is some curve as in Fig. 14-6. An observer notes the position of the particle at times t and $t + dt$. This defines two points A and B , as in the figure.

The *proper time interval* dt_0 between A and B is defined as the time interval between these two events, as measured in the *unaccelerated frame S' occupied momentarily by the particle*:

$$dt_0 = \frac{dt}{\gamma} = dt \left(1 - \frac{v^2}{c^2}\right)^{1/2}, \quad (14-14)$$

as in Eq. 14-5, where v is the average velocity of the particle during that time interval.

We have simply applied the rule for time dilation that we found in Sec. 14.2 for *inertial* frames. This seems illogical because a reference frame attached to the particle does accelerate. This peculiar procedure follows from the experimental fact that the half-life of a fast particle of a given velocity is the same whether the trajectory is a straight line or a circle.

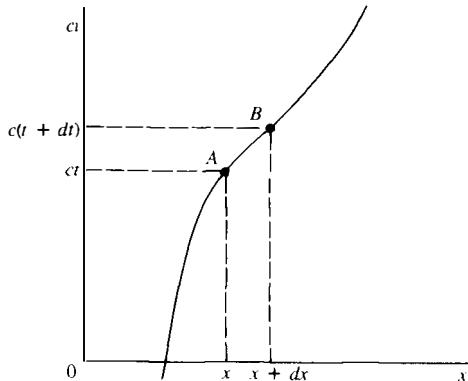


Fig. 14-6. The world line of a particle moving along the x -axis. In the fixed reference frame S , the time interval between events A and B is dt . In the frame S' of constant velocity \mathcal{V} occupied momentarily by the particle, the proper time interval is dt_0 , which is equal to dt/γ .

We use the increment of proper time repeatedly in the following chapters on relativity.

14.4 TRANSFORMATION OF A VELOCITY

Observer o' on reference frame S' notes that an object moves at some velocity v' . The velocity need not be constant. What is the velocity of this same object for observer o on reference frame S ?

From the vector form of the Lorentz transformation (Table 13-1),

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \frac{\gamma(d\mathbf{r}'_{||} + \mathcal{V}dt') + d\mathbf{r}'_{\perp}}{\gamma(dt' + \mathcal{V} \cdot d\mathbf{r}'/c^2)}. \quad (14-15)$$

Dividing numerator and denominator by $\gamma dt'$ gives

$$\mathbf{v} = \frac{\mathbf{v}'_{||} + \mathcal{V} + \mathbf{v}'_{\perp}/\gamma}{1 + v'_{||}\mathcal{V}/c^2}. \quad (14-16)$$

The Galilean relation $v = v' + \mathcal{V}$ is therefore valid only if $\gamma \approx 1$ and if $v'\mathcal{V} \ll c^2$. It follows from Sec. 13.4 that

$$\mathbf{v}' = \frac{\mathbf{v}_{||} - \mathcal{V} + \mathbf{v}_{\perp}/\gamma}{1 - v'_{||}\mathcal{V}/c^2}. \quad (14-17)$$

Example

TRANSFORMING THE SPEED OF LIGHT

A photon travels at the speed c in some arbitrary direction in reference frame S' . Then

$$v'^2_{\parallel} + v'^2_{\perp} = c^2 \quad (14-18)$$

and

$$v^2_{\parallel} + v^2_{\perp} = \left(\frac{v'_{\parallel} + \mathcal{V}}{1 + v'_{\parallel} \mathcal{V}/c^2} \right)^2 + \left(\frac{v'_{\perp}/\gamma}{1 + v'_{\parallel} \mathcal{V}/c^2} \right)^2. \quad (14-19)$$

Expanding and using Eq. 14-18 twice, we find that the speed of the photon in S is also c , whatever the value of \mathcal{V} !

This is in agreement with the first example in Sec. 13.4, where we discussed the invariance of $x^2 + y^2 + z^2 - c^2 t^2$.

14.5 SUMMARY

An object has a proper length l_0 , oriented in some arbitrary direction. Then, in another reference frame moving at some constant velocity \mathcal{V} with respect to the object,

$$l = \frac{l_{0\parallel}}{\gamma} + l_{0\perp}. \quad (14-2)$$

Lengths parallel to the motion are shorter by the factor γ . This is the *Lorentz contraction*. Lengths orthogonal to the motion are unaffected.

A process lasts a proper time T_0 in its own reference frame. In another frame as above, the time interval is γ times larger:

$$T = \gamma T_0. \quad (14-5)$$

This is *time dilation*.

If a source of electromagnetic waves of proper frequency f_0 moves away at a velocity $\mathcal{V} = \beta c$ from an observer, the apparent frequency is

$$f = \left(\frac{1 - \beta}{1 + \beta} \right)^{1/2} f_0 < f_0. \quad (14-11)$$

For an *approaching* source, the sign before β changes,

$$f = \left(\frac{1 + \beta}{1 - \beta} \right)^{1/2} f_0 > f_0. \quad (14-12)$$

This is the *Doppler effect*. Only the *relative* velocity between source and detector matters.

A particle describes some arbitrary motion in space-time. If the time interval between two successive positions is dt , then the corresponding *proper time interval* is

$$dt_0 = \frac{dt}{\gamma}. \quad (14-14)$$

A *velocity* transforms as follows:

$$\mathbf{v} = \frac{\mathbf{v}'_{||} + \mathcal{V} + \mathbf{v}'_{\perp}/\gamma}{1 + v'_{||}\mathcal{V}/c^2}, \quad (14-16)$$

$$\mathbf{v}' = \frac{\mathbf{v}_{||} - \mathcal{V} + \mathbf{v}_{\perp}/\gamma}{1 - v_{||}\mathcal{V}/c^2}. \quad (14-17)$$

PROBLEMS

14-1. (14.1) Lorentz contraction for a length

A one-meter ruler moves at a speed $c/2$. In its own reference frame it forms an angle of 45° with its velocity.

What is its length, as measured by a fixed observer?

14-2. (14.1) Two successive events at a given point.

Two events occur at the same place in the laboratory at an interval of 3 seconds.

What is the spatial distance between these two events in a moving frame with respect to which the events occur 5 seconds apart, and what is the relative speed of the moving and laboratory frames?

14-3. (14.1.1) Transformation of an element of area

A small rigid parallelogram of sides dl_{10} and dl_{20} , in its own reference frame, has an area

$$dA_0 = dl_{10} \times dl_{20}.$$

Show that, with respect to another reference frame, the element of area is given by

$$dA = dl_1 \times dl_2 = dA_{0||} + \frac{dA_{0\perp}}{\gamma}.$$

14-4. (14.2) The red shift

The radio galaxy 3C295 has a red shift of 46%. Astronomers mean by this that the observed wavelength is 1.46 times the wavelength of the same radiation originating in the laboratory.

(a) Calculate the radial velocity of the galaxy.

(b) Some quasars have red shifts of 200%. What is their radial velocity?

14-5. (14.2) The case of the speeding physicist

A physicist is arrested for going through a red traffic light. In court she pleads that she approached at such a speed that the red light appeared green. The judge, a graduate of a physics class, changes the charge to speeding and fines the defendant \$1 for every kilometer per hour she exceeded the speed limit of 50 kilometers/hour.

What is the fine ($\lambda_{\text{green}} \approx 5.3 \times 10^{-7}$ meter, $\lambda_{\text{red}} \approx 6.5 \times 10^{-7}$ meter)?

14-6. (14.2) The twin paradox

On his twenty-first birthday, Peter leaves his twin Paul behind on the earth and goes off in a straight line for 7 years of his time at a speed of $0.96c$. Peter then reverses direction and returns at the same speed.

(a) What are the ages of Peter and Paul at the moment of reunion?

(b) Peter and Paul, expecting a strange result, perform the following experiment during Peter's trip. They both observe a distant variable star whose light alternates from dim to bright at a frequency f when observed from the earth. The variable star is in a direction perpendicular to Peter's trajectory. They, of course, both count the same number of pulsations during the trip.

Use the expression for the Doppler shift to verify the difference in age between Peter and Paul at the end of the trip. See the next problem.

14-7. (14.2) Doppler effect for a source moving at $y = \text{constant}$

Figure 14-7 shows a source of electromagnetic radiation of proper frequency f_0 situated at the origin O' of reference frame S' moving at the velocity $V\hat{x}$ with respect to S . As usual, O' is at $x = 0$ at $t = 0$, $t' = 0$. We wish to calculate the frequency as measured by observer o situated at the origin O of S .

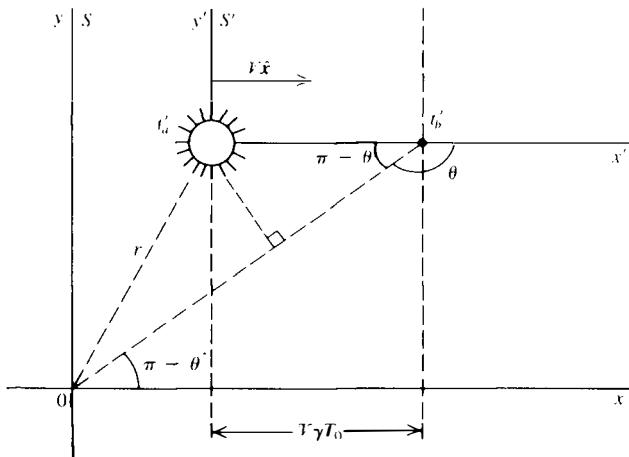


Fig. 14-7.

The figure shows S' at two successive beats of the source, separated by a time

$$t'_b - t'_a = \frac{1}{f_0} = T_0,$$

if $\gamma T_0 \ll r$.

- (a) Show that the period measured at O is $T = \gamma(1 - \beta \cos \theta)T_0$ and that

$$f = f_0 / [\gamma(1 - \beta \cos \theta)] = f' / [\gamma(1 - \beta \cos \theta)].$$

Thus, if $\theta = \pi/2$, then $f = f_0/\gamma$, $f = f'/\gamma$.

- (b) Compare this result with the Doppler effect calculated in Sec. 14.2.1 and in Prob. 14-6.

- (c) Show that

$$f' = \frac{f}{\gamma(1 + \beta \cos \theta')}.$$

Thus, if $\theta' = \pi/2$, then $f' = f/\gamma$ and $f = \gamma f'$.

- (d) From (a) and (c), $\gamma^2(1 - \beta \cos \theta)(1 + \beta \cos \theta') = 1$.

Check the validity of this equation at $\theta = \theta' = 0$ and π , at $\theta = \pi/2$, and at $\theta' = \pi/2$. Refer to Prob. 14-10 on the headlight effect.

14-8. (14.2) Transforming visible light to high-energy radiation

Visible light can be transformed into high-energy radiation by reflecting a laser beam backward on a high-energy electron beam. Say the initial photon energy is 2 electronvolts, and the electron energy is 6 gigaelectronvolts.

- (a) Calculate the photon energy $h\nu'$ in the reference frame S' of the electrons.

- (b) Now calculate the energy $h\nu''$ of the reflected photons in the laboratory frame.

As a first approximation, you can neglect the recoil of the electrons, but this gives too large a value for $h\nu''$.

14-9. (14.4) The speed of light in a moving medium

Light moves more slowly through a material medium than through a vacuum, its phase velocity v being c/n , where n is the index of refraction of the medium.

If now the medium itself moves at a velocity $\mathcal{V} \ll c$ with respect to the laboratory, show that the phase velocity of the light with respect to the laboratory is approximately $c/n + \mathcal{V}(1 - 1/n^2)$.

14-10. (14.4) The headlight effect

A source of light moves at a velocity $\mathcal{V}\hat{x}$. Consider a ray that forms an angle θ' with respect to the x -axis.

- (a) Show that in the reference frame S , $\tan \theta = \sin \theta' / [\gamma(\cos \theta' + \beta)]$. Then $\tan \theta' = \sin \theta / [\gamma(\cos \theta - \beta)]$, from Sec. 13.4.

- (b) Plot θ as a function of θ' for $\beta = 0, 0.5, 0.9, 0.9999$.

Observe that, for large values of β , θ is much smaller than θ' , except for values of θ' near π . If the source radiates isotropically in its own reference

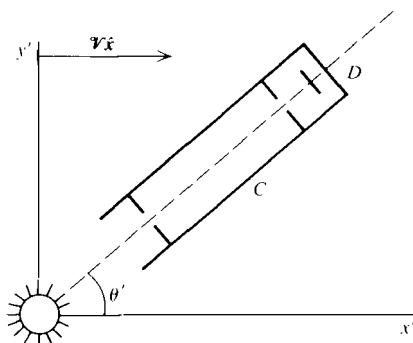


Fig. 14-8.

frame, then, for a stationary observer, the source radiates mostly in the forward direction. This is the *headlight effect*.

(c) An isotropic source of light moves at a speed $c/3$ with respect to an observer. Calculate the solid angle defined by a cone that points forward and that contains 25% of the total light flux.

14-11. (14.4) The collimator paradox

Figure 14-8 shows a source of light and a collimator C fixed in a reference frame S' that moves at a velocity $\mathcal{V}\hat{x}$ with respect to a fixed frame S . The detector D measures the light that goes through the collimator.

According to Prob. 14-10, the angle θ formed by the beam of light is less than θ' because of the headlight effect. However, the Lorentz contraction on the collimator makes its angle θ larger than θ' . But this is absurd! If light reaches D in S' , then it does so in S .

You can solve this paradox by using the Lorentz transformation.

14-12. (14.4) Three reference frames

We have three reference frames A, B, C . Frames B and C move, respectively, at velocities $\mathcal{V}\hat{x}/2$ and $\mathcal{V}\hat{x}$ with respect to A . Use subscripts to identify the velocities: $v_{BA} = \mathcal{V}/2$, $v_{CA} = \mathcal{V}$.

Show that v_{CB} (velocity of C with respect to B) is larger than $\mathcal{V}/2$. Thus, with respect to B , C moves away faster than A .

CHAPTER 15

*RELATIVITY III *Mass, Momentum, Force, and Energy*

15.1 THE FOUR-VECTOR \mathbf{r}	265
15.2 FOUR-VECTORS	267
15.2.1 THE SCALAR PRODUCT OF TWO FOUR-VECTORS	268
15.2.2 THE NORM OF THE FOUR-VECTOR \mathbf{r}	268
15.2.3 THE NORM OF A FOUR-VECTOR \mathbf{a}	269
15.3 THE RELATIVISTIC MASS m	270
15.4 THE RELATIVISTIC MOMENTUM \mathbf{p}	270
15.5 THE RELATIVISTIC FORCE \mathbf{F}	270
<i>Example</i>	271
15.6 THE FOUR-MOMENTUM \mathbf{p}	271
15.7 THE RELATIVISTIC ENERGY $\mathcal{E} = mc^2$	273
<i>Examples</i>	274
<i>Example:</i> THE RELATION $\mathcal{E}^2 = m_0^2c^4 + \mathbf{p}^2c^2$	275
15.8 KINETIC ENERGY	276
15.9 THE LAW OF CONSERVATION OF FOUR-MOMENTUM \mathbf{p}	276
15.10 TRANSFORMATION OF A FORCE	276
<i>Examples</i>	277
15.11 THE PHOTON	277
15.12 SUMMARY	278
PROBLEMS	280

There are five chapters on relativity; the first three provide the basic ideas that are prerequisite for the other two. This third chapter concerns mechanics. As you will see, relativity makes havoc of mechanics. All its basic concepts, such as mass, momentum, force, and energy, crumble, and classical mechanics sinks to the rank of an approximate theory. It is a curious fact that classical electromagnetism, by contrast, is completely compatible with relativity and is thus spared.

15.1 THE FOUR-VECTOR \mathbf{r}

An event E , such as the collision between two particles or the emission of a photon, occurs at a given point (x, y, z) and at a given time t . The event

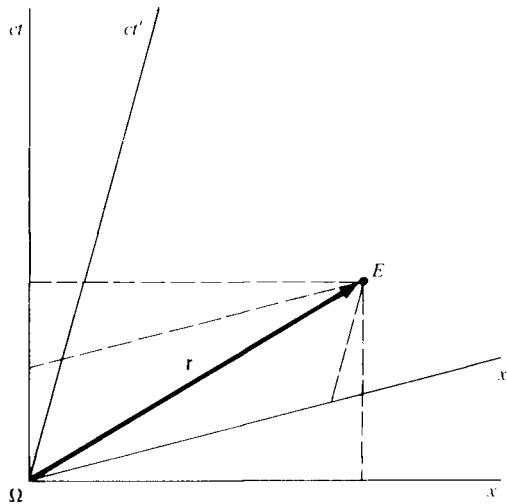


Fig. 15-1. Event E occurs at $(x, 0, 0, ct)$ in S , or at $(x', 0, 0, ct')$ in S' . The four-vector \mathbf{r} defines the coordinates of E with respect to the event Ω that occurs at $(0, 0, 0, 0)$.

is thus said to occur at the *space-time coordinates* (x, y, z, ct) . These are the coordinates of event E with respect to the datum event Ω that occurs at the origin $(0, 0, 0, 0)$.

We can thus imagine a *four-vector* \mathbf{r} pointing from Ω to E in space-time. In particular, if $y = 0$ and $z = 0$, we can draw \mathbf{r} as in Fig. 15-1. We denote four-vectors by means of boldface sans-serif type: \mathbf{r} . It is the custom to write out the components of \mathbf{r} in one of two forms:

$$\mathbf{r} = (x, y, z, ct) = (\mathbf{r}, ct). \quad (15-1)$$

The four-vector \mathbf{r} transforms as in Table 15-1, because of the Lorentz transformation (Sec. 13.4).

Observe that the four-vector \mathbf{r} relates one event to another. It thus possesses its own identity, irrespective of the choice of reference frame:

Table 15-1 Transformation of the four-vector $\mathbf{r} = (\mathbf{r}, ct)$

$\mathbf{r} = \gamma(\mathbf{r}'_{\parallel} + \mathcal{V}t') + \mathbf{r}'_{\perp}$	$\mathbf{r}' = \gamma(\mathbf{r}_{\parallel} - \mathcal{V}t) + \mathbf{r}_{\perp}$
$ct = \gamma\left(ct' + \frac{\mathcal{V}'\mathbf{r}'_{\parallel}}{c}\right)$	$ct' = \gamma\left(ct - \frac{\mathcal{V}\mathbf{r}_{\parallel}}{c}\right)$

$$\mathbf{r} = (x, y, z, ct)_S = (x', y', z', ct')_{S'} = \dots, \quad (15-2)$$

where x, y, z, ct and x', y', z', ct' satisfy the Lorentz transformation. This four-vector is the *position vector in space-time*.

This is not a new situation. For example, say point O in ordinary three-dimensional space is the origin, and point A is 1 meter above O . The vector \mathbf{r} has a magnitude of unity and points upward. But the components of \mathbf{r} depend on the orientation of the coordinate axes.

In three dimensions, a vector \mathbf{r} can serve to relate a point P_2 to a point P_1 , instead of to the origin:

$$\mathbf{r}_{12} = (x_2 - x_1)\hat{\mathbf{x}} + (y_2 - y_1)\hat{\mathbf{y}} + (z_2 - z_1)\hat{\mathbf{z}}. \quad (15-3)$$

Similarly, the four-vector

$$\mathbf{r}_{12} = \mathbf{r}_2 - \mathbf{r}_1 = [(\mathbf{r}_2 - \mathbf{r}_1), c(t_2 - t_1)] \quad (15-4)$$

relates event 2 to event 1.

15.2 FOUR-VECTORS

In three dimensions, the vector $\mathbf{r}(\hat{x}, y, z)$ denotes the position of the point (x, y, z) with respect to the origin of coordinates.

Other vector quantities in three dimensions, such as velocities, accelerations, forces, etc., transform in the same way as the position vector \mathbf{r} . Indeed, any three quantities that transform as the components of the position vector are the components of a three-dimensional vector, by definition.

Similarly, in four dimensions, any four quantities a_1, a_2, a_3, a_4 that transform as the components of the position vector in space-time are the components of a four-vector

$$\mathbf{a} = (a_1, a_2, a_3, a_4) = (\mathbf{a}, a_4). \quad (15-5)$$

Table 15-2 states the rules for transforming the components of a four-vector.

Table 15-2 Transformation of a four-vector $\mathbf{a} = (\mathbf{a}, a_4)$

$\mathbf{a} = \gamma(\mathbf{a}'_{ } + \mathcal{V}t') + \mathbf{a}'_{\perp}$	$\mathbf{a}' = \gamma(\mathbf{a}_{ } - \mathcal{V}t) + \mathbf{a}_{\perp}$
$a_4 = \gamma\left(a'_4 + \frac{\mathcal{V}'a'_{ }}{c}\right)$	$a'_4 = \gamma\left(a_4 - \frac{\mathcal{V}a_{ }}{c}\right)$

15.2.1 The Scalar Product of Two Four-Vectors

Given the two four-vectors

$$\mathbf{a} = (a_1, a_2, a_3, a_4) \quad \mathbf{b} = (b_1, b_2, b_3, b_4), \quad (15-6)$$

their *scalar product* is

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 - a_4 b_4. \quad (15-7)$$

Note the negative sign before the last term on the right. Otherwise, this product is similar to the scalar product of two vectors.

For example, if

$$\mathbf{r} = (x, y, z, ct) = (\mathbf{r}, ct) \quad (15-8)$$

and

$$\mathbf{R} = (X, Y, Z, cT) = (\mathbf{R}, cT), \quad (15-9)$$

then

$$\mathbf{r} \cdot \mathbf{R} = xX + yY + zZ - c^2tT = \mathbf{r} \cdot \mathbf{R} - c^2tT. \quad (15-10)$$

The scalar product of two four-vectors is an invariant. This is a consequence of the Lorentz transformation.

15.2.2 The Norm of the Four-Vector \mathbf{r}

The quantity

$$|\mathbf{r}| = |\mathbf{r} \cdot \mathbf{r}|^{1/2} = |x^2 + y^2 + z^2 - c^2t^2|^{1/2} = |\mathbf{r}^2 - c^2t^2|^{1/2} \quad (15-11)$$

is the *norm* of the four-vector \mathbf{r} , by analogy with ordinary vectors. Observe again the *minus* sign. This is not Euclidean geometry! Then observe the *absolute value bars*; by definition, the norm of a four-vector is real and positive.

The norm of a four-vector is an invariant, like a scalar product.

Let us pause a bit to reflect on the meaning of the norm of the four-vector \mathbf{r}_E that defines the position of an event E in space-time with respect to a given reference frame S . We set $y_E = 0$, $z_E = 0$, so as to be able to show \mathbf{r}_E on a space-time diagram as in Fig. 15-2(a) and (b). Then

$$|\mathbf{r}_E| = |x_E^2 - c^2t_E^2|^{1/2}. \quad (15-12)$$

Clearly, the norm of \mathbf{r}_E is *not* its length.

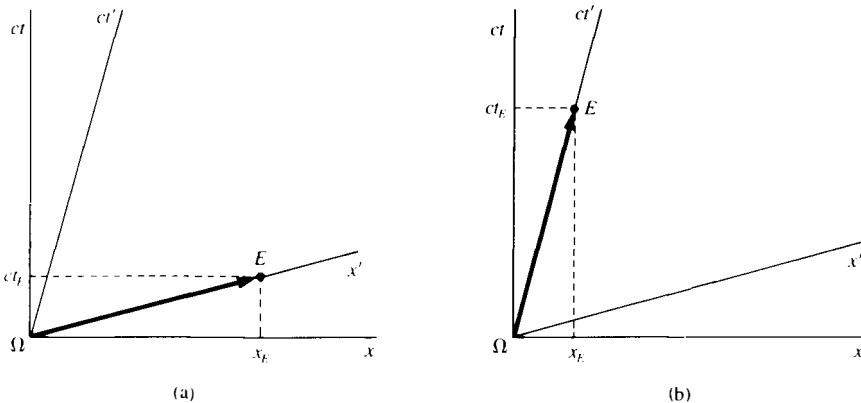


Fig. 15-2. (a) Event E occurs at $y_E = 0$, $z_E = 0$, and $x_E > ct_E$. It occurs at $t' = 0$ in a certain frame S' . (b) Event E occurs at $y_E = 0$, $z_E = 0$, and $x_E < ct_E$. It occurs at $x' = 0$ in some other frame S' .

Let us set, successively, $x_E > ct_E$, $x_E < ct_E$, and $x_E = ct_E$. First, we choose an event E such that $x_E > ct_E$, as in Fig. 15-2(a). Then we can imagine axes x' and ct' of a reference frame S' in which E occurs at $t'_E = 0$. In S' , $|\mathbf{r}_E| = x'_E$. This means that if $x_E > ct_E$, then the norm of \mathbf{r}_E is equal to the spatial distance between events Ω and E in a reference frame with respect to which Ω and E are simultaneous. This result applies even if $y_E \neq 0$, $z_E \neq 0$.

Now suppose that $x_E < ct_E$, as in Fig. 15-2(b). We now select axes x' and ct' so that E occurs at $x' = 0$. Then, in S' , event E occurs at the time

$$t'_E = \frac{|\mathbf{r}_E|}{c}. \quad (15-13)$$

Therefore, if $x_E < ct_E$, the norm of \mathbf{r}_E divided by c is equal to the time interval between Ω and E in a frame S' where both events occur at the same place. This is the proper time interval between Ω and E . Remember that we have set $y_E = 0$ and $z_E = 0$.

If now $x_E = ct_E$, then $|\mathbf{r}_E| = 0$, despite the fact that $\mathbf{r}_E \neq 0$. For example, suppose a photon travels along \mathbf{r}_E in reference frame S . Then Ω and E represent two points on the world line (Sec. 13.5) of a photon with $x_E = ct_E$, and there does not exist a reference frame with respect to which Ω and E take place either at the same time or at the same x .

15.2.3 The Norm of a Four-Vector \mathbf{a}

The norm of an arbitrary four-vector \mathbf{a} is defined like that of \mathbf{r} :

$$|\mathbf{a}| = |\mathbf{a} \cdot \mathbf{a}|^{1/2} = |a_1^2 + a_2^2 + a_3^2 - a_4^2|^{1/2}. \quad (15-14)$$

Since all four-vectors transform as \mathbf{r} , it follows that the norm of a four-vector is invariant.

15.3 THE RELATIVISTIC MASS m

In its own reference frame, an object has a *rest mass* m_0 . If the object moves at a velocity v with respect to an observer, then, for *that* observer,

$$m = \frac{m_0}{(1 - v^2/c^2)^{1/2}} = \gamma m_0 > m_0. \quad (15-15)$$

We take this fairly well-known result for granted.[†] The quantity m is the *relativistic mass*. All mass measurements are in agreement with the above equation. For example, the mass of a high-velocity electron in an accelerator is γ times that of a slow electron.

The relativistic mass tends to infinity as v approaches c . Then what about the photon? It has a relativistic mass m , but a (presumably) zero rest mass m_0 . See Sec. 15.11.

We shall see how to transform a relativistic mass in Sec. 15.6 (Table 15-3).

15.4 THE RELATIVISTIC MOMENTUM p

The *relativistic momentum* of a mass m moving at a velocity \mathbf{v} is defined as in classical mechanics, except that m is the relativistic mass:

$$\mathbf{p} = m\mathbf{v} = \frac{m_0\mathbf{v}}{(1 - v^2/c^2)^{1/2}} = \gamma m_0\mathbf{v}. \quad (15-16)$$

It is this quantity that is conserved in collisions, and not the momentum $m_0\mathbf{v}$ of classical physics.

We shall find how to transform a relativistic momentum in Sec. 15.6 (Table 15-3).

15.5 THE RELATIVISTIC FORCE F

The *relativistic force* is also defined as in classical mechanics:

[†] See, for example, David Bohm, *The Special Theory of Relativity*, W. A. Benjamin, New York, 1965, p. 84.

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}, \quad (15-17)$$

where \mathbf{p} is now the relativistic momentum. We shall find the transformation equations for \mathbf{F} later, in Sec. 15.10.

Example A high-energy particle of rest mass m_0 , velocity \mathbf{v} , and charge Q crosses a region where the electric field strength is \mathbf{E} . The electric force is

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{d}{dt}(m\mathbf{v}) = \frac{d}{dt}\left[\frac{m_0}{(1 - v^2/c^2)^{1/2}}\mathbf{v}\right] = QE. \quad (15-18)$$

15.6 THE FOUR-MOMENTUM \mathbf{p}

Consider a particle of rest mass m_0 and velocity \mathbf{v} at

$$\mathbf{r} = (x, y, z, ct) = (\mathbf{r}, ct) \quad (15-19)$$

with respect to a reference frame S , as in Fig. 15-3.

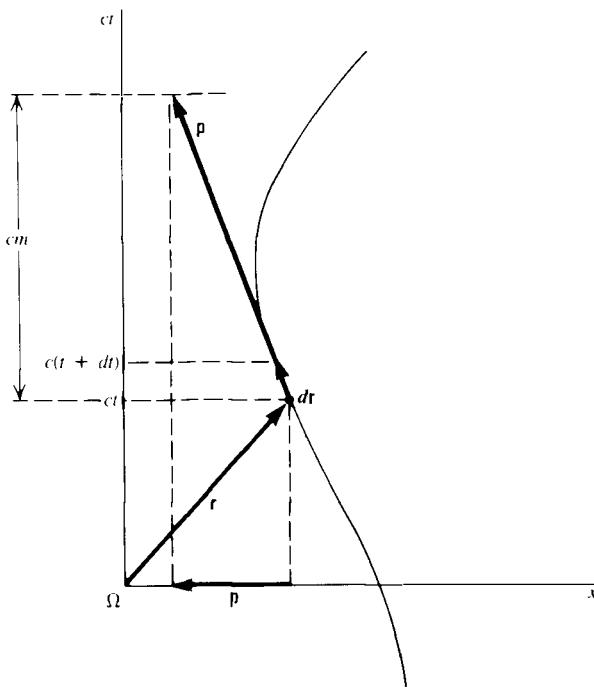


Fig. 15-3. World line of a particle moving in the (x, ct) plane. The four-momentum \mathbf{p} is along $d\mathbf{r}$, and is thus tangent.

Over the time interval dt , as measured in S ,

$$\mathbf{dr} = (\mathbf{dr}, c dt). \quad (15-20)$$

This increment \mathbf{dr} of \mathbf{r} is tangent to the world line of the particle. Now

$$|\mathbf{dr}| = |dr^2 - c^2 dt^2|^{1/2} = c dt \left| 1 - \frac{1}{c^2} \left(\frac{dr}{dt} \right)^2 \right|^{1/2} \quad (15-21)$$

$$= c dt \left| 1 - \frac{v^2}{c^2} \right|^{1/2} = \frac{c dt}{\gamma} = c dt_0, \quad (15-22)$$

from Sec. 14.2, where dt_0 is the increment of proper time *in the inertial frame occupied momentarily by the particle*. The velocity \mathbf{v} need not be constant.

For a given dt_0 , say 1 second, the corresponding dt depends on the reference frame of the observer, but dt_0 is the same for all observers, and hence dt_0 is an invariant.

The *four-momentum* of the particle is defined as follows:

$$\mathbf{p} = m_0 \frac{\mathbf{dr}}{dt_0}. \quad (15-23)$$

This is a four-vector because \mathbf{dr} is a four-vector, while both m_0 and dt_0 are invariants. The four-momentum is tangent to the world line of the particle.

To find the components of \mathbf{p} , we write out the components of \mathbf{r} and differentiate:

$$\mathbf{p} = m_0 \frac{\mathbf{dr}}{dt_0} = m_0 \frac{d}{dt_0} (\mathbf{r}, ct) = \left(m_0 \frac{\mathbf{dr}}{dt} \frac{dt}{dt_0}, m_0 c \frac{dt}{dt_0} \right). \quad (15-24)$$

But, from Sec. 14.2,

$$\frac{dt}{dt_0} = \gamma = \frac{1}{(1 - v^2/c^2)^{1/2}}. \quad (15-25)$$

Thus

$$\mathbf{p} = (m_0 \gamma \mathbf{v}, m_0 \gamma c) = (m \mathbf{v}, mc) = m(\mathbf{v}, c) = (\mathbf{p}, mc). \quad (15-26)$$

The norm of \mathbf{p} has a peculiar value:

$$|\mathbf{p}| = |\mathbf{p} \cdot \mathbf{p}|^{1/2} = m |v^2 - c^2|^{1/2} = mc \left| 1 - \frac{v^2}{c^2} \right|^{1/2} = m_0 c. \quad (15-27)$$

Table 15-3 Transformation of a four-momentum $\mathbf{P} = (\mathbf{p}, mc)$

$\mathbf{p} = \gamma(\mathbf{p}'_{ } + m'\mathcal{V}) + \mathbf{p}'_{\perp}$	$\mathbf{p}' = \gamma(\mathbf{p}_{ } - m\mathcal{V}) + \mathbf{p}_{\perp}$
$mc = \gamma\left(m'c + \frac{\mathcal{V}p'_{ }}{c}\right)$	$m'c = \gamma\left(mc - \frac{\mathcal{V}p_{ }}{c}\right)$
$m = \gamma m'\left(1 + \frac{\mathcal{V}v'_{ }}{c^2}\right)$	$m' = \gamma m\left(1 - \frac{\mathcal{V}v_{ }}{c^2}\right)$

So the norm of the four-momentum of a given object is independent of the velocity! It is an invariant.

Of course, a particle that is stationary in space is not stationary in space-time: it travels parallel to the ct -axis at the speed of light.

As to the photon, see Sec. 15.11.

The four-momentum has another peculiar property. Since $\mathbf{p} \cdot \mathbf{p}$ is invariant,

$$d(\mathbf{p} \cdot \mathbf{p}) = 2\mathbf{p} \cdot d\mathbf{p} = 0. \quad (15-28)$$

This means that $d\mathbf{p}$ is “orthogonal” to \mathbf{p} . But \mathbf{p} is “parallel” to $d\mathbf{r}$, from Eq. 15-23. Thus

$$d\mathbf{r} \cdot d\mathbf{p} = 0, \quad (15-29)$$

and $d\mathbf{p}$ is “orthogonal” to $d\mathbf{r}$, in agreement with the fact that \mathbf{p} is invariant.

To summarize, the four-momentum \mathbf{p} of a particle is a four-vector (1) that is tangent to the world line, (2) whose space component is \mathbf{p} and whose time component is mc , (3) whose norm is m_0c , and (4) whose differential $d\mathbf{p}$ is “orthogonal” to both \mathbf{p} and $d\mathbf{r}$.

Since \mathbf{p} is a four-vector, it transforms like \mathbf{r} , as in Table 15-3.

If v^2 , v'^2 , \mathcal{V}^2 are all negligible compared to c^2 , then we revert to classical mechanics:

$$\mathbf{p} = \mathbf{p}' + m'\mathcal{V}, \quad m = m'. \quad (15-30)$$

15.7 THE RELATIVISTIC ENERGY $\mathcal{E} = mc^2$

We have just seen that $d\mathbf{r} \cdot d\mathbf{p} = 0$. Thus

$$d\mathbf{r} \cdot d\mathbf{p} = (d\mathbf{r}, c dt) \cdot (d\mathbf{p}, c dm) = d\mathbf{r} \cdot d\mathbf{p} - c^2 dt dm = 0, \quad (15-31)$$

or

$$d\mathbf{r} \cdot \frac{d\mathbf{p}}{dt} = c^2 dm. \quad (15-32)$$

Now we are concerned with an object of relativistic mass m , velocity \mathbf{v} , and relativistic momentum $\mathbf{p} = m\mathbf{v}$. So $d\mathbf{p}/dt$ is the force \mathbf{F} applied to m , and the term on the left is the energy expended by \mathbf{F} on m over the distance $d\mathbf{r}$. An increment of mass dm accompanies this increment of energy.

More generally, in any physical process, an increase of energy $d\mathcal{E}$ results in an increase of mass dm such that

$$d\mathcal{E} = c^2 dm, \quad (15-33)$$

and

$$\mathcal{E} = mc^2 \quad (15-34)$$

is the *relativistic energy* of an object of mass m .

If a force \mathbf{F} acts on a mass m moving at a velocity \mathbf{v} ,

$$\mathbf{F} \cdot \mathbf{v} = \frac{d(mc^2)}{dt} = \frac{d\mathcal{E}}{dt}. \quad (15-35)$$

A particle at rest possesses a *rest energy* m_0c^2 .

A relativistic energy transforms like a mass (Table 15-3).

Examples

Say a mixture of hydrogen and oxygen explodes. What happens to the mass? Before the reaction, the molecules may be assumed to be at rest. Let their total rest mass be m_{0b} . After the reaction, the relativistic mass is unaltered because no *external* energy has been fed into the gas. Then

$$m_a = m_{0b} = m_{0a} + KE, \quad (15-36)$$

where KE is the kinetic energy of the high-temperature steam. The new *rest* mass is smaller than the initial *rest* mass.

After a while, the steam cools and becomes water at room temperature, KE tends to zero, and m_a decreases to m_{0a} . In the process the kinetic energy has spread out to neighboring bodies, thereby increasing their masses. The mass of the water is less than that of the original mixture, but the relativistic mass of the universe is unchanged.

Fission and fusion reactions are qualitatively similar to exothermic chemical reactions: there is a loss of rest mass and a release of thermal energy.

What happens when you drop a brick? The rest mass of a brick depends on its position, because of its potential energy: the higher it is, the larger is its rest mass. As the brick falls, its rest mass decreases and its kinetic energy increases. When the brick hits the ground, its kinetic energy becomes thermal energy. It loses rest mass, but, again, the relativistic mass of the universe is

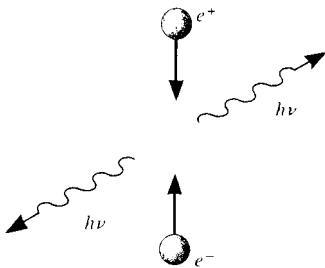


Fig. 15-4. A positive and a negative electron annihilate to form two photons. If the kinetic energies of the electrons are low, each photon has an energy $h\nu$ equal to the rest energy of an electron, 5.11×10^5 electronvolts.

unchanged. Strictly, one should, of course, think in terms of the potential energy of the brick-earth system and take into account the upward motion of the earth as the brick falls.

When a positron meets an electron, both particles disappear to form a pair of photons, as in Fig. 15-4. Charge, relativistic mass, energy, and momentum are all conserved.

Inversely, a photon can create a positron-electron pair. There is conservation of charge, but the conservation of mass, energy, and momentum requires the presence of another particle.

Example

$$\text{THE RELATION } \mathcal{E}^2 = m_0^2 c^4 + p^2 c^2$$

An object of rest mass m_0 lies at rest in a reference frame S' . With respect to another reference frame S , its mass is m and its momentum is \mathbf{p} . Since the norm of the four-momentum is invariant,

$$p^2 - m^2 c^2 = -m_0^2 c^2. \quad (15-37)$$

Therefore

$$\mathcal{E}^2 = m^2 c^4 = m_0^2 c^4 + p^2 c^2, \quad (15-38)$$

as in Fig. 15-5.

The term $p^2 c^2$ is negligible when $\gamma^2 \beta^2 \ll 1$, or when $2\beta^2 \ll 1$. Then $\mathcal{E} \approx m_0 c^2$. On the other hand, if $\gamma^2 \beta^2 \gg 1$ or if $\beta^2 \approx 1$ and $v \approx c$, then

$$\mathcal{E} = mc^2 \approx pc. \quad (15-39)$$

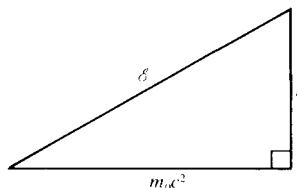


Fig. 15-5. The relativistic energy \mathcal{E} is related to the rest energy $m_0 c^2$ and to the relativistic momentum \mathbf{p} as in this right-angled triangle.

15.8 KINETIC ENERGY

By definition, the *kinetic energy* of a point mass m moving at a velocity v with respect to a given reference frame is equal to the energy expended in increasing its velocity from zero to v in that frame. This is

$$mc^2 - m_0 c^2 = m_0 c^2 \left[\frac{1}{(1 - v^2/c^2)^{1/2}} - 1 \right] = m_0 c^2 \left(\frac{v^2}{2c^2} + \frac{3v^4}{8c^4} + \dots \right) \quad (15-40)$$

$$= \frac{m_0 v^2}{2} + \frac{3m_0 v^4}{8c^2} + \dots \quad (15-41)$$

The first term on the right is the kinetic energy of classical mechanics. It is equal to the relativistic kinetic energy if $v^2 \ll c^2$.

15.9 THE LAW OF CONSERVATION OF FOUR-MOMENTUM \mathbf{p}

Say two particles interact, yielding two or three other particles. There is conservation of the total four-momentum

$$\sum \mathbf{p} = \left(\sum \mathbf{p}, c \sum m \right) \quad (15-42)$$

for the simple reason that both $\sum \mathbf{p}$ and $\sum m$ are conserved. So the conservation of four-momentum groups the conservation of momentum, of mass, and of energy!

The conservation and the invariance of the four-momentum imply that if \mathbf{p} and m are conserved in one inertial reference frame, then they are conserved in any other such frame.

15.10 TRANSFORMATION OF A FORCE

A force (Sec. 15.5) transforms as follows. From Tables 15-3 and 13-1,

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \frac{\gamma(d\mathbf{p}'_{||} + \mathcal{V} d\mathcal{E}'/c^2) + d\mathbf{p}'_{\perp}}{\gamma(dt' + \mathcal{V} dr'_{||}/c^2)}, \quad (15-43)$$

Dividing above and below by $\gamma dt'$ yields

$$\mathbf{F} = \frac{\mathbf{F}'_{||} + (\mathcal{V}/c^2) d\mathcal{E}'/dt' + \mathbf{F}'_{\perp}/\gamma}{1 + v_{||}\mathcal{V}/c^2}, \quad (15-44)$$

Table 15-4 Transformation of a force

$$\mathbf{F} = \mathbf{F}'_{\parallel} + \frac{\gamma(\mathcal{V}/c^2)\mathbf{F}'_{\perp} \cdot \mathbf{v}'_{\perp} + \mathbf{F}'_{\perp}}{\gamma(1 + v'_{\parallel}\mathcal{V}/c^2)} \quad \mathbf{F}' = \mathbf{F}_{\parallel} + \frac{-\gamma(\mathcal{V}/c^2)\mathbf{F}_{\perp} \cdot \mathbf{v}_{\perp} + \mathbf{F}_{\perp}}{\gamma(1 - v_{\parallel}\mathcal{V}/c^2)}$$

where \mathbf{v}' is the velocity of the point of application of \mathbf{F}' in reference frame S' .

Now $d\mathcal{E}'/dt'$ is the rate at which the relativistic energy builds up in S' under the action of \mathbf{F}' . Then

$$\frac{d\mathcal{E}'}{dt'} = \mathbf{F}' \cdot \mathbf{v}' = \mathbf{F}'_{\parallel} v'_{\parallel} + \mathbf{F}'_{\perp} \cdot \mathbf{v}'_{\perp}. \quad (15-45)$$

Substituting into Eq. 15-44 and simplifying leads to the transformation equations of Table 15-4. The transformation does *not* involve the coordinates of the point of application of the force.

Note that $\mathbf{F}'_{\perp} \neq \mathbf{F}_{\perp}$. Since a four-vector has the same perpendicular component in all inertial frames, the relativistic force is *not* the spatial component of a four-vector.

Examples

If \mathbf{F} is parallel to \mathcal{V} , then $\mathbf{F} = \mathbf{F}'$.	Two forces that are equal and opposite in one frame are not necessarily so in another frame. They remain equal and opposite only if their points of application have equal <i>velocities</i> .
--	--

15.11 THE PHOTON

The photon has presumably a zero rest mass. Its speed[†] is c , and its energy is

$$\mathcal{E} = h\nu, \quad (15-46)$$

where h is Planck's constant, 6.626×10^{-34} joule-second, and ν is the frequency of the associated wave. From Sec. 15.7, its mass is

$$m = \frac{\mathcal{E}}{c^2}, \quad (15-47)$$

[†] Photons always travel at the speed c , even inside matter. See Sec. 37.4. For the moment, we can think of photons traveling in a vacuum.

and its momentum is

$$p = mc = \frac{\mathcal{E}}{c} = \frac{h\nu}{c} = \frac{h}{\lambda}, \quad (15-48)$$

where λ is the wavelength of the associated wave. The relation between p and \mathcal{E} for a photon has been verified by measuring both the radiation pressure and the energy flux of a light beam.

The four-momentum of a photon is

$$\mathbf{p} = (p, mc) = (mc, mc). \quad (15-49)$$

Its spatial component is mc , where the vector \mathbf{c} points in the direction of propagation. Also,

$$|\mathbf{p}| = |mc^2 - mc^2|^{1/2} = 0. \quad (15-50)$$

According to Eq. 15-27, this implies that the photon has a zero rest mass.

Also, from Sec. 15.6,

$$|\mathbf{p}| = m \frac{|\mathbf{dr}|}{dt} = 0, \quad (15-51)$$

and $|\mathbf{dr}| = 0$. The increment of proper time (Sec. 14.3) over \mathbf{dr} is zero:

$$dt_0 = \frac{|\mathbf{dr}|}{c} = 0. \quad (15-52)$$

A photon's clock always reads the same time!

The kinetic energy of a photon is $(m - m_0)c^2 = mc^2$, and *not* $mc^2/2$, as we would expect from classical mechanics. Although a given photon has a speed c for all observers, its energy \mathcal{E} and its momentum are not invariant, because of the Doppler effect. For example, if S' moves in the same direction as the photon, then $\nu' < \nu$, and thus

$$h\nu' < h\nu, \quad \mathcal{E}' < \mathcal{E}, \quad p' < p. \quad (15-53)$$

15.12 SUMMARY

A *four-vector* has four components, the first three of which are the space components of a three-dimensional vector. One specifies the location of an event in space-time with respect to the datum event $\Omega(0, 0, 0, 0)$ by means of the four-vector

$$\mathbf{r} = (x, y, z, ct) = (\mathbf{r}, ct). \quad (15-1)$$

The four-vector

$$\mathbf{r}_{12} = \mathbf{r}_2 - \mathbf{r}_1 = [(r_2 - r_1), c(t_2 - t_1)] \quad (15-4)$$

relates an event E_2 with respect to another event E_1 . A four-vector transforms as in Tables 15-1 and 15-2. Any set of four quantities that transform as the components of a four-vector are the components of a four-vector.

The *scalar product* of two four-vectors \mathbf{a} and \mathbf{b} is

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 - a_4 b_4. \quad (15-7)$$

This product is invariant.

The *norm of a four-vector* is also invariant:

$$|\mathbf{a}| = |\mathbf{a} \cdot \mathbf{a}|^{1/2} = |a_1^2 + a_2^2 + a_3^2 - a_4^2|^{1/2}. \quad (15-14)$$

The *mass* of an object depends on its velocity with respect to the observer:

$$m = \frac{m_0}{(1 - v^2/c^2)^{1/2}} = \gamma m_0 > m_0. \quad (15-15)$$

This quantity is also called the *relativistic mass*.

The *relativistic momentum* of a point mass is

$$\mathbf{p} = m\mathbf{v} = \gamma m_0 \mathbf{v}, \quad (15-16)$$

and the *relativistic force* is

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}. \quad (15-17)$$

The *four-momentum* of a particle is the four-vector

$$\mathbf{p} = m_0 \frac{d\mathbf{r}}{dt_0} = (\mathbf{p}, mc) \quad (15-23), (15-26)$$

whose norm is $m_0 c$.

Four-momenta and masses transform as in Table 15-3.

The *relativistic energy* is

$$\mathcal{E} = mc^2. \quad (15-34)$$

The *rest energy* is m_0c^2 . *Kinetic energy* is relativistic energy minus rest energy, $mc^2 - m_0c^2$.

In any interaction there is conservation of the total four-momentum \mathbf{p} , and hence conservation of momentum, mass, and energy.

Forces transform as in Table 15-4.

For the photon, the rest mass is presumably zero and

$$m = \frac{\mathcal{E}}{c^2}, \quad (15-47)$$

$$p = mc = \frac{\mathcal{E}}{c} = \frac{h\nu}{c} = \frac{h}{\lambda}, \quad (15-48)$$

where h is Planck's constant, 6.626×10^{-34} joule-second, ν is the frequency, and λ the wavelength of the associated wave.

PROBLEMS

15-1. (15.3) Burning hydrogen

You ignite a mixture of hydrogen and oxygen inside a closed vessel, and then allow the water vapor to cool.

Sketch graphs of mc^2 and of m_0c^2 as functions of time.

15-2. (15.3) Relativistic effects with 40-GeV electrons

A linear accelerator accelerates electrons up to energies of 40 gigaelectronvolts (40×10^9 electronvolts).

(a) Calculate the mass of an electron that has the full energy. How does this mass compare with that of a proton at rest? See the page facing the back cover.

(b) What is the length of the accelerator in the reference frame of an electron that has the full energy? The length of the accelerator, as measured on the ground, is 3000 meters.

(c) How much time would such an electron take to go from one end of the accelerator to the other (i) in the laboratory frame and (ii) in the electron's frame of reference?

15-3. (15.6) Transformation of a mass density

We use the symbol τ for a volume in this problem.

A small element in an object has a proper mass dm_0 , a proper volume $d\tau_0$, and a proper mass density $\rho_0 = dm_0/d\tau_0$. With respect to S and S' the mass densities are $\rho = dm/d\tau$ and $\rho' = dm'/d\tau'$, respectively, and the velocities of the element are v and v' .

Show that, if \mathcal{V} is the speed of S' with respect to S , then

PROBLEMS

$$(a) \ dm = \gamma \left(1 + v'_x \frac{\mathcal{V}}{c^2}\right) dm', \quad (b) \ d\tau = \frac{d\tau'}{\gamma(1 + v'_x \mathcal{V}/c^2)},$$

$$(c) \ \rho = \gamma^2 \left(1 + \frac{v'_x \mathcal{V}}{c^2}\right)^2 \rho', \quad \rho' = \gamma^2 \left(1 - \frac{v_x \mathcal{V}}{c^2}\right)^2 \rho.$$

15-4. (15.5) The relativistic force

In classical mechanics, $\mathbf{F} = m\mathbf{a}$ if the mass is constant.
Show that with relativity,

$$\mathbf{F} = \frac{m_0}{(1 - v^2/c^2)^{3/2}} \mathbf{a}_{\parallel} + \frac{m_0}{(1 - v^2/c^2)^{1/2}} \mathbf{a}_{\perp} = \gamma^2 m \mathbf{a}_{\parallel} + m \mathbf{a}_{\perp},$$

where \mathbf{a}_{\parallel} and \mathbf{a}_{\perp} are the components of the acceleration that are, respectively, parallel and perpendicular to the velocity \mathbf{v} of the point of application of the force.

A force that is perpendicular to \mathbf{v} changes the direction of \mathbf{v} , but not the mass, and hence not the speed, as we could expect because the force does no work. Then, if \mathbf{F} is perpendicular to \mathbf{v} , $\mathbf{F} = m\mathbf{a}$ applies!

However, a force parallel to \mathbf{v} changes the magnitude of \mathbf{v} and hence the mass also. The resistance to acceleration is larger because of the γ^2 term.

The quantity m is sometimes called the *transverse inertial mass*, and $\gamma^2 m$ the *longitudinal inertial mass*.

15-5. (15.7) The gravitational red shift

A photon of energy $h\nu_0$ leaves the surface of a star of radius R and mass M .

(a) Show that, after the photon has escaped to infinity, $\Delta\nu/\nu_0$ is equal to $GM/(Rc^2)$, where G is the gravitational constant. This change is so small that you can set $v = v_0$ in your calculation of the change in potential energy. What is the sign of $\Delta\nu$? This change of frequency is the *gravitational red shift*.

(b) Calculate $\Delta\nu/\nu$ for the sun and for the earth. See the page facing the back cover.

(c) Calculate $\Delta\nu/\nu$ for a photon that travels from the surface of the sun to the surface of the earth, taking into account both gravitational fields.

(d) Sirius and a smaller star revolve around each other. The mass of the smaller star is about equal to that of the sun, but its light has a $\Delta\nu/\nu$ of 7×10^{-4} . What is its average density?

(e) The period of rotation of the sun is 24.7 days. What is the Doppler shift for 500-nanometer light emitted from the edge of the sun's disk, at its equator? Compare this Doppler shift with the gravitational red shift.

(f) The sun ejects ionized hydrogen. How does the mass of a proton vary as it flies away from the sun?

15-6. (15.8) The mass of a high-energy proton

A proton has a kinetic energy of 500 million electronvolts. Find its mass and velocity.

15-7. (15.9) The conservation laws for colliding particles

In the course of a collision between two particles there is conservation of energy and conservation of momentum.

(a) Show that, if these conservation laws apply in one inertial reference frame, then they apply in any other inertial frame.

(b) Show that if, in a given reaction, relativistic mass is conserved in all inertial frames, then \mathbf{p} is also conserved, and inversely.

15-8. (15.10) Transformation of a force

Show that

$$\mathbf{F} = \mathbf{F}'_{\parallel} + \gamma \left[\frac{\mathbf{V} \times (\mathbf{F}' \times \mathbf{V})}{V^2} - \frac{\mathbf{v} \times (\mathbf{F}' \times \mathbf{V})}{c^2} \right].$$

15-9. (15.10) Force and power

Starting from Eq. 15-17, show that

$$\mathbf{F} \cdot \mathbf{v} = \frac{d\mathcal{E}}{dt}.$$

15-10. (15.11) The ultimate spaceship

The thrust of a spaceship engine is equal to the product $m'v$, where m' is the mass of propellant ejected per second and v is the exhaust velocity with respect to the ship. The ultimate spaceship would transform all its propellant into radiation and eject photons backward at the speed of light. The mass of the propellant would then be minimum.

(a) Show that the power-to-thrust ratio P/F for a photon engine is c .

Since P/F and dM/dt are independent of the frequency, the source of radiation need not be monochromatic.

(b) Then a photon ship burning 1 gram of matter per second would have a thrust of 3×10^5 newtons. The difficulty is to transform an appreciable fraction of the propellant mass into radiation, as the following example will show.

An ordinary flashlight has a capacity of about 2 ampere-hours at about 2 volts. Show that its terminal velocity is of the order of 10^{-4} meter/second.

15-11. (15.11) Is interstellar travel possible?

(a) First, time should be dilated by, say, a factor of 10. Then $\gamma = 10$. Calculate v/c .

(b) Imagine a spaceship equipped with a photon motor. See Prob. 15-10. You can find the fraction f of the initial mass that remains, after the ship has attained the proper β , from the conservation of energy and the conservation of momentum. Take into account the energy and momentum of the radiation. You should find that $f = 0.05$.

The spaceship must then brake to a stop. This requires 95% of the remaining mass. At the end of the return trip we are left with $(0.05)^4 = 6.25 \times 10^{-6}$ of the initial mass. If the mass of the ship and its payload is 1 ton, then the propellant has a mass of about 200,000 tons.

(c) In principle, the spaceship could collect and annihilate interstellar matter. There is about one atom of hydrogen per cubic centimeter.

Calculate the mass of hydrogen collected during 1 year if the ship sweeps out a volume 1000 square meters in cross-section at the speed of light.

(d) Now this hydrogen must first be brought up to speed. This slows the ship. Show that the net gain is positive up to $\beta = 0.707$ and negative afterward.

15-12. (15.11) The Doppler effect again

Refer to Fig. 14-7. An observer at the origin O of the reference frame S measures the frequency of a source of electromagnetic radiation situated at the origin O' of S' . In Prob. 14-9 we found that $f = \gamma f'(1 + \beta \cos \theta')$.

Check this equation by transforming the four-momentum of a photon.

15-13. (15.11) The Mössbauer effect

An excited nucleus of ^{57}Fe formed by the radioactive decay of ^{57}Co emits a gamma ray of 1.44×10^4 electronvolts. In the process, there is conservation of energy and $m_0 c^2 = m_o c^2 + h\nu$, where m_0 is the initial mass of the nucleus and m_o is its mass after the emission of the gamma ray. There is also conservation of momentum, $h\nu/c = m_o u$, where u is the recoil velocity of the iron nucleus. Let m_{a0} be the rest mass of the nucleus after the reaction. Then the energy released by the reaction is $\mathcal{E} = (m_0 - m_{a0})c^2$.

(a) Rewrite the first equation after subtracting $h\nu$ on both sides, and square. Then square the second equation and substitute. You should find that

$$h\nu = \frac{\mathcal{E}(m_0 + m_{a0})}{2m_0} = \left(1 - \frac{\mathcal{E}}{2m_0 c^2}\right) \mathcal{E}.$$

So $h\nu < \mathcal{E}$: part of \mathcal{E} goes to the photon, and the other part supplies kinetic energy to the recoiling nucleus.

(b) Set $m_0 \approx 57 \times 1.7 \times 10^{-27}$, and show that $\mathcal{E}/(2m_0 c^2) \approx 1.3 \times 10^{-7}$. Thus the fraction of the available energy \mathcal{E} that appears as recoil is small.

(c) Mössbauer discovered in 1958 that, with solid iron, a significant fraction of the atoms recoil as if they were locked rigidly to the rest of the solid. This is the *Mössbauer effect*. If the sample has a mass of 1 gram, by what fraction is the gamma ray energy shifted in the recoil process?

(d) A sample of normal ^{57}Fe absorbs gamma rays of 14.4 kiloelectronvolts by the inverse recoilless process much more strongly than it absorbs gamma rays of any nearby energy. The excited nuclei thus formed reemit 14.4-kiloelectronvolt radiation in random directions some time later. This is *resonant scattering*.

If a sample of activated ^{57}Fe moves in the direction of a sample of normal ^{57}Fe , what must be the value of the velocity v that will shift the frequency of the gamma rays, as seen by the normal nuclei, by 3 parts in 10^{13} ? This is one line width.

(e) A Doppler shift in the gamma ray results in a much lower absorption by a nucleus if the shift is of the order of one line width or more. What happens to the counting rate of a gamma-ray detector placed behind the sample of normal ^{57}Fe when the source of activated ^{57}Fe moves (i) toward the normal ^{57}Fe , (ii) away from it?

(f) If a 14.4-kiloelectronvolt gamma ray travels 22.5 meters vertically upward, by what fraction will its energy decrease?

(g) A normal ^{57}Fe absorber located at this height must move in what direction and at what speed in order for resonant scattering to occur?

CHAPTER 16

*RELATIVITY IV *The Field of a Moving Electric Charge*

16.1	INVARIANCE OF ELECTRIC CHARGE	284
<i>Example</i>		285
16.2	THE FOUR-CURRENT DENSITY \mathbf{J}	285
<i>Example: THE CONDUCTION CURRENT IN A WIRE</i>		287
16.3	THE FOUR-DIMENSIONAL OPERATOR \square	288
16.4	THE CONSERVATION OF CHARGE	290
16.5	THE FIELD OF A POINT CHARGE Q MOVING AT A CONSTANT VELOCITY	291
16.5.1	THE FORCE $\mathbf{F}_{\mathcal{O}q}$	292
16.5.2	THE FORCE \mathbf{F}_{Qq}	292
16.5.3	THE LORENTZ FORCE	293
16.5.4	THE ELECTRIC AND MAGNETIC FIELDS	294
<i>Example: THE FIELD OF A 10-GIGAELECTRONVOLT ELECTRON</i>		295
<i>Example: THE MAGNETIC FIELD NEAR A STRAIGHT WIRE CARRYING A STEADY CURRENT</i>		298
16.5.5	THE FORCE \mathbf{F}_{qQ}	300
16.6	TRANSFORMATION OF \mathbf{E} AND \mathbf{B}	301
<i>Example: THE PARALLEL-PLATE CAPACITOR</i>		303
16.7	SUMMARY	304
PROBLEMS		306

In this chapter we first examine two fundamental properties of electric charge, namely, invariance and conservation. However, our main objective here is to calculate the electric and magnetic fields of a charge moving at a constant velocity. We do, of course, avail ourselves of the mathematical apparatus that we developed in the last three chapters.

The simplest field of all is that of the stationary point charge. The next one, in order of complexity, is the one to which we address ourselves here.

16.1 INVARIANCE OF ELECTRIC CHARGE

Electric charge is invariant: a body carries the same electric charge for all observers. It is an experimental fact that, in an accelerator, the

charge-to-mass ratio e/m for a particle moving at an increasing velocity v varies as follows:

$$\frac{e}{m} = \frac{e}{m_0} \left(1 - \frac{v^2}{c^2}\right)^{1/2}. \quad (16-1)$$

The elementary charge e therefore remains invariant, while the mass m varies with velocity as in Sec. 15.3. This relation applies up to the highest energies attained to date.

Another proof of the invariance of electric charge is the fact that a metal object does not acquire an electric charge when it is either heated or cooled (excluding thermionic emission), despite the fact that the average kinetic energy of its conduction electrons is much less affected than that of its atoms.[†] It is because the enormous positive and negative charges in a piece of matter (Sec. 3.1) cancel perfectly at all temperatures that ordinary matter remains neutral when its temperature changes.

Example Ten kilograms of copper contain about 10^{26} atoms and $10^{26} \times 1.6 \times 10^{-19} = 1.6 \times 10^7$ colombs of conduction electrons. If the positive charge increased or decreased by only 1 part in 10^{15} upon heating, the copper would acquire a net charge of 1.6×10^{-8} coulomb. A 10-kilogram copper sphere, with a radius of about 65 millimeters, would then acquire a potential of about 2 kilovolts. Such an effect has never been observed.

16.2 THE FOUR-CURRENT DENSITY \mathbf{J}

We now show that \mathbf{J} and $c\rho$ are the components of a four-vector \mathbf{J} , called the *four-current density*.

Imagine a macroscopic charge Q moving at some arbitrary velocity \mathbf{v} with respect to a reference frame S at time t . In the inertial reference frame S_0 occupied momentarily by Q , the volume of the charge is V_0 . In S , the charge is also Q , but the volume is shorter in the direction of \mathbf{v} and

$$V = V_0 \left(1 - \frac{v^2}{c^2}\right)^{1/2}. \quad (16-2)$$

In S_0 the volume charge density is

$$\rho_0 = \frac{Q}{V_0}, \quad (16-3)$$

[†]See, for example, Charles Kittel, *Introduction to Solid State Physics*, 5th ed., John Wiley, New York, 1976, p. 166.

while in S

$$\rho = \frac{Q}{V} = \frac{\rho_0}{(1 - v^2/c^2)^{1/2}} = \gamma \rho_0 > \rho_0. \quad (16-4)$$

Now consider two nearby points A and B on the world line (Sec. 13.5) of Q , separated by the four-vector

$$\mathbf{d}\mathbf{r} = (\mathbf{d}\mathbf{r}, c dt) \quad (16-5)$$

in S , with $\mathbf{d}\mathbf{r} = \mathbf{v} dt$. As we saw in Sec. 15.6, the proper time interval between A and B is

$$dt_0 = \frac{|\mathbf{d}\mathbf{r}|}{c}. \quad (16-6)$$

Also,

$$\frac{dt}{dt_0} = \gamma = \frac{1}{(1 - v^2/c^2)^{1/2}}. \quad (16-7)$$

Thus

$$\rho = \gamma \rho_0 = \frac{dt}{dt_0} \rho_0. \quad (16-8)$$

The current density in S_0 is zero but, in S ,

$$\mathbf{J} = \rho \mathbf{v} = \rho \frac{d\mathbf{r}}{dt} = \gamma \rho_0 \frac{d\mathbf{r}}{dt} = \rho_0 \frac{dt}{dt_0} \frac{d\mathbf{r}}{dt} = \rho_0 \frac{d}{dt_0} \mathbf{r}. \quad (16-9)$$

Comparing now Eqs. 16-9 and 16-8, we see that \mathbf{J} and $c\rho$ stand out as the components of the four-vector

$$\mathbf{J} = (\mathbf{J}, c\rho) = \rho_0 \frac{d}{dt_0} (\mathbf{r}, ct) = \rho_0 \frac{\mathbf{d}\mathbf{r}}{dt_0} \quad (16-10)$$

called the *four-current density*. Here both ρ_0 and dt_0 are invariants.

Therefore \mathbf{J} and ρ transform as in Table 16-1, and the norm of \mathbf{J} is invariant:

$$|\mathbf{J}| = |J^2 - c^2 \rho^2|^{1/2} = |J'^2 - c^2 \rho'^2|^{1/2}. \quad (16-11)$$

Table 16-1 Transformation of a four-current density
 $\mathbf{J} = (\mathbf{J}, cp)$

$\mathbf{J} = \gamma(\mathbf{J}_{\parallel}' + \mathcal{V}\rho') + \mathbf{J}_{\perp}$	$\mathbf{J}' = \gamma(\mathbf{J}_{\parallel} - \mathcal{V}\rho) + \mathbf{J}_{\perp}$
$\rho = \gamma\left(\rho' + \frac{\mathcal{V}\mathbf{J}_{\parallel}'}{c^2}\right)$	$\rho' = \gamma\left(\rho - \frac{\mathcal{V}\mathbf{J}_{\parallel}}{c^2}\right)$

The analogy between the four-current density \mathbf{J} and the four-momentum \mathbf{p} is striking. You will recall from Sec. 15.6 that

$$\mathbf{p} = (p, mc), \quad |\mathbf{p}| = m_{\bullet}c, \quad (16-12)$$

and that the four-momentum is tangent to the world line of the mass m , pointing in the direction of motion. Similarly, the four-current density is

$$\mathbf{J} = (\mathbf{J}, cp), \quad (16-13)$$

its norm is

$$|\mathbf{J}| = \rho_0 \frac{|\mathbf{dr}|}{dt_0} = \rho_0 c, \quad (16-14)$$

it is tangent to the world line of the charge Q , and it points in the direction of motion if the charge is positive.

Example THE CONDUCTION CURRENT IN A WIRE

A wire that is stationary in reference frame S carries a current density J . The net volume charge density in S is zero:

$$\rho = \rho_p + \rho_n = 0. \quad (16-15)$$

There are surface charges, but we can disregard them because they just superpose another electric field over the one that we are interested in here. This extra electric field depends on the resistivity and geometry of the wire, as well as on the current flowing through it.

The wire is parallel to the x -axis, and the current flows in the negative direction, as in Fig. 16-1. Then the conduction electrons flow in the positive direction of the x -axis.

For an observer on S' moving to the right at a velocity \mathcal{V} , the Lorentz contraction for ρ_n is less than for ρ_p and the wire appears to be positively charged. From the transformations of the four-vector \mathbf{J} (Table 16-1),

$$\rho' = -\gamma \frac{\mathcal{V}}{c^2} J = \gamma \frac{\mathcal{V}}{c^2} |J|, \quad J' = \gamma J. \quad (16-16)$$

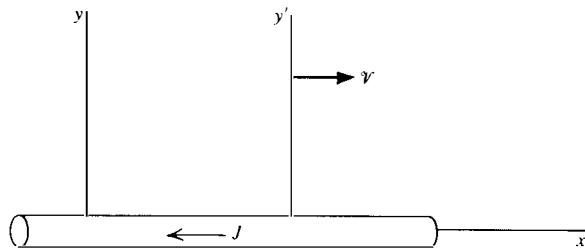


Fig. 16-1. Conducting wire carrying a current density \mathbf{J} . The conduction electrons, of volume charge density ρ_n , flow to the right at the velocity v . Reference frame S' has some velocity \mathbf{v}' parallel to the wire.

Multiplying both equations by the cross section of the wire, which is the same in S and S' , we find that the primed linear charge density is

$$\lambda' = \gamma \frac{\mathcal{V}}{c^2} |I|, \quad I' = \gamma I. \quad (16-17)$$

Another way of explaining the existence of a positive charge is the following. Imagine that, for an observer in the reference frame of the wire, the current starts to flow everywhere along the wire at $t = 0$. Then, for the observer on S' , the current flow starts at

$$t' = -\gamma \frac{\mathcal{V}}{c^2} x, \quad (16-18)$$

according to the Lorentz transformation (Sec. 13.4). That is, the electrons start flowing to the right at the right-hand end of the wire and the motion propagates to the left. This has the effect of depleting the population of conduction electrons in the wire.

16.3 THE FOUR-DIMENSIONAL OPERATOR \square

In three-dimensional space, the three operators $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$ are the components of the del operator:

$$\nabla = \hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y} + \hat{z} \frac{\partial}{\partial z}. \quad (16-19)$$

As we saw in Sec. 1.3, this operator transforms as a vector.

There exists a corresponding four-dimensional operator

$$\square = \left(\hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y} + \hat{z} \frac{\partial}{\partial z}, -\frac{\partial}{\partial ct} \right) = \left(\nabla, -\frac{\partial}{\partial ct} \right) \quad (16-20)$$

called *quad*. Note the negative sign before the fourth component. Remember that the fourth component of \mathbf{r} is $+ct$ (Sec. 15.1).

The space and time derivatives transform as follows:

$$\frac{\partial}{\partial x} = \frac{\partial x'}{\partial x} \frac{\partial}{\partial x'} + \frac{\partial t'}{\partial x} \frac{\partial}{\partial t'} = \gamma \left(\frac{\partial}{\partial x'} - \frac{\mathcal{V}}{c^2} \frac{\partial}{\partial t'} \right), \quad (16-21)$$

$$\frac{\partial}{\partial y} = \frac{\partial}{\partial y'}, \quad \frac{\partial}{\partial z} = \frac{\partial}{\partial z'}, \quad (16-22)$$

$$\frac{\partial}{\partial t} = \frac{\partial t'}{\partial t} \frac{\partial}{\partial t'} + \frac{\partial x'}{\partial t} \frac{\partial}{\partial x'} = \gamma \left(\frac{\partial}{\partial t'} - \mathcal{V} \frac{\partial}{\partial x'} \right). \quad (16-23)$$

In vector form,

$$\nabla = \gamma \left(\nabla'_{||} - \frac{\mathcal{V}}{c^2} \frac{\partial}{\partial t'} \right) + \nabla'_{\perp}, \quad (16-24)$$

$$\frac{\partial}{\partial t} = \gamma \left(\frac{\partial}{\partial t'} - \mathcal{V} |\nabla'_{||}| \right). \quad (16-25)$$

The operator \square transforms as a four-vector, as in Table 16-2, and acts like a four-vector.

If F is a scalar function of x, y, z, t , then its *four-gradient* is

$$\square F = \left(\nabla F, -\frac{\partial F}{\partial ct} \right) = \left(\frac{\partial F}{\partial x} \hat{x} + \frac{\partial F}{\partial y} \hat{y} + \frac{\partial F}{\partial z} \hat{z}, -\frac{\partial F}{\partial ct} \right). \quad (16-26)$$

The four-gradient is a four-vector if F is invariant. We shall not require the four-gradient.

The *four-divergence* of a four-vector is a scalar:

$$\square \cdot \mathbf{a} = \left(\nabla, -\frac{\partial}{\partial ct} \right) \cdot (\mathbf{a}, ca_t) = \nabla \cdot \mathbf{a} + \frac{\partial a_t}{\partial t} \quad (16-27)$$

Table 16.2 Transformation of $\square = \left(\nabla, -\frac{\partial}{\partial ct} \right)$

$$\nabla = \gamma \left(\nabla'_{||} - \frac{\mathcal{V}}{c^2} \frac{\partial}{\partial t'} \right) + \nabla'_{\perp} \quad \nabla' = \gamma \left(\nabla_{||} + \frac{\mathcal{V}}{c^2} \frac{\partial}{\partial t} \right) + \nabla_{\perp}$$

$$\frac{\partial}{\partial t} = \gamma \left(\frac{\partial}{\partial t'} - \mathcal{V} |\nabla'_{||}| \right) \quad \frac{\partial}{\partial t'} = \gamma \left(\frac{\partial}{\partial t} + \mathcal{V} |\nabla_{||}| \right)$$

$$\square \cdot \mathbf{a} = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} + \frac{\partial a_t}{\partial t}. \quad (16-28)$$

Remember, from Sec. 15.2.1, that the fourth term in the scalar product of two four-vectors is *minus* the product of the fourth components of the four-vectors.

The four-divergence is invariant because it is the scalar product of two four-vectors. Then

$$\square \cdot \mathbf{a} = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} + \frac{\partial a_t}{\partial t} = \frac{\partial a_{x'}}{\partial x'} + \frac{\partial a_{y'}}{\partial y'} + \frac{\partial a_{z'}}{\partial z'} + \frac{\partial a_{t'}}{\partial t'}. \quad (16-29)$$

We shall not require the four-dimensional curl, which is a second-order tensor.

The scalar product of \square with itself is the *d'Alembertian*:

$$\square \cdot \square = \square^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} = \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}. \quad (16-30)$$

For example, if F is either a scalar or a vector function of x, y, z, t , then

$$\square^2 F = \nabla^2 F - \frac{1}{c^2} \frac{\partial^2 F}{\partial t^2}. \quad (16-31)$$

The *wave equation* for F is $\square^2 F = 0$ (App. C).

The d'Alembertian is invariant, like the square of the norm of a four-vector.

16.4 THE CONSERVATION OF CHARGE

As we saw in Sec. 4.2, the law of conservation of charge states that

$$\nabla \cdot \mathbf{J} = - \frac{\partial \rho}{\partial t}. \quad (16-32)$$

Thus, in any frame of reference, there is never creation or annihilation of the net charge of a closed system.

Do not confuse this law with charge invariance. Charge invariance means that the electric charge carried by an object is independent of the velocity of the object with respect to the observer. In other words, the charge is the same in all inertial frames of reference. Mass is conserved in the course of an interaction (Sec. 15.9), but it is not invariant.

The law of conservation of charge can also be written as $\square \cdot \mathbf{J} = 0$:

$$\square \cdot \mathbf{J} = \left(\nabla, -\frac{\partial}{\partial ct} \right) \cdot (\mathbf{J}, c\rho) = \nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0. \quad (16-33)$$

Now $\square \cdot \mathbf{J}$ is invariant, like the scalar product of two four-vectors. Then the conservation of charge in given circumstances applies to any inertial reference frame.

16.5 THE FIELD OF A POINT CHARGE Q MOVING AT A CONSTANT VELOCITY

We consider two point charges, as in Fig. 16-2. The charge Q is at the origin O' of the inertial reference frame S' and thus moves at the constant velocity $\gamma \hat{x}$ with respect to the fixed frame S . Charge q moves at some unspecified velocity v with respect to S . That velocity need not be constant. There are no other electric charges or currents in the vicinity.

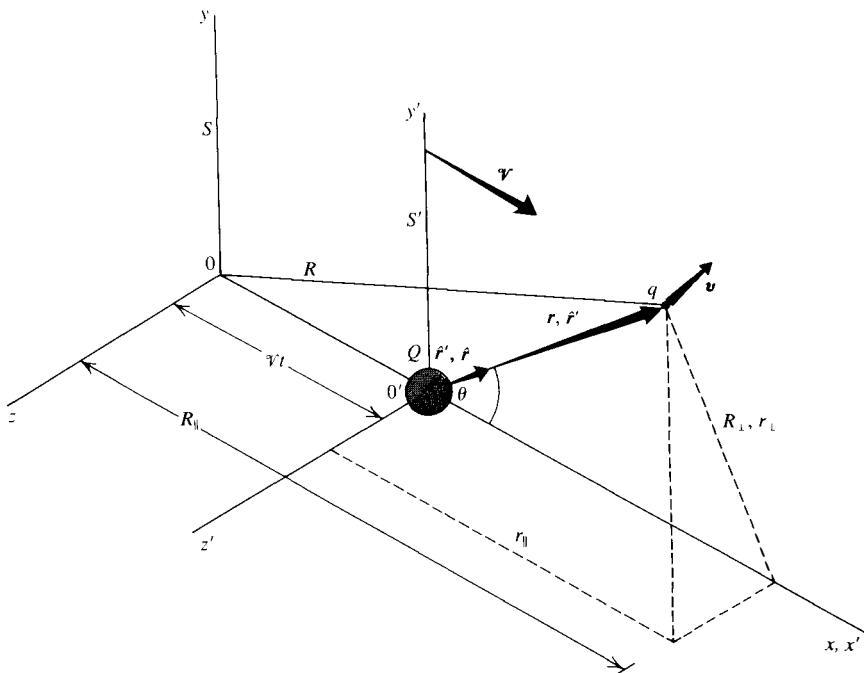


Fig. 16-2. The velocity of the charge Q at the origin O' of S' is $\gamma \hat{x}$ with respect to S . The velocity v of the charge Q with respect to S is arbitrary. All the unprimed variables shown are measured with respect to S .

The force \mathbf{F}_{Qq} exerted by Q on q , as measured in S , will give us the field of Q at the position of q , in S . We calculate this force, first in S' and then in S .

16.5.1 The Force \mathbf{F}'_{Qq}

In S' , charge Q is stationary. Now it is a well-established experimental fact that the force exerted by a stationary charge on a moving charge is independent of the latter's velocity (Sec. 3.1). We may therefore calculate \mathbf{F}'_{Qq} from Coulomb's law:

$$\mathbf{F}'_{Qq} = \frac{Qq\hat{\mathbf{r}}'}{4\pi\epsilon_0 r'^3} = \frac{Qqr'}{4\pi\epsilon_0 r'^3}, \quad (16-34)$$

where \mathbf{r}' and $\hat{\mathbf{r}}'$ are defined as in Fig. 16-2.

16.5.2 The Force \mathbf{F}_{Qq}

Referring now to Table 15-4, we can deduce \mathbf{F}_{Qq} from \mathbf{F}'_{Qq} :

$$\mathbf{F}_{Qq} = \frac{Qq}{4\pi\epsilon_0 r'^3} \left[\mathbf{r}'_{\parallel} + \frac{\gamma(\mathcal{V}/c^2)\mathbf{r}'_{\perp} \cdot \mathbf{v}'_{\perp} + \mathbf{r}'_{\perp}}{\gamma(1 + v'_{\parallel}\mathcal{V}/c^2)} \right]. \quad (16-35)$$

We still have to transform the primed quantities on the right. From Table 13-1 and from Fig. 16-2,

$$\mathbf{r}'_{\parallel} = \gamma(\mathbf{R}_{\parallel} - \mathcal{V}t) = \gamma\mathbf{r}_{\parallel}, \quad \mathbf{r}'_{\perp} = \mathbf{R}_{\perp} = \mathbf{r}_{\perp}, \quad (16-36)$$

$$r'^2 = \gamma^2(R_{\parallel} - \mathcal{V}t)^2 + R_{\perp}^2 = \gamma^2[(R_{\parallel} - \mathcal{V}t)^2 + R_{\perp}^2(1 - \beta^2)]. \quad (16-37)$$

Now

$$r^2 = (R_{\parallel} - \mathcal{V}t)^2 + R_{\perp}^2. \quad (16-38)$$

Thus

$$r'^2 = \gamma^2(r^2 - \beta^2 R_{\perp}^2) = \gamma^2 r^2 (1 - \beta^2 \sin^2 \theta), \quad (16-39)$$

where θ is measured in reference frame S . Also, from Sec. 14-4,

$$\mathbf{v}'_{\parallel} = \frac{\mathbf{v}_{\parallel} - \mathcal{V}}{1 - v'_{\parallel}\mathcal{V}/c^2}, \quad \mathbf{v}'_{\perp} = \frac{\mathbf{v}_{\perp}}{\gamma(1 - v'_{\parallel}\mathcal{V}/c^2)}. \quad (16-40)$$

Substituting and simplifying, we find that

$$\mathbf{F}_{Qq} = \frac{Qq}{4\pi\epsilon_0} \frac{\mathbf{r} + (\mathcal{V}/c^2)\mathbf{r}_\perp \cdot \mathbf{v}_\perp - (\mathcal{V}/c^2)v_\parallel \mathbf{r}_\perp}{\gamma^2 r^3 (1 - \beta^2 \sin^2 \theta)^{3/2}}, \quad (16-41)$$

where, from Identity 2 on the page inside the front cover,

$$\mathcal{V}(\mathbf{r}_\perp \cdot \mathbf{v}_\perp) - \mathcal{V}v_\parallel \mathbf{r}_\perp = \mathcal{V}(\mathbf{v} \cdot \mathbf{r}_\perp) - \mathbf{r}_\perp(\mathbf{v} \cdot \mathcal{V}) \quad (16-42)$$

$$= \mathbf{v} \times (\mathcal{V} \times \mathbf{r}_\perp) = \mathbf{v} \times (\mathcal{V} \times \mathbf{r}). \quad (16-43)$$

Finally,

$$\mathbf{F}_{Qq} = \frac{Qq}{4\pi\epsilon_0} \frac{\mathbf{r} + \mathbf{v} \times (\mathcal{V} \times \mathbf{r})/c^2}{\gamma^2 r^3 (1 - \beta^2 \sin^2 \theta)^{3/2}}. \quad (16-44)$$

16.5.3 The Lorentz Force

The force \mathbf{F}_{Qq} comprises two terms. The first term is independent of \mathbf{v} and is the *electric force*. The second term does depend on \mathbf{v} and is the *magnetic force*. So we have the *Lorentz force*

$$\mathbf{F}_{Qq} = q(\mathbf{E}_Q + \mathbf{v} \times \mathbf{B}_Q), \quad (16-45)$$

where \mathbf{v} is the velocity of q and where

$$\mathbf{E}_Q = \frac{Q}{4\pi\epsilon_0 \gamma^2 r^2 (1 - \beta^2 \sin^2 \theta)^{3/2}} \hat{\mathbf{r}} \quad (16-46)$$

is the *electric field strength* in the field of Q at the position of q , $\beta = \mathcal{V}/c$, and

$$\mathbf{B}_Q = \frac{Q(\mathcal{V} \times \hat{\mathbf{r}})/c^2}{4\pi\epsilon_0 \gamma^2 r^2 (1 - \beta^2 \sin^2 \theta)^{3/2}} \quad (16-47)$$

is similarly the *magnetic flux density* in the field of Q at the position of q . The unit of magnetic flux is the tesla.

In these three equations, all quantities concern the *same* reference frame S .

More generally, if a charge q moves at a velocity \mathbf{v} in a field \mathbf{E} , \mathbf{B} , the Lorentz force is

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (16-48)$$

The velocity \mathbf{v} need *not* be constant.

Note that, for an observer in S , the charge Q exerts on q both an electric force $q\mathbf{E}$ and a magnetic force $q\mathbf{v} \times \mathbf{B}$. However, for an observer

in the reference frame S' of Q , the charge Q is stationary and the force exerted by Q on q is purely electric.

The fact that the magnetic force exerted on an electric charge is proportional to the vector product of its velocity \mathbf{v} by the local magnetic flux density \mathbf{B} has four obvious consequences. In a given reference frame, the magnetic force on a point charge (1) exists only if the charge moves with respect to that frame, (2) is independent of the component of \mathbf{v} that is parallel to \mathbf{B} , (3) is perpendicular to both \mathbf{v} and \mathbf{B} , (4) does not affect the kinetic energy of q .

Then, if \mathcal{E} is the relativistic energy mc^2 of a particle of mass m and charge Q moving at a velocity \mathbf{v} in a field \mathbf{E} , \mathbf{B} ,

$$\frac{d\mathcal{E}}{dt} = \mathbf{F} \cdot \mathbf{v} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{v} = q\mathbf{E} \cdot \mathbf{v}. \quad (16-49)$$

Only an electric field can affect the kinetic energy of a charged particle; a magnetic field can deflect such a particle, but cannot change its speed.

16.5.4 The Electric and Magnetic Fields

It is convenient to set

$$\frac{1}{\epsilon_0 c^2} = \mu_0. \quad (16-50)$$

This is the *permeability of free space*. By definition, $\mu_0 = 4\pi \times 10^{-7}$.

We now rewrite the fields \mathbf{E} and \mathbf{B} at a point P as in Fig. 16-3 near a charge Q moving at a constant velocity \mathbf{V} :

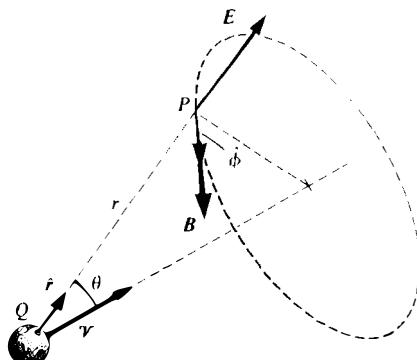


Fig. 16-3. The \mathbf{E} and \mathbf{B} vectors at a point P in the field of a charge Q moving at a constant velocity \mathbf{V} with respect to a reference frame S . The distance r and the angle θ are both measured in S .

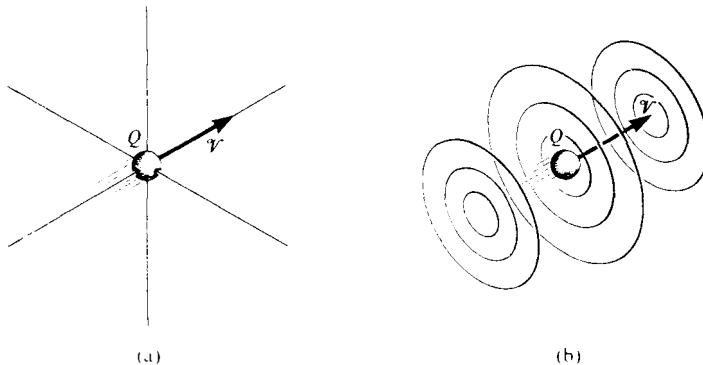


Fig. 16-4. (a) Lines of E and (b) lines of B for a positive charge Q moving at a constant velocity V , as seen by a stationary observer.

$$\mathbf{E} = \frac{Q\hat{r}}{4\pi\epsilon_0\gamma^2 r^2(1 - \beta^2 \sin^2 \theta)^{3/2}}, \quad (16-51)$$

$$\mathbf{B} = \frac{\mu_0 Q V \sin \theta \hat{\phi}}{4\pi\gamma^2 r^2(1 - \beta^2 \sin^2 \theta)^{3/2}} = \frac{\mu_0 Q V \times \hat{r}}{4\pi\gamma^2 r^2(1 - \beta^2 \sin^2 \theta)^{3/2}}. \quad (16-52)$$

See Figs. 16-4 to 16-7 and Prob. 16-3.

Observe that E is *radial*, as if the information concerning the position of the charge traveled at an infinite velocity! Actually, it is only when the velocity of the charge is constant that the electric field is radial. If the charge accelerates, the lines are not radial. See Fig. 38-1.

The lines of B are circles centered on the trajectory of Q .

At a given distance r , both E and B are maximum at $\theta = \pi/2$ and minimum at $\theta = 0$, $\theta = \pi$.

Example

THE FIELD OF A 10-GIGAELECTRONVOLT ELECTRON

What are the maximum values of E and of B 10 millimeters from the path of a single 10-gigaelectronvolt electron? One gigaelectronvolt is 10^9 electronvolts.

Both E and B are maximum at $\theta = 90^\circ$, or when the line joining the particle to the point of observation is perpendicular to the trajectory in the reference frame S of the laboratory. Then $\sin \theta = 1$,

$$\mathbf{E} = \frac{Q\hat{r}}{4\pi\epsilon_0 r^2(1/\gamma)}, \quad (16-53)$$

or γ times larger than if the electron were stationary,

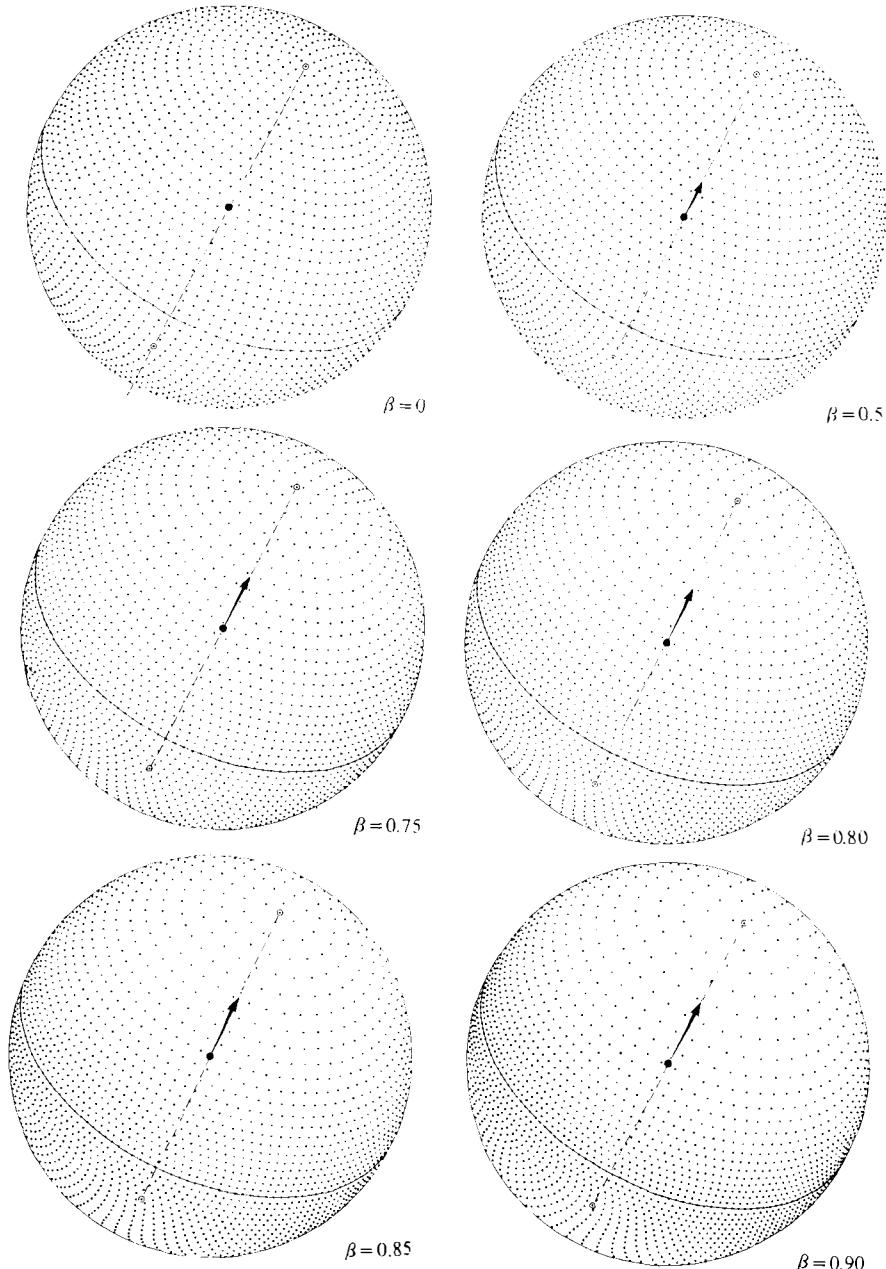


Fig. 16-5. Lines of \mathbf{E} for a charge Q moving along the diameter of an imaginary stationary sphere. The dots show where the lines emerge from the sphere at the instant when the charge goes through its center. The density of the dots is a measure of the magnitude of \mathbf{E} . The total number of dots is the same in all six figures, to satisfy Gauss's law (Sec. 3.7). Note how the field shifts toward $\theta = 90^\circ$ as the velocity increases. For $\beta = 1$ the field is all concentrated at $\theta = 90^\circ$.

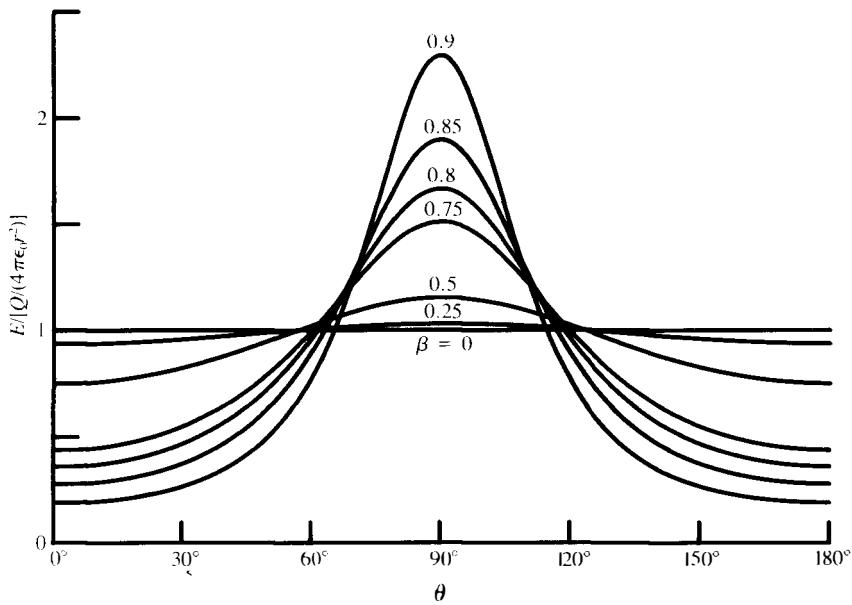


Fig. 16-6. The E of a moving point charge as a function of the polar angle θ of Figs. 16-3 and 16-5, for seven values of $\beta = V/c$. The observer is stationary and sees the charge moving at the uniform speed V . For $\beta = 0$ the field is isotropic. It is hardly disturbed at $\beta = 0.25$. As the speed increases, the field increases near $\theta = 90^\circ$ and decreases both ahead of the charge (near $\theta = 0$) and behind it (near $\theta = 180^\circ$). At extremely high velocities, most of the electric field concentrates near $\theta = 90^\circ$. These curves explain qualitatively the validity of Gauss's law for moving charges: as the speed increases, the flux of E shifts from the region where $\theta \approx 0$ and $\theta \approx 180^\circ$ to $\theta \approx 90^\circ$, and the total flux of E remains constant. (Then why are the areas under the curves not equal?) Note that the electric field is symmetric about 90° . Thus there is no way of telling from the shape of the field whether the charge is moving to the right or to the left.

$$E_{\max} = \frac{\gamma Q}{4\pi\epsilon_0 r^2} \quad (16-54)$$

in the radial direction and

$$B_{\max} = \frac{\mu_0 \gamma Q V}{4\pi r^2} \quad (16-55)$$

in the azimuthal direction.

Since the relativistic kinetic energy is 10 gigaelectronvolts,

$$(m - m_0)c^2 = m_0(\gamma - 1)c^2 = 10^{10} \text{ electronvolts}, \quad (16-56)$$

with $m_0c^2 = 5.11 \times 10^5$ electronvolts. Thus

$$\gamma = \frac{10^{10}}{5.11 \times 10^5} + 1 = 2.0 \times 10^4 \quad (16-57)$$

and $V \approx c$. Then

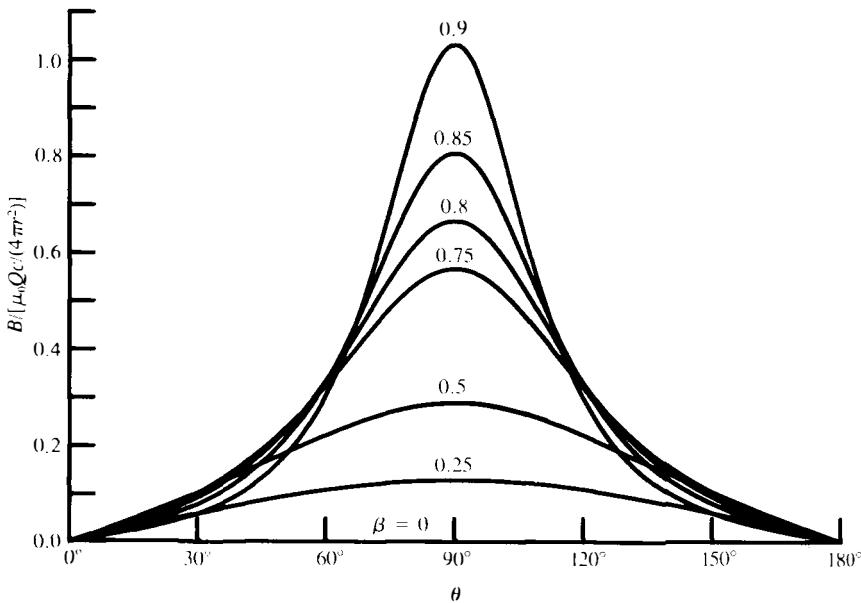


Fig. 16-7. The B of a moving point charge as a function of the polar angle θ for seven values of $\beta = v/c$. For $\beta = 0$ there is no magnetic field. As β increases, B first increases at all angles. Then B continues to increase near $\theta = 90^\circ$, while decreasing both ahead of the charge and behind it. At extremely high speeds, most of the magnetic field concentrates near the plane $\theta = 90^\circ$. The magnetic field, like the electric field, is symmetric about 90° . The maximum ordinate on any curve is β .

$$E_{\max} = \frac{(9 \times 10^9)(2 \times 10^4)(1.6 \times 10^{-19})}{10^{-4}} = 0.29 \text{ volt/meter}, \quad (16-58)$$

$$B_{\max} = \frac{(10^{-7})(2 \times 10^4)(1.6 \times 10^{-19})(3 \times 10^8)}{10^{-4}} = 9.6 \times 10^{-10} \text{ tesla}. \quad (16-59)$$

Example

THE MAGNETIC FIELD NEAR A STRAIGHT WIRE CARRYING A STEADY CURRENT

We return to the current-carrying wire of the example in Sec. 16.2. Figure 16-8 shows a positive charge Q moving at a velocity v_Q parallel to a stationary wire carrying a current I at a distance y . Let us calculate the force exerted by the wire on Q .

In the reference frame S of the wire, the net linear charge density in the wire is zero:

$$\lambda_p + \lambda_n = 0, \quad (16-60)$$

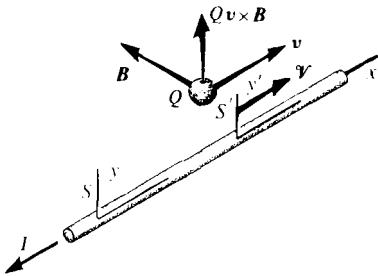


Fig. 16-8. A positive charge moves at a velocity \mathbf{v} parallel to a straight wire carrying a current I . The magnetic force $Q\mathbf{v} \times \mathbf{B}$ is in the direction shown because, for the charge Q , the wire is positively charged. If Q is positive, the force is repulsive.

where p refers to the lattice of positive charges and n to the conduction electrons. The conduction electrons drift at the velocity v_n , and their linear charge density, in their own reference frame S' , is λ'_n . Because of the Lorentz contraction,

$$\lambda_n = \gamma_n \lambda'_n, \quad \gamma_n = \left(1 - \frac{v_n^2}{c^2}\right)^{-1/2}. \quad (16-61)$$

In S , the positive charges are stationary, and exert on Q a repulsive force in the y direction

$$F_p = QE = Q\left(\frac{\lambda_p}{2\pi\epsilon_0 y}\right). \quad (16-62)$$

Similarly, the conduction electrons exert in their own frame an attractive force

$$F'_n = \frac{Q\lambda'_n}{2\pi\epsilon_0 y} = \frac{Q\lambda_n}{\gamma_n 2\pi\epsilon_0 y}, \quad (16-63)$$

where λ'_n is negative and $y' = y$. This force also points in the y direction. From Table 15-4,

$$F_n = F'_n \gamma_n \left(1 - \frac{v_n v_Q}{c^2}\right) = \frac{Q\lambda_n}{2\pi\epsilon_0 y} \left(1 - \frac{v_n v_Q}{c^2}\right). \quad (16-64)$$

This force does not quite cancel F_p .

The net force exerted by the wire on the charge Q , in the reference frame S of the wire, is

$$F = F_p + F_n = \frac{Q\lambda_p}{2\pi\epsilon_0 y} + \frac{Q\lambda_n}{2\pi\epsilon_0 y} - \frac{Q(\lambda_n v_n) v_Q}{2\pi\epsilon_0 y c^2}. \quad (16-65)$$

Now $\lambda_n v_n$ is equal to the current that flows to the right, and $-\lambda_n v_n = I$, as in the figure. So substituting μ_0 for $1/\epsilon_0 c^2$,

$$F = Qv_Q \frac{\mu_0 I}{2\pi y}. \quad (16-66)$$

Expressing this as a magnetic force $Q\mathbf{v} \times \mathbf{B}$, we find that

$$\mathbf{B} = \frac{\mu_0 I}{2\pi y} \hat{\phi} \quad (16-67)$$

in the direction shown in the figure. We shall rediscover this law in Sec. 18.2.2.

This magnetic force is infinitesimal compared to the force between Q and the conduction electrons in the wire and to the force between Q and the lattice of positive ions.

Say we have a copper wire with a cross-sectional area of 1 millimeter² and carrying a current of 1 ampere. Then suppose that Q is an electron traveling at the drift velocity of conduction electrons in the wire. We can calculate these forces as follows.

We found in Sec. 4.3.2 that the drift velocity is about 10^{-4} meter/second or about 40 centimeters/hour. Then γ is equal to unity within 1 part in 10^{25} .

Copper contains 10^{29} atoms per cubic meter and one conduction electron per atom. Then 1 meter of the wire contains 1.6×10^4 coulombs of conduction electrons. If y is 1 centimeter, then the force of attraction between the electron and the positive lattice is

$$F = \frac{\lambda_p}{2\pi\epsilon_0 y} Q = \frac{1.6 \times 10^4}{2\pi \times 8.85 \times 10^{-12} \times 10^{-2}} 1.6 \times 10^{-19} \quad (16-68)$$

$$= 4.6 \times 10^{-3} \text{ newton.} \quad (16-69)$$

This force, if acting alone, would impart to the electron an acceleration of 5×10^{27} meters/second²! The force of repulsion between Q and the conduction electrons in the wire is slightly larger.

The *net* force on an electron of velocity 10^{-4} meter/second is the magnetic force of Eq. 16-66:

$$F = \frac{1.6 \times 10^{-19} \times 10^{-4}}{4\pi \times 10^{-7} \times 1/(2\pi \times 10^{-2})} = 3.2 \times 10^{-28} \text{ newton.} \quad (16-70)$$

So the magnetic force, in this particular instance, is smaller than the electrostatic forces by 25 orders of magnitude! This is a purely relativistic effect that takes place for $V/c = 10^{-4}/(3 \times 10^8) \approx 3 \times 10^{-13}$.

Of course, the force exerted on a second *wire* is appreciable, for the simple reason that it also contains an enormous number of conduction electrons.

16.5.5 The Force \mathbf{F}_{qQ}

To calculate the force exerted by q on Q in Fig. 16-2, we interchange the roles of q and Q in Eq. 16-44. This means: (1) replacing $\hat{\mathbf{r}}$ by $-\hat{\mathbf{r}}$, as in

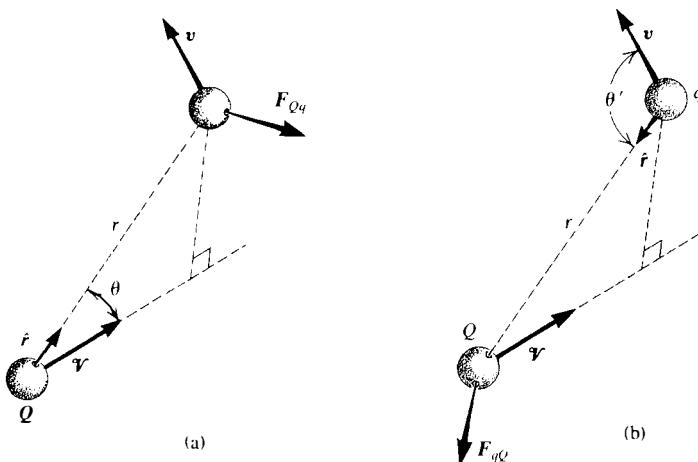


Fig. 16-9. (a) The force \mathbf{F}_{Qq} exerted by Q on q . (b) The force \mathbf{F}_{qQ} exerted by q on Q . These two forces are not equal and opposite, as one would expect from classical mechanics.

Fig. 16-9; (2) interchanging \mathcal{V} and v ; (3) replacing θ by θ' ; and (4) replacing $\beta = \mathcal{V}/c$ by $\beta' = v/c$ and γ by γ' .

Thus

$$\mathbf{F}_{qQ} = -\frac{qQ}{4\pi\epsilon_0} \frac{\mathbf{r} + \mathcal{V} \times (\mathbf{v} \times \mathbf{r})/c^2}{\gamma'^2 r^3 (1 - \beta'^2 \sin^2 \theta')^{3/2}}. \quad (16-71)$$

Clearly, $\mathbf{F}_{qQ} \neq -\mathbf{F}_{Qq}$. The electric forces are in opposite directions, but their magnitudes are not the same because of the γ' , β' , θ' terms. The magnetic forces are totally different. The forces \mathbf{F}_{Qq} and \mathbf{F}_{qQ} are equal if $\mathcal{V}^2 \ll c^2$ and if $\mathcal{V}v \ll c^2$.

There is a difference between \mathbf{F}_{qQ} and \mathbf{F}_{Qq} that is noteworthy. The expression for \mathbf{F}_{Qq} is valid if the velocity \mathcal{V} of Q is constant, while that for \mathbf{F}_{qQ} applies if \mathbf{v} is constant.

The fact that $\mathbf{F}_{qQ} \neq -\mathbf{F}_{Qq}$ is not peculiar to electrical phenomena. It is a purely relativistic effect that occurs with any type of force.

16.6 TRANSFORMATION OF \mathbf{E} AND \mathbf{B}

We have just seen that a purely electric field in one reference frame becomes both an electric and a magnetic field in another frame. We now deduce the general rule for transforming electric and magnetic fields by again transforming the force on a moving charge.

A point charge Q moves at a velocity \mathbf{v} in a region where there exists an \mathbf{E} and a \mathbf{B} , all measured with respect to the same inertial reference frame S . Then the Lorentz force on Q is

$$\mathbf{F} = Q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (16-72)$$

This same force, when it is measured in a different inertial frame S' moving at a velocity \mathcal{V} with respect to S , is

$$\mathbf{F}' = Q(\mathbf{E}' + \mathbf{v}' \times \mathbf{B}'). \quad (16-73)$$

Since the charge Q is invariant, it has the same numerical value in both frames, from Sec. 16.1.

We can find expressions for \mathbf{E} and for \mathbf{B} in terms of \mathbf{E}' , \mathbf{B}' , and \mathcal{V} in the following way. Starting with the above \mathbf{F}' , we find \mathbf{F} and then transform \mathbf{v}' to \mathbf{v} . This gives \mathbf{F} as a function of \mathbf{v} , \mathbf{E}' , \mathbf{B}' , and \mathcal{V} . Comparing this \mathbf{F} with the one above gives \mathbf{E} and \mathbf{B} as functions of \mathbf{E}' , \mathbf{B}' , and \mathcal{V} .

Since the calculation is straightforward and fairly similar to that of Sec. 16.5, we run through only part of it. We set $\mathbf{B}' = 0$. Then

$$\mathbf{F}' = QE' = Q(\mathbf{E}'_{\parallel} + \mathbf{E}'_{\perp}) \quad (16-74)$$

and, from Table 15-4,

$$\mathbf{F} = QE'_{\parallel} + \frac{\gamma(\mathcal{V}/c^2)\mathbf{v}'_{\perp} \cdot QE'_{\perp} + QE'_{\perp}}{\gamma(1 + v'_{\parallel}\mathcal{V}/c^2)}. \quad (16-75)$$

Now, from Table 14-1,

$$\mathbf{v}'_{\parallel} = \frac{\mathbf{v}_{\parallel} - \mathcal{V}}{1 - v'_{\parallel}\mathcal{V}/c^2}, \quad \mathbf{v}'_{\perp} = \frac{\mathbf{v}_{\perp}}{\gamma(1 - v'_{\parallel}\mathcal{V}/c^2)}, \quad (16-76)$$

and it is a simple matter to show that

$$1 + \frac{v'_{\parallel}\mathcal{V}}{c^2} = \frac{1}{\gamma^2(1 - v'_{\parallel}\mathcal{V}/c^2)}. \quad (16-77)$$

Substituting and simplifying, we find that

$$\mathbf{F} = Q \left(\mathbf{E}'_{\parallel} + \gamma \frac{\mathcal{V}}{c^2} \mathbf{v}_{\perp} \cdot \mathbf{E}'_{\perp} + \gamma \mathbf{E}'_{\perp} - \gamma \frac{v'_{\parallel}\mathcal{V}}{c^2} \mathbf{E}'_{\perp} \right). \quad (16-78)$$

Grouping the second and fourth terms on the right yields

$$\gamma \frac{\mathcal{V}}{c^2} \mathbf{v}_{\perp} \cdot \mathbf{E}'_{\perp} - \gamma \frac{v'_{\parallel}\mathcal{V}}{c^2} \mathbf{E}'_{\perp} = \frac{\gamma}{c^2} \mathbf{v} \times (\mathcal{V} \times \mathbf{E}'). \quad (16-79)$$

So, finally,

Table 16-3 Transformation of \mathbf{E} and \mathbf{B}

$$\begin{aligned}\mathbf{E} &= \mathbf{E}'_{\parallel} + \gamma(\mathbf{E}'_{\perp} - \mathbf{V} \times \mathbf{B}') & \mathbf{E}' &= \mathbf{E}_{\parallel} + \gamma(\mathbf{E}_{\perp} + \mathbf{V} \times \mathbf{B}) \\ \mathbf{B} &= \mathbf{B}'_{\parallel} + \gamma\left(\mathbf{B}'_{\perp} + \frac{\mathbf{V} \times \mathbf{E}'}{c^2}\right) & \mathbf{B}' &= \mathbf{B}_{\parallel} + \gamma\left(\mathbf{B}_{\perp} - \frac{\mathbf{V} \times \mathbf{E}}{c^2}\right)\end{aligned}$$

$$\mathbf{F} = Q\left(\mathbf{E}'_{\parallel} + \gamma\mathbf{E}'_{\perp} + \mathbf{v} \times \frac{\gamma}{c^2} \mathbf{V} \times \mathbf{E}'\right). \quad (16-80)$$

Comparing with Eq. 16-72,

$$\mathbf{E} = \mathbf{E}'_{\parallel} + \gamma\mathbf{E}'_{\perp}, \quad \mathbf{B} = \frac{\gamma}{c^2} \mathbf{V} \times \mathbf{E}'. \quad (16-81)$$

The purely electric field \mathbf{E}' in S' becomes both an \mathbf{E} and a \mathbf{B} field in reference frame S .

Table 16-3 shows the transformation equations that apply to any electromagnetic field.

It is fairly simple to show that

$$B^2 - \frac{E^2}{c^2} = B'^2 - \frac{E'^2}{c^2} \quad (16-82)$$

and that

$$\mathbf{E} \cdot \mathbf{B} = \mathbf{E}' \cdot \mathbf{B}'. \quad (16-83)$$

Therefore both $B^2 - E^2/c^2$ and $\mathbf{E} \cdot \mathbf{B}$ are invariants.

Example

THE PARALLEL-PLATE CAPACITOR

In its own reference frame a charged parallel-plate capacitor possesses only an electric field. But suppose the capacitor moves at a velocity \mathbf{V} parallel to its plates. What field does a fixed observer perceive?

Choose axes as in Fig. 16-10, and call S' the reference frame of the capacitor and S the laboratory frame. Then

$$E'_x = 0, \quad E'_y = E', \quad E'_z = 0, \quad B'_x = B'_y = B'_z = 0. \quad (16-84)$$

In S , from Table 16-3,

$$E_x = 0, \quad E_y = \gamma E', \quad E_z = 0, \quad (16-85)$$

$$B_z = 0, \quad B_y = 0, \quad B_x = \gamma \frac{\mathcal{V}}{c^2} E'. \quad (16-86)$$

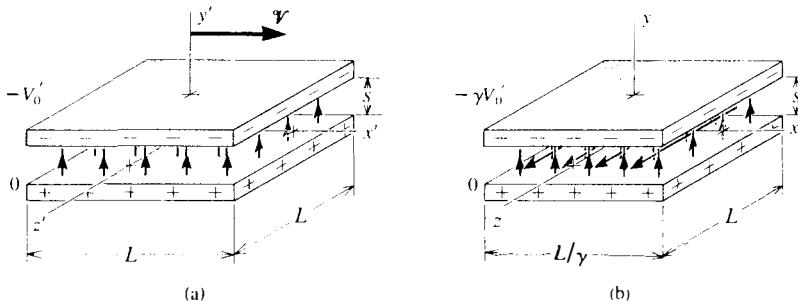


Fig. 16-10. (a) Parallel-plate capacitor, as seen in its own reference frame. The electric field strength is E' , and $B' = 0$. (b) The same capacitor moves to the right at the velocity \mathcal{V} . Here $E = \gamma E'$, and $B = \gamma(\mathcal{V}/c^2)E'$ in the direction of the z -axis. The origins of S and S' are on the lower plate. We have disregarded edge effects.

This is not difficult to explain. First, the electric field. In S' the electric charge densities are $\pm\sigma' = \pm\epsilon_0 E'$. In S the charges are the same, the plates are shorter by the factor $1/\gamma$, and $\sigma = \gamma\sigma'$. So $E = \gamma E'$. (We shall see in Sec. 17.1 that Gauss's law applies to moving charges.)

In S' the potential difference between the plates is $V'_0 = E's$. In S , the potential difference V_0 is $Es = \gamma V'_0$. The spacing s is the same in both frames because that length is perpendicular to the velocity \mathcal{V} . We return to this question of voltages in Sec. 17.8.

Now the magnetic field. An observer in the “fixed” reference frame S sees a current flowing to the right in the lower plate and an equal current flowing to the left in the upper plate. From Table 16-3,

$$B_z = \gamma \frac{\mathcal{V}}{c^2} E' = \mu_0 (\gamma \epsilon_0 E') \mathcal{V} = \mu_0 \sigma \mathcal{V} = \mu_0 I, \quad (16-87)$$

where I is the current per unit width that results from the motion of the charged plates

Finally, let us check the invariance of $E^2 - c^2 B^2$ and of $\mathbf{E} \cdot \mathbf{B}$:

$$E^2 - c^2 B^2 = \gamma^2 E'^2 - c^2 \gamma^2 \left(\frac{\mathcal{V}}{c^2} \right)^2 E'^2 \quad (16-88)$$

$$= \gamma^2 \left(1 - \frac{\mathcal{V}^2}{c^2} \right) E'^2 = E'^2 = E^2 - c^2 B'^2. \quad (16-89)$$

Also, $\mathbf{E} \cdot \mathbf{B} = \mathbf{E}' \cdot \mathbf{B}' = 0$.

16.7 SUMMARY

An electric charge is invariant: its numerical value is the same for all observers.

The *four-current density* is a four-vector that relates the current density to the charge density:

$$\mathbf{J} = (J, c\rho) = \rho_0 \frac{d\mathbf{r}}{dt}, \quad (16-10)$$

where \mathbf{J} is the usual current density, ρ is the electric charge density for a fixed observer, ρ_0 is the charge density in the reference frame of the moving charges, and \mathbf{r} is the four-vector defining the position of the charges. The four-current density transforms as in Table 16-1.

The *four-dimensional operator* \square corresponds to the del operator ∇ :

$$\square = \left(\hat{\mathbf{x}} \frac{\partial}{\partial x} + \hat{\mathbf{y}} \frac{\partial}{\partial y} + \hat{\mathbf{z}} \frac{\partial}{\partial z}, - \frac{\partial}{\partial ct} \right) = \left(\nabla, - \frac{\partial}{\partial ct} \right). \quad (16-20)$$

It transforms as in Table 16-2. The *four-divergence* of a four-vector is a scalar:

$$\square \cdot \mathbf{a} = \nabla \cdot \mathbf{a} + \frac{\partial a_t}{\partial t}. \quad (16-27)$$

The d'Alembertian operator is

$$\square \cdot \square = \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}. \quad (16-30)$$

The *law of conservation of charge* can be written as

$$\nabla \cdot \mathbf{J} = - \frac{\partial \rho}{\partial t}. \quad (16-32)$$

In the field of a charge Q moving at a constant velocity \mathbf{V} , the electric field strength and the magnetic flux density are given, respectively, by

$$\mathbf{E} = \frac{Q}{4\pi\epsilon_0\gamma^2 r^2 (1 - \beta^2 \sin^2 \theta)^{3/2}} \hat{\mathbf{r}}, \quad (16-46)$$

$$\mathbf{B} = \frac{Q(\mathbf{V} \times \hat{\mathbf{r}})/c^2}{4\pi\epsilon_0\gamma^2 r^2 (1 - \beta^2 \sin^2 \theta)^{3/2}}, \quad (16-47)$$

as in Fig. 16-3.

The *transformation equations* for \mathbf{E} and \mathbf{B} , in any field, are those of Table 16-3.

PROBLEMS

16-1. (16.1) The invariance of electric charge

Imagine that electric charge is not invariant and that $Q = Q_0[1 - (v^2/c^2)]^{1/2}$. (Remember that charge is, in fact, invariant, according to all experiments performed to date). The charge Q_0 is that measured by an observer moving with the charge, and Q is the charge for an observer moving at a velocity \mathcal{V} with respect to it.

If the electrons in a given sample have an average kinetic energy of 100 electronvolts, what percentage increase in their charge must we expect if their velocity increases by 1%?

16-2. (16.2) Conduction and convection currents in a moving ring

A square conducting ring carries a current I' in its own reference frame, as in Fig. 16-11. Its cross section is \mathcal{A}' .

(a) The ring moves at a velocity \mathcal{V} in the direction normal to its plane.

Find the current and the charge density with respect to a fixed reference frame.

(b) The ring moves to the right at a velocity $\mathcal{V}\hat{x}$.

Find the currents and the charge densities in a fixed reference frame.

(c) The motion of the space charge at a velocity $\mathcal{V}\hat{x}$ gives a *convection current*. Calculate the convection and conduction currents in the four sides.

16-3. (16.5.3) Alternate expressions for \mathbf{E} and \mathbf{B}

Show that

$$\mathbf{E} = \frac{\gamma Q \mathbf{r}}{4\pi\epsilon_0(\gamma^2x^2 + y^2 + z^2)^{3/2}}, \quad \mathbf{B} = \frac{\mu_0\gamma Q \mathbf{V} \times \mathbf{r}}{4\pi(\gamma^2x^2 + y^2 + z^2)^{3/2}}.$$

16-4. (16.5.3) The field of a 10-megaelectronvolt proton

Plot E and B as functions of the time at a point P one centimeter away from the path of a 10 megaelectronvolt proton. Set P at $(0, 0.01, 0)$, with the charge at $(\mathcal{V}t, 0, 0)$.

16-5. (16.5) The force between electrons moving side by side

Calculate the force, as observed in the laboratory, between two electrons

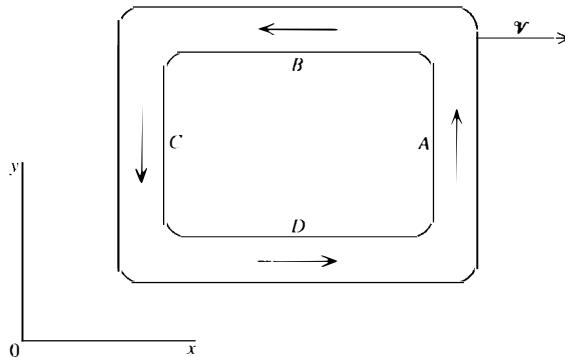


Fig. 16-11.

moving side by side along parallel paths 1 millimeter apart if they each have a kinetic energy of (a) 1 electronvolt and (b) 1 megaelectronvolt. Use Table 15-4.

16-6. (16.6) $\mathbf{E} \cdot \mathbf{B}$ is invariant

Show that $\mathbf{E} \cdot \mathbf{B}$ is invariant under a Lorentz transformation.

16-7. (16.6) $B^2 - E^2/c^2$ is invariant

Show that $B^2 - E^2/c^2$ is invariant under a Lorentz transformation.

16-8. (16.6) The angle between \mathbf{E} and \mathbf{B} is *not* invariant

Show that the angle between \mathbf{E} and \mathbf{B} is not invariant.

16-9. (16.6) Transformation of a relative permittivity ϵ_r

A dielectric-filled parallel-plate capacitor moves at a velocity $\gamma\hat{x}$ with its plates (a) parallel to the xz plane (Fig. 16-10) and (b) parallel to the yz plane. Show that $\epsilon_r = \epsilon'_r$ in both cases.

16-10. (16.6) The transformation of \mathbf{P}^\dagger

A dielectric situated in frame S' contains N'' atoms per cubic meter, each atom possessing a dipole moment $p' = Qs'$. So $P' = N''Qs'$.

Show that, with respect to frame S , $\mathbf{P} = \mathbf{P}'_{||} + \gamma\mathbf{P}'_\perp$.

16-11. (16.6) Time-independent magnetic field.

In reference frame S we have a constant magnetic field and no electric field.

Show that $\mathbf{E}' = \mathbf{V} \times \mathbf{B}'$ in S' . Note the prime on the right-hand side. So \mathbf{E}' is perpendicular to both \mathbf{V} and \mathbf{B}' .

[†] The transformation for \mathbf{M} is more difficult to prove. See Paul Penfield and Hermann A. Haus, *Electrodynamics of Moving Media*, M.I.T. Press, Cambridge, Mass., 1967, p. 209. The transformation is the same as for \mathbf{P} .

CHAPTER 17

*RELATIVITY V *Maxwell's Equations. The Four-Potential A*

17.1	THE DIVERGENCE OF \mathbf{E}	309
17.2	THE DIVERGENCE OF \mathbf{B}	310
17.3	THE CURL OF \mathbf{E}	311
17.4	THE CURL OF \mathbf{B}	312
17.5	MAXWELL'S EQUATIONS	313
17.6	THE VECTOR POTENTIAL \mathbf{A}	314
17.6.1	THE VECTOR POTENTIAL IN THE FIELD OF A CHARGE MOVING AT A CONSTANT VELOCITY	314
	<i>Example: THE VECTOR POTENTIAL NEAR THE TRAJECTORY OF A 10-GIGAELECTRONVOLT ELECTRON</i>	316
17.7	THE SCALAR POTENTIAL V . THE ELECTRIC FIELD STRENGTH \mathbf{E} EXPRESSED IN TERMS OF V AND \mathbf{A}	316
	<i>Example: THE SCALAR POTENTIAL IN THE FIELD OF A CHARGE MOVING AT A CONSTANT VELOCITY</i>	317
	<i>Example: THE VALUES OF V, ∇V, AND $\partial \mathbf{A} / \partial t$ NEAR THE TRAJECTORY OF A 10-GIGAELECTRONVOLT ELECTRON</i>	319
17.8	THE FOUR-POTENTIAL \mathbf{A} . TRANSFORMING V AND \mathbf{A}	319
	<i>Example: THE PARALLEL-PLATE CAPACITOR</i>	321
17.9	THE LORENTZ CONDITION	321
17.10	SUMMARY	322
	PROBLEMS	323

This is our last chapter on relativity. We first check the validity of two of Maxwell's equations for the field of a point charge moving at a constant velocity. The simple process of transforming these two equations then yields the other two. We shall not discuss Maxwell's equations any further for the moment; that will come in Chap. 27.

The latter part of this chapter concerns the four-potential \mathbf{A} and its components, the vector and scalar potentials \mathbf{A} and V .

17.1 THE DIVERGENCE OF \mathbf{E}

In discussing electrostatic fields in Chap. 9 we arrived at the equation

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad (17-1)$$

where ρ is the total charge density, free plus bound, $\rho_f + \rho_b$. This is Gauss's law, and it is one of Maxwell's equations. In integral form,

$$\int_{\mathcal{A}} \mathbf{E} \cdot d\mathcal{A} = \frac{Q}{\epsilon_0}, \quad (17-2)$$

where \mathcal{A} is the area of a closed surface enclosing the total charge $Q = Q_f + Q_b$.

We shall check the validity of Gauss's law for moving charges in the following special case. A charge Q moves at a constant velocity $\mathcal{V}\hat{x}$. At the instant it goes through the origin, as in Fig. 17-1, the flux of \mathbf{E} over an imaginary sphere of radius r and area \mathcal{A} , centered at the origin and stationary in the reference frame S of an observer, should be equal to Q/ϵ_0 .

At that instant, $t = 0$, and, from Eq. 16-46,

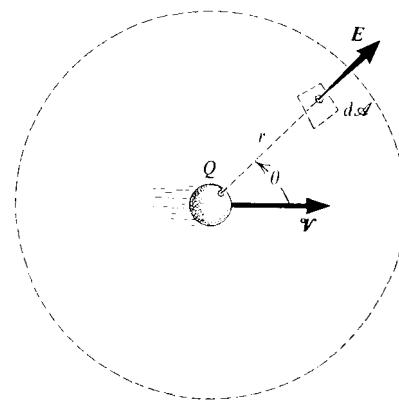


Fig. 17-1. Gauss's law applies to a moving charge: the surface integral of the normal component of the \mathbf{E} field of a moving charge Q , evaluated over a fixed sphere, is equal to Q/ϵ_0 , as if the charge were stationary. The excess field near $\theta = \pi/2$ just compensates the weak field near $\theta = 0$ and $\theta = \pi$. See Figs. 16-5 and 16-6.

$$\int_{\mathcal{A}} \mathbf{E} \cdot d\mathcal{A} = \frac{Q}{4\pi\epsilon_0\gamma^2} \int_0^\pi \frac{2\pi r^2 \sin \theta \, d\theta}{r^2(1 - \beta^2 \sin^2 \theta)^{3/2}} \quad (17-3)$$

$$= \frac{Q}{2\epsilon_0\gamma^2} \int_0^\pi \frac{\sin \theta \, d\theta}{(\beta^2 \cos^2 \theta + 1/\gamma^2)^{3/2}}. \quad (17-4)$$

To integrate on the right we set $\cos \theta = u$. Then

$$\int_{\mathcal{A}} \mathbf{E} \cdot d\mathcal{A} = \frac{Q}{2\epsilon_0\gamma^2} \int_1^{-1} \frac{-du}{(\beta^2 u^2 + 1/\gamma^2)^{3/2}} = \frac{Q}{\epsilon_0}. \quad (17-5)$$

Gauss's law therefore applies to the fields of charges moving at constant velocities. We have already utilized this fact in Figs. 16-5 and 16-6. The law applies even if the charges accelerate.

17.2 THE DIVERGENCE OF \mathbf{B}

We disregard the fields of magnetic monopoles (Sec. 18.1), and we consider only the fields of electric currents. Such a magnetic field is the sum of the magnetic fields of the individual moving charges.

We further restrict our discussion to currents in conductors. Conduction electrons (1) move in all directions because of thermal agitation, (2) have kinetic energies that vary randomly, (3) drift in the direction opposite to \mathbf{E} through a lattice of fixed positive charges, and (4) accelerate as they go around bends in the wire.

The time-averaged \mathbf{B} that results from thermal agitation is zero. So we need think of only the drift velocity.

We can dispose of the centripetal acceleration of the drifting cloud of electrons at bends. We have not discussed the fields of accelerated charges, and all that we can say is that these accelerations give rise to magnetic fields that are of the order of v/c , typically about 3×10^{-13} , times the one that we calculated in the second example in Sec. 16.5.4.

Consider Fig. 16-4. It shows that the lines of \mathbf{B} for a given charge are circles centered on, and orthogonal to, the trajectory, with \mathbf{B} uniform all around a given circle. Then imagine a volume of arbitrary shape situated in this field. Clearly the net flux of \mathbf{B} emerging through its surface is zero. In other words, for any closed surface of area \mathcal{A} ,

$$\int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} = 0 \quad (17-6)$$

for the magnetic field of one individual charge, and hence for the

magnetic field of any steady current distribution. This equation, in fact, applies to any magnetic field, excluding the fields of magnetic monopoles.

This is *Gauss's law for magnetic fields*, another one of Maxwell's equations. Applying the divergence theorem (Sec. 1.6) yields the same law in differential form:

$$\boxed{\nabla \cdot \mathbf{B} = 0.} \quad (17-7)$$

These two equations apply in any inertial reference frame.

17.3 THE CURL OF \mathbf{E}

We have just seen that $\nabla \cdot \mathbf{B} = 0$ in any inertial frame, say S . Then, in a frame S' moving at a constant velocity \mathcal{V} with respect to S ,

$$\nabla' \cdot \mathbf{B}' = 0. \quad (17-8)$$

Transforming both ∇' and \mathbf{B}' as in Tables 16-2 and 16-3, we find an equation that is valid in any frame S :

$$\left[\gamma \left(\nabla_{\parallel} + \frac{\mathcal{V}}{c^2} \frac{\partial}{\partial t} \right) + \nabla_{\perp} \right] \cdot \left[\mathbf{B}_{\parallel} + \gamma \left(\mathbf{B}_{\perp} - \frac{\mathcal{V} \times \mathbf{E}}{c^2} \right) \right] = 0. \quad (17-9)$$

Of course, the scalar product of a parallel component and a perpendicular component is zero. This leaves only four terms on the left. Expanding and then dividing by γ leads to

$$\nabla_{\parallel} \cdot \mathbf{B}_{\parallel} + \frac{\mathcal{V}}{c^2} \frac{\partial B_{\parallel}}{\partial t} + \nabla_{\perp} \cdot \mathbf{B}_{\perp} - \nabla_{\perp} \cdot \frac{\mathcal{V} \times \mathbf{E}}{c^2} = 0. \quad (17-10)$$

The sum of the first and third terms is $\nabla \cdot \mathbf{B}$, which is zero. Multiplying now by c^2 ,

$$\mathcal{V} \frac{\partial B_{\parallel}}{\partial t} = \nabla_{\perp} \cdot (\mathcal{V} \times \mathbf{E}). \quad (17-11)$$

Now the orientation of \mathcal{V} is arbitrary and

$$\mathcal{V} \cdot \frac{\partial \mathbf{B}}{\partial t} = \nabla \cdot (\mathcal{V} \times \mathbf{E}) = \mathbf{E} \cdot (\nabla \times \mathcal{V}) - \mathcal{V} \cdot (\nabla \times \mathbf{E}) = -\mathcal{V} \cdot (\nabla \times \mathbf{E}). \quad (17-12)$$

We have used Identity 7 inside the front cover, as well as the fact that \mathcal{V} is independent of the coordinates. Thus

$$(\nabla \times \mathbf{E})_{\parallel} = - \left(\frac{\partial \mathbf{B}}{\partial t} \right)_{\parallel}. \quad (17-13)$$

But, again, the direction of \mathcal{V} , to which the subscript \parallel refers, is arbitrary. It follows that

$$\nabla \times \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t}.$$

(17-14)

This is still another one of Maxwell's equations.

Applying Stokes's theorem, we find that

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = - \frac{d}{dt} \int_{\mathcal{A}} \mathbf{B} \cdot d\mathbf{A} = - \frac{d\Phi}{dt}, \quad (17-15)$$

where C is a closed curve, \mathcal{A} is the area of any open surface bounded by C , and Φ is the *magnetic flux* linking C .

17.4 THE CURL OF \mathbf{B}

We can proceed in a similar manner to relate the curl of \mathbf{B} to the other field quantities.

We saw above that

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}. \quad (17-16)$$

This equation is valid in any inertial frame S , for either stationary or moving charges. Then, in another frame S' moving at the velocity \mathcal{V} with respect to S ,

$$\nabla' \cdot \mathbf{E}' = \frac{\rho'}{\epsilon_0}. \quad (17-17)$$

Transforming again all the primed terms according to Tables 16-1, 16-2 and 16-3, we find that

$$\left(\gamma \nabla_{\parallel} + \frac{\gamma \mathcal{V}}{c^2} \frac{\partial}{\partial t} + \nabla_{\perp} \right) \cdot [\mathbf{E}_{\parallel} + \gamma (\mathbf{E}_{\perp} + \mathcal{V} \times \mathbf{B})] = \frac{\gamma}{\epsilon_0} \left(\rho - \frac{\mathcal{V} J_{\parallel}}{c^2} \right). \quad (17-18)$$

We now expand the left-hand side, as in the preceding section, and then divide by γ . Then

$$\nabla \cdot \mathbf{E} + \frac{\mathcal{V}}{c^2} \cdot \frac{\partial \mathbf{E}}{\partial t} - \mathcal{V} \cdot (\nabla \times \mathbf{B}) = \frac{1}{\epsilon_0} \left(\rho - \frac{\mathcal{V} J_{||}}{c^2} \right). \quad (17-19)$$

The first terms on either side cancel, and

$$\frac{\mathcal{V}}{c^2} \cdot \frac{\partial \mathbf{E}}{\partial t} - \mathcal{V} \cdot (\nabla \times \mathbf{B}) = \frac{\mathcal{V} \cdot \mathbf{J}}{\epsilon_0 c^2}. \quad (17-20)$$

The orientation of \mathcal{V} being arbitrary, we can again disregard the scalar multiplication by \mathcal{V} . Also, since $\epsilon_0 \mu_0 c^2 = 1$, from Sec. 16.5.4, we have the last of Maxwell's equations,

$$\nabla \times \mathbf{B} = \frac{1}{\epsilon_0 c^2} \mathbf{J} + \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}, \quad (17-21)$$

or

$$\nabla \times \mathbf{B} - \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{J}. \quad (17-22)$$

In these equations \mathbf{J} is the *total* current density at a point, including polarization currents in dielectrics (Sec. 9.3.3) and equivalent currents in magnetic materials (Sec. 20.3).

Applying Stokes's theorem leads to the integral form of the above equation:

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_0 \int_{\mathcal{A}} \left(\mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \cdot d\mathbf{A}, \quad (17-23)$$

where \mathcal{A} is the area of any open surface bounded by C .

17.5 MAXWELL'S EQUATIONS

Table 17-1 groups Maxwell's equations. *These are the four fundamental equations of electromagnetism.* We shall use them constantly throughout the remaining chapters. We have already used the first pair repeatedly in Chaps. 3 to 12. *These equations apply in any inertial reference frame.* Of course, x , y , z , t , ρ , \mathbf{J} , \mathbf{E} , and \mathbf{B} all refer to the same reference frame.

These equations are invariant: if we transform all the variables to primed quantities, then we obtain identical equations that are valid in another inertial reference frame S' .

Table 17-1 *Maxwell's equations*

DIFFERENTIAL FORM	INTEGRAL FORM
$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$	$\int_{\mathcal{A}} \mathbf{E} \cdot d\mathcal{A} = \frac{Q}{\epsilon_0}$
$\nabla \cdot \mathbf{B} = 0$	$\int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} = 0$
$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0$	$\oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{d}{dt} \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} = -\frac{d\Phi}{dt}$
$\nabla \times \mathbf{B} - \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{J}$	$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_0 \int_{\mathcal{A}} \left(\mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \cdot d\mathcal{A}$

The equation $\nabla \cdot \mathbf{J} = 0$ (Sec. 16.4) for the conservation of charge follows from the equation for the curl of \mathbf{B} , simply by taking the divergence of both sides.

We return to Maxwell's equations in Chap. 27.

Observe that the equations for $\nabla \cdot \mathbf{B}$ and for $\nabla \times \mathbf{E}$ are closely related. Indeed, transforming one gives the other. We deduced the second from the first, but the inverse operation is equally effective. These two equations are sometimes called the *first pair*.

The equations for $\nabla \cdot \mathbf{E}$ and $\nabla \times \mathbf{B}$ are similarly related. They form the *second pair*.

17.6 THE VECTOR POTENTIAL \mathbf{A}

You will remember from Chap. 3 that, in electrostatic fields, $\mathbf{E} = -\nabla V$. The magnetic flux density \mathbf{B} can be expressed in an analogous fashion as

$$\mathbf{B} = \nabla \times \mathbf{A}, \quad (17-24)$$

where \mathbf{A} is the *vector potential*, expressed in tesla-meters. It can be shown mathematically that if $\nabla \cdot \mathbf{B} = 0$, then there exists a function \mathbf{A} that satisfies the above equation.

The vector potential \mathbf{A} is not uniquely defined by \mathbf{B} : for a given $\mathbf{B}(x, y, z)$ there exists an infinite number of possible $\mathbf{A}(x, y, z)$'s.

17.6.1 The Vector Potential in the Field of a Charge Moving at a Constant Velocity

We now verify that, at a point P in the field of a point charge Q moving at a constant velocity \mathbf{V} ,

$$\mathbf{A} = \frac{\mu_0 Q \mathbf{V}}{4\pi r (1 - \beta^2 \sin^2 \theta)^{1/2}}, \quad (17-25)$$

with \mathbf{V} , \mathbf{r} , and θ defined as in Fig. 17-2. The curl of this quantity should equal the \mathbf{B} of Sec. 16.5.4.

First recall from Sec. 16.5.2 that

$$r' = \gamma r (1 - \beta^2 \sin^2 \theta)^{1/2} = [\gamma^2(x - \mathcal{V}t)^2 + y^2 + z^2]^{1/2}. \quad (17-26)$$

Thus

$$A = \frac{\mu_0 \gamma Q \mathbf{V}}{4\pi r'} = \frac{\mu_0 Q \mathbf{V}}{4\pi [\gamma^2(x - \mathcal{V}t)^2 + y^2 + z^2]^{1/2}}. \quad (17-27)$$

The four partial derivatives of $1/r'$ are as follows:

$$\frac{\partial}{\partial x} \frac{1}{r'} = -\frac{1}{2r'^3} [2\gamma^2(x - \mathcal{V}t)] = -\frac{\gamma^2(x - \mathcal{V}t)}{r'^3}, \quad (17-28)$$

$$\frac{\partial}{\partial y} \frac{1}{r'} = -\frac{1}{2r'^3} 2y = -\frac{y}{r'^3}, \quad \frac{\partial}{\partial z} \frac{1}{r'} = -\frac{z}{r'^3}, \quad (17-29)$$

$$\frac{\partial}{\partial t} \frac{1}{r'} = -\frac{1}{2r'^3} [2\gamma^2(x - \mathcal{V}t)(-\mathcal{V})] = \frac{\gamma^2 \mathcal{V}(x - \mathcal{V}t)}{r'^3}. \quad (17-30)$$

Then

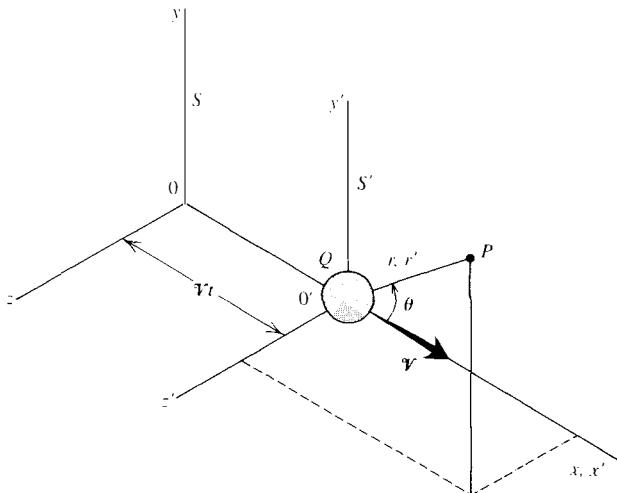


Fig. 17-2. The charge Q moves to the right at a velocity \mathbf{V} . We calculate \mathbf{A} and V at the point P . The distance QP is r in the fixed reference frame S , and it is r' in S' .

$$\nabla \times \mathbf{A} = \nabla \times (A\hat{\mathbf{x}}) = \frac{\partial A}{\partial z}\hat{\mathbf{y}} - \frac{\partial A}{\partial y}\hat{\mathbf{z}} = \frac{\mu_0}{4\pi} \gamma Q \mathcal{V} \left(-\frac{z\hat{\mathbf{y}}}{r'^3} + \frac{y\hat{\mathbf{z}}}{r'^3} \right) \quad (17-31)$$

$$= \frac{\mu_0 \gamma Q \mathcal{V} \times \mathbf{r}}{4\pi r'^3} = \frac{\mu_0 Q (\mathcal{V} \times \hat{\mathbf{r}})}{4\pi \gamma^2 r^2 (1 - \beta^2 \sin^2 \theta)^{3/2}}, \quad (17-32)$$

which is the \mathbf{B} that we found in Sec. 16.5.4.

Example

THE VECTOR POTENTIAL NEAR THE TRAJECTORY OF A 10-GIGAELECTRONVOLT ELECTRON

As in the first example of Sec. 16.5.4, $\gamma = 2.0 \times 10^4$ and $\mathcal{V} \approx c$. The vector potential is maximum at $\theta = 90^\circ$ and, at 10 millimeters from the path,

$$A_{\max} = \frac{\mu_0 \gamma Q \mathcal{V}}{4\pi r} = \frac{4\pi \times 10^{-7} \times 2 \times 10^4 \times 1.6 \times 10^{-19} \times 3 \times 10^8}{4\pi \times 10^{-2}} \quad (17-33)$$

$$= 9.6 \times 10^{-12} \text{ tesla-meter.} \quad (17-34)$$

17.7 THE SCALAR POTENTIAL V . THE ELECTRIC FIELD STRENGTH \mathbf{E} EXPRESSED IN TERMS OF V AND \mathbf{A}

We have seen in Sec. 17.3 that $\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t$. Now the curl of a gradient is zero. It follows that the equation $\mathbf{E} = -\nabla V$ of Chap. 3 cannot be general.

Electric fields result from two phenomena that we may consider, for the moment, to be distinct. First, accumulations of charge give rise to both an \mathbf{E} and a V , with $\mathbf{E} = -\nabla V$. However, if some of the charges move, there is also a \mathbf{B} and an \mathbf{A} . If \mathbf{A} is time-dependent, then there exists a further electric field $\mathbf{E} = -\partial \mathbf{A} / \partial t$. So the general expression for the electric field strength is

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t}, \quad (17-35)$$

as in Fig. 17-3. In this context, V is termed the *scalar potential*. This is an important equation. We shall return to it in Sec. 23.5.

As we shall see in Sec. 17-8, the relative values of the two terms on the right-hand side of the above equation depend on the reference frame of the observer. So, in that sense, the two phenomena that we have referred to above are not distinct.

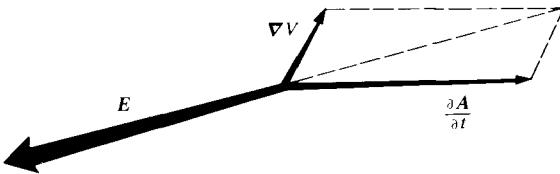


Fig. 17-3. \mathbf{E} is the vector sum of $-\nabla V$ and $-\partial \mathbf{A}/\partial t$.

Example

THE SCALAR POTENTIAL IN THE FIELD OF A CHARGE MOVING AT A CONSTANT VELOCITY

We first write down the expression for the scalar potential V in the field of a single point charge Q moving at a constant velocity \mathcal{V} , as in Fig. 17-2, and then we check the above equation.

In this field,

$$\begin{aligned} V &= \frac{Q}{4\pi\epsilon_0 r(1 - \beta^2 \sin^2 \theta)^{1/2}} \\ &= \frac{\gamma Q}{4\pi\epsilon_0 r'} = \frac{\gamma Q}{4\pi\epsilon_0 [\gamma^2(x - \mathcal{V}t)^2 + y^2 + z^2]^{1/2}}. \end{aligned} \quad (17-36)$$

Observe the analogy with the expression for \mathbf{A} given in Eq. 17-25.

Figure 17-4 shows equipotentials surrounding a moving charge, as seen by a fixed observer, for six values of $\beta = \mathcal{V}/c$. These equipotentials are spheres that are elongated by a factor of $\gamma = (1 - \beta^2)^{-1/2}$ in the directions perpendicular to the trajectory (Prob. 17-1).

We now check the validity of the above equation by proceeding as in Sec. 17.6.1:

$$\mathbf{E} = -\frac{\partial V}{\partial x} \hat{x} - \frac{\partial V}{\partial y} \hat{y} - \frac{\partial V}{\partial z} \hat{z} - \frac{\partial \mathbf{A}}{\partial t} \quad (17-37)$$

$$= -\frac{\gamma Q}{4\pi\epsilon_0 r'^3} \left[-\frac{\gamma^2(x - \mathcal{V}t)}{r'^3} \hat{x} - \frac{y}{r'^3} \hat{y} - \frac{z}{r'^3} \hat{z} \right] - \frac{\mu_0 \gamma Q \mathcal{V} \gamma^2 \mathcal{V} (x - \mathcal{V}t)}{4\pi\mathcal{V} r'^3} \hat{x}. \quad (17-38)$$

Recalling that $\epsilon_0 \mu_0 c^2 = 1$, we find that

$$\mathbf{E} = \frac{\gamma Q}{4\pi\epsilon_0 r'^3} \left[(x - \mathcal{V}t) \gamma^2 \left(1 - \frac{\mathcal{V}^2}{c^2} \right) \hat{x} + y \hat{y} + z \hat{z} \right] \quad (17-39)$$

$$= \frac{\gamma Q}{4\pi\epsilon_0 r'^3} [(x - \mathcal{V}t) \hat{x} + y \hat{y} + z \hat{z}] = \frac{\gamma Q \mathbf{r}}{4\pi\epsilon_0 r'^3} \quad (17-40)$$

$$= \frac{Q \hat{\mathbf{r}}}{4\pi\epsilon_0 \gamma^2 r^2 (1 - \beta^2 \sin^2 \theta)^{3/2}}, \quad (17-41)$$

as in Sec. 16.5.4.

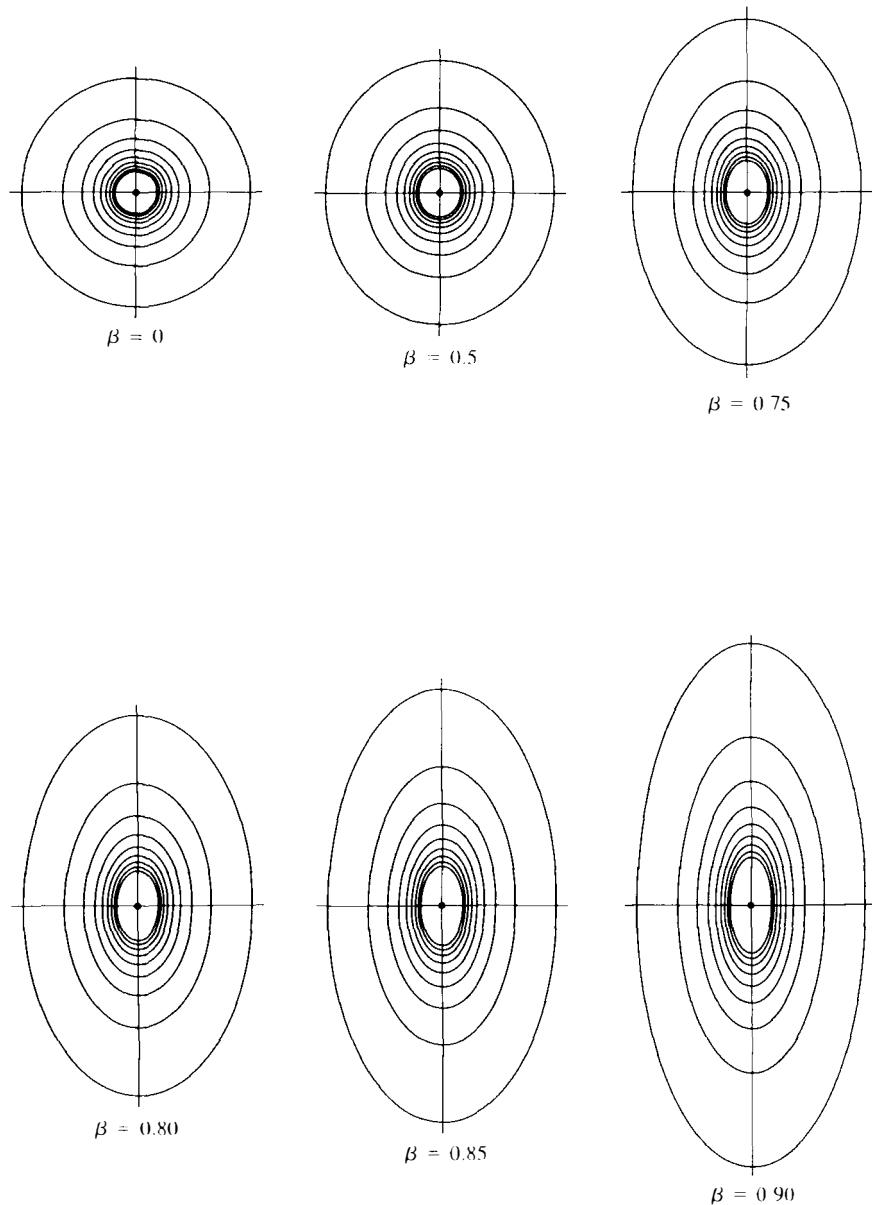


Fig. 17-4. Equipotentials for a point charge moving either to the right or to the left. The equipotentials near Q are too close together to be shown. Remember that E equals $-\nabla V - \partial A / \partial t$, not $-\nabla V$.

Example

THE VALUES OF V , ∇V , AND $\partial \mathbf{A} / \partial t$ NEAR THE TRAJECTORY OF A 10-GIGAELECTRONVOLT ELECTRON

From the first example in Sec. 16.5.4, $\gamma = 2.0 \times 10^4$. At a fixed distance r from the trajectory,

$$V_{\max} = \frac{\gamma Q}{4\pi\epsilon_0 r} \quad (17-42)$$

and is γ times larger than if the electron were stationary. At 10 millimeters from the trajectory,

$$V_{\max} = \frac{2 \times 10^4 \times 1.6 \times 10^{-19}}{4\pi \times 8.85 \times 10^{-12} \times 10^{-2}} = 2.9 \text{ millivolts.} \quad (17-43)$$

To calculate ∇V , we use Eq. 17-36 and polar coordinates:

$$\nabla V = \frac{\partial V}{\partial r} \hat{\mathbf{r}} + \frac{1}{r} \frac{\partial V}{\partial \theta} \hat{\boldsymbol{\theta}} \quad (17-44)$$

$$= \frac{Q}{4\pi\epsilon_0 r^2 (1 - \beta^2 \sin^2 \theta)^{1/2}} \left(-\hat{\mathbf{r}} + \frac{\beta^2 \sin \theta \cos \theta}{1 - \beta^2 \sin^2 \theta} \hat{\boldsymbol{\theta}} \right) \quad (17-45)$$

$$= -\frac{\gamma Q \hat{\mathbf{r}}}{4\pi\epsilon_0 r^2} \quad \left(\theta = \frac{\pi}{2} \right) \quad (17-46)$$

$$= -0.29 \hat{\mathbf{r}} \quad \text{volt/meter} \quad (r = 10 \text{ mm}). \quad (17-47)$$

This is also γ times larger than if the charge were stationary.

The vector potential is parallel to the velocity but points in the opposite direction because electrons are negative. At $\theta = 90^\circ$, the distance r is minimum and the angle θ is maximum, so \mathbf{A} is maximum, $\partial \mathbf{A} / \partial t$ is zero, and

$$\mathbf{E} = -\nabla V = 0.29 \hat{\mathbf{r}} \quad \text{volt/meter} \quad (r = 10 \text{ mm}). \quad (17-48)$$

17.8 THE FOUR-POTENTIAL \mathbf{A} . TRANSFORMING V AND \mathbf{A}

We now demonstrate that \mathbf{A} and V/c are, respectively, the space and time components of a four-vector \mathbf{A} .

We are interested in the potentials at a point P in space-time defined by the four-vector

$$\mathbf{r} = (x, y, z, ct) \quad (17-49)$$

in a reference frame S . The source of the field is a charge Q moving at a velocity \mathbf{V} . We found above that

$$\mathbf{A} = \frac{\mu_0 \gamma Q \mathbf{V}}{4\pi r'}, \quad V = \frac{\gamma Q}{4\pi \epsilon_0 r'}, \quad (17-50)$$

where r' is the distance from Q to P in the reference frame S' of Q .

For a given four-vector \mathbf{r} the distance r' is a specific distance in a specific frame. Then we may rewrite \mathbf{A} as follows:

$$\mathbf{A} = \frac{\mu_0 \gamma Q}{4\pi r'} \frac{d\mathbf{r}_Q}{dt} = \frac{\mu_0 Q}{4\pi r'} \frac{d\mathbf{r}_Q}{dt_0}, \quad (17-51)$$

where \mathbf{r}_Q defines the position of Q at the time t , and t_0 is the time in the reference frame S' of the charge Q . We saw in Sec. 14.3 that $dt = \gamma dt_0$.

Similarly,

$$V = \frac{\gamma Q}{4\pi \epsilon_0 r'} = \frac{c^2 \mu_0 Q}{4\pi r'} \frac{dt}{dt_0}. \quad (17-52)$$

We have used the relation $\epsilon_0 \mu_0 c^2 = 1$.

We can now see that \mathbf{A} and V/c are the components of the *four-potential*

$$\mathbf{A} = \left(\mathbf{A}, \frac{V}{c} \right) \quad (17-53)$$

$$= \frac{\mu_0 Q}{4\pi r'} \frac{d}{dt_0} (\mathbf{r}_Q, ct) = \frac{\mu_0 Q}{4\pi r'} \frac{d\mathbf{r}_Q}{dt_0}. \quad (17-54)$$

Equation 17-53 is general, while Eqs. 17-54 apply to a single charge Q .

Note the analogy with the four-momentum of an object of proper mass m_0 (Sec. 15.6)

$$\mathbf{p} = \left(\mathbf{p}, \frac{\mathcal{E}}{c} \right) = m_0 \frac{d\mathbf{r}}{dt_0} \quad (17-55)$$

and with the four-current density resulting from the motion of a charge of proper density ρ_0 (Sec. 16.2):

$$\mathbf{J} = (\mathbf{J}, c\rho) = \rho_0 \frac{d\mathbf{r}}{dt_0}. \quad (17-56)$$

We have therefore shown that, in the field of a point charge moving at a constant velocity, \mathbf{A} and V/c are the components of a four-vector. This is true for any electromagnetic field.

Since \mathbf{A} is a four-vector, it transforms like \mathbf{r} . Thus \mathbf{A} and V transform as in Table 17-2.

Table 17-2 Transformation of the four-potential $\mathbf{A} = (\mathbf{A}, V/c)$

$\mathbf{A} = \gamma \left(\mathbf{A}'_{\parallel} + \frac{\mathcal{V}}{c^2} \mathbf{V}' \right) + \mathbf{A}'_{\perp}$	$\mathbf{A}' = \gamma \left(\mathbf{A}_{\parallel} - \frac{\mathcal{V}}{c^2} \mathbf{V} \right) + \mathbf{A}_{\perp}$
$\frac{V}{c} = \gamma \left(\frac{V'}{c} + \frac{\mathcal{V} A'_{\parallel}}{c} \right)$	$\frac{V'}{c} = \gamma \left(\frac{V}{c} - \frac{\mathcal{V} A_{\parallel}}{c} \right)$

*Example***THE PARALLEL-PLATE CAPACITOR**

In the example in Sec. 16.6 we transformed the \mathbf{E} of a parallel-plate capacitor. Let us now transform V . In reference frame S' of the capacitor, we may set $A' = 0$, since there is zero magnetic field.

Let the potential of the lower plate be zero and that of the top plate be $-V_0$, as in Fig. 16-10. Then

$$V' = -\frac{yV'_0}{s} = -yE'. \quad (17-57)$$

From Table 17-2,

$$A_x = -\gamma \frac{\mathcal{V}}{c^2} y E', \quad A_y = 0, \quad A_z = 0, \quad V = -\gamma y E'. \quad (17-58)$$

So

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t} = -\frac{\partial V}{\partial y} \hat{\mathbf{y}} - \frac{\partial \mathbf{A}}{\partial t} = -\frac{\partial V}{\partial y} \hat{\mathbf{y}} = \gamma E', \quad (17-59)$$

$$\mathbf{B} = \nabla \times \mathbf{A} = \frac{\partial A}{\partial z} \hat{\mathbf{y}} - \frac{\partial A}{\partial y} \hat{\mathbf{z}} = -\frac{\partial A}{\partial y} \hat{\mathbf{z}} = \gamma \frac{\mathcal{V}}{c^2} E' \hat{\mathbf{z}}, \quad (17-60)$$

as in the example in Sec. 16.6

You can easily check that $|\mathbf{A}| = |\mathbf{A}'|$, or that $A^2 - V^2/c^2 = A'^2 - V'^2/c^2$.

17.9 THE LORENTZ CONDITION

The expressions for \mathbf{A} and V of Secs. 17.6 and 17.7 are so closely related that one suspects the existence of some general relation linking them. Such a relation does exist (Prob. 17-8), and it is called the *Lorentz condition*:

$$\square \cdot \mathbf{A} = 0, \quad (17-61)$$

where \square is defined in Sec. 16.3 and \mathbf{A} in Sec. 17.8. Thus

$$\nabla \cdot \mathbf{A} + \epsilon_0 \mu_0 \frac{\partial V}{\partial t} = 0, \quad (17-62)$$

which is the more usual form of the Lorentz condition.

Since the divergence of a four-vector is invariant, the Lorentz condition applies in any inertial reference frame.

Again, this result is valid for any electromagnetic field: one can always select V and \mathbf{A} so as to satisfy the Lorentz condition. Observe that the Lorentz condition is mathematically similar to the law of conservation of charge (Sec. 16.4).

We shall return to the Lorentz condition on various occasions.

17.10 SUMMARY

Gauss's law for electric fields states that

$$\boxed{\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}}, \quad (17-1)$$

where ρ is the total charge density, free plus bound, or $\rho_f + \rho_b$. This law applies even to moving charges.

Gauss's law for magnetic fields is

$$\boxed{\nabla \cdot \mathbf{B} = 0.} \quad (17-7)$$

We also found that

$$\boxed{\nabla \times \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t}}, \quad (17-14)$$

$$\boxed{\nabla \times \mathbf{B} - \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{J}.} \quad (17-22)$$

These are the *four equations of Maxwell*. They apply in any inertial reference frame.

The *vector potential* \mathbf{A} is defined by

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (17-24)$$

The general expression for the *electric field strength* is

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t}. \quad (17-35)$$

The first term on the right results from charge accumulations, and the second from changing magnetic fields.

The *four-potential* \mathbf{A} groups together the scalar potential V and the vector potential \mathbf{A} :

$$\mathbf{A} = \left(\mathbf{A}, \frac{V}{c} \right). \quad (17-53)$$

The *Lorentz condition* relates \mathbf{A} to V :

$$\square \cdot \mathbf{A} = 0, \quad (17-61)$$

or

$$\nabla \cdot \mathbf{A} + \epsilon_0 \mu_0 \frac{\partial V}{\partial t} = 0. \quad (17-62)$$

PROBLEMS

- 17-1. (17.7) The equipotentials of a moving point charge are foreshortened spheres

Show that the equipotentials of a moving point charge are foreshortened spheres, as in Fig. 17-4, for a stationary observer. Set $V=1$ and $\gamma Q/4\pi\epsilon_0 = 1$.

- 17-2. (17.8) The integral for \mathbf{A}

Verify that

$$\mathbf{A} = \frac{\mu_0}{4\pi} \int_{v'} \frac{\mathbf{J}}{r} dv'.$$

- 17-3. (17.8) Transforming the field of a parallel-plate capacitor

A charged parallel-plate capacitor moves at a velocity $\gamma \hat{x}$ in the direction normal to its plates. The capacitor plates have an area \mathcal{A} and are separated by a distance s . The vector \mathbf{E}' points in the positive direction of the x -axis, and the positive plate is at $x' = 0$.

Find V , \mathbf{A} , \mathbf{E} , and \mathbf{B} with respect to a stationary reference frame.

- 17-4. (17.8) Transformation of the field of a solenoid moving in a direction perpendicular to its axis

Inside a long solenoid, $B = \mu_0 N' I$ if N' is the number of turns per meter and the current is I . The solenoid moves at a velocity \mathbf{V} in a direction perpendicular to its length.

(a) Calculate \mathbf{E} and \mathbf{B} , both inside and outside the solenoid, as measured by a stationary observer. The axis of the solenoid is the z' -axis, and $\mathbf{V} = V\hat{\mathbf{x}}$.

(b) You can also calculate this field by transforming the potentials. First show that, in the frame of the solenoid, the vector potential

$$\mathbf{A}' = -\frac{B'y'}{2}\hat{\mathbf{x}}' + \frac{B'x'}{2}\hat{\mathbf{y}}'$$

gives the correct \mathbf{B}' . Note that there exists an infinite number of possible expressions for \mathbf{A}' . For example, we could have set $\mathbf{A}' = B'x'\hat{\mathbf{y}}$.

(c) Set $V' = 0$. Now calculate \mathbf{A} , V , \mathbf{E} , and \mathbf{B} inside the solenoid. Both V and \mathbf{A} depend on the expression that we chose arbitrarily for \mathbf{A}' . Nonetheless, the relations $\mathbf{E} = -\nabla V - \partial \mathbf{A} / \partial t$ and $\mathbf{B} = \nabla \times \mathbf{A}$ always apply.

At points outside the solenoid, in its own frame, $\mathbf{A} \neq 0$, as we shall see in the example in Sec. 19.1.

- 17-5. (17.8) Transformation of the field of a solenoid moving parallel to its axis

The solenoid of Prob. 17-4 moves at a velocity \mathbf{V} in the direction of its axis. Find \mathbf{E} and \mathbf{B} inside and outside the solenoid, as measured by a stationary observer.

- 17-6. (17.8) The paradox of the perpendicular capacitors

Figure 17-5 shows two identical capacitors set at right angle, one parallel to the velocity \mathbf{V} and the other perpendicular. In the reference frame of the capacitors $s'_a = s'_b$, $V'_{0a} = V'_{0b}$, and $E'_a = E'_b$.

For a stationary observer, $E_a = \gamma E'_a$ and $s_a = s'_a$. So $V_{0a} = E_a s_a = \gamma E'_a s'_a = \gamma V'_{0a}$. However, $E_b = E'_b$, $s_b = s'_b/\gamma$, and $V_{0b} = E_b s_b = E'_b s'_b = V'_{0b}/\gamma = V_{0a}/\gamma^2$.

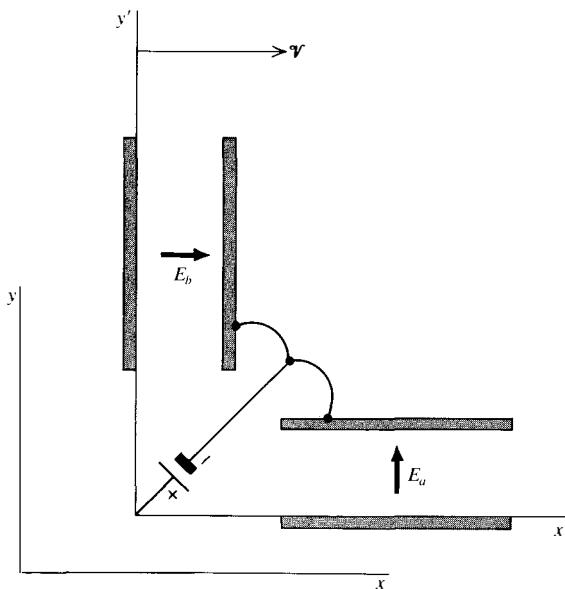
This is absurd because the capacitors are in parallel and V_{0a} must equal V_{0b} ! You can solve this paradox if you transform the potentials and the fields carefully.

- 17-7. (17.8) How the magnetic force $Q\mathbf{v} \times \mathbf{B}$ becomes an electric force $Q(-\nabla V' - \partial \mathbf{A}' / \partial t')$

A charge Q moves at a velocity \mathbf{v} in a constant, but not necessarily uniform, magnetic field \mathbf{B} . The magnetic force is $Q\mathbf{v} \times \mathbf{B}$. All three variables refer to a stationary frame S . There is no electric field.

The charge accelerates. However, at a given instant, it occupies an inertial frame S' that travels at the instantaneous velocity \mathbf{v} of the particle. With respect to S' , Q is at rest and $\mathbf{F}' = QE'$. The charge has the same value in both frames. From Prob. 16-11,

$$\mathbf{E}' = \mathbf{V} \times \mathbf{B}' = \mathbf{V} \times (\nabla' \times \mathbf{A}').$$

**Fig. 17-5.**

Show that

$$\mathbf{E}' = -\nabla' V' - \frac{\partial \mathbf{A}'}{\partial t'}.$$

You will have to show that $\nabla'(\partial/\partial x') = \partial/\partial t'$ for this particular field.

17-8. (17.9) The Lorentz condition

Show that the Lorentz condition

$$\square \cdot \mathbf{A} = 0, \quad \text{or} \quad \nabla \cdot \mathbf{A} + \epsilon_0 \mu_0 \frac{\partial V}{\partial t} = 0,$$

applies to the field of a point charge moving at a velocity \mathbf{V} with respect to the observer.

CHAPTER 18

MAGNETIC FIELDS I

*The Magnetic Flux Density **B** and the Vector Potential **A***

*18.1	MAGNETIC MONOPOLES	327
18.2	THE MAGNETIC FLUX DENSITY B . THE BIOT-SAVART LAW	327
18.2.1	THE PRINCIPLE OF SUPERPOSITION	329
	<i>Example: A LONG STRAIGHT WIRE</i>	330
	<i>Example: THE CIRCULAR LOOP</i>	331
	<i>Example: THE SOLENOID</i>	331
18.3	THE DIVERGENCE OF B	333
18.4	THE VECTOR POTENTIAL A	333
	<i>Example: A AND B NEAR A LONG, STRAIGHT WIRE</i>	335
	<i>Example: PAIR OF LONG PARALLEL CURRENTS</i>	336
	<i>Example: A IN THE FIELD OF A MAGNETIC DIPOLE</i>	337
	<i>Example: B IN THE FIELD OF A MAGNETIC DIPOLE</i>	340
18.5	THE MAGNETIC DIPOLE MOMENT OF AN ARBITRARY CURRENT DISTRIBUTION	340
18.6	SUMMARY	341
	PROBLEMS	343

Imagine a set of charges moving around in space.[†] At any point \mathbf{r} in space and at any time t there exists an electric field strength $\mathbf{E}(\mathbf{r}, t)$ and a magnetic flux density $\mathbf{B}(\mathbf{r}, t)$ that are defined as follows. If a charge Q moves at velocity \mathbf{v} at (\mathbf{r}, t) in this field, then it suffers a Lorentz force

$$\mathbf{F} = Q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (18-1)$$

The *electric force* QE is proportional to Q but independent of \mathbf{v} , while the *magnetic force* $Q\mathbf{v} \times \mathbf{B}$ is orthogonal to both \mathbf{v} and \mathbf{B} .

[†]If you have not studied Chaps. 13 to 17 on relativity, simply disregard references to them from here on.

In Chaps. 3 to 12 we studied the \mathbf{E} fields of charges that are fixed in position or that move slowly. Fixed charges have no magnetic field; a magnetic field exists only if there are moving charges.

In this chapter we study the magnetic fields of constant electric currents. We assume that the electric charge density ρ is also constant. Thus $\partial\rho/\partial t = 0$, and hence, from Sec. 4.2, $\nabla \cdot \mathbf{J} = 0$. We also assume that there are no magnetic materials, and no moving materials, in the field.

*18.1 MAGNETIC MONOPOLES

We assume here that magnetic fields arise solely from the motion of electric charges.

However, Dirac postulated in 1931 that magnetic fields can also arise from magnetic ‘‘charges,’’ called *magnetic monopoles*. Such particles have not been observed to date (1987). The theoretical value of the elementary magnetic charge is

$$\frac{h}{e} = 4.1356692 \times 10^{-15} \text{ weber.}^{\dagger} \quad (18-2)$$

where h is Planck’s constant and e is the charge of the electron. See the table inside the back cover.

At a distance r from a stationary magnetic monopole of ‘‘charge’’ Q^* , we would have that

$$\mathbf{B} = \frac{Q^*}{4\pi r^2} \hat{\mathbf{r}}. \quad (18-3)$$

Also, the force of attraction or repulsion between two monopoles Q_a^* and Q_b^* would be

$$\mathbf{F} = \frac{Q_a^* Q_b^*}{4\pi\mu_0 r^2} \hat{\mathbf{r}}. \quad (18-4)$$

A magnetic field would exert a force $Q^* \mathbf{B}/\mu_0$ on a monopole in free space.

18.2 THE MAGNETIC FLUX DENSITY \mathbf{B} . THE BIOT-SAVART LAW

In the neighborhood of an electric circuit C carrying a steady current I there exists a magnetic field and, at a point P in space, as in Fig. 18-1,

[†] This is the Dirac charge; the Schwinger charge is twice as large.

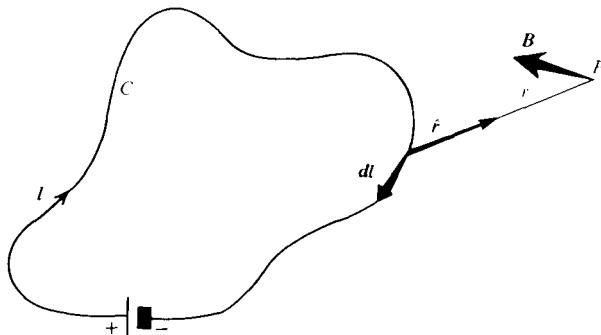


Fig. 18-1. Circuit C carrying a current I and a point P in its field. At P the magnetic flux density is \mathbf{B} .

$$\mathbf{B} = \frac{\mu_0 I}{4\pi} \oint_C \frac{dl' \times \hat{r}}{r^2}. \quad (18-5)$$

As usual, the unit vector \hat{r} points *from* the source *to* the point of observation P . This is the *Biot-Savart law*. The integration can be carried out analytically only for the simplest geometries. See below for the definition of μ_0 .

This integral applies to the fields of alternating currents, as long as the time r/c , where c is the speed of light, is a small fraction of one period (Sec. 37.4).

The unit of magnetic flux density is the *tesla*. We can find the dimensions of the tesla as follows. As we saw in the introduction to this chapter, vB has the dimensions of E . Then

$$\text{Tesla} = \frac{\text{volt second}}{\text{meter meter}} = \frac{\text{weber}}{\text{meter}^2}. \quad (18-6)$$

One volt-second is defined as 1 weber.

By definition,

$$\mu_0 = 4\pi \times 10^{-7} \text{ weber/ampere-meter}. \quad (18-7)$$

This is the *permeability of free space*.

We have assumed a current I flowing through a thin wire. If the current flows over a finite volume, we substitute $J dA'$ for I , \mathbf{J} being the current density in amperes per square meter at a point and dA' an element of area, as in Fig. 18-2. Then $J dA' dl'$ is $J dv'$ and, at a point P ,

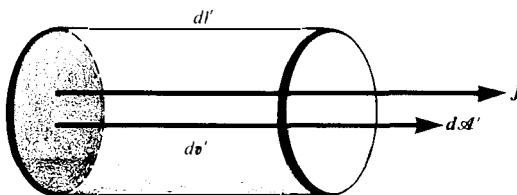


Fig. 18-2. At a given point in a volume distribution of current, the current density is \mathbf{J} . The vector $d\mathcal{A}'$ specifies the magnitude and orientation of the shaded area. Shifting this element of area to the right by the distance dl' along \mathbf{J} sweeps out a volume $d\mathcal{A}' dl' = dv'$.

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int_{v'} \frac{\mathbf{J} \times \hat{\mathbf{r}}}{r^2} dv', \quad (18-8)$$

in which v' is any volume enclosing all the currents and r is the distance between the element of volume dv' and the point P .

The current density \mathbf{J} encompasses moving free charges, polarization currents in dielectrics (Sec. 9.3.3), and equivalent currents in magnetic materials (Sec. 20.3).

Can this integral serve to calculate \mathbf{B} at a point *inside* a current distribution? The integral appears to diverge because r goes to zero when dv' is at P . The integral does not, in fact, diverge: it does apply even if the point P lies inside the conducting body. We encountered the same problem when we calculated the value of \mathbf{E} inside a charge distribution in Sec. 3.5.

Lines of \mathbf{B} point everywhere in the direction of \mathbf{B} . They prove to be just as useful as lines of \mathbf{E} . The density of lines of \mathbf{B} is proportional to the magnitude of \mathbf{B} .

As with electric fields again, a great deal of convenience attends the use of the concept of flux. The *magnetic flux* through a surface of area \mathcal{A} is

$$\Phi = \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} \text{ webers.} \quad (18-9)$$

The surface is usually open; if it is closed, then $\Phi = 0$, as we shall see below.

18.2.1 The Principle of Superposition

The above integrals for \mathbf{B} imply that the net magnetic flux density at a point is the sum of the \mathbf{B} 's of the elements of current $I dl'$, or $\mathbf{J} dv'$. The

principle of superposition applies to magnetic fields as well as to electric fields (Sec. 3.3): if there exist several current distributions, then the net \mathbf{B} is the vector sum of the individual \mathbf{B} 's

Example | **A LONG STRAIGHT WIRE**

An element dl' of a long, straight wire carrying a current I , as in Fig. 18-3, gives, at the point $P(r, \theta, \phi)$, a magnetic flux density

$$d\mathbf{B} = \frac{\mu_0 I dl' \sin \theta}{4\pi r^2} \hat{\phi} = \frac{\mu_0 I dl' \cos \alpha}{4\pi r^2} \hat{\phi}. \quad (18-10)$$

The relative orientations of I and \mathbf{B} satisfy the right-hand screw rule. Here

$$l = \rho \tan \alpha, \quad dl' = \frac{r d\alpha}{\cos \alpha} = \frac{r^2 d\alpha}{\rho}. \quad (18-11)$$

Thus

$$\mathbf{B} = \frac{\mu_0 I}{4\pi\rho} \int_{-\pi/2}^{+\pi/2} \cos \alpha d\alpha \hat{\phi} = \frac{\mu_0 I}{2\pi\rho} \hat{\phi}. \quad (18-12)$$

Lines of \mathbf{B} are circles lying in a plane perpendicular to the wire and centered on it. The magnitude of \mathbf{B} falls off as $1/\rho$.

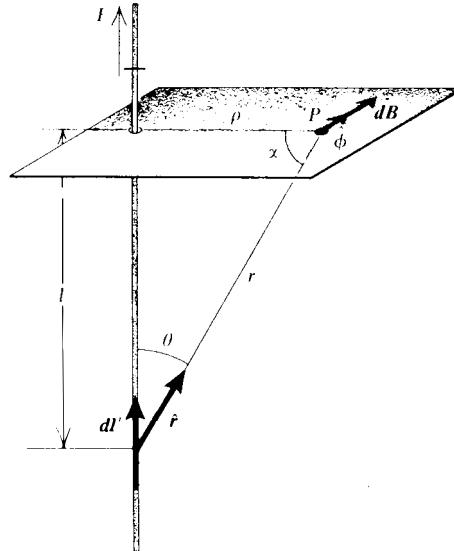


Fig. 18-3. Long, straight wire carrying a current I . At the point P the element $I dl'$ contributes a $d\mathbf{B}$ in the direction shown. A line of \mathbf{B} is a circle centered on the wire.

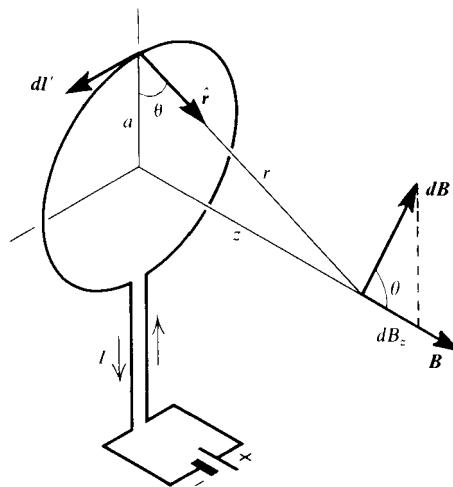


Fig. 18-4. Coil of wire of radius R carrying a current I , the field $d\mathbf{B}$ that originates in the element $I dl'$, and the total field \mathbf{B} .

Example

THE CIRCULAR LOOP

To calculate the value of \mathbf{B} on the axis of a circular loop of radius a , refer to Fig. 18-4. The figure shows the $d\mathbf{B}$ of an element of current $I dl'$. By symmetry, the total \mathbf{B} points along the axis and

$$dB_z = \frac{\mu_0}{4\pi} \frac{I dl'}{r^2} \cos \theta, \quad (18-13)$$

$$B_z = \frac{\mu_0}{4\pi} \frac{2\pi a I}{r^2} \cos \theta = \frac{\mu_0 I a^2}{2(a^2 + z^2)^{3/2}}. \quad (18-14)$$

Along the axis, $B = \mu_0 I / 2a$ at $z = 0$ and falls off as $1/z^3$ for $z^2 \gg a^2$.

Example

THE SOLENOID

The above result can serve to calculate \mathbf{B} on the axis of the solenoid of Fig. 18-5 by summing the $d\mathbf{B}$'s of the individual turns. The solenoid is close-wound, of length L , with N' turns per meter, and its radius is R . At the center,

$$B = \frac{\mu_0}{2} \int_{-L/2}^{+L/2} \frac{R^2 N' I dz}{(R^2 + z^2)^{3/2}} \quad (18-15)$$

$$= \frac{\mu_0}{2} N' I \frac{L}{(R^2 + L^2/4)^{1/2}} = \mu_0 N' I \sin \theta_m. \quad (18-16)$$

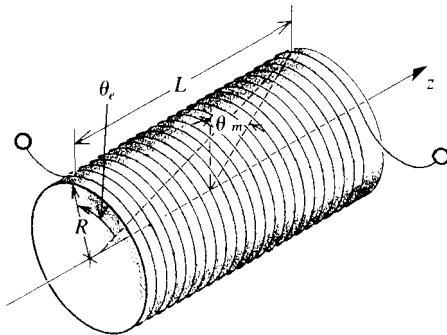


Fig. 18-5. Solenoid.

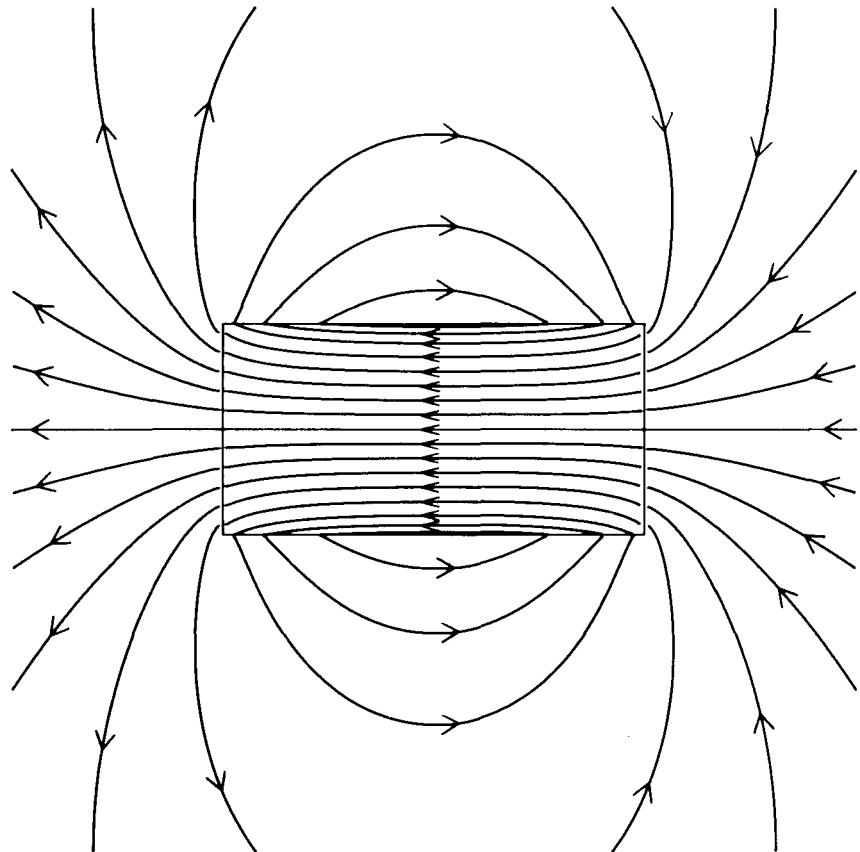


Fig. 18-6. Lines of B for a solenoid whose length is equal to twice its diameter.

See Fig. 18-5 for the definitions of θ_m and θ_e . At one end, again on the axis,

$$B = \frac{\mu_0 N' I \sin \theta_e}{2}. \quad (18-17)$$

The magnetic flux density is larger at the center than at the ends because the lines of \mathbf{B} flare out at the ends, as in Fig. 18-6. Inside a *long* solenoid, at points remote from the ends, $B \approx \mu_0 N' I$.

Calculating \mathbf{B} at a point off the axis would be much more difficult.

18.3 THE DIVERGENCE OF \mathbf{B}

Assuming that magnetic monopoles do not exist (Sec. 18.1), or at least that the net magnetic charge density is everywhere zero, all magnetic fields result from electric currents, and the lines of \mathbf{B} for each element of current are circles, as in Fig. 18-3. Thus the net outward flux of \mathbf{B} through any closed surface is zero:

$$\int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} = 0. \quad (18-18)$$

Applying the divergence theorem, it follows that

$$\nabla \cdot \mathbf{B} = 0. \quad (18-19)$$

These are alternate forms of one of Maxwell's equations. Observe that Eq. 18-19 establishes a relation between the space derivatives of \mathbf{B} at a given *point*. Equation 18-18, on the contrary, concerns the magnetic flux over a *closed surface*.

18.4 THE VECTOR POTENTIAL \mathbf{A}

We have just seen that $\nabla \cdot \mathbf{B} = 0$. It is convenient to set

$$\mathbf{B} = \nabla \times \mathbf{A}, \quad (18-20)$$

where \mathbf{A} is the *vector potential*, as opposed to V , which is the *scalar potential*. The divergence of \mathbf{B} is then automatically equal to zero because the divergence of a curl is zero.

It is immediately apparent that, for a given \mathbf{B} , there exist an infinite number of possible \mathbf{A} 's. Indeed, one can add to \mathbf{A} any quantity whose curl is zero, for example $25\hat{x}$, without affecting \mathbf{B} . The magnetic flux

density is a measurable quantity, but \mathbf{A} is known only within an additive term.[†]

Note the analogy with the relation

$$\mathbf{E} = -\nabla V \quad (18-21)$$

of electrostatics.

The vector potential is an important quantity; we shall use it as often as V .

Notice also that \mathbf{B} is a function of the space derivatives of \mathbf{A} , just as \mathbf{E} is a function of the space derivatives of V . Thus, to deduce the value of \mathbf{B} from \mathbf{A} at a given point P , one must know the value of \mathbf{A} in the *region* around P .

We now deduce the integral for \mathbf{A} , starting from the Biot-Savart law of Sec. 18.2:

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int_{v'} \frac{\mathbf{J} \times \hat{\mathbf{r}}}{r^2} dv' = \frac{\mu_0}{4\pi} \int_{v'} \left(\nabla \frac{1}{r} \right) \times \mathbf{J} dv', \quad (18-22)$$

from Identity 16 inside the back cover. Applying now Identity 11, we find that

$$\left(\nabla \frac{1}{r} \right) \times \mathbf{J} = \nabla \times \frac{\mathbf{J}}{r} - \frac{\nabla \times \mathbf{J}}{r}, \quad (18-23)$$

where the second term on the right is zero because \mathbf{J} is a function of x' , y' , z' , while ∇ involves derivatives with respect to x , y , z . Thus

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int_{v'} \left(\nabla \times \frac{\mathbf{J}}{r} \right) dv' = \nabla \times \left(\frac{\mu_0}{4\pi} \int_{v'} \frac{\mathbf{J}}{r} dv' \right), \quad (18-24)$$

and

$$\mathbf{A} = \frac{\mu_0}{4\pi} \int_{v'} \frac{\mathbf{J}}{r} dv'. \quad (18-25)$$

This expression for \mathbf{A} has a definite value for a given current distribution.

This integral, like that for \mathbf{B} , appears to diverge inside a current-carrying conductor, because of the r in the denominator. Actually, it is well behaved, like the integral for V inside a charge distribution.

If a current I flows in a circuit C that is not necessarily closed, then, at a point $P(x, y, z)$ in space,

[†] See Richard P. Feynman, Robert B. Leighton, and Matthew Sands, *Lectures on Physics*, Vol. 2, Sec. 15-5, Addison-Wesley, Reading, Mass., 1964, for a discussion of some quantum mechanical aspects of \mathbf{A} .

$$\mathbf{A} = \frac{\mu_0 I}{4\pi} \int_C \frac{d\mathbf{l}'}{r}, \quad (18-26)$$

where the element $d\mathbf{l}'$ of circuit C is at $P'(x', y', z')$, and r is the distance between P and P' .

These two integrals apply to the fields of alternating currents if the time delay r/c is a small fraction of one period.

Example | **A AND \mathbf{B} NEAR A LONG, STRAIGHT WIRE**

We first calculate \mathbf{A} and then deduce \mathbf{B} in the field of the long, straight, current-carrying wire of Fig. 18-7. We should find the same value of \mathbf{B} as in the first example in Sec. 18.2.1.

At a distance ρ

$$d\mathbf{A} = \frac{\mu_0 I d\mathbf{l}'}{4\pi r}. \quad (18-27)$$

The vector \mathbf{A} is parallel to the wire and points in the direction of the current.

For an infinitely long conductor, $d\mathbf{A}$ is proportional to $d\mathbf{l}'/l$ for large values of r where $r \approx l$. Then \mathbf{A} tends to infinity logarithmically. However, the fact that a function is infinite does *not* mean that its derivatives are infinite; that is, \mathbf{B} can be finite even though \mathbf{A} is infinite.

We can circumvent this infinite value of \mathbf{A} by first calculating \mathbf{A} and \mathbf{B} for a wire of finite length L and then setting $L \gg \rho$ at the end of the calculation. Referring to Fig. 18-7, we see that

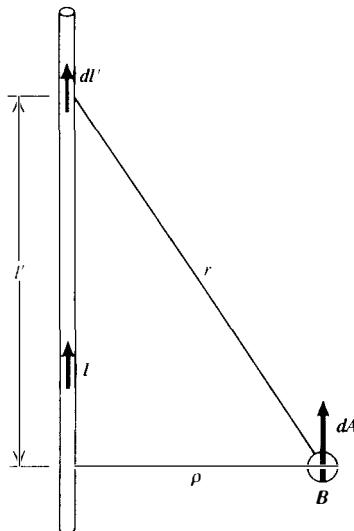


Fig. 18-7. The element of vector potential $d\mathbf{A}$ due to the element $I d\mathbf{l}'$. The vector \mathbf{B} is azimuthal, as in Fig. 18-3.

$$A = \frac{\mu_0 I}{4\pi} \int_{-L/2}^{L/2} \frac{dl}{(\rho^2 + l'^2)^{1/2}} = \frac{\mu_0 I}{2\pi} \left(\ln \frac{l' + (\rho^2 + l'^2)^{1/2}}{\rho} \right) \Big|_0^{L/2} \quad (18-28)$$

$$\begin{aligned} &= \frac{\mu_0 I}{2\pi} \ln \frac{(L/2)[1 + (1 + 4\rho^2/L^2)^{1/2}]}{\rho} \approx \frac{\mu_0 I}{2\pi} \ln \frac{L}{\rho} \quad (4\rho^2 \ll L^2) \\ &\approx \frac{\mu_0 I}{2\pi} \ln \frac{\mathcal{R}}{\rho} \quad (4\rho^2 \ll L^2) \end{aligned} \quad (18-29)$$

In this last expression we have neglected a term in $\ln(L/\mathcal{R})$ and we have arbitrarily set $A = 0$ at the radius \mathcal{R} . See also Prob. 18-13.

To calculate $\mathbf{B} = \nabla \times \mathbf{A}$, we use cylindrical coordinates, keeping in mind that A is parallel to the z -axis and independent of both ϕ and z . From the expression for $\nabla \times \mathbf{A}$ on the back of the front cover,

$$\mathbf{B} = \nabla \times \mathbf{A} = \frac{\mu_0 I}{2\pi\rho} \hat{\phi} \quad (4\rho^2 \ll L^2), \quad (18-30)$$

as in Secs. 16.5.4 and 18.2.1.

Example

PAIR OF LONG PARALLEL CURRENTS

Figure 18-8 shows two long parallel wires separated by a distance D and carrying equal currents I in opposite directions. To

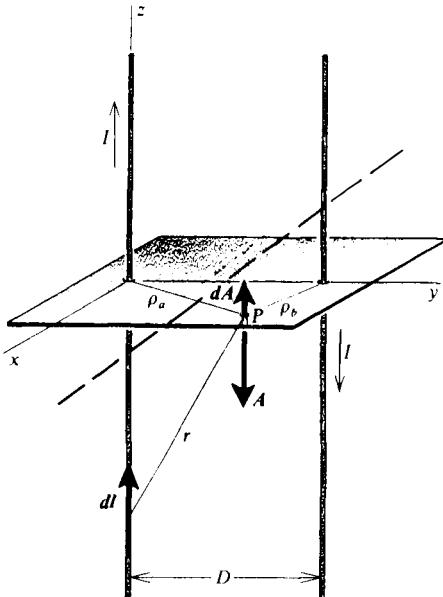


Fig. 18-8. Pair of long parallel wires carrying currents of the same magnitude in opposite directions. The vector \mathbf{A} is zero in the vertical plane that passes through the dashed line, and it points upward on the left and downward on the right.

calculate \mathbf{A} , we use the above result for the \mathbf{A} of a single wire and add the two vector potentials:

$$A = \frac{\mu_0 I}{2\pi} \left(\ln \frac{L}{\rho_b} - \ln \frac{L}{\rho_a} \right) = \frac{\mu_0 I}{2\pi} \ln \frac{\rho_a}{\rho_b} = \frac{\mu_0 I}{4\pi} \ln \frac{x^2 + (D-y)^2}{x^2 + y^2}. \quad (18-31)$$

The vector \mathbf{A} points in the direction of the current that is closer to P ; it is zero in the plane $\rho_a = \rho_b$. Then

$$B_x = \frac{\partial A}{\partial y} = -\frac{\mu_0 I}{2\pi} \left(\frac{D-y}{\rho_b^2} + \frac{y}{\rho_a^2} \right), \quad (18-32)$$

$$B_y = -\frac{\partial A}{\partial x} = \frac{\mu_0 I}{2\pi} \left(\frac{1}{\rho_a^2} - \frac{1}{\rho_b^2} \right), \quad (18-33)$$

$$B_z = 0. \quad (18-34)$$

Along the line midway between the two wires,

$$B_x = -\frac{2\mu_0 I}{\pi D}, \quad B_y = 0, \quad B_z = 0. \quad (18-35)$$

Example

A IN THE FIELD OF A MAGNETIC DIPOLE

A magnetic dipole is a loop of wire carrying a current I , as in Fig. 18-9. We calculate \mathbf{A} in this section and \mathbf{B} in the next. This will lead us to an interesting relationship between the \mathbf{B} field of a magnetic dipole and the \mathbf{E} field of an electric dipole.

We calculated the field on the axis of a circular loop in the second example in Sec. 18.2.1. We now calculate \mathbf{A} and \mathbf{B} at any remote point in space, at distances r that are much larger than the radius a of the loop. Figure 18-10(b) shows the field close to and inside the loop. The field in that region, away from the axis, is difficult to calculate.

At the point P in Fig. 18-9,

$$\mathbf{A} = \frac{\mu_0 I}{4\pi} \oint_C \frac{d\mathbf{l}'}{r}. \quad (18-36)$$

By symmetry, \mathbf{A} is azimuthal: for any value of r' we have two symmetric $d\mathbf{l}'$'s whose y -components add and whose x -components cancel. Then we need calculate only the y -component of the \mathbf{A} in the figure, and

$$A = \frac{\mu_0 I}{4\pi} \int_0^{2\pi} \frac{a d\phi \cos \phi}{r'}. \quad (18-37)$$

We can express r' as a function of r and ϕ in the following way. Refer to the figure. First,

$$r'^2 = r^2 + a^2 - 2ar \cos \psi. \quad (18-38)$$

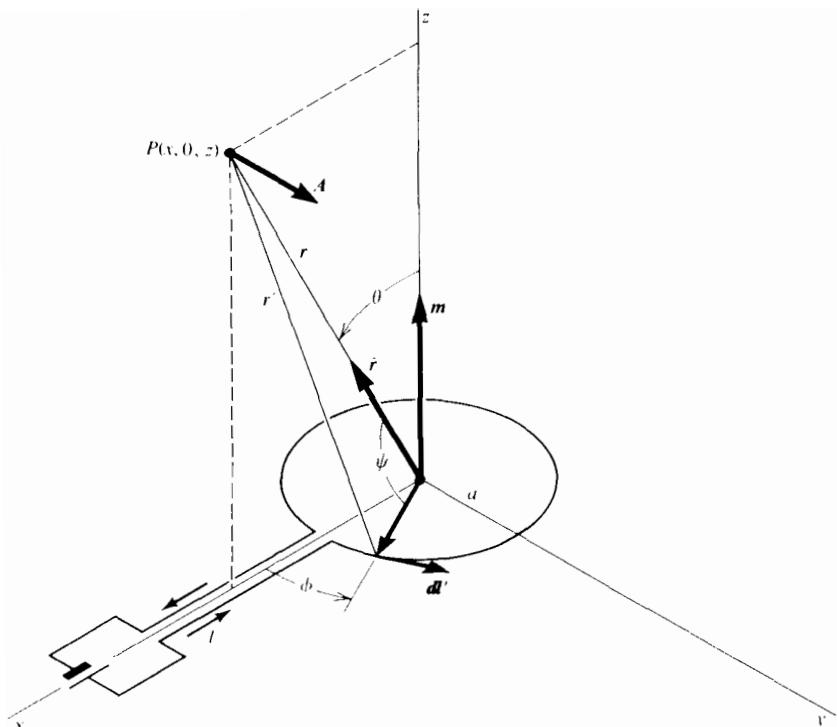


Fig. 18-9. Magnetic dipole. The vector A is azimuthal.

Now

$$x \cos \phi = r \cos \psi \quad (18-39)$$

and

$$r'^2 = r^2 + a^2 - 2ax \cos \phi, \quad (18-40)$$

$$r' = r \left\{ 1 + \left[\frac{a^2}{r^2} - 2 \frac{a}{r} \left(\frac{x}{r} \cos \phi \right) \right] \right\}^{1/2} = r \{ 1 + [] \}^{1/2}. \quad (18-41)$$

Observe that $x/r \cos \phi \leq 1$.

Since we are interested in the field only at points where $r \gg a$, we expand $1/r'$ as an infinite series and disregard terms involving higher powers of a/r . Thus

$$\frac{1}{r'} = \frac{1}{r} \left\{ 1 - \frac{1}{2} [] + \frac{3}{8} []^2 - \dots \right\}. \quad (18-42)$$

Setting

$$\left(\frac{x}{r} \cos \phi \right) = (), \quad (18-43)$$

we find that

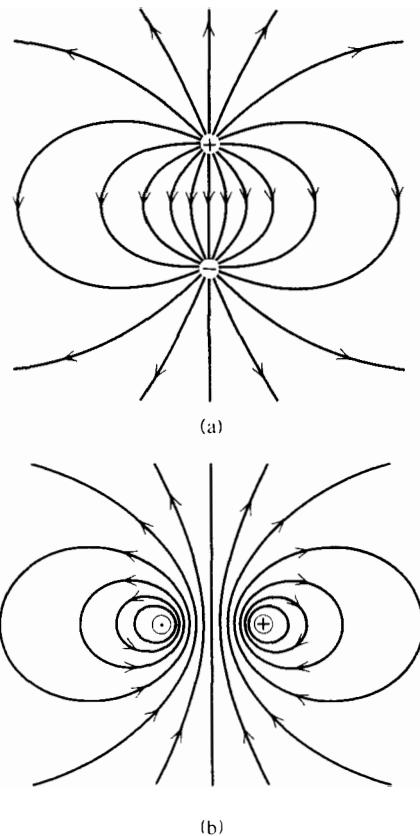


Fig. 18-10. The fields (a) of an electric dipole and (b) of a magnetic dipole, in the immediate vicinity of the dipoles.

$$\frac{1}{r'} = \frac{1}{r} \left\{ 1 - \frac{1}{2} \left[\frac{a^2}{r^2} - 2 \frac{a}{r} \right] + \frac{3}{8} \left[\frac{a^4}{r^4} - 4 \frac{a^3}{r^3} \right] + 4 \frac{a^2}{r^2} (\)^2 \right\} - \dots \quad (18-44)$$

Discarding now all terms containing the third and higher powers of a/r ,

$$\frac{1}{r'} = \frac{1}{r} \left\{ 1 + \frac{a}{r} (\) - \left[\frac{1}{2} - \frac{3}{2} (\)^2 \right] \frac{a^2}{r^2} \right\}. \quad (18-45)$$

Finally, substituting into Eq. 18-37 yields

$$A = \frac{\mu_0 I a}{4\pi r} \int_0^{2\pi} \left[1 + \frac{a}{r} \left(\frac{x}{r} \cos \phi \right) - \left(\frac{1}{2} - \frac{3}{2} \frac{x^2}{r^2} \cos^2 \phi \right) \frac{a^2}{r^2} \right] \cos \phi \, d\phi. \quad (18-46)$$

Only the second term between the brackets survives, and

$$A = \frac{\mu_0 I a^2 x}{4r^3} = \frac{\mu_0 I a^2 \sin \theta}{4r^2} \quad (r^3 \gg a^3). \quad (18-47)$$

By definition,

$$\mathbf{m} = \pi a^2 I \hat{z} \quad (18-48)$$

is the *magnetic dipole moment* of the loop. If there are N turns, then \mathbf{m} is N times larger.

Since \mathbf{A} is azimuthal,

$$\mathbf{A} = \frac{\mu_0 \mathbf{m} \times \hat{\mathbf{r}}}{4\pi r^3} \quad (r^3 \gg a^3). \quad (18-49)$$

The condition $r^3 \gg a^3$ is easy to satisfy: at $r = 5a$, $a^3/r^3 = \frac{1}{125}$.

Example

B IN THE FIELD OF A MAGNETIC DIPOLE

The value of $\mathbf{B} = \nabla \times \mathbf{A}$ follows immediately:

$$\mathbf{B} = \frac{\mu_0 \mathbf{m}}{4\pi r^3} (2 \cos \theta \hat{\mathbf{r}} + \sin \theta \hat{\theta}). \quad (18-50)$$

The analogy with the field of the electric dipole of Sec. 5.1 is obvious. The analogy, however, applies solely at distances r that are large compared to the sizes of the dipoles. Figure 18-10 shows the near fields.

18.5 THE MAGNETIC DIPOLE MOMENT OF AN ARBITRARY CURRENT DISTRIBUTION

Assume first a plane loop of arbitrary shape carrying a current I . Then we can set

$$\mathbf{m} = \mathcal{A} I \hat{z}, \quad (18-51)$$

where \mathcal{A} is the area of the loop, and where the unit vector \hat{z} is normal to the loop and satisfies the right-hand screw rule. The above expressions for \mathbf{A} and for \mathbf{B} apply as long as $a^3 \ll r^3$, where a is now the longest dimension of the loop.

According to Prob. 1-3, we can also write that

$$\mathbf{m} = \frac{1}{2} I \oint_C \mathbf{r} \times d\mathbf{l}', \quad (18-52)$$

where the line integral runs in the direction of the current I . The origin of \mathbf{r} can be anywhere.

What if loop C does not lie in a plane? Imagine an arbitrary surface bounded by C , as in Fig. 18-11, divided into infinitesimal plane cells,

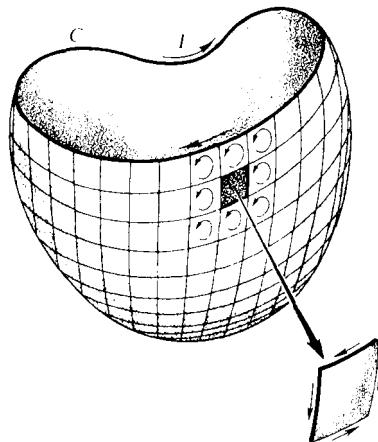


Fig. 18-11. Loop C carrying a current I . The loop is not plane. We have divided an arbitrary surface bounded by C into infinitesimal plane cells, each carrying a current I .

each one carrying a current I around its periphery. Adjoining currents cancel, and the magnetic dipole moment of C is the vector sum of the magnetic dipole moments of the individual cells. Thus

$$\mathbf{m} = \sum \frac{1}{2} I \oint_{\text{cell}} \mathbf{r} \times d\mathbf{l}' = \frac{1}{2} I \oint_C \mathbf{r} \times d\mathbf{l}', \quad (18-53)$$

and Eq. 18-52 applies to any closed circuit.

An arbitrary current distribution possesses a magnetic dipole moment

$$\mathbf{m} = \frac{1}{2} \int_v \mathbf{r} \times \mathbf{J} dv'. \quad (18-54)$$

We have replaced $I d\mathbf{l}'$ by $\mathbf{J} dA' d\mathbf{l}'$ and $dA' d\mathbf{l}'$ by dv' . The origin of \mathbf{r} can be anywhere.

18.6 SUMMARY

A charge Q moving at a velocity \mathbf{v} in the field of an arbitrary distribution of charges and currents is subjected to a *Lorentz force*

$$\mathbf{F} = Q(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \quad (18-1)$$

where QE is the *electric force*, $Q\mathbf{v} \times \mathbf{B}$ is the *magnetic force*, and \mathbf{B} is the *magnetic flux density*, expressed in *teslas*.

In the field of a current I flowing through a circuit C ,

$$\mathbf{B} = \frac{\mu_0 I}{4\pi} \oint_C \frac{d\mathbf{l}' \times \hat{\mathbf{r}}}{r^2}. \quad (18-5)$$

In the field of a volume current distribution,

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int_{v'} \frac{\mathbf{J} \times \hat{\mathbf{r}}}{r^2} dv'. \quad (18-8)$$

Lines of \mathbf{B} point everywhere in the direction of \mathbf{B} .

The *magnetic flux* through an area \mathcal{A} is

$$\Phi = \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} \quad \text{webers.} \quad (18-9)$$

The *principle of superposition* applies to magnetic fields.

The net magnetic flux through a closed surface is zero:

$$\int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} = 0. \quad (18-18)$$

Hence

$$\nabla \cdot \mathbf{B} = 0. \quad (18-19)$$

These are, respectively, the integral and the differential forms of one of Maxwell's equations.

The equation

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (18-20)$$

defines the *vector potential* \mathbf{A} . For a volume current distribution,

$$\mathbf{A} = \frac{\mu_0}{4\pi} \int_{v'} \frac{\mathbf{J}}{r} dv' \quad (18-25)$$

while, for a current I flowing through a circuit C ,

$$\mathbf{A} = \frac{\mu_0 I}{4\pi} \int_C \frac{d\mathbf{l}'}{r}. \quad (18-26)$$

In the field of a *magnetic dipole*,

$$\mathbf{A} = \frac{\mu_0}{4\pi} \frac{\mathbf{m} \times \hat{\mathbf{r}}}{r^2}, \quad (18-50)$$

where the magnitude of the *magnetic dipole moment* \mathbf{m} is the area of the

loop, times the number of turns, times the current. The vector \mathbf{m} is normal to the plane of the loop, in the direction defined by the right-hand screw rule. Refer to Fig. 18-9. Also,

$$\mathbf{B} = \frac{\mu_0 m}{4\pi r^3} (2 \cos \theta \hat{\mathbf{r}} + \sin \theta \hat{\theta}). \quad (18-50)$$

The magnetic dipole moment of an arbitrary circuit C carrying a current I is

$$\mathbf{m} = \frac{1}{2} I \oint_C \mathbf{r} \times d\mathbf{l}'. \quad (18-53)$$

For an arbitrary current distribution occupying a volume v ,

$$\mathbf{m} = \frac{1}{2} \int_{v'} \mathbf{r} \times \mathbf{J} dv', \quad (18-54)$$

where \mathbf{J} is the current density at a point and \mathbf{r} is the position vector for that point. The origin of \mathbf{r} is arbitrary.

PROBLEMS

- 18-1. (18.1) The force on a magnetic monopole situated in a magnetic field
 Show that the equation $F = Q_a^* Q_b^* / (4\pi\mu_0 r^2)$ is dimensionally correct.
 This means that $\mathbf{F} = Q^* \mathbf{B} / \mu_0$, and not $Q^* \mathbf{B}$, as stated by some authors.
- 18-2. (18.2.1) The field of two parallel wires
 Two parallel wires of radius R and separated by a distance $2D$ carry a current I in opposite directions.
 (a) Find B along a perpendicular line passing through the wires.
 (b) Plot B for $R = 1.00$ millimeter, $D = 10.0$ millimeters, and $I = 1$ ampere.
- 18-3. (18.2.1) Saddle coils
 Figure 18-12(a) shows two views of a pair of saddle coils. In (a) we have shown just one turn in each coil, and (b) shows a cross section C . More generally, we could have the current distribution of Fig. 18-12(b), where the two parts carry equal current *densities*. There is zero current in the central region. We could also have the current distribution of Fig. 18-12(c). As we shall see, the magnetic fields in the cavities are uniform.
 (a) Show that $\mathbf{B} = \mu_0 \mathbf{J} \times \mathbf{r} / 2$ inside a conductor of circular cross section. The origin of \mathbf{r} is at the center of the cross section.
 (b) In Fig. 18-12(a), \mathbf{B} is the same as if each conductor occupied a full circle, with opposite currents in the central region.
 Find \mathbf{B} in the central region.
 (c) Find \mathbf{B} in the cavities of Figs. 18-12(b) and (c).

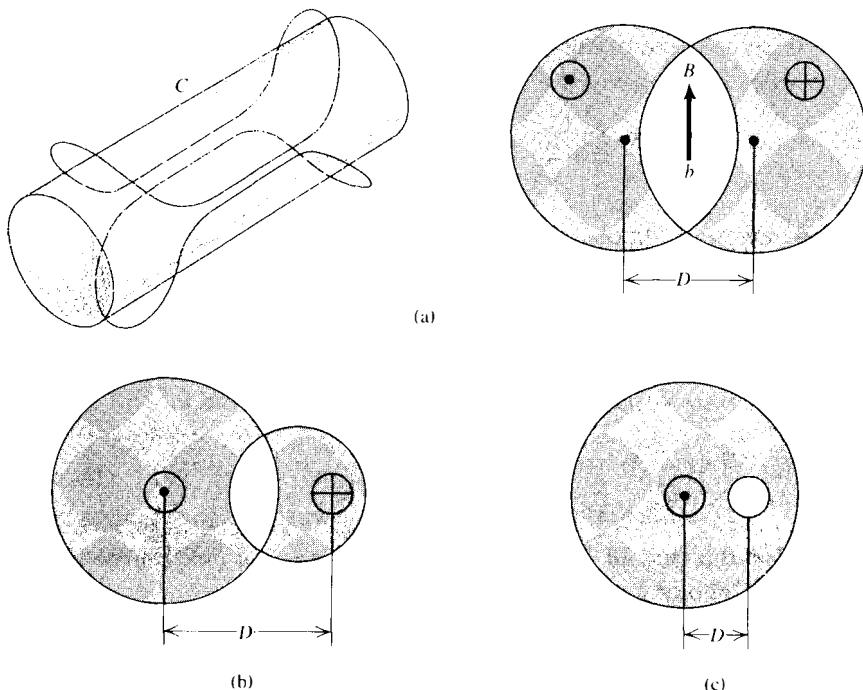


Fig. 18-12.

18-4. (18.2.1) The magnetic flux density at the center of a sunspot

The Zeeman effect observed in the spectra of sunspots reveals the existence of magnetic fields as large as 0.4 tesla. These fields are associated with pancake-shaped current distributions in the plasma near the surface. In effect, one has a disk of electrons, with a radius of, say, 10^7 meters, rotating at an angular velocity of the order of 3×10^{-2} radian/second. The thickness of the disk is small compared to its radius.

(a) Calculate the surface density of electrons required to achieve a B of 0.4 tesla at the center.

(b) Calculate the current.

18-5. (18.2.1) The Bohr magneton

According to the old Bohr model of the atom, electrons describe orbits around the nucleus. Atomic and molecular magnetic moments are expressed in Bohr magnetons.

(a) Find the magnetic moment of an electron on a circular orbit of radius r .

(b) According to the Bohr postulate, the angular momentum is quantized: $mvr = n\hbar = nh/2\pi = n \times 1.0546 \times 10^{-34}$, where n is an integer and a quantum number.

Calculate the value of the *Bohr magneton* μ_B , which is the magnetic moment of an electron orbit for which $n = 1$. The number of Bohr magnetons per atom or per molecule is of the order of a few and is, in fact, not an even number.

18-6. (18.2.1) Rotating magnetic field

Three identical coils, oriented as in Fig. 18-13(a), carry three-phase alternating current. Their magnetic fields are

$$B_a = B_m \cos \omega t, \quad B_b = B_m \cos \left(\omega t + \frac{2\pi}{3} \right), \quad B_c = B_m \cos \left(\omega t + \frac{4\pi}{3} \right),$$

and point as in Fig. 18-13(b).

(a) Show that the resulting field has a magnitude of $1.5B_m$ and rotates at an angular velocity ω . This is the method used to generate rotating magnetic fields in large electric motors.

(b) Does the field rotate clockwise or anticlockwise?

18-7. (18.2.1) The Fabry equation for solenoids

A solenoid has an inner radius R_1 , an outer radius R_2 , and a length $2L$. The current is I .

(a) Show that at the center

$$B = \mu_0 n I L \ln \frac{\alpha + (\alpha^2 + \beta^2)^{1/2}}{1 + (1 + \beta^2)^{1/2}},$$

where n is the number of turns per square meter ($\approx 1/\text{cross section of the wire}$), $\alpha = R_2/R_1$, and $\beta = L/R_1$.

(b) Show that the length of the wire is

$$l = nV = 2\pi n(\alpha^2 - 1)\beta R_1^3,$$

where V is the volume of the winding.

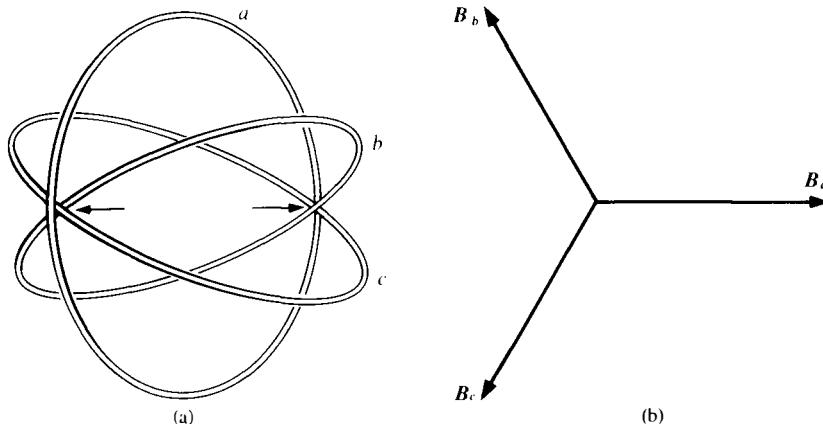


Fig. 18-13.

- (c) Check the *Fabry equation*, which states that at the center of any solenoid

$$B = G \left(\frac{P\lambda\sigma}{R_1} \right)^{1/2}.$$

Here G depends on the geometry, P is the dissipated power, $\lambda = n\pi r^2$ is the *filling factor*, or the fraction of the coil cross section occupied by the conductor, r is the radius of the wire, and σ the resistivity.

18-8. (18.2.1) A short, thick solenoid

Figure 18-14 shows the cross section of a coil. The dimensions shown are in millimeters. The wire has a square cross section of 2 millimeters² and a resistance of 8.93 ohms/kilometer. The current is 1 ampere. See the preceding problem.

- (a) Calculate B at the center. Use the formulas given in Prob. 18-7.
- (b) Calculate the power and the applied voltage.
- (c) Plot B as a function of z along the axis, from $z = -0.3$ to $z = 0.3$ meter.

18-9. (18.2.1) Helmholtz coils provide a uniform field

The *Helmholtz coils* of Fig. 18-15(a) provide a simple means of obtaining a uniform magnetic field over a given volume. Roughly speaking, B_z is uniform within 10% inside a sphere of radius $0.1a$.

- (a) Find B as a function of z along the axis.

If you have the patience to expand this expression about $z = 0$, you will find that

$$B = B_0 \left(1 - \frac{144z^4}{125a^4} + \dots \right).$$

This means that the first, second, and third derivatives of B with respect to z are zero at $z = 0$. So the curve of $B(z)$ is exceptionally flat near the center.

- (b) Plot $B/(\mu_0 NI/a)$ as a function of z/a from $z/a = -0.5$ to $z/a = 0.5$. Figure 18-16(b) shows $B_z(z)$ for values of r ranging up to $0.16a$.

18-10. (18.2.1) A Maxwell pair provides a uniform gradient of B

Plot $B/(\mu_0 NI/2a)$ as a function of z/a for a pair of coils like those of Fig.

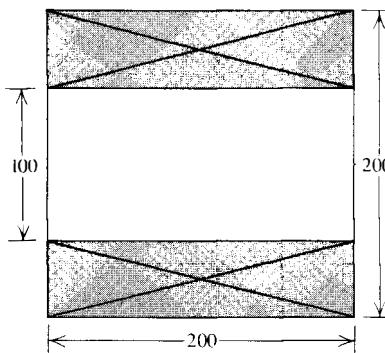


Fig. 18-14.

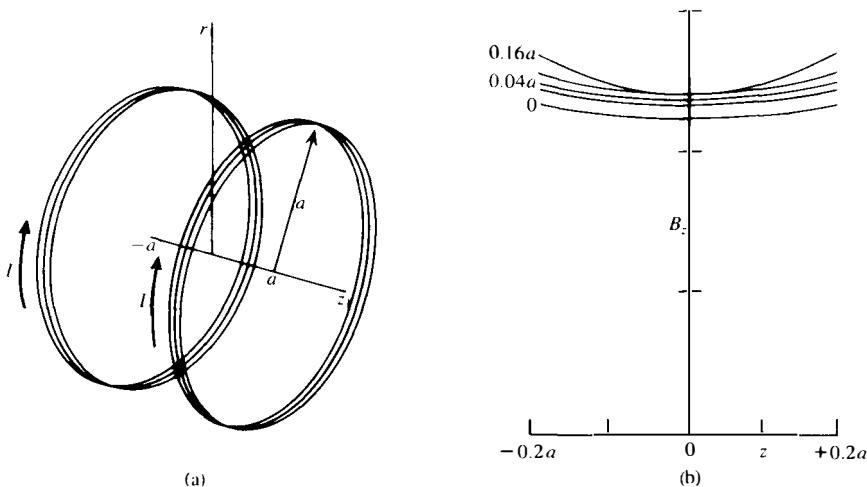


Fig. 18-15.

18-15(a) but with a spacing of $2a$, instead of a , and with currents flowing in opposite directions. This is a *Maxwell pair*. You will find that dB/dz is surprisingly linear between about $z = -0.7a$ and $z = 0.7a$.

18-11. (18.4) The vector potential \mathbf{A}

In a given region, $\mathbf{B} = B\hat{\mathbf{z}}$. Suggest possible \mathbf{A} 's and a characteristic of the corresponding current distribution.

18-12. (18.4) In two-dimensional magnetic fields a line of constant A is a line of \mathbf{B}

A certain magnetic field has a zero z -component.

(a) Show that $\mathbf{A} = A\hat{\mathbf{z}}$ is one possible value of \mathbf{A} .

(b) Show that a line of constant A is a line of \mathbf{B} .

(c) Show that this applies to the field of a straight current-carrying wire.

18-13. (18.5) The magnetic field of a spinning electrically charged sphere

A conducting sphere of radius R is charged to a potential V and spun about a diameter at an angular velocity ω .

(a) Show that the surface current density is $\alpha = \epsilon_0 \omega V \sin \theta = M \sin \theta$, where M is $\epsilon_0 \omega V$.

(b) Find that the magnetic flux density B_0 at the center.

(c) What is the numerical value of the B_0 for a sphere 100 millimeters in radius, charged to 10.0 kilovolts, and spinning at 10,000 turns per minute?

(d) Show that the dipole moment is $\frac{4}{3}\pi R^3 M\hat{\mathbf{z}}$, where $\hat{\mathbf{z}}$ is a unit vector along the axis, related to the direction of rotation by the right-hand screw rule.

(e) What is the dipole moment of the above rotating sphere?

(f) What current flowing through a loop 100 millimeters in diameter would have the same dipole moment?

CHAPTER 19

MAGNETIC FIELDS II

The Vector Potential \mathbf{A} . Ampère's Circuital Law

19.1 THE LINE INTEGRAL OF $\mathbf{A} \cdot d\mathbf{l}$ AROUND A CLOSED CURVE	348
<i>Example: THE VECTOR POTENTIAL \mathbf{A} IN FIELD OF A LONG SOLENOID</i>	349
19.2 THE LAPLACIAN OF \mathbf{A}	351
19.3 THE DIVERGENCE OF \mathbf{A}	351
19.4 THE CURL OF \mathbf{B}	352
19.5 AMPÈRE'S CIRCUITAL LAW	352
<i>Example: LONG CYLINDRICAL CONDUCTOR</i>	353
<i>Example: THE LONG SOLENOID</i>	355
<i>Example: THE REFRACTION OF LINES OF \mathbf{B} AT A CURRENT SHEET</i>	356
19.6 THE LAPLACIAN OF \mathbf{B}	357
19.7 SUMMARY	357
PROBLEMS	358

In this second chapter on magnetic fields, we first derive a direct consequence of the definition of the vector potential \mathbf{A} : the line integral of $\mathbf{A} \cdot d\mathbf{l}$ over a closed curve is equal to the encircled magnetic flux. This result is general. However, the rest of the chapter applies only to static fields. The expressions that we shall find here for $\nabla^2 \mathbf{A}$, $\nabla \cdot \mathbf{A}$, $\nabla \times \mathbf{B}$, and $\nabla^2 \mathbf{B}$ are all truncated: they all lack time-dependent terms. It is only in Chap. 27 that we shall find the full-fledged expressions.

19.1 THE LINE INTEGRAL OF $\mathbf{A} \cdot d\mathbf{l}$ AROUND A CLOSED CURVE

Consider first a simple closed curve, as in Fig. 19-1(a). The line integral of $\mathbf{A} \cdot d\mathbf{l}$ around C is equal to the magnetic flux linking C :

$$\oint_C \mathbf{A} \cdot d\mathbf{l} = \int_{\mathcal{A}} (\nabla \times \mathbf{A}) \cdot d\mathcal{A} = \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} = \Phi, \quad (19-1)$$

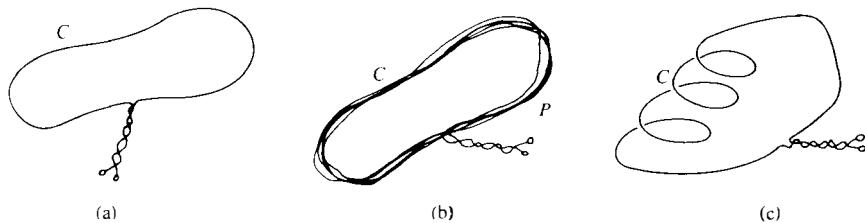


Fig. 19-1. (a) A simple closed circuit C . (b) An N -turn coil. The turns are all close together. (c) A more complex closed circuit.

where \mathcal{A} is the area of any surface bounded by C . We have used Stokes's theorem.

Now suppose the coil has N turns wound close together, as in Fig. 19-1(b). Over any cross section of the coil, say at P , the various turns are all exposed to approximately the same \mathbf{A} . Then

$$\oint_C \mathbf{A} \cdot d\mathbf{l} = N \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} = N\Phi = \Lambda, \quad (19-2)$$

where Λ is the *flux linkage* and \mathcal{A} is the area of any surface bounded by the coil.

The unit of flux linkage is the *weber turn*.

What if one has a circuit such as that of Fig. 19-1(c)? Then

$$\oint_C \mathbf{A} \cdot d\mathbf{l} = \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} = \Lambda, \quad (19-3)$$

except that now it is difficult to devise a surface bounded by C . Luckily enough, this surface is of no interest because the flux linkage Λ is easily measurable (Sec. 24.2).

Example

THE VECTOR POTENTIAL \mathbf{A} IN THE FIELD OF A LONG SOLENOID

Let us first see, qualitatively, how \mathbf{A} varies with position, both inside and outside a long solenoid. Remember the integral for \mathbf{A} that we found in Sec. 18.4:

$$\mathbf{A} = \frac{\mu_0 I}{4\pi} \oint_C \frac{d\mathbf{l}'}{r}. \quad (19-4)$$

At a point on the axis of the solenoid, the $d\mathbf{A}$ of an element $I d\mathbf{l}'$ situated somewhere on the solenoid cancels the $d\mathbf{A}$ of the element $I d\mathbf{l}'$ situated diametrically opposite the first one. On the axis of a long solenoid, \mathbf{A} is therefore zero.

At a point P inside the solenoid, but off the axis, the elements Idl' closest to P will contribute most. So \mathbf{A} is azimuthal, as in Fig. 19-2, and it increases with the radius r .

At a point P' outside the solenoid, \mathbf{B} is zero, as we shall see in the second example in Sec. 19.5. But \mathbf{A} is clearly not zero because the elements Idl' closest to P' contribute most.

A vector potential \mathbf{A} can therefore exist in a region where \mathbf{B} is zero. This simply means that $\nabla \times \mathbf{A} = 0$ for $\mathbf{A} \neq 0$, which is perfectly sensible. For example, if $\mathbf{A} = k\hat{\mathbf{r}}$, where k is independent of the coordinates, then $\nabla \times \mathbf{A} = 0$. We are already familiar with a similar situation in electrostatics: V can take any uniform value in a region where $\mathbf{E} = -\nabla V = 0$.

Outside a long solenoid, the vector \mathbf{A} is again azimuthal, but now it decreases with r as in the figure.

Now let us calculate \mathbf{A} .

First, consider the field inside the solenoid. At a point remote from the ends, \mathbf{B} is parallel to the axis, uniform, and equal to $\mu_0 N' I$, where N' is the number of turns per meter (Sec. 18.2.4). Then, from Sec. 19.1 and at a radius r as in the figure,

$$2\pi r A_i = \pi r^2 \mu_0 N' I, \quad A_i = \frac{\mu_0 N' I r}{2}. \quad (19-5)$$

Outside a solenoid of radius R , at the radius r' ,

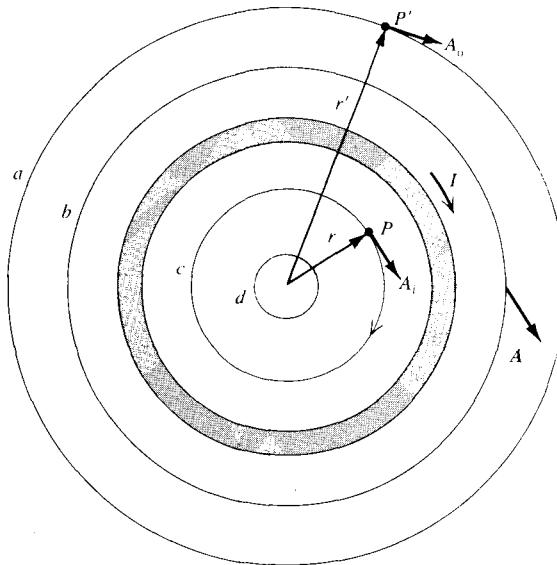


Fig. 19-2. Long solenoid seen endwise and lines of \mathbf{A} inside and outside. The magnitude of \mathbf{A} is proportional to r inside and inversely proportional to r outside.

$$2\pi r' A_0 = \pi R^2 \mu_0 N' I, \quad A_0 = \frac{\mu_0 N' I R^2}{2r'}. \quad (19-6)$$

Think how laborious it would be to calculate \mathbf{A} by integrating $I \mathbf{dl}/r$ over the winding!

19.2 THE LAPLACIAN OF \mathbf{A}

You will recall from Secs. 3.4.1 and 4.1 that

$$V = \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\rho}{r} dv', \quad \nabla^2 V = -\frac{\rho}{\epsilon_0}. \quad (19-7)$$

The first equation relates the potential V at the point $P(x, y, z)$ to the complete charge distribution, ρ being the total volume charge density at $P'(x', y', z')$ and r the distance PP' . The second equation expresses the relation between the *space derivatives* of V at any point to the volume charge density ρ *at that point*.

There exists an analogous pair of equations for the vector potential \mathbf{A} . We have already found the integral for \mathbf{A} in Sec. 18.4:

$$\mathbf{A} = \frac{\mu_0}{4\pi} \int_{v'} \frac{\mathbf{J}}{r} dv', \quad (19-8)$$

where v' is any volume enclosing all the currents. The x component of this equation is

$$A_x = \frac{\mu_0}{4\pi} \int_{v'} \frac{J_x}{r} dv'. \quad (19-9)$$

Then, by analogy with Eq. 19-7,

$$\nabla^2 A_x = -\mu_0 J_x. \quad (19-10)$$

Of course, similar equations apply to the y - and z -components, and

$$\nabla^2 \mathbf{A} = -\mu_0 \mathbf{J}. \quad (19-11)$$

This equation applies only to *static* fields.

19.3 THE DIVERGENCE OF \mathbf{A}

We can prove that, for *static* fields and for currents of finite extent, the divergence of \mathbf{A} is zero. First,

$$\nabla \cdot \mathbf{A} = \nabla \cdot \frac{\mu_0}{4\pi} \int_{v'} \frac{\mathbf{J}}{r} dv' = \frac{\mu_0}{4\pi} \int_{v'} \nabla \cdot \left(\frac{\mathbf{J}}{r} \right) dv', \quad (19-12)$$

where the del operator acts on the unprimed coordinates (x, y, z) of the field point, while \mathbf{J} is a function of the source point (x', y', z') . The integral operates on the primed coordinates. As usual, r is the distance between these two points, and the integration covers any volume enclosing all the currents.

We now use successively Identities 15, 16, and 6 from the back of the front cover:

$$\nabla \cdot \mathbf{A} = \frac{\mu_0}{4\pi} \int_{v'} \left(\nabla \frac{1}{r} \right) \cdot \mathbf{J} dv' = - \frac{\mu_0}{4\pi} \int_{v'} \left(\nabla' \frac{1}{r} \right) \cdot \mathbf{J} dv' \quad (19-13)$$

$$= \frac{\mu_0}{4\pi} \int_{v'} \left(-\nabla' \cdot \frac{\mathbf{J}}{r} + \frac{\nabla' \cdot \mathbf{J}}{r} \right) dv'. \quad (19-14)$$

In a time-independent field, $\partial\rho/\partial t = 0$ and, from the conservation of charge (Sec. 4.2), $\nabla' \cdot \mathbf{J} = 0$. Therefore

$$\nabla \cdot \mathbf{A} = - \frac{\mu_0}{4\pi} \int_{v'} \nabla' \cdot \frac{\mathbf{J}}{r} dv' = - \frac{\mu_0}{4\pi} \int_{\mathcal{A}'} \frac{\mathbf{J}}{r} \cdot d\mathcal{A}' \equiv 0, \quad (19-15)$$

where \mathcal{A}' is the area of the surface enclosing the volume v' . We have used the divergence theorem to transform the first integral into the second. The second integral is zero because, over \mathcal{A}' , \mathbf{J} is either zero or tangential.

19.4 THE CURL OF \mathbf{B}

From Definitions 5, 10, and 15 on the back of the front cover

$$\nabla \times \mathbf{B} = \nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}. \quad (19-16)$$

Thus, from Secs. 19.2 and 19.3,

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}. \quad (19-17)$$

This equation is valid only for *static fields*.

19.5 AMPÈRE'S CIRCUITAL LAW

The line integral of $\mathbf{B} \cdot d\mathbf{l}$ around a closed curve C is important:

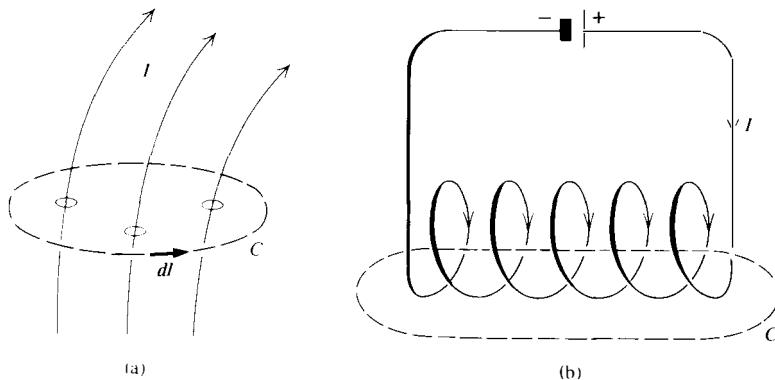


Fig. 19-3. (a) Closed path of integration C linked by a current I . Ampère's circuital law states that the line integral of $\mathbf{B} \cdot d\mathbf{l}$ over C is equal to $\mu_0 I$. (b) Here the line integral of $\mathbf{B} \cdot d\mathbf{l}$ over the dashed curve is equal to $6\mu_0 I$.

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \int_S (\nabla \times \mathbf{B}) \cdot d\mathcal{A} = \mu_0 \int_S \mathbf{J} \cdot d\mathcal{A} = \mu_0 I. \quad (19-18)$$

In this set of equations we first used Stokes's theorem, \mathcal{A} being the area of any surface bounded by C . Then we used the relation $\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$ that we found above. Finally, I is the net current that crosses any surface bounded by the closed curve C . The right-hand screw rule applies to the direction of I and to the direction of integration around C , as in Fig. 19-3(a).

This is *Ampère's circuital law*: the line integral of $\mathbf{B} \cdot d\mathbf{l}$ around a closed curve C is equal to μ_0 times the current linking C . This result is again valid only for constant fields.

Sometimes the same current crosses the surface bounded by C several times. For example, with a solenoid, the closed curve C could follow the axis and return outside the solenoid, as in Fig. 19-3(b). The total current linking C is then the current in one turn, multiplied by the number of turns, or the number of *ampere-turns*.

The circuital law can be used to calculate \mathbf{B} , when \mathbf{B} is uniform along the path of integration. This law is analogous to Gauss's law, which we used to calculate an \mathbf{E} that is uniform over a surface.

Example | LONG CYLINDRICAL CONDUCTOR

Let us apply Ampère's circuital law to calculate \mathbf{B} inside and outside the long, straight cylindrical conductor of Fig. 19-4 carrying a current I uniformly distributed over its cross section.

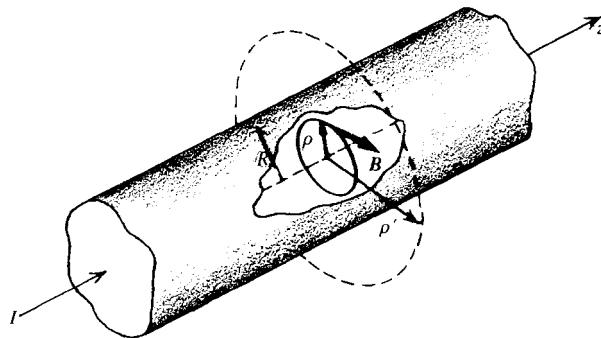


Fig. 19-4. Long cylindrical conductor of circular cross section carrying a current I . The circles are paths of integration for calculating \mathbf{B} .

We use cylindrical coordinates with the z -axis along the conductor.

Outside the conductor, \mathbf{B} is azimuthal and independent of ϕ . Then

$$B = \frac{\mu_0 I}{2\pi\rho}. \quad (19-19)$$

Inside the conductor, for a circuital path of radius ρ ,

$$B = \mu_0 \frac{[I/(\pi R^2)]\pi\rho^2}{2\pi\rho} = \frac{\mu_0 I \rho}{2\pi R^2}. \quad (19-20)$$

See Fig. 19-5.

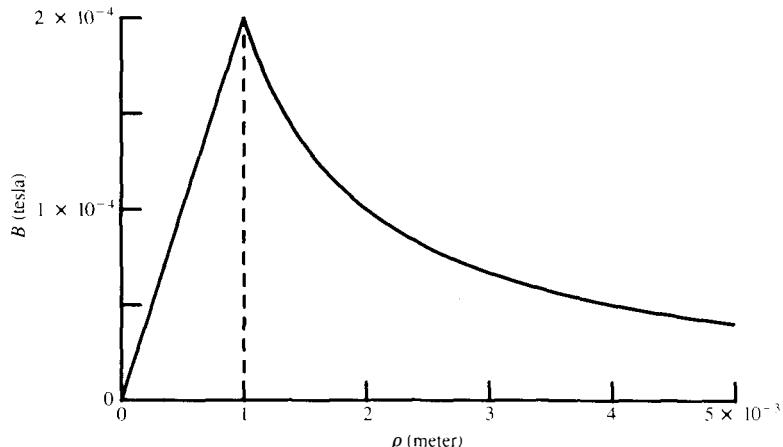


Fig. 19-5. B as a function of ρ for a wire 1 millimeter in radius carrying a current of 1 ampere.

Example

THE LONG SOLENOID

We return to the long solenoid and recalculate \mathbf{B} inside, in a region remote from the ends. This renders end effects negligible. See Fig. 19-6. We assume that the pitch of the winding is small. We again use cylindrical coordinates. The figure shows a solenoid of circular cross section. However, our main conclusions will be valid for any cross section.

First note that \mathbf{B} possesses the following general characteristics:

- (1) By symmetry, \mathbf{B} is everywhere independent of z and of ϕ .
- (2) Imagine an axial cylinder as in the figure. Its radius is either smaller or larger than that of the solenoid. The integrals of $\mathbf{B} \cdot d\mathcal{A}$ over the end faces cancel. Then the integral of $\mathbf{B} \cdot d\mathcal{A}$ over the cylindrical surface is zero, from Gauss's law for \mathbf{B} (Sec. 18.3). Then, both inside and outside, $B_\rho = 0$.
- (3) The curl of \mathbf{B} is zero everywhere except inside the wire, where $\mathbf{J} \neq 0$. Then, from the expression for the curl in cylindrical coordinates, outside the wire, $\partial B_z / \partial \rho = 0$. By symmetry, $\partial B_z / \partial \phi$ is also zero. Then B_z is uniform inside the solenoid and it is also uniform outside, neglecting end effects.

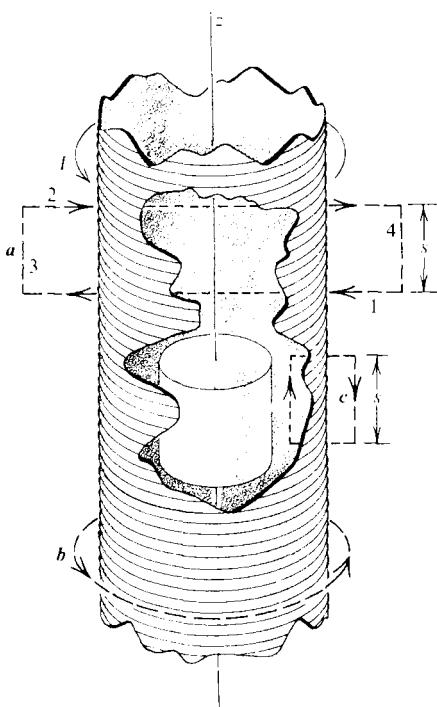


Fig. 19-6. Long solenoid carrying a current I , with paths of integration a , b , c .

Now consider the field outside the solenoid.

(1) We can show that $B_z = 0$ by considering path *a* in the figure. The net current linking this path is zero, and the line integral of $\mathbf{B} \cdot d\mathbf{l}$ around it is therefore also zero. Now, since the line integrals along sides 1 and 2 are zero ($B_\rho = 0$), the line integrals along sides 3 and 4 cancel. But sides 3 and 4 can each be situated at any distance from the solenoid, so B_z is either zero, or nonzero and independent of ρ . Now the flux outside is equal to the finite flux inside. Therefore, *outside*, B_z tends to zero.

(2) A path such as *b* is linked once by the current. Thus, *outside the solenoid*, $B_\phi = \mu_0 I / 2\pi\rho$. This flux is usually negligible.

Now let us look inside the solenoid.

(1) There is zero flux in the ϕ -direction inside because the line integral of $B_\phi dl$ over a circle of radius ρ , say the top edge of the small cylinder shown in the figure, is $2\pi\rho B_\phi$; and this is zero, according to Sec. 19.5 because the path encloses zero current.

(2) Consider now path *c* in the figure. Remembering that $B_\rho = 0$ both inside and outside, and that $B_z = 0$ outside, we see that, if there are N' turns/meter, $B_z s = \mu_0 N' / I_s$ and $B_z = \mu_0 N' I / L$.

Example

THE REFRACTION OF LINES OF \mathbf{B} AT A CURRENT SHEET

A thin conducting sheet carries a surface current density of α amperes/meter, as in Fig. 19-7. In passing through the sheet, the lines of \mathbf{B} bend as follows.

Since $\nabla \cdot \mathbf{B} = 0$, the normal component of \mathbf{B} is the same on the two sides: $B_{1n} = B_{2n}$.

Applying Ampère's circuital law to the path of length L shown in the figure,

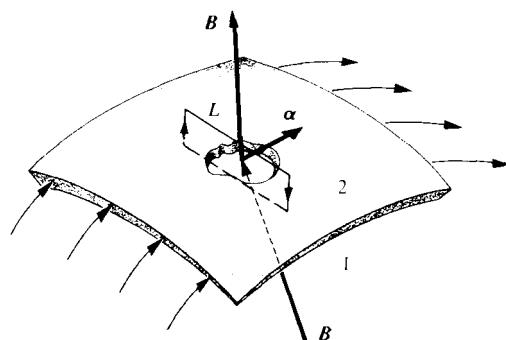


Fig. 19-7. Refraction of a line of \mathbf{B} crossing a current sheet.

$$B_{1t}L - B_{2t}L = \mu_0\alpha L, \quad B_{2t} = B_{1t} - \mu_0\alpha. \quad (19-21)$$

A line of \mathbf{B} therefore deflects in the clockwise direction for an observer looking in the direction of α .

We can also arrive at this result in another way. The magnetic flux density \mathbf{B} results from the existence of a current in the sheet *and* to currents flowing elsewhere. According to Ampère's circuital law, the magnetic field of the sheet, just below the sheet, is $\mu_0\alpha/2$ and points left. Just above, the field is again $\mu_0\alpha/2$, but it points right. Adding this field to that of the other currents leads to tangential components that differ as above.

19.6 THE LAPLACIAN OF \mathbf{B}

We can deduce the value of the Laplacian of \mathbf{B} from that of the Laplacian of \mathbf{A} (Sec. 19.2). Since

$$\nabla^2 \mathbf{A} = -\mu_0 \mathbf{J}, \quad (19-22)$$

then

$$\nabla \times (\nabla^2 \mathbf{A}) = -\mu_0 \nabla \times \mathbf{J}. \quad (19-23)$$

Now the curl of a Laplacian is equal to the Laplacian of a curl and thus

$$\nabla^2(\nabla \times \mathbf{A}) = -\mu_0 \nabla \times \mathbf{J}. \quad (19-24)$$

Finally,

$$\nabla^2 \mathbf{B} = -\mu_0 \nabla \times \mathbf{J}, \quad (19-25)$$

again for *static* fields.

19.7 SUMMARY

The line integral of $\mathbf{A} \cdot d\mathbf{l}$ around a closed curve C is equal to the magnetic flux linking C :

$$\oint_C \mathbf{A} \cdot d\mathbf{l} = \int_A \mathbf{B} \cdot d\mathcal{A} = \Lambda, \quad (19-3)$$

where \mathcal{A} is the area of a surface bounded by C .

For *static* fields,

$$\nabla^2 \mathbf{A} = -\mu_0 \mathbf{J}, \quad (19-11)$$

$$\nabla \cdot \mathbf{A} = 0, \quad (19-15)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}, \quad (19-17)$$

$$\nabla^2 \mathbf{B} = -\mu_0 \nabla \times \mathbf{J}. \quad (19-25)$$

Ampère's circuital law states that

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_0 I, \quad (19-18)$$

where I is the net current that crosses any open surface bounded by the curve C , in the direction given by the right-hand screw rule.

PROBLEMS

19-1. (19.1) The vector potential inside a current-carrying conductor

Show that, inside a straight current-carrying conductor of radius R ,

$$A = \frac{\mu_0 I}{4\pi} \left(1 - \frac{\rho^2}{R^2} \right),$$

if A is set equal to zero at $\rho = R$.

19-2. (19.5) Van de Graaff high-voltage generator

In a Van de Graaff generator, a charged insulating belt transports electric charge to the high-voltage electrode.

(a) Calculate the current carried by a 500-millimeter-wide belt driven by a 100-millimeter-diameter pulley that rotates at 60 revolutions/second. if $E = 2 \times 10^6$ volts/meter at the surface of the belt.

(b) Calculate the B close to the belt.

19-3. (19.5) The toroidal coil

Figure 19-8 shows a toroidal coil of square cross section. There are N turns, and the current is I .

Find (a) the azimuthal field along paths a and c , (b) B inside the toroid, and (c) the line integral of B along path d .

19-4. (19.5) \mathbf{B} near a conducting sheet

A conducting sheet carries a current density of α amperes/meter. There are no other currents in the vicinity.

(a) What is the value of B , close to the sheet?

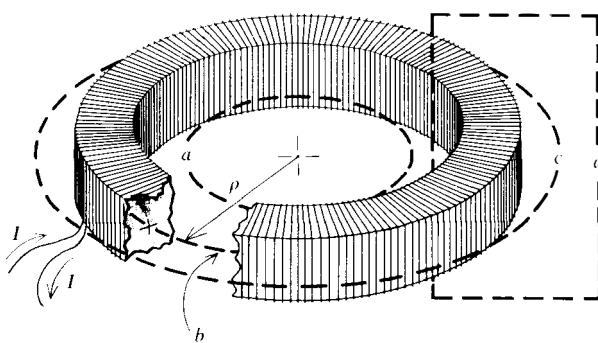


Fig. 19-8.

- (b) How is \mathbf{B} oriented with respect to α ?
(c) A conducting body carries a high-frequency current that is confined near the surface. The surface current density is α amperes/meter. Show that, in the air near the conductor, $\mathbf{B} = \mu_0 \alpha \times \hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is a unit vector normal to the surface and pointing outward.
- 19-5. (19.5) The magnetic field near the axis of a circular loop
A circular loop carries a current I . Choose the axis of symmetry as the z -axis, and calculate B_ρ and B_z near the axis.
- 19-6. (19.5) The average \mathbf{B} over a sphere is equal to \mathbf{B} at the center
Refer to Prob. 3-16 concerning the average \mathbf{E} over a spherical volume.
Show that, in a region where there are no currents, the average \mathbf{B} over a spherical volume is equal to the \mathbf{B} at the center.
- 19-7. (19.5) The field of a short thick solenoid, compared to that of a long solenoid
The value of B at the center of a short, thick solenoid given in Prob. 18-7 can be written as $B = \mu_0 N' I g$, where $\mu_0 N' I$ is the field of a long solenoid. Find g .

CHAPTER 20

MAGNETIC FIELDS III

Magnetic Materials A: The Magnetic Flux Density \mathbf{B} and the Magnetic Field Strength \mathbf{H}

20.1	TYPES OF MAGNETIC MATERIAL	361
20.2	THE MAGNETIZATION \mathbf{M}	361
20.3	THE MAGNETIC FIELD OF A MAGNETIZED BODY	362
	<i>Example: THE EQUIVALENT CURRENTS IN A UNIFORMLY MAGNETIZED ROD</i>	364
20.4	THE DIVERGENCE OF \mathbf{B} IN THE PRESENCE OF MAGNETIC MATERIAL	364
20.5	THE MAGNETIC FIELD STRENGTH \mathbf{H} , THE CURL OF \mathbf{H}	365
20.5.1	DIELECTRIC AND MAGNETIC MATERIALS COMPARED	366
20.6	AMPÈRE'S CIRCUITAL LAW IN THE PRESENCE OF MAGNETIC MATERIAL	367
	<i>Example: SOLENOID WOUND ON A MAGNETIC CORE</i>	367
20.7	THE MAGNETIC SUSCEPTIBILITY χ_m AND THE RELATIVE PERMEABILITY μ_r	368
*20.7.1	THE MAGNETIZATION CURVE OF A FERROMAGNETIC MATERIAL	369
*20.7.2	FOUR DEFINITIONS OF THE RELATIVE PERMEABILITY μ_r	369
20.8	BOUNDARY CONDITIONS	370
20.9	SUMMARY	372
	PROBLEMS	373

Thus far we have studied magnetic fields associated with moving charges, although we did allude to the magnetic materials that we shall study in this chapter and the next.

All atoms contain spinning electrons that give rise to magnetic fields. It is our purpose in these two chapters to express these fields in macroscopic terms.

In dielectric materials, individual atoms or molecules can possess electric dipole moments which, when properly oriented, confer a net

electric moment to a macroscopic body. Magnetic materials are analogous in that their atoms can act as magnetic dipoles that can also be oriented. The body is then said to be *magnetized*. Magnetic effects are weak in all but ferromagnetic substances, and those are grossly nonlinear.

This first chapter on magnetic materials sets forth a few fundamental ideas. Then, in the end, it provides a glimpse of the complexities of ferromagnetism.

Chapter 21 concerns ferromagnetism and some of the methods that are available for designing devices incorporating ferromagnetic materials.

20.1 TYPES OF MAGNETIC MATERIAL

There exist three main types of magnetic material.

- (1) All materials are *diamagnetic*. This magnetism originates from the fact that the application of an external magnetic field induces moments according to the Faraday induction law (Sec. 23.4). This effect is usually imperceptible, and it disappears upon removal of the external field.
- (2) In most atoms the magnetic moments resulting from the orbital and spinning motions of the electrons cancel. If the cancellation is not complete, the material is *paramagnetic*. Thermal agitation causes the individual moments to be randomly oriented, but the application of a magnetic field brings about a partial orientation.
- (3) In *ferromagnetic* materials such as iron, the magnetization can be orders of magnitude larger than in either diamagnetic or paramagnetic substances. This effect comes from electron spin, together with group phenomena that align the moments throughout a small region called a *domain*.

20.2 THE MAGNETIZATION \mathbf{M}

The *magnetization* \mathbf{M} is the magnetic moment per unit volume of magnetized material at a point. If there are N atoms per unit volume, each possessing a magnetic dipole moment \mathbf{m} oriented in a given direction, then

$$\mathbf{M} = N\mathbf{m}. \quad (20-1)$$

The magnetization \mathbf{M} in magnetic media corresponds to the polarization \mathbf{P} in dielectrics. The unit of magnetization is the ampere per meter.

20.3 THE MAGNETIC FIELD OF A MAGNETIZED BODY

To find the magnetic field of a magnetized body, we shall integrate the expression for the vector potential \mathbf{dA} of a magnetic dipole $\mathbf{M} d\mathbf{v}$ over the volume of the material. This calculation will show that the field is the same as if we had an equivalent volume current density $\nabla \times \mathbf{M}$, plus an equivalent surface current density $\mathbf{M} \times \hat{\mathbf{n}}$, situated in a vacuum.

You will recall that we arrived at a similar situation when we discussed dielectrics: the electric field of a polarized dielectric is the same as if we had volume and surface charge distributions ρ_b and σ_b situated in a vacuum.

This result is interesting, but it is of no use for calculating magnetic fields because \mathbf{M} is itself a function of \mathbf{B} , and unknown.

In practice, one calculates \mathbf{B} either by rather crude, semiempirical methods such as those of Chap. 21 or by means of elaborate computer codes.

We shall find that the equivalent volume current density \mathbf{J}_e is equal to $\nabla \times \mathbf{M}$. Thus $\nabla \cdot \mathbf{J}_e$ is zero and charge cannot accumulate at a point by virtue of \mathbf{J}_e . Furthermore, the equivalent currents do not dissipate energy because they do not involve electron drift and scattering processes like conduction currents.

Let us calculate \mathbf{B} at a point outside a magnetized body, as in Fig. 20-1.

From the third example in Sec. 18.4, the vector potential at a point P located at a distance r from a current loop of magnetic moment \mathbf{m} is

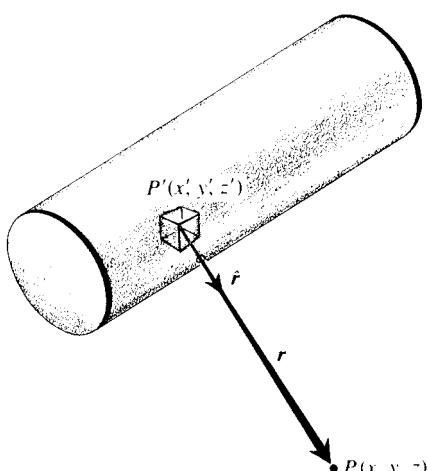


Fig. 20-1. Element of volume inside a magnetized body and an external point P .

$$\mathbf{A} = \frac{\mu_0}{4\pi} \frac{\mathbf{m} \times \hat{\mathbf{r}}}{r^2}. \quad (20-2)$$

The unit vector $\hat{\mathbf{r}}$ points in the direction of P , from the center of the loop, and r is large compared to the largest dimension of the loop. Then, for a volume v' of magnetized material,

$$\mathbf{A} = \frac{\mu_0}{4\pi} \int_{v'} \frac{\mathbf{M} \times \hat{\mathbf{r}}}{r^2} dv' = \frac{\mu_0}{4\pi} \int_{v'} \mathbf{M} \times \nabla' \left(\frac{1}{r} \right) dv'. \quad (20-3)$$

We have used Identity 15 from the back of the front cover. The volume of magnetized material is v' , and its surface has an area \mathcal{A}' .

Then, from Identity 11 and from Stokes's theorem,

$$\mathbf{A} = -\frac{\mu_0}{4\pi} \int_{v'} \left(\nabla' \times \frac{\mathbf{M}}{r} \right) dv' + \frac{\mu_0}{4\pi} \int_{v'} \frac{\nabla' \times \mathbf{M}}{r} dv' \quad (20-4)$$

$$= \frac{\mu_0}{4\pi} \int_{\mathcal{A}'} \frac{\mathbf{M} \times \hat{\mathbf{n}}}{r} d\mathcal{A}' + \frac{\mu_0}{4\pi} \int_{v'} \frac{\nabla' \times \mathbf{M}}{r} dv', \quad (20-5)$$

$\hat{\mathbf{n}}$ being the unit vector normal to the surface of area \mathcal{A} of the magnetized material and pointing *outward*. We may omit the prime on the del that operates on \mathbf{M} , since that del clearly operates on the coordinates x' , y' , z' of the point where the magnetization is \mathbf{M} .

These expressions for \mathbf{A} are all equivalent, but the last one lends itself to a simple physical interpretation. It is clear that the vector potential in the neighborhood of a piece of magnetized material is the same as if one had, instead, volume and surface densities

$$\mathbf{J}_e = \nabla' \times \mathbf{M} \quad \text{and} \quad \boldsymbol{\alpha}_e = \mathbf{M} \times \hat{\mathbf{n}}. \quad (20-6)$$

Inside a magnetized material none of the above integrals diverge, and the field is again the same as if the magnetized material were replaced by its equivalent currents.

More generally,[†]

[†] The question is sometimes raised as to whether the term $\epsilon_0 \partial \mathbf{E} / \partial t$ should be included under the integral sign. Then one would replace the polarization current density by the displacement current density $\partial \mathbf{D} / \partial t$ (Sec. 9.10). The term $\epsilon_0 \partial \mathbf{E} / \partial t$ does *not* belong here because magnetic fields arise solely from the motion of charge. This is confirmed by the fact that one can calculate the field of a transmitting antenna in free space from the currents flowing in it, disregarding displacement currents in free space. See also Sec. 37.1.

$$\mathbf{A} = \frac{\mu_0}{4\pi} \int_{v'} \left(\mathbf{J}_f + \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M} \right) dv', \quad (20-7)$$

where \mathbf{J}_f is the current density of free charges and $\partial \mathbf{P}/\partial t$ is the polarization current density of Sec. 9.3.3.

Thus a more general form of the Biot-Savart law (Sec. 18.2) is

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int_{v'} \frac{(\mathbf{J}_f + \partial \mathbf{P}/\partial t + \nabla \times \mathbf{M}) \times \hat{\mathbf{r}}}{r^2} dv'. \quad (20-8)$$

Example

THE EQUIVALENT CURRENTS IN A UNIFORMLY MAGNETIZED ROD

Suppose the magnetization \mathbf{M} is uniform and parallel to the axis. This is an idealized situation because the elementary dipoles tend to orient themselves along \mathbf{B} , which is only approximately axial. Since \mathbf{M} is uniform, $\nabla \times \mathbf{M} = 0$, and there are no equivalent volume currents. Also, since \mathbf{M} is parallel to the axis, the current density on the cylindrical surface is M , in the direction shown in Fig. 20-2, and there are no currents on the end faces. The rod therefore acts as a solenoid with $N'I = M$, where N' is the number of turns per meter. Observe that, inside, \mathbf{B} and \mathbf{M} point in the same general direction.

20.4 THE DIVERGENCE OF \mathbf{B} IN THE PRESENCE OF MAGNETIC MATERIAL

Magnetic fields originate either in the macroscopic motion of charge or in equivalent currents. The relation

$$\boxed{\nabla \cdot \mathbf{B} = 0} \quad (20-9)$$

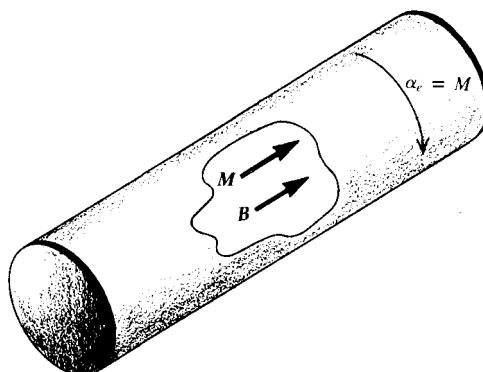


Fig. 20-2. A uniformly magnetized rod acts as a solenoid carrying a surface current density $\alpha_e = M$.

that we found in Sec. 18.3 applies even in the presence of magnetic materials. This is one of Maxwell's equations.

20.5 THE MAGNETIC FIELD STRENGTH \mathbf{H} . THE CURL OF \mathbf{H}

In Sec. 19.4 we found that, for static fields in the absence of magnetic materials,

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}_f. \quad (20-10)$$

Henceforth we shall use \mathbf{J}_f , instead of the unadorned \mathbf{J} , for the current density related to the motion of free charges.

In the presence of magnetized materials,

$$\nabla \times \mathbf{B} = \mu_0 (\mathbf{J}_f + \mathbf{J}_e). \quad (20-11)$$

This equation, of course, applies only in regions where the space derivatives exist, that is, inside magnetized materials, but not at their surfaces. Then

$$\nabla \times \mathbf{B} = \mu_0 (\mathbf{J}_f + \nabla \times \mathbf{M}), \quad (20-12)$$

$$\nabla \times \left(\frac{\mathbf{B}}{\mu_0} - \mathbf{M} \right) = \mathbf{J}_f. \quad (20-13)$$

The vector within the parentheses, whose curl equals the free current density, is the *magnetic field strength*:

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \mathbf{M}. \quad (20-14)$$

Both \mathbf{H} and \mathbf{M} are expressed in amperes/meter. Thus

$$\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M}) \quad (20-15)$$

and, even inside magnetized materials,

$$\nabla \times \mathbf{H} = \mathbf{J}_f \quad (20-16)$$

for *static* fields.

20.5.1 Dielectric and Magnetic Materials Compared

Compare the above equation for \mathbf{B} with the corresponding Eq. 9-20,

$$\mathbf{E} = \frac{1}{\epsilon_0} (\mathbf{D} - \mathbf{P}). \quad (20-17)$$

Note the difference in sign: minus \mathbf{P} instead of plus \mathbf{M} .

Figure 20-3 illustrates the difference. In Fig. 20-3(a) the capacitor plates carry free charges that are not affected by the presence of the dielectric, neglecting edge effects. So $D = \sigma$ is fixed. All three vectors point to the right. Without the dielectric, E would be equal to D/ϵ_0 . With the dielectric, E is *smaller* because the field of the bound charges *opposes* that of the free charges.

In Fig. 20-3(b), the coil applies a given \mathbf{H} . All three vectors point to the right again. Without the magnetic core, $B = \mu_0 H$. With the core, B is *larger* because the field of the equivalent currents *aids* that of the free currents.

Remember that we are concerned here solely with the space- and time-averaged fields inside matter. Remember also that the similarity between the fields of electric and magnetic dipoles exists only at points remote from the dipoles. Closer in, the fields are totally different, as shown in Fig. 18-10.

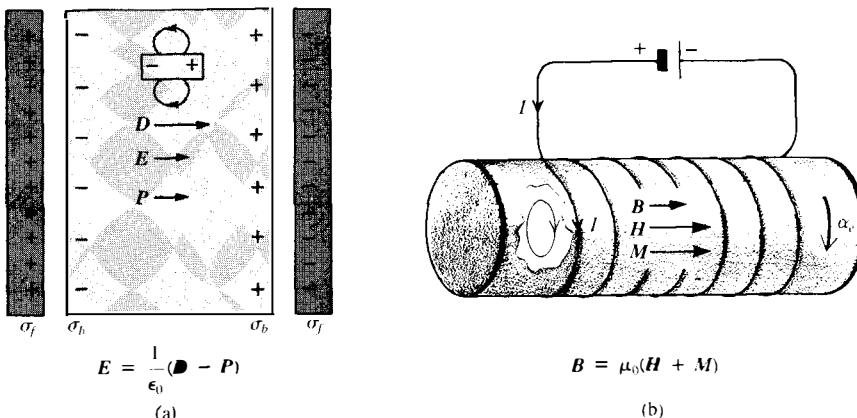


Fig. 20-3. (a) Plane-parallel capacitor. The plates carry fixed surface charge densities σ . Introducing the dielectric reduces E by the factor of ϵ_r , neglecting edge effects. Note the orientation of the small dipole and its field. (b) Solenoid carrying a fixed current I . Introducing the magnetic core increases B by a factor of μ_r , neglecting end effects. Note again the orientation of the small dipole and of its field.

20.6 AMPÈRE'S CIRCUITAL LAW IN THE PRESENCE OF MAGNETIC MATERIAL

Let us integrate Eq. 20-16 over an open surface of area \mathcal{A} bounded by a curve C :

$$\int_{\mathcal{A}} (\nabla \times \mathbf{H}) \cdot d\mathcal{A} = \int_{\mathcal{A}} \mathbf{J}_f \cdot d\mathcal{A}, \quad (20-18)$$

or, using Stokes's theorem on the left-hand side,

$$\oint_C \mathbf{H} \cdot d\mathbf{l} = I_f, \quad (20-19)$$

where I_f is the current of free charges linking C . The right-hand screw rule applies to the direction of integration and to the direction of z . Note that I_f does not include the equivalent currents. The term on the left is the *magnetomotance*.

This is a more general form of *Ampère's circuital law* of Sec. 19.5, in that it can serve to calculate \mathbf{H} even in the presence of magnetic materials. It is rigorously valid, however, only for steady currents.

Example SOLENOID WOUND ON A MAGNETIC CORE

Imagine a long solenoid wound on a magnetic core, as in Fig. 20-4. At points remote from the ends, \mathbf{H} is parallel to the axis inside the solenoid, and essentially zero outside. Then, from

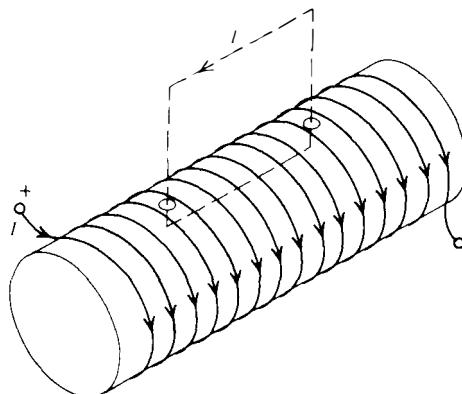


Fig. 20-4. Long solenoid wound on a magnetic core. According to Ampère's circuit law, $H = N'I$ inside the core.

Ampère's circuital law applied to the path shown, if the number of turns per meter is N' and the current is I ,

$$Hl = N'I l, \quad H = N'I \quad (20-20)$$

inside the core, whatever it is made of.

20.7 THE MAGNETIC SUSCEPTIBILITY χ_m AND THE RELATIVE PERMEABILITY μ_r

It is convenient to define a *magnetic susceptibility* χ_m such that[†]

$$\mathbf{M} = \chi_m \mathbf{H}. \quad (20-21)$$

Then

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}) = \mu_0(1 + \chi_m)\mathbf{H} = \mu_0\mu_r \mathbf{H} = \mu \mathbf{H}, \quad (20-22)$$

where

$$\mu_r = 1 + \chi_m \quad (20-23)$$

is the *relative permeability* and

$$\mu = \mu_r \mu_0 \quad (20-24)$$

is the *permeability* of a material. Both χ_m and μ_r are pure numbers.

Thus

$$\mathbf{M} = \chi_m \frac{\mathbf{B}}{\mu}. \quad (20-25)$$

The magnetic susceptibility of purely diamagnetic materials is negative and of the order of 10^{-5} . In paramagnetic materials, χ_m varies from about 10^{-5} to 10^{-3} . The susceptibility of ferromagnetic substances can be as large as 10^6 .

The equation $\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M})$ is general, but equations involving either μ_r or χ_m assume that the material is both isotropic and linear. In other words, they assume that \mathbf{M} is proportional to \mathbf{H} and in the same direction.

In ferromagnetic materials, \mathbf{B} and \mathbf{H} do not always point in the same direction, and when they do, μ_r can vary by orders of magnitude, depending on the value of \mathbf{H} and on the previous history of the material (Sec. 21.2). In a permanent magnet, \mathbf{B} and \mathbf{H} point in roughly *opposite* directions.

[†] Compare with $P = \chi_e \epsilon_0 E$ (Sec. 9.8).

*20.7.1 The Magnetization Curve of a Ferromagnetic Material

If an unmagnetized ferromagnetic material is subjected to a gradually increasing H , then B increases along a roughly S-shaped curve, as in Fig. 20-5. After a while, saturation sets in and M increases no further. For most ferromagnetic materials saturation occurs at about 2 teslas.

*20.7.2 Four Definitions of the Relative Permeability μ_r

The behavior of ferromagnetic materials is so complex that several different types of relative permeability have been defined. We mention only four. All definitions relate to the magnetization curve. Relative permeability is often loosely called permeability. This is what we do here.

- (1) The word *permeability*, left unqualified, simply means the order of magnitude of the ratio $B/\mu_0 H$ on the magnetization curve, possibly over some specified range of B or of H .
- (2) The meaning of the term *maximum permeability* is obvious.
- (3) The *initial permeability* is the ratio $B/\mu_0 H$ at very weak fields.

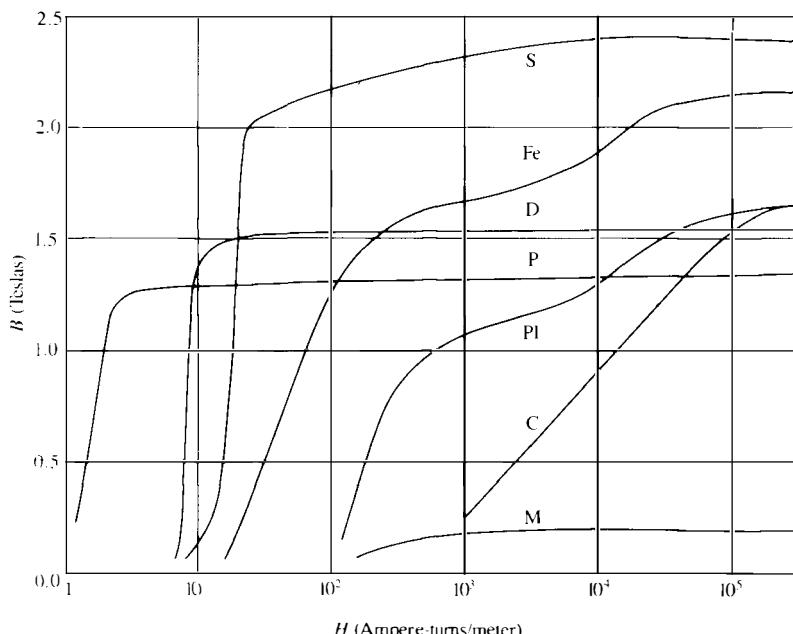


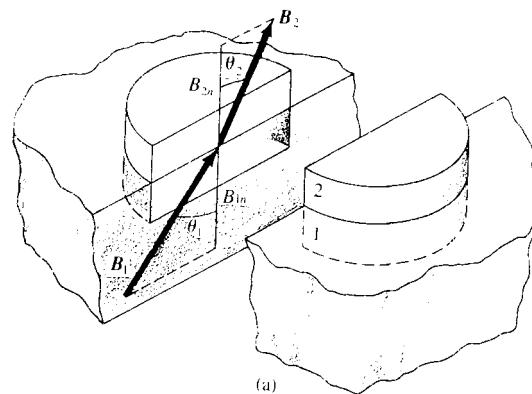
Fig. 20-5. Magnetization curves for various materials: S, Supermendur; Fe, annealed pure iron; D, Deltamax; P, Permalloy; PI, powdered iron; C, gray cast iron; M, Monel.

- (4) The *differential permeability* is the ratio $dB/\mu_0 dH$, or the slope of the magnetization curve divided by μ_0 , at a given point on a hysteresis loop.

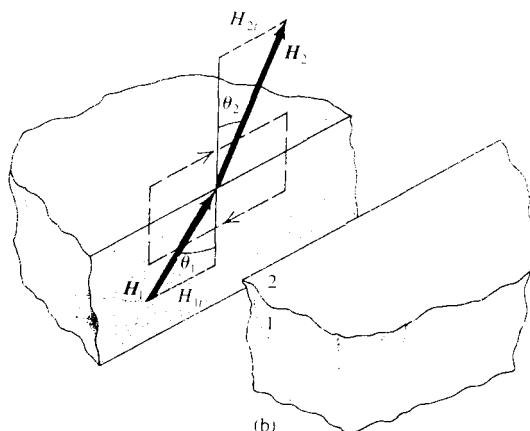
20.8 BOUNDARY CONDITIONS

Both \mathbf{B} and \mathbf{H} obey boundary conditions at the interface between two media. We proceed as in Sec. 10.2.

Figure 20-6(a) shows a short Gaussian volume at an interface. From Gauss's law, the flux leaving through the top equals that entering the bottom and



(a)



(b)

Fig. 20-6. (a) Gaussian surface straddling the interface between media 1 and 2. The normal components of the \mathbf{B} 's are equal. (b) Closed path piercing the interface. The tangential components of the \mathbf{H} 's are equal.

$$B_{1n} = B_{2n}. \quad (20-26)$$

The normal component of \mathbf{B} is therefore continuous across an interface.

Consider now Fig. 20-6(b). The small rectangular path pierces the interface. From the circuital law of Sec. 20.6, the line integral of $\mathbf{H} \cdot d\mathbf{l}$ around the path is equal to the current I linking the path. With the two long sides of the path infinitely close to the interface, I is zero and the tangential component of \mathbf{H} is continuous across the interface:

$$H_{1t} = H_{2t}. \quad (20-27)$$

These two equations are general.

Setting $\mathbf{B} = \mu \mathbf{H}$ for both media, the permeabilities being those that correspond to the actual fields, and assuming that the materials are isotropic, then the above two equations imply that

$$\frac{\tan \theta_1}{\tan \theta_2} = \frac{\mu_{r1}}{\mu_{r2}}. \quad (20-28)$$

We therefore have the following rule for linear and isotropic media: lines of \mathbf{B} lie farther away from the normal in the medium possessing the larger permeability. In other words, the lines “prefer” to pass through the more permeable medium, as in Fig. 20-7. You will recall from Sec. 10.2.4 that we had a similar situation with dielectrics.

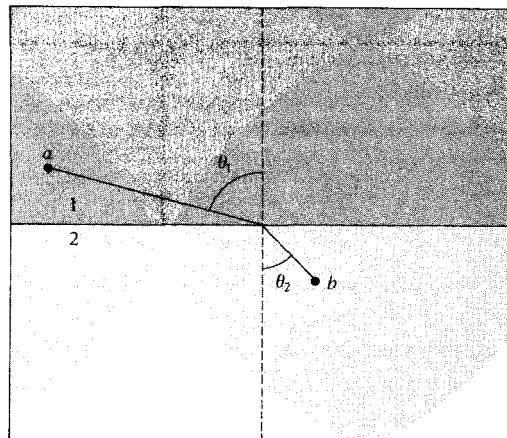


Fig. 20-7. Line of \mathbf{B} crossing the interface between linear and isotropic media 1 and 2. The permeability of medium 1 is larger than that of medium 2. Points a and b are at equal distances from the interface. The path in higher-permeability medium 1 is longer than the path in 2: the line of \mathbf{B} “prefers” to be in the higher-permeability material.

20.9 SUMMARY

The *magnetization* \mathbf{M} is the magnetic dipole moment per unit volume in magnetized material. The magnetic field, both outside and inside, is the same as if the material were replaced by its *equivalent volume and surface current densities*:

$$\mathbf{J}_e = \nabla \times \mathbf{M} \quad \text{and} \quad \mathbf{a}_e = \mathbf{M} \times \hat{\mathbf{n}}. \quad (20-6)$$

The *divergence of \mathbf{B}* is zero even in the presence of magnetic materials:

$$\boxed{\nabla \cdot \mathbf{B} = 0.} \quad (20-9)$$

The *magnetic field strength \mathbf{H}* has the same dimensions as \mathbf{M} , and

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \mathbf{M}, \quad \mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}), \quad (20-14), (20-15)$$

$$\nabla \times \mathbf{H} = \mathbf{J}_f, \quad (20-16)$$

where \mathbf{J}_f is the current density attributable to free charges. This last equation applies only to *static* fields. It follows that, for *static* fields,

$$\oint_C \mathbf{H} \cdot d\mathbf{l} = I_f, \quad (20-19)$$

where C is a closed curve linked by the current I_f . This is *Ampère's circuital law* in a more general form.

As with dielectrics, it is convenient to define a *magnetic susceptibility* χ_m and a *relative permeability* μ_r as follows

$$M = \chi_m H, \quad (20-21)$$

$$\mathbf{B} = \mu_0 \mu_r \mathbf{H}, \quad \mu_r = 1 + \chi_m. \quad (20-22), (20-23)$$

Then $\mu = \mu_r \mu_0$ is the *permeability*. In ferromagnetic materials the value of μ_r can vary by orders of magnitude, and even in sign, depending on the value of H and on the previous magnetic history of the sample.

At the interface between two media the normal component of \mathbf{B} and the tangential component of \mathbf{H} are continuous. There is *refraction of the lines of \mathbf{B}* at an interface, the larger angle with the normal being in the medium with the larger permeability.

PROBLEMS

- 20-1. (20.2) The number of Bohr magnetons per atom in iron

The magnetization M in iron can contribute as much as 2 teslas to B . If one electron contributes one Bohr magneton, how many electrons per atom, on average, can contribute to M ? See Prob. 18-5.

- 20-2. (20.2) The field of a disk magnet

A disk of iron of radius a and thickness s is magnetized parallel to its axis. Calculate B on the axis, outside the iron.

- 20-3. (20.2) The field inside a tubular magnet

You are asked to design a permanent magnet that would supply a magnetic field in a cylindrical volume about 20 millimeters in length and 20 millimeters in diameter. Someone suggests a tubular magnet, magnetized along its length, that would surround this volume. What do you think of this suggestion?

- 20-4. (20.5) The divergence of \mathbf{H} is not always zero

Show that $\nabla \cdot \mathbf{H}$ is not zero in a nonhomogeneous magnetic material.

- 20-5. (20.5) The field in a cavity inside magnetic material

Find B and H inside a thin, disk-shaped cavity whose axis is parallel to \mathbf{B} , inside magnetic material.

- 20-6. (20.7) The free and equivalent volume current densities

Show that, in a homogeneous, isotropic, and linear magnetic material,

$$\mathbf{J}_e = (\mu_r - 1)\mathbf{J}_f.$$

- 20-7. (20.7) Current-carrying wire along the axis of an iron tube

A wire carrying a current I is situated on the axis of a hollow iron cylinder.

(a) Find H , B , and M in the inner region, in the iron, and in the outer region.

(b) Find the equivalent currents.

- 20-8. (20.7) The magnetization in terms of χ_m and B

Show that in a linear and isotropic magnetic medium,

$$M = \frac{\chi_m B}{\mu_0(1 + \chi_m)}.$$

CHAPTER 21

MAGNETIC FIELDS IV[†]

Magnetic Materials B. Ferromagnetism and Magnetic Circuits

*21.1	FERROMAGNETIC MATERIALS	374
*21.2	HYSTERESIS	375
*21.2.1	ENERGY DISSIPATED IN A HYSTERESIS CYCLE	377
	<i>Example: HYSTERESIS LOSSES IN A TRANSFORMER IRON</i>	377
*21.3	MAGNETIC FIELD CALCULATIONS	377
*21.3.1	THE BAR MAGNET	377
	<i>Example: THE BAR MAGNET AND THE BAR ELECTRET COMPARED</i>	379
21.4	MAGNETIC CIRCUITS	381
	<i>Example: MAGNETIC CIRCUIT WITH AN AIR GAP</i>	382
*21.5	SUMMARY	384
	PROBLEMS	384

In this chapter we first describe briefly the properties of ferromagnetic materials. These materials possess high permeabilities, but they are grossly nonlinear and lossy.

Magnetic fields involving ferromagnetic materials do not therefore lend themselves to rigorous mathematical analyses. Elaborate computer codes are available for performing numerical calculations and drawing field lines. Otherwise, one uses approximate methods like the ones that we survey briefly here.

*21.1 FERROMAGNETIC MATERIALS

Ferromagnetic and ferroelectric (Sec. 10.1.4) materials are, to a certain extent, similar. *Ferromagnetic materials* are partitioned into microscopic *domains* within which the spins are all spontaneously aligned in a given

[†]This chapter may be omitted without losing continuity, except that the concept of reluctance (Sec. 21.4) is a prerequisite for Sec. 25.4.

direction, even in the absence of an external field. In the unmagnetized state the spins of the various domains are randomly oriented and the net macroscopic field is zero.

Upon application of a magnetic field, those domains that happen to be oriented in about the correct direction grow at the expense of their neighbors by the migration of the domain walls. Eventually, near saturation, the magnetization rotates to the imposed direction.

Ferromagnetic substances are usually *anisotropic*.

*21.2 HYSTERESIS

One can measure B as a function of H with a *Rowland ring*, as in Fig. 21-1.[†] Winding a has N_a turns and bears a current I . From Sec. 20-6, it applies an azimuthal magnetic field strength

$$H = \frac{N_a I}{2\pi r}, \quad (21-1)$$

where r is the major radius. We have assumed that the minor radius of the toroid is much smaller than r .

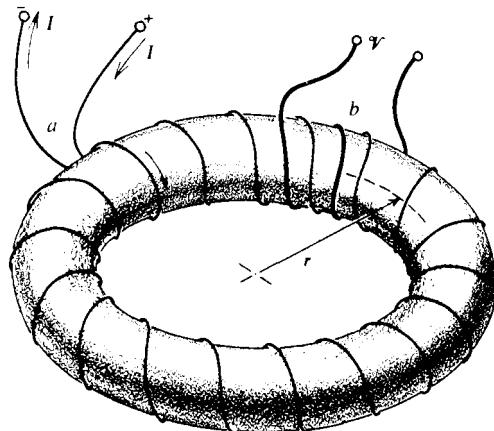


Fig. 21-1. Rowland ring for measuring B as a function of H in a ferromagnetic material. Winding a applies a uniform H . The voltage V across winding b gives B as in the first example in Sec. 23.4.2.

[†]The Rowland ring geometry has long been abandoned for obvious reasons: it is relatively difficult to machine a ring, especially if the material is brittle, and winding two coils on each sample takes time. Instead, one places a cylindrical sample of the material in the air gap of an electromagnet. However, the calculation of H in the sample is then relatively difficult. The Rowland ring remains an instructive exercise.

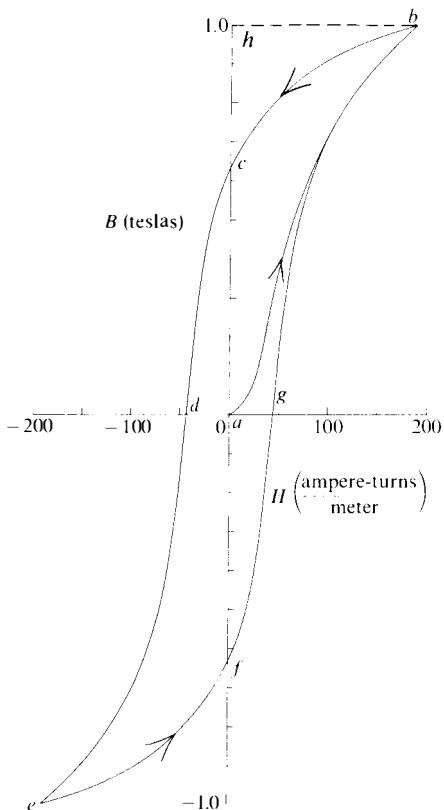


Fig. 21-2. Magnetization curve ab and hysteresis loop $bcdefgb$ for one type of transformer iron.

The magnetic flux density B is Φ/\mathcal{A} , where Φ is the magnetic flux and \mathcal{A} is the cross-sectional area of the core. One measures Φ by a method described in the first example in Sec. 23.4.2.

If we start with an unmagnetized sample of ferromagnetic material and increase the current in coil a , then B increases along a curve such as ab of Fig. 21-2. This is the magnetization curve of Sec. 20.7.1.

Reducing now the current in winding a to zero, B decreases along bc . The magnitude of B at c is the *remanence*, or the *retentivity*. If the current then reverses in direction and increases, B decreases to zero at d . The magnitude of H at this point is the *coercive force*. On further increasing the current in the same direction, a point e , symmetric to point b , is reached. If the current now decreases, then reverses and increases, point b is again reached. The closed curve $bcdefgb$ is a *hysteresis loop*.

Although hysteresis loops are always of the same general shape, they take many forms. They can be narrow as in the figure, or broad, or even rectangular with nearly vertical and horizontal sides.

Soft magnetic materials possess a high permeability and a narrow hysteresis loop. They serve in electromagnets, transformers, motors, etc. *Hard* materials are characterized by broad hysteresis loops. They are used mainly for permanent magnets.

As a rule, one records a hysteresis loop with \mathbf{H} and \mathbf{B} either parallel or antiparallel. In most applications however, \mathbf{H} is only approximately collinear with \mathbf{B} , because ferromagnetic materials are usually anisotropic.

One can demagnetize an object by exposing it to the magnetic field of a coil carrying an alternating current and then gradually removing the object from the field, so as to take the material round smaller and smaller hysteresis loops.

*21.2.1 Energy Dissipated in a Hysteresis Cycle

A ferromagnetic material dissipates energy in going around a hysteresis loop, as we shall see in the second example in Sec. 23.4.2. The energy dissipated per cubic meter of material and per cycle is equal to the area of the loop, measured in tesla-ampere-turns per meter, or in weber-ampere-turns per cubic meter, or in joules per cubic meter.

Example HYSTERESIS LOSSES IN A TRANSFORMER IRON

The loop of Fig. 21-2 encloses an area of 150 joules/meter³ cycle, or 1.1 watts/kilogram at 60 hertz. This iron is suitable for power transformers, but there exist less lossy alloys. Hysteresis losses in some amorphous alloys called *metallic glasses* are lower by an order of magnitude.

*21.3 MAGNETIC FIELD CALCULATIONS

As noted above, magnetic fields involving ferromagnetic materials do not lend themselves to rigorous mathematical analyses.

Simple-minded calculations, such as the one in the example that follows, are useful. However, they yield no more than the general features of a field.

The concept of magnetic circuit (Sec. 21.4), when applicable, usually offers a somewhat more realistic approach.

*21.3.1 The Bar Magnet

As we saw in the example in Sec. 20.3, the \mathbf{B} field of an idealized bar magnet that is magnetized uniformly parallel to its length is the same as that of a solenoid of identical dimensions and bearing a current density N/I equal to the magnetization M . Figure 21-3 shows the general features

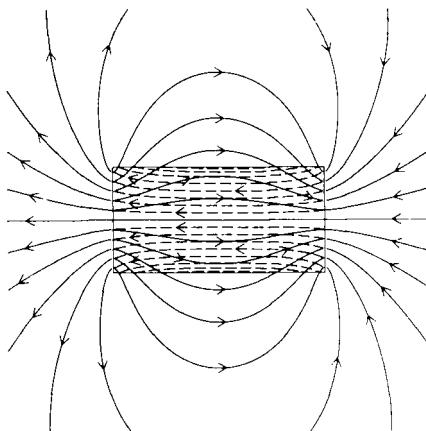


Fig. 21-3. Lines of \mathbf{H} (solid) for a uniformly magnetized cylinder, with \mathbf{M} uniform, parallel to the axis, and pointing to the left. Lines of \mathbf{B} are shown broken inside the magnet; outside they follow the lines of \mathbf{H} . This figure is identical to Fig. 9-4.

of this field. Observe how the lines of \mathbf{B} break at the cylindrical surface. However, the lines of \mathbf{H} pierce the cylindrical surface undisturbed. The normal component of \mathbf{B} and the tangential component of \mathbf{H} are continuous at the cylindrical surface and at the ends, as in Sec. 20.8.

Outside, $\mathbf{H} = \mathbf{B}/\mu_0$, while inside $\mathbf{H} = \mathbf{B}/\mu_0 - \mathbf{M}$.

Observe that inside the magnet \mathbf{B} and \mathbf{H} point in approximately opposite directions.

Where is the point (H, B) situated on the hysteresis loop? Suppose one has a ring of the material with windings as in Fig. 21-1. One varies the current in winding a so as to go, on the hysteresis loop of Fig. 21-2, from a to b and then to c . At that point $\mathbf{H} = 0$ and $\mathbf{B} = \mu_0\mathbf{M}$. This is the situation near the center of a long bar magnet.

One now cuts out a section of the ring, as in Fig. 21-4, to form a bar magnet. From Ampère's circuital law, the line integral of $\mathbf{H} \cdot d\mathbf{l}$ around path C is zero, because there are no free currents linking C . Outside, \mathbf{H} points in the same direction as \mathbf{B} , since $\mathbf{B} = \mu_0\mathbf{H}$. Therefore, inside, \mathbf{H} and \mathbf{B} point in roughly opposite directions. In other words, upon removing a section of the ring as in Fig. 21-4, there appears inside an \mathbf{H}

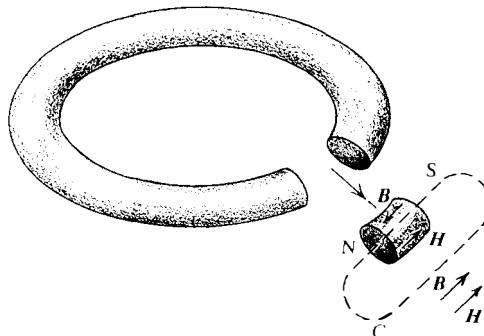


Fig. 21-4. Thought experiment in which one cuts out a portion of a magnetized ring to obtain a bar magnet. Inside the bar, \mathbf{B} and \mathbf{H} point in opposite directions. Lines of \mathbf{B} emerge from the North pole.

that points in the direction opposite to \mathbf{B} . The operating point of a permanent magnet therefore lies in the second quadrant of the hysteresis loop.

The operating point of a long slender magnet is close to c , and that of a stubby magnet approaches d , on the hysteresis loop.

The field of a real bar magnet is not that simple. Since the magnetic moments of the individual atoms tend to align themselves with the \mathbf{B} field, the magnetization \mathbf{M} , and hence the equivalent current density on the cylindrical surface, are weaker near the ends. Moreover, since \mathbf{M} is not uniform, there are equivalent currents inside the magnet. The end faces also carry equivalent currents since $\mathbf{M} \times \hat{\mathbf{n}}$ at the faces is zero only on the axis. The net result is that there are “poles” near the ends of the magnet from which lines of \mathbf{B} , outside the magnet, appear to radiate in all directions. The poles are most conspicuous if the bar magnet is long and thin.

Example

THE BAR MAGNET AND THE BAR ELECTRET COMPARED

It is instructive to compare the field of our uniformly magnetized cylinder of magnetic material, Fig. 21-5(a), with that of its electrical equivalent, the uniformly polarized cylinder of dielectric

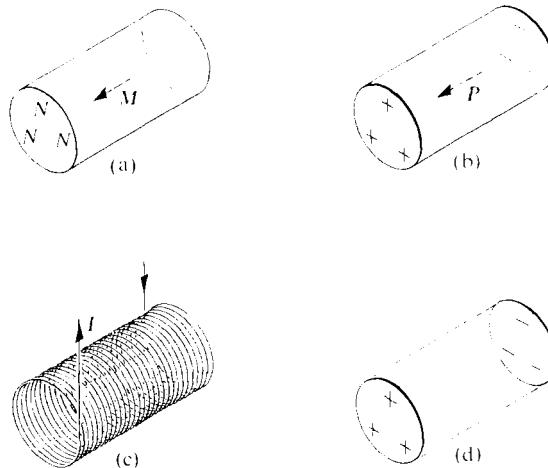


Fig. 21-5. (a) Bar magnet. (b) Bar electret. (c) Solenoid whose \mathbf{B} field is the same as that of the bar magnet. (d) Pair of electrically charged plates whose \mathbf{E} field is the same as that of the bar electret.

of Fig. 21-5(b). We discussed the field of the bar electret in the second example of Sec. 9.7.

The \mathbf{B} field of the magnet is that of an equivalent solenoid, Fig. 21-5(c), that carries a current density $N'I = M$. Then, to find the \mathbf{H} field, we use the relation

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}). \quad (21-2)$$

The \mathbf{E} field of the electret is that of the bound charges on the end faces, as in Fig. 21-5(d), or of a pair of parallel and oppositely charged disks carrying uniform charge densities $+P$ and $-P$. The \mathbf{D} field then follows from

$$\mathbf{E} = \frac{1}{\epsilon_0}(\mathbf{D} - \mathbf{P}). \quad (21-3)$$

Mathematically, the fields obey similar equations:

Electret

Magnet

$$\mathbf{E} = \frac{1}{\epsilon_0}(\mathbf{D} - \mathbf{P}) \quad (21-4), \quad \mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}) \quad (21-5),$$

$$\nabla \cdot \mathbf{D} = 0 \quad (21-6), \quad \nabla \cdot \mathbf{B} = 0 \quad (21-7),$$

$$\nabla \times \mathbf{E} = 0 \quad (21-8), \quad \nabla \times \mathbf{H} = 0 \quad (21-9).$$

Indeed, the \mathbf{H} field of the magnet is the same as if one had magnetic pole densities $+M$ and $-M$ on the end faces.

21.4 MAGNETIC CIRCUITS

Imagine a ferromagnetic ring carrying a short N -turn coil. The current is I . We wish to calculate the magnetic flux Φ through a cross section of the core.

We expect the B inside the core to be much larger near the winding than on the opposite side. On the contrary, B is, in fact, of the same order of magnitude at all points around the ring.

This is not difficult to understand. The magnetic field of the current I magnetizes the core in the region near the coil, and this magnetization gives equivalent currents that both increase B and extend it along the core. This further increases and extends the magnetization, and hence B .

Some of the lines of \mathbf{B} escape into the air and then return to the core to pass again through the coil. This constitutes the *leakage flux* that may, or may not, be negligible. For example, if the ring is a long, thin iron wire, most of the flux leaks across from one side of the ring to the other.

If the winding extends all around the ring,

$$H = \frac{NI}{2\pi r}, \quad (21-10)$$

as in Sec. 21.2, and

$$B = \frac{\mu NI}{2\pi r}. \quad (21-11)$$

The flux is the same as if the coil had no core and μ_r times more ampere-turns. In other words, for each ampere-turn in the coil, there are $\mu_r - 1$ ampere-turns of equivalent currents in the core. The amplification can be as high as 10^6 .

If R_1 is the radius where B has its average value and if R_2 is the minor radius of the ring, then

$$\Phi = \frac{\mu NI}{2\pi R_1} \pi R_2^2 = \frac{NI}{2\pi R_1 / (\mu \pi R_2^2)}. \quad (21-12)$$

This relation is reminiscent of Ohm's law. Here NI is the *magnetomotance*, while

$$\mathcal{R} = \frac{L}{\mu \mathcal{A}} = \frac{2\pi R_1}{\mu \pi R_2^2} \quad (21-13)$$

is the *reluctance* of the *magnetic circuit*, where L is its length and \mathcal{A} is its cross section. Thus

$$\text{Magnetic flux} = \frac{\text{magnetomotance}}{\text{reluctance}}. \quad (21-14)$$

The inverse of a reluctance is a *permeance*.

The corresponding quantities in electric and magnetic circuits are as follows:

Current I	Magnetic flux Φ
Current density \mathbf{J}	Magnetic flux density \mathbf{B}
Conductivity σ	Permeability μ
Applied voltage V	Magnetomotance NI
Electric field strength \mathbf{E}	Magnetic field strength \mathbf{H}
Resistance R	Reluctance \mathcal{R}
Conductance $G = 1/R$	Permeance $1/\mathcal{R}$

There is one important difference between electric and magnetic circuits: the magnetic flux cannot be made to follow a magnetic circuit as an electric current follows a conducting path. A magnetic circuit behaves much as an electric circuit would if it were submerged in salt water: part of the current would flow through the components, and the rest would flow through the water.

The leakage flux can easily be one order of magnitude larger than the useful flux if the circuit is not designed carefully.

Example

MAGNETIC CIRCUIT WITH AN AIR GAP

Figure 21-6 shows an electromagnet with a soft-iron yoke. Each winding provides $NI/2$ ampere-turns. We shall see that the magnetic flux in the air gap is equal to the magnetomotance NI divided by the sum of the reluctances of the iron yoke and of the air gap, assuming that the leakage flux is negligible.

This is a general rule: reluctances and permeances in a magnetic circuit add in the same way as resistances and conductances in an electric circuit.

From Ampère's circuital law,

$$NI = H_i L_i + H_g L_g, \quad (21-15)$$

where the subscript i refers to the iron and g to the air gap. The quantity L , can be measured along the center of the cross section all around the yoke.

We assume that the flux of \mathbf{B} is the same over any cross section of the circuit:

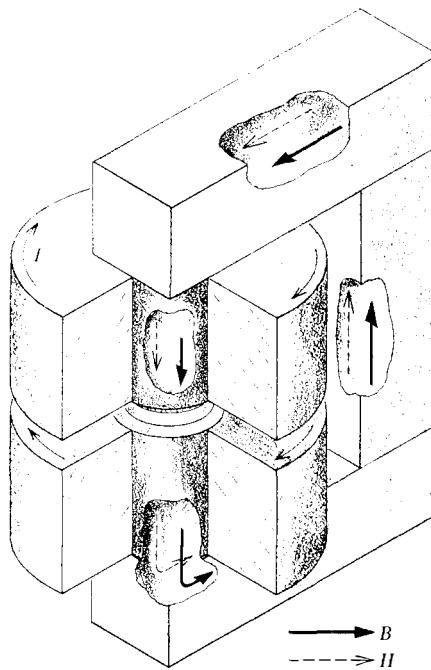


Fig. 21-6. Electromagnet. The artist has removed part of the coil to expose the left-hand side of the yoke.

$$\Phi = B_i \mathcal{A}_i = B_g \mathcal{A}_g, \quad (21-16)$$

where \mathcal{A} is a cross section.

Combining these two equations,

$$B_g \mathcal{A}_g \left(\frac{L_i}{\mu \mathcal{A}_i} + \frac{L_g}{\mu_0 \mathcal{A}_g} \right) = NI \quad (21-17)$$

and

$$\Phi = B_g \mathcal{A}_g = \frac{NI}{L_i / (\mu \mathcal{A}_i) + L_g / (\mu_0 \mathcal{A}_g)} \approx \frac{NI \mu_0 \mathcal{A}_g}{L_g}, \quad (21-18)$$

on the assumption that the reluctance of the yoke is negligible compared to that of the gap. You can now show that $H_i L_i \ll H_g L_g$ if this assumption is correct.

Since we have neglected the leakage flux, this equation can only serve to provide an upper limit for Φ and for B_g .

There exist empirical formulas for calculating leakage fluxes.[†]

[†] See Malcolm McCuaig, *Permanent Magnets in Theory and in Practice*, Wiley, New York, 1977, Chap. 6.

Let $N = 10,000$, $I = 1.00$ ampere, $\mathcal{A} = 100$ centimeters², $\mathcal{A}_g = 50.0$ centimeters², $\mu_r = 1000$, $L_i = 900$ millimeters, $L_g = 10.0$ millimeters. Then

$$\mathcal{R} = \frac{0.9}{10^3 \times 4\pi \times 10^{-7} \times 10^{-2}} + \frac{10^{-2}}{4\pi \times 10^{-7} \times 5 \times 10^{-3}} \quad (21-19)$$

$$= 1.7 \times 10^6 \text{ ampere-turns/weber}, \quad (21-20)$$

$$\Phi = \frac{10^4}{1.7 \times 10^6} = 6.0 \times 10^{-3} \text{ weber}, \quad (21-21)$$

$$B_g = \frac{5.9 \times 10^{-3}}{5 \times 10^{-3}} = 1.2 \text{ teslas}. \quad (21-22)$$

With this particular design the leakage flux is approximately 70% of the flux in the gap. In other words, one can expect the B in the gap to be only $1.2/1.7 = 0.7$ tesla.

*21.5 SUMMARY

Ferromagnetic materials spontaneously magnetize over microscopic *domains* that can grow, one at the expense of another, upon application of a magnetic field.

If one plots B as a function of H for an initially unmagnetized sample, B first increases with H along the *magnetization curve*. Over a complete cycle of H one obtains a closed curve called a *hysteresis loop*. The area of the loop is equal to the energy lost per cubic meter of material and per cycle.

The concept of *magnetic circuit* is useful when the magnetic flux is mostly circumscribed to magnetic material. The magnetic equivalent of Ohm's law is then

$$\text{Magnetic flux} = \frac{\text{magnetomotance}}{\text{reluctance}}, \quad (21-14)$$

with corresponding quantities as in Sec. 21.4. The *magnetomotance* of a coil is NI , where N is the total number of turns.

PROBLEMS

21-1. (21.4) A permanent-magnet loudspeaker

Figure 21-7 shows a cross section of a common form of magnetic circuit for a loudspeaker. The permanent magnet is a cylinder marked NS. The yoke has the form of a cup with a top plate, and the voice coil lies in the air gap with its axis vertical.

The top plate has a diameter of 25.0 millimeters, and the gap is 2.50 millimeters wide. The magnet is made of Alnico V, it is

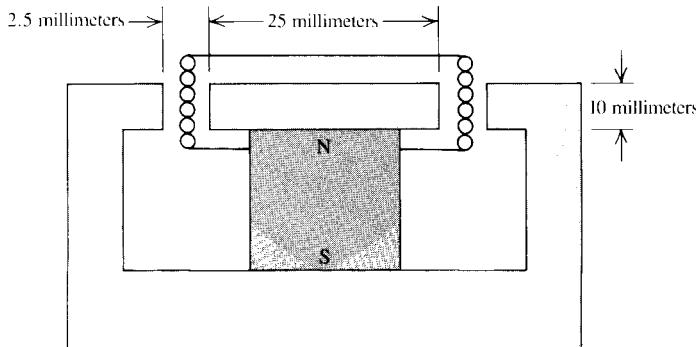


Fig. 21-7.

25.0 millimeters long, its diameter is 20.0 millimeters, and it operates at its optimum H of 4×10^4 ampere-turns/meter.

Calculate B in the gap. The reluctance of the yoke is negligible. The magnetomotance of the permanent magnet is 4×10^4 times its length. The value of B in loudspeakers is usually in the range of 0.3 to 2 teslas.

21-2. (21.4) Magnetic circuit

A magnetic circuit comprises an air gap, as in Fig. 21-8, with $R_2 - R_1 \ll a$. Find an approximate expression for the reluctance of the gap.

21-3. (21.4) Iron ring with a thin air gap

An iron ring carries a 300-turn coil. The ring has a mean diameter of 400 millimeters, a cross section of 1000 millimeters², and a relative permeability of 500.

(a) Calculate B when the current in the coil is 1 ampere.

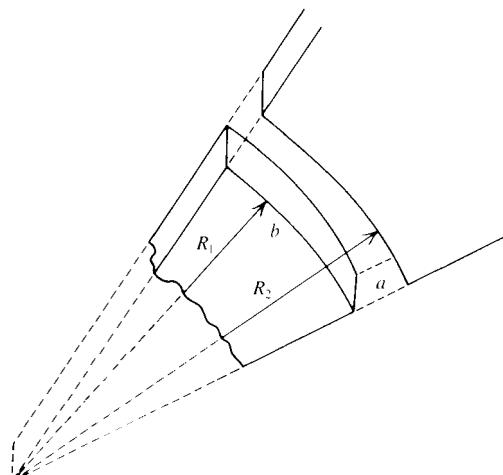


Fig. 21-8.

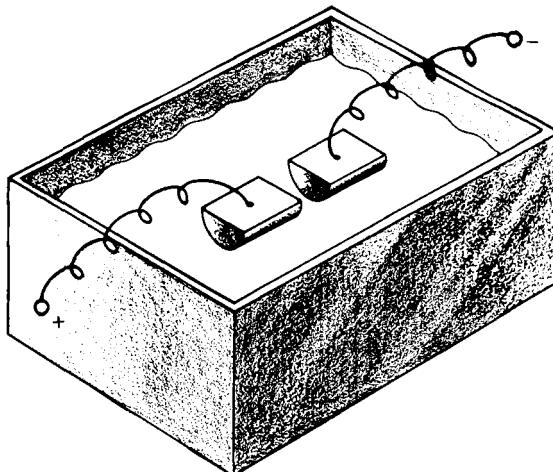


Fig. 21-9.

- (b) Calculate B for the same current when the ring has a gap of 1.00 millimeter.

21-4. (21.4) Plotting a magnetic field with an electrolytic tank.

Figure 21-9 shows one example of the use of an electrolytic tank for plotting a magnetic field in a region where there are no currents and no magnetic materials. Here the electrodes are shaped like the pole pieces of an electromagnet. The model is strictly valid if the relative permittivity of the pole pieces is infinite.

The lines of E for the model are identical to the lines of B for the electromagnet. Let us see why.

(a) Which three differential equations does B satisfy in free space? The equation $\mathbf{B} = \nabla \times \mathbf{A}$ is not useful here. Which three differential equations does E satisfy in the model?

(b) Show that B is derivable from a potential: $\mathbf{B} = -\nabla u$. The function u is the *scalar magnetic potential*.

(c) Show that a surface of constant u is orthogonal to the lines of B . Thus a surface of constant u corresponds to an equipotential.

(d) Show that $\nabla^2 u = 0$.

(e) On the model, $\int_C \mathbf{E} \cdot d\mathbf{l} = V$, where C is any curve that goes from one electrode to the other and V is the applied voltage. What is the corresponding equation for B ?

(f) On the model, $I = \int \sigma \mathbf{E} \cdot d\mathcal{A}$, where I is the current between the electrodes, σ the conductivity of the electrolyte, and $d\mathcal{A}$ an element of area on an electrode. What is the corresponding equation for the magnetic flux?

One could also deduce the reluctance from the resistance V/I .

CHAPTER 22

MAGNETIC FIELDS V *The Magnetic Forces on Charges and Currents*

22.1	THE LORENTZ FORCE	387
	<i>Example: THE HALL EFFECT</i>	388
	<i>Example: THE MAGNETOHYDRODYNAMIC GENERATOR</i>	389
*22.1.1	THE MAGNETIC FORCE INSIDE FERROMAGNETIC MATERIALS	392
22.2	THE MAGNETIC FORCE ON A CURRENT-CARRYING WIRE	393
	<i>Example: THE FLOATING-WIRE HODOSCOPE</i>	394
22.3	THE MAGNETIC FORCE BETWEEN TWO CLOSED CIRCUITS	396
	<i>Example: THE FORCE BETWEEN TWO PARALLEL CURRENTS</i>	397
22.3.1	THE DEFINITIONS OF μ_0 , THE AMPERE, THE COULOMB, AND ϵ_0	398
22.4	THE MAGNETIC FORCE ON A VOLUME DISTRIBUTION OF CURRENT	398
	<i>Example: THE HOMOPOLAR MOTOR</i>	399
	<i>Example: THE HOMOPOLAR GENERATOR</i>	403
22.5	SUMMARY	404
	PROBLEMS	405

We shall study magnetic forces in two chapters, the present one and Chap. 26. This is because we require the force $Q\mathbf{v} \times \mathbf{B}$ now, but we must defer a more general discussion of magnetic forces to a later stage.

The $Q\mathbf{v} \times \mathbf{B}$ force manifests itself most clearly on electron and ion beams, say in television sets and certain ion accelerators. It is also the force that drives electric motors.

22.1 THE LORENTZ FORCE

Experiments show that the force exerted on a particle of charge Q moving in a vacuum at an instantaneous velocity \mathbf{v} in a region where there exist both an electric and a magnetic field is

$$\mathbf{F} = Q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (22-1)$$

This is the *Lorentz force*. This equation is valid, even if v approaches the speed of light. The variables E , B , and v can be space- and time-dependent, but they all concern the *same* reference frame.

The term $Qv \times B$ is the *magnetic force*. Observe that the magnetic force is perpendicular to v . It can therefore change the direction of v , but it cannot alter its magnitude, nor can it alter the kinetic energy of the particle. It can nonetheless do useful work, as we shall see.

Example

THE HALL EFFECT

In a bar of conducting material subjected to electric and magnetic fields as in Fig. 22-1(a) or (b) the charge carriers drift longitudinally. They also drift sideways because of the $Qv \times B$ force. As a result, there appears a potential difference between the upper and lower electrodes. This tendency of charge carriers to drift sideways in a transverse magnetic field is called the *Hall effect*.

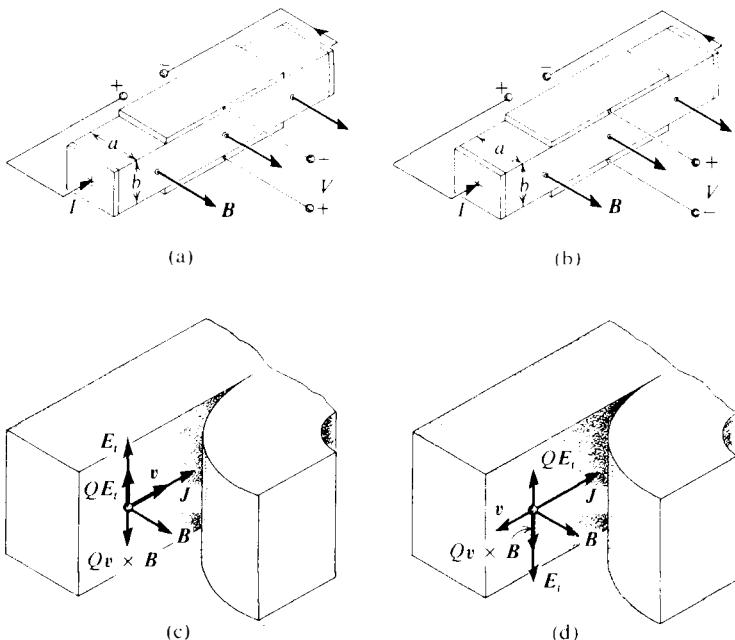


Fig. 22-1. The Hall effect in semiconductors. (a) In *p*-type material the charge carriers are holes, and the Hall voltage is as shown. (b) In *n*-type material the carriers are electrons, and the Hall voltage has the opposite polarity. (c) In *p*-type material the transverse forces QE_i and $Qv \times B$ are equal and opposite. (d) The transverse forces in *n*-type material.

If a voltmeter connected between the upper and lower electrodes draws a negligible current, then the plates charge until their transverse electric field halts the transverse drift. This transverse field is termed the *Hall field*. The net transverse force on the charge carriers is then zero, as in the figure.

It is a simple matter to calculate the voltage V if we disregard end effects and assume that the externally applied \mathbf{B} is both uniform and much larger than that of the current I of the figure. Once the transverse drift has subsided,

$$QE_r = \frac{QV}{b} = QvB, \quad (22-2)$$

where b is the sample thickness as in the figure, and v is the longitudinal drift velocity. Then

$$V = vBb. \quad (22-3)$$

If the mobility of the carriers is M (Sec. 4.3.3) and the longitudinal electric field strength is E_{long} , then

$$V = E_{\text{long}}Mb. \quad (22-4)$$

The voltage is proportional to the mobility.

Note that in the figures the charge carriers tend to drift downward, whether they are positive or negative. Thus the polarity of V depends on the sign of the carriers.

The Hall effect serves to measure B , but it also serves several other purposes. See Prob. 22-8.

The mobilities of semiconductors being larger than those of good conductors by orders of magnitude, Hall devices invariably make use of semiconductors. Some devices are microscopic and form part of integrated circuits.

Example

THE MAGNETOHYDRODYNAMIC GENERATOR

The *magnetohydrodynamic (MHD) generator* is a large-scale application of the Hall effect. It converts part of the kinetic energy of a hot gas directly into electric energy. Figure 22-2 shows its principle of operation. A hot gas enters on the left at a velocity of the order of 1000 meters/second. It contains a salt such as K_2CO_3 that ionizes readily at high temperature, forming positive ions and electrons. The temperature approaches 3000 kelvins and the conductivity is about 100 siemens/meter. (The conductivity of copper is 5.8×10^7 siemens/meter.) The gas remains neutral.

Positive ions curve downward, and electrons upward. The resulting current I flows through the load resistance R . This establishes an electric field \mathbf{E} as in the figure.

One obvious advantage of the MHD generator is that it comprises no moving parts, except for the gas. Another is that it

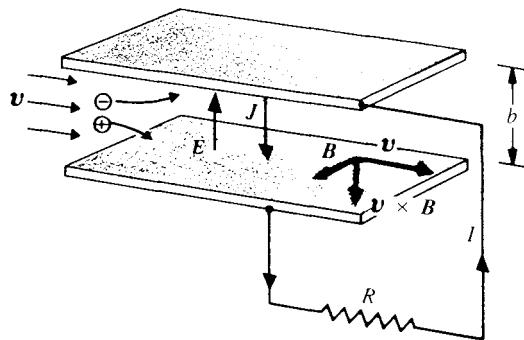


Fig. 22-2. Schematic diagram of a magnetohydrodynamic (MHD) generator. Part of the kinetic energy of a very hot gas injected on the left at a velocity v is converted directly into electric energy. The magnetic field B is that of a pair of superconducting coils situated outside the chamber and not shown. The moving ions deflect either up or down, according to their signs.

can operate at such a high input temperature that the overall thermodynamic efficiency

$$\mathcal{E} = \frac{T_{\text{in}} - T_{\text{out}}}{T_{\text{in}}} \quad (22-5)$$

exceeds 50% if the hot output gas feeds a conventional turbine generator. The fuel is either coal or oil. The thermodynamic efficiencies of conventional thermal plants range from about 30% to 33%.

One transportable MHD generator serves to generate current pulses of 10,000 amperes for geophysical exploration. The largest MHD generator at this time is under construction and will produce 500 megawatts of electricity.

Part of the kinetic energy associated with the bulk motion of the gas becomes electric energy in the following way. The magnetic force on a charged particle is normal to the velocity and has no effect on the kinetic energy. The function of the magnetic forces is to compel the positive particles to go to the positive electrode and the negative particles to go to the negative electrode. So the ions and the electrons both move uphill, so to speak, under the action of the forces shown in Fig. 22-3, and slow down. Since the gas pressure is above atmospheric, the mean free path between collisions is infinitesimal and the charged particles are embedded in the gas. Thus, slowing the particles slows the gas and runs down its bulk kinetic energy.

Ideally, the charged particles should arrive at the electrodes at zero velocity. They arrive in fact at a finite velocity, the electrodes heat up, and only part of the kinetic energy of the *particles* becomes electric energy. An even smaller fraction of the bulk

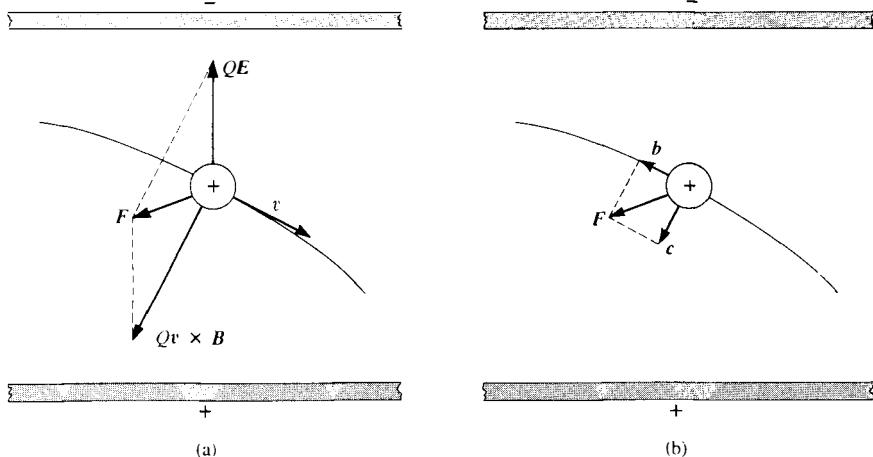


Fig. 22-3. (a) The electric force QE and the magnetic force $Q\mathbf{v} \times \mathbf{B}$ acting on a positively charged ion in the MHD generator of Fig. 22-2. The sum \mathbf{F} of those two forces points downward and to the left. (b) The force \mathbf{F} has two components: a tangential braking force \mathbf{b} and a normal centripetal force \mathbf{c} .

kinetic energy of the gas serves to generate electric energy. Also, some of the kinetic energy associated with the transverse motion of the charged particles only increases the random thermal energy of the gas.

We suppose that \mathbf{E} , \mathbf{B} , and the particle velocity \mathbf{v} are uniform and mutually perpendicular inside the generator, as in the figure. These are crude assumptions indeed.

The Lorentz force $Q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ acts on a charge Q as if the electric field strength were $\mathbf{E} + \mathbf{v} \times \mathbf{B}$. So, for a gas of conductivity σ ,

$$|\mathbf{J}| = |\sigma(\mathbf{E} + \mathbf{v} \times \mathbf{B})| = \sigma(vB - E) = \sigma\left(vB - \frac{V}{b}\right), \quad (22-6)$$

$$I = \mathcal{A}J = \mathcal{A}\sigma\left(vB - \frac{IR}{b}\right), \quad (22-7)$$

where I and R are as in the figure and where \mathcal{A} is the area of one of the electrodes. Solving,

$$I = \frac{vBb}{b/(\sigma\mathcal{A}) + R}. \quad (22-8)$$

Observe that vBb is the open-circuit output voltage and that $b/\sigma\mathcal{A}$ is the resistance of the gas in the transverse direction. We could, therefore, have arrived at this result directly from Thévenin's theorem (Sec. 8.3).

The output voltage is IR . After cross multiplication,

$$V = IR = vBb - \frac{Ib}{\sigma\mathcal{A}}. \quad (22-9)$$

The voltage V decreases linearly with I .

Now let us look into the efficiency. We are not prepared to calculate the efficiency with which the bulk kinetic energy of the gas becomes electric energy. However, we can compare the Joule losses in the load resistance R to those inside the gas. So let us define the efficiency as

$$\mathcal{E} = \frac{\text{Joule losses in } R}{\text{Joule losses in } R + \text{Joule losses in the gas}} \quad (22-10)$$

$$= \frac{I^2R}{I^2R + I^2b/(\sigma\mathcal{A})} = \frac{R}{R + b/(\sigma\mathcal{A})}, \quad (22-11)$$

as we would expect from Thévenin's theorem.

From the above expressions for I and for IR ,

$$\mathcal{E} = R \frac{I}{vBb} = \frac{IR}{vBb} = 1 - \frac{I}{\sigma\mathcal{A}vB}. \quad (22-12)$$

The efficiency is therefore equal to unity when $I = 0$, or when $R \rightarrow \infty$. It is equal to zero when $I = \sigma\mathcal{A}vB$, or when $R = 0$, $V = 0$.

*22.1.1 The Magnetic Force Inside Ferromagnetic Materials

Imagine a charged particle, for example, a conduction electron or a high-energy cosmic ray proton, moving through magnetized iron. The particle "sees" an exceedingly inhomogeneous magnetic field because each individual electron in the material acts as a small coil. The magnetic force $Q\mathbf{v} \times \mathbf{B}$ thus varies erratically with time, both in magnitude and in direction.

It would be pointless to attempt a detailed description of the motion. All that matters is that, on a macroscopic scale, the magnetic force corresponds to some effective \mathbf{B} that is not necessarily the same as the macroscopic \mathbf{B} that we have been thinking of until now.

For slow particles, such as conduction electrons in ferromagnetic materials, the magnetic force is $Q\mathbf{v} \times \mu_0\mathbf{H}$, where \mathbf{H} is the macroscopic magnetic field strength and \mathbf{v} is the drift velocity of the particle. However, Rasetti showed, many years ago[†] that with high-energy particles the magnetic force is $Q\mathbf{v} \times \mathbf{B}$. So the effective flux density in relation to the magnetic force in ferromagnetic bodies depends on the velocity of the particle.[‡]

[†] F. Rasetti, *Physical Review*, vol. 66, p. 1 (1944).

[‡] G. H. Wanier, *Physical Review*, vol. 72, p. 4 (1947).

We should really write $Q\mathbf{v} \times \mathbf{B}_{\text{eff}}$ for the magnetic force, but we refrain from doing so, to avoid complicating the notation. The important point is that the effective \mathbf{B} for either conduction electrons or holes is $\mu_0 \mathbf{H}$, and not $\mu_r \mu_0 \mathbf{H}$.

22.2 THE MAGNETIC FORCE ON A CURRENT-CARRYING WIRE

A stationary wire of cross section \mathcal{A} carries a current I in a region where there exists a magnetic field \mathbf{B} originating elsewhere. The wire contains N charge carriers per cubic meter drifting at a velocity \mathbf{v} , each one of charge Q .

An element of length dl of the wire contains $\mathcal{A}Nd l$ charge carriers. Then the magnetic force on dl is

$$\mathbf{dF} = \mathcal{A}N dl Q\mathbf{v} \times \mathbf{B} = (\mathcal{A}NQ\mathbf{v}) dl \times \mathbf{B} = I dl \times \mathbf{B}, \quad (22-13)$$

since I is equal to the charge contained in a length v of the wire.

The magnetic force per unit length on a wire bearing a current I is therefore $I \times \mathbf{B}$.

Now we have calculated the force on the charge carriers. How is this force transmitted to the wire? The charge carriers move sidewise, slightly, as in Fig. 22-4, which sets up a Hall field, and the resultant electric force on the stationary positive lattice pulls the wire sideways.[†]

The total magnetic force on a closed circuit C carrying a current I and lying in a magnetic field is

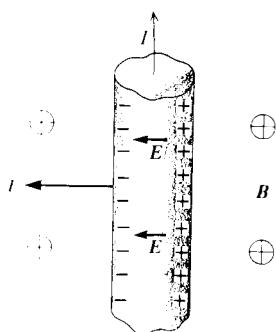


Fig. 22-4. Section through a wire bearing a current I in a magnetic field \mathbf{B} . We have grossly exaggerated the charge separation. The magnetic force on the electrons pushes them to the left, leaving an excess of positive charge on the right. The resultant Hall field \mathbf{E} inside the wire pulls the lattice of positive charges to the left.

[†] See W. R. McKinnon, S. P. McAlister, and C. M. Hurd, *American Journal of Physics*, Vol. 49, p. 493 (1981), for a more rigorous discussion of this force.

$$\mathbf{F} = I \oint_C d\mathbf{l} \times \mathbf{B}. \quad (22-14)$$

If \mathbf{B} is uniform, the net magnetic force \mathbf{F} is zero.

Example

THE FLOATING-WIRE HODOSCOPE

The *floating-wire hodoscope*[†] is a device that simulates the trajectory of a charged particle in a magnetic field. Say a charged particle of mass m , charge Q , and velocity \mathbf{v} follows a certain trajectory in going from a point a to a point b in a magnetic field. Then a light wire carrying a current I , fixed at a and going over a pulley at b , in the *same* magnetic field, as in Fig. 22-5, will follow that trajectory if

$$\frac{mv}{Q} = \frac{T}{I}, \quad (22-15)$$

where T is the tension in the wire. This statement is by no means

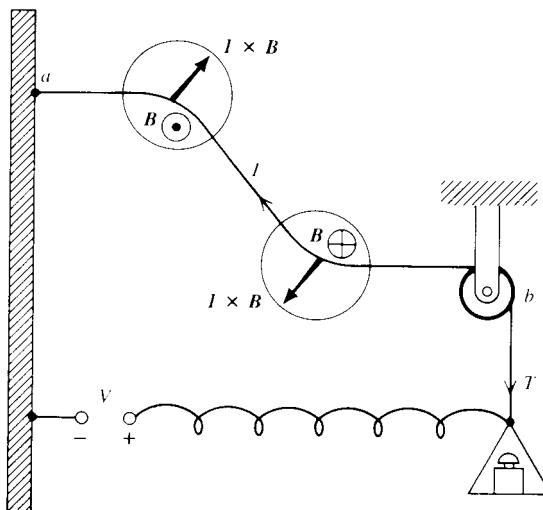


Fig. 22-5. Floating-wire hodoscope. A light wire bearing a current I fixed at a goes over a pulley at b . A small weight provides a tension T . This device simulates the trajectory of an ion beam travelling through the same magnetic field. With the magnetic fields shown the wire adopts an S-shaped posture.

[†] The term *hodoscope* also designates various devices that record the trajectory followed by a high-energy particle. The principles involved in those devices bear no relation to the material in this section.

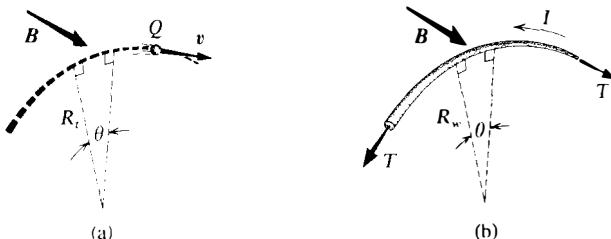


Fig. 22-6. (a) A positive charge Q moving at a velocity v in a magnetic field B . The radius of curvature of the trajectory is R_t . (b) Light wire carrying a current I in the opposite direction in the same magnetic field. The tension in the wire is T , and the radius of curvature is R_w . The two radii are equal if mv/Q is equal to T/I .

obvious, but it is easy to prove, as we shall see. The advantage of the floating-wire hodoscope is that it is much easier to experiment with a wire than with an ion beam.

Suppose the *beam* is normal to B as in Fig. 22-6(a). Then, if R_t is the radius of curvature of the trajectory,

$$QvB = \frac{mv^2}{R_t}, \quad R_t = \frac{mv}{QB}. \quad (22-16)$$

Now suppose the *wire* is also normal to B , as in Fig. 22-6(b). The element dl , with a radius of curvature R_w , is in equilibrium if the outward magnetic force $BI dl$ just compensates the inward component of the tension forces T :

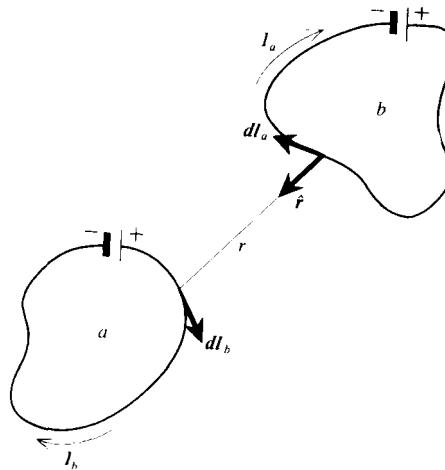


Fig. 22-7. Two closed circuits *a* and *b*.

$$BI dl = 2T \sin \frac{d\theta}{2} = T \frac{dl}{R_w}, \quad R_w = \frac{T}{IB}. \quad (22-17)$$

The two radii are equal if

$$\frac{mv}{Q} = \frac{T}{I}. \quad (22-18)$$

Notice that the particle deflects downward if the magnetic force is downward, but the wire curves downward if the force is upward.

If \mathbf{B} is not uniform and if the beam is not perpendicular to \mathbf{B} , then the wire does not always follow the trajectory. For example, a magnetic field can serve to both focus and deflect an ion beam. Then focusing forces on the beam can become defocusing forces on the wire, which then moves away from the trajectory.

22.3 THE MAGNETIC FORCE BETWEEN TWO CLOSED CIRCUITS

We saw above that the magnetic force exerted on a stationary circuit carrying a current I is I times the line integral of $d\mathbf{l} \times \mathbf{B}$. Then, applying the Biot-Savart law of Sec. 18.2, the magnetic force exerted by a current I_a on a current I_b , as in Fig. 22-7, is given by

$$\mathbf{F}_{ab} = I_b \oint_b d\mathbf{l}_b \times \frac{\mu_0}{4\pi} I_a \oint_a \frac{d\mathbf{l}_a \times \hat{\mathbf{r}}}{r^2}, \quad (22-19)$$

$$= \frac{\mu_0}{4\pi} I_a I_b \oint_a \oint_b d\mathbf{l}_b \times \frac{d\mathbf{l}_a \times \hat{\mathbf{r}}}{r^2}, \quad (22-20)$$

where r is the distance between $d\mathbf{l}_a$ and $d\mathbf{l}_b$, and $\hat{\mathbf{r}}$ points from $d\mathbf{l}_a$ to $d\mathbf{l}_b$.

The fact that $d\mathbf{l}_a$ and $d\mathbf{l}_b$ do not play symmetric roles in this integral is disturbing. The asymmetry appears to indicate that $\mathbf{F}_{ab} \neq \mathbf{F}_{ba}$, which is contrary to Newton's law.[†] That impression is false. We can transform the double line integral to a symmetric one and show that $\mathbf{F}_{ab} = -\mathbf{F}_{ba}$ as follows. First, we expand the triple vector product:

$$d\mathbf{l}_b \times (d\mathbf{l}_a \times \hat{\mathbf{r}}) = d\mathbf{l}_a (d\mathbf{l}_b \cdot \hat{\mathbf{r}}) - \hat{\mathbf{r}} (d\mathbf{l}_a \cdot d\mathbf{l}_b). \quad (22-21)$$

Then, rearranging terms,

$$\oint_a \oint_b \frac{d\mathbf{l}_a (d\mathbf{l}_b \cdot \hat{\mathbf{r}})}{r^2} = \oint_a d\mathbf{l}_a \oint_b \frac{d\mathbf{l}_b \cdot \hat{\mathbf{r}}}{r^2}. \quad (22-22)$$

[†] Action is equal to reaction at speeds much less than the speed of light.

The second integral on the right is zero for the following reason. It is the ordinary integral of dr/r^2 , with identical upper and lower limits, because circuit b is closed, by hypothesis. So the double integral on the left is zero, and

$$\mathbf{F}_{ab} = -\frac{\mu_0}{4\pi} I_a I_b \oint_a \oint_b \hat{\mathbf{r}} \frac{d\mathbf{l}_a \cdot d\mathbf{l}_b}{r^2}. \quad (22-23)$$

It follows that $\mathbf{F}_{ab} = -\mathbf{F}_{ba}$ because $\hat{\mathbf{r}}$ points toward the circuit on which the force acts.

The above double line integral is usually difficult to calculate analytically, unless you have access to a computer that can perform symbolic calculations. We shall find more useful expressions in Chap. 26.

Example

THE FORCE BETWEEN TWO PARALLEL CURRENTS

We can calculate the force per unit length between two long parallel wires bearing currents as in Fig. 22-8, without having to carry out the integration. At the position of I_b , B_a is $\mu_0 I_a / 2\pi D$ in the direction shown in the figure. See the first example in Sec. 18.4. The force on a unit length of wire b is thus

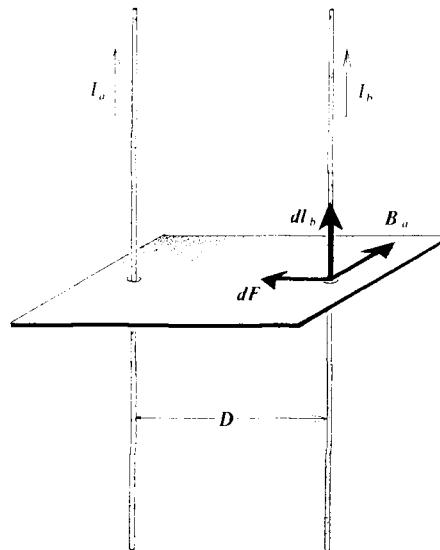


Fig. 22-8. Two long parallel wires bearing currents I_a and I_b . The force is attractive when the currents flow in the same direction.

$$F' = B_a I_b = \frac{\mu_0 I_a I_b}{2\pi D}. \quad (22-24)$$

The force is attractive if the currents flow in the same direction and repulsive otherwise. The force is normally negligible.

22.3.1 The Definitions of μ_0 , the Ampere, the Coulomb, and ϵ_0

As stated previously in Sec. 18.2,

$$\mu_0 = 4\pi \times 10^{-7} \text{ weber/ampere-meter.} \quad (22-25)$$

If $I_a = I_b = I$ amperes, then the force per unit length is

$$F' = \frac{2 \times 10^{-7} I^2}{D} \text{ newtons/meter.} \quad (22-26)$$

This equation defines the *ampere*.

The definition of the *coulomb* follows: it is the charge carried by a current of 1 ampere during 1 second.

So the definitions of μ_0 , I , and Q are arbitrary and related.

But Coulomb's law relates the force of attraction between two electric charges to their magnitudes, and force is defined in mechanics. Coulomb's law must therefore involve a constant of proportionality whose value must be measured.

We could, in principle, deduce the value of ϵ_0 from the measurement of F , the Q 's, and r in Coulomb's law. However, this would not make much sense because none of those measurements can be very accurate. Instead, we use the fact that

$$\epsilon_0 = \frac{1}{\mu_0 c^2}, \quad (22-27)$$

where c , the speed of light, is defined to 9 significant figures. Thus

$$\epsilon_0 = 8.854187817 \times 10^{-12} \text{ farad/meter.} \quad (22-28)$$

22.4 THE MAGNETIC FORCE ON A VOLUME DISTRIBUTION OF CURRENT

In this chapter we have assumed that the conduction currents flow in thin wires. What if, instead, we have a volume distribution of current? We

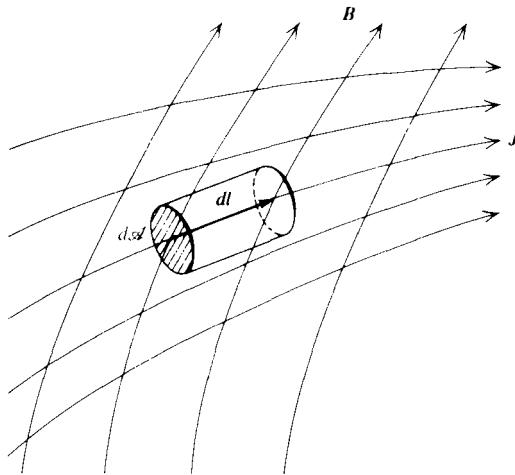


Fig. 22-9. Element of volume $dv = dl da$ in a current distribution \mathbf{J} .

can find the magnetic force density without much effort. Consider a small element of volume of length dl parallel to \mathbf{J} and of cross section $d\mathcal{A}$, as in Fig. 22-9. The magnetic force exerted on the element is

$$\mathbf{dF} = (\mathbf{J} d\mathcal{A}) dl \times \mathbf{B} = \mathbf{J} \times \mathbf{B} dv \quad (22-29)$$

and the force per unit volume is

$$\mathbf{F}' = \mathbf{J} \times \mathbf{B}. \quad (22-30)$$

The total magnetic force on a given distribution of conduction currents occupying a volume v is thus

$$\mathbf{F} = \int_v \mathbf{J} \times \mathbf{B} dv. \quad (22-31)$$

Example

THE HOMOPOLAR MOTOR

A *homopolar motor* comprises a copper disk that rotates in a more or less uniform axial magnetic field, with contacts on the axis and on the periphery, as in Fig. 22-10. The magnetic force on the current flowing radially through the disk provides the driving torque.

Homopolar motors are inherently low-voltage, high-current devices. They usually operate with direct current. They can provide high torques at low velocities and are suitable, for example, for ship propulsion. The current is then supplied by a

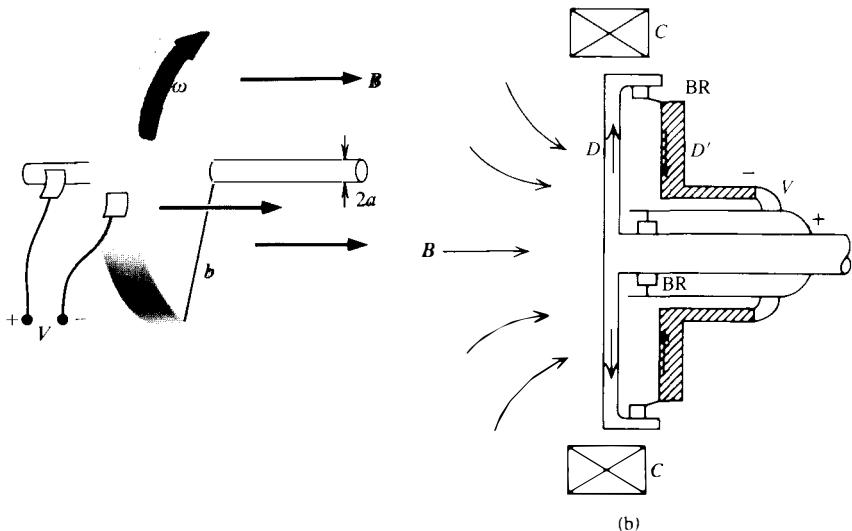


Fig. 22-10. Homopolar motor. (a) Principle of operation. (b) Cross section showing the coil C , the disk D , the brush contacts BR , and the terminals V . The torque on the rotating disk D is equal and opposite to that on the fixed disk D' . Moreover, the field of the current through D and D' is azimuthal. Thus there is no force on coil C . This is important because C is superconducting and enclosed in a cryostat.

diesel motor-generator through a step-down transformer and a rectifying circuit. Superconducting coils provide the axial \mathbf{B} .

We assume that the electric field in the disk is purely radial. This requires contacts all around the periphery, which is the usual configuration.

Refer to Fig. 22-11. (1) The charge carrier Q has an azimuthal velocity ωr . (2) It is therefore subjected to the radial magnetic force $Q\omega rB$. (3) The charge Q is also subjected to the radial force QE . (4) Its radial velocity is therefore $M(E - \omega rB)$, where M is the mobility (Sec. 4.3.3). The mobility of a charge carrier is equal to the force divided by Qu . This radial velocity is smaller than ωr by many orders of magnitude. (5) This small radial velocity gives an azimuthal force $Q[M(E - \omega rB)]B$.

The azimuthal force per unit volume at the radius r is thus

$$F' = NQ[M(E - \omega rB)]B = \sigma(E - \omega rB)B, \quad (22-32)$$

from Sec. 4.3.3, N being the number of charge carriers per unit volume. This force density is simply $|\mathbf{J} \times \mathbf{B}|$. It points in the counterclockwise direction in the figure.

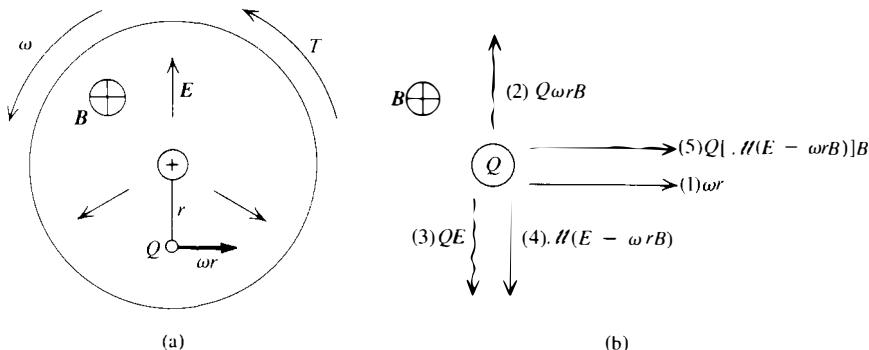


Fig. 22-11. (a) The rotor of a homopolar motor. The electric field is radial, while the magnetic field is axial. A charge carrier Q has an azimuthal velocity ωr . (b) The velocity components of Q (straight arrows) and the various forces exerted on Q (wavy arrows).

For a disk of thickness s , the torque per unit volume exerted on the charge carriers is $\mathbf{r} \times \mathbf{F}'$, and the total torque is

$$T = \int_a^b r [2\pi rs\sigma(E - \omega r B) B] dr = 2\pi s \sigma B \int_a^b r^2 (E - \omega r B) dr \quad (22-33)$$

in the counterclockwise direction. Here a and b are, respectively, the inner and outer radii of the region of the disk where the current is radial.

We must still find E as a function of r . That is easy because the radial current is independent of r :

$$I = 2\pi r s J = 2\pi r s \sigma (E - \omega r B). \quad (22-34)$$

Then

$$E - \omega r B = \frac{I}{2\pi r s \sigma}. \quad (22-35)$$

Finally,

$$T = BI \int_a^b r dr = \frac{BI(b^2 - a^2)}{2}. \quad (22-36)$$

This is the counterclockwise torque exerted on the charge carriers. It is also the torque exerted on the disk because the carriers continually collide with the atoms of the crystal lattice.

There are three interesting points to note here.

(1) The magnetic force does useful work here because it possesses an azimuthal component. The fact that it does not affect the kinetic energy of the charge carriers is irrelevant.

(2) Equation 22-35 shows that E is a peculiar function of the radius:

$$E = (\omega B)r + \frac{I}{2\pi s \sigma r} \frac{1}{r}. \quad (22-37)$$

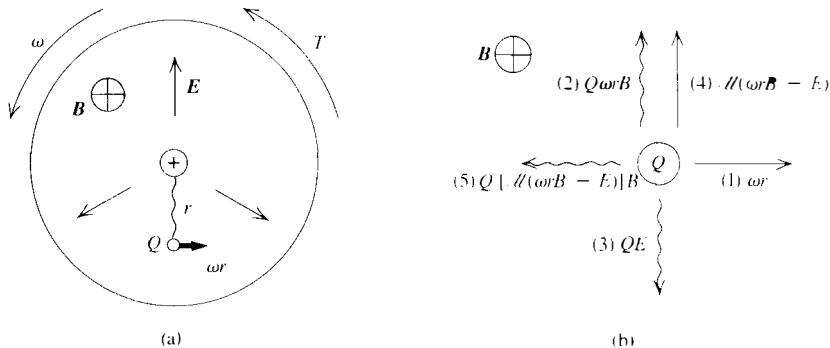


Fig. 22-12. This figure is similar to Fig. 22-11, except that it applies to a homopolar generator. (a) With ω and \mathbf{B} in the same directions as for the motor, the axle is positive and \mathbf{E} is also in the same direction. (b) The product $\omega r B$ is larger than E , and there is now a braking force (5).

The volume charge density is

$$\rho = \epsilon_0 \nabla \cdot \mathbf{E} = \epsilon_0 \frac{1}{r} \frac{\partial}{\partial r} (rE) = 2\epsilon_0 \omega B. \quad (22-38)$$

This charge density is uniform. It is positive with ω and B as in Fig. 22-12(a). It results from the $\mathbf{v} \times \mathbf{B}$ field. See Sec. 4.3.5 and Prob. 22-9.

(3) The voltage applied between the axis of radius a and the periphery at b is

$$V = \int_a^b \mathbf{E} \cdot d\mathbf{l} = \int_a^b \left(\omega Br + \frac{I}{2\pi s\sigma} \frac{1}{r} \right) dr \quad (22-39)$$

$$= \omega B \frac{b^2 - a^2}{2} + \frac{I}{2\pi s\sigma} \ln \frac{b}{a}. \quad (22-40)$$

The coefficient of I is the resistance of the disk between radii a and b (Sec. 4.3.2). Also, the coefficient of ω is the torque that we found above, divided by I . So

$$V = \frac{\omega T}{I} + IR. \quad (22-41)$$

We could have expected this: multiplying both sides by I gives

$$VI = \omega T + I^2 R. \quad (22-42)$$

This equation simply says that the power VI supplied by the source is equal to the mechanical power ωT plus the thermal power loss in the disk $I^2 R$. We have disregarded other losses.

Equation 22-42 shows that, for a given V , the current I is maximum when $\omega = 0$. From Eq. 22-36, T is maximum when I is maximum, and thus when $\omega = 0$. Then the mechanical power is zero. Conversely, from Eq. 22-40, I and T are both zero when

$$\omega = \omega_{\max} = \frac{2V}{B(b^2 - a^2)}. \quad (22-43)$$

This is the maximum angular velocity. The mechanical power is then also zero. Beyond this angular velocity, the device acts as a generator, with the axis as positive terminal, feeding current into the source V .

Under what condition is the mechanical power maximum? Set

$$\frac{d}{dl}(\omega T) = \frac{d}{dl}(VI - I^2R) = 0. \quad (22-44)$$

Then

$$I = \frac{V}{2R}, \quad VI = 2I^2R, \quad (22-45)$$

and the mechanical power is equal to the Joule power loss.

The mechanical power is maximum when half the energy supplied by the source dissipates as heat. Then the efficiency is 50%, and, from Prob. 22-13,

$$\omega_{\text{maximum power}} = \frac{V}{B(b^2 - a^2)}, \quad (22-46)$$

or $\omega_{\max}/2$.

Example

THE HOMOPOLAR GENERATOR

The homopolar generator, like the homopolar motor, is a low-voltage, high-current device. One application is the generation of the large currents required for purifying metals by electrolysis on an industrial scale.

Figure 22-11 applies, except that now ωrB is larger than E , as in Fig. 22-12. The above calculation also applies, with the same proviso.

The force density \mathbf{F}' is now clockwise. The magnetic torque exerted on the charge carriers

$$T = \frac{BI(b^2 - a^2)}{2} \quad (22-47)$$

is also clockwise and therefore opposes the motion.

Conservation of energy now requires that

$$\omega T = VI + I^2R = I^2(R_{\text{load}} + R), \quad (22-48)$$

where ωT is the mechanical power fed into the generator, VI is

the electric power fed to the load resistance, and I^2R is the thermal power dissipation. We have again disregarded other losses. Substituting the value of T and simplifying,

$$I = \frac{\omega B(b^2 - a^2)}{2(R_{\text{load}} + R)}. \quad (22-49)$$

The homopolar generator therefore acts as an ideal voltage source $\omega B(b^2 - a^2)/2$, with an output resistance equal to the resistance R of the disk.

22.5 SUMMARY

The force on a charge Q moving at a velocity \mathbf{v} in a field \mathbf{E} , \mathbf{B} is

$$\mathbf{F} = Q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (22-1)$$

This is the *Lorentz force*. The term $Q\mathbf{v} \times \mathbf{B}$ is the *magnetic force*.

Charge carriers flowing along a conductor situated in a magnetic field tend to drift sideways because of the magnetic force. This is the *Hall effect*.

The *magnetic force* per unit length on a current-carrying wire is $\mathbf{I} \times \mathbf{B}$. The total net force on a closed circuit is thus

$$\mathbf{F} = I \oint_C d\mathbf{l} \times \mathbf{B} \quad (22-14)$$

The magnetic force exerted by a closed circuit a on a closed circuit b is

$$\mathbf{F}_{ab} = \frac{\mu_0}{4\pi} I_a I_b \oint_a \oint_b d\mathbf{l}_b \times \frac{d\mathbf{l}_a \times \hat{\mathbf{r}}}{r^2} \quad (22-20)$$

$$= -\frac{\mu_0}{4\pi} I_a I_b \oint_a \oint_b \hat{\mathbf{r}} \frac{d\mathbf{l}_a \cdot d\mathbf{l}_b}{r^2}. \quad (22-23)$$

By definition, $\mu_0 = 4\pi \times 10^{-7}$ weber/ampere-meter. Then the force per meter between parallel wires carrying the same current I and separated by a distance D is $2 \times 10^{-7} I^2/D$. The force is attractive if the currents flow in the same direction.

The *magnetic force exerted on a volume distribution of current* is

$$\mathbf{F} = \int_v \mathbf{J} \times \mathbf{B} dv. \quad (22-31)$$

PROBLEMS

22-1. (22.1) Electrons in the Crab nebula[†]

In the Crab nebula there is a magnetic field of about 2×10^{-8} tesla and electrons whose energy is about 2×10^{14} electronvolts.

- (a) Find the radius of gyration. Compare this radius with that of the earth's orbit. (See the page inside the back cover).
- (b) How long does an electron take to complete one turn, in days?

22-2. (22.1) The pinch effect

A beam of charged particles of charge Q , mass m , and velocity v has a radius r . The beam current is I . Assume that the charge density is uniform. (This is a poor approximation; the current density as a function of the radius follows, in fact, a Gaussian curve.)

Find the outward force on an ion situated at the periphery of the beam. You will find that there is an outward electric force and an inward magnetic force. The magnetic force tends to "pinch" the beam, or to concentrate it along the axis.

If you cancel the electric force by adding ions of the opposite sign, then the magnetic force acts alone and the beam contracts. This phenomenon is easy to observe with positive ion accelerators. Residual gas in the path of the beam ionizes by impact, and the resulting low-energy electrons remain in the beam, while the positive ions drift away. If the pressure increases somewhat, the focusing improves. This is *gas focusing*.

At higher gas pressures the low-energy ions remain mostly inside the beam because their mean free path between collisions is shorter. The beam then becomes unstable because of phenomena that are not well understood at this time.

22-3. (22.1) The acceleration of an electron in a field $\mathbf{E}, \mathbf{B}^{\dagger}$

- (a) Show that the equation of motion for a particle of rest mass m_0 , charge Q , and velocity \mathbf{v} in a field \mathbf{E}, \mathbf{B} is

$$\gamma m_0 \frac{d\mathbf{v}}{dt} = Q \left(\mathbf{E} + \mathbf{v} \times \mathbf{B} - \frac{\mathbf{v}}{c^2} \mathbf{v} \cdot \mathbf{E} \right).$$

If \mathbf{E} is zero and \mathbf{B} is static,

$$\frac{d\mathbf{v}}{dt} = \mathbf{v} \times \frac{QB}{m} = \mathbf{v} \times \boldsymbol{\omega}_c,$$

and the electron describes a circle at the angular velocity $\boldsymbol{\omega}_c = QB/m$, which is called the *cyclotron frequency*.

- (b) A 12.0-MeV (million electronvolt) electron moves in the positive direction of the z -axis in a field $\mathbf{E} = 1.00 \times 10^6 \hat{z}$, $\mathbf{B} = 1.00 \hat{x}$.

Calculate its acceleration. The rest mass of an electron is 5.11×10^5 electronvolts.

[†] This problem requires a knowledge of relativity.

- 22-4. (22.1) The motion of a charged particle in uniform and perpendicular \mathbf{E} and \mathbf{B} fields

A particle of charge Q starts from rest at the origin in a region where $\mathbf{E} = E\hat{\mathbf{y}}$ and $\mathbf{B} = B\hat{\mathbf{z}}$.

- Find two simultaneous differential equations for v_x and v_y .
- Find v_x and v_y . Set $\omega_c = BQ/m$. This is the *cyclotron frequency*. Also, set $v_x = 0$, $v_y = 0$ at $t = 0$. There will be two constants of integration.
- Find $x(t)$ and $y(t)$. Set $x = 0$, $y = 0$ at $t = 0$.
- Describe the motion of the charge.
- Sketch the trajectory for $\omega_c = 1$, $E/B = 1$.

(f) Show that the particle drifts at the velocity $\mathbf{E} \times \mathbf{B}/B^2$. Note that this velocity is independent of the nature of the particle and of its energy. In a plasma, charges of both signs drift at the same velocity, and the net drift current is zero.

(g) Calculate the drift velocity of a proton at the equator under the combined actions of gravity and the \mathbf{B} of the earth. Assume that $B = 4 \times 10^{-5}$ tesla, in the horizontal direction. Remember that in the region of the *north geographic pole* there is a *south magnetic pole*. Thus, at the equator, \mathbf{B} points north.

- In which direction does an electron drift at the equator?

- 22-5. (22.1) The crossed-field photomultiplier

Figure 22-13 shows the principle of operation of a crossed-field photomultiplier. A sealed and evacuated enclosure contains two parallel plates called *dynodes*. They provide the electric field \mathbf{E} . An external permanent magnet superimposes the magnetic field \mathbf{B} .

A photon ejects a low-energy photoelectron. The electron accelerates upward, but the magnetic field deflects it back to the negative dynode. At this point it ejects a few secondary electrons, and the process repeats itself. Eventually, the electrons impinge on the collector C .

It is possible to obtain in this way about 10^5 electrons per photon and output currents as large as 100 microamperes. Since the time of flight is nearly the same for all the electrons, this type of photomultiplier can be used at data rates in excess of 1 megabit/second.

In actual practice there is an alternating voltage between the electrodes, but we simplify the problem by assuming a steady voltage.

Let us find the value of a .

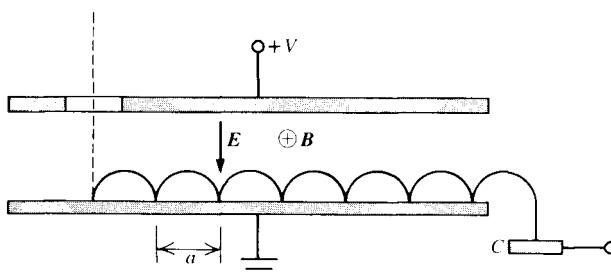


Fig. 22-13.

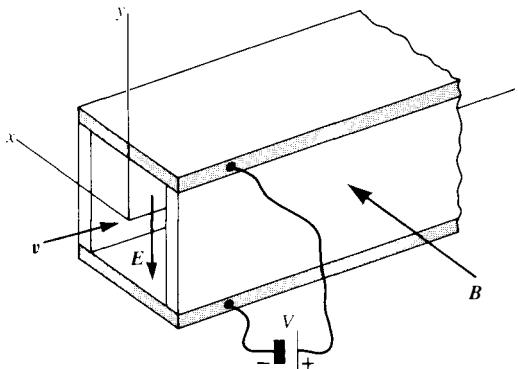


Fig. 22-14.

(a) Find the differential equations for v_x and v_y . The trajectory is not circular. You can simplify the calculation by setting $Be/m = \omega_c$, the cyclotron frequency.

(b) Find v_x as a function of y .

(c) You can now find y , and then x , as functions of t . Set $t = 0$, $dx/dt = 0$, and $dy/dt = 0$ at $x = 0$, $y = 0$. You should find that the trajectory is a cycloid.

(d) What is the maximum value of y ?

(e) What is the value of a ?

22-6. (22.1) Fermi acceleration[†]

Fermi proposed the following mechanism, now called *Fermi acceleration*, to explain the existence of very high-energy particles in space. Imagine a clump of plasma traveling at some velocity $v_c\hat{x}$. The plasma carries a current and thus has a magnetic field. Imagine now a particle traveling in the opposite direction at a velocity $-v_a\hat{x}$, both v_c and v_a being positive quantities. The particle is deflected in the magnetic field and acquires a velocity $v_b\hat{x}$, where v_b is also a positive quantity.

Set $v_c = v_a = c/2$. Calculate the initial and final values of γ .

It is now believed that cosmic rays acquire their energy, not in the interstellar medium, but in stars and in sources outside our galaxy.

22-7. (22.1) Electromagnetic pumps

Electromagnetic pumps are convenient for pumping highly conducting fluids, for example, liquid sodium in certain nuclear reactors. Their great advantage is that they have no moving parts, except the fluid. See Fig. 22-14.

The conduction current density in a liquid metal of conductivity σ that moves at a velocity \mathbf{v} in a field \mathbf{E} , \mathbf{B} is $\mathbf{J} = \sigma(\mathbf{E} + \mathbf{v} \times \mathbf{B})$. All quantities are measured with respect to a fixed reference frame. Then the magnetic force per unit volume is $\mathbf{F}' = \mathbf{J} \times \mathbf{B} = \sigma(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \times \mathbf{B}$. As a rule, \mathbf{E} , \mathbf{B} , and \mathbf{v} are orthogonal.

[†]This problem requires a knowledge of relativity.

(a) Show that $\mathbf{F}' = \sigma B^2(\mathbf{u} - \mathbf{v})$, where $\mathbf{u} = \mathbf{E} \times \mathbf{B}/B^2$ and \mathbf{v} is the fluid velocity, which is perpendicular to \mathbf{B} . This means that the magnetic force tries to make \mathbf{v} equal to \mathbf{u} .

(b) Calculate the efficiency, on the assumption that a permanent magnet supplies the magnetic field. Neglect edge and end effects. Is the efficiency high, or low?

22-8. (22.1) The Hall effect

Let us investigate the Hall effect more closely. We assume again that the charge carriers are electrons of charge $-e$. Their effective mass is m^* . The *effective mass* takes into account the periodic forces exerted on the electrons as they travel through the crystal lattice. As a rule, the effective mass is *smaller* than the mass of an isolated electron.

The force on an electron is $\mathbf{F} = -e(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, where \mathbf{E} has two components, the applied field E_x and the Hall field E_y . The average drift velocity is

$$\mathbf{v} = \frac{\mathcal{M}\mathbf{F}}{e} = -\mathcal{M}(\mathbf{E} + \mathbf{v} \times \mathbf{B}),$$

where \mathcal{M} is the mobility (Sec. 4.3.3). The law $\mathbf{F} = m^* \mathbf{a}$ applies only between collisions with the crystal lattice.

(a) Show that

$$v_x = -\mathcal{M}(E_x + v_y B), \quad v_y = -\mathcal{M}(E_y - v_x B), \quad v_z = 0.$$

(b) Show that

$$J_x = Ne\mathcal{M} \frac{E_x - \mathcal{M}E_y B}{1 + \mathcal{M}^2 B^2}, \quad J_y = Ne\mathcal{M} \frac{E_y + \mathcal{M}E_x B}{1 + \mathcal{M}^2 B^2}.$$

Thus, if $J_y = 0$,

$$E_y = -\mathcal{M}E_x B \quad \text{or} \quad V_y = \frac{b}{a} \mathcal{M}V_x B.$$

Note that the Hall voltage is proportional to the *product* of the applied voltage V_x and B . The Hall effect is thus useful for multiplying one variable by another.

When it is connected in this way, the Hall element has four terminals and is called a *Hall generator*, or a *Hall probe*.

(c) Calculate V_y for $b = 1$ millimeter, $a = 5$ millimeters, $\mathcal{M} = 7$ meters $^2/\text{volt-second}$ (indium antimonide), $V_x = 1$ volt, $B = 10^{-4}$ tesla.

(d) Show that, if $E_y = 0$, then

$$\frac{\Delta R}{R_0} = \mathcal{M}^2 B^2,$$

where R_0 is the resistance of the probe in the x -direction when $B = 0$, and ΔR is the increase in resistance upon application of the magnetic field.

The Hall field E_y can be made equal to zero by making c small, say a few micrometers, and plating conducting strips parallel to the y -axis, as in Fig.

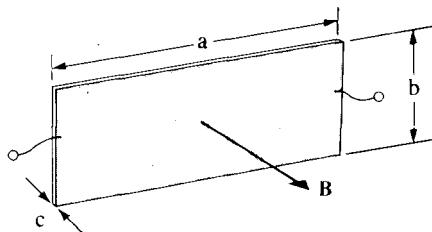


Fig. 22-15.

22-15. The element then has only two terminals and is called a *magnetoresistor*. Magnetoresistors are useful for measuring magnetic flux densities.

22-9. (22.1) The electromagnetic flowmeter

The *electromagnetic flowmeter* is the inverse of the electromagnetic pump (Prob. 22-7). It operates as follows. See Fig. 22-16. A conducting fluid flows in a nonconducting tube between the poles of a magnet. Electrodes on either side of the tube and in contact with the fluid measure the $\mathbf{v} \times \mathbf{B}$ field, and thus the quantity of fluid that flows through the tube per second. This is a Hall effect, except that here ions of both signs move with the fluid in the same direction.

Faraday attempted to measure the velocity of the Thames River in this way in 1832. The magnetic field was, of course, that of the earth.

In the absence of turbulence, the fluid velocity in a tube of radius R is of the form $v = v_0(1 - r^2/R^2)$. The $\mathbf{v} \times \mathbf{B}$ field in the fluid is therefore not uniform. This gives rise to circulating currents with $\mathbf{J} = \sigma(-\nabla V + \mathbf{v} \times \mathbf{B})$. The potential V_0 results from the charges that accumulate on the electrodes.

(a) Sketch a cross section of the tube, showing qualitatively, by means of arrows of various lengths, the magnitude and direction of $\mathbf{v} \times \mathbf{B}$.

(b) Sketch another cross section, showing the lines of current flow. The current drawn by the electrodes is negligible.

(c) Neglect end effects by setting $\partial/\partial z = 0$. Use the fact that $\nabla \cdot \mathbf{J} = 0$ to show that

$$\nabla^2 V = B \frac{\partial v}{\partial y} = B \frac{\partial v}{\partial r} \sin \phi.$$

Since this Laplacian is equal to $-\rho/\epsilon_0$, the volume charge density ρ is zero

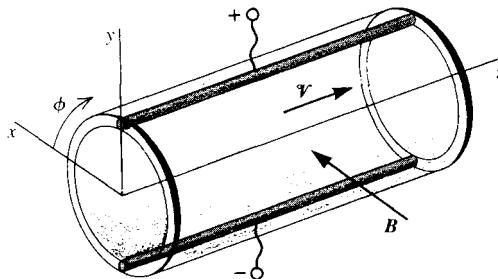


Fig. 22-16.

on the axis, where $\partial v / \partial r = 0$ and at $\phi = 0$, and at π . See page 402.

(d) Set $V = V' \sin \phi$, where V' is independent of ϕ , and show that

$$\frac{d^2V'}{dr^2} + \frac{1}{r} \frac{dV'}{dr} - \frac{V'}{r^2} = B \frac{dv}{dr}.$$

(e) You can solve this differential equation as follows: (1) express the left-hand side as a derivative, (2) integrate, (3) multiply both sides by r , (4) express both sides as derivatives, (5) integrate. This will leave you with two constants of integration, one of which is easy to dispose of. You can find the value of the other by remembering that $J_y = 0$ at $x = 0$, $y = R$ if the voltmeter draws zero current.

Note that the output voltage is independent of the conductivity of the fluid, if one assumes that the voltmeter draws zero current.

Find the output voltage as a function of the volume Q of fluid that flows in one second.

(f) We have neglected edge effects in the region where the fluid enters, and emerges from, the magnetic field. Sketch lines of current flow for these two regions. These currents reduce the output voltage somewhat.

22-10. (22.2) Improving (?) electric motors

(a) Someone suggests that, if the rotors of electric motors were wound with iron wire instead of copper wire, the torque could increase by a factor of 1000.

Show that the torque would indeed increase, but by a negligible amount. The increased Joule losses render the substitution undesirable.

(b) Here is another suggestion. The wires are located, not at the surface of the rotor, but rather in slots. Moving the wires to the surface of the rotor would place them in a stronger field and increase the torque by a factor of maybe 3 or 4.

You can show that this is incorrect if you are careful to distinguish between the magnetic field of the stator and that of the iron of the rotor.

22-11. (22.3) The force between two parallel currents.

Two long, straight, parallel wires of length $2L$ separated by a distance D carry equal currents I flowing in the same direction.

Calculate the force of attraction.

22-12. (22.3) The magnetic force law does not apply to the forces between single particles.[†]

We found that the force exerted between two charges Q_a and Q_b moving together at the same velocity v and a distance s apart is

$$F = \frac{(1 - \beta^2)^{1/2} Q_a Q_b}{4\pi\epsilon_0 s^2}.$$

Now calculate this same force from Coulomb's law and from the magnetic force law, substituting $Q_a v$ for $I_a dl_a$ and $Q_b v$ for $I_b dl_b$.

[†]This problem requires a knowledge of relativity.

22-13. (22.4) The pinch effect on a conductor

We have seen that the magnetic force density on a conductor is $\mathbf{J} \times \mathbf{B}$. If the force density is enormous, then any solid can be treated as a fluid, and if p is the pressure, $\nabla p = \mathbf{J} \times \mathbf{B}$.

(a) A conducting wire of radius R carries a current I . At the surface, $p = 0$. Show that at the radius r , $p = [\mu_0 I^2 / (4\pi^2 R^2)](1 - r^2/R^2)$. The magnetic force compresses the wire.

(b) Calculate the instantaneous pressure on the axis for a current of 30 kiloamperes in a wire 1 millimeter in radius. The wire, of course, vaporizes. Such large currents are obtained by discharging large, low-inductance capacitors.

(c) Show that, with a tubular conductor of inner radius R_1 and outer radius R_2 , the pressure in the cavity is given by

$$p = \frac{\mu_0 I^2}{4\pi^2 R_2^2} \frac{1 - (R_1/R_2)^2 [1 + 2 \ln(R_2/R_1)]}{[1 - (R_1/R_2)^2]^2}.$$

Disregard the skin effect. Transient pressures approaching 10^6 atmospheres have been obtained in this way.

One type of x-ray source implodes thin, aluminized plastic tubes by discharging 1 megajoule in them. The resulting plasma generates a 150-kilojoule pulse of radiation.

22-14. (22.4) The homopolar motor

(a) Find the mechanical power as a function of ω for a homopolar motor, for a given applied voltage V .

You can do this by first writing $P = \omega T = \omega B(b^2 - a^2)I/2 = \omega AI$, where A is a constant, and then expressing I as a function of ω .

(b) Sketch a curve of P as a function of ω .

(c) Show that the mechanical power is maximum at the angular velocity $\omega_{\text{maximum power}} = V/[B(b^2 - a^2)]$.

22-15. (22.4) The homopolar motor

Show that the centrifugal force on the conduction electrons in the rotor of a homopolar motor is completely negligible. Set $B = 1$, $\omega = 2\pi$. The ratio of the magnetic force to the centrifugal force is enormous because the ratio e/m is equal to about 2×10^{11} .

CHAPTER 23

MAGNETIC FIELDS VI

The Faraday Induction Law

23.1	THE $\mathbf{v} \times \mathbf{B}$ FIELD INSIDE A NONCONDUCTOR	413
23.2	MOTIONAL ELECTROMOTANCE. THE FARADAY INDUCTION LAW FOR $\mathbf{v} \times \mathbf{B}$ FIELDS	413
	<i>Example: A SIMPLE-MINDED GENERATOR</i>	414
	<i>Example: AN ALTERNATING-CURRENT GENERATOR</i>	418
23.3	LENZ'S LAW	419
23.4	FARADAY'S INDUCTION LAW FOR TIME-DEPENDENT \mathbf{B} 'S. THE CURL OF \mathbf{E}	420
23.4.1	LENZ'S LAW AGAIN	421
23.4.2	FLUX LINKAGE	421
	<i>Example: MEASURING B IN A ROWLAND RING</i>	422
	<i>Example: THE ENERGY DISSIPATED IN DESCRIBING A HYSTERESIS LOOP</i>	423
23.5	THE ELECTRIC FIELD STRENGTH \mathbf{E} EXPRESSED IN TERMS OF THE POTENTIALS V AND \mathbf{A}	424
	<i>Example: EDDY CURRENTS</i>	425
	<i>Example: THE INDUCED ELECTROMOTANCE IN A RIGID CIRCUIT</i>	426
23.6	THE \mathbf{E} , $-\nabla V$, $-\partial \mathbf{A}/\partial t$, AND $\mathbf{v} \times \mathbf{B}$ FIELDS	427
*23.7	RELATING THE TWO FORMS OF THE FARADAY INDUCTION LAW	428
*23.7.1	TRANSFORMATION OF A MAGNETIC FLUX	429
*23.7.2	TRANSFORMATION OF AN ELECTROMOTANCE	429
23.8	SIX KEY EQUATIONS	430
23.9	SUMMARY	431
	PROBLEMS	432

In the previous chapter we studied the magnetic forces exerted on moving charged particles and on currents situated in magnetic fields. These forces were of the form $Q\mathbf{v} \times \mathbf{B}$.

We now investigate the effect of a $\mathbf{v} \times \mathbf{B}$ field inside a moving macroscopic body. We shall see that in conducting bodies $\mathbf{v} \times \mathbf{B}$ acts like \mathbf{E} . We shall also see that a time-dependent magnetic field gives rise to an electric field of strength $-\partial \mathbf{A}/\partial t$, where \mathbf{A} is the vector potential, as

usual, and that $\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t$. The Faraday induction law neatly groups both phenomena.

23.1 THE $\mathbf{v} \times \mathbf{B}$ FIELD INSIDE A NONCONDUCTOR

(1) Suppose that a nonmagnetic nonconductor moves in some arbitrary fashion in a constant magnetic field. Then a charge Q carried along inside the body at a velocity \mathbf{v} , in a region where the magnetic flux density is \mathbf{B} , experiences a magnetic force $Q\mathbf{v} \times \mathbf{B}$.

As we noted in Sec. 22.1, the velocity \mathbf{v} can be any function of the coordinates and of the time. It can be uniform throughout the body, or it can vary from one point to another and with time.

Of course, \mathbf{B} can itself be any function of the coordinates. However, \mathbf{B} is constant, by hypothesis; we shall come to time-dependent magnetic fields in due time.

Now $\mathbf{v} \times \mathbf{B}$ has the dimensions of \mathbf{E} , because $Q\mathbf{v} \times \mathbf{B}$ is a force. Indeed, $\mathbf{v} \times \mathbf{B}$ adds to any \mathbf{E} that may be present. The polarization is therefore given by

$$\mathbf{P} = \epsilon_0 \chi_e (\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (23-1)$$

(2) If the nonconductor is magnetic, its equivalent currents, of course, follow the moving medium, but they can be time-dependent if the ambient \mathbf{B} is nonuniform. So the situation can be complex.

23.2 MOTIONAL ELECTROMOTANCE. THE FARADAY INDUCTION LAW FOR $\mathbf{v} \times \mathbf{B}$ FIELDS

Consider a closed circuit C that moves as a whole and distorts in some arbitrary way in a constant magnetic field, as in Fig. 23-1. Then, by definition, the *induced*, or *motional, electromotance* is

$$\mathcal{V} = \oint_C (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l} = -\oint_C \mathbf{B} \cdot (\mathbf{v} \times d\mathbf{l}). \quad (23-2)$$

The negative sign comes from the fact that we have altered the cyclic order of the terms under the integral sign.

Now $\mathbf{v} \times d\mathbf{l}$ is the area swept by the element $d\mathbf{l}$ in 1 second. Thus $\mathbf{B} \cdot (\mathbf{v} \times d\mathbf{l})$ is the rate at which the magnetic flux linking the circuit increases because of the motion of the element $d\mathbf{l}$. Integrating over the

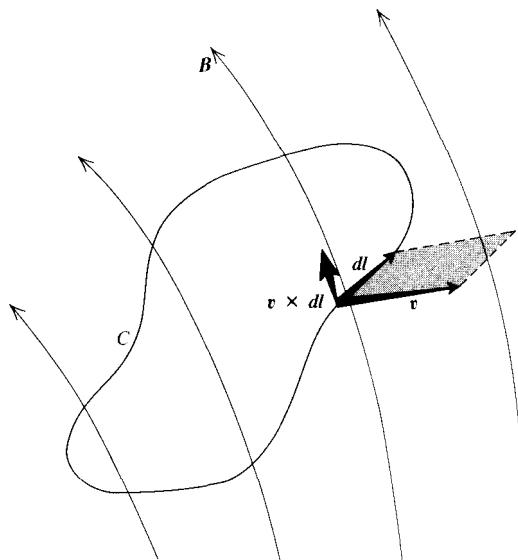


Fig. 23-1. Closed circuit C that moves and distorts in some arbitrary way in a constant magnetic field \mathbf{B} . The element $d\mathbf{l}$ moves at a velocity \mathbf{v} and sweeps an area $\mathbf{v} \times d\mathbf{l}$ in 1 second.

complete circuit, we find that the induced electromotive force is proportional to the time rate of change of the magnetic flux linking the circuit:

$$\mathcal{V} = -\frac{d\Phi}{dt}. \quad (23-3)$$

The positive directions for \mathcal{V} and for Φ satisfy the right-hand screw rule. The current is the same as if the circuit comprised a battery of voltage \mathcal{V} .

This is the *Faraday induction law for $\mathbf{v} \times \mathbf{B}$ fields*. This law is important. As far as our demonstration goes, it applies only to constant B 's, but it is, in fact, general, as we see in Sec. 23.4. Quite often Φ is difficult to define; then we can integrate $\mathbf{v} \times \mathbf{B}$ around the circuit to obtain \mathcal{V} .

If C is open, as in Fig. 23-2, then current flows until the electric field resulting from the accumulations of charge exactly cancels the $\mathbf{v} \times \mathbf{B}$ field.

Example

A SIMPLE-MINDED GENERATOR

An electric generator transforms mechanical energy to electric energy, usually by moving conducting wires in a direction perpendicular to a magnetic field.

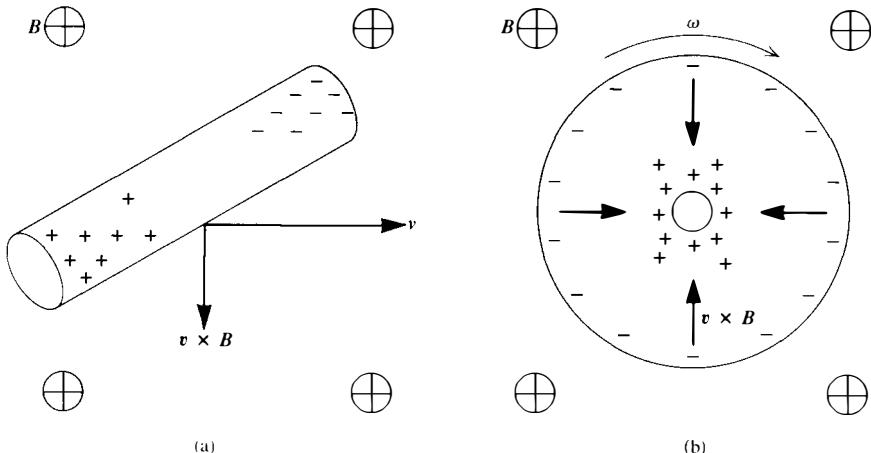


Fig. 23-2. (a) Conducting rod moving at a velocity v in a magnetic field B . The $v \times B$ field drives conduction electrons upward. Current flows until the electric field E exactly cancels the $v \times B$ field. (b) Disk rotating in a magnetic field B . Here the $v \times B$ field causes conduction electrons to move outward. A radial current flows until the electric field cancels the $v \times B$ field.

The simplest (and most impractical!) type of generator is that of Fig. 23-3(a). The link slides to the right at a speed v such that $v^2 \ll c^2$, where c is the speed of light, in a uniform \mathbf{B} that is normal to the paper. The resistance at the left-hand end of the line is R , and that of the link is R_l . The horizontal wires have zero resistance.

The electromotance is

$$\mathcal{V} = -\frac{d\Phi}{dt} = BDv. \quad (23-4)$$

We have disregarded the magnetic flux resulting from the current I itself. In other words, the resistance R is large. Then

$$I = \frac{BDv}{R + R_l}. \quad (23-5)$$

A conduction electron inside the link follows along at the velocity v , and the $v \times B$ field prods it downward. Its velocity v_r , relative to the link is strictly vertical, under steady-state conditions, because otherwise negative charges would accumulate indefinitely along one edge of the link, leading to an infinite potential. See Fig. 23-3(b).

In a fixed reference frame S , the force on a conduction electron of charge Q inside the link is $Q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$. Thus, in the link,

$$\mathbf{J} = \sigma(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = \sigma(-\nabla V + \mathbf{v} \times \mathbf{B}). \quad (23-6)$$

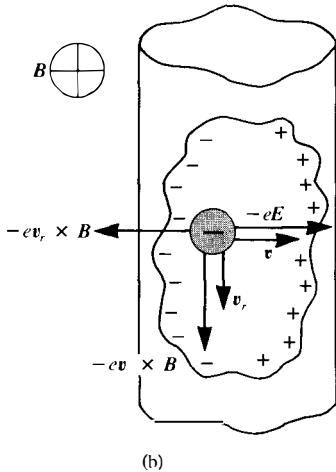
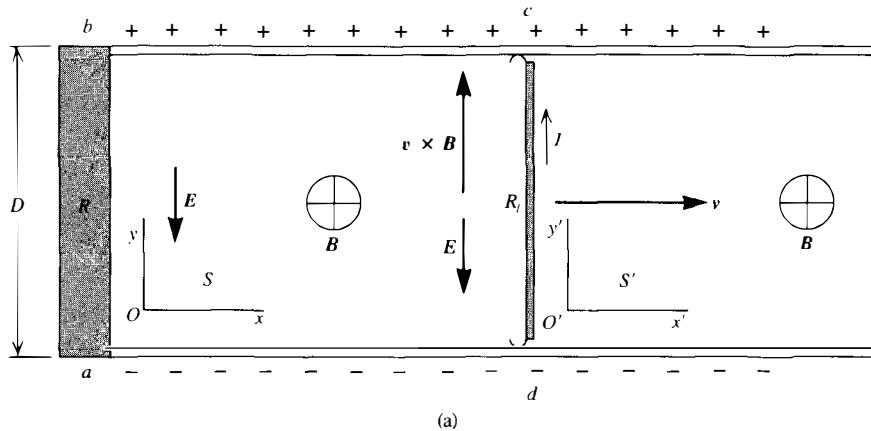


Fig. 23-3. (a) Simple-minded electric generator. Sliding the link to the right at a velocity v generates a current I in the direction shown. The magnetic field is constant and uniform. The resistance R is the load. (b) Section through the link: v is the velocity of the wire, v_r the velocity of a conduction electron relative to the link, E the Hall field. The Hall field exactly cancels the $v_r \times B$ field.

Let us calculate V in S .

At b in Fig. 23-3(a), $V_b = IR$. In either horizontal wire, $\mathbf{J} = \sigma \mathbf{E}$ is finite. Since $\sigma \rightarrow \infty$, by hypothesis, then $\mathbf{E} = 0$, $\nabla V = 0$, and

$$V_d = V_a = 0, \quad V_c = V_b = IR. \quad (23-7)$$

Inside R and R_l , with the y -axis as in the figure,

$$V = IR \frac{y}{D} = \frac{\nu BR}{R + R_l} y. \quad (23-8)$$

The voltage V_c across R_l is IR :

$$V_c = IR = I(R + R_l) - IR_l = \nu BD - IR_l. \quad (23-9)$$

This means that the motion generates a voltage νBD in the link, while its current causes a voltage drop IR_l .

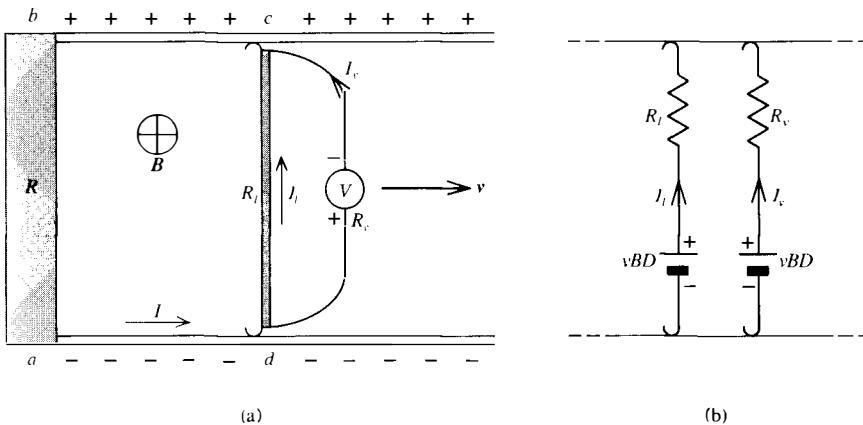


Fig. 23-4. (a) Voltmeter V connected across the link in the generator of Fig. 23-3.
 (b) Equivalent circuit.

Suppose we connect a voltmeter across the link as in Fig. 23-4(a). Call the resistance R_v , with $R_v \gg R_l$. This hardly affects the current I . What will be the reading on the voltmeter? If the current through the voltmeter is I_v , then it will read a voltage $I_v R_v$, with the polarity shown in the figure.

Now refer to Fig. 23-4(b). Clearly, $I_v R_v = I_t R_t$ and the voltmeter reads the voltage drop $I_t R_t$.

Chapter 17, on relativity, is a prerequisite for the rest of this example.

All we know about the magnetic field is that $\mathbf{B} = -B\hat{\mathbf{z}}$; the current distribution that generates B is unspecified. Let us set

$$A_x = nBy, \quad A_y = (n+1)Bx, \quad (23-10)$$

where n is a pure number. It is a simple matter to check that $\mathbf{B} = \nabla \times \mathbf{A}$.

If $n = 0$, then the currents supplying B are all vertical. If $n = -1$, they are all horizontal. With a solenoid whose axis coincides with the z -axis, $n = -\frac{1}{2}$. Therefore, inside R and R_1 ,

$$E = -\nabla V - \frac{\partial A}{\partial t} = -\nabla V \quad (23-11)$$

$$= - \frac{vBR}{R + R_i} \hat{y}. \quad (23-12)$$

Let us now see what happens inside the link, in its own reference frame S' . We assume that $v^2 \ll c^2$, which makes $\gamma \approx 1$.

From Sec. 17.8,

$$A'_x = A_x - \frac{vV}{c^2} = nBy - \frac{v^2}{c^2} \frac{BR}{R + R'} y \approx nBy, \quad (23-13)$$

$$A'_y = A_y = (n + 1)Bx = (n + 1)Bvt, \quad (23-14)$$

$$V' = V - vA_x = \frac{vBR}{R + R_l} y - vnBy = \left(\frac{R}{R + R_l} - n \right) vBy. \quad (23-15)$$

$$V'_c = \left(\frac{R}{R + R_l} - n \right) vBD. \quad (23-16)$$

Note that the values of A' and of V' depend on the value of n . In other words, they depend on the particular geometry of the coils selected for generating \mathbf{B} . Observe also the appearance of a $\partial A'/\partial t$ term in S' .

Now

$$\mathbf{E}' = -\nabla' V' - \frac{\partial \mathbf{A}'}{\partial t'}, \quad (23-17)$$

where

$$\nabla' = \frac{\partial}{\partial y'}, \hat{\mathbf{y}} = \frac{\partial}{\partial y'} \hat{\mathbf{y}}', \quad t' = t - \frac{v}{c^2} x = t - \frac{v^2}{c^2} t \approx t. \quad (23-18)$$

Thus

$$\mathbf{E}' = -\frac{\partial V'}{\partial y'} \hat{\mathbf{y}} - \frac{\partial \mathbf{A}'}{\partial t'} = -\frac{\partial V'}{\partial y'} \hat{\mathbf{y}} - \frac{\partial A'_y}{\partial t} \hat{\mathbf{y}} \quad (23-19)$$

$$= \left[-\left(\frac{R}{R + R_l} - n \right) vB - (n + 1)vB \right] \hat{\mathbf{y}} \quad (23-20)$$

$$= -\left(\frac{R}{R + R_l} - 1 \right) vB \hat{\mathbf{y}} = \frac{R_l}{R + R_l} vB \hat{\mathbf{y}}. \quad (23-21)$$

In general, $-\partial \mathbf{A}'/\partial t'$ is *not* equal to $\mathbf{v} \times \mathbf{B}$.

The quantity n has disappeared! We could have expected this because, clearly, \mathbf{E}' must be independent of the configuration of the coils that generate the given magnetic field.

We could also have found \mathbf{E}' directly, by simply transforming \mathbf{E} , with $v^2 \ll c^2$:

$$\mathbf{E}' = \mathbf{E}_\parallel + (\mathbf{E}_\perp + \mathbf{v} \times \mathbf{B}) = \mathbf{E}_\perp + \mathbf{v} \times \mathbf{B} \quad (23-22)$$

$$= \left(-\frac{vBR}{R + R_l} + vB \right) \hat{\mathbf{y}} = \frac{R_l}{R + R_l} vB \hat{\mathbf{y}}, \quad (23-23)$$

as above. This shows that, in the moving reference frame of the link, \mathbf{E}' is equal to \mathbf{E} plus $\mathbf{v} \times \mathbf{B}$.

The current I' is equal to I , and the voltmeter reading is the same for an observer in S' as for an observer in S .

Example

AN ALTERNATING-CURRENT GENERATOR

The loop of Fig. 23-5 rotates at an angular velocity ω in a uniform, constant \mathbf{B} . We calculate the induced electromotance \mathcal{V} , first through $\mathbf{v} \times \mathbf{B}$ and then through $d\Phi/dt$.

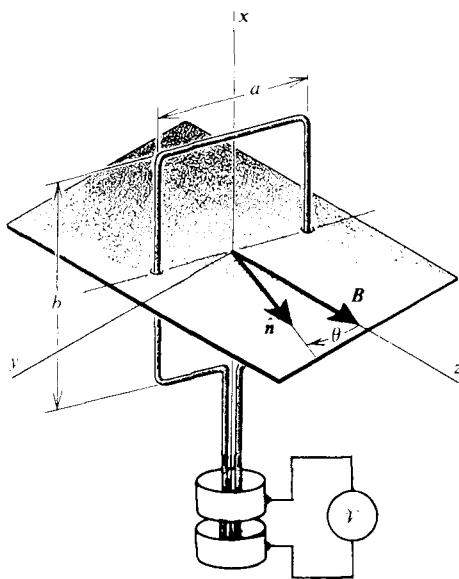


Fig. 23-5. Loop rotating in a constant and uniform magnetic field \mathbf{B} . The slip rings provide contacts between the voltmeter and the loop.

(a) Along the right-hand side of the loop,

$$b\mathbf{v} \times \mathbf{B} = \frac{\omega a}{2} B b \sin \theta \hat{x} = \frac{wabB}{2} \sin \omega t \hat{x}. \quad (23-24)$$

Along the left-hand side we have the same induced electromotance, but directed downward as in the figure. Along the upper and lower sides, $\mathbf{v} \times \mathbf{B}$ is perpendicular to the wire. This crowds the conduction electrons sideways, thereby increasing the resistance imperceptibly, but $\mathbf{v} \times \mathbf{B}$ contributes nothing to the electromotance. So

$$\mathcal{V} = abB\omega \sin \omega t. \quad (23-25)$$

Notice that there is zero electromotance when $\omega t = n\pi$, where n is a whole number. Then \mathbf{v} and \mathbf{B} are parallel, and $\mathbf{v} \times \mathbf{B}$ is zero.

(b) The time rate of change of the magnetic flux gives the same result:

$$\mathcal{V} = -\frac{d\Phi}{dt} = -\frac{d}{dt} abB \cos \omega t = abB\omega \sin \omega t. \quad (23-26)$$

23.3 LENZ'S LAW

Now let us return to Fig. 23-3(a). Observe that sliding the bar to the right increases the linking flux, but the induced current tends to decrease it. So the induced electromotance drives a current whose field opposes a *change*

in the net magnetic flux linking the circuit. This is *Lenz's law*. If the circuit was superconducting, the enclosed flux would remain constant.

We return to Lenz's law later.

23.4 FARADAY'S INDUCTION LAW FOR TIME-DEPENDENT \mathbf{B} 's. THE CURL OF \mathbf{E}

Imagine now two closed and rigid circuits as in Fig. 23-6. The active circuit a is stationary, while the passive circuit b moves in some arbitrary way, say in the direction of a as in the figure. The current I_a is constant.

From Sec. 23.2, the electromotance induced in circuit b is

$$\mathcal{V} = \oint_b (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l} = -\frac{d\Phi}{dt}, \quad (23-27)$$

where Φ is the magnetic flux linking b . This seems trivial, but it is not, because $d\Phi/dt$ could be the same if both circuits were stationary and if I_a changed appropriately. This means that the *Faraday induction law*

$$\mathcal{V} = -\frac{d\Phi}{dt} \quad (23-28)$$

applies whether there are moving conductors in a constant \mathbf{B} or stationary conductors in a time-varying \mathbf{B} . However, our argument is no more than plausible. A proper demonstration follows at the end of this chapter. It requires relativity.

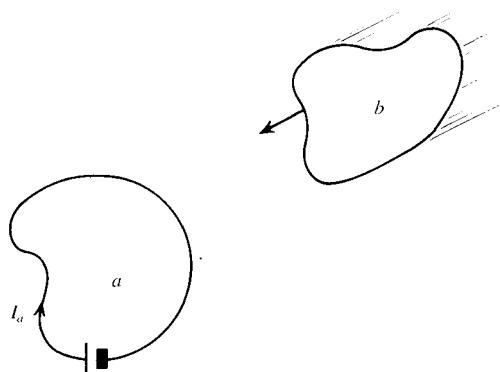


Fig. 23-6. Circuit a is active and fixed in position. Circuit b is passive and moves in the field of a .

Assuming the correctness of the above result, the electromotance induced in a rigid and stationary circuit C lying in a time-varying magnetic field is

$$\mathcal{V} = \oint_C \mathbf{E} \cdot d\mathbf{l} = \int_A (\nabla \times \mathbf{E}) \cdot d\mathcal{A} = - \frac{d\Phi}{dt} = - \int_A \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathcal{A}. \quad (23-29)$$

We have used Stokes's theorem in going from the first to the second integral, \mathcal{A} being an arbitrary surface bounded by C . Also, we have a partial derivative under the last integral sign, to take into account the fact that the magnetic field can be a function of the coordinates as well as of the time. The right-hand screw rule applies.

The path of integration need not lie in conducting material.

Observe that the above equation involves only the integral of $\mathbf{E} \cdot d\mathbf{l}$. It does *not* give \mathbf{E} as a function of the coordinates, except for simple geometries, and only after integration.

Since the surface of area \mathcal{A} chosen for the surface integrals is arbitrary, the equality of the third and last terms above means that

$$\nabla \times \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t}. \quad (23-30)$$

This is yet another of Maxwell's equations. This equation, like the other two (Eqs. 9-15 and 18-19), is valid on the condition that all the variables relate to the same reference frame.

23.4.1 Lenz's Law Again

The negative sign in Eqs. 23-28 and 23-30 is important. If Φ points into the paper and increases, then $d\Phi/dt$ points into the paper. Then, according to the right-hand screw rule, the negative sign means that the induced electromotance is counterclockwise.

Observe that the induced electromotance tends to generate a magnetic field that counters the imposed *change* in flux. So Lenz's law of Sec. 23.3 applies here also.

23.4.2 Flux Linkage

If the closed circuit comprises N turns, each intercepting the same magnetic flux, then the electromotances add and the net electromotance is N times larger. Then the quantity $N\Phi$ is termed the *flux linkage*:

$$\Lambda = N\Phi \quad \text{and} \quad \mathcal{V} = - \frac{d\Lambda}{dt}. \quad (23-31)$$

Of course, the geometry of the circuit and the configuration of the field can be quite complex. Then this equation still applies and the geometric meaning of Λ becomes obscure, but \mathcal{V} , and hence Λ , are measurable quantities.

Example

MEASURING B IN A ROWLAND RING

We are now in a position to understand how to measure B in the Rowland ring of Sec. 21.2 to observe the hysteresis loop of a sample of ferromagnetic material.

Let the low-frequency alternating current through winding a in Fig. 23-7 be

$$I = I_m \cos \omega t. \quad (23-32)$$

Then, from Sec. 20.6, H is a sinusoidal function of the time and, if winding a has N_a turns,

$$H = \frac{N_a I_m \cos \omega t}{2\pi r}. \quad (23-33)$$

We assume that the minor radius of the ring is small compared to its major radius. Then r is the mean radius.

If the magnetic flux density, averaged over the cross section, is B , then the magnetic flux is $B\mathcal{A}$ and, disregarding signs, the electromotance induced in the N_b turns of winding b is

$$\mathcal{V} = N_b \frac{d\Phi}{dt} = N_b \mathcal{A} \frac{dB}{dt}. \quad (23-34)$$

This is the voltage at the terminals of winding b if the measuring device draws essentially zero current.

To obtain B , we connect an integrating circuit (Prob. 7-9) to the terminals of winding b , as in Fig. 23-7. Then, with an operational amplifier of gain $A \gg 1$, and if $A\omega RC \gg 1$,

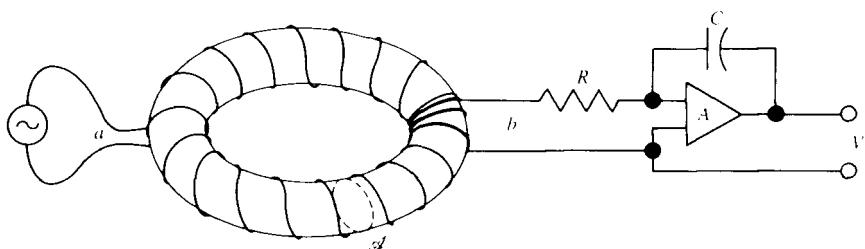


Fig. 23-7. The Rowland ring. This is basically a transformer with a primary a and a secondary b . The secondary is connected to an integrating circuit. A is an operational amplifier. The output alternating voltage V is proportional to the magnetic flux density B in the core.

$$V = -\frac{1}{RC} \int_0' N_b \mathcal{A} \frac{dB}{dt} dt = -\frac{N_b \mathcal{A}}{RC} B. \quad (23-35)$$

We have set $V = 0$ at $t = 0$.

So the voltage at the output of the integrating circuit is equal to a known constant times B .

Example

THE ENERGY DISSIPATED IN DESCRIBING A HYSTERESIS LOOP

We can now also calculate the energy dissipated in describing the hysteresis loop of Fig. 21-2. We refer to the Rowland ring of Sec. 21.2. When the current I_a increases, the flux Φ in the core increases, and the electromotance $-N_a d\Phi/dt$ induced in winding a opposes this increase, according to Lenz's law. The extra power spent by the source is then

$$\frac{d\mathcal{E}}{dt} = I_a \left(N_a \frac{d\Phi}{dt} \right) = I_a N_a \mathcal{A} \frac{dB}{dt}, \quad (23-36)$$

if \mathcal{A} is the cross-sectional area of the sample. Also,

$$\frac{d\mathcal{E}}{dt} = \frac{I_a N_a}{2\pi r} 2\pi r \mathcal{A} \frac{dB}{dt} = Hv \frac{dB}{dt}, \quad (23-37)$$

where $v = 2\pi r \mathcal{A}$ is the volume of the sample and

$$\mathcal{E}_1 = v \int_g^b H dB \quad (23-38)$$

is the energy supplied by the source in going from g to b on the loop of Fig. 21-2. This integral is equal to the area defined by the points $agbh$ and to the energy supplied by the source per unit volume of core.

When the current is in the same direction but decreasing, the polarity of the induced electromotance reverses, and in going from b to c , the energy

$$\mathcal{E}_2 = v \int_b^c H dB \quad (23-39)$$

returns to the source.

Proceeding in this way all around the loop, we find that, in the course of one complete cycle, the source supplies an energy

$$\mathcal{E} = v \oint H dB, \quad (23-40)$$

which is v times the area enclosed by the hysteresis loop.

23.5 THE ELECTRIC FIELD STRENGTH \mathbf{E} EXPRESSED IN TERMS OF THE POTENTIALS V AND \mathbf{A}

An arbitrary, rigid, and stationary closed circuit C lies in a time-dependent \mathbf{B} . Then, from Sec. 23.4,

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{d}{dt} \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A}, \quad (23-41)$$

where \mathcal{A} is the area of any open surface bounded by C .

Now, from Sec. 19.1, we can replace the surface integral on the right by the line integral of the vector potential \mathbf{A} around C :

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{d}{dt} \oint_C \mathbf{A} \cdot d\mathbf{l} = -\oint_C \frac{\partial \mathbf{A}}{\partial t} \cdot d\mathbf{l}. \quad (23-42)$$

There is no objection to inserting the time derivative under the integral sign, but then it becomes a partial derivative because \mathbf{A} is normally a function of the coordinates as well as of the time.

Thus

$$\oint_C \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) \cdot d\mathbf{l} = 0, \quad (23-43)$$

where C is a closed curve, as stated above. Then, from Sec. 1.9.1, the expression enclosed in parentheses is equal to the gradient of some function:

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\nabla V, \quad (23-44)$$

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t}, \quad (23-45)$$

where V is, of course, the electric potential.

So \mathbf{E} is the sum of two terms, $-\nabla V$ that results from accumulations of charge and $-\partial \mathbf{A} / \partial t$ whenever there are time-dependent fields in the given reference frame.

This is an important equation; we shall use it repeatedly. Observe that it expresses \mathbf{E} itself, *not* its derivatives or its integral, at a given point in terms of the derivatives of V and of \mathbf{A} in the region of that point. Its magnetic equivalent is $\mathbf{B} = \nabla \times \mathbf{A}$ (Sec. 18.4).

The Faraday induction law, in differential form (Eq. 23-30), relates space derivatives of \mathbf{E} to the time derivative of \mathbf{B} at a given point.

Observe that ∇V is a function of V , which depends on the positions of the charges. However, $\partial \mathbf{A} / \partial t$ is a function of the time derivative of the current density \mathbf{J} , hence of the acceleration of the charges.

The relations

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t} \quad \text{and} \quad \mathbf{B} = \nabla \times \mathbf{A} \quad (23-46)$$

are always valid in any given inertial reference frame.[†]

In a time-dependent \mathbf{B} , the electromotance induced in a circuit C is

$$\mathcal{V} = - \int_C \frac{\partial \mathbf{A}}{\partial t} \cdot d\mathbf{l}. \quad (23-47)$$

Example

EDDY CURRENTS

Imagine a sheet of copper lying inside a solenoid, in a plane perpendicular to the axis. The solenoid carries an alternating current.

According to Lenz's law, the $-\partial \mathbf{A} / \partial t$ electric field induces currents in the copper that tend to cancel the changes in the net \mathbf{B} . These currents are azimuthal because \mathbf{A} and $\partial \mathbf{A} / \partial t$ are azimuthal (example in Sec. 19.1).

Currents induced in bulk conductors by changing magnetic fields are termed *eddy currents*.

Eddy currents can be useful. For example, they dissipate energy in various damping mechanisms.

They are harmful in transformers because they cause Joule losses in the core. Transformer cores are usually assembled from thin sheets of transformer iron, called *laminations*, a fraction of a millimeter thick in small units, insulated from each other by a thin layer of oxide. With a solid core the eddy currents would largely cancel changes in magnetic flux. Also, the Joule losses would be excessive.

In audio transformers the iron alloy is sometimes in the form of a powder molded in an insulating binder. The transformer is then said to have a *powdered iron core*.

Ferrites serve at audio frequencies and above. These are ceramic-type materials that are molded from oxides of iron and of various other metals. Their main advantage is that their electric conductivity is low, of the order of 1 siemens/meter, or 10^{-6} times that of transformer iron. Eddy-current losses in ferrites are thus easily manageable.

[†] Unless stated otherwise, all the reference frames that we refer to are inertial: they do not accelerate, and they not rotate.

Example

THE INDUCED ELECTROMOTANCE IN A RIGID CIRCUIT

In the rigid circuit of Fig. 23-8, we assume that (1) the horizontal wires have zero resistance, (2) the resistance R_a is a long distance away from I' , and (3) both R_a and R_b are large enough to render the magnetic field of I negligible compared to that of I' . Thus \mathbf{B} and \mathbf{A} are essentially those of I' and point in the directions shown. According to Lenz's law (Sec. 23.4.1), an increase in I' induces an electromotance and a current I in the counterclockwise direction.

One may ascribe the induced electromotance either to the changing magnetic flux Φ or to the electric field strength $-\partial\mathbf{A}/\partial t$ in R_b , with

$$\Phi = \oint_C \mathbf{A} \cdot d\mathbf{l} \approx A_b D. \quad (23-48)$$

The line integral runs clockwise because of the right-hand screw rule. The induced electromotance is thus

$$\mathcal{V} = -\frac{\partial\Phi}{\partial t} = -\frac{\partial A_b}{\partial t} D. \quad (23-49)$$

This is counterclockwise if Φ points into the paper and increases, or if I' increases. Thus

$$I = \frac{(\partial A_b / \partial t) D}{R_a + R_b} \quad (23-50)$$

in the direction shown, again if I' increases.

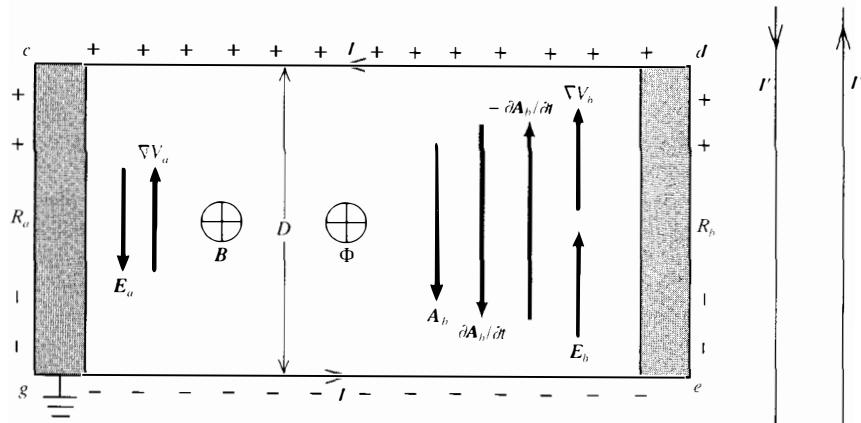


Fig. 23-8. Rigid circuit $gcde$ terminated by resistances R_a and R_b and situated near a pair of wires carrying a current I' . The lengths of the arrows show the relative magnitudes of \mathbf{E} , ∇V , and $\partial\mathbf{A}/\partial t$, with $R_a = R_b$.

Inside the resistance R_a , $A \approx 0$ by hypothesis and

$$\mathbf{E}_a = -\nabla V_a = -\frac{IR_a}{D}\hat{\mathbf{y}}, \quad V_c = IR_a, \quad (23-51)$$

with \mathbf{E}_a and ∇V_a pointing as in the figure.

The potential at e is zero for the following reason. Inside the horizontal wires, $\mathbf{E} = \mathbf{J}/\sigma$, where \mathbf{J} has some finite value and σ is infinite, by hypothesis. So

$$E_x = 0, \quad -\frac{\partial V}{\partial x} - \frac{\partial A_x}{\partial t} = 0. \quad (23-52)$$

But A_x is everywhere zero because the currents I' have no x -component. So

$$\frac{\partial V}{\partial x} = 0, \quad V_e = 0. \quad (23-53)$$

Similarly,

$$V_c = V_d = IR_a. \quad (23-54)$$

This potential on the cd wire results from the accumulation of positive surface charges on the upper half of the circuit, and negative charges on the lower half, as in the figure.

Inside R_b , therefore,

$$\nabla V_b = \frac{IR_b}{D}\hat{\mathbf{y}} \quad (23-55)$$

points upward, and

$$\mathbf{E}_b = -\nabla V_b - \frac{\partial \mathbf{A}_b}{\partial t} = \frac{IR_b}{D}\hat{\mathbf{y}} \quad (23-56)$$

also points upward. Of course, \mathbf{E} points in the same direction as \mathbf{J} .

If $R_a = R_b$, then

$$|\mathbf{E}_a| = |\mathbf{E}_b|, \quad \left| \frac{\partial \mathbf{A}_b}{\partial t} \right| = 2 |\nabla V_b|. \quad (23-57)$$

23.6 THE \mathbf{E} , $-\nabla V$, $-\partial \mathbf{A}/\partial t$, AND $\mathbf{v} \times \mathbf{B}$ FIELDS

In any given inertial reference frame, say S , the equation

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t} \quad (23-58)$$

always applies.

If a charge Q moves at a velocity \mathbf{v} with respect to S , then for an observer on S the force is

$$\mathbf{F} = Q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = Q\left(-\nabla V - \frac{\partial \mathbf{A}}{\partial t} + \mathbf{v} \times \mathbf{B}\right). \quad (23-59)$$

All the variables are measured with respect to the *same* reference frame S . These equations are valid even if v approaches the speed of light.

For an observer on the moving body, say in reference frame S' , the body is at rest and

$$\mathbf{F}' = QE' = Q\left(-\nabla' V' - \frac{\partial \mathbf{A}'}{\partial t'}\right). \quad (23-60)$$

*23.7 RELATING THE TWO FORMS OF THE FARADAY INDUCTION LAW[†]

We found above that if a rigid circuit moves in a constant \mathbf{B} , then the induced electromotance follows Faraday's induction law. Then we concluded that the same law applies to a stationary circuit lying in a time-dependent \mathbf{B} .

Passing from one form of the law to the other requires relativity.

In the rest of this chapter we shall call an electromotance \mathcal{V}_{em} , and the velocity of S' with respect to S \mathcal{V}_{vel} .

Equation 23-28 refers to the induced electromotance, as measured in the fixed reference frame S . From our experience with relativistic calculations, it is by no means evident that, in the frame S' of the moving circuit,

$$\mathcal{V}'_{em} = -\frac{d\Phi'}{dt'}. \quad (23-61)$$

That is, in fact, true because, as we shall see,

$$\mathcal{V}'_{em} = \gamma \mathcal{V}_{em}, \quad \Phi' = \Phi, \quad dt' = \frac{dt}{\gamma}. \quad (23-62)$$

We can immediately accept the third equation for the following reason. For any point in S' ,

$$t = \gamma\left(t' + \frac{\mathcal{V}_{vel}x'}{c^2}\right), \quad (23-63)$$

[†] Relativity is a prerequisite for Sec. 23.7.

from Sec. 13.4. Thus, at any given point in S' , x' is fixed and

$$dt = \gamma dt'. \quad (23-64)$$

Proving the other two equations takes a bit longer.

23.7.1 Transformation of a Magnetic Flux

The magnetic flux linking a given closed circuit bounding an area \mathcal{A} is

$$\Phi = \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} \quad (23-65)$$

in a reference frame S and

$$\Phi' = \int_{\mathcal{A}'} \mathbf{B}' \cdot d\mathcal{A}' \quad (23-66)$$

in S' .

Of course, the surface of area \mathcal{A} in frame S has a different shape in frame S' , because of the Lorentz contraction, and a different area \mathcal{A}' . A given element of area (say it is painted red) carries a flux $\mathbf{B} \cdot d\mathcal{A}$ in frame S , and $\mathbf{B}' \cdot d\mathcal{A}'$ in S' . Thus

$$\frac{d\Phi'}{d\Phi} = \frac{\mathbf{B}' \cdot d\mathcal{A}'}{\mathbf{B} \cdot d\mathcal{A}}. \quad (23-67)$$

But, from Secs. 14.1.1 and 16.6, and setting $\mathbf{E} = 0$,

$$\mathbf{B}' \cdot d\mathcal{A}' = (\mathbf{B}_{||} + \gamma \mathbf{B}_{\perp}) \cdot \left(d\mathcal{A}_{||} + \frac{d\mathcal{A}_{\perp}}{\gamma} \right) \quad (23-68)$$

$$= \mathbf{B}_{||} \cdot d\mathcal{A}_{||} + \mathbf{B}_{\perp} \cdot d\mathcal{A}_{\perp} = \mathbf{B} \cdot d\mathcal{A}, \quad (23-69)$$

and $\Phi' = \Phi$. *The magnetic flux linking a rigid closed curve is invariant.*

*23.7.2 Transformation of an Electromotance

Refer again to Fig. 23-6 and call S' the reference frame of the moving circuit. In S' ,

$$\mathcal{V}'_{em} = \oint \mathbf{E}' \cdot dl'. \quad (23-70)$$

Since we are only interested in motional electromotance for the moment, we may assume that \mathbf{B} is constant and \mathbf{E} is zero in S . Then, from Sec. 16-6,

$$\mathcal{V}'_{\text{em}} = \oint \gamma (\mathcal{V}_{\text{vel}} \times \mathbf{B}) \cdot d\mathbf{l}'. \quad (23-71)$$

Now the vector product is perpendicular to \mathcal{V}_{vel} . It is therefore only the perpendicular component of $d\mathbf{l}'$ that matters, $d\mathbf{l}'_{\perp} = d\mathbf{l}_{\perp}$, and

$$\mathcal{V}'_{\text{em}} = \gamma \oint (\mathcal{V}_{\text{vel}} \times \mathbf{B}) \cdot d\mathbf{l} = \gamma \mathcal{V}_{\text{em}}. \quad (23-72)$$

We have therefore proved all three equations 23-62. As a consequence, we have shown that, under any circumstance, but in a single reference frame, the electromotance induced in a closed circuit associated with a changing magnetic flux is given by

$$\mathcal{V}_{\text{em}} = - \frac{d\Phi}{dt}. \quad (23-73)$$

The positive directions chosen for \mathcal{V}_{em} and for Φ follow the right-hand screw rule.

23.8 SIX KEY EQUATIONS

It is useful at this stage to group the following six equations:

$$(G) \quad \mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t}, \quad (23-46)$$

$$(G) \quad \oint_C \mathbf{E} \cdot d\mathbf{l} = - \int_A \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{A}, \quad (23-29)$$

$$(G) \quad \nabla \times \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t}, \quad (23-30)$$

$$(G) \quad \mathbf{B} = \nabla \times \mathbf{A}, \quad (\text{Sec. 18.4}) \quad \text{and} \quad (23-46)$$

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 \int_A \mathbf{J} \cdot d\mathbf{A}, \quad (\text{Sec. 19.5})$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}. \quad (\text{Sec. 19.4})$$

The four equations preceded by (G) are general, while the other two apply only to slowing varying fields (Sec. 27.1). In each equation all the terms concern the *same* reference frame.

23.9 SUMMARY

A body moves in some arbitrary fashion in a constant, but not necessarily uniform, magnetic field. At a point fixed to the moving body, a charge Q experiences a force $Q\mathbf{v} \times \mathbf{B}$. The *motional electromotance* is

$$\mathcal{V} = \int_C (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l}. \quad (23-2)$$

Here \mathbf{v} is the velocity of a charge fixed in the body, and \mathbf{B} is the magnetic flux density at that point in space.

When the curve C is closed, the motional electromotance is also given by

$$\mathcal{V} = -\frac{d\Phi}{dt}, \quad (23-3)$$

where Φ is the enclosed flux. This is the *Faraday induction law*. The right-hand screw rules applies. This law also applies to a fixed circuit situated in a time-varying magnetic field.

At any point in space, in a given reference frame,

$$\mathbf{v} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}.$$

(23-30)

This is one of Maxwell's equations.

Lenz's law states that the electromotance induced in a closed circuit tends to oppose *changes* in the magnetic flux linking the circuit.

With a multturn closed circuit,

$$\mathcal{V} = -\frac{d\Lambda}{dt}, \quad (23-31)$$

where Λ is the *flux linkage*.

In any given inertial reference frame,

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t} \quad \text{and} \quad \mathbf{B} = \nabla \times \mathbf{A}. \quad (23-46)$$

A charge Q moving at a velocity \mathbf{v} in superposed electric and magnetic fields is subjected to a force

$$\mathbf{F} = Q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = Q\left(-\nabla V - \frac{\partial \mathbf{A}}{\partial t} + \mathbf{v} \times \mathbf{B}\right). \quad (23-59)$$

PROBLEMS

23-1. (23.2) The thought experiment of Fig. 23-3

Show that there is conservation of energy in the thought experiment of Fig. 23-3(a).

23-2. (23.2) Tides and the magnetic field of the earth

Discuss how tides affect the magnetic field of the earth by considering the case of a river flowing into the sea in the east-to-west direction in the northern hemisphere. Remember that the magnetic pole situated at the *north geographic pole* is a *south magnetic pole*. The vector \mathbf{B} points downward in the *northern hemisphere*.

23-3. (23.2) The magnetic braking force on a satellite

A natural satellite whose diameter is 10^4 meters moves at a velocity of 1 kilometer/second in the direction normal to the magnetic field of a planet in a region where $B = 10^{-7}$ tesla. The satellite has an appreciable conductivity.

(a) The satellite moves in a perfect vacuum. What happens?

(b) The ambient gas has a density of the order of 10^{10} particles per cubic meter, the particles being either electrons or singly charged ions. Each half of the satellite collects particles of the correct sign in sweeping through space. Calculate the order of magnitude of the current.

(c) Calculate the order of magnitude of the braking force.

(d) Someone suggests that this current could provide power for an artificial satellite traveling in the same field at the same velocity. Inversely, a current in the opposite direction could serve to propel the satellite. What is your opinion?

Artificial satellite velocities range from about 4 to 8 kilometers/second, and $\mathbf{v} \times \mathbf{B}$ in the ionosphere and magnetosphere ranges from about 100 microvolts/meter to 320 millivolts/meter.

23-4. (23.3) Eddy-current damping

Figure 23-9 shows one common type of eddy-current damper. Motion of the copper plate in the field of the permanent magnet induces currents that tend to oppose the motion, according to Lenz's law. Joule losses in the plate dissipate its kinetic energy.

Dampers of this general type are used mostly, but not exclusively, in low-power devices such as watt-hour meters and balances. As you will see, the braking force is proportional to the velocity, as in a viscous fluid.

(a) Explain qualitatively, but in greater detail, the origin of the braking force.

(b) How could you design an automobile speedometer that uses eddy currents?

(c) Say B is uniform over the pole face. The path followed by the current is complex; set $R \approx 3a/\sigma s$. This quantity is of the order 3. The plate has a thickness s and a conductivity σ . Calculate the current.

(d) Calculate the braking force F . This is proportional to the conductivity. So the plate should be either copper or aluminum. An even better solution is to use an iron plate faced with copper.

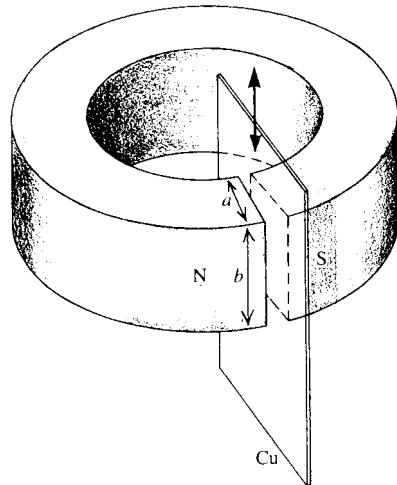


Fig. 23-9.

(e) Calculate the power IV dissipated in the plate. This should be equal to Fv .

(f) Estimate the value of B required in eddy-current disk brakes for a small bus. The conductivity of copper is 5.8×10^7 siemens/meter. Estimate the power dissipated at each wheel. Why does a vehicle equipped with eddy-current brakes still need conventional brakes?

In mountainous regions some buses and trucks are equipped with dynamos that brake by generating electric power that is dissipated in a large resistance on the roof.

23-5. (23.4) Detecting flaws in metal tubing

Figure 23-10 shows the principle of operation of a device for detecting flaws in metal tubing, or rod. The coils a provide a large gradient of magnetic field along the axis, as in Prob. 18-10. Coil b is connected to a monitor. The tubing T moves at a constant velocity v along the axis of symmetry. A voltage appears across coil b when a flaw passes through. Explain.

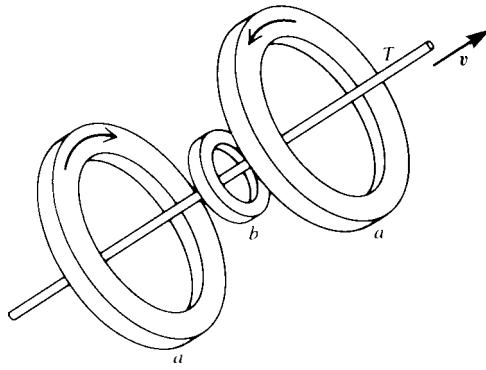


Fig. 23-10.

23-6. (23.4) The flux-gate magnetometer

A *magnetometer* measures B . One common type is the *flux-gate magnetometer*, which puts to use the hysteresis curve. There exist many forms, one of which is shown in Fig. 23-11(a). The two rods are made of a ferromagnetic material such as a ferrite, whose hysteresis curve is shown in Fig. 23-11(b). The twin coils are in series and are wound as in the figure so as to magnetize the rods in opposite directions. The current through these coils is sufficient to carry the material through a complete hysteresis loop.

In the absence of an external field H_{ex} the magnetic fluxes through the rods cancel, and $V = 0$.

(a) Sketch $\Phi(t)$ and $V(t)$ for each rod when $H_{ex} = 0$.

(b) Sketch the same quantities for $H_{ex} \neq 0$. You will notice that if the oscillator operates at a frequency f , the fundamental frequency of V is $2f$.

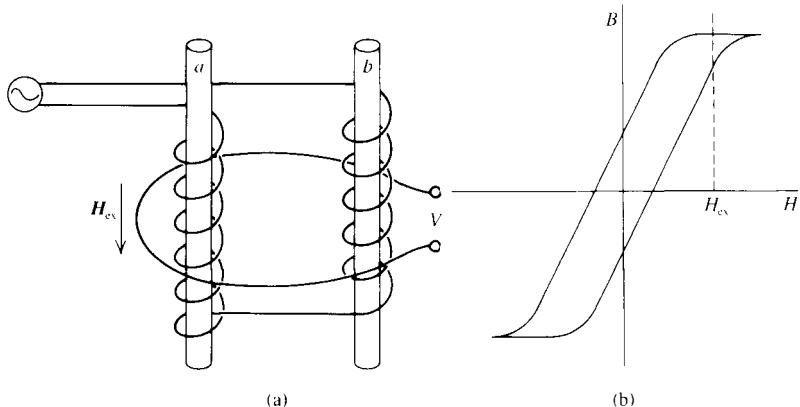


Fig. 23-11.

This facilitates the measurement because the detector can be made to reject the frequency f .

Flux-gate magnetometers can measure fields down to a few nanoteslas.

23-7. (23.4) The peaking strip

A *peaking strip* serves to measure B . It consists of a fine wire of permalloy (see below) oriented in the direction of B with a small pickup coil of a few thousand turns near the center, on the axis of a solenoid, as in Fig. 23-12.

To measure the ambient B , the solenoid carries a direct current that just cancels B , plus a small alternating current. Then the H on the axis of the solenoid is that of the alternating current, and the strip goes through a hysteresis loop at every cycle.

With molybdenum permalloy the loop is approximately rectangular, and the voltage induced in the small coil has two sharp peaks, one positive and one negative, which can be observed on an oscilloscope.

When the oscilloscope sweep is synchronized with the alternating current in the solenoid, the two peaks are symmetric if the time-averaged H on the axis of the solenoid is zero. Then the steady field of the solenoid exactly

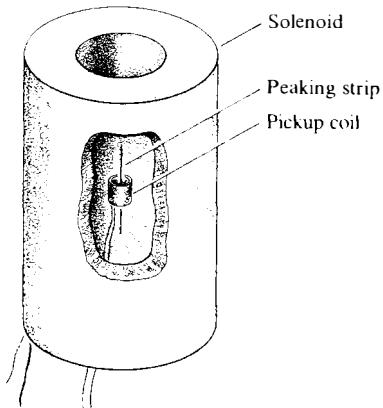


Fig. 23-12.

cancels the ambient B and the current in the solenoid is then a measure of B .

The peaking strip has a rather limited range of applications. (1) The solenoid has to be at least about 10 centimeters long because it must be at least a few times longer than the strip, to avoid excessive end effects. But the length of the strip must be much larger than its diameter, again to reduce end effects, and a decrease in the strip cross section decreases the signal proportionately. (2) The ambient B cannot be larger than a few hundredths of a tesla, for otherwise the power dissipated in the solenoid becomes excessive. (3) If one measures B in the neighborhood of a pole-piece, the field of the solenoid alters the permeability of the iron locally.

Calculate the peak voltage induced in the pickup coil under the following conditions: strip diameter, 25 micrometers; number of turns in the pickup coil, 1000; maximum value of μ_r , 75,000; frequency, 60 hertz; amplitude of the alternating H , 7 ampere-turns/meter.

13.8. (23.4) Measuring a resistivity without contacts

It is useful to be able to measure the resistivity of a sample without having to cement contacts to it. One method involves placing a disk of the material inside a solenoid carrying an alternating current, with the two axes parallel, and measuring the power absorbed by the disk. The disk has a radius a , a thickness s , and a conductivity σ . The magnetic field is uniform, and $B = B_m \cos \omega t$. We neglect the magnetic field of the induced currents. We therefore restrict ourselves to low-conductivity materials.

Find the relation between σ and the average dissipated power P .

13.9. (23.5) The induction linear accelerator

Figure 23-13 shows a schematic diagram of an induction linear accelerator. It consists of a series of ferrite toroids linked by the ion beam and by one-turn loops that carry large pulsed currents.

One such accelerator comprises 200 toroids and accelerates a 10-kiloampere pulsed electron beam to 50 million electronvolts. Its total length is 80 meters, and the pulses are 70 nanometers wide.

Explain its operation qualitatively.

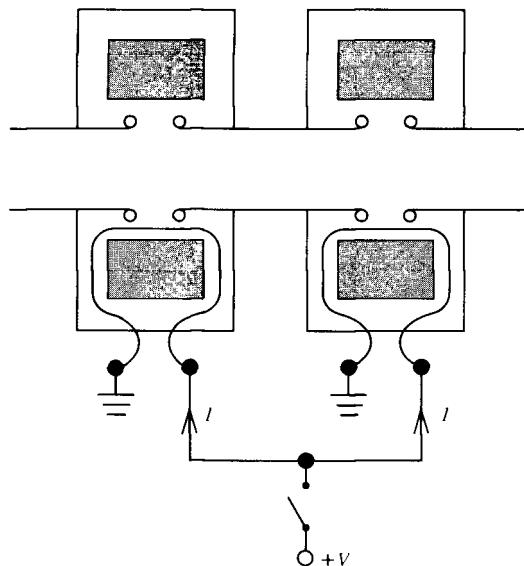


Fig. 23-13.

23-10. (23.5) A magnetometer that uses eddy currents

Figure 23-14 shows the principle of operation of a magnetometer that can measure magnetic fields as small as 10^{-8} tesla and up to 10^{-2} tesla. The aluminum plate P turns on the axis AA' in the ambient field B_0 that we wish to measure. The fluctuating eddy currents induced in P produce a fluctuating magnetic flux through the fixed coil C , which has N turns, and the voltage V is a measure of B_0 .

The plate is 10 millimeters square and is cemented inside the plastic rotor of a small air turbine that operates at 1000 revolutions/second. The only metallic parts are the plate and the coil.

An exact calculation of V as a function of geometry, of ω , and of B_0 would be difficult. But this is unnecessary because we can calibrate the instrument with Helmholtz coils (Prob. 18-9).

(a) How does V vary with B_0 and with ω ? Set $\omega t = 0$ when the plate lies in the plane of C .

(b) What is the frequency of V ?

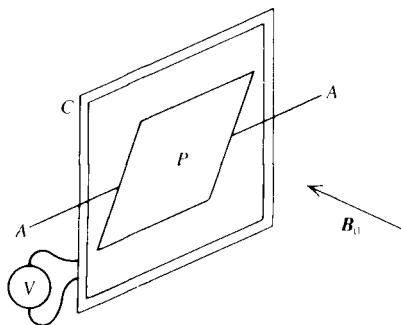


Fig. 23-14.

CHAPTER 24

*MAGNETIC FIELDS VII *Electric Circuits C. Mutual and Self-Inductance*

24.1	MUTUAL INDUCTANCE M	437
24.1.1	THE NEUMANN EQUATION	437
	<i>Example: THE MUTUAL INDUCTANCE BETWEEN TWO COAXIAL SOLENOIDS</i>	439
24.1.2	THE SIGN OF M	440
24.2	SELF-INDUCTANCE L	441
	<i>Example: THE SELF-INDUCTANCE OF A LONG SOLENOID</i>	442
	<i>Example: THE SELF-INDUCTANCE OF A TOROIDAL COIL</i>	442
24.2.1	THE IMPEDANCE OF AN INDUCTOR	443
24.2.2	THE POWER FACTOR λ	445
	<i>Example: MINIMIZING LINE LOSSES</i>	445
24.3	THE COUPLING COEFFICIENT k	446
	<i>Example: COAXIAL SOLENOIDS</i>	447
24.4	SUMMARY	448
	PROBLEMS	448

This chapter concerns the electromotance induced in a circuit when its magnetic flux linkage changes. The change in flux can occur either because of a change in the current flowing through the circuit itself or in currents flowing elsewhere or because of a change in the geometry of the circuits.

24.1 MUTUAL INDUCTANCE M

24.1.1 The Neumann Equation

In Fig. 24-1 the active circuit a carries a current I_a . The magnitude flux Φ_{ab} that originates in a and links b is

$$\Phi_{ab} = \oint_b \mathbf{A}_a \cdot d\mathbf{l}_b = \oint_b \left(\frac{\mu_0 I_a}{4\pi} \oint_a \frac{d\mathbf{l}_a}{r} \right) \cdot d\mathbf{l}_b, \quad (24-1)$$

where r is the distance between the elements $d\mathbf{l}_a$ and $d\mathbf{l}_b$. Thus

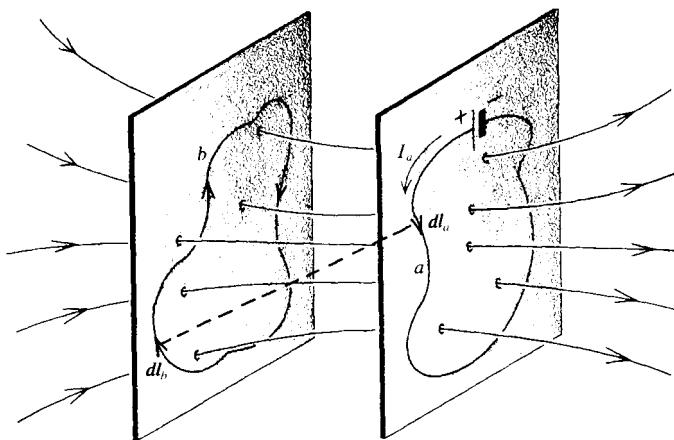


Fig. 24-1. The active circuit *a* bears a current I_a that increases. Part of its magnetic flux links the passive circuit *b*. The electromotance \mathcal{V}_b induced in *b* is in the direction shown.

$$\Phi_{ab} = \frac{\mu_0}{4\pi} \oint_a \oint_b \frac{dI_a \cdot dI_b}{r} = M_{ab} I_a, \quad (24-2)$$

where

$$M_{ab} = \frac{\mu_0}{4\pi} \oint_a \oint_b \frac{dI_a \cdot dI_b}{r} \quad (24-3)$$

is the *mutual inductance* between the two circuits. Mutual inductance is expressed in webers per ampere, or in *henrys*. This is the *Neumann equation*.

If the geometry is more complex than that of Fig. 24-1, the above reasoning still applies, except that the flux Φ_{ab} becomes the flux linkage Λ_{ab} (Sec. 23.4.2), and

$$\Lambda_{ab} = M_{ab} I_a. \quad (24-4)$$

The electromotance induced in *b* by a change in I_a is

$$\mathcal{V}_b = -\frac{d\Lambda_{ab}}{dt} = -M_{ab} \frac{dI_a}{dt} - I_a \frac{dM_{ab}}{dt}. \quad (24-5)$$

The mutual inductance is usually constant. Then the last term on the right vanishes.

Therefore the mutual inductance between two circuits is 1 henry if a current changing at the rate of 1 ampere/second in one circuit induces an electromotance of 1 volt in the other.

If current I_a is sinusoidal, then

$$\mathcal{V}_b = -j\omega M_{ab} I_a. \quad (24-6)$$

A device comprising two circuits designed to possess mutual inductance is termed a *mutual inductor*, or a *transformer*.

The Neumann equation is seldom useful because the double integral is difficult to evaluate, even for simple geometries. This is not a matter for concern, because mutual inductances are easily measured with impedance bridges. One can also calculate M_{ab} from the ratio \mathcal{V}_b to dI_a/dt .

The Neumann equation is nonetheless interesting. It shows that mutual inductance depends solely on the geometry of the system. We had a similar situation with respect to capacitance.

Also, we can interchange the subscripts in the Neumann equation without altering the mutual inductance. Therefore

$$M_{ab} = M_{ba} = M. \quad (24-7)$$

This is surprising because the circuits can have different shapes and different numbers of turns. This is a striking example of the *reciprocity theorem* of Sec. 8.7. That is,

$$\text{if } \mathcal{V}_b = -M \frac{dI_a}{dt}, \quad \text{then } \mathcal{V}_a = -M \frac{dI_b}{dt}. \quad (24-8)$$

Example

THE MUTUAL INDUCTANCE BETWEEN TWO COAXIAL SOLENOIDS

In the third example in Sec. 18.2.1 we found that, inside a long solenoid with N' turns per meter and bearing a current I (ignoring end effects),

$$B = \mu_0 N' I \quad (24-9)$$

We add a second winding over the solenoid, as in Fig. 24-2, and we assume that both windings are long compared to their common diameter, in order to render end effects negligible.

First, we assume that solenoid a , of radius R and number of turns N_a , bears a current I_a . The magnetic flux of a that links solenoid b , of the same radius and with N_b turns, is then, from the third example in Sec. 18.2.1,

$$\Phi_{ab} = \pi R^2 \mu_0 \frac{N_a}{a} I_a, \quad (24-10)$$

and

$$M_{ab} = \frac{N_b \Phi_{ab}}{I_a} = \frac{\mu_0 \pi R^2 N_a N_b}{a}. \quad (24-11)$$

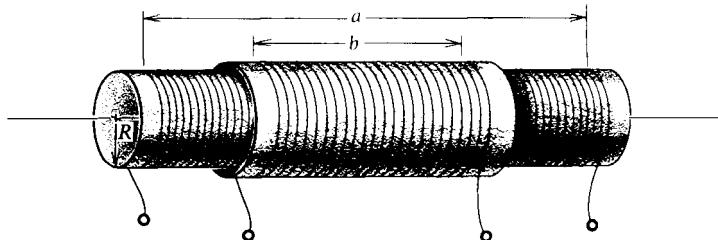


Fig. 24-2. Two coaxial solenoids. We have shown different radii for clarity, but coil b is wound directly over a .

Alternatively, we assume a current I_b in solenoid b . Then

$$\Phi_{ba} = \pi R^2 \mu_0 \frac{N_b}{b} I_b. \quad (24-12)$$

This flux links only $(b/a)N_a$ turns of coil a , since B falls rapidly to zero beyond the end of a long solenoid. Thus

$$M_{ba} = \frac{b}{a} \frac{N_a \Phi_{ba}}{I_b} = \frac{\mu_0 \pi R^2 N_a N_b}{a}, \quad (24-13)$$

and $M_{ab} = M_{ba}$, as expected.

It is paradoxical that a varying current in the inner solenoid should induce an electromotance in the outer one, since $B \approx 0$ outside a long solenoid. The reason for this is that the induced \mathbf{E} is equal to minus the time derivative of \mathbf{A} , not \mathbf{B} , and \mathbf{A} does not vanish outside along solenoid, even though $\nabla \times \mathbf{A}$ is zero.

24.1.2 The Sign of M

As a rule, the sign of M is immaterial. There are occasions, however, when phases matter, and then one requires the sign of M . Also, the directions of forces and torques between circuits depend on the sign of M , as we shall see in Chap. 26.

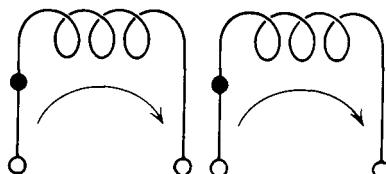


Fig. 24-3. Convention for showing the sign of M on a circuit diagram. If the positive directions chosen for the currents in the two windings are such that current enters through the wire carrying a dot, then M is positive, as it is here.

The sign of M is defined as follows. First we select arbitrary positive directions for the currents in the two circuits. Then M is positive if a positive current in a gives in b a flux linkage of the same sign as that resulting from a positive current in b . See Fig. 24-3.

24.2 SELF-INDUCTANCE L

A simple circuit carrying a current I is, of course, linked by its own magnetic flux, as in Fig. 24-4. The ratio

$$L = \frac{\Lambda}{I} \quad (24-14)$$

is termed the *self-inductance* of the circuit. As for mutual inductance, self-inductance depends solely on the geometry of the circuit and is measured in henrys. Self-inductance is always positive.

If the current I in a self-inductance L changes, then a voltage

$$\mathcal{V} = -\frac{d\Lambda}{dt} = -L \frac{dI}{dt} \quad (24-15)$$

appears between its terminals. The negative sign means that this voltage opposes the change in current. See Fig. 24-4.

Therefore the self-inductance of a zero-resistance circuit is 1 henry if the current increases at the rate of 1 ampere/second when the difference of potential applied between the terminals is 1 volt.

If both I and L are time-dependent, then

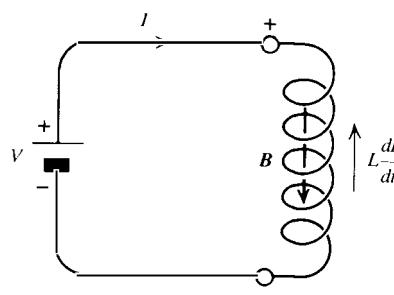


Fig. 24-4. Circuit linked by its own flux. If I increases, the voltage across L opposes I .

$$\mathcal{V} = -\frac{d\Lambda}{dt} = -\frac{d}{dt}(LI) = -L \frac{dI}{dt} - I \frac{dL}{dt}. \quad (24-16)$$

As a rule, L is constant.

One can calculate a self-inductance, at least in principle, from the Neumann equation (Sec. 24.1.1), with both line integrals running over the same circuit. If the conductor cross section is infinitely small, then the linking flux and the self-inductance turn out to be infinite. This arises from the fact that B tends to infinity in the immediate neighborhood of the wire. The region where B tends to infinity is itself infinitely small, but the flux tends to infinity logarithmically. With currents distributed over an infinitely thin *surface*, as in the next example, B remains finite and L is also finite.

Self-inductance is easily measured. A circuit designed to possess self-inductance is called an *inductor*.

Example

THE SELF-INDUCTANCE OF A LONG SOLENOID

If l is the length of a long solenoid, N the number of turns, and R its radius, then

$$L = \frac{\Lambda}{I} = \frac{N\Phi}{I} = \frac{N\pi R^2 \mu_0 (N/l)I}{I} = \frac{\mu_0 N^2 \pi R^2}{l}. \quad (24-17)$$

The self-inductance of a short solenoid is smaller by a factor that is a function of the ratio R/l .

A high-permeability nonconducting magnetic core increases B , and hence the self inductance, by a factor of about 1.3 (length/diameter)^{1/7}.

Example

THE SELF-INDUCTANCE OF A TOROIDAL COIL

The magnetic flux density inside the air-core toroidal coil of Fig. 24-5 follows from the circuital law (Sec. 20.6):

$$B = \frac{\mu_0 NI}{2\pi}. \quad (24-18)$$

Thus

$$\Phi = \frac{\mu_0 NI}{2\pi} \int_{a-b/2}^{a+b/2} \frac{b}{r} dr = \frac{\mu_0 NI b}{2\pi} \ln \frac{2a+b}{2a-b}, \quad (24-19)$$

$$L = \frac{N\Phi}{I} = \frac{\mu_0 N^2 b}{2\pi} \ln \frac{2a+b}{2a-b}. \quad (24-20)$$

If the relative permeability of the core is μ_r , then the self-inductance is μ_r times larger.

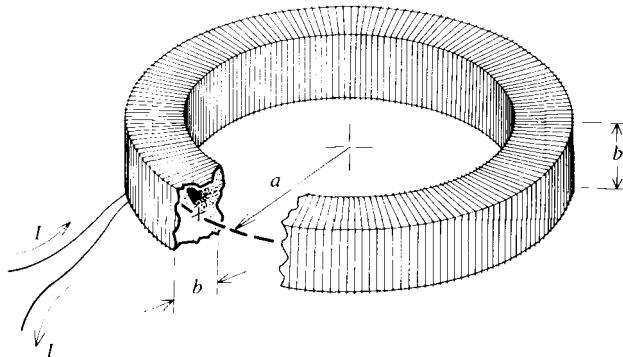


Fig. 24-5. Toroidal coil of square cross section.

24.2.1 The Impedance of an Inductor

The voltage V across an ideal zero-resistance inductor determines the rate of change of the current flowing through it. Disregarding the negative sign of Eq. 24-15,

$$V = L \frac{dI}{dt}. \quad (24-21)$$

If I is a sinusoidal function of time, then

$$V = j\omega LI \quad (24-22)$$

and the impedance of an inductor is

$$Z = j\omega L. \quad (24-23)$$

Real inductors are not that simple. Unless an inductor is superconducting, one must usually take its resistance into account and write

$$Z = R + j\omega L. \quad (24-24)$$

Increasing the frequency increases the ratio $\omega L/R$, and this lessens the relative importance of R . However, this brings in another phenomenon. Because of the voltage difference between turns, the inductor also acts as a capacitor, and its equivalent circuit is that of Fig. 24-6, where C is its *stray capacitance*. Also, R increases with frequency because the current flows closer and closer to the surface of the wire as the frequency increases. This is the *skin effect*.

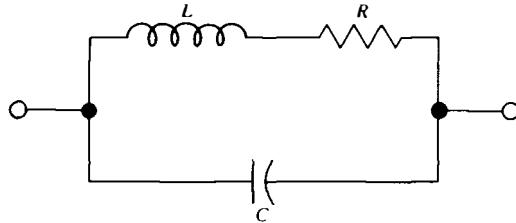


Fig. 24-6. The equivalent circuit of a real inductor. The parameters R , L , and C depend on the geometry and on the nature of the materials.

The impedance of a real inductor is thus

$$Z = \frac{(R + j\omega L)/(j\omega C)}{R + j\omega L + 1/(j\omega C)} = \frac{R + j\omega L}{(R + j\omega L)j\omega C + 1} \quad (24-25)$$

$$= \frac{R + j\omega L}{1 - \omega^2 LC + Rj\omega C} = \frac{(R + j\omega L)(1 - \omega^2 LC - Rj\omega C)}{(1 - \omega^2 LC)^2 + (R\omega C)^2} \quad (24-26)$$

$$= \frac{R + j\omega [L - C(R^2 + \omega^2 L^2)]}{(1 - \omega^2 LC)^2 + (R\omega C)^2} R' + jX. \quad (24-27)$$

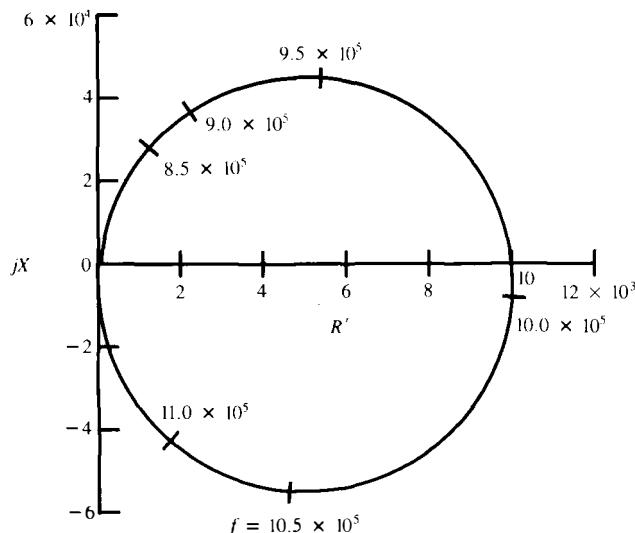


Fig. 24-7. Locus of the impedance $Z = R' + jX$ of a certain inductor as a function of frequency. The parameters R , L , and C of Fig. 24-6 are, respectively, 1000 ohms, 1.00 millihenry, and 100 picofarads. We have disregarded the fact that R increases with frequency.

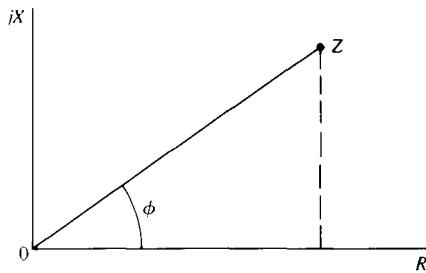


Fig. 24.8. Impedance Z in the complex plane. The power factor is equal to $\cos \phi$.

For $\omega = 0$, $Z = R$, as one must expect. For large values of ω , only its higher powers remain and $Z = 1/(j\omega C)$: all the current then flows through the stray capacitance C , and the inductor is capacitive. As ω tends to infinity, Z tends to zero. See Fig. 24-7.

Inductors are so designed that $j\omega L$, or at worst $R + j\omega L$, is a good approximation for Z over a given frequency range.

24.2.2 The Power Factor λ

The ratio $R/|Z|$ is the *power factor* of an impedance. This quantity is usually denoted by λ and expressed as a percentage. With Z plotted in the complex plane as in Fig. 24-8, $\lambda = \cos \phi$. If ϕ is 45° , then λ is 70.7%.

Example

MINIMIZING LINE LOSSES

Figure 24-9 shows a load

$$Z = R_{\text{load}} + jX = |Z| \exp j\phi, \quad (24-28)$$

$$\phi = \arctan \frac{X}{R_{\text{load}}}, \quad (24-29)$$

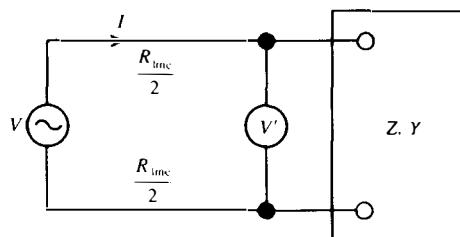


Fig. 24.9. A load impedance $Z = R_{\text{load}} + jX$, or an admittance $Y = G_{\text{load}} + jB$, connected to a source through a line of resistance R_{line} . The line resistance is equal to twice the resistance of one wire.

$$Y = \frac{1}{Z} = \frac{1}{|Z|} \exp(-j\phi) = G + jB \quad (24-30)$$

connected to a source

$$V = V_m \exp j\omega t \quad (24-31)$$

through a line of resistance $R_{\text{line}} \ll |Z|$. We wish to minimize the losses in the line, for a given line, a given V_m , and a given power dissipation in the load.

Defining V' as in the figure, the current is

$$I = V' Y \approx VY = V(G + jB). \quad (24-32)$$

Then the power in the load is

$$P_{\text{load}} = \frac{1}{2} \operatorname{Re} VI^* = \frac{1}{2} \operatorname{Re} [VV^*(G - jB)] = \frac{1}{2} V_m^2 G = V_{\text{rms}}^2 G, \quad (24-33)$$

while the power lost in the line is

$$W_{\text{line}} = \frac{1}{2} \operatorname{Re} (V - V')I^* = \frac{1}{2} \operatorname{Re} [(IR_{\text{line}})I^*] \quad (24-34)$$

$$= \frac{1}{2} R_{\text{line}} V_m^2 (G^2 + B^2) = R_{\text{line}} V_{\text{rms}}^2 (G^2 + B^2) \quad (24-35)$$

$$= R_{\text{line}} (GW_{\text{load}} + V_{\text{rms}}^2 B^2). \quad (24-36)$$

The power lost in the line is minimum when $B = 0$. Then ϕ is zero and the power factor $\cos \phi$ is 100%.

Stated otherwise, the component of the current that is in quadrature (phase lead or lag of 90°) with the applied voltage yields zero useful power in the load, but nonetheless gives rise to a power loss and to a voltage drop in the line.

Electric motors are inductive loads. Therefore, in large installations, one connects capacitor banks across the line, close to the motors, to make B equal to zero.

24.3 THE COUPLING COEFFICIENT k

Consider a *loop* of wire a bearing a current I_a and linked by its own flux Φ_{aa} . Another loop of wire b , nearby, intercepts a fraction of this flux

$$\Phi_{ab} = k_{ab} \Phi_{aa}, \quad (24-37)$$

where $|k_{ab}| \leq 1$. The self-inductance of loop a is

$$L_a = \frac{\Phi_{aa}}{I_a}, \quad (24-38)$$

and the mutual inductance between the two loops is

$$M_{ab} = \frac{\Phi_{ab}}{I_a} = \frac{k_{ab}\Phi_{aa}}{I_a} = k_{ab}L_a. \quad (24-39)$$

Likewise,

$$M_{ba} = k_{ba}L_b. \quad (24-40)$$

Now $M_{ab} = M_{ba} = M$, from Sec. 24.1.1. Thus

$$M^2 = k_{ab}k_{ba}L_aL_b, \quad (24-41)$$

$$M = \pm |k_{ab}k_{ba}|^{1/2} (L_aL_b)^{1/2} = k(L_aL_b)^{1/2}, \quad (24-42)$$

$$|k| = |k_{ab}k_{ba}|^{1/2}, \quad (24-43)$$

where k is the *coupling coefficient* for the two loops. Superposing the two loops makes k equal to unity. The sign of k is that of M .

For *coils* of arbitrary shapes and sizes, one replaces the fluxes by the flux linkages

$$k_{ab} = \frac{\Lambda_{ab}}{\Lambda_{aa}}, \quad k_{ba} = \frac{\Lambda_{ba}}{\Lambda_{bb}}, \quad (24-44)$$

and Eq. 24-42 applies. However, either k_{ab} or k_{ba} can now be larger than unity.

The maximum possible magnitude of k is unity. We can prove this statement easily for two solenoids of the same length and nearly the same radius, one inside the other. Then k_{ab} and k_{ba} are both maximum. Say solenoid a has N_a turns, and solenoid b has N_b turns. Then

$$k_{ab} = \frac{N_b}{N_a}, \quad k_{ba} = \frac{N_a}{N_b}. \quad (24-45)$$

Either one of these ratios can be larger than unity, but their product is equal to unity and $k = 1$.

In Sec. 24.2 we saw that the self-inductance of a circuit tends to infinity as the wire diameter tends to zero. However, mutual inductance remains well behaved even if both circuits are filamentary. The reason is that the flux close to a thin wire does not link another circuit some distance away. As the diameter of the wire tends to zero, L tends to infinity, k tends to zero, and M remains finite.

Example COAXIAL SOLENOIDS

The coaxial solenoids of the example in Sec. 24.1.1 are of different lengths. Then, from Eq. 24-42,

$$k = \frac{M}{(L_a L_b)^{1/2}} = \left(\frac{b}{a}\right)^{1/2}, \quad (24-46)$$

with M as in Eq. 24-110 and the L 's as in Eq. 24-17.

24.4 SUMMARY

The *mutual inductance* M between two circuits a and b is given by the *Neumann equation*

$$M_{ab} = \frac{\mu_0}{4\pi} \oint_a \oint_b \frac{\mathbf{dl}_a \cdot \mathbf{dl}_b}{r}, \quad (24-3)$$

where the line integrals run around each circuit and r is the distance between the elements \mathbf{dl}_a and \mathbf{dl}_b . Because of the symmetry of the integral,

$$M_{ab} = M_{ba} = M. \quad (24-7)$$

If Λ_{ab} is the flux that originates in a and links b , then

$$\Lambda_{ab} = M_{ab} I_a. \quad (24-4)$$

The electromotance induced in b by the current in a is

$$\mathcal{V}_b = -\frac{d\Lambda_{ab}}{dt} = -M \frac{dI_a}{dt} - I_a \frac{dM}{dt}. \quad (24-5)$$

The *self-inductance* of a circuit is

$$L = \frac{\Lambda}{I}, \quad (24-13)$$

where Λ is the flux linkage when the current is I . The impedance of an ideal inductor is $j\omega L$.

The *coupling coefficient* between two circuits is

$$k = \frac{M}{(L_a L_b)^{1/2}}. \quad (24-46)$$

This coefficient takes the sign of M , and its magnitude is at most unity.

The *power factor* of an impedance is $R/|Z|$.

PROBLEMS

- 24-1. (24.1) The mutual inductance between a solenoid and a short coaxial coil
A long solenoid of radius R and N' turns per meter carries a short coil of
 N turns near its center.

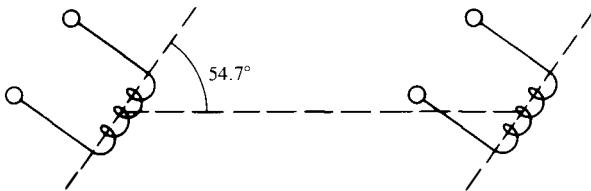


Fig. 24-10.

- (a) Calculate the mutual inductance M .
 (b) Does the radius of the short coil affect M ?

Note how much more difficult it would be to calculate the flux linking the solenoid for a given current in the short coil.

24-2. (24.1) The mutual inductance between a toroid and an axial wire

A long straight wire lies along the axis of a toroid of N turns, major radius a , and square cross section of side b , with $a \gg b$.

Calculate the mutual inductance (a) assuming a current I in the wire and
 (b) assuming a current I in the toroid.

24-3. (24.1) The mutual inductance between a straight wire and a loop

A loop of wire of radius R is centered at a distance $2R$ from a long straight wire. The wire is in the plane of the loop.

Calculate the mutual inductance.

24-4. (24.1) A zero-mutual-inductance magnetic dipole pair

A certain device for geophysical exploration comprises two short coils in the position shown in Fig. 24-10. The manufacturer states that the mutual inductance is zero. Is that true?

24-5. (24.1) Current transformer

Figure 24-11 shows a *side-look current transformer* for measuring large current pulses. Show that for a single-turn coil

$$V = \frac{\mu_0 a}{\pi} \ln \left(\frac{b+a}{b-a} \right) \frac{dI}{dt}.$$

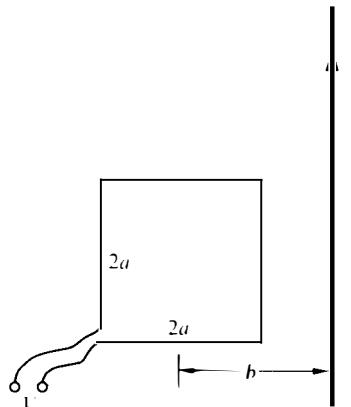


Fig. 24-11.

One can obtain $I(t)$ with an integrating circuit (Prob. 7-9).

24-6. (24.2) A conducting shield for fluctuating magnetic fields

It is often necessary to shield instruments from stray magnetic fields. If the only disturbing field is that of the earth, then one can set up a pair of Helmholtz coils (Prob. 18-9) to oppose the earth's field. If the field is static but not uniform, then one must use a shield made of high-permeability material. Multiple shields, one inside the other, are better than a single thick shield.

Could a conducting enclosure be a good shield against fluctuating magnetic fields? The answer is yes, as we shall see, but only at quite high frequencies.

Imagine a simple situation where the external magnetic field \mathbf{B}_{ex} is uniform, with $B_{ex} = B_{ex,m} \exp j\omega t$. The shield is a long tube, parallel to the lines of \mathbf{B} , a few times longer than its diameter $2a$, and a few times longer than the shielded region.

We assume that the current induced in the shield is uniformly distributed throughout its thickness b . In other words, we disregard the skin effect (Sec. 29.1). If this assumption is not valid, then the shielding is better than our calculation would indicate.

(a) Calculate the resistance R' of the tube, per unit length, in the azimuthal direction. Calculate L .

(b) Let $B_{in} = B_{in,m} \exp j\omega t$ be the value of B inside the tube, away from the ends. Find the ratio $B_{in,m}/B_{ex,m}$.

(c) Show that, if the skin effect is negligible, this ratio cannot be smaller than 0.5. A conducting enclosure therefore acts as a shield only through the skin effect.

24-7. (24.2) Inductance and reluctance

An N -turn coil links a magnetic circuit. Show that $L = N^2 \mathcal{R}$.

24-8. (24.2) The Maxwell bridge

See Prob. 7-12. Figure 24-12 shows a *Maxwell bridge*. This circuit serves to measure the inductance L and the resistance R of an inductor. One adjusts the values of R_b , R_c , R_d , and C until V equals zero.

Find L and R in terms of the other components.

24-9. (24.2) Electromagnet operating on alternating current

An electromagnet with a variable gap length operates on alternating current. How does the rms value of the magnetic flux depend on the gap length, for a given applied voltage, and neglecting leakage flux?

24-10. (24.2) Power-factor correction

A load is inductive, has a power factor of 65%, and draws a current of 100 amperes at 600 volts.

(a) Calculate the magnitude of Z , its phase angle, and its real and imaginary parts.

(b) Calculate the in-phase and quadrature components of the current.

(c) What size capacitor should be placed in parallel with the load to cancel the reactive current at 60 hertz?

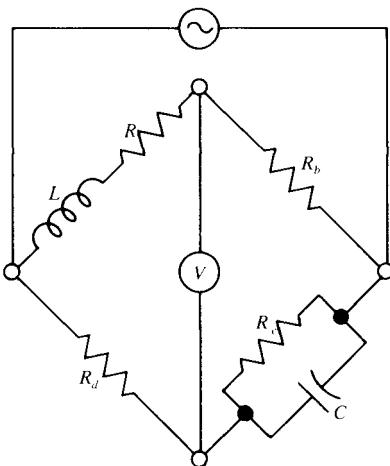


Fig. 24-12.

- (d) What is the current supplied by the source now?
Note that Z' is not equal to the real part of Z .

24-11. (24.2.1) Power-factor correction with fluorescent lamps

A *fluorescent lamp* consists of an evacuated glass tube containing mercury vapor and coated on the inside with a fluorescent mixture. A discharge occurs between electrodes situated at each end. The discharge emits most of its energy at 253.7 nanometers, in the ultraviolet. The fluorescent coating absorbs this radiation and reemits visible light.

The discharge operates correctly only when it is connected in series with an impedance. A resistor would dissipate energy, so it is the custom to use a series inductor. Several types of circuit are in use.

One particular fluorescent fixture operates at 120 volts and dissipates 80 watts. Its power factor is 50%.

- Find the reactive current.
- What is the size of the capacitor connected in parallel with the discharge tube and its inductor that will make the power factor equal to 100%?

CHAPTER 25

*MAGNETIC FIELDS VIII *Electric Circuits D: Inductive Circuits and Transformers*

25.1	CIRCUITS COMPRISING SELF-INDUCTANCES	453
	<i>Example: SERIES RESONANCE</i>	453
	<i>Example: PARALLEL RESONANCE</i>	455
25.2	CIRCUITS COMPRISING SELF-INDUCTANCE AND MUTUAL INDUCTANCE	457
	<i>Example: SIMPLE CIRCUIT COMPRISING A TRANSFORMER</i>	457
25.3	TRANSFORMATION OF A MUTUAL INDUCTANCE	458
	<i>Example: SOLVING THE TRANSFORMER CIRCUIT IN ANOTHER WAY</i>	459
25.4	MAGNETIC-CORE TRANSFORMERS	460
25.4.1	THE IDEAL TRANSFORMER	461
25.4.2	THE RATIO V_1/V_2	461
25.4.3	THE RATIO L_1/L_2	462
25.4.4	THE INPUT IMPEDANCE Z_{in}	463
25.4.5	THE RATIO I_1/I_2	463
	<i>Example: THE AUTOTRANSFORMER</i>	463
25.4.6	POWER TRANSFER FROM SOURCE TO LOAD THROUGH A TRANSFORMER	464
25.5	SUMMARY	465
	PROBLEMS	466

This is the last chapter on electric circuits. We still have to discuss how to apply Kirchhoff's laws to circuits comprising self- and mutual inductances. Self-inductance is easy. Mutual inductance is less simple because it consists of an interaction between branches. However, with either one of two simple transformations, the application of Kirchhoff's laws becomes straightforward.

A careful discussion of magnetic-core transformers is well beyond the scope of this book. We therefore limit ourselves to a crude, but useful, approximation called the ideal transformer.

25.1 CIRCUITS COMPRISING SELF-INDUCTANCES

If a circuit comprises self-inductance—we exclude mutual inductance for the moment—then both the Kirchhoff current law (KCL) and the Kirchhoff voltage law (KVL) (Sec. 7.8) apply as previously, the voltage drop across an inductor being $L \frac{dI}{dt}$.

For sinusoidal currents, the impedance of an inductor is $j\omega L$, and the delta-star transformations of Sec. 8.9 apply. Ohm's law becomes $I = V/Z$.

Example | **SERIES RESONANCE**

In Fig. 25-1, a source of alternating voltage feeds a series *LRC* circuit. Then $I = V/Z$, where

$$Z = R + j\left(\omega L - \frac{1}{\omega C}\right) = R + j\omega L\left(1 - \frac{1}{\omega^2 LC}\right). \quad (25-1)$$

At *resonance*,

$$\omega L = \frac{1}{\omega C} \quad \text{or} \quad \omega^2 LC = 1. \quad (25-2)$$

The impedances of the inductance and of the capacitance then cancel, Z is real and equal to R , and $I = V/R$. The resonance circular frequency is

$$\omega_r = \frac{1}{(LC)^{1/2}}. \quad (25-3)$$

It is convenient to use the two dimensionless numbers

$$\omega' = \frac{\omega}{\omega_r} \quad \text{and} \quad Q = \frac{(L/C)^{1/2}}{R} = \frac{\omega_r L}{R}. \quad (25-4)$$

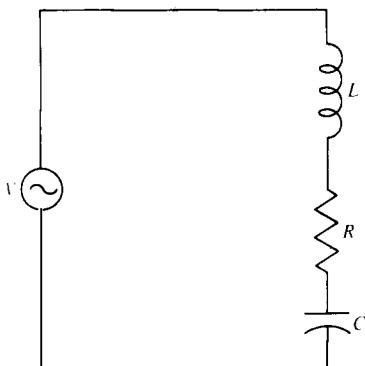
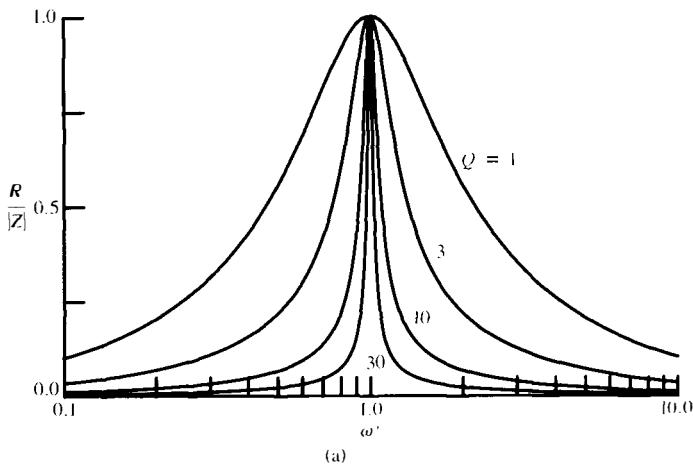


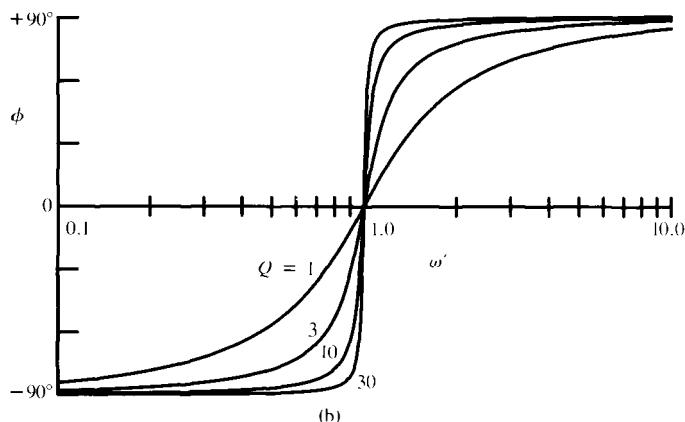
Fig. 25-1. Series-resonant circuit connected to a source of alternating voltage.

The second one is a measure of the “quality” of a circuit: the higher the Q , the lower the dissipation.[†] Then

$$\frac{Z}{R} = 1 + jQ \left(\omega' - \frac{1}{\omega'} \right). \quad (25-5)$$



(a)



(b)

Fig. 25-2. (a) $R/|Z|$ as a function of $\omega' = \omega/\omega_0$ for the circuit of Fig. 25-1 and for various values of Q . (b) The phase ϕ of Z as a function of ω' for the same values of Q .

[†] Estill I. Green has recounted “The Story of Q ” in *The American Scientist*, vol. 43, p. 584 (1955).

Figure 25-2(a) shows curves of $R/|Z| = R |Y|$ as a function of ω' for various values of Q . By definition, the *width of the resonance peak* is the difference Δf between the frequencies for which

$$\frac{R}{|Z|} = \frac{1}{2^{1/2}} = 0.7071. \quad (25-6)$$

and

$$\Delta f = \frac{f_r}{Q}. \quad (25-7)$$

The higher the Q , the sharper the resonance peak.

The phase angle of Z is

$$\phi = \arctan \left[Q \left(\omega' - \frac{1}{\omega'} \right) \right] \quad (25-8)$$

and varies from $-\pi/2$ to $\pi/2$, as in Fig. 25-2(b).

Example

PARALLEL RESONANCE

Figure 25-3 shows a parallel-resonant circuit fed by a source V . There is a resistance R in series with L , but none in series with C because real inductors are lossy, whereas real capacitors are nearly lossless. Here

$$Y = j\omega C + \frac{1}{R + j\omega L} = j\omega C + \frac{R - j\omega L}{R^2 + \omega^2 L^2} \quad (25-9)$$

and after a fair amount of algebra, we find that

$$R |Y| = \left[\frac{(\omega'^2 - 1)^2 + \omega'^2/Q^2}{1 + Q^2\omega'^2} \right]^{1/2}. \quad (25-10)$$

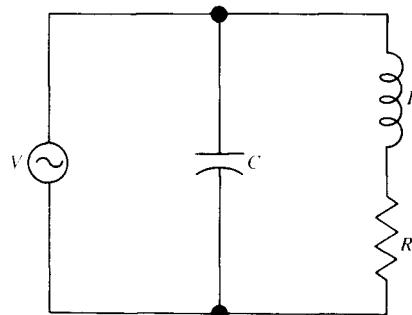


Fig. 25-3. Parallel-resonant circuit.

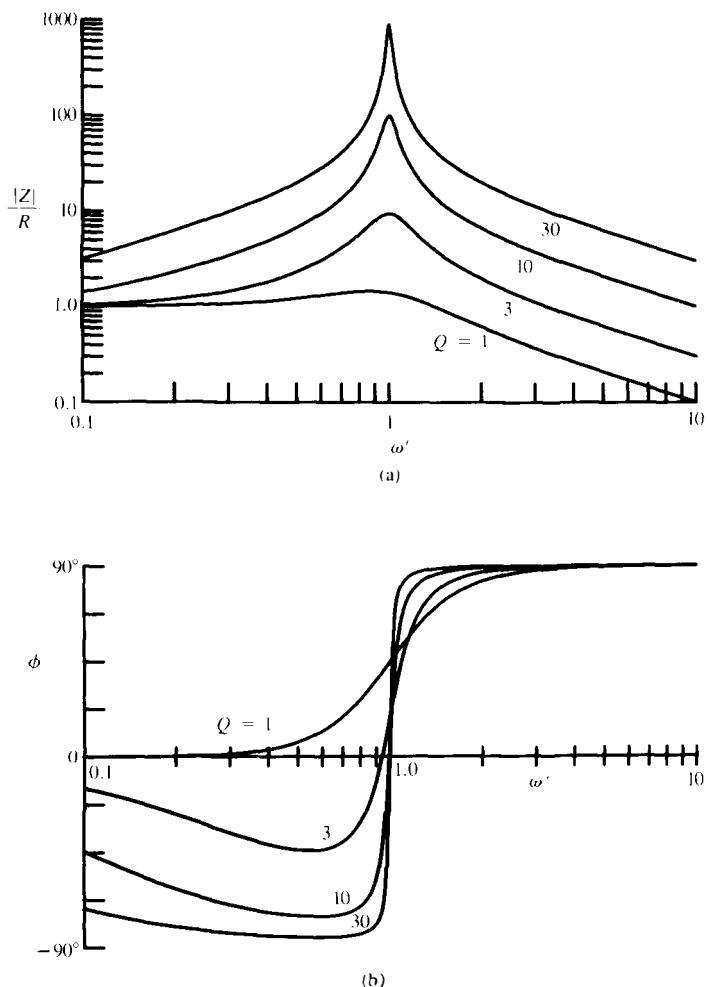


Fig. 25-4. (a) $|Z|/R$ as a function of ω' for the circuit of Fig. 25-3 for various values of Q . (b) The phase angle ϕ of Z as a function of ω' for the same values of Q .

Figure 25-4(a) shows the inverse of this quantity, namely $|Z|/R$, as a function of ω' for various values of Q . The maximum occurs slightly below $\omega' = 1$. Figure 25-4(b) shows the phase angle ϕ of Z .

By definition, the width of the resonance peak is again the difference Δf between the two frequencies where $|Z|/R$ is equal to $1/2^{1/2}$, or 0.7071, and $\Delta f \approx f_r/Q$. This relation is now approximate. It is valid within 1% for $Q \geq 1.7$.

25.2 CIRCUITS COMPRISING SELF-INDUCTANCE AND MUTUAL INDUCTANCE

There exists a mutual inductance between branches a and b in the circuit of Fig. 25-5(a). Then, to apply Kirchhoff's laws, one adds a voltage source $-j\omega MI_b$ in branch a , as in Fig. 25-5(b), and a voltage source $-j\omega MI_a$ in branch b . One chooses the sign of M as in Sec. 24.1.2. If branches a and b have a common terminal, then one can also proceed as in Sec. 25.3.

We use signs and arrows on circuit diagrams, as with direct currents. Then we apply Kirchoff's laws as we would with direct currents.

Example | SIMPLE CIRCUIT COMPRISING A TRANSFORMER

Figure 25-6(a) shows a transformer fed by a source V and feeding a load resistance. Figure 25-6(b) shows a circuit that is equivalent but without mutual inductance. The Kirchhoff voltage law yields two equations, one for the primary mesh and one for the secondary:

$$V_1 - j\omega Ml_2 - Z_1 l_1 = 0, \quad -j\omega Ml_1 - Z_2 l_2 = 0. \quad (25-11)$$

Here Z_1 is the impedance $R_1 + j\omega L_1$ of the primary winding, and $Z_2 = R_2 + j\omega L_2$ is the impedance of the complete secondary circuit.

Solving,

$$l_1 = \frac{V_1}{Z_1 + \omega^2 M^2 / Z_2}, \quad l_2 = -\frac{j\omega M V_1}{Z_1 Z_2 + \omega^2 M^2}. \quad (25-12)$$

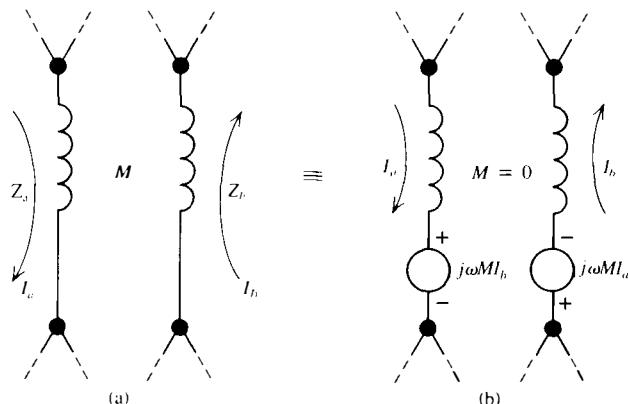


Fig. 25-5. (a) Part of a circuit in which there exists a mutual inductance M between two branches. (b) Equivalent circuit.

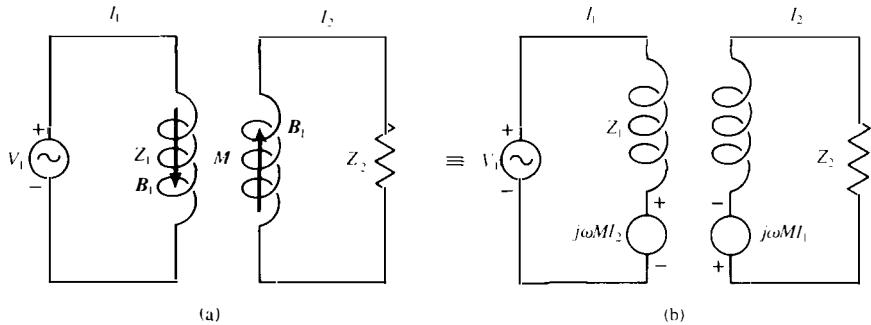


Fig. 25-6. (a) Transformer feeding a resistance. (b) Equivalent circuit.

The current in the primary is the same as if one removed the secondary and added the *reflected impedance* $\omega^2 M^2 / Z_2$ in the primary. So the *input impedance* is

$$Z_{in} = \frac{V_1}{I_1} = Z_1 + \frac{\omega^2 M^2}{Z_2}. \quad (25-13)$$

25.3 TRANSFORMATION OF A MUTUAL INDUCTANCE

Figure 25-7(a) shows a mutual inductance M with its two coils connected at C . We can transform this mutual inductance to the equivalent star of Fig. 25-7(b) in the following way. Assume that C is at ground potential.

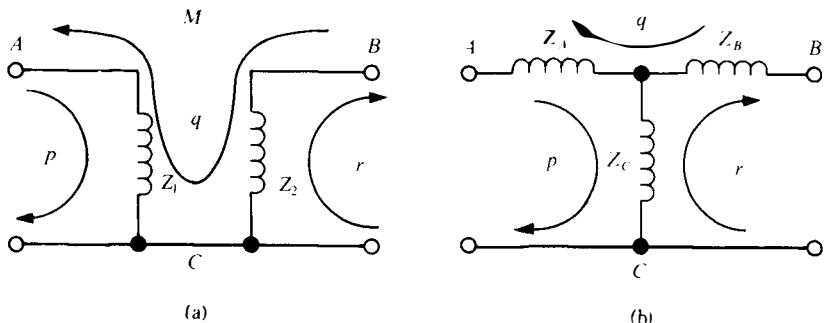


Fig. 25-7. Transformation of a mutual inductance. (a) The actual circuit. Observe that, for this transformation, the two inductors must have one common terminal. (b) In the equivalent circuit $Z_A = Z_1 + j\omega M$, $Z_B = Z_2 + j\omega M$, $Z_C = -j\omega M$.

Since the two circuits are equivalent, they carry the same mesh currents p, q, r . We assume that the currents are sinusoidal, and we thus use phasors.

In Fig. 25-7(a),

$$V_A = (p - q)Z_1 + j\omega M(r - q). \quad (25-14)$$

The second term on the right is the voltage induced in branch AC . We have assumed that clockwise currents are positive. For the sign of M , see Sec. 24.1.2. Also,

$$V_B = (q - r)Z_2 - j\omega M(p - q). \quad (25-15)$$

In Fig. 25-7(b),

$$V_A = (p - q)Z_A + (p - r)Z_C, \quad (25-16)$$

$$V_B = (q - r)Z_B + (p - r)Z_C. \quad (25-17)$$

After equating the two values of V_A , then the two values of V_B , and simplifying, we find that

$$p(Z_1 - Z_A - Z_C) + q(Z_A - Z_1 - j\omega M) + r(Z_C + j\omega M) = 0, \quad (25-18)$$

$$p(Z_C + j\omega M) + q(Z_B - Z_2 - j\omega M) + r(Z_2 - Z_B - Z_C) = 0. \quad (25-19)$$

These equations are valid for any set of arbitrary values of p, q, r . Then the six parenthetical expressions are zero, and

$$Z_A = Z_1 + j\omega M, \quad Z_B = Z_2 + j\omega M, \quad Z_C = -j\omega M. \quad (25-20)$$

If M is positive, then the reactance $-j\omega M$ is capacitive. If M is negative, $-j\omega M$ is the reactance of a pure self-inductance $|M|$.

Example

SOLVING THE TRANSFORMER CIRCUIT IN ANOTHER WAY

The circuit of Fig. 25-8(a) is the same as that of Fig. 25-6(a). We now find I_1 , I_2 , and Z_{in} by transforming the mutual inductance. Note that Z_2 is the secondary impedance plus R_L . This leads to the equivalent circuit of Fig. 25-8(b).

Applying the Kirchhoff voltage law to the two meshes in succession,

$$V_1 - (Z_1 + j\omega M)I_1 - (-j\omega M)(I_1 - I_2) = 0, \quad (25-21)$$

$$-(-j\omega M)(I_2 - I_1) - (Z_2 + j\omega M)I_2 = 0. \quad (25-22)$$

These equations yield Eqs. 25-12.

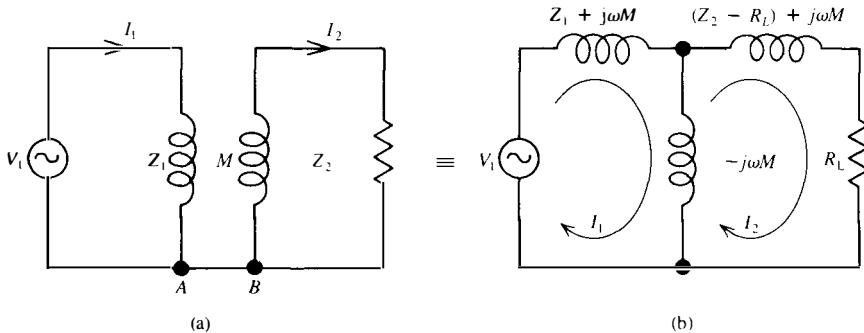


Fig. 25-8. (a) A simple transformer circuit. The impedance Z_2 is that of the complete secondary circuit. (b) The equivalent circuit.

25.4 MAGNETIC-CORE TRANSFORMERS

Magnetic-core transformers serve at frequencies ranging from a few hertz to about 1 megahertz. Figure 25-9 shows one common type that is used at 60 hertz.

A magnetic-core transformer possesses two essential features. (1) For a given cross section, the magnetic flux per ampere-turn in the primary is larger than with an air-core transformer by a few orders of magnitude. This permits the design of transformers that are smaller and have fewer

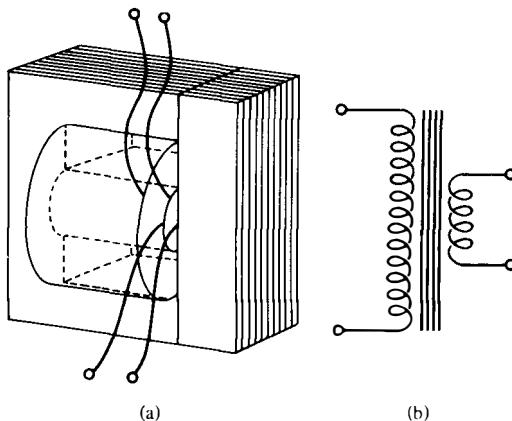


Fig. 25-9. (a) Common type of magnetic-core transformer for use at 60 hertz. The core has the shape of a figure eight. Both primary and secondary windings surround the center leg. The core comprises two types of lamination, one shaped like an E and the other like an I. The laminations are assembled after the coils have been wound. (b) Schematic diagram for a magnetic-core transformer.

turns. (2) The magnetic flux through the secondary is nearly equal to that through the primary because the magnetic flux follows the core. Thus the coupling coefficient is close to unity and two such transformers can operate close together with little interaction.

The core design minimizes the effects of eddy currents. (See the example of Sec. 23.5.)

The analysis of magnetic-core transformers is difficult for several reasons. First, the relationship between B and H in ferromagnetic materials is not linear (Sec. 21.2). For example, if the flux linkage in a circuit is Λ when it carries a current I , then the self-inductance L is Λ/I , as in Sec. 24.2. If there are no ferromagnetic materials in the field, Λ is proportional to I and L depends solely on the geometry of the circuit. However, in the presence of magnetic materials, Λ is not proportional to I , and the value of the self-inductance $L = \Lambda/I$ can only be approximate.

Moreover, the losses in a magnetic-core transformer are complex: there are eddy-current losses (Sec. 23.5), both in the iron core and in the copper windings, hysteresis losses in the iron core (Sec. 21.2), and Joule losses resulting from the currents flowing in the windings. All these losses can be expressed as an I^2R loss in the primary but, for a given transformer, R depends on the voltage applied to the primary, on the current drawn from the secondary, and on the frequency.

25.4.1 The Ideal Transformer

The ideal transformer is a coarse, but useful, approximation. We make the following assumptions. (1) There are no Joule or eddy-current losses. (2) The hysteresis loop for the core is a straight line through the origin. Then B is proportional to H , and there are no hysteresis losses either. (3) All the magnetic flux is confined to the core. Then the coupling coefficient is equal to unity, and the flux through the primary is equal to that through the secondary.

As a consequence of the first two assumptions, the transformer is lossless. The efficiency is, in fact, close to 100% for large transformers, but only of the order of 75% for small power transformers supplying tens of watts.

The assumption that B is proportional to H , and hence to I , is not realistic, but it is difficult to avoid.

25.4.2 The Ratio V_1/V_2

As a rule, the load connected to the secondary terminals is a resistance, which we denote R_L .

With an ideal transformer the voltage V_1 applied to the primary is proportional to the magnetic flux in the core:

$$V_1 = N_1 \frac{d\Phi}{dt} = j\omega N_1 \Phi. \quad (25-23)$$

The voltage across the secondary is

$$V_2 = N_2 \frac{d\Phi}{dt} = N_2 j\omega \Phi = I_2 R_L, \quad (25-24)$$

where N_1 and N_2 are the numbers of turns in the primary and in the secondary windings, respectively, and I_2 is the current in the secondary. Thus

$$\frac{V_1}{V_2} = \frac{N_1}{N_2}. \quad (25-25)$$

In an ideal transformer and for a given V_1 , V_2 is independent of the load current. Also V_2 is either in phase with V_1 or π radians out of phase. Of course, one can change the phase of the voltage on R_L by π radians by interchanging the connections to the secondary winding.

Note that V_1 is equal to $N_1 j\omega \Phi$. Thus, for a given applied voltage, the magnetic flux is independent of the current drawn from the secondary.

Also, Φ is B times the cross section of the core. Since the maximum possible value for B depends on the core material, Eq. 25-67 shows that, for a given V_1 , an increase in frequency permits the use of a smaller Φ , and hence of a core of smaller cross section.

25.4.3 The Ratio L_1/L_2

Let Φ_1 be the magnetic flux when I_2 is zero, and, similarly, let Φ_2 be the magnetic flux when I_1 is zero. The reluctance (Sec. 21.4) of the core is \mathcal{R} . Then

$$L_1 = N_1 \frac{\Phi_1}{I_1} = N_1 \frac{N_1 I_1 / \mathcal{R}}{I_1} = \frac{N_1^2}{\mathcal{R}}. \quad (25-26)$$

Similarly,

$$L_2 = \frac{N_2^2}{\mathcal{R}} \quad (25-27)$$

and

$$\frac{L_1}{L_2} = \left(\frac{N_1}{N_2} \right)^2. \quad (25-28)$$

The principle of superposition applies because, by hypothesis, B/H is constant. So $\Phi = \Phi_1 + \Phi_2$.

25.4.4 The Input Impedance Z_{in}

From the example in Sec. 25.2, with $M^2 = L_1 L_2$, $R_1 \approx 0$, $R_2 \approx R_L$,

$$Z_{in} = j\omega L_1 + \frac{\omega^2 L_1 L_2}{R_L + j\omega L_2} = j\omega L_1 \left(1 - \frac{j\omega L_2}{R_L + j\omega L_2} \right) \quad (25-29)$$

$$= \frac{R_L j\omega L_1}{R_L + j\omega L_2} \approx \frac{L_1}{L_2} R_L \quad (R_L \ll \omega L_2). \quad (25-30)$$

Then

$$Z_{in} = \left(\frac{N_1}{N_2} \right)^2 R_L \quad (R_L \ll \omega L_2), \quad (25-31)$$

and the input impedance is real.

25.4.5 The Ratio I_1/I_2

The voltage induced in the secondary is equal to $j\omega M I_1$. Therefore

$$j\omega M I_1 = (R_L + j\omega L_2) I_2, \quad (25-32)$$

$$\frac{I_2}{I_1} = \frac{j\omega M}{R_L + j\omega L_2} \approx \frac{M}{L_2} = \left(\frac{L_1}{L_2} \right)^{1/2} = \frac{N_1}{N_2}. \quad (25-33)$$

Thus the transformer is lossless, as we assumed at the beginning:
 $I_1 V_1 = I_2 V_2$.

Example | THE AUTOTRANSFORMER

Figure 25-10(a) shows a schematic diagram of an autotransformer. Its single winding serves as both primary and secondary. The winding is often wound over a toroidal core.

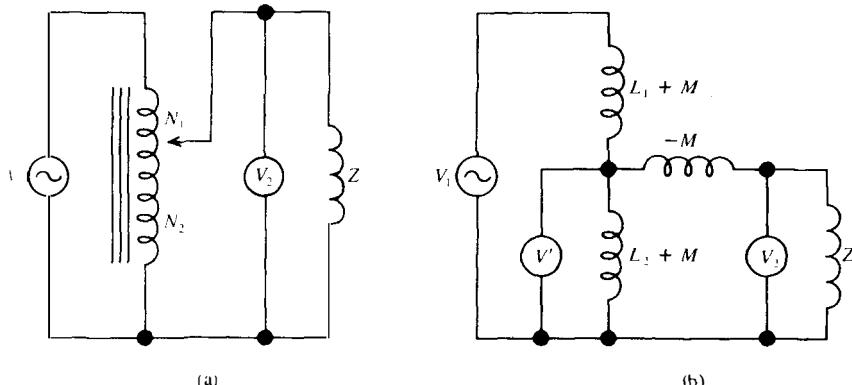


Fig. 25-10. (a) Schematic diagram of an autotransformer. (b) The equivalent circuit.

Let us calculate the ratio V_2/V_1 . We assume a coupling coefficient of unity, and we disregard the resistance of the winding. These assumptions are reasonable in practice. Let the numbers of turns on either side of the tap be N_1 and N_2 , with corresponding inductances L_1 and L_2 as in Fig. 25-10(b). Then, if A is a constant of proportionality,

$$L_1 = AN_1^2, \quad L_2 = AN_2^2, \quad M = (L_1 L_2)^{1/2} = AN_1 N_2. \quad (25-34)$$

We set

$$\frac{V_2}{V_1} = \frac{V_2}{V'} \frac{V'}{V_1}, \quad (25-35)$$

where

$$\frac{V_2}{V'} = \frac{Z}{Z - j\omega M} \quad (25-36)$$

and

$$\frac{V'}{V_1} = \frac{\frac{(Z - j\omega M)j\omega(L_2 + M)}{Z - j\omega M + j\omega(L_2 + M)}}{\frac{(Z - j\omega M)j\omega(L_2 + M)}{Z - j\omega M + j\omega(L_2 + M)} + j\omega(L_1 + M)}. \quad (25-37)$$

After multiplying these two ratios, substituting the values of L_a , L_b , M , and simplifying, we are left with

$$\frac{V_2}{V_1} = \frac{N_2}{N_1 + N_2}. \quad (25-38)$$

So V_2 is simply proportional to that fraction of the winding that is spanned by the secondary. The output voltage is even independent of the load impedance Z ! Recall that we have disregarded the resistance of the winding.

25.4.6 Power Transfer from Source to Load Through a Transformer

As we saw in Sec. 8.8, the power dissipated in a load is maximum when its resistance R_L is equal to the output resistance of the source R_s , and when $X_L = -X_s$. As a rule, the reactances are zero.

If it is impossible to vary either R_L or R_s , then one can still achieve optimum power transfer by inserting a transformer between source and load, as in Fig. 25-11(a). Let us see how this comes about.

Assume that the transformer has a magnetic core, that the approximations of Sec. 25.4 are satisfactory, and that $X_s = X_L = 0$. Then, from Sec. 25.4.4, the transformer has an input impedance of $(N_1/N_2)^2 R_L$ ohms. In other words, the current and the power supplied by the source are precisely the same as if the transformer and its load resistance R_L were

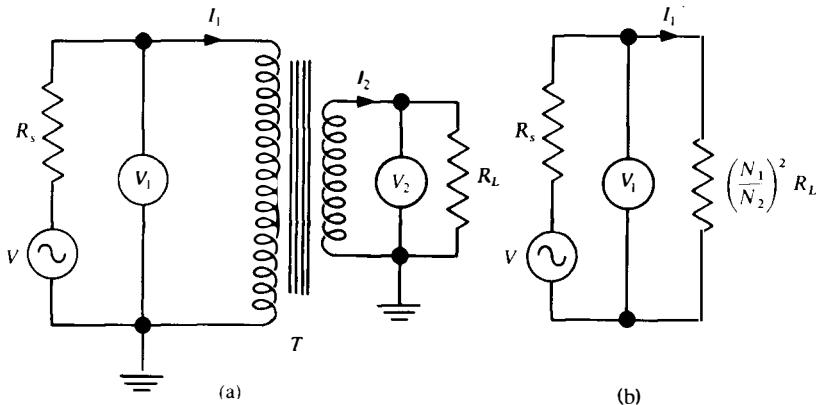


Fig. 25-11. (a) Transformer T inserted between a source of internal resistance R_s and a load resistance R_L . (b) Equivalent circuit.

replaced by the resistance $(N_1/N_2)^2 R_L$, as in Fig. 25-11(b). Remember that with our approximation the transformer has an efficiency of 100%.

Then, with a transformer as in Fig. 25-16, the power transfer is optimum when

$$R_s = \left(\frac{N_1}{N_2}\right)^2 R_L, \quad \text{or} \quad \frac{N_1}{N_2} = \left(\frac{R_s}{R_L}\right)^{1/2}. \quad (25-39)$$

The transformer is then said to be used for *impedance matching*. When this condition applies, the power dissipation in the load is maximum, but the efficiency is still only 50%.

25.5 SUMMARY

Kirchhoff's voltage and current laws apply to circuits comprising self-inductances. The voltage drop across an inductance is $L dI/dt$, or $j\omega LI$ with sinusoidal currents.

If a circuit comprises a *mutual inductance* M between two branches, then one can apply Kirchhoff's laws by disregarding M and adding to each branch a voltage source $-j\omega M$ times the current in the other branch. If the two branches have one common terminal, one can transform them into a Y circuit, as in Fig. 25-7.

In an *air-core transformer*, the *input impedance* is

$$Z_{in} = Z_1 + \frac{\omega^2 M^2}{Z_2}, \quad (25-13)$$

where Z_1 is the impedance of the primary *winding* and Z_2 is the impedance of the *complete* secondary. Also,

$$I_2 = -\frac{j\omega MV_1}{Z_1 Z_2 + \omega^2 M^2} \quad (25-12)$$

and V_2 is I_2 times the load impedance Z_L .

The *ideal transformer* approximates, in a crude way, the behavior of a magnetic-core transformer. We found that

$$V_1 = j\omega N_1 \Phi, \quad (25-23)$$

$$\frac{I_1}{I_2} = \frac{V_2}{V_1} = \frac{N_2}{N_1} \quad (R_L \ll \omega L_2) \quad (25-25), (25-33)$$

$$Z_{in} = \left(\frac{N_1}{N_2} \right)^2 R_L, \quad (R_L \ll \omega L_2). \quad (25-31)$$

One can achieve maximum power transfer between a source and a load by interposing a transformer with a turns ratio $N_1/N_2 = (R_s/R_L)^{1/2}$.

PROBLEMS

25-1. (25.1) Impedance

- (a) Calculate the impedance Z of the circuit shown in Fig. 25-12. What is the value of Z when (i) $f = 0$, (ii) $f \rightarrow \infty$?
- (b) Calculate the magnitude and the phase angle of Z at 1 kilohertz.
- (c) Calculate the amplitude and the phase angle of $Y = 1/Z$ at that frequency.
- (d) Calculate the power dissipation when the current is 100 milliamperes, again at 1 kilohertz.
- (e) Can the real part of the impedance become negative?
- (f) For what frequency ranges is the circuit equivalent to (i) a resistor in series with an inductor, (ii) a resistor in series with a capacitor?
- (g) At what frequency is the circuit equivalent to a pure resistance?

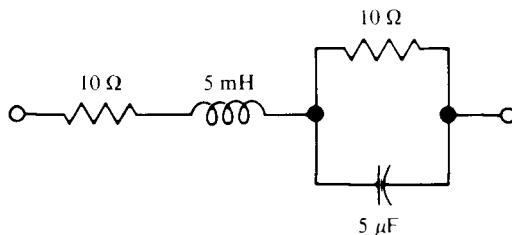


Fig. 25-12.

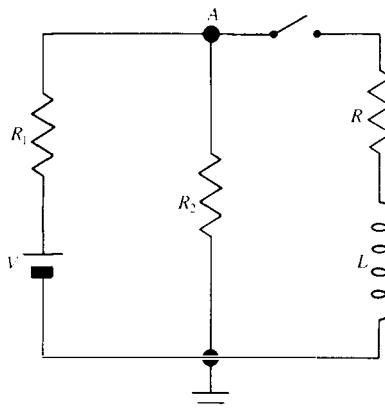


Fig. 25-13.

25-2. (25.1) The magnetic energy stored in an inductor

A voltage source V is connected through a switch to an inductor of inductance L and resistance R . The switch closes at $t = 0$.

Show that at any time T the energy that has been supplied by the source, minus the energy dissipated in the resistance, is equal to the magnetic energy $I^2L/2$.

25-3. (25.1) RL circuit

Find the current in the inductance L of the circuit of Fig. 25-13. The switch closes at $t = 0$. If you have studied Chap. 8, use the substitution theorem and then Millman's theorem.

25-4. (25.1) The star-delta transformation with a self-inductance

Section 8.9 is a prerequisite for this problem. Show that the star and the delta of Fig. 25-14 are equivalent at 1 kilohertz.

25-5. The coefficient of coupling

The coefficient of coupling k between two single-turn coils was defined in Sec. 24.3. In general, $k_a \neq k_b$. Show that $k_{ab}/k_{ba} = L_b/L_a$.

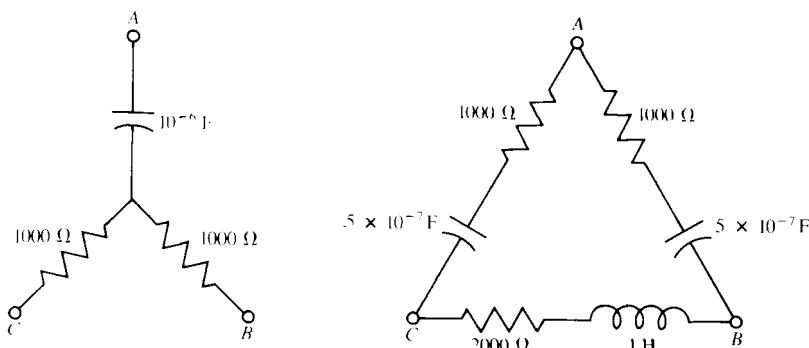


Fig. 25-14.

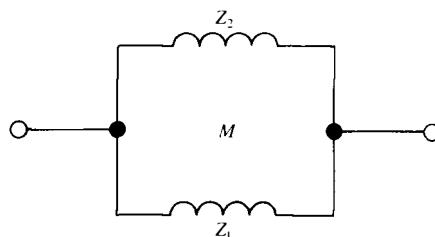


Fig. 25-15.

25-6. (25.2) Measurement of the coefficient of coupling k

A transformer has a primary inductance L_1 , a secondary inductance L_2 , and a mutual inductance M . The winding resistances are negligible.

Show that $Z_0/Z_\infty = 1 - k^2$, where Z_0 and Z_∞ are the impedances measured at the terminals of the primary, when the secondary is short-circuited and when it is open-circuited.

25-7. (25.3) Impedances in parallel, with mutual inductance

Calculate the impedance of the circuit shown in Fig. 25-15.

25-8. (25.4) Improving (?) iron-core transformers

In a iron-core transformer, the windings are outside the core, where B is orders of magnitude smaller than inside. Why not put them inside?

25-9. (25.4) Eddy-current losses in transformer laminations

Eddy-current losses in magnetic cores are minimized by assembling them from laminations. Consider a core of rectangular cross section as in Fig. 25-16. The eddy-current loss is proportional to \mathcal{V}^2/R , where \mathcal{V} is the electromotance induced around a typical current path such as the one shown by a dashed curve. The resistance is also difficult to define, but it is of the order of twice the resistance of the upper half, or $2a/[\sigma(b/a)L]$.

(a) Show that splitting the core into n laminations reduces eddy-current losses by a factor of n^2 .

(b) Show that these losses increase as the square of the frequency.

25-10. (25.4) Hysteresis losses

Hysteresis losses are proportional to the operating frequency f , while eddy-current losses increase as f^2 , as we saw above. You are given a number of transformer laminations. Can you devise an experiment that will permit you to evaluate the relative importance of the two types of loss?

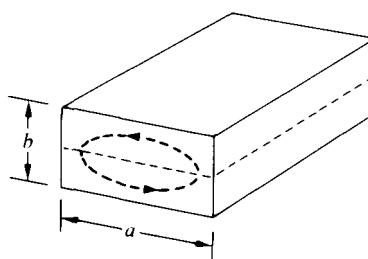


Fig. 25-16.

25-11. (25.4.4) The reflected impedance

Show that a positive (inductive) reactance in the secondary of a transformer is equivalent to a negative (capacitive) reactance in the primary, and inversely.

25-12. (25.4.4) Electromagnetic crack detectors and metal detectors

It is possible to detect cracks in metallic objects as follows. If the part to be examined is placed in the vicinity of a coil fed with alternating current, the inductance measured at the coil terminals is lowest when there are no cracks. Such instruments can detect cracks only 10 micrometers deep. The coil forms part of a resonant circuit.[†] Metal detectors operate similarly.

Consider the following simpler situation. A single-layer close-wound solenoid has a length l , a radius a , and N turns. Let us calculate how its impedance changes when one introduces into the solenoid a thin brass tube of wall thickness b .

When an alternating current flows in the solenoid, the changing magnetic flux induces a current in the tube, which thus acts as a secondary winding. According to Lenz's law, the induced current tends to cancel $d\Phi/dt$, and hence $j\omega\Phi$, and hence Φ . The presence of the tube thus reduces the inductance at the solenoid terminals. The effective inductance of the solenoid decreases when the resistance of the tube decreases.

We disregard the skin effect (Sec. 29.1) in the tube and the stray capacitance of the coil. Also, we set $l \gg a$ so as to disregard end effects. The coefficient of coupling is nearly equal to unity.

(a) Calculate the resistance R_1 of the winding of the solenoid. Set the conductivity of copper equal to σ_c .

(b) Calculate the impedance Z_1 of the solenoid without the brass tube.

(c) Calculate the resistance R_2 of the brass tube in the azimuthal direction. Set its radius equal to a , and call its conductivity σ_b .

(d) Calculate its inductance L_2 and impedance Z_2 .

(e) Now calculate the impedance at the solenoid terminals with the tube in place.

(f) Calculate impedances, without and with the brass tube, when $N = 1000$, $l = 200$ millimeters, $a = 20.0$ millimeters, $b = 0.5$ millimeter, $f = 1000$ hertz, $\sigma_c = 5.8 \times 10^7$ siemens/meter, $\sigma_b = 1.6 \times 10^7$ siemens/meter. Note how the presence of the tube increases R (more dissipation) and decreases L (less flux).

25-13. (25.4.5) Soldering gun

A soldering gun consists of a step-down transformer that feeds a large current through a length of copper wire. One type dissipates 100 watts in a piece of copper wire ($\sigma = 5.8 \times 10^7$ siemens/meter) having a cross section of 4 millimeters² and a length of 100 millimeters.

(a) Find V and I in the secondary.

(b) Find the current in the primary if it is fed at 120 volts, assuming an efficiency of 100%.

[†] See Prob. 17-15 in *Electromagnetism: Principles and Applications* by the same authors and the same publisher.

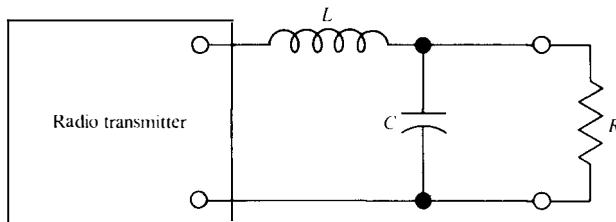


Fig. 25-17.

25-14. (25.4.7) Impedance matching with an LC circuit

Figure 25-17 shows a radio transmitter connected to a resistance R that represents an antenna. Set the impedance Z_L seen by the transmitter equal to $R_L + jX_L$.

- (a) Under what condition is $X_L = 0$?
- (b) Then what is the value of R/R_L ?
- (c) Calculate the values of C and L for $R = 50$ ohms and $f = 14$ megahertz if R/R_L must equal 12.5.
- (d) Now plot R_L and X_L as functions of the frequency between 13.5 and 14.5 megahertz.

This LC circuit is inexpensive, compared to a transformer, but the impedance match applies only at the design frequency.

CHAPTER 26

MAGNETIC FIELDS IX

Magnetic Energy and Macroscopic Magnetic Forces

- 26.1 ENERGY STORAGE IN AN INDUCTIVE CIRCUIT. THE MAGNETIC ENERGY \mathcal{E}_m EXPRESSED IN TERMS OF L AND I 472
Example 474
- 26.2 THE MAGNETIC ENERGY DENSITY \mathcal{E}'_m EXPRESSED IN TERMS OF \mathbf{J} AND \mathbf{A} 474
- 26.3 THE MAGNETIC ENERGY DENSITY \mathcal{E}'_m EXPRESSED IN TERMS OF \mathbf{H} AND \mathbf{B} 475
Example: THE LONG SOLENOID 476
- *26.4 THE SELF-INDUCTANCE OF A VOLUME DISTRIBUTION OF CURRENT 477
**Example: THE SELF-INDUCTANCE OF A COAXIAL LINE* 477
- *26.5 THE FORCE BETWEEN TWO CURRENT-CARRYING CIRCUITS EXPRESSED IN TERMS OF M , I_a , AND I_b 478
- *26.6 THE FORCE BETWEEN TWO CURRENT-CARRYING CIRCUITS EXPRESSED IN TERMS OF THE MAGNETIC ENERGY \mathcal{E}_m 481
Example: THE FORCE BETWEEN TWO LONG COAXIAL SOLENOIDS 481
- *26.7 MAGNETIC FORCES AND LINES OF \mathbf{B} 482
- *26.8 MAGNETIC PRESSURE 483
Example: THE MAGNETIC PRESSURE INSIDE A LONG SOLENOID 484
- *26.9 MAGNETIC FORCES WITHIN AN ISOLATED CIRCUIT 485
- *26.10 MAGNETIC TORQUE 485
Example: THE MAGNETIC TORQUE EXERTED ON A CURRENT LOOP 485
- *26.11 SUMMARY 486
- PROBLEMS 487

This is the last chapter dealing specifically with magnetic fields. It concerns energy and macroscopic forces. We shall find several expressions for the magnetic energy stored in a field, and then deduce the forces exerted on a current-carrying body situated in a magnetic field that originates elsewhere. As in most of the material that we have studied until now, we restrict ourselves to low frequencies and to currents that result from the motion of free charges.

26.1 ENERGY STORAGE IN AN INDUCTIVE CIRCUIT. THE MAGNETIC ENERGY \mathcal{E}_m EXPRESSED IN TERMS OF L AND I

The circuit of Fig. 26-1 illustrates the storage of magnetic energy in inductive circuits. The wire is of uniform cross section \mathcal{A} and uniform conductivity σ . All materials are nonferromagnetic, and we assume that $\epsilon_r \approx 1$ for the wire. The circuit is rigid.

(1) The current is constant. Inside the wire the volume charge density is zero, at least in the absence of a $\mathbf{v} \times \mathbf{B}$ field, and

$$E = |\nabla V| = \frac{J}{\sigma}. \quad (26-1)$$

If there are no sharp bends, both E and J are uniform. The lines of E , inside, follow the wire, parallel to its axis.

At the surface of the wire, \mathbf{E} has both a normal and a tangential component. If the wire is situated in a vacuum, the free surface charge density is $\epsilon_0 E_n$, where E_n is the normal component of \mathbf{E} just outside. If it lies in a dielectric, then the free surface charge density is $D_n = \epsilon_r \epsilon_0 E_n$.

At any point in space \mathbf{E} is proportional to V_0 . Of course, \mathbf{E} also depends on the geometry of the circuit and on the neighboring objects.

In principle, one could calculate V and \mathbf{E} everywhere from the surface charge density all around the circuit, source included, and on the neighboring bodies. However, σ is itself a function of \mathbf{E} .

If l is the length of the wire, \mathcal{A} its cross section, and R its resistance $l/(\sigma \mathcal{A})$, then

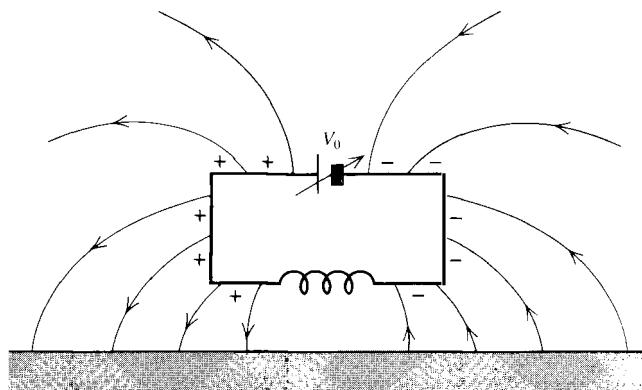


Fig. 26-1. Coil fed by a battery.

$$I |\nabla V| = IE = \frac{I}{\sigma} = \frac{I}{A\sigma} = IR = V_0. \quad (26-2)$$

Part of the power supplied by the source serves to establish the magnetic field, and the rest dissipates as heat.

(2) Now set the source voltage V_0 at zero and increase it slowly. Then

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t}. \quad (26-3)$$

The wire now lies in its own $\partial \mathbf{A}/\partial t$ field.

The relation $\mathbf{J} = \sigma \mathbf{E}$ still applies, and both \mathbf{J} and \mathbf{E} are uniform throughout the wire, as previously. This means that the surface charges distribute themselves so as to maintain a uniform \mathbf{E} inside, despite the presence of the $\partial \mathbf{A}/\partial t$ term.

Inside the wires that go from the source to the coil, \mathbf{A} is weak and $\mathbf{E} \approx -\nabla V$. Inside the coil wire, we have the same \mathbf{E} , but it comes partly from $-\nabla V$, which points in the direction of \mathbf{J} , and partly from $-\partial \mathbf{A}/\partial t$, which points in the opposite direction because \mathbf{A} points in the direction of \mathbf{J} and increases.

Now we have just seen that $IE = IR$. Integrating \mathbf{E} inside the wire from the positive to the negative terminal of the source,

$$\int_{+}^{-} \left(-\nabla V - \frac{\partial \mathbf{A}}{\partial t} \right) \cdot d\mathbf{l} = IR, \quad (26-4)$$

$$-\int_{+}^{-} \nabla V \cdot d\mathbf{l} - \int_{+}^{-} \frac{\partial \mathbf{A}}{\partial t} \cdot d\mathbf{l} = IR. \quad (26-5)$$

Since the first integral equals V_0 ,

$$V_0 = IR + \int_{+}^{-} \frac{\partial \mathbf{A}}{\partial t} \cdot d\mathbf{l} = IR + \frac{d}{dt} \int_{+}^{-} \mathbf{A} \cdot d\mathbf{l} = IR + \frac{d\Lambda}{dt}. \quad (26-6)$$

We have used the fact that the line integral of \mathbf{A} is equal to the flux linkage (Sec. 19.1).

At any instant the power supplied by the source is

$$IV_0 = I^2 R + I \frac{d\Lambda}{dt}. \quad (26-7)$$

The first term on the right is the power dissipated as heat, and the second is the rate of increase of the magnetic energy.

Thus, if \mathcal{E}_m is the stored magnetic energy at a given instant,

$$\frac{d\mathcal{E}_m}{dt} = I \frac{d\Lambda}{dt} = I \frac{d(LI)}{dt} = \frac{1}{2}L \frac{dI^2}{dt}. \quad (26-8)$$

The inductance L is a constant because we have assumed that the circuit is rigid and that there are no ferromagnetic materials in the vicinity. Clearly, $\mathcal{E}_m = 0$ when $I = 0$. Then

$$\mathcal{E}_m = \frac{LI^2}{2} = \frac{I\Lambda}{2}. \quad (26-9)$$

If we have two circuits a and b , then

$$\mathcal{E}_{ma} + \mathcal{E}_{mb} = \frac{1}{2}(I_a\Lambda_a + I_b\Lambda_b) \quad (26-10)$$

$$= \frac{1}{2}[I_a(L_a I_a + MI_b) + I_b(L_b I_b + MI_a)] \quad (26-11)$$

$$= \frac{1}{2}L_a I_a^2 + \frac{1}{2}L_b I_b^2 + MI_a I_b. \quad (26-12)$$

Example

The magnetic energy stored in a long solenoid follows from the value of the self-inductance that we found in the first example in Sec. 24.2

$$\mathcal{E}_m = \frac{LI^2}{2} = \frac{\pi\mu_0 N^2 R^2 I^2}{2l}. \quad (26-13)$$

26.2 THE MAGNETIC ENERGY DENSITY \mathcal{E}'_m EXPRESSED IN TERMS OF \mathbf{J} AND \mathbf{A}

We can rewrite the time derivative of \mathcal{E}_m as follows. Since $\Lambda = LI$,

$$\begin{aligned} \frac{d\mathcal{E}_m}{dt} &= \frac{1}{2} \left(I \frac{d\Lambda}{dt} + \Lambda \frac{dI}{dt} \right) = I \frac{d\Lambda}{dt} = I \oint_C \frac{\partial \mathbf{A}}{\partial t} \cdot d\mathbf{l} \\ &= \int_A \mathbf{J} \cdot d\mathcal{A} \oint_C \frac{\partial \mathbf{A}}{\partial t} \cdot d\mathbf{l} \end{aligned} \quad (26-14)$$

$$= \int_V \mathbf{J} \cdot \frac{\partial \mathbf{A}}{\partial t} dv, \quad (26-15)$$

where \mathcal{A} is the cross-sectional area of the wire, C is the curve defined by the wire, and v is the volume of the wire. The volume v is finite.

Note in passing that $\mathbf{J} \cdot (\partial \mathbf{A} / \partial t) dv$ is the work done in moving the conduction charges situated in the element of volume dv against the electric field $-\partial \mathbf{A} / \partial t$ during 1 second.

Now consider the identity

$$\frac{d}{dt} \int_v \mathbf{J} \cdot \mathbf{A} dv = \int_v \mathbf{J} \cdot \frac{\partial \mathbf{A}}{\partial t} dv + \int_v \frac{\partial \mathbf{J}}{\partial t} \cdot \mathbf{A} dv. \quad (26-16)$$

The first integral on the right is equal to $I d\Lambda/dt$, as we saw above. Similarly, the second integral is equal to $\Lambda dI/dt$, which is the same as $I d\Lambda/dt$ if the inductance is constant. Therefore

$$\frac{d\mathcal{E}_m}{dt} = \frac{1}{2} \frac{d}{dt} \int_v \mathbf{J} \cdot \mathbf{A} dv \quad (26-17)$$

and

$$\mathcal{E}_m = \frac{1}{2} \int_v \mathbf{J} \cdot \mathbf{A} dv, \quad (26-18)$$

where, again, v is the volume of the conductor. This equation applies only if the source is of finite size.

If there are several circuits, then the integral runs over all of them and the vector potential in one circuit is the sum of the \mathbf{A} 's of all the circuits. One can add to \mathbf{A} any quantity whose curl is zero without affecting this integral (Prob. 26-4).

The magnetic energy density at a point can therefore be taken to be

$$\mathcal{E}'_m = \frac{1}{2} \mathbf{J} \cdot \mathbf{A}. \quad (26-19)$$

26.3 THE MAGNETIC ENERGY DENSITY \mathcal{E}'_m EXPRESSED IN TERMS OF \mathbf{H} AND \mathbf{B}

To express the magnetic energy in terms of \mathbf{H} and \mathbf{B} , we use Eq. 26-9 and apply it to the loop of Fig. 26-2. The loop lies in a homogeneous,

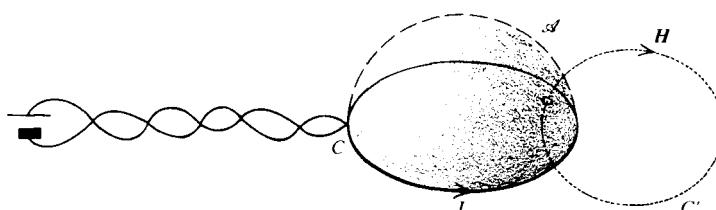


Fig. 26-2. Single-turn loop of wire C bearing a current I . The dotted line is a typical line of \mathbf{H} . The open surface, of area \mathcal{A} , is bounded by C , and it is everywhere orthogonal to \mathbf{H} .

isotropic, linear, and stationary (HILS) magnetic medium. This excludes ferromagnetic media. From Ampère's circuital law,

$$I = \oint_{C'} H \, dl, \quad (26-20)$$

where C' is any line of \mathbf{H} .

Also, let \mathcal{A} be the area of any open surface bounded by the loop C and orthogonal to the lines of \mathbf{H} and of \mathbf{B} . Then

$$\Lambda = \Phi = \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} \quad (26-21)$$

and

$$\mathcal{E}_m = \frac{1}{2} I \Lambda = \frac{1}{2} \oint_{C'} H \, dl \int_{\mathcal{A}} B \, d\mathcal{A}. \quad (26-22)$$

Now the lines of \mathbf{H} and the set of open surfaces define a coordinate system in which $d\mathbf{l} \cdot d\mathcal{A}$ is an element of volume with $d\mathbf{l}$ and $d\mathcal{A}$ both parallel to \mathbf{H} . Also, for each element $d\mathbf{l}$ along the chosen line of \mathbf{H} , one integrates over all the corresponding surface. Since the field extends to infinity, this double integral is the volume integral of $\mathbf{H} \cdot \mathbf{B}$ over all space, and

$$\mathcal{E}_m = \frac{1}{2} \int_{\infty} \mathbf{H} \cdot \mathbf{B} \, dv. \quad (26-23)$$

The *magnetic energy density* in nonferromagnetic media is thus

$$\mathcal{E}'_m = \frac{\mathbf{H} \cdot \mathbf{B}}{2} = \frac{\mathbf{B}^2}{2\mu} = \frac{\mu H^2}{2}. \quad (26-24)$$

The magnetic energy density varies as B^2 . Thus, after superposing several fields, the total field energy is *not* equal to the sum of the individual energies. See Eq. 26-12.

Compare this section with Sec. 6.2.

Example

THE LONG SOLENOID

Neglecting end effects, we find that the magnetic energy stored in the field of a long solenoid in air is

$$\mathcal{E}_m = \frac{1}{2\mu_0} (\mu_0 N' I)^2 \pi R^2 l = \frac{\pi R^2 \mu_0 N^2 I^2}{2l}, \quad (26-25)$$

as in the example in Sec. 26.1.

*26.4 THE SELF-INDUCTANCE OF A VOLUME DISTRIBUTION OF CURRENT

We saw in Sec. 24.2 that the self-inductance of a circuit comprising infinitely thin wires is infinite. A *real* circuit comprises conductors of finite cross section and its self-inductance L , by definition, is proportional to the stored energy:

$$\frac{1}{2}LI^2 = \mathcal{E}_m = \frac{1}{2\mu_0} \int_{\infty} B^2 dv. \quad (26-26)$$

Thus

$$L = \frac{1}{\mu_0 I^2} \int_{\infty} B^2 dv. \quad (26-27)$$

*Example THE SELF-INDUCTANCE OF A COAXIAL LINE

Assume that the frequency is low enough to ensure that the currents spread uniformly throughout the cross sections of the conductors, and neglect end effects.

We calculate successively the magnetic energies per unit length of the line in regions 1, 2, 3, 4 as in Fig. 26-3, and then we set the sum equal to $L'I^2/2$, where L' is the inductance per meter. Lines of \mathbf{B} are circles centered on the axis.

(1) Region 1. From Ampère's circuital law, Secs. 19.5 and 20.6, at the radius p ,

$$2\pi p B_1 = \mu_0 I \frac{\pi p^2}{\pi a^2}, \quad (26-28)$$

and the magnetic energy per unit length is

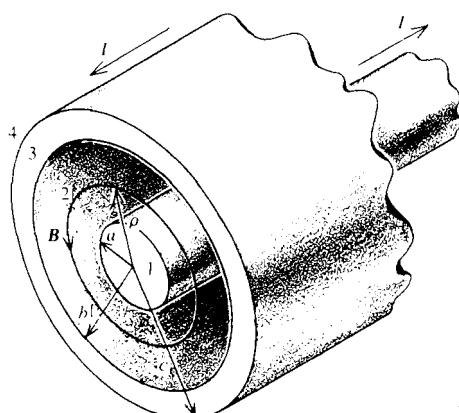


Fig. 26-3. Coaxial line.

$$\mathcal{E}'_{m1} = \frac{1}{2\mu_0} \int_0^a \left(\frac{\mu_0 I \rho}{2\pi a^2} \right)^2 2\pi \rho d\rho = \frac{\mu_0}{16\pi} I^2. \quad (26-29)$$

(2) Region 2. Here

$$B_2 = \frac{\mu_0 I}{2\pi \rho}, \quad \mathcal{E}'_{m2} = \frac{\mu_0 I^2}{4\pi} \ln \frac{b}{a}. \quad (26-30)$$

(3) Region 3. We require the net current that flows within a circular path of radius ρ . This is I minus that part of the current in the outer conductor that flows between radii b and ρ . Thus

$$B_3 = \frac{\mu_0 I}{2\pi \rho} \left(1 - \frac{\rho^2 - b^2}{c^2 - b^2} \right) = \frac{\mu_0 I}{2\pi \rho} \frac{c^2 - \rho^2}{c^2 - b^2}, \quad (26-31)$$

$$\mathcal{E}'_{m3} = \frac{\mu_0 I^2}{4\pi} \left[\frac{c^4}{(c^2 - b^2)^2} \ln \frac{c}{b} - \frac{3c^2 - b^2}{4(c^2 - b^2)} \right]. \quad (26-32)$$

(4) Region 4. From Ampère's circuital law, there is zero field in this region.

Finally, the self-inductance per unit length of the coaxial line is

$$L' = \mu_0 \left\{ \frac{1}{8\pi} + \frac{1}{2\pi} \ln \frac{b}{a} + \frac{1}{2\pi} \left[\frac{c^4}{(c^2 - b^2)^2} \ln \frac{c}{b} - \frac{3c^2 - b^2}{4(c^2 - b^2)} \right] \right\}. \quad (26-33)$$

The second term between the braces comes from \mathcal{E}'_2 and is normally the most important.

*26.5 THE FORCE BETWEEN TWO CURRENT-CARRYING CIRCUITS EXPRESSED IN TERMS OF M , I_a , AND I_b

We already found an integral for the force between two current-carrying circuits in Sec. 22.3. However, as we noted at the time, the integral is difficult to evaluate.

Here we express the force in terms of the mutual inductance between the two circuits. Mutual inductance is just as difficult to calculate as the force, but it is easy to measure, much more so than the force itself. In the process, we shall find that whenever one changes the geometry of a circuit or of a pair of circuits, precisely one-half of the energy furnished by the source, exclusive of Joule and other losses, becomes magnetic energy, and the other half becomes mechanical work.

We consider two circuits carrying currents I_a and I_b in the same direction, as in Fig. 26-4. The magnetic force is such that the loops tend

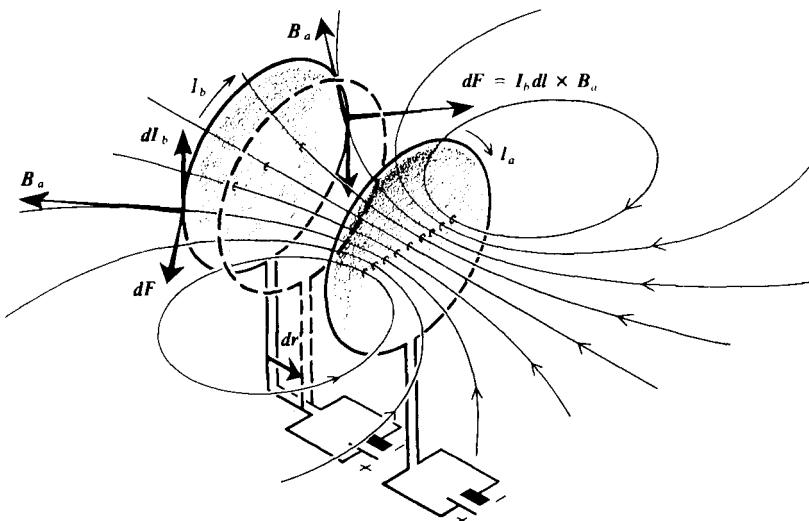


Fig. 26-4. Two parallel loops bearing currents I_a and I_b , with typical lines of \mathbf{B} originating in loop a . The element of force $d\mathbf{F}$ possesses a component in the direction of a , and \mathbf{F} is attractive.

to move toward each other. They are fixed in position by opposing mechanical forces. All materials are nonmagnetic.

Now assume a small virtual translation (Sec. 6.6) of one circuit, without any rotation. Since there is conservation of energy, the energy expended by the sources is equal to the increase in magnetic energy plus the mechanical work done. The displacement takes place slowly so as to avoid taking kinetic energy into account.

To simplify the calculation, we assume that the currents are constant. This assumption will not affect our result. We had a similar situation in electrostatics. See Prob. 6-12.

Loop b moves a distance dr toward loop a . Only M changes and, from Eq. 26-12, the magnetic energy increases by

$$d\mathcal{E}_m = I_a I_b dM = I_a d\Lambda_{ba} = I_b d\Lambda_{ab}, \quad (26-34)$$

Λ_{ba} being the flux originating in b and linking a , and similarly for Λ_{ab} . Since M is positive (Sec. 24.1.2) and increases, $d\mathcal{E}_m$ is positive.

Now consider the extra energy supplied by the sources. In loop b , the linking flux increases and the induced electromotance tends to generate a magnetic field that opposes this increase. Therefore the electromotance induced in b tends to oppose I_b . To keep that current constant, its source supplies the extra voltage $d\Lambda_{ab}/dt$ and the extra energy

$$d\mathcal{E}_{sb} = I_b \frac{d\Lambda_{ab}}{dt} dt = I_b d\Lambda_{ab} = I_b I_a dM. \quad (26-35)$$

By symmetry, $d\mathcal{E}_{sa}$ is the same, and the extra energy supplied by the two sources is

$$d\mathcal{E}_s = 2I_a I_b dM, \quad (26-36)$$

which is exactly twice the increase in magnetic energy. The remainder has gone into mechanical work. In other words, the energy supplied by the sources divides equally between mechanical energy and magnetic energy. See Sec. 6.6.1.

If \mathbf{F}_{ab} is the force that coil a exerts on coil b , then the mechanical work done is

$$\mathbf{F}_{ab} \cdot d\mathbf{r} = I_a I_b dM. \quad (26-37)$$

Since the quantity on the right is positive, \mathbf{F}_{ab} points toward coil a , like $d\mathbf{r}$, which is correct.

The x component of the force is

$$F_{abx} = I_a I_b \frac{\partial M}{\partial x}, \quad (26-38)$$

where dx is the x component of $d\mathbf{r}$. This equation applies to any pair of circuits.

This is an alternate expression for the force that we found in Sec. 22.3. We can show that the two expressions are equal as follows. Let circuit b move as a whole, without rotating, parallel to the x -axis. Then

$$F_{abx} = -I_a I_b \frac{\partial}{\partial x} \left(\frac{\mu_0}{4\pi} \oint_a \oint_b \frac{\mathbf{dl}_a \cdot \mathbf{dl}_b}{r} \right) \quad (26-39)$$

where r is the distance from $d\mathbf{l}_a$ to $d\mathbf{l}_b$. The derivative with respect to x acts only on the $1/r$ term under the integral sign because the vectors $d\mathbf{l}$ are not affected by a *translation* of circuit b . Thus

$$F_{abx} = -\frac{\mu_0}{4\pi} I_a I_b \oint_a \oint_b x \frac{\mathbf{dl}_a \cdot \mathbf{dl}_b}{r^3} \quad (26-40)$$

and, more generally,

$$\mathbf{F}_{ab} = -\frac{\mu_0}{4\pi} I_a I_b \int_a \int_b \mathbf{r} \frac{\mathbf{dl}_a \cdot \mathbf{dl}_b}{r^3} = -\frac{\mu_0}{4\pi} I_a I_b \int_a \int_b \hat{\mathbf{r}} \frac{\mathbf{dl}_a \cdot \mathbf{dl}_b}{r^2}, \quad (26-41)$$

as in Sec. 22.3.

*26.6 THE FORCE BETWEEN TWO CURRENT-CARRYING CIRCUITS EXPRESSED IN TERMS OF THE MAGNETIC ENERGY \mathcal{E}_m

Since the term on the right in Eq. 21-37 is the increase in magnetic energy, we could also write that

$$\mathbf{F}_{ab} \cdot d\mathbf{r} = d\mathcal{E}_m, \quad (26-42)$$

remembering that the force pulls in the direction that *increases* the magnetic energy. Also,

$$F_{abx} = \frac{\partial \mathcal{E}_m}{\partial x}. \quad (26-43)$$

Example

THE FORCE BETWEEN TWO LONG COAXIAL SOLENOIDS

Figure 26-5 shows two coaxial solenoids, one of which extends a distance l inside the other. The mutual inductance is positive. The net force is axial, and it is attractive, as can be seen from Fig. 26-6. Remember that the force between two parallel currents flowing in the same direction is attractive.

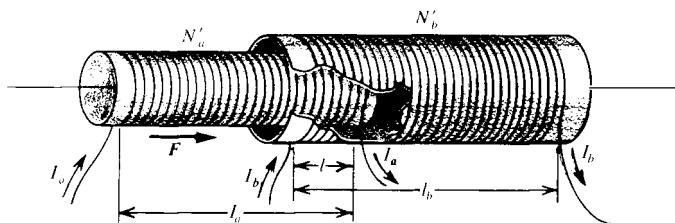


Fig. 26-5. Two coaxial solenoids of approximately equal diameters. The force \mathbf{F} is attractive when the currents flow in the direction shown. The solenoids have the same number of turns per meter N' .

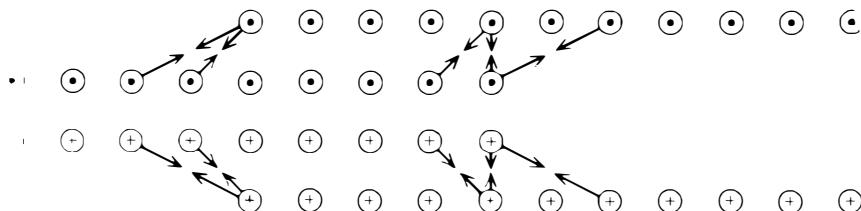


Fig. 26-6. Section through part of the solenoids of Fig. 26-5. There is a force of attraction at the ends of the solenoids.

We cannot perform a rigorous calculation of the force because it clearly depends on the very end effect that we have disregarded until now. However, we can find an approximate expression by applying the formulas that we found above.

We assume that the field of a solenoid stops abruptly at the end. Then

$$M = \frac{N'l\Phi_{ba}}{I_b} = \frac{N'l\Phi_{ab}}{I_a}. \quad (26-44)$$

For a cross-sectional area \mathcal{A} ,

$$M = \frac{(N'l)(\mu_0 N' I_b \mathcal{A})}{I_b} = \mu_0 N'^2 \mathcal{A} l, \quad (26-45)$$

$$F = I_a I_b \frac{\partial M}{\partial l} = \mu_0 N'^2 \mathcal{A} I_a I_b. \quad (26-46)$$

The force is attractive because M increases with l .

Now let us calculate the force from $\partial \mathcal{E}_m / \partial l$:

$$\mathcal{E}_m = \frac{1}{2\mu_0} \int_{\infty} B^2 dv = \frac{1}{2\mu_0} [B_a^2(l_a - l) + B_b^2(l_b - l) + (B_a + B_b)^2 l] \mathcal{A} \quad (26-47)$$

$$= \frac{1}{2\mu_0} (B_a^2 l_a + B_b^2 l_b + 2B_a B_b l) \mathcal{A}, \quad (26-48)$$

where $B_a = \mu_0 N' I_a$ originates in solenoid a , and similarly for b . Then

$$F = \frac{d\mathcal{E}_m}{dl} = \frac{B_a B_b \mathcal{A}}{\mu_0} = \mu_0 N'^2 \mathcal{A} I_a I_b. \quad (26-49)$$

Observe that the force would be zero if \mathcal{E}_m were proportional to B .

*26.7 MAGNETIC FORCES AND LINES OF \mathbf{B}

Lines of \mathbf{B} are useful for visualizing magnetic forces between current-carrying wires. Figure 26-7(a) shows lines of \mathbf{B} for two wires carrying equal currents flowing in the same direction. The lines of \mathbf{B} that cross the midplane are “under tension,” and the force of attraction per unit area in the midplane is $B^2/(2\mu_0)$. Figure 26-7(b) shows lines of \mathbf{B} for currents flowing in opposite directions. Now the lines of \mathbf{B} “repel laterally,” and the force of repulsion per unit area in the midplane is $B^2/(2\mu_0)$.

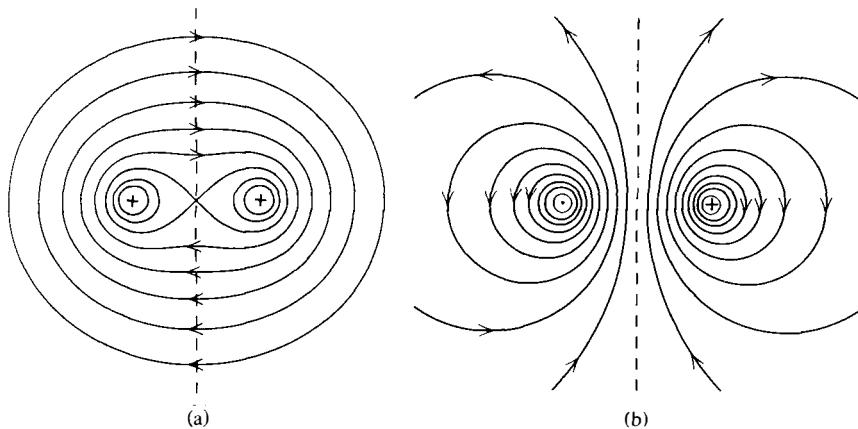


Fig. 26-7. (a) The \mathbf{B} field of two parallel currents flowing in the same direction. The magnetic flux between two successive lines of \mathbf{B} is constant. We have not shown the lines near the wires because they are too close together. (b) The \mathbf{B} field of two parallel currents flowing in opposite directions.

These lines of \mathbf{B} are identical to the equipotentials in Fig. 6-7 because, in a two-dimensional field, a line of \mathbf{B} is a line of constant A (Prob. 18-12), and both V and A vary as 1 over the distance from the wire.

This is a general rule: lines of \mathbf{B} are “under tension” and “repel laterally,” just like lines of \mathbf{E} . See the next section.

*26.8 MAGNETIC PRESSURE

If the current flows through a conducting sheet, then it is appropriate to think in terms of *magnetic pressure*. Imagine a conducting sheet, in air, carrying α amperes/meter and situated in a uniform tangential magnetic field $\mathbf{B}/2$ originating in currents flowing elsewhere, as in Fig. 26-8(a), with α normal to \mathbf{B} . The force per unit area is $\alpha\mathbf{B}/2$.

Now increase α until its field cancels the ambient field on one side and doubles it on the other, as in Fig. 26-8(b). Then, from Prob. 19-3, $\alpha' = B/\mu_0$, and the force per unit area, or the pressure, is

$$P = \frac{B^2}{2\mu_0} . \quad (26-50)$$

This applies to any current sheet with zero \mathbf{B} on one side.

The pressure is equal to the energy density, as with electric fields (Sec. 6.5). This pressure pushes the current sheet away from the field. On the field side, the lines of \mathbf{B} are parallel to the sheet and repel laterally.

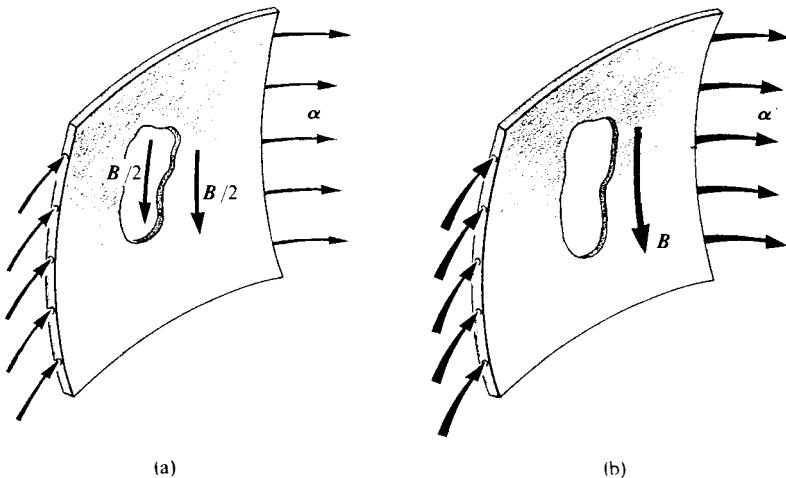


Fig. 26-8. (a) A conducting sheet carries a current density of α amperes/meter of width and lies in a uniform magnetic field $B/2$ originating elsewhere. (b) The current density α' is now such that the total B on the near side is twice what it was, while the total B on the far side is zero.

Example

THE MAGNETIC PRESSURE INSIDE A LONG SOLENOID

Figure 26-9 shows an end view of a solenoid. We use the method of virtual work (Sec. 6.6) to show that the magnetic pressure is $B^2/(2\mu_0)$. Imagine that the current remains constant, while the magnetic pressure increases the radius from R to $R + dR$. Then, for a solenoid of length l , the magnetic energy increases by

$$d\mathcal{E}_m = \frac{B_2}{2\mu_0} l 2\pi R dR. \quad (26-51)$$

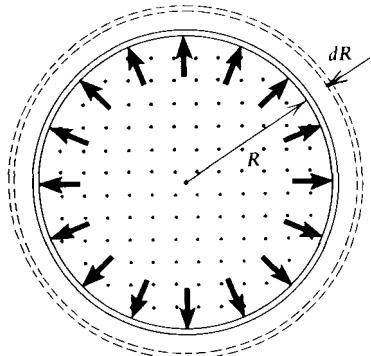


Fig. 26-9. End view of a solenoid. The dots represent lines of \mathbf{B} , normal to the paper. The arrows show how the magnetic pressure pushes the winding away from the field.

This is equal to the mechanical work performed by the magnetic pressure p over an area $2\pi Rl dR$, and

$$p = \frac{B^2}{2\mu_0}. \quad (26-52)$$

*26.9 MAGNETIC FORCES WITHIN AN ISOLATED CIRCUIT

Within a single isolated circuit one still has magnetic forces, because the current in one part flows in the magnetic field of the rest of the circuit. For example, if the circuit is a simple loop, then the magnetic force on the wire tends to expand the loop.

26.10 MAGNETIC TORQUE

By analogy with Sec. 26.5, a circuit a that forms an angle θ with another circuit b exerts on b a torque

$$T = I_a I_b \frac{\partial M}{\partial \theta} = \frac{\partial \mathcal{E}_m}{\partial \theta}. \quad (26-53)$$

The torque tends to *increase* both the mutual inductance M and the magnetic energy \mathcal{E}_m .

Example

THE MAGNETIC TORQUE EXERTED ON A CURRENT LOOP

A rectangular loop of wire carrying a current I lies in a uniform \mathbf{B} in air, as in Fig. 26-10. We calculate the torque at the angle θ .

(1) The simplest procedure here is to calculate the torque from the magnetic force $I\mathbf{dl} \times \mathbf{B}$ on the element $d\mathbf{l}$ of the wire, as in Sec. 22.2:

$$T = 2a \sin \theta BIb = BI\mathcal{A} \sin \theta, \quad (26-54)$$

where \mathcal{A} is the area of the loop. The torque is in the direction of the curved arrow shown in the figure.

(2) Now let us use Eq. 26-53. Let I_a be the unknown current that provides the field \mathbf{B} . Then

$$M = -\frac{B\mathcal{A} \cos \theta}{I_a}. \quad (26-55)$$

We require a negative sign here because the flux $B\mathcal{A} \cos \theta$ links

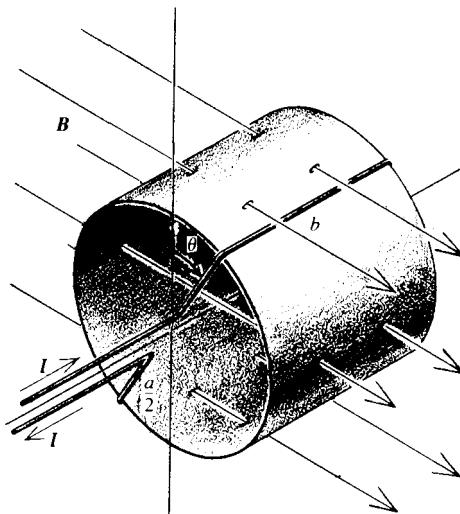


Fig. 26-10. Loop carrying a current I in a constant \mathbf{B} . The magnetic torque tends to turn the loop in the direction of the arrow.

the loop in the direction opposite to that of the flux of a current I flowing in the loop as in the figure. Then

$$T = I_a I_b \frac{\partial M}{\partial \theta} = -I \frac{d}{d\theta} (B \mathcal{A} \cos \theta) = BI \mathcal{A} \sin \theta. \quad (26-56)$$

The positive sign means that the torque is in the direction shown, as we found above. The torque tends to increase θ and hence to increase M .

26.11 SUMMARY

When a current I flows in a circuit of self-inductance L , the *magnetic energy* stored in the field can be expressed in various ways:

$$\mathcal{E}_m = \frac{1}{2} LI^2 = \frac{1}{2} I \Lambda = \frac{1}{2} \int_{\infty} \mathbf{H} \cdot \mathbf{B} dv, \quad (26-9), (26-23)$$

where Λ is the magnetic flux linkage and where the integration runs over all space. Also, for a finite current distribution,

$$\mathcal{E}_m = \frac{1}{2} \int_v \mathbf{J} \cdot \mathbf{A} dv, \quad (26-18)$$

where v is the volume occupied by the current.

Thus the *energy density* can be taken to be

$$\mathcal{E}' = \frac{1}{2} \mathbf{H} \cdot \mathbf{B} = \frac{B^2}{2\mu} \quad (26-24)$$

We define the *self-inductance* of a real circuit comprising currents distributed over a finite volume in terms of the magnetic energy stored in the field:

$$L = \frac{1}{\mu_0 I^2} \int_v B^2 dv. \quad (26-27)$$

The x -component of the *force* exerted on a circuit b situated in the field of another circuit a is

$$F_{abx} = I_a I_b \frac{\partial M}{\partial x} = \frac{\partial \mathcal{E}_m}{\partial x}. \quad (26-38), (26-43)$$

The force pulls in the direction that increases both M and the magnetic energy.

Similarly, the *torque* is given by

$$T = I_a I_b \frac{\partial M}{\partial \theta} = \frac{\partial \mathcal{E}_m}{\partial \theta}. \quad (26-53)$$

If the current flowing through a conducting sheet situated in a magnetic field is such that the magnetic flux density is \mathbf{B} on one side and zero on the other, then the *magnetic pressure* on the sheet is $B^2/(2\mu_0)$. The pressure tends to push the sheet away from the field.

PROBLEMS

26-1. (26.1) The energy stored in a magnetic circuit

Show that the energy stored in a magnetic circuit is $\Phi^2 \mathcal{R}/2$, where Φ is the magnetic flux and \mathcal{R} is the reluctance.

26-2. (26.1) The energy theorem for linear passive circuits

(a) This theorem follows from Tellegen's theorem (Sec. 8.6). Suppose one has a passive circuit comprising resistances, self-inductances, and capacitances. One applies an alternating voltage V_p to an input port.

Show that

$$V_p I_p^* = P + 2j\omega (\mathcal{E}_{\text{mag,av}} - \mathcal{E}_{\text{el,av}}),$$

where the left-hand side is the input complex power, P is the power dissipated in the circuit, $\mathcal{E}_{\text{mag,av}}$ is the average magnetic stored energy, and

$\mathcal{E}_{el,av}$ is the average electric stored energy. This is the *energy theorem*.

(b) It is shown in Prob. 25-4 that the star and the delta of Fig. 25-19 are equivalent. Now, if one applies an alternating voltage V (rms) between terminals B and C of the star, then VI^* is real and equal to $V^2/2000$. According to the energy theorem, the average energy stored in the capacitors of the delta must be equal to the average energy stored in the inductor, at any frequency. Show that this is correct.

- 26-3. (26.1) The average stored energies in capacitors and in inductors

The average stored energies in capacitors and in inductors are $CV^2/2$ and $LI^2/2$, respectively, where V and I are rms values.

Show that

$$\mathcal{E}_{C,av} = \frac{I^2}{2\omega^2 C}, \quad \mathcal{E}_{L,av} = \frac{V^2}{2\omega^2 L}.$$

- 26-4. (26.2) The magnetic energy in terms of \mathbf{J} and \mathbf{A}

We saw that

$$\mathcal{E}_m = \frac{1}{2} \int_v \mathbf{J} \cdot \mathbf{A} dv,$$

where v' is any volume that encloses all the conductors.

Show that any \mathbf{A} such that $\mathbf{B} = \nabla \times \mathbf{A}$ is satisfactory here.

- 26-5. (26.4) The inductance of a coaxial line is slightly frequency-dependent

High-frequency currents do not penetrate a conductor as do low-frequency currents. This is the skin effect (Sec. 29.1). Does the self-inductance of a coaxial line increase or decrease with frequency?

- 26-6. (26.5) The electromagnetic levitation of high-speed tracked vehicles

The suspension and the propulsion of tracked vehicles become major problems at speeds of several hundred kilometers per hour. Wheels are then impractical because vehicle vibration, track damage, and power loss become excessive. The tractive force also deteriorates with increasing speed.

An air cushion provides a satisfactory suspension at high speeds, but it consumes a large amount of power. Propulsion then requires either a propeller or a linear electric motor, with the stator in the track.

It is also possible to support a vehicle by means of magnetic forces, and several methods have been developed. In one of these, superconducting coils in the vehicle generate a magnetic field that extends down into the track, which is a sheet of aluminum. At rest and at low speeds, the vehicle uses wheels. As the speed increases, the eddy currents induced in the track by the traveling magnetic field exert a force of repulsion on the currents in the vehicle coils, and the vehicle flies about 10 centimeters above the track. There are, of course, problems of stability. Also, the suspension is not lossless because there are Joule losses in the track.

Let us consider a simplified form of levitation. A pair of parallel and coaxial coils of radius R and N turns are separated by a distance D . The lower coil simulates the track. For $D \approx 0.1R$, the mutual inductance is given by $N^2\{2.154 - 12.04[(D/R) - 0.1]\}R$ microhenrys.

(a) Calculate the number of ampere-turns required in each coil to support a mass of 1 metric ton when $R = 1$ meter.

(b) Draw a sketch showing the two coils and lines of \mathbf{B} . Can you explain the force of repulsion qualitatively?

26-7. (26.7) The force between two parallel bus bars of finite cross section.

Two parallel bus bars have equal circular cross sections and carry equal currents I . The currents are equally distributed over the cross sections.

Show, without any calculation, that the force is the same as if the bus bars were thin wires.

26-8. (26.7) A superconducting power transmission line

A superconducting dc power transmission line has been proposed that would carry 100 gigawatts of power at 200 kilovolts over 1000 kilometers. The conductors would have a diameter of 25 millimeters and be separated by a center-to-center distance of 50 millimeters.

(a) Calculate the magnetic force per meter. See the previous problem. It is clearly preferable to use a coaxial line.

(b) Calculate the stored energy in kilowatthours. The self-inductance per meter is $[\mu_0/(4\pi)][1 + 4 \ln(D/R)]$.

26-9. (26.8) Large-scale energy storage in inductors and in capacitors

Much work has been done on the large-scale storage of energy in inductors, for public utilities. One author proposes a huge, underground, cryogenized inductor that would operate at a field of 14 teslas.

(a) Calculate the energy density in kilowatthours/meter³.

(b) Calculate the magnetic pressure in atmospheres.

(c) It seems more reasonable to store energy in a capacitor, because a capacitor need not be cryogenized and because the force points inward, not outward as in an inductor. Calculate the energy density in kilowatthours/meter³ if $\epsilon_0 = 3$ and the dielectric strength is 1.5×10^8 volts/meter.

Gasoline can store over 100 kilowatthours/meter³, and flywheels over 200.

26-10. (26.9) The mechanical work performed by mechanical forces on an isolated, active, and deformable circuit

We showed in Sec. 26.5 that, if one active circuit moves with respect to another, the mechanical work performed by the sources is equal to the increase in magnetic energy if the currents are maintained constant. Hence the force between two active circuits is given by the rate of increase of magnetic energy.

Show that, similarly, if the geometry of an isolated active circuit changes, the energy supplied by the sources divides in the same way. Assume again that the current is constant. It follows that, on this assumption, the force on an element of an active circuit is equal to the rate of increase of magnetic energy.

26-11. (26.9) The axial compression force on a solenoid

(a) Show qualitatively, in two different ways, that the turns of a solenoid tend to squeeze together.

(b) Calculate the axial compression force on a long solenoid.

26-12. (26.9) Magnetic shutter

Magnetic fields can perform mechanical tasks that require a high power level for a very short time. For example, magnetic pressure can crush a light aluminum tube that acts as a shutter to turn off a beam of light or of soft x-rays. The tube is placed inside a coil, parallel to the axis. When the coil is suddenly connected to a large capacitor, the change in flux induces a large current in the tube, which collapses under the magnetic pressure.

Let us calculate the pressure. If the current I in the solenoid increases gradually from zero to some large value, the induced current is small and the magnetic pressure is negligible. Let us assume that dI/dt in the coil is so large that the induced current in the tube maintains zero magnetic field inside it. Then there is a magnetic field B only in the annular region between the solenoid and the conducting tube.

(a) Calculate the pressure on the tube in atmospheres at 1 tesla.

(b) What would be the pressure if the conducting tube were parallel to the axis but off the axis?

26-13. (26.9) Flux compression

Flux compression is one method of obtaining large magnetic fields. For example, one can insert a light conducting tube in the field B_0 of a solenoid and then implode the tube by means of an annular explosive charge situated between the tube and the solenoid. Currents flow in the tube, and the magnetic pressure builds up until it is equal to the external gas pressure. The solenoid is fed by a constant-current source.

(a) Show that, if the radius of the tube shrinks very rapidly, the B inside is about $B_0(R_0^2/R^2)$ at the instant when the radius is equal to R . For example, if B_0 is 10 teslas and if $R_0/R = 10$, then $B = 10^3$ teslas.

(b) Calculate the surface current density in the tube in amperes/meter.

(c) Calculate the change in magnetic energy, the energy absorbed by the constant-current source feeding the solenoid, and the explosive energy required to compress the field. Assume that the tube is 200 millimeters long, $R_0 = 50$ millimeters, and neglect end effects.

26-14. (26.10) The torque on a current-carrying coil

(a) Show that a current-carrying coil tends to orient itself in a magnetic field in such a way that the total magnetic flux linking the coil is maximum.

(b) Show that the torque on the coil is $\mathbf{m} \times \mathbf{B}$, where \mathbf{m} is the magnetic moment of the coil and \mathbf{B} is the magnetic flux density when the current in the coil is zero.

26-15. (26.10) The torque on a cylindrical permanent magnet

Show that the torque exerted on a small, cylindrical permanent magnet of dipole moment \mathbf{m} situated in a magnetic field is $\mathbf{m} \times \mathbf{B}$. See the preceding problem.

26-16. (26.10) High-gradient magnetic separation

It is possible to separate magnetic particles in suspension in a fluid by passing the mixture through steel wool subjected to a strong magnetic

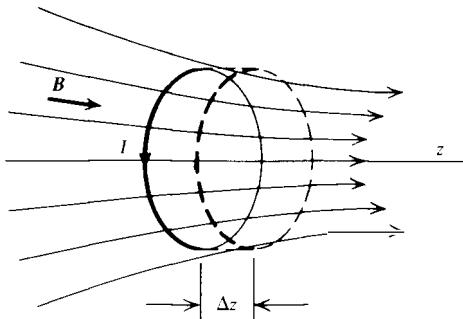


Fig. 26-11.

field. The magnetic particles cling to the steel wires where the field gradient is large. Arrays of fine steel wires normal to \mathbf{B} are also used.

With a field of the order of several teslas supplied by superconducting coils, the separation occurs even with materials that are only slightly magnetic. The method is also applicable in air for removing magnetic particles, say in pulverized coal.

Let us see how a small magnetic dipole behaves in a nonuniform \mathbf{B} . The dipole first orients itself. Then, as we shall see, it tends to move in the direction in which the applied B increases.

Figure 26-11 shows a small current loop of radius R that is already oriented in a field \mathbf{B} that increases symmetrically about the positive direction of the z -axis.

- Show, without any calculation, that the magnetic force points to the right. Note that this force tends to *increase* the linking flux.
- Show that $F = 2\pi RIB_\rho$, where B_ρ is the component of \mathbf{B} that is normal to the z -axis.
- Now consider a small volume of thickness Δz , as in the figure. Use the fact that the net outward flux of \mathbf{B} is zero to find B_ρ and F .
- Calculate the force from the rate of increase of magnetic energy.

CHAPTER 27

MAXWELL'S EQUATIONS

27.1 MAXWELL'S EQUATIONS IN DIFFERENTIAL FORM	492
27.2 MAXWELL'S EQUATIONS IN INTEGRAL FORM	495
27.3 FURTHER COMMENTS ON MAXWELL'S EQUATIONS	498
27.4 THE LAW OF CONSERVATION OF CHARGE	500
27.5 MAXWELL'S EQUATIONS ARE REDUNDANT	500
27.6 DUALITY	501
<i>Example: THE FIELDS OF ELECTRIC AND MAGNETIC DIPOLES</i>	501
27.7 LORENTZ'S RECIPROCITY THEOREM	502
<i>Examples</i>	503
27.8 THE WAVE EQUATIONS FOR E AND FOR B	504
27.9 SUMMARY	505
PROBLEMS	507

This chapter concerns the four fundamental equations of electromagnetism that bear the name of James Clerk Maxwell (1831–1879). He was the first to state them clearly and to recognize their importance, but it was Oliver Heaviside (1850–1925) who first expressed them in the form that we know today. These equations are so fundamental that we shall express them in various ways and then discuss their physical meaning.

27.1 MAXWELL'S EQUATIONS IN DIFFERENTIAL FORM

Let us group Maxwell's four equations; we discuss them at length below. We found them successively in Secs. 9.5, 23.4, 20.4, and 17.4:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad (27-1)$$

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, \quad (27-2)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (27-3)$$

$$\nabla \times \mathbf{B} - \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{J}.^+ \quad (27-4)$$

[†] If you have not studied Chap. 17, you will have to take this equation for granted at this stage. You will find a proof in Sec. 28.6.

The above equations are general in that the media can be *nonhomogeneous*, *nonlinear*, and *nonisotropic*. However, (1) they apply only to media that are stationary with respect to the coordinate axes,[†] and (2) the coordinate axes must not accelerate and must not rotate.

These are the four fundamental equations of electromagnetism. They form a set of simultaneous partial differential equations relating certain time and space derivatives at a point to the charge and current densities at that point. They apply, whatever be the number or diversity of the sources.

We have followed the usual custom of writing the field terms on the left and the source terms on the right. However, this is somewhat illusory because ρ and \mathbf{J} are themselves functions of \mathbf{E} and \mathbf{B} . As usual,

\mathbf{E} is the electric field strength, in volts/meter;

$\rho = \rho_f + \rho_b$ is the total electric charge density, in coulombs/meter³;

ρ_f is the free charge density;

$\rho_b = -\nabla \cdot \mathbf{P}$ is the bound charge density;

\mathbf{P} is the electric polarization, in coulombs/meter²;

\mathbf{B} is the magnetic flux density, in teslas;

$\mathbf{J} = \mathbf{J}_f + \partial \mathbf{P} / \partial t + \nabla \times \mathbf{M}$ is the total current density, in amperes/meter²;[‡]

\mathbf{J}_f is the current density resulting from the motion of free charge;

$\partial \mathbf{P} / \partial t$ is the polarization current density in a dielectric;

$\nabla \times \mathbf{M}$ is the equivalent current density in magnetized matter;

\mathbf{M} is the magnetization, in amperes/meter;

c is the speed of light, about 300 megameters per second;

ϵ_0 is the permittivity of free space, about 8.85×10^{-12} farad/meter.

In isotropic, linear, and stationary media,

$$\mathbf{J}_f = \sigma \mathbf{E}, \quad \mathbf{P} = \epsilon_0 \chi_e \mathbf{E}, \quad \mathbf{M} = \chi_m \mathbf{H}, \quad (27-5)$$

where σ is the conductivity, χ_e is the electric susceptibility, and χ_m is the magnetic susceptibility. Also,

$$\epsilon_r = 1 + \chi_e, \quad \mu_r = 1 + \chi_m, \quad (27-6)$$

[†] See Paul Penfield, Jr., and Herman A. Haus, *Electrodynamics of Moving Media*, Research Monograph 40, M.I.T. Press, Cambridge, Mass., 1967.

[‡] Until now we were concerned solely with free current densities, and we used \mathbf{J} instead of \mathbf{J}_f , to simplify the notation.

where ϵ_r is the relative permittivity and μ_r is the relative permeability. Inside a source, such as a battery or a Van de Graaff generator, electric charges are “pumped” by the locally generated electric field \mathbf{E}_s , against the electric field \mathbf{E} of other sources, and $\mathbf{J} = \sigma(\mathbf{E} + \mathbf{E}_s)$.

Writing out ρ and \mathbf{J} in full, Maxwell's equations become

$$\nabla \cdot \mathbf{E} = \frac{\rho_f - \nabla \cdot \mathbf{P}}{\epsilon_0}, \quad (27-7)$$

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, \quad (27-8)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (27-9)$$

$$\nabla \times \mathbf{B} - \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \left(\mathbf{J}_f + \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M} \right). \quad (27-10)$$

This *Amperian formulation* expresses the field in terms of the four vectors \mathbf{E} , \mathbf{B} , \mathbf{P} , and \mathbf{M} .

With homogeneous, isotropic, linear, and stationary (HILS) media,

$$\rho = \frac{\rho_f}{\epsilon_r} \quad (\text{Sec. 9.9}) \quad (27-11)$$

$$\mathbf{P} = (\epsilon_r - 1)\epsilon_0 \mathbf{E} \quad (\text{Sec. 9.9}) \quad (27-12)$$

$$\mathbf{M} = \frac{(\mu_r - 1)}{\mu_r \mu_0} \mathbf{B} \quad (\text{Sec. 20.7}) \quad (27-13)$$

and

$$\nabla \cdot \mathbf{E} = \frac{\rho_f}{\epsilon}, \quad (27-14) \qquad \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, \quad (27-15)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (27-16) \qquad \nabla \times \mathbf{B} - \epsilon \mu \frac{\partial \mathbf{E}}{\partial t} = \mu \mathbf{J}_f. \quad (27-17)$$

Recall that $\epsilon = \epsilon_r \epsilon_0$ and $\mu = \mu_r \mu_0$, ϵ_r is frequency-dependent, and μ_r is hardly definable in ferromagnetic materials. The expressions for \mathbf{P} and for \mathbf{M} are not symmetrical, but \mathbf{P} , \mathbf{E} , and \mathbf{D} point in the same direction, like \mathbf{M} , \mathbf{H} , and \mathbf{B} , in isotropic and linear media.

Observe that the above set of equations follows from Eqs. 27-1 to 27-4 with the following substitutions:

$$\epsilon_0 \rightarrow \epsilon, \quad \mu_0 \rightarrow \mu, \quad (27-18)$$

$$\rho \rightarrow \rho_f, \quad \mathbf{J} \rightarrow \mathbf{J}_f. \quad (27-19)$$

This is a general rule for transforming an equation in terms of ϵ_0 , μ_0 , ρ , \mathbf{J}_f to another one in terms of ϵ , μ , ρ_f , \mathbf{J}_f .

The *Minkowski formulation* of Maxwell's equations is often useful. It expresses the same relations, but in terms of the four vectors \mathbf{E} , \mathbf{D} , \mathbf{B} , \mathbf{H} :

$$\nabla \cdot \mathbf{D} = \rho_f, \quad (27-20) \qquad \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, \quad (27-21)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (27-22) \qquad \nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J}_f. \quad (27-23)$$

In the following chapters we shall be mostly concerned with electric and magnetic fields that are sinusoidal functions of the time. Then, for isotropic, linear, and stationary media, not necessarily homogeneous,

$$\nabla \cdot \epsilon \mathbf{E} = \rho_f, \quad (27-24) \qquad \nabla \times \mathbf{E} + j\omega \mu \mathbf{H} = 0, \quad (27-25)$$

$$\nabla \cdot \mu \mathbf{H} = 0, \quad (27-26) \qquad \nabla \times \mathbf{H} - j\omega \epsilon \mathbf{E} = \mathbf{J}_f. \quad (27-27)$$

It is worthwhile to discuss Maxwell's equations further, but first let us rewrite them in integral form.

27.2 MAXWELL'S EQUATIONS IN INTEGRAL FORM

Integrating Eq. 27-1 over a finite volume v and then applying the divergence theorem, we find the integral form of Gauss's law (Sec. 9.5):

$$\int_{\mathcal{A}} \mathbf{E} \cdot d\mathcal{A} = \frac{1}{\epsilon_0} \int_v \rho dv = \frac{Q}{\epsilon_0}, \quad (27-28)$$

where \mathcal{A} is the area of the surface bounding the volume v and Q is the total charge enclosed within v . See Fig. 27-1.

Similarly, Eq. 27-3 says that the net outward flux of \mathbf{B} through any closed surface is zero, as in Fig. 27-2:

$$\int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} = 0. \quad (27-29)$$

Equation 27-2 is the differential form of the Faraday induction law for time-dependent magnetic fields. Integrating over an open surface of area \mathcal{A} bounded by a curve C gives the integral form, as in Sec. 23.4:

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{d}{dt} \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} = -\frac{d\Lambda}{dt}, \quad (27-30)$$

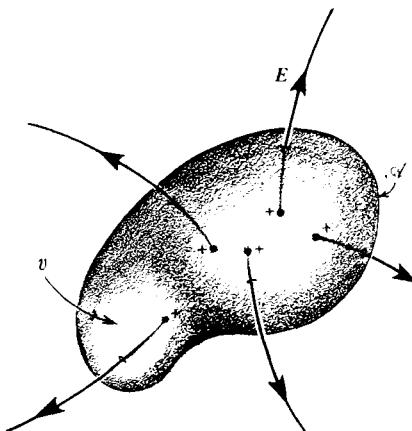


Fig. 27-1. Lines of \mathbf{E} emerging from a volume containing a net charge Q . The outward flux of \mathbf{E} is equal to Q/ϵ_0 .

where Λ is the linking flux. See Fig. 27-3. The electromotance induced around a closed curve C is equal to minus the time derivative of the flux linkage. The positive directions for Λ and around C satisfy the right-hand screw convention.

Finally, Eq. 27-4 is Ampère's circuital law in integral form:

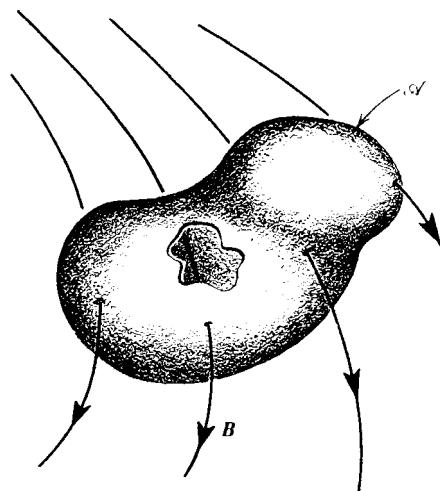


Fig. 27-2. Lines of \mathbf{B} passing through a closed surface. The net outward flux of \mathbf{B} is equal to zero.

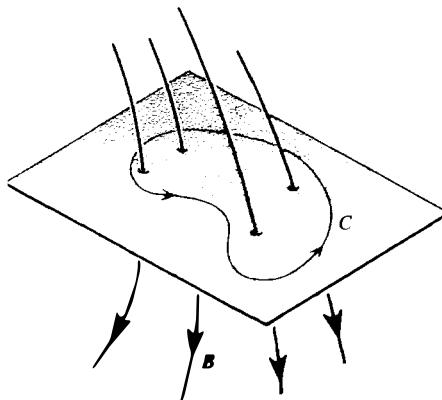


Fig. 27-3. If the magnetic flux linking C increases, it induces an electromotance around C in the direction of the arrow. The electromotance points in the same direction if \mathbf{B} points upward and decreases.

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_0 \int_{\mathcal{A}} \left(\mathbf{J} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \cdot d\mathcal{A}. \quad (27-31)$$

We found two less general forms of this law in Secs. 19.5 and 20.6. The closed curve C bounds a surface of area \mathcal{A} through which flows a current of density $\mathbf{J} + \epsilon_0 \partial \mathbf{E} / \partial t$. See Fig. 27-4.

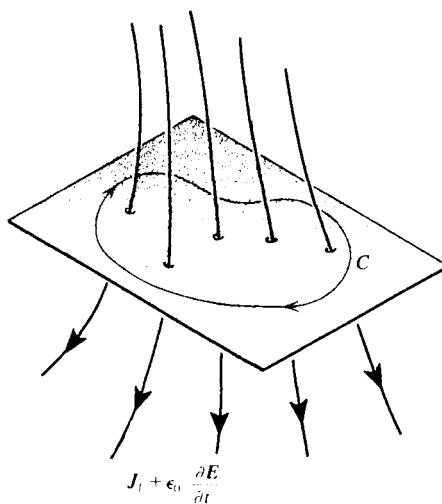


Fig. 27-4. The line integral of $\mathbf{B} \cdot d\mathbf{l}$ around C is positive if the integration runs in the direction of the arrows with a current density $\mathbf{J} + \epsilon_0 \partial \mathbf{E} / \partial t$ pointing downward.

27.3 FURTHER COMMENTS ON MAXWELL'S EQUATIONS

(1) Maxwell's equations 27-1 to 27-4 are linear. This means that they do not contain products of two or more of the variables or of their derivatives. It follows that if the field \mathbf{E}_1 , \mathbf{B}_1 satisfies Maxwell's equations when $\rho = \rho_1$ and $\mathbf{J} = \mathbf{J}_1$, and if the field \mathbf{E}_2 , \mathbf{B}_2 , corresponds similarly to $\rho = \rho_2$ and $\mathbf{J} = \mathbf{J}_2$, then the total field is $\mathbf{E}_1 + \mathbf{E}_2$, $\mathbf{B}_1 + \mathbf{B}_2$ when $\rho = \rho_1 + \rho_2$ and $\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2$. This is the principle of superposition (Secs. 3.3 and 18.2.1). In other words, each source of ρ and \mathbf{J} acts independently of all the others.

In linear media, \mathbf{P} and \mathbf{M} are proportional, respectively, to \mathbf{E} and \mathbf{B} , and Eqs. 27-7 to 27-10 are also linear. The principle of superposition applies.

However, in nonlinear media, \mathbf{P} and \mathbf{M} are more or less complicated functions of \mathbf{E} and \mathbf{B} . Equations 27-7 to 27-10 are then nonlinear and the principle of superposition does not apply to ρ_f and \mathbf{J}_f . The principle of superposition continues to apply to ρ and \mathbf{J} .

All materials become nonlinear at very high field strengths.

(2) We deduced Gauss's law for electric fields

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad \text{or} \quad \int_{\mathcal{A}} \mathbf{E} \cdot d\mathcal{A} = \frac{Q}{\epsilon_0}, \quad (27-32)$$

from Coulomb's law for stationary electric charges situated in a vacuum in Sec. 3.7. In Sec. 9.5 we found that the same relation applies to charges lying inside matter, if ρ and Q include both free and bound charges. Nonetheless, Gauss's law is more fundamental than Coulomb's law because it applies even to moving charges (Sec. 17.1), while Coulomb's law is strictly valid only for stationary charges (Secs. 3.1 and 16.5).

(3) One of Maxwell's equations states that

$$\nabla \cdot \mathbf{B} = 0, \quad \text{or} \quad \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} = 0, \quad (27-33)$$

where \mathcal{A} is the area of any closed surface. We deduced this relation in Sec. 18.3 from the law of Biot-Savart for the \mathbf{B} field of a time-independent current distribution. This relation is, in fact, general (Sec. 17.2), but on one condition: the density of magnetic monopoles (Sec. 18.1) must be zero. Since magnetic monopoles have never been observed to date, we may assume that the above equations are valid for any \mathbf{E} , \mathbf{B} field and for any current distribution, even if these are time-dependent.

(4) We found Faraday's law

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \text{or} \quad \oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{d}{dt} \int_{\mathcal{A}} \mathbf{B} \cdot d\mathbf{A}, \quad (27-34)$$

in Sec. 17.3, and we discussed it in Sec. 23.4. These equations are equally general. Here the closed curve C must be fixed in the coordinate system with respect to which both \mathbf{E} and \mathbf{B} are measured. However, the open surface bounded by C may move.

(5) In Sec. 20.5 we found that, for static fields,

$$\nabla \times \mathbf{H} = \mathbf{J}_f. \quad (27-35)$$

This equation is not general, for the following reason. Take the divergence of both sides. Then the divergence of the left-hand side is zero because the divergence of a curl is identically equal to zero. However, the divergence of the right-hand side is not necessarily zero:

$$\nabla \cdot \mathbf{J}_f = -\frac{\partial \rho_f}{\partial t} = -\frac{\partial}{\partial t}(\nabla \cdot \mathbf{D}) = -\nabla \cdot \left(\frac{\partial \mathbf{D}}{\partial t} \right). \quad (27-36)$$

Here we used first the conservation of free electric charge (Sec. 4.2) and then Eq. 27-20.

It therefore occurred to Maxwell that the correct relation must be

$$\nabla \times \mathbf{H} = \mathbf{J}_f + \frac{\partial \mathbf{D}}{\partial t}, \quad (27-37)$$

as above.

(6) Observe that the line integral of $\mathbf{B} \cdot d\mathbf{l}$ is related to the current density \mathbf{J} plus $\epsilon_0 \partial \mathbf{E} / \partial t$ in Eq. 27-31. This is remarkable because the integral for \mathbf{B} (Sec. 20.3) does *not* involve $\epsilon_0 \partial \mathbf{E} / \partial t$.[†]

(7) The displacement current density $\partial \mathbf{D} / \partial t$ of Sec. 9.10 is the sum of two terms:

$$\frac{\partial \mathbf{D}}{\partial t} = \frac{\partial}{\partial t}(\epsilon_0 \mathbf{E} + \mathbf{P}) = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \frac{\partial \mathbf{P}}{\partial t}. \quad (27-38)$$

In a material of conductivity σ and relative permittivity ϵ , subjected to an alternating electric field,

[†]The magnetic field associated with the displacement current in a parallel-plate capacitor has been observed. See D. F. Bartlett and P. R. Corle, *Physical Review Letters*, vol. 55, p. 59 (1985).

$$\mathbf{J}_f = \sigma \mathbf{E}_m \exp j\omega t, \quad \frac{\partial \mathbf{D}}{\partial t} = j\omega \epsilon_r \epsilon_0 \mathbf{E}_m \exp j\omega t, \quad (27-39)$$

$$\left| \frac{\partial D / \partial t}{J_f} \right| = \frac{\omega \epsilon_r \epsilon_0}{\sigma}. \quad (27-40)$$

In a good conductor, and at frequencies lower than roughly 1 gigahertz, $\sigma \approx 10^7$ siemens/meter, $\epsilon_r \approx 1$ (Sec. 9.9), the above ratio is of the order of $f/10^{17}$ at a frequency f , and the displacement current is negligible compared to the conduction current. At higher frequencies both ϵ_r and σ vary erratically with frequency because of atomic and molecular resonances.

27.4 THE LAW OF CONSERVATION OF CHARGE

In Sec. 4.2 we saw that free charges are conserved. At that time we were using the symbol \mathbf{J} for the current density of free charges instead of \mathbf{J}_f .

Let us calculate the divergence of \mathbf{J} as defined in Sec. 27.1. We will need the value of this divergence in the next section. First,

$$\nabla \cdot \mathbf{J} = \nabla \cdot \left(\mathbf{J}_f + \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M} \right) = \nabla \cdot \mathbf{J}_f + \frac{\partial}{\partial t} (\nabla \cdot \mathbf{P}), \quad (27-41)$$

the divergence of a curl being equal to zero. Thus

$$\nabla \cdot \mathbf{J} = - \frac{\partial \rho_f}{\partial t} - \frac{\partial \rho_b}{\partial t} = - \frac{\partial (\rho_f + \rho_b)}{\partial t} = - \frac{\partial \rho}{\partial t}. \quad (27-42)$$

This is a more general form of the law of conservation of charge of Sec. 4.2.

27.5 MAXWELL'S EQUATIONS ARE REDUNDANT

Maxwell's four equations are redundant. We saw in Secs. 17.3 and 17.4 that the equation for $\nabla \times \mathbf{E}$ follows from the one for $\nabla \cdot \mathbf{B}$, and the equation for $\nabla \times \mathbf{B}$ from the one for $\nabla \cdot \mathbf{E}$. These are, respectively, the first and second pairs.

The two equations of the first pair are also related as follows. If we take the divergence of Eq. 27-2 and remember that the divergence of a curl is zero, we find that

$$\nabla \cdot \frac{\partial \mathbf{B}}{\partial t} = 0 \quad \text{or} \quad \frac{\partial}{\partial t} \nabla \cdot \mathbf{B} = 0. \quad (27-43)$$

So $\nabla \cdot \mathbf{B}$ is a constant at every point in space. Then we can set $\nabla \cdot \mathbf{B} = 0$ everywhere and for all time if we assume that, for each point in space, $\nabla \cdot \mathbf{B}$ is zero at some time, in the past, at present, or in the future. With this assumption, Eq. 27-3 follows from Eq. 27-2.

Similarly, taking the divergence of Eq. 27-4 and applying the law of conservation of charge, we find

$$\epsilon_0 \nabla \cdot \frac{\partial \mathbf{E}}{\partial t} = -\nabla \cdot \mathbf{J} = \frac{\partial \rho}{\partial t}, \quad (27-44)$$

$$\frac{\partial}{\partial t} (\nabla \cdot \mathbf{E}) = \frac{\partial}{\partial t} \left(\frac{\rho}{\epsilon_0} \right), \quad \nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} + C. \quad (27-45)$$

The constant of integration C can be a function of the coordinates.

If we now assume that, at every point in space, at some time, $\nabla \cdot \mathbf{E}$ and ρ are simultaneously equal to zero, then C is zero and we have Eq. 27-1.

So there are really only two independent equations.

27.6 DUALITY

Imagine a field \mathbf{E} , \mathbf{B} that satisfies Maxwell's equations with $\rho_f = 0$, $\mathbf{J}_f = 0$ in a given region. The medium is homogeneous, isotropic, linear, and stationary (HILS). Now imagine a different field

$$\mathbf{E}' = -K\mathbf{B} = -K\mu\mathbf{H}, \quad (27-46)$$

$$\mathbf{H}' = +K\mathbf{D} = +K\epsilon\mathbf{E}, \quad (27-47)$$

where the constant K has the dimensions of a velocity and is independent of x , y , z , t . This other field *also* satisfies Maxwell's equations, as you can check by substitution into Eqs. 27-20 to 27-23.

Figure 27-5 illustrates this duality property of electromagnetic fields. One field is said to be the *dual*, or the *dual field*, of the other. Therefore, if one field can exist, then its dual can also exist.

Example THE FIELDS OF ELECTRIC AND MAGNETIC DIPOLES

In Sec. 5.1 we found that in the field of an electric dipole of moment \mathbf{p}

$$\mathbf{E} = \frac{\mathbf{p}}{4\pi\epsilon_0 r^3} (2 \cos \theta \hat{\mathbf{r}} + \sin \theta \hat{\mathbf{\theta}}). \quad (27-48)$$

Later, in the fourth example in Sec. 18.4, we showed that in the field of a magnetic dipole of moment \mathbf{m}

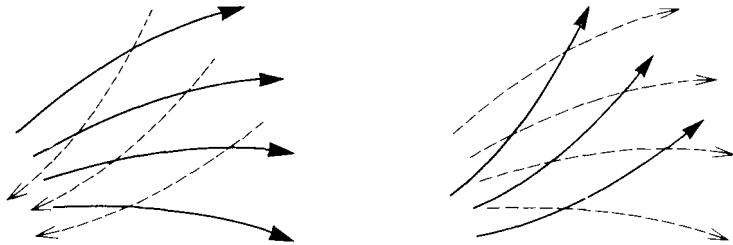


Fig. 27-5. Pair of dual fields. Lines of \mathbf{E} are solid, and lines of \mathbf{H} dashed.

$$\mathbf{B} = \frac{\mu_0 m}{4\pi r^3} (2 \cos \theta \hat{r} + \sin \theta \hat{\theta}), \quad (27-49)$$

or

$$\mathbf{H} = \frac{m}{4\pi r^3} (2 \cos \theta \hat{r} + \sin \theta \hat{\theta}). \quad (27-50)$$

We have set $\epsilon_0 = 1$, $\mu_0 = 1$. If the field of the electric dipole is the unprimed field, and that of the magnetic dipole the primed field, then $K = m/p$.

27.7 LORENTZ'S RECIPROCITY THEOREM

Consider two fields \mathbf{E}_a , \mathbf{B}_a , and \mathbf{E}_b , \mathbf{B}_b in a linear and isotropic medium. Because of the principle of superposition, these two fields can either exist separately or be superimposed without disturbing each other, giving a third field $\mathbf{E}_a + \mathbf{E}_b$, $\mathbf{B}_a + \mathbf{B}_b$.

We now use the vector identity

$$\begin{aligned} \nabla \cdot (\mathbf{E}_a \times \mathbf{B}_b - \mathbf{E}_b \times \mathbf{B}_a) &= \mathbf{B}_b \cdot (\nabla \times \mathbf{E}_a) - \mathbf{E}_a \cdot (\nabla \times \mathbf{B}_b) \\ &\quad - \mathbf{B}_a \cdot (\nabla \times \mathbf{E}_b) + \mathbf{E}_b \cdot (\nabla \times \mathbf{B}_a). \end{aligned} \quad (27-51)$$

Then, from the Maxwell equations for the curls of \mathbf{E} and \mathbf{B} ,

$$\begin{aligned} \nabla \cdot (\mathbf{E}_a \times \mathbf{B}_b - \mathbf{E}_b \times \mathbf{B}_a) &= -\mathbf{B}_b \cdot \frac{\partial \mathbf{B}_a}{\partial t} - \mathbf{E}_a \cdot \mu_0 \left(\mathbf{J}_b + \epsilon_0 \frac{\partial \mathbf{E}_b}{\partial t} \right) \\ &\quad + \mathbf{B}_a \cdot \frac{\partial \mathbf{B}_b}{\partial t} + \mathbf{E}_b \cdot \mu_0 \left(\mathbf{J}_a + \epsilon_0 \frac{\partial \mathbf{E}_a}{\partial t} \right). \end{aligned} \quad (27-52)$$

If the two fields are harmonic functions of the time and if they are of the same frequency, the $\partial/\partial t$ operators can be replaced by $j\omega$, and the time derivatives cancel. Substituting the value of the total current density \mathbf{J} ,

$$\begin{aligned}\nabla \cdot (\mathbf{E}_a \times \mathbf{B}_b - \mathbf{E}_b \times \mathbf{B}_a) &= -\mu_0 \mathbf{E}_a \cdot \left(\mathbf{J}_{fb} + \frac{\partial \mathbf{P}_b}{\partial t} + \nabla \times \mathbf{M}_b \right) \\ &\quad + \mu_0 \mathbf{E}_b \cdot \left(\mathbf{J}_{fa} + \frac{\partial \mathbf{P}_a}{\partial t} + \nabla \times \mathbf{M}_a \right).\end{aligned}\quad (27-53)$$

Since, by hypothesis the medium is linear and isotropic, \mathbf{P} is $\epsilon_0 \chi_e \mathbf{E}$ and the above time derivatives cancel. Finally, if the point considered is not inside a source, Ohm's law applies, $\mathbf{J}_f = \sigma \mathbf{E}$, and the \mathbf{J}_f terms cancel. We are then left with the $\nabla \times \mathbf{M}$ terms. If the medium is nonmagnetic, then \mathbf{M} is zero and

$$\nabla \cdot (\mathbf{E}_a \times \mathbf{B}_b - \mathbf{E}_b \times \mathbf{B}_a) = 0.\quad (27-54)$$

If the medium is magnetic, we can perform a similar calculation by using \mathbf{H} 's instead of \mathbf{B} 's and

$$\nabla \cdot (\mathbf{E}_a \times \mathbf{H}_b - \mathbf{E}_b \times \mathbf{H}_a) = 0.\quad (27-55)$$

In summary, therefore, if two fields a and b are sinusoidal and of the same frequency, if also the medium is linear and isotropic, and if the point considered is not inside a source, then Eq. 27-54 applies. If the medium is magnetic, then we have Eq. 27-55.

Applying the divergence theorem yields

$$\int_{\mathcal{A}} (\mathbf{E}_a \times \mathbf{H}_b - \mathbf{E}_b \times \mathbf{H}_a) \cdot d\mathcal{A} = 0,\quad (27-56)$$

where \mathcal{A} is the area of any closed surface, with the above restrictions. This result is known as *Lorentz's reciprocity theorem*. It is paradoxical because it establishes a relation between two unrelated electromagnetic fields.

Examples

- (1) If the a field is purely electric and the b field purely magnetic, then

$$\nabla \cdot (\mathbf{E}_a \times \mathbf{H}_b) = 0,\quad (27-57)$$

subject *only* to the above limitations. This surprising statement becomes obvious after expanding the divergence:

$$\nabla \cdot (\mathbf{E}_a \times \mathbf{H}_b) = \mathbf{H}_b \cdot (\nabla \times \mathbf{E}_a) - \mathbf{E}_a \cdot (\nabla \times \mathbf{H}_b)\quad (27-58)$$

$$= -\mathbf{H}_b \cdot \frac{\partial \mathbf{B}_a}{\partial t} - \mathbf{E}_a \cdot \left(\mathbf{J}_{fb} + \frac{\partial \mathbf{D}_b}{\partial t} \right).\quad (27-59)$$

This last quantity is identically equal to zero because, by hypothesis, \mathbf{B}_a and \mathbf{E}_b are both zero, $\mathbf{J}_b = \sigma\mathbf{E}_b = 0$, and $\mathbf{D}_b = \epsilon\mathbf{E}_b = 0$.

(2) The field of a small permanent magnet is that of a magnetic dipole, as in the last example in Sec. 18.4. If the magnet carries an electric charge Q , it also has an electric field

$$\mathbf{E}_r = \frac{Q\hat{\mathbf{r}}}{4\pi\epsilon_0 r^2}, \quad (27-60)$$

and

$$\nabla \cdot (\mathbf{E}_a \times \mathbf{H}_b) = \nabla \cdot \begin{vmatrix} \hat{\mathbf{r}} & \hat{\theta} & \hat{\phi} \\ \mathbf{E}_{ar} & 0 & 0 \\ H_{br} & H_{b\theta} & 0 \end{vmatrix} = \nabla \cdot (E_{ar} H_{b\theta} \hat{\phi}) \quad (27-61)$$

$$= \nabla \cdot \left(\frac{Q}{4\pi\epsilon_0 r^2} \frac{\mu_0 m}{4\pi r^3} \sin \theta \hat{\phi} \right) = 0. \quad (27-62)$$

27.8 THE WAVE EQUATIONS FOR \mathbf{E} AND FOR \mathbf{B}

Taking the curl of Eq. 27-2 and remembering that

$$\nabla \times \nabla \times \mathbf{E} = -\nabla^2 \mathbf{E} + \nabla(\nabla \cdot \mathbf{E}), \quad (27-63)$$

then, from Eq. 27-4,

$$\nabla^2 \mathbf{E} - \nabla(\nabla \cdot \mathbf{E}) = \frac{\partial}{\partial t} \nabla \times \mathbf{B} = \frac{\partial}{\partial t} \left(\mu_0 \mathbf{J} + \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t} \right). \quad (27-64)$$

Substituting now the value of the divergence of \mathbf{E} from Eq. 27-1 and rearranging,

$$\nabla^2 \mathbf{E} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{\nabla \rho}{\epsilon_0} + \mu_0 \frac{\partial \mathbf{J}}{\partial t}. \quad (27-65)$$

This is the *nonhomogeneous wave equation for \mathbf{E}* . The source terms are on the right.

Outside the sources,

$$\nabla^2 \mathbf{E} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0. \quad (27-66)$$

This is the usual *wave equation*. The speed of propagation, which is the *speed of light*, is

$$c = \frac{1}{(\epsilon_0 \mu_0)^{1/2}}. \quad (27-67)$$

Similarly, taking the curl of Eq. 27-4 and substituting Eqs 27-2 and 27-3, we find that

$$\nabla^2 \mathbf{B} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{B}}{\partial t^2} = -\mu_0 \nabla \times \mathbf{J}, \quad (27-68)$$

which is the *nonhomogeneous wave equation for \mathbf{B}* . The source term is again on the right.

Outside the sources,

$$\nabla^2 \mathbf{B} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{B}}{\partial t^2} = 0. \quad (27-69)$$

According to the rule given in Sec. 27-1, the wave equations for a HILS medium are as follows:

$$\nabla^2 \mathbf{E} - \epsilon \mu \frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{\nabla \rho_f}{\epsilon} + \mu \frac{\partial \mathbf{J}_f}{\partial t}, \quad (27-70)$$

$$\nabla^2 \mathbf{B} - \epsilon \mu \frac{\partial^2 \mathbf{B}}{\partial t^2} = -\mu \nabla \times \mathbf{J}_f. \quad (27-71)$$

We therefore have a wave equation for the field \mathbf{E} , and a separate wave equation for \mathbf{B} . Within the wave, however, \mathbf{E} and \mathbf{B} are inextricably related through Maxwell's equations. In other words, purely electric, or purely magnetic, waves are impossible. The fact remains that, in some waves, the energy density can be either mostly magnetic or mostly electric.

If σ is constant,

$$\nabla^2 \mathbf{E} - \epsilon \mu \frac{\partial^2 \mathbf{E}}{\partial t^2} - \sigma \mu \frac{\partial \mathbf{E}}{\partial t} = \frac{\nabla \rho_f}{\epsilon}, \quad (27-72)$$

$$\nabla^2 \mathbf{B} - \epsilon \mu \frac{\partial^2 \mathbf{B}}{\partial t^2} - \sigma \mu \frac{\partial \mathbf{B}}{\partial t} = 0. \quad (27-73)$$

27.9 SUMMARY

We expressed *Maxwell's equations* in several different forms. See the back cover.

All the quantities that appear there are defined in Sec. 27.1.

The *displacement current density* is

$$\frac{\partial \mathbf{D}}{\partial t} = \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \frac{\partial \mathbf{P}}{\partial t}. \quad (27-38)$$

The second term on the right is the polarization current density. The first term can exist even in a vacuum.

The *law of conservation of charge*, in its general form, takes into account both free and bound charges:

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t}, \quad (27-42)$$

where ρ and \mathbf{J} are defined in Sec. 27.1.

The equations for the curl of \mathbf{E} and for the divergence of \mathbf{B} are redundant since one follows from the other. They are occasionally called the *first pair*. The other two are similarly related and form the *second pair*.

For every field \mathbf{E} , \mathbf{H} there can exist a *dual field* \mathbf{E}' , \mathbf{H}' such that

$$\mathbf{E}' = -K\mathbf{B} = -K\mu\mathbf{H}, \quad \mathbf{H}' = +K\mathbf{D} = +K\epsilon\mathbf{E}. \quad (27-46), (27-47)$$

This is the *duality property* of electromagnetic fields.

The *Lorentz reciprocity theorem* states that, for any two fields \mathbf{E}_a , \mathbf{H}_a and \mathbf{E}_b , \mathbf{H}_b that are sinusoidal and of the same frequency, in linear and isotropic media away from sources,

$$\nabla \cdot (\mathbf{E}_a \times \mathbf{H}_b - \mathbf{E}_b \times \mathbf{H}_a) = 0, \quad (27-55)$$

The *nonhomogeneous wave equations* for \mathbf{E} and for \mathbf{B} are as follows:

$$\nabla^2 \mathbf{E} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{\nabla \rho}{\epsilon_0} + \mu_0 \frac{\partial \mathbf{J}}{\partial t}, \quad (27-65)$$

$$\nabla^2 \mathbf{B} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{B}}{\partial t^2} = -\mu_0 \nabla \times \mathbf{J}. \quad (27-71)$$

Outside the sources, the terms on the right-hand side are zero. The *speed of light* is thus

$$c = \frac{1}{(\epsilon_0 \mu_0)^{1/2}}. \quad (27-67)$$

In a HILS medium, if σ is constant,

$$\nabla^2 \mathbf{E} - \epsilon \mu \frac{\partial^2 \mathbf{E}}{\partial t^2} - \sigma \mu \frac{\partial \mathbf{E}}{\partial t} = \frac{\nabla \rho_f}{\epsilon}, \quad (27-72)$$

$$\nabla^2 \mathbf{B} - \epsilon \mu \frac{\partial^2 \mathbf{B}}{\partial t^2} - \sigma \mu \frac{\partial \mathbf{E}}{\partial t} = 0. \quad (27-73)$$

PROBLEMS

27-1. (27.1) Superconductivity

A superconductor offers zero resistance to the motion of superconducting charge carriers. These are pairs of electrons that move as a unit.

(a) If there are N such carriers per cubic meter of mass m' and charge e' , show that

$$\mathbf{E} = \frac{m'}{Ne'^2} \frac{d\mathbf{J}}{dt}.$$

This is the *first London equation*. Note that \mathbf{E} is zero only if \mathbf{J} is constant.

(b) Set $K = m'/(Ne'^2)$. Show that in an alternating field

$$\sigma = \frac{1}{j\omega K} = -\frac{j}{\omega K}.$$

(c) Show that

$$\nabla \times \left(K \frac{\partial \mathbf{J}}{\partial t} \right) = -\frac{\partial \mathbf{B}}{\partial t}.$$

The equation

$$\nabla \times K\mathbf{J} = -\mathbf{B}$$

is the *second London equation*. It does *not* follow mathematically from the first.

(d) Show that under steady-state conditions

$$\nabla^2 \mathbf{B} = \frac{\mu_0}{K} \mathbf{B}.$$

In one dimension, this means that

$$\frac{d^2 \mathbf{B}}{dx^2} = \frac{\mu_0}{K} \mathbf{B},$$

or that

$$\mathbf{B} = \mathbf{B}_0 \exp \frac{x}{(K/\mu_0)^{1/2}},$$

where $(K/\mu_0)^{1/2}$ is the *depth of penetration* of the field.

(e) Calculate the value of the depth of penetration, setting m' equal to twice the mass of an electron, $e' = 2e$, and $N = 10^{29}$. The depth of penetration is, in fact, a few times larger.

(f) Much beyond the depth of penetration, $\mathbf{E} = 0$, $\mathbf{J} = 0$, $\mathbf{B} = 0$.

Show that just outside a superconductor \mathbf{B} is tangential to the surface and equal in magnitude to the surface current density.

27-2. (27.1) Experiments on models

A problem that cannot be solved on paper can often be solved in the laboratory. However, there are instances where a full-scale experiment

would be too costly. One such problem is that of the design of an MHD generator (Sec. 22.1).

In such cases it is sometimes useful to perform experiments on a *model* of convenient size. One then has the real system, for which the variables are $x, y, z, t, \mathbf{E}, \mathbf{B}, \epsilon_r, \mu_r, \sigma, R, C, L$, etc., and the model, whose variables are x', y', z', t' , etc. The ratios $x/x', y/y', z/z'$, etc. are the *scale factors*. Not all these factors can be chosen arbitrarily because both the unprimed and the primed variables must satisfy Maxwell's equations. The number of arbitrary scale factors is equal to 4, the number of fundamental units (meter, kilogram, second, ampere). With mechanical systems there are only three arbitrary scale factors.

Let us set

$$\frac{x}{x'} = \frac{y}{y'} = \frac{z}{z'} = l, \quad \frac{t}{t'} = \tau, \quad \frac{E}{E'} = e, \quad \frac{H}{H'} = h.$$

The other scale factors follow from Maxwell's equations and from other relations.

(a) Use Maxwell's equations to show that

$$(i) \frac{\mu_r}{\mu'_r} = \frac{\epsilon\tau}{lh}, \quad (ii) \frac{\epsilon_r}{\epsilon'_r} = \frac{\tau h}{el}, \quad (iii) \frac{\sigma}{\sigma'} = \frac{h}{el}.$$

(b) Show that

$$(i) \frac{\mathcal{S}}{\mathcal{S}'} = eh, \quad (ii) \frac{J}{J'} = \frac{h}{l}, \quad (iii) \frac{I}{I'} = hl, \quad (iv) \frac{R}{R'} = \frac{e}{h}, \quad (v) \frac{V}{V'} = el.$$

(c) Show that

$$(i) \frac{\Phi}{\Phi'} = e\tau l, \quad (ii) \frac{B}{B'} = \frac{e\tau}{l}, \quad (iii) \frac{C}{C'} = \tau \frac{h}{\epsilon}, \quad (iv) \frac{L}{L'} = \frac{e\tau}{h}.$$

(d) Show that

$$(i) \frac{f}{f'} = \frac{1}{\tau}, \quad (ii) \frac{Q}{Q'} = hl\tau, \quad (iii) \frac{\rho}{\rho'} = \frac{\tau h}{l^2}, \quad (iv) \frac{P}{P'} = hel^2.$$

(e) Show that, if L is a length, then

$$\frac{f\mu\sigma L^2}{f'\mu'\sigma'L'^2} = 1.$$

(f) One author states that

$$\frac{f^2\epsilon\mu}{f'^2\epsilon'\mu'} = 1.$$

Is he right?

In practice, these relations are simplified by the exclusion of ferromagnetic materials because of their nonlinearity. Then $\mu_r/\mu'_r = 1$, $e\tau = lh$, and

there are only three independent scale factors. Then $\sigma/\sigma' = \tau/l^2$ and $\sigma\omega L^2 = \sigma'\omega'L'^2$, where L is a length.

If the fields are in a vacuum, $\epsilon_0/\epsilon'_0 = 1$ and $\tau h = el$. Then $e = h$ and $\tau = 1$. Note that if $\epsilon_0/\epsilon'_0 = 1$, $\mu_r/\mu'_r = 1$, and $\sigma/\sigma_r = 1$, then

$$\frac{\tau h}{el} = \frac{e\tau}{lh} = \frac{h}{el} = 1$$

and $e = h$, $\tau = 1$, $l = 1$. Then the model is the same size as the original! If the model is to be a different size, then either ϵ_0 , or σ , or both, must be different. This condition is often impossible to satisfy.

At low frequencies one can disregard the displacement current, and hence attribute any value to the ratio ϵ_0/ϵ'_0 , or to $\tau h/el$.

- 27-3. (27.2) Parallel-plate capacitor fed at one end by a time-dependent source

Figure 27-6 shows a parallel-plate capacitor connected at one end to a source whose voltage increases slowly and linearly with time: $dV_0/dt = k$. Edge effects are negligible: $a \gg s$, $b \gg s$.

- (a) Find the current I as a function of x .
- (b) Find \mathbf{B} inside in two different ways. Find \mathbf{A} inside.
- (c) Find \mathbf{B} and \mathbf{A} outside. The capacitor plates are thin.
- (d) Draw a large cross section of the capacitor in the midplane parallel to the xz -plane, showing I and the vectors \mathbf{A} , \mathbf{B} , $\nabla \times \mathbf{B}$, \mathbf{E} , $\partial \mathbf{E}/\partial t$ near both ends. Use arrows of different sizes to indicate qualitatively how these vectors vary with x and with z .

- 27-4. (27.1) A transformation that leaves Maxwell's equations invariant

- (a) Show that Maxwell's equations for free space are invariant under the transformation

$$\mathbf{E}' = a\mathbf{E} + bc\mathbf{B}, \quad \mathbf{B}' = -\left(\frac{b}{c}\right)\mathbf{E} + a\mathbf{B},$$

where a and b are constants and c is the speed of light.

- (b) Under what condition are the energy density $\epsilon_0 E^2/2 + B^2/(2\mu_0)$ and the Poynting vector $\mathbf{E} \times \mathbf{H}$ invariant?

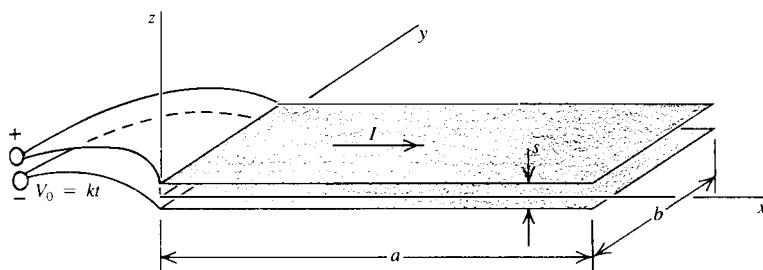


Fig. 27-6.

27-5. (27.3) The skin effect

As we shall see in Sec. 29.1, a high-frequency field does not penetrate significantly into the body of a good conductor. Also, both \mathbf{E} and \mathbf{B} inside are approximately tangent to the surface. Let the z -axis be normal to the surface, pointing outward, with \mathbf{E} in the direction of the x -axis and \mathbf{B} in the direction of the y -axis. Set $\partial/\partial x = 0$, $\partial/\partial y = 0$.

- Show that, inside, $\partial E/\partial z = -\partial B/\partial t$.
- Show that, just outside the conductor, \mathbf{B} is tangent, or nearly so.
- Let the current density near the surface be \mathbf{a} amperes/meter.
Show that, just outside the conductor, $\mathbf{B} = \mu_0 \mathbf{a} \times \hat{\mathbf{z}}$
- Does this last result depend on how the current varies with depth inside the conductor?

27-6. (27.3) The magnetic field of a leaky capacitor

A charged capacitor whose electrodes are parallel and circular lies in a large volume of dielectric that is slightly conducting. The capacitor discharges.

- Calculate the value of the ratio $J_f/(\partial D/\partial t)$ at any point in the dielectric in terms of the resistance and the capacitance.
- Show that this ratio is equal to -1 at the surface of an electrode.
- Show that \mathbf{B} is zero everywhere in the dielectric. This means that the magnetic field of the conduction and polarization (*not* displacement) currents in the fringing field exactly cancels the magnetic field of the conduction and polarization currents in the region between the plates.
- Show that \mathbf{B} is also zero for electrodes of any shape.

If the dielectric occupies only part of the field of a capacitor, say the region between the plates of a parallel-plate capacitor, then the value of the ratio calculated under (a) applies. However, at the surface of an electrode, this ratio is not equal to -1 because charge migrates from the outer surface of an electrode to the inner surface, where it leaks out. Then, in the dielectric, $|J_f| > |\partial D/\partial t|$, $\mathbf{J}_f - \partial \mathbf{D}/\partial t$ points in the direction of \mathbf{J}_f and thus of \mathbf{E} , and there is an azimuthal magnetic field.

27-7. (27.3) Magnetic monopoles

If monopoles exist, then Maxwell's equations require two more terms, to take into account magnetic charges and magnetic currents. It is the custom to write the equations in the following form:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} - \mathbf{J}^*, \quad \nabla \cdot \mathbf{B} = \rho^*, \quad \nabla \times \mathbf{B} = \mu_0 \left(\epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J} \right),$$

where ρ^* is the magnetic charge density, expressed in webers/meter³, and \mathbf{J}^* is the magnetic current density, in webers/second-meter².

- Show that

$$\nabla \cdot \mathbf{J}^* = -\frac{\partial \rho^*}{\partial t}.$$

This is the equation of conservation for magnetic monopoles.

- Show that, by analogy with electric fields, near a point magnetic charge Q^*

$$\mathbf{B} = \frac{Q^*}{4\pi r^2} \hat{\mathbf{r}}, \quad \mathbf{H} = \frac{Q^*}{4\pi\mu_0 r^2} \hat{\mathbf{r}}.$$

(c) Calculate the energy acquired by a magnetic monopole that accelerates over a distance of 1000 kilometers in the earth's magnetic field ($\approx 10^{-5}$ tesla).

(d) Magnetic monopoles go through a loop of copper wire.

Show that the induced electromotance is equal to minus the magnetic current, with the right-hand screw convention. This is one method of detecting magnetic monopoles.

27-8. (27.3) The continuous-creation theory and Maxwell's equations

Imagine an expanding spherically symmetric universe in which there is continuous creation of charge at the rate of q coulombs/meter³-second. Creation of electric charge occurs through the creation of hydrogen atoms carrying a slight excess charge ye as in Prob. 3-15.

The rate of mass creation Q is proportional to q : $Q = [m/(ye)]q$, where m is the mass of the proton, the universe being mostly hydrogen.

(a) By symmetry, the vector potential can only be radial.

Show that under steady-state conditions the current density \mathbf{J} is everywhere zero, according to Maxwell's equations.

Lyttleton and Bondi (see Prob. 3-15) suggested that, if continuous creation does exist, then Maxwell's equations must be modified as follows:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} - \left[\frac{1}{l^2} \mathbf{A} \right], \quad \nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} - \left[\frac{1}{l^2} V \right],$$

where the new terms are enclosed in brackets. The quantities V and \mathbf{A} are the usual scalar and vector potentials:

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t}, \quad \mathbf{B} = \nabla \times \mathbf{A}.$$

The other two equations of Maxwell for $\nabla \times \mathbf{E}$ and $\nabla \cdot \mathbf{B}$ remain unchanged. Lyttleton and Bondi suggested that the constant l , which has the dimensions of a length, would be of the order of the radius of the universe. The new terms would therefore be negligible in all but cosmological problems.

(b) If these modified Maxwell equations are correct, are V and \mathbf{A} measurable, in principle? Remember that, with the above equations for \mathbf{E} and \mathbf{B} in terms of V and \mathbf{A} , only the rates of change of V and \mathbf{A} determine \mathbf{E} and \mathbf{B} .

(c) Write out the equation for the conservation of the total charge (Sec. 27.4).

(d) Would the Lorentz condition (Secs. 17.9 and 37.1) still be valid?

(e) Now set $\mathbf{A} = A' \mathbf{r}$, where A' is a constant, and assume V to be constant. Show that $\mathbf{B} = 0$, $\mathbf{E} = 0$, $\mathbf{J} = (q/3)\mathbf{r}$, $\rho = \epsilon_0 V/l^2$.

Assuming that the velocity of the outward flow of matter is the same as that of the charge, namely \mathbf{J}/ρ , it follows that the radial velocity is

proportional to r , which is consistent with the linear velocity-distance relation observed by astronomers: $v = r/T$, where $T \approx 3 \times 10^{17}$ seconds is the *Hubble constant*.

(f) Show that $\rho = qT/3$.

(g) Now the space-charge density ρ is $\eta ye/m$, where η , the mass density of the universe, is about 10^{-26} kilogram/meter³.

Show that, if this theory is correct, then $Q \approx 1/(2 \times 10^{16})$ hydrogen atom/meter³-second.

27-9. (27.3) Another transformation that leaves Maxwell's equations invariant

(a) Show that Maxwell's equations for free space are invariant under the transformation

$$\mathbf{E}' = \mathbf{E} \cos \theta + c\mathbf{B} \sin \theta, \quad \mathbf{B}' = -\frac{\mathbf{E}}{c} \sin \theta + \mathbf{B} \cos \theta.$$

The transformation $\mathbf{E}' = -K\mathbf{B}$, $\mathbf{H}' = K\mathbf{D}$ of Sec. 27.7 and the transformation $\mathbf{E}' = -\mathbf{E}$, $\mathbf{B}' = -\mathbf{B}$ are special cases corresponding to $\theta = \pi/2$ and $\theta = \pi$, respectively.

(b) Show that the energy density $\epsilon_0 E^2/2 + B^2/(2\mu_0)$ and the Poynting vector $\mathbf{E} \times \mathbf{H}$ are also invariant under this transformation.

27-10. (27.3) The magnetic field of a point charge that moves at a constant velocity

Figure 27-7 shows a point charge Q that travels along the x -axis at a velocity $V\hat{x}$. Its position at time t is $(Vt, 0, 0)$.

(a) Find \mathbf{B} at a point $P(X, Y, 0)$, not at the origin, at the instant that the charge passes through the origin. The particle travels in a vacuum, and $v^2 \ll c^2$, where c is the speed of light.

(b) If you have studied Chap. 16, compare your result with that of Sec. 16.5.4.

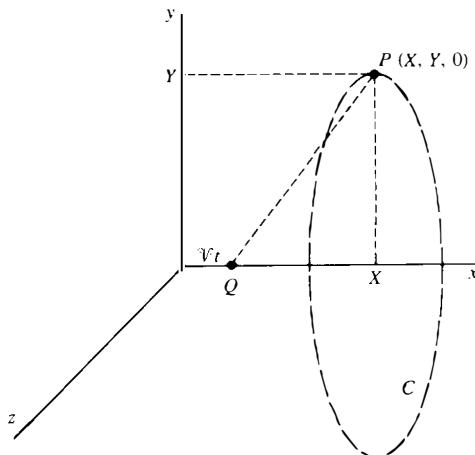


Fig. 27-7.

(c) Sketch the value of

$$\frac{d}{dt} \int_{\mathcal{A}} \mathbf{D} \cdot d\mathcal{A},$$

where \mathcal{A} is the area of the spherical segment of radius Y , as a function of $\mathcal{V}t$. When the charge is just to the left of $\mathcal{V}t = X$, the flux of \mathbf{D} through \mathcal{A} points to the right and is $+Q/2$. Immediately afterward, the flux points to the left and is $-Q/2$, so there is a discontinuity in this curve.

(d) The integral of $\mathbf{H} \cdot d\mathbf{l}$ around the circle shown in the figure is equal to $2\pi YB/\mu_0$. Calculate the integral of this quantity over time from minus infinity to plus infinity. Set $X = 0$ to simplify the calculation. Explain your result.

27-11. (27.9) The Watson theory of continuous charge creation

Problems 3-15 and 27-8 sketch the Lyttleton theory, according to which hydrogen atoms are continuously created in the universe, each atom bearing a slight positive charge. W. H. Watson had proposed a similar theory several years before.[†] Watson postulated a scalar potential N such that

$$\nabla \cdot \mathbf{D} = \rho_f + \epsilon_0 \mu_0 \frac{\partial N}{\partial t}, \quad \nabla \times \mathbf{H} = \mathbf{J}_f + \frac{\partial \mathbf{D}}{\partial t} - \nabla N.$$

(a) Find the nonhomogeneous wave equation for N from the equation for the nonconservation of charge. Set the rate of charge creation equal to q coulombs/meter³-second.

(b) Find the nonhomogeneous wave equations for \mathbf{E} and for \mathbf{H} .

[†] See L. G. Chambers, *Journal of Mathematical Physics*, vol. 4, p. 1373 (1963).

CHAPTER 28

PLANE ELECTROMAGNETIC WAVES I

Uniform Plane Waves in Free Space, in Nonconductors, and in Conductors

28.1	THE ELECTROMAGNETIC SPECTRUM	515
28.2	UNIFORM PLANE ELECTROMAGNETIC WAVES IN A GENERAL MEDIUM	515
28.2.1	THE RELATIVE ORIENTATIONS OF \mathbf{E} , \mathbf{H} AND \mathbf{k}	517
28.2.2	THE CHARACTERISTIC IMPEDANCE Z OF A MEDIUM	518
28.2.3	THE WAVE NUMBER k	518
28.2.4	THE WAVE EQUATIONS	518
28.3	UNIFORM PLANE WAVES IN FREE SPACE	520
	<i>Example: THE \mathbf{E} AND \mathbf{B} FIELDS IN A LASER BEAM</i>	522
28.4	UNIFORM PLANE WAVES IN NONCONDUCTORS	522
	<i>Example: THE \mathbf{E} AND \mathbf{B} FIELDS OF A LASER BEAM IN GLASS</i>	524
28.5	UNIFORM PLANE WAVES IN CONDUCTORS	524
28.5.1	THE COMPLEX WAVE NUMBER $k = \beta - j\alpha$	524
28.5.2	THE CHARACTERISTIC IMPEDANCE Z OF A CONDUCTOR	526
28.5.3	THE ENERGY DENSITIES	527
28.6	THE POYNTING THEOREM	528
28.7	SUMMARY	529
	PROBLEMS	530

We now go on to the study of electromagnetic waves. This chapter concerns the propagation of uniform plane waves in unbounded media, first a general medium, then free space, then nonconductors, and then conductors. Good conductors and ionized gases follow in the next chapter. Later we shall study reflection and refraction in Chaps. 30 to 32, guided waves in Chaps. 33 to 36, and finally the process of radiation in Chaps. 37 to 39.

If you are not familiar with wave propagation you would be well advised to read App. C now.

28.1 THE ELECTROMAGNETIC SPECTRUM

Maxwell's equations impose no limit on the frequency of electromagnetic waves. The known spectrum extends continuously from the long radio waves to the very high-energy gamma rays of cosmic radiation, as in Fig. 28-1. In the former, the frequencies are of the order of 100 hertz and the wavelengths about 3 megameters; in the latter, the frequencies are of the order of 10^{24} hertz and the wavelengths less than 1 femtometer.

The known spectrum thus covers a range of 22 orders of magnitude. Radio, heat waves, light, x-rays, and gamma rays are all electromagnetic, although the sources and the detectors, as well as the modes of interaction with matter, vary widely as the frequency changes by orders of magnitude.

Many experiments demonstrate the fundamental identity of all these waves. In particular, they are all transverse, and they all travel at the speed c in free space, except in special circumstances. For example, simultaneous radio and optical observations on stars show that the velocity of propagation is the same, within experimental error, for wavelengths differing by more than 6 orders of magnitude.

We use \mathbf{H} rather than \mathbf{B} in discussing electromagnetic waves, in spite of the fact that until now we have used \mathbf{H} only for magnetic materials. There are two reasons for using \mathbf{H} , instead of \mathbf{B} , in dealing with electromagnetic waves: one is that $\mathbf{E} \times \mathbf{H}$ is a power density, and the other is that E/H is an impedance. These two concepts have great practical value.

28.2 UNIFORM PLANE ELECTROMAGNETIC WAVES IN A GENERAL MEDIUM

A *wave front* is a surface of uniform phase. The wave fronts of a *plane wave* are planar. A wave is *uniform* if a wave front is a surface of uniform phase and uniform amplitude. We shall not be concerned with nonuniform waves until Chap. 31.

Uniform plane electromagnetic waves in unbounded media possess several general properties that apply whether the wave travels in free space or in matter. To avoid needless repetition, we start with a general medium ϵ_r , μ_r , σ that is homogeneous, isotropic, linear, and stationary (HILS).

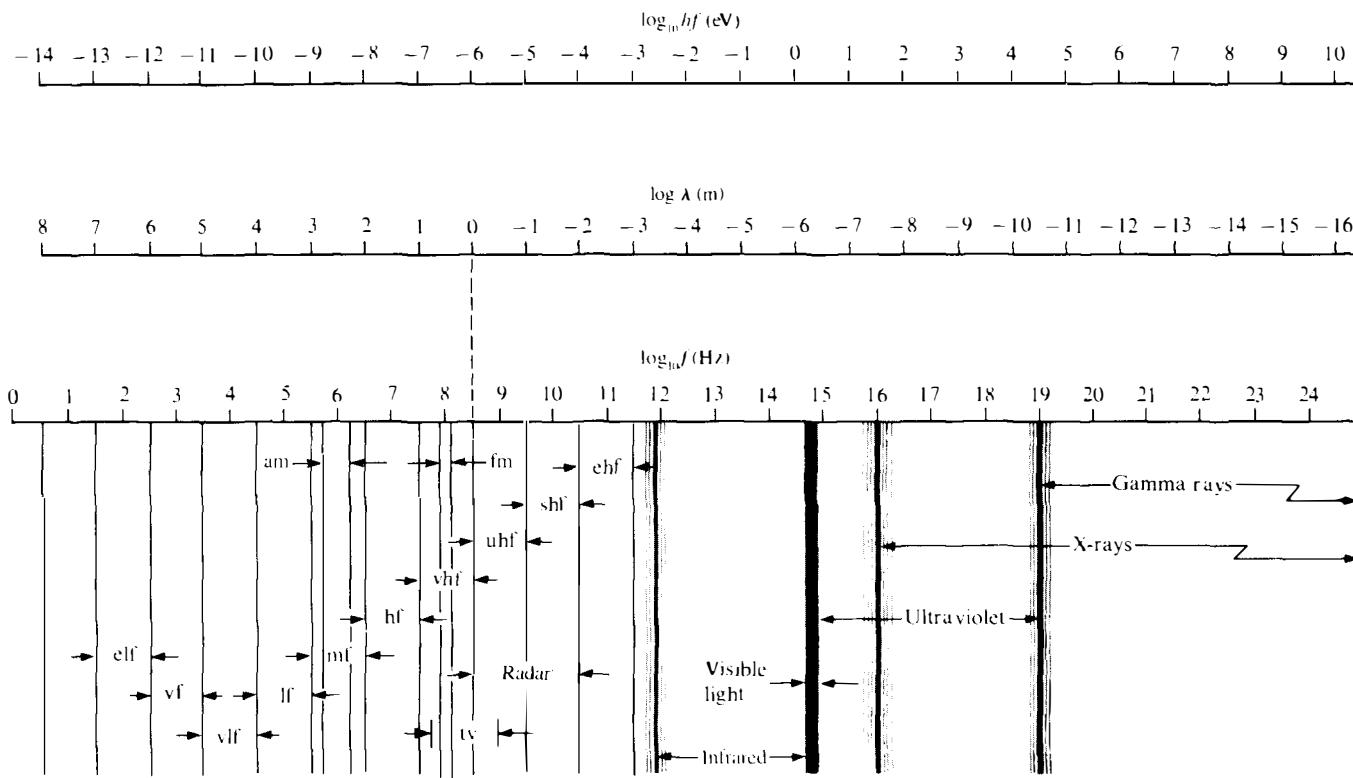


Fig. 28-1. The spectrum of electromagnetic waves. The abbreviations elf, vf, vlf, ... mean, respectively, extremely low frequency, voice frequency, very low frequency, low frequency, medium frequency, high frequency, very high frequency, ultrahigh frequency, super high frequency, and extremely high frequency. The limits indicated by the shaded regions are approximate. The energy hf , where h is Planck's constant (6.63×10^{-34} joule-second) and f is the frequency, is that of a photon or quantum of radiation.

We assume a sinusoidal wave traveling in the positive direction of the z -axis. We also assume that the \mathbf{E} vectors are all parallel to a given direction. In other words, we assume that the wave is *linearly polarized*. If the plane wave is not linearly polarized, then it is the sum of linearly polarized waves.[†] The *plane of polarization* is parallel to \mathbf{E} .

In a linearly polarized plane wave, \mathbf{E} and \mathbf{H} are of the form

$$\mathbf{E} = \mathbf{E}_m \exp j(\omega t - kz), \quad \mathbf{H} = \mathbf{H}_m \exp j(\omega t - kz)^{\ddagger} \quad (28-1)$$

where \mathbf{E}_m and \mathbf{H}_m are vectors that are independent of the time and of the coordinates. If there is no attenuation, the *wave number* is real:

$$k = \frac{\omega}{v} = \frac{2\pi}{\lambda} = \frac{1}{\bar{\lambda}}, \quad (28-2)$$

where v is the *phase velocity*, λ is the *wavelength*, and $\bar{\lambda}$ (pronounced “lambda bar”) is the *radian length*. You can easily show by substitution that Eqs. 28-1 are solutions of Eqs. 27-66 and 27-69.

28.2.1 The Relative Orientations of \mathbf{E} , \mathbf{H} , and \mathbf{k}

For this particular field,

$$\frac{\partial}{\partial t} = j\omega, \quad \nabla = \frac{\partial}{\partial x} \hat{x} + \frac{\partial}{\partial y} \hat{y} + \frac{\partial}{\partial z} \hat{z} = \frac{\partial}{\partial z} \hat{z} = -jk\hat{z}. \quad (28-3)$$

We set $\rho_f = 0$. We also set

$$\mathbf{J}_f = \sigma \mathbf{E}, \quad (28-4)$$

on the assumption that $\mathbf{v} \times \mathbf{B}$ is negligible compared to \mathbf{E} , where \mathbf{v} is the velocity of a conduction electron.

[†] For example, one can add two linearly polarized waves \mathbf{E}_1 and \mathbf{E}_2 , with \mathbf{E}_1 perpendicular to \mathbf{E}_2 , that differ in phase. Then, at a given point, the maxima of \mathbf{E}_1 and of \mathbf{E}_2 do not occur at the same time, and their sum \mathbf{E} describes an ellipse about the z -axis. We then have an *elliptically polarized wave*. If \mathbf{E}_1 and \mathbf{E}_2 have equal amplitudes but are $\pi/2$ radians out of phase, the ellipse becomes a circle and the wave is *circularly polarized*. The polarization is *right-* or *left-handed* according to whether the vectors \mathbf{E} and \mathbf{H} rotate clockwise or counterclockwise for an observer looking toward the *source*.

[‡] We shall use this particular notation for a wave traveling in the positive direction of the z -axis because it is the most common. Some authors use the opposite sign in the exponent and write $\exp i(kz - \omega t)$, usually with i instead of j for the square root of -1 . This expression has the advantage of reducing to $\exp ikz$ after the $\exp(-i\omega t)$ factor is suppressed. With that convention, one must substitute $-i$ for j in all phasor calculations. For example, the impedance of an inductor becomes $R - i\omega L$, instead of $R + j\omega L$. Electrical engineers use $\exp(j\omega t - \gamma z)$, where γ , the *propagation constant*, is equal to jk .

Then Maxwell's equations 27-24 to 27-27 reduce to

$$-jk\hat{z} \cdot \mathbf{E} = 0, \quad -jk\hat{z} \times \mathbf{E} = -j\omega\mu\mathbf{H}, \quad (28-5)$$

$$-jk\hat{z} \cdot \mathbf{H} = 0, \quad -jk\hat{z} \times \mathbf{H} = \sigma\mathbf{E} + j\omega\epsilon\mathbf{E} \quad (28-6)$$

and then to

$$\hat{z} \cdot \mathbf{E} = 0, \quad \mathbf{E} = -\frac{k}{\omega\epsilon + j\sigma}\hat{z} \times \mathbf{H}, \quad (28-7)$$

$$\hat{z} \cdot \mathbf{H} = 0, \quad \mathbf{H} = \frac{k}{\omega\mu}\hat{z} \times \mathbf{E}. \quad (28-8)$$

It follows that \mathbf{E} and \mathbf{H} are transverse and orthogonal. Figure 28-2 shows the relative orientations of \mathbf{E} , \mathbf{H} , and $\mathbf{k} = k\hat{z}$. Observe that $\mathbf{E} \times \mathbf{H}$ points in the direction of propagation.

28.2.2 The Characteristic Impedance Z of a Medium

The ratio E/H is the *characteristic impedance* Z of the medium of propagation:

$$Z = \frac{E}{H} = \frac{k}{\omega\epsilon - j\sigma} = \frac{\omega\mu}{k}. \quad (28-9)$$

28.2.3 The Wave Number k

The value of k^2 follows from the above equation:

$$k^2 = \omega^2\epsilon\mu - j\omega\sigma\mu = \omega^2\epsilon\mu\left(1 - j\frac{\sigma}{\omega\epsilon}\right), \quad (28-10)$$

$$= \omega^2\epsilon_0\mu_0\epsilon_r\mu_r\left(1 - j\frac{\sigma}{\omega\epsilon}\right). \quad (28-11)$$

The σ terms account for Joule losses and attenuation.

28.2.4 The Wave Equations

We found the nonhomogenous wave equations for \mathbf{E} and \mathbf{B} in Sec. 27.9:

$$\nabla^2\mathbf{E} - \epsilon_0\mu_0\frac{\partial^2\mathbf{E}}{\partial t^2} = \frac{\nabla\rho}{\epsilon_0} + \mu_0\frac{\partial\mathbf{J}}{\partial t}, \quad (28-12)$$

$$\nabla^2\mathbf{B} - \epsilon_0\mu_0\frac{\partial^2\mathbf{B}}{\partial t^2} = \mu_0\nabla \times \mathbf{J}. \quad (28-13)$$

We now apply the rule of Sec. 27.1 and Eq. 28-4 to obtain the equivalent equations for a medium ϵ , μ , σ . We again set $\rho_f = 0$. Then

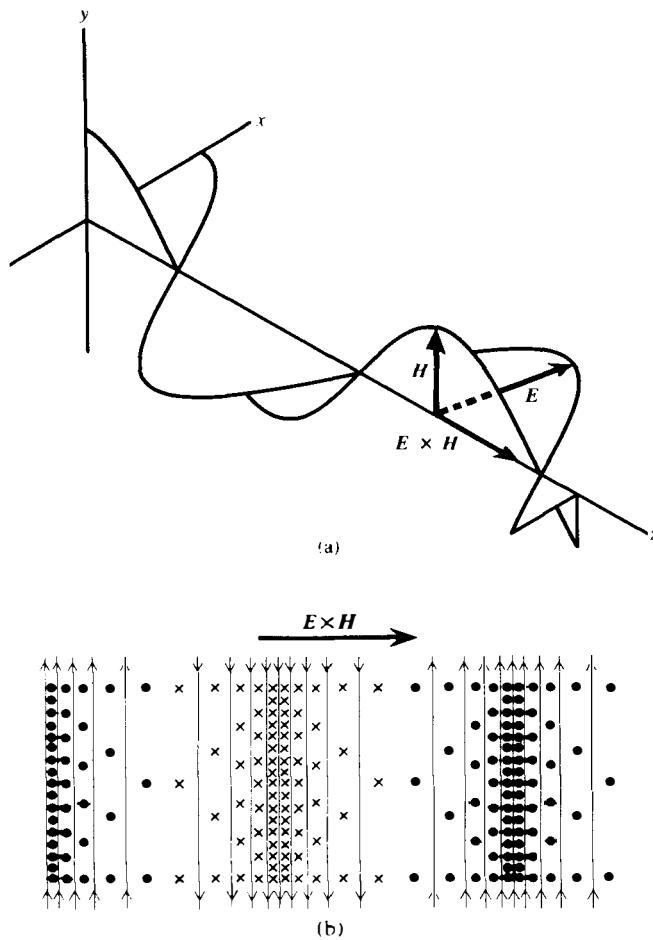


Fig. 28-2. The \mathbf{E} and \mathbf{H} vectors for a plane electromagnetic wave traveling in free space in the positive direction along the z -axis. (a) The fields \mathbf{E} and \mathbf{H} as functions of z at a particular moment. The two vectors are orthogonal and in phase. (b) Lines of \mathbf{E} (arrowheads), as seen when looking down on the xz -plane. The dots represent lines of \mathbf{H} coming out of the paper, and the crosses lines going into the paper. The vector $\mathbf{E} \times \mathbf{H}$ points everywhere in the direction of propagation.

$$\nabla^2 \mathbf{E} - \epsilon \mu \frac{\partial^2 \mathbf{E}}{\partial t^2} = \mu \frac{\partial \mathbf{J}_f}{\partial t} = \mu \sigma \frac{\partial \mathbf{E}}{\partial t}, \quad (28-14)$$

$$\nabla^2 \mathbf{B} - \epsilon \mu \frac{\partial^2 \mathbf{B}}{\partial t^2} = -\mu \sigma \nabla \times \mathbf{E} = \mu \sigma \frac{\partial \mathbf{B}}{\partial t}. \quad (28-15)$$

It is the custom to write these wave equations in the form

$$\nabla^2 \mathbf{E} - \epsilon\mu \frac{\partial^2 \mathbf{E}}{\partial t^2} - \mu\sigma \frac{\partial \mathbf{E}}{\partial t} = 0, \quad (28-16)$$

$$\nabla^2 \mathbf{B} - \epsilon\mu \frac{\partial^2 \mathbf{B}}{\partial t^2} - \mu\sigma \frac{\partial \mathbf{B}}{\partial t} = 0, \quad (28-17)$$

or,

$$\nabla^2 \mathbf{H} - \epsilon\mu \frac{\partial^2 \mathbf{H}}{\partial t^2} - \mu\sigma \frac{\partial \mathbf{H}}{\partial t} = 0. \quad (28-18)$$

Then, from Sec. 28.2.1,

$$(-k^2 + \omega^2 \epsilon\mu - j\omega\sigma\mu) \mathbf{E} = 0, \quad (28-19)$$

and similarly for \mathbf{H} . The expression enclosed in parentheses is equal to zero, from Eq. 28-10.

28.3 UNIFORM PLANE WAVES IN FREE SPACE

In free space, $\epsilon_r = 1$, $\mu_r = 1$, $\sigma = 0$, there is no attenuation, and from Eq. 28-19,

$$k = \frac{1}{\lambda_0} \quad (28-20)$$

$$= \omega(\epsilon_0\mu_0)^{1/2}. \quad (28-21)$$

From Eq. 28-2 the speed of light is

$$c = \frac{\omega}{k} = \frac{1}{(\epsilon_0\mu_0)^{1/2}} = 2.99792458 \times 10^8 \text{ meters/second.} \quad (28-22)$$

This equation is remarkable. It links three basic constants of electromagnetism: the speed of light c , the permittivity of free space ϵ_0 that appears in the expression for the Coulomb force, and the permeability of free space μ_0 from the magnetic force law.

Since μ_0 is, by definition, *exactly* equal to $4\pi \times 10^{-7}$, the value of ϵ_0 follows from the value of c :

$$\epsilon_0 = \frac{1}{\mu_0 c^2} = 8.854187817 \times 10^{-12} \text{ farad/meter.} \quad (28-23)$$

The *characteristic impedance of the vacuum* is

$$Z_0 = \frac{E}{H} = \frac{k}{\omega\epsilon_0} = \frac{\omega\mu_0}{k} = \frac{1}{\epsilon_0 c} = \mu_0 c = \left(\frac{\mu_0}{\epsilon_0}\right)^{1/2} \quad (28-24)$$

$$= 3.767303 \times 10^2 \approx 377 \text{ ohms.} \quad (28-25)$$

Thus, since $B = \mu_0 H$ in free space,

$$\frac{E}{B} = \frac{1}{(\epsilon_0 \mu_0)^{1/2}} = c, \quad \text{or} \quad E = Bc. \quad (28-26)$$

The \mathbf{E} and \mathbf{H} vectors in free space are in phase because the characteristic impedance of free space is real.

The electric and magnetic energy densities[†] are equal:

$$\frac{\epsilon_0 E^2/2}{\mu_0 H^2/2} = \frac{\epsilon_0}{\mu_0} \left(\frac{\mu_0}{\epsilon_0} \right) = 1. \quad (28-27)$$

At any instant the total energy density fluctuates with z as in Fig. 28-3, and its time-averaged value at any point is

$$\mathcal{E}' = \frac{\epsilon_0 E_{\text{rms}}^2}{2} + \frac{\mu_0 H_{\text{rms}}^2}{2} = \epsilon_0 E_{\text{rms}}^2 = \mu_0 H_{\text{rms}}^2. \quad (28-28)$$

Abandoning the phasor notation for a moment,

$$\mathbf{E} = \mathbf{E}_m \cos(\omega t - kz), \quad \mathbf{H} = \mathbf{H}_m \cos(\omega t - kz). \quad (29-29)$$

The magnitude of the *Poynting vector* is

$$|\mathcal{S}| = |\mathbf{E} \times \mathbf{H}| = E_m H_m \cos^2(\omega t - kz). \quad (28-30)$$

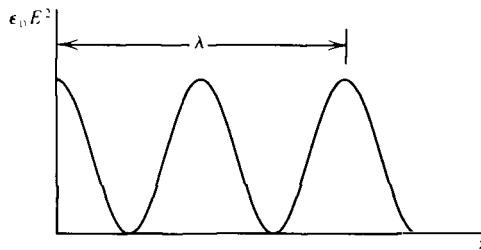


Fig. 28-3. The energy density $\epsilon_0 E^2$, or $\mu_0 H^2$, as a function of z , at $t = 0$, for a plane wave travelling along the z -axis in free space.

[†] We have shown that the electric and magnetic energy densities are $\epsilon_0 E^2/2$ and $\mu_0 H^2/2$, respectively, for slowly varying phenomena. We have not shown that these expressions apply to any field. They are, in fact, in agreement with experiments on the energy flux in electromagnetic waves.

We shall see in Sec. 28.6 that the Poynting vector, when integrated over a surface, yields the power flow through that surface. Power flows in the direction of \mathcal{S} .

Returning to phasors, the time-averaged Poynting vector is (Sec. 2.4)

$$\mathcal{S}_{av} = \frac{1}{2} \operatorname{Re} (\mathbf{E} \times \mathbf{H}^*) \quad (28-31)$$

and, for a uniform plane wave in free space,

$$\mathcal{S}_{av} = \frac{1}{2} \operatorname{Re} (EH^*) \hat{\mathbf{z}} \quad (28-32)$$

$$= \frac{1}{2} c \epsilon_0 |E_m|^2 \hat{\mathbf{z}} = c \epsilon_0 E_{rms}^2 \hat{\mathbf{z}} = \frac{E_{rms}^2}{Z_0} \hat{\mathbf{z}} \quad (28-33)$$

$$\approx \frac{E_{rms}^2}{377} \hat{\mathbf{z}} \quad \text{watts/meter}^2. \quad (28-34)$$

This is the time-averaged total energy density $\epsilon_0 E_{rms}^2$, multiplied by the speed of light c .

Example

THE \mathbf{E} AND \mathbf{B} FIELDS IN A LASER BEAM

Most lasers operate at powers of the order of milliwatts. However, there exist a few lasers that are vastly more powerful. One of these supplies a pulsed beam of 27 terawatts (27,000 gigawatts!) of *peak* power over a circular area 0.1 millimeter in radius. Then

$$\mathcal{S} = \frac{27 \times 10^{12}}{\pi \times (10^{-4})^2} \approx 9 \times 10^{20} \text{ watts/meter}^2, \quad (28-35)$$

$$E_{rms} = (377 \times 9 \times 10^{20})^{1/2} \approx 6 \times 10^{11} \text{ volts/meter}. \quad (28-36)$$

This is an enormous field: 60 volts over the diameter of an atom ($\approx 10^{-10}$ meter)! Air breaks down at fields of about 3×10^6 volts/meter. Also,

$$B_{rms} = \frac{E_{rms}}{c} \approx 2 \times 10^3 \text{ teslas}, \quad (28-37)$$

or about 2000 times the field between the pole pieces of a powerful electromagnet.

28.4 UNIFORM PLANE WAVES IN NONCONDUCTORS

The situation here is the same as in free space, with ϵ and μ replacing ϵ_0 and μ_0 . The phase velocity is now

$$v = \frac{1}{(\epsilon\mu)^{1/2}} = \frac{c}{(\epsilon_r\mu_r)^{1/2}} = \frac{c}{n}, \quad (28-38)$$

where n is the *index of refraction*:

$$n = (\epsilon_r\mu_r)^{1/2}. \quad (28-39)$$

The phase velocity v is less than in free space, since both ϵ_r and μ_r are larger than unity. In nonmagnetic media,

$$n = \epsilon_r^{1/2}. \quad (28-40)$$

As we saw in Sec. 10.1.2, ϵ_r is a function of the frequency, so n is also frequency-dependent. As a rule, tables of n apply to optical frequencies ($\approx 10^{15}$ hertz), whereas tables of ϵ_r apply at much lower frequencies, at best up to about 10^{10} hertz. Pairs of values drawn from such tables do not therefore satisfy the above equation.

The characteristic impedance of the medium is

$$Z = \frac{E}{H} = \left(\frac{\mu}{\epsilon}\right)^{1/2} = 377 \left(\frac{\mu_r}{\epsilon_r}\right)^{1/2} \text{ ohms.} \quad (28-41)$$

The electric and magnetic energy densities are again equal:

$$\frac{\epsilon E^2/2}{\mu H^2/2} = 1, \quad (28-42)$$

and the time-averaged energy density is

$$\mathcal{E}'_{av} = \frac{\epsilon E_{rms}^2}{2} + \frac{\mu H_{rms}^2}{2} = \epsilon E_{rms}^2 = \mu H_{rms}^2. \quad (28-43)$$

The Poynting vector $\mathbf{E} \times \mathbf{H}$ points again in the direction of propagation, and

$$\mathcal{S}_{av} = \frac{1}{2} \operatorname{Re}(EH^*)\hat{z} = \left(\frac{\epsilon}{\mu}\right)^{1/2} E_{rms}^2 \hat{z} \quad (28-44)$$

$$\mathcal{S}_{av} \approx \frac{(\epsilon_r/\mu_r)^{1/2} E_{rms}^2}{377} \hat{z} \quad \text{watts/meter}^2 \quad (28-45)$$

$$= \frac{1}{\epsilon\mu} \epsilon E_{rms}^2 \hat{z} = v \epsilon E_{rms}^2 \hat{z}. \quad (28-46)$$

The time-averaged Poynting vector is again equal to the phase velocity multiplied by the time-averaged energy density.

Example

THE \mathbf{E} AND \mathbf{B} FIELDS OF A LASER BEAM IN GLASS

There is no point in referring here to the laser beam of Sec. 28.3 because such a beam would instantly vaporize glass. Say we have a 1.0-milliwatt beam with a diameter of 1.0 millimeter in glass whose index of refraction is 1.5. Then

$$\mathcal{S} = \frac{10^{-3}}{\pi \times (5 \times 10^{-4})^2} = 1.3 \times 10^3 \text{ watts/meter}^2, \quad (28-47)$$

$$E_{\text{rms}} = \left[\left(\frac{\mu}{\epsilon} \right)^{1/2} \mathcal{S} \right]^{1/2} = \left[\left(\frac{4\pi \times 10^{-7}}{1.5^2 \times 8.85 \times 10^{-12}} \right)^{1/2} \times 1.3 \times 10^3 \right]^{1/2} \quad (28-48)$$

$$= 5.7 \times 10^2 \text{ volts/meter}, \quad (28-49)$$

$$B_{\text{rms}} = \mu_0 H_{\text{rms}} = \mu_0 \left(\frac{\epsilon}{\mu_0} \right)^{1/2} E_{\text{rms}} = (\epsilon \mu_0)^{1/2} E_{\text{rms}} \quad (28-50)$$

$$= (1.5^2 \times 8.85 \times 10^{-12} \times 4\pi \times 10^{-7})^{1/2} \times 5.7 \times 10^2 \\ = 2.9 \times 10^{-6} \text{ tesla}. \quad (28-51)$$

28.5 UNIFORM PLANE WAVES IN CONDUCTORS

We start at the point where we left off in Sec. 28.2.4.

28.5.1 The Complex Wave Number $k = \beta - j\alpha$

In Sec. 28.2.3 we found that in a conducting medium

$$k^2 = \frac{\epsilon_r \mu_r}{\lambda_0^2} \left(1 - j \frac{\sigma}{\omega \epsilon} \right), \quad (28-52)$$

so k is complex. It is the custom to set

$$k = \beta - j\alpha \quad \text{and then} \quad \mathbf{E} = \mathbf{E}_m \exp(-\alpha z) \exp j(\omega t - \beta z), \quad (28-53)$$

where both α and β are positive.

The quantity $1/\alpha$ is the *attenuation distance* or the *skin depth* δ over which the amplitude decreases by a factor of e . The real part β of k is the inverse of λ :

$$\alpha = \frac{1}{\delta}, \quad (28-54)$$

$$\beta = \frac{1}{\lambda} = \frac{2\pi}{\lambda}, \quad (28-55)$$

and the phase velocity is

$$v = \frac{\omega}{\beta}. \quad (28-56)$$

Let us find α and β in terms of ϵ_r , μ_r , σ , and λ_0 . First we set

$$\mathcal{D} = \frac{\sigma}{\omega\epsilon} = \left| \frac{\sigma E}{\epsilon \partial E / \partial t} \right| = \left| \frac{\sigma E}{\partial D / \partial t} \right| \approx 377 \frac{\sigma \lambda_0}{\epsilon_r}. \quad (28-57)$$

This is the magnitude of the conduction current density, divided by the magnitude of the displacement current density. As a rule, \mathcal{D} (for “dissipation”), is written $\tan l$, as in Sec. 10.1.1:

$$\mathcal{D} = \tan l, \quad (28-58)$$

where l is here the *loss angle* of the medium, but we use \mathcal{D} for conciseness.

The permittivity ϵ that appears above is the real part $\epsilon' \epsilon_0$ (Sec. 10.1.1). One can account for conductivity *either* by means of a complex permittivity $(\epsilon'_r - j\epsilon''_r)\epsilon_0$ or by means of a real permittivity and a conductivity σ , where $\sigma = \omega \epsilon''_r \epsilon_0$, again as in Sec. 10.1.1. Thus

$$\mathcal{D} = \frac{\epsilon''_r}{\epsilon'_r}. \quad (28-59)$$

This quantity, like ϵ' and ϵ'' , is always positive.

If $\mathcal{D} \ll 1$, the medium is a good dielectric; if $\mathcal{D} \gg 1$, the medium is a good conductor. For common types of good conductor, $\sigma \approx 10^7$ ($\sigma = 5.8 \times 10^7$ for copper) and $\epsilon_r \approx 1$ (Sec. 4.3.6). You will remember from Sec. 4.3.6 that ϵ/σ is the relaxation time of a medium.

Thus

$$k^2 = (\beta - j\alpha)^2 = \left(\frac{\epsilon_r \mu_r}{\lambda_0^2} \right) (1 - j\mathcal{D}), \quad (28-60)$$

and

$$\alpha = \frac{1}{\lambda_0} \left(\frac{\epsilon_r \mu_r}{2} \right)^{1/2} [(1 + \mathcal{D}^2)^{1/2} - 1]^{1/2}, \quad (28-61)$$

$$\beta = \frac{1}{\lambda_0} \left(\frac{\epsilon_r \mu_r}{2} \right)^{1/2} [(1 + \mathcal{D}^2)^{1/2} + 1]^{1/2}, \quad (28-62)$$

$$k = \frac{(\epsilon_r \mu_r)^{1/2}}{\lambda_0} (1 + \mathcal{D}^2)^{1/4} \exp \left(-j \arctan \frac{\alpha}{\beta} \right). \quad (28-63)$$

The argument of the exponential function is correct because β is positive (Sec. 2.1).

In a *low-loss dielectric* \mathcal{D} is small, and

$$\alpha \approx \frac{(\epsilon_r \mu_r)^{1/2} \mathcal{D}}{2\lambda_0} = \left(\frac{\mu_r}{\epsilon_r} \right)^{1/2} \frac{\sigma c \mu_0}{2}, \quad (28-64)$$

$$\beta \approx \frac{(\epsilon_r \mu_r)^{1/2}}{\lambda_0}, \quad v = \frac{\omega}{\beta} \approx \frac{c}{(\epsilon_r \mu_r)^{1/2}}. \quad (28-65)$$

In such media the conductivity hardly affects the phase velocity, but it gives rise to an attenuation that is independent of the frequency.

In a *good conductor* $\mathcal{D} \gg 1$ and

$$k^2 = -j\mathcal{D} \frac{\epsilon_r \mu_r}{\lambda_0^2} = -j \frac{\sigma}{\omega \epsilon} \epsilon_r \mu_r \omega^2 \epsilon_0 \mu_0 = -j \sigma \mu \omega, \quad (28-66)$$

$$k = \left(\frac{\sigma \mu \omega}{2} \right)^{1/2} (1-j), \quad (28-67)$$

$$\alpha = \beta = \left(\frac{\sigma \mu \omega}{2} \right)^{1/2}. \quad (28-68)$$

The index of refraction of a *good conductor*

$$n = \frac{c}{\omega/\beta} = \frac{c\beta}{\omega} = c \left(\frac{\sigma \mu}{2\omega} \right)^{1/2} \quad (28-69)$$

is a large quantity. It is 1.1×10^8 for copper at 1 megahertz.

28.5.2 The Characteristic Impedance Z of a Conductor

The characteristic impedance of a conducting medium is complex:

$$Z = \frac{E}{H} = \frac{k}{\omega \epsilon - j\sigma} = \frac{\omega \mu}{k} \quad (28-70)$$

$$= \left(\frac{\mu}{\epsilon} \right)^{1/2} \frac{\exp j \arctan (\alpha/\beta)}{(1 + \mathcal{D}^2)^{1/4}} \approx 377 \left(\frac{\mu_r}{\epsilon_r} \right)^{1/2} \frac{\exp j \arctan (\alpha/\beta)}{(1 + \mathcal{D}^2)^{1/4}} \text{ ohms} \quad (28-71)$$

as we saw in Sec. 28.2.2. This means that E and H are not in phase:

$$\frac{E}{H} = \frac{\omega \mu}{\beta - j\alpha}, \quad (28-72)$$

where α and β are both positive. So E leads H by the angle

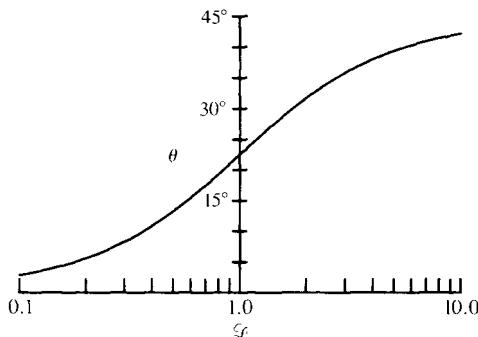


Fig. 28-4. The phase of E with respect to H as a function of \mathcal{D} in a wave traveling through a conducting medium.

$$\theta = \arctan \frac{\alpha}{\beta} \quad (28-73)$$

Figure 28-4 shows θ as a function of \mathcal{D} .

Therefore,

$$\mathbf{E} = \mathbf{E}_m \exp(-\alpha z) \exp j(\omega t - \beta z), \quad (28-74)$$

$$\mathbf{H} = \mathbf{H}_m \exp(-\alpha z) \exp j(\omega t - \beta z - \theta), \quad (28-75)$$

with

$$\frac{\mathbf{E}_m}{\mathbf{H}_m} = \frac{\omega \mu}{k} = \left(\frac{\mu}{\epsilon} \right)^{1/2} \frac{1}{(1 + \mathcal{D}^2)^{1/4}} = \left(\frac{\mu_0}{\epsilon_0} \right)^{1/2} \left(\frac{\mu_r}{\epsilon_r} \right)^{1/2} \frac{1}{(1 + \mathcal{D}^2)^{1/4}} \quad (28-76)$$

$$\approx 377 \left(\frac{\mu_r}{\epsilon_r} \right)^{1/2} \frac{1}{(1 + \mathcal{D}^2)^{1/4}} \text{ ohms.} \quad (28-77)$$

From Eq. 28-8, \mathbf{E} and \mathbf{H} are orthogonal in a linearly polarized wave. If the wave is *not* linearly polarized, then the vectors \mathbf{E} and \mathbf{H} are not necessarily orthogonal.

28.5.3 The Energy Densities

The time-averaged electric and magnetic energy densities are in the ratio

$$\frac{\mathcal{E}'_e}{\mathcal{E}'_m} = \frac{\epsilon E_{\text{rms}}^2 / 2}{\mu H_{\text{rms}}^2 / 2} = \frac{1}{(1 + \mathcal{D}^2)^{1/2}}. \quad (28-78)$$

There is less electric energy than magnetic energy because the conductivity both decreases \mathbf{E} and adds a conduction current to the displacement current, which increases \mathbf{H} .

The time-averaged total energy density is

$$\frac{1}{2}(\epsilon E_{\text{rms}}^2 + \mu H_{\text{rms}}^2) \exp(-2\alpha z) = \frac{1}{2}(\epsilon E_{\text{rms}}^2)[1 + (1 + \mathcal{D}^2)^{1/2}] \exp(-2\alpha z). \quad (28-79)$$

28.6 THE POYNTING THEOREM

We referred to the Poynting vector

$$\mathcal{S} = \mathbf{E} \times \mathbf{H} \quad (28-80)$$

in previous sections, but we said very little about it. We only stated that it is equal to the power density in an electromagnetic wave, and that it points in the direction of propagation.

The Poynting vector is of great theoretical and practical interest. Its significance follows from the Poynting theorem that we now prove.

First, we have the vector identity

$$\nabla \cdot (\mathbf{E} \times \mathbf{H}) = \mathbf{H} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{H}). \quad (28-81)$$

In a HILS medium, Eqs. 27-20 to 27-23 apply, and then

$$\nabla \cdot (\mathbf{E} \times \mathbf{H}) = -\mathbf{H} \cdot \mu \frac{\partial \mathbf{H}}{\partial t} - \mathbf{E} \cdot \left(\epsilon \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J}_f \right) \quad (28-82)$$

$$= -\frac{\partial}{\partial t} \left(\frac{\epsilon E^2}{2} + \frac{\mu H^2}{2} \right) - \mathbf{E} \cdot \mathbf{J}_f. \quad (28-83)$$

We now change the signs, integrate over a volume v of finite extent and of surface area \mathcal{A} , and finally apply the divergence theorem on the left. This yields the *Poynting theorem*:

$$-\int_{\mathcal{A}} (\mathbf{E} \times \mathbf{H}) \cdot d\mathcal{A} = \frac{d}{dt} \int_v \left(\frac{\epsilon E^2}{2} + \frac{\mu H^2}{2} \right) dv + \int_v \mathbf{E} \cdot \mathbf{J}_f dv. \quad (28-84)$$

The first integral on the right gives the increase in the electric and magnetic energy densities inside the volume v , per unit time. The second gives that part of the field energy that dissipates as heat, again per unit time. Then the term on the left, with its negative sign, must represent the rate at which electromagnetic energy flows *into* the volume v .

Then the integral

$$\int_{\mathcal{A}} \mathcal{S} \cdot d\mathcal{A} = \int_{\mathcal{A}} (\mathbf{E} \times \mathbf{H}) \cdot d\mathcal{A} \quad (28-85)$$

is the total power flowing *out* of a closed surface of area \mathcal{A} .

The Poynting theorem therefore simply states that there is conservation of energy in electromagnetic fields. It is a proof of the validity of Eq. 27-23, and hence of Eq. 27-4.

For a uniform, plane, and linearly polarized wave in conducting material, the time-averaged magnitude of the Poynting vector is

$$\mathcal{S}_{av} = \frac{1}{2} \operatorname{Re} \{ [E_m \exp(-\alpha z) \exp j(\omega t - \beta z) \\ \times H_m \exp(-\alpha z) \exp j(-\omega t + \beta z + \theta)] \} \quad (28-86)$$

$$= \frac{1}{2} E_m H_m \cos \theta \exp(-2\alpha z), \quad (28-87)$$

where θ is defined as in Sec. 28.5.2 and

$$\cos \theta = \frac{\beta}{(\alpha^2 + \beta^2)^{1/2}}. \quad (28-88)$$

We found the ratio E_m/H_m in the previous section. If we eliminate H_m , then

$$\mathcal{S}_{av} = \left(\frac{\epsilon}{\mu} \right)^{1/2} (1 + \mathcal{D}^2)^{1/4} E_{rms}^2 \cos \theta \exp(-2\alpha z) \quad (28-89)$$

$$\approx \frac{1}{377} \left(\frac{\epsilon_r}{\mu_r} \right)^{1/2} (1 + \mathcal{D}^2)^{1/4} E_{rms}^2 \cos \theta \exp(-2\alpha z). \quad (28-90)$$

You can easily show that

$$\mathcal{S}_{av} = (\text{time-averaged energy density}) \times (\text{phase velocity}) \quad (28-91)$$

28.7 SUMMARY

In a *uniform plane wave* the wave fronts, which are surfaces of uniform phase, are also surfaces of uniform amplitude.

Linearly polarized uniform plane electromagnetic waves traveling either in free space or in a HILLS nonconductor or conductor possess the following properties.

- (1) The vectors \mathbf{E} and \mathbf{H} are *transverse* and *orthogonal*.
- (2) The *Poynting vector* $\mathbf{E} \times \mathbf{H}$ points in the direction of propagation.
- (3) The magnitude of the Poynting vector, averaged over time, gives the power flow per square meter in the wave:

$$\mathcal{S}_{av} = \frac{1}{2} \operatorname{Re} (\mathbf{E} \times \mathbf{H}^*). \quad (28-31)$$

(4) The power density is equal to the energy density multiplied by the phase velocity.

(5) The ratio E/H is equal to the *characteristic impedance* Z of a medium.

In *free space*,

$$c = \frac{1}{(\epsilon_0 \mu_0)^{1/2}} = 2.99792458 \times 10^8 \text{ meters/second}, \quad (28-22)$$

$$Z_0 \approx 377 \text{ ohms}, \quad (28-25)$$

$$\mathcal{S}_{av} \approx \frac{E_{rms}^2}{377} \hat{z} \quad \text{watts/meter}^2. \quad (28-34)$$

In *nonconductors*,

$$v = \frac{c}{(\epsilon_r \mu_r)^{1/2}} = \frac{c}{n}, \quad (28-38)$$

where n is the *index of refraction*. Also,

$$Z \approx 377 \left(\frac{\mu_r}{\epsilon_r} \right)^{1/2} \text{ ohms}. \quad (28-41)$$

In *conductors* there is attenuation and $k = \beta - j\alpha$, where

$$\alpha = \frac{1}{\lambda_0} \left(\frac{\epsilon_r \mu_r}{2} \right)^{1/2} [(1 + \mathcal{D}^2)^{1/2} - 1]^{1/2}, \quad (28-61)$$

$$\beta = \frac{1}{\lambda_0} \left(\frac{\epsilon_r \mu_r}{2} \right)^{1/2} [(1 + \mathcal{D}^2)^{1/2} + 1]^{1/2}, \quad (28-62)$$

$$\mathcal{D} = \frac{\sigma}{\omega \epsilon}, \quad (28-57)$$

$$Z = \frac{\omega \mu}{k}. \quad (28-70)$$

The *electric* and *magnetic energy densities* are equal in free space and in nonconductors. In conductors the magnetic energy density is larger than the electric energy density.

The *Poynting theorem* is a statement of the conservation of energy in an electromagnetic field.

PROBLEMS

- 28-1. (28.2.1) A general theorem for electromagnetic fields in free space
 (a) Show that, for any electromagnetic field in a vacuum

$$\nabla \cdot (\mathbf{E} \times \dot{\mathbf{E}} + c^2 \mathbf{B} \times \dot{\mathbf{B}}) = -\frac{\partial}{\partial t} (\dot{\mathbf{E}} \cdot \mathbf{B} - \mathbf{E} \cdot \dot{\mathbf{B}}),$$

where the dots above \mathbf{E} and \mathbf{B} indicate partial differentiation with respect to time.

This equation has the form of a conservation law. The spatial density of the conserved quantity appears between parentheses on the right. This is expressed in volts squared per cubic meter. The flux density of this conserved quantity is the quantity between parentheses on the left.

(b) Show that the flux vanishes in the field of a linearly polarized wave.

(c) Show that, in the field of a circularly polarized wave, the flux does not vanish, that it is proportional to the frequency, and that it contrary to the direction of propagation if the \mathbf{E} and \mathbf{B} vectors rotate clockwise for an observer who looks at the source, and in the direction of propagation if \mathbf{E} and \mathbf{B} rotate in the opposite direction.

28-2. (28.4) The phase and group indices of refraction

The *phase index of refraction*, often called the *phase index*, is c/v_p , where v_p is the phase velocity. The *group index* is similarly c/v_g , where $v_g = 1/(d\beta/d\omega)$ is the group velocity (Sec. 29.2.6).

Show that $m = n + \omega dn/d\omega$.

28-3. (28.5) The skin depth as a function of frequency in low-conductivity materials

(a) Plot on a single graph the log-log curves of the skin depth as a function of the frequency from $f = 1$ to $f = 10^5$, for σ equal to 10^{-2} , 10^{-4} , 10^{-6} and for $\epsilon_r = 1$ and $\epsilon_r = 10$. Set $\mu_r = 1$. The skin depth will vary from about 10 to 10^6 meters.

(b) Show that, in nonmagnetic good conductors for which $\sigma \geq 50\omega\epsilon_r$, $\delta \approx 503/(f\sigma)^{1/2}$.

(c) Show that, in nonmagnetic poor conductors for which $\sigma \leq 0.1\omega\epsilon_r$, $\delta \approx 5.3 \times 10^{-3}\epsilon_r^{1/2}/\sigma$.

28-4. (28.5.1) The optical properties of metals

At optical frequencies ($f \approx 10^{15}$ hertz) and above, the values of σ , ϵ_r , and μ_r bear no relation to the values measured at lower frequencies. For metals, both β and α are of the order of $3/\lambda_0$, within approximately a factor of 10 either way, and $\beta \neq \alpha$. For aluminum at $\lambda_0 = 650$ nanometers, $\beta\lambda_0 = 1.3$ and $\alpha\lambda_0 = 7.11$.

(a) Calculate λ and δ .

(b) Calculate λ and δ from Sec. 28.5.1 and Table 29-1.

28-5. (28.5.2) λ , δ , and Z in poor conductors

Show that, in a medium that is only slightly conducting ($\mathcal{D} \ll 1$),

(a) $\lambda = (1 - \mathcal{D}^2/8)\lambda_{\mathcal{D}=0}$

(b) $\delta = 2^{1/2}\chi_{\mathcal{D}=0}/\mathcal{D}$

(c) $Z = (1 - \mathcal{D}^2/4) \exp j[\arctan(\mathcal{D}/2)] Z_{\mathcal{D}=0}$

28-6. (28.5.2) Alternate expressions for the characteristic impedance of a conducting medium

Show that the characteristic impedance E/H of a conducting medium is also given by these two other expressions:

$$Z = \frac{\beta + j\alpha}{(\sigma^2 + \omega^2\epsilon^2)^{1/2}} = \left(\frac{j\omega\mu}{\sigma + j\omega\epsilon} \right)^{1/2}.$$

28-7. (28.5.2) The characteristic impedance of ultra-low-loss polyethylene

The ultra-low-loss polyethylene that serves as insulator in submarine coaxial cables has a loss angle of 50 microradians and a relative permittivity of 2.26. Calculate its characteristic impedance at 45 megahertz.

28-8. (28.6) The Poynting vector in the field of a resistive wire carrying a current

A long, straight wire of radius a and resistance R' ohms/meter carries a current I .

(a) Calculate the Poynting vector at the surface, and explain.

(b) Calculate the Poynting vector both outside and inside the wire. Explain.

28-9. (28.6) The Poynting vector in a capacitor

A thin, air-insulated parallel-plate capacitor has circular plates of radius R , separated by a distance s . A constant current I charges the plates through thin wires along the axis of symmetry.

(a) Find the value of E between the plates as a function of the time. Assume a uniform E . Show the direction of E on a figure.

(b) The magnetic field is the sum of two terms, H_w , related to the current in the wire, and H_p , related to the current in the plates. The latter current deposits charges on the inside surfaces of the plates.

Find H_w , H_p , and H . Use cylindrical coordinates with the z -axis along the wire and in the direction of the current. To calculate H_p , apply Ampère's circuital law to each plate. You should find that the magnetic fields tend to infinity as $\rho \rightarrow 0$. This is simply because we have assumed infinitely thin wires and plates. Show the directions of H_w , H_p , and H on your figure.

(c) Do E and H satisfy Maxwell's equations? You should find that one of our assumptions is incorrect

(d) Find $E \times H$.

(e) Find the electric and magnetic energy densities inside a radius ρ .

You should find that the magnetic energy density is negligible if $\rho^2/t^2 \ll c^2$. This condition applies because we have assumed that the capacitor charges up slowly. If it charged very quickly, then there would be a wave of E and H in the capacitor, E would not be uniform, and the above calculation would be invalid.

(f) Now relate the Poynting vector at ρ to the electric energy inside ρ .

(g) Draw a sketch showing E , H , and $E \times H$ vectors at various points inside and around the capacitor.

28-10. (28.6) The Poynting vector in a solenoid

A long solenoid of radius R and N' turns per meter carries a current I .

(a) The current increases. Calculate $S = E \times H$. (See example in Sec. 19.1.)

Sketch a cross section of the solenoid, showing the direction of the current and of S . Explain.

(b) Repeat with a decreasing current.

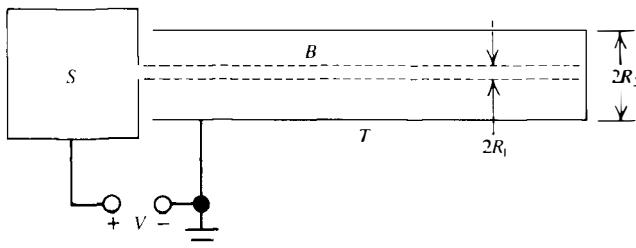


Fig. 28-5.

28-11. (28.6) Energy and power in a proton beam

Figure 28-5 shows a highly simplified diagram of a proton accelerator. A gas discharge within the source S ionizes hydrogen gas to produce protons. Some of the protons emerge through a hole and are focused into a beam B of radius R_1 inside a conducting tube of radius R_2 . The source is at a potential V , and the target is grounded.

To avoid needless complications, we assume that the charge density in the beam is uniform. We also assume that the velocity of the protons is much less than c : $v^2 \ll c^2$.

Calculate, in terms of the current I and the velocity v :

- the electric energy per meter \mathcal{E}'_E ;
- the magnetic energy per meter \mathcal{E}'_m ;
- the energy flux associated with the Poynting vector P_P ;
- the kinetic power P_k , or the flux of kinetic energy, disregarding P_P .

The existence of this Poynting vector is interesting. Because of the radial E , the voltage inside the beam is slightly positive. So the protons are not accelerated to the full voltage V , and the kinetic energy in the beam is slightly lower than VI . Most of the power flows down the tube as kinetic energy, and the rest flows as electromagnetic energy. The total power at any point along the tube and on the target is VI .

(e) Find the numerical values of these quantities for a 1.00-millampere, 1.00-MeV (megaelectronvolt) proton beam, with $R_1 = 1.00$ millimeter and $R_2 = 50.00$ millimeters.

28-12. (28.6) The solar wind

The *solar wind* is formed of highly ionized, and hence highly conducting, hydrogen that evaporates from the surface of the sun. In the plane of the earth's orbit, the magnetic field of the sun is approximately radial, pointing outward in certain regions and inward in others. Since the sun rotates (period of 27 days), while the plasma has a radial velocity, the lines of B are Archimedes spirals. This is the *garden-hose effect*. At the earth, the lines of B form an angle of about 45° with the sun-earth direction.

At the orbit of the earth the solar wind has a density of about 10^7 proton masses per cubic meter and a velocity of about 4×10^5 meters/second. The magnetic field of the sun is about 5×10^{-9} tesla.

(a) Show that, in a neutral ($\rho = 0$) plasma of conductivity σ and velocity v , Maxwell's equations become

$$\nabla \cdot \mathbf{E} = 0, \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \cdot \mathbf{B} = 0,$$

$$\nabla \times \mathbf{B} = \mu_0 \left[\sigma(\mathbf{E} + \mathbf{v} \times \mathbf{B}) + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right].$$

In a medium of infinite conductivity σ , $\mathbf{E} = -\mathbf{v} \times \mathbf{B}$. This is a satisfactory approximation for the solar wind.

(b) Show that the component of the plasma velocity \mathbf{v} that is normal to \mathbf{B} is given by $v_{\perp} = [\mathbf{B} \times (\mathbf{v} \times \mathbf{B})]/B^2$.

(c) Show that the Poynting vector is given by $\mathcal{S} = B^2 v_{\perp}/\mu_0$, or about 6 microwatts/meter². This is about 4×10^{-9} times the Poynting vector of solar radiation, which is about 1.4 kilowatts/meter². The Poynting vector of the solar wind is normal to the local \mathbf{B} .

(d) Show that the kinetic, magnetic, and electric energy densities of the solar wind are related as follows: $\mathcal{E}'_k \gg \mathcal{E}'_m \gg \mathcal{E}'_E$.

28-13. (28.6) The Poynting vector in an induction motor

In an induction motor, the stator generates a magnetic field that is perpendicular to, and that rotates about, the axis of symmetry (Prob. 18-6). The rotor is a cylinder of laminated iron (Prob. 25-7), with copper bars parallel to the axis and set in grooves in the cylindrical surface. Copper rings at each end of the rotor connect all the copper bars. The rotor is *not* connected to the source of electric current that feeds the motor.

As we shall see, the rotor tends to follow the rotating magnetic field. Figure 28-6 shows the principle of operation. To simplify the analysis, we suppose that the rotor is stationary and that a rotating electromagnet, represented here by its poles N and S, provides the rotating magnetic field.

(a) Draw a larger figure with wide air gaps, showing the direction of the induced currents in the bars and the direction of \mathbf{E} in the air gaps.

(b) The current in the rotor generates a magnetic field. Add arrows showing the direction of *that* \mathbf{H} , inside the rotor and in the air gaps.

(c) Now show Poynting vectors $\mathbf{E} \times \mathbf{H}$ in the air gaps. The field feeds power into the rotor.

(d) Now draw another figure showing the currents in the bars and a line of \mathbf{B} for the sum of the two magnetic fields.

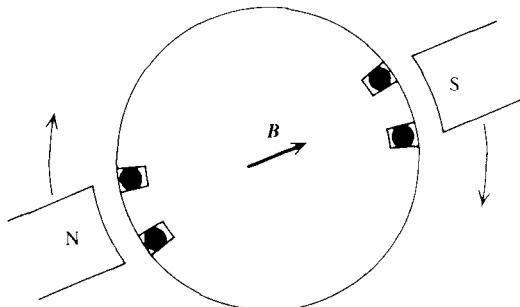


Fig. 28-6.

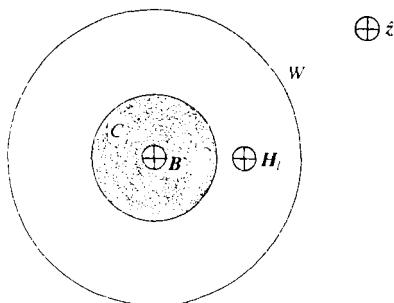


Fig. 28-7.

- (e) Show the direction of the magnetic forces on the bars.
- 28-14. (28.6) The Poynting vector in a transformer
- Figure 28-7 shows, in simplified form, a cross section of a transformer secondary. The field \mathbf{B} inside the core C , and the leakage field \mathbf{H}_l outside, both result from the currents in the windings and from the equivalent currents in the core. The secondary winding is W . Assume that \mathbf{B} and \mathbf{H}_l increase.
- Draw a larger figure and show, at one point between C and W , vectors \mathbf{A} , $\partial\mathbf{A}/\partial t$, and $\mathbf{E} = -\partial\mathbf{A}/\partial t$, disregarding the current in W . Show a vector \mathbf{E} at one point outside W .
 - Assume that the impedance of the secondary is a pure resistance R . Show the direction of the current I in W .
 - Show the direction of its field \mathbf{H} at points between C and W and outside W .
 - Show vectors $\mathbf{E} \times \mathbf{H}$.
 - How would the directions of the $\mathbf{E} \times \mathbf{H}$ vectors be affected if \mathbf{B} and \mathbf{H}_l decreased?
 - What is the time-averaged value of a vector $\mathbf{E} \times \mathbf{H}_l$?
 - Now let us calculate the power flow into the secondary. Assume that the secondary is a long solenoid of N turns and of length L . Disregard \mathbf{H}_l and set $\Phi = \Phi_m \exp j\omega t$ in the core. Integrate the Poynting vector over a cylindrical surface situated between the core and the winding, and show that the power flowing into the winding is $(N\omega\Phi_{rms})^2/R = V_{rms}^2/R$, where V is the voltage induced in the secondary winding.

CHAPTER 29

*PLANE ELECTROMAGNETIC WAVES II

Waves in Good Conductors and in Plasmas

29.1	UNIFORM PLANE ELECTROMAGNETIC WAVES IN GOOD CONDUCTORS. THE SKIN EFFECT	537
	<i>Example: PROPAGATION IN COPPER AT 1 MEGAHERTZ</i>	540
	<i>Example: JOULE LOSSES IN GOOD CONDUCTORS</i>	540
29.2	PLANE ELECTROMAGNETIC WAVES IN PLASMAS	542
29.2.1	THE CONDUCTIVITY OF A PLASMA	542
	<i>Example: THE IMPEDANCE OF A UNIT CUBE OF PLASMA</i>	543
29.2.2	THE CONDUCTION AND DISPLACEMENT CURRENT DENSITIES AND THE PLASMA ANGULAR FREQUENCY ω_p	544
	<i>Example: PLASMA-FILLED PARALLEL-PLATE CAPACITOR</i>	545
29.2.3	THE WAVE NUMBER k	545
29.2.4	PROPAGATION AT $f > f_p$	546
29.2.5	THE FIELD AT $f = f_p$	547
29.2.6	THE FIELD AT $f < f_p$	547
	<i>Example: THE TELECOMMUNICATION BLACKOUT UPON THE REENTRY OF A SPACE SHUTTLE</i>	548
	<i>Example: THE IONOSPHERE</i>	548
29.3	SUMMARY	549
	PROBLEMS	550

In Chap. 28 we deduced the main characteristics of a uniform plane electromagnetic wave traveling through a medium ϵ_r , μ_r , σ . We now apply this knowledge to two simple media, namely good conductors and low-density plasmas.

The propagation of electromagnetic waves in good conductors is peculiar in that the amplitude of the wave decreases by a factor of e in one radian length $\lambda = \lambda/2\pi$. The attenuation is so large that the wave is hardly discernible.

We study plasmas of low density so as to be able to disregard energy losses arising from collisions.

29.1 UNIFORM PLANE ELECTROMAGNETIC WAVES IN GOOD CONDUCTORS. THE SKIN EFFECT

Recall from Sec. 28.5.1 that, in a linearly polarized, uniform plane wave propagating in a conductor in the positive direction of the z -axis,

$$\mathbf{E} = \mathbf{E}_m \exp(-\alpha z) \exp j(\omega t - \beta z). \quad (29-1)$$

We define a good conductor as a material such that, in the expressions for α and β given in Sec. 28.5.1,

$$[(1 + \mathcal{D}^2)^{1/2} + 1]^{1/2} \approx \mathcal{D}^{1/2}. \quad (29-2)$$

This condition is satisfied within 1% if

$$\mathcal{D} = \frac{\sigma}{\omega\epsilon} = \left| \frac{\partial E}{\partial D/\partial t} \right| \geq 50, \quad (29-3)$$

or if the conduction current density is at least 50 times larger than the displacement current density. But note here that σ and ϵ are functions of ω , especially at optical and x-ray frequencies. So \mathcal{D} does not decrease indefinitely as $1/f$, as the above equation appears to indicate.

In good conductors the wave equation 28-16 reduces to[†]

$$\nabla^2 \mathbf{E} - \mu\sigma \frac{\partial \mathbf{E}}{\partial t} = 0, \quad (29-4)$$

and Eq. 28-10 for the wave number to

$$k^2 = -j\omega\sigma\mu. \quad (29-5)$$

Thus

$$k = \beta - j\alpha = \left(\frac{\omega\sigma\mu}{2} \right)^{1/2} (1 - j) = \frac{1 - j}{\delta}, \quad n = \frac{c}{v} = \frac{c}{\omega\lambda} = \lambda_0 k = \frac{\lambda_0}{\delta} (1 - j) \quad (29-6)$$

$$\beta = \frac{1}{\lambda} = \alpha = \frac{1}{\delta} = \left(\frac{\omega\sigma\mu}{2} \right)^{1/2}, \quad (29-7)$$

where n is the index of refraction, $\lambda = \lambda/2\pi$, as usual, and where δ is the attenuation distance, defined in Sec. 28.5.1 as the distance over which the amplitude decreases by a factor of e .

[†] An equation of the same form applies to heat conduction. See Prob. 29-4.

From Eq. 28-70, the characteristic impedance of a good conductor is

$$Z = \frac{E}{H} = \frac{\omega\mu}{k} = \left(\frac{\omega\mu}{\sigma}\right)^{1/2} \exp\frac{j\pi}{4}, \quad (29-8)$$

and \mathbf{E} leads \mathbf{H} by $\pi/4$ radian. Compare with nonconductors in which \mathbf{E} and \mathbf{H} are in phase (Sec. 28.4). The difference comes from the fact that the current that is associated with \mathbf{H} in good conductors is the conduction current, which is in phase with \mathbf{E} , and not the displacement current of nonconductors, which leads \mathbf{E} by 90° .

Therefore

$$\mathbf{E} = E_m \exp\left[j\left(\omega t - \frac{z}{\delta}\right) - \frac{z}{\delta}\right], \quad (29-9)$$

$$\mathbf{H} = \left(\frac{\sigma}{\omega\mu}\right)^{1/2} E_m \exp\left[j\left(\omega t - \frac{z}{\delta} - \frac{\pi}{4}\right) - \frac{z}{\delta}\right]. \quad (29-10)$$

The vectors \mathbf{E} and \mathbf{H} are transverse and orthogonal, say \mathbf{E} is parallel to the x -axis and \mathbf{H} to the y -axis. In terms of cosine functions,

$$\mathbf{E} = E_m \exp\left(-\frac{z}{\delta}\right) \cos\left(\omega t - \frac{z}{\delta}\right), \quad (29-11)$$

$$\mathbf{H} = \left(\frac{\sigma}{\omega\mu}\right)^{1/2} E_m \exp\left(-\frac{z}{\delta}\right) \cos\left(\omega t - \frac{z}{\delta} - \frac{\pi}{4}\right) \quad (29-12)$$

$$= H_m \exp\left(-\frac{z}{\delta}\right) \cos\left(\omega t - \frac{z}{\delta} - \frac{\pi}{4}\right). \quad (29-13)$$

Figure 29-1 shows E/E_m and H/H_m as functions of z/λ at $t = 0$.

The amplitude of the wave decreases by a factor of $(1/e)^{2\pi} \approx 2 \times 10^{-3}$ in one wavelength, and the Poynting vector by $(1/e)^{4\pi} \approx 3 \times 10^{-6}$. This is the *skin effect*.

The attenuation distance δ in conductors is termed the *skin depth*, or the *depth of penetration*. The skin depth *decreases* if the conductivity, the relative permeability, or the frequency *increases*. Good conductors are therefore opaque to light, except in the form of extremely thin films. It does not follow, however, that substances that are nonconducting at low frequencies are transparent at optical frequencies.

Table 29-1 shows the skin depth δ for various conductors at four typical frequencies. The attenuation in iron is much larger than in silver, despite the fact that iron is a relatively poor conductor.

The phase velocity

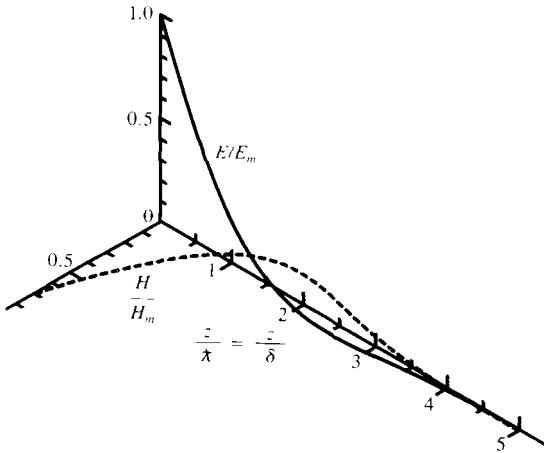


Fig. 29-1. The ratios E/E_m and H/H_m at $t = 0$ as functions of z/λ for an electromagnetic wave propagating in a good conductor.

$$v_p = \frac{\omega}{\beta} = \omega\lambda = \left(\frac{2\omega}{\sigma\mu}\right)^{1/2} \quad (29-14)$$

is proportional to the square root of the frequency.

In good conductors the group velocity (App. C) is twice as large as the phase velocity:

$$v_g \equiv \frac{1}{d\beta/d\omega} = 2v_p, \quad (29-15)$$

If σ and μ are not frequency-dependent.

The ratio of the time-averaged electric to the time-averaged magnetic energy densities is

$$\frac{\epsilon E_{\text{rms}}^2/2}{\mu H_{\text{rms}}^2/2} = \frac{\omega\epsilon}{\sigma} = \frac{1}{\mathcal{D}} \leq \frac{1}{50}. \quad (29-16)$$

The energy is thus essentially all magnetic. This results from the large conductivity σ , which causes E/J_f to be small. The electric field strength is weak, but the current density and hence H are relatively large.

From Eqs. 29-11 and 29-13, the time-averaged value of the Poynting vector is

$$\mathcal{S}_{\text{av}} = \frac{1}{2} \operatorname{Re} (\mathbf{E} \times \mathbf{H}^*) = \frac{1}{2} \left(\frac{\sigma}{2\omega\mu} \right)^{1/2} \exp\left(-\frac{2z}{\delta}\right) E_m^2 \hat{\mathbf{z}}. \quad (29-17)$$

Table 29-1 Skin depths δ for conductors

CONDUCTOR	σ	μ_r	$\delta f^{1/2}$
Aluminum	3.54×10^7	1.00	0.0846
Brass (65.8 Cu, 34.2 Zn)	1.59×10^7	1.00	0.126
Chromium	3.8×10^7	1.00	0.0816
Copper	5.80×10^7	1.00	0.0661
Gold	4.50×10^7	1.00	0.0750
Graphite	1.0×10^5	1.00	1.59
Soft iron	1.0×10^7	$2 \times 10^{2+}$	0.011
Mumetal	1.6×10^6	$2 \times 10^{4+}$	0.0028
Nickel	1.3×10^7	$1 \times 10^{2+}$	0.014
Sea water	5	1.00	2.3×10^2
Silver	6.15×10^7	1.00	0.0642
Tin	8.70×10^6	1.00	0.171
Zinc	1.86×10^7	1.00	0.117

[†] At $B = 0.002$ tesla.

Example

PROPAGATION IN COPPER AT 1 MEGAHERTZ

Copper has a conductivity of 5.80×10^7 siemens/meter. Then, at 1 megahertz,

$$\lambda = \frac{5.80 \times 10^7}{2\pi \times 10^6 \times 8.85 \times 10^{-12}} \approx 10^{12} \quad (29-18)$$

and

$$\begin{aligned} \delta = \lambda &= \left(\frac{2}{2\pi \times 10^6 \times 5.8 \times 10^7 \times 4\pi \times 10^{-7}} \right)^{1/2} \\ &= 66 \text{ micrometers.} \end{aligned} \quad (29-19)$$

The wavelength is about 0.4 millimeter in copper, while it is 300 meters in air. The phase velocity is correspondingly low:

$$v_p = \omega \lambda = 415 \text{ meters/second,} \quad (29-20)$$

which is about 10 times less than the velocity of sound in copper (3.6 kilometers/second).

The characteristic impedance is

$$|Z| = \left| \frac{E}{H} \right| = \left(\frac{2\pi \times 10^6 \times 4\pi \times 10^{-7}}{5.8 \times 10^7} \right)^{1/2} = 3.7 \times 10^{-4} \text{ ohm,} \quad (29-21)$$

versus 377 ohms in free space (Sec. 28.3).

Example

JOULE LOSSES IN GOOD CONDUCTORS

Let us compare the power lost by the wave and that gained by the medium through Joule losses. We consider a thin slice inside a

δ			
60 HERTZ mm	1 KILOHERTZ mm	1 MEGAHERTZ μm	3 GIGAHERTZ μm
10.9	2.67	84.6	1.54
16.3	3.99	126	2.30
10.5	2.58	81.6	1.49
8.53	2.09	66.1	1.21
9.69	2.37	75.0	1.37
2.1×10^2	50	1.6×10^3	29
1	0.4	10	0.2
0.4	0.09	3	0.05
2	0.4	10	0.3
3×10^4	7×10^3	2×10^5	$4 \times 10^{3\dagger}$
8.29	2.03	64.2	1.17
22.0	5.40	171	3.62
15.1	3.69	117	2.13

[†] At this frequency, $\epsilon_r \approx 35$, $\sigma \approx 1$, and sea water is not a good conductor.

conductor, perpendicular to the direction of propagation, as in Fig. 29-2. If the amplitude of E on the left-hand face is E_m , then on the right it is $E_m \exp(-\Delta z/\delta)$ and, within the slice, the time-averaged Poynting vector decreases by

$$\Delta S_{av} = \frac{1}{2} \left(\frac{\sigma}{2\omega\mu} \right)^{1/2} E_m^2 \left[1 - \exp \left(-\frac{2\Delta z}{\delta} \right) \right]. \quad (29-22)$$

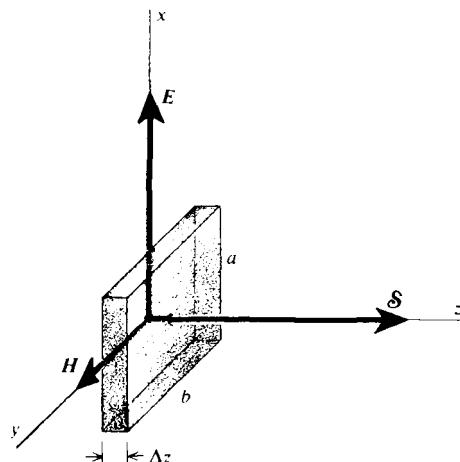


Fig. 29-2. Element of volume normal to the direction of propagation in a good conductor.

The power P_L lost by the wave is ab times this. If $2\Delta z/\delta \ll 1$,

$$P_L = \frac{ab}{2} \left(\frac{\sigma}{2\omega\mu} \right)^{1/2} E_m^2 \frac{2\Delta z}{\delta} = \sigma E_{\text{rms}}^2 ab \Delta z. \quad (29-23)$$

This is just the power dissipated per unit volume by Joule losses (Sec. 4.3.7).

29.2 PLANE ELECTROMAGNETIC WAVES IN PLASMAS

A plasma is an ionized gas. As a rule, plasmas contain free electrons and positive ions. Since the ions are more massive than the electrons, the current is carried almost exclusively by the electrons. We therefore disregard ionic currents.

We assume that the gas pressure is low. This will permit us to ignore collisions between the free electrons and the gas molecules, and hence to ignore energy losses. Then the gas offers no resistance to the motion of free electrons and acts somewhat like a superconductor. Also, we may set $\epsilon_r = 1$.

Finally, we disregard thermal agitation. In effect, we set the temperature equal to zero.

29.2.1 The Conductivity of a Plasma

For a fixed observer, a free electron situated in an electromagnetic wave is subjected to a Lorentz force $-e(\mathbf{E} + \mathbf{v} \times \mathbf{B})$, and the electron current density in a plasma is

$$\mathbf{J}_f = \sigma(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (29-24)$$

We can disregard the $\mathbf{v} \times \mathbf{B}$ term in this equation for the following reason. As we shall see below, the ratio E/H for a wave in a plasma is always larger than in free space, where $E/H = (\mu_0/\epsilon_0)^{1/2}$ (Sec. 28.3). So, in a plasma, $E/H > (\mu_0/\epsilon_0)^{1/2}$ and $E/B > c$. Also, we may safely assume that the velocity v of a free electron will be less than c by many orders of magnitude. So we may set

$$\mathbf{J}_f = -Nev = \sigma\mathbf{E}, \quad \sigma = -\frac{Nev}{E}, \quad (29-25)$$

where N is the number of free electrons per cubic meter, e is the magnitude of the electronic charge, and σ is the conductivity.

We now require the velocity v of a free electron of mass m subjected to an alternating electric field. Since

$$m \frac{dv}{dt} = mj\omega v = -eE_m \exp j\omega t, \quad (29-26)$$

then

$$v = j \frac{e}{\omega m} E_m \exp j\omega t. \quad (29-27)$$

The velocity leads the field by 90° . Substituting this value of v in the above expression for the conductivity gives[†]

$$\sigma = -\frac{jNe^2}{\omega m} = -\frac{4.48 \times 10^{-9} jN}{f} \quad \text{siemens/meter.} \quad (29-28)$$

The average power dissipation per cubic meter is

$$P' = \frac{1}{2} \operatorname{Re} (\mathbf{E} \cdot \mathbf{J}_f^*) = \frac{1}{2} \operatorname{Re} (\sigma E E^*) = 0, \quad (29-29)$$

in accordance with our assumption of zero losses.

Example

THE IMPEDANCE OF A UNIT CUBE OF PLASMA

First, imagine a cube of some resistive material such as carbon, with copper electrodes deposited on opposite faces. If the cube has a volume of 1 cubic meter, then the resistance between opposite faces is

$$R' = \frac{l}{\sigma A} = \frac{1}{\sigma}. \quad (29-30)$$

Thus the conductivity of a medium is the conductance $G' = 1/R'$, or, more generally, the admittance Y' , or $1/Z'$, between opposite faces of a unit cube.

With a plasma,

$$Z' = \frac{1}{\sigma} = j\omega \left(\frac{m}{Ne^2} \right) = j\omega L', \quad (29-31)$$

where L' is the “inductance” of 1 cubic meter of plasma. This quantity is expressed in henry-meters and is independent of the frequency.

[†] This expression is meaningless at zero frequency where the conduction electrons would take an infinite time to reach an infinite velocity. Equations 29-25, however, apply to steady fields, as long as $|\mathbf{v} \times \mathbf{B}| \ll E$.

29.2.2 The Conduction and Displacement Current Densities and the Plasma Angular Frequency ω_p

In a plasma,

$$\mathbf{J} = \frac{\partial \mathbf{D}}{\partial t} + \sigma \mathbf{E} = j\omega\epsilon_0 \mathbf{E} + \sigma \mathbf{E} = j\omega\epsilon_0 \mathbf{E} - \frac{jNe^2}{\omega m} \mathbf{E} \quad (29-32)$$

$$= j\omega\epsilon_0 \mathbf{E} \left(1 - \frac{Ne^2}{\omega^2 \epsilon_0 m} \right). \quad (29-33)$$

The displacement current density $\partial \mathbf{D} / \partial t$ leads \mathbf{E} by 90° , while the conduction current density $\sigma \mathbf{E}$ lags by 90° . The term between parentheses is the *equivalent relative permittivity* of the plasma because, in a dielectric,

$$\frac{\partial \mathbf{D}}{\partial t} = j\omega\epsilon_r \epsilon_0 \mathbf{E}. \quad (29-34)$$

The equivalent relative permittivity can be either positive or negative.

Rewriting Eq. 29-33,

$$\frac{\partial \mathbf{D}}{\partial t} + \sigma \mathbf{E} = j\omega\epsilon_0 \mathbf{E} \left(1 - \frac{\omega_p^2}{\omega^2} \right), \quad (29-35)$$

where the second term between the parentheses is the ratio of conduction current density to displacement current density, and where

$$\omega_p = \left(\frac{Ne^2}{\epsilon_0 m} \right)^{1/2} \quad (29-36)$$

is the *plasma angular frequency*. More simply,

$$f_p = \frac{\omega_p}{2\pi} = 8.98N^{1/2} \approx 9N^{1/2} \quad \text{hertz.} \quad (29-37)$$

Also, from the value of σ that we found above,

$$\frac{j\sigma}{\omega\epsilon_0} = \frac{\omega_p^2}{\omega^2}. \quad (29-38)$$

Figure 29-3 shows $\sigma \mathbf{E}$, $\partial \mathbf{D} / \partial t$, and their sum as functions of ω / ω_p .

In the plasma of a gas discharge, N varies widely but is typically of the order of 10^{18} electrons per cubic meter, and f_p is about 10 gigahertz. In the ionosphere, N is of the order of 10^{11} electrons per cubic meter, and f_p about 3 megahertz. See below.

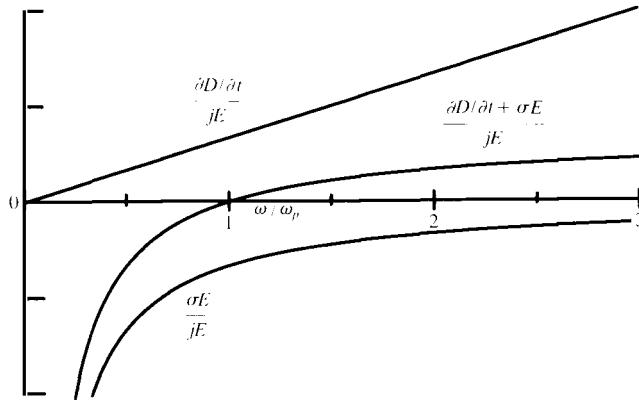


Fig. 29-3. The current densities $\partial\mathbf{D}/\partial t$, $\sigma\mathbf{E}$, and their sum, as functions of ω/ω_p in a plasma.

Example

PLASMA-FILLED PARALLEL-PLATE CAPACITOR

If the capacitor plates have an area \mathcal{A} and are separated by a distance s , then, upon application of a voltage \mathcal{V} ,

$$I = \left(j\omega C + \sigma \frac{\mathcal{A}}{s}\right) \mathcal{V} = \left(j\omega C + \frac{\mathcal{A}/s}{j\omega L'}\right) \mathcal{V}. \quad (29-39)$$

The presence of the plasma therefore has the same effect as if one had an inductance sL'/\mathcal{A} connected in parallel with the capacitor.

The net current is zero for

$$\omega^2 \frac{sL'}{\mathcal{A}} C = 1. \quad (29-40)$$

In actual fact, the plasma would not be uniform and N would be smaller near the electrodes than in the body of the plasma. Moreover, we have ignored edge effects. So this example is not realistic.

29.2.3 The Wave Number k

Recall from Sec. 28.2.3 that, in a conductor, the wave number k is given by

$$k^2 = \frac{\epsilon_r \mu_r}{\chi_0^2} \left(1 - j \frac{\sigma}{\omega \epsilon}\right). \quad (29-41)$$

Then, from Eq. 29-38 and setting $\epsilon_r = 1$, $\mu_r = 1$,

$$k^2 = \frac{1}{\lambda_0^2} \left(1 - \frac{\omega_p^2}{\omega^2} \right). \quad (29-42)$$

The wave number is either real or imaginary in a lossless plasma.

29.2.4 Propagation at $f > f_p$

At high frequencies such that $\omega > \omega_p$, or $f > f_p$, k is real. Then there is zero attenuation, as we assumed at the beginning.

The \mathbf{E} and \mathbf{H} vectors for a uniform, plane, linearly polarized wave are transverse and orthogonal, as in any homogeneous and isotropic medium.

Also, from the general expression for the characteristic impedance E/H that we found in Sec. 28.2.2 for any linearly polarized uniform plane wave,

$$Z = \frac{E}{H} = \frac{\omega \mu}{k} = \left(\frac{\mu_0 / \epsilon_0}{1 - \omega_p^2 / \omega^2} \right)^{1/2} = \frac{377}{(1 - \omega_p^2 / \omega^2)^{1/2}} > 377 \text{ ohms}. \quad (29-43)$$

The wavelength $\lambda = 2\pi/k$ is longer than in free space, and the phase velocity is larger than c :

$$v_p = \frac{c}{(1 - \omega_p^2 / \omega^2)^{1/2}} \quad (29-44)$$

$$= \frac{c}{(1 - 80.6N/f^2)^{1/2}} \approx \frac{c}{(1 - 81N/f^2)^{1/2}} \quad \text{meters/second}, \quad (29-45)$$

as in Fig. 29-4.

The phase velocity increases with increasing electron density. Waves therefore tend to bend away from more highly ionized regions, just as light tends to bend away from low-density air in a mirage.

This requires some discussion because, according to relativity, a signal cannot travel at a velocity greater than c .

The phase velocity is the velocity of propagation of a given phase, and not the velocity of propagation of a signal. The reason is that a wave can transmit a signal only if it is modulated, in amplitude, in frequency, or otherwise. Now any modulation, and hence any signal, involves frequencies other than the carrier frequency. The phase velocity in an ionized gas being frequency-dependent, the various frequency components of a signal travel at different velocities. The net result is that the envelope of the wave, which carries the signal, travels at a velocity that is different from those of the various components of the wave, and this signal velocity is less than c .

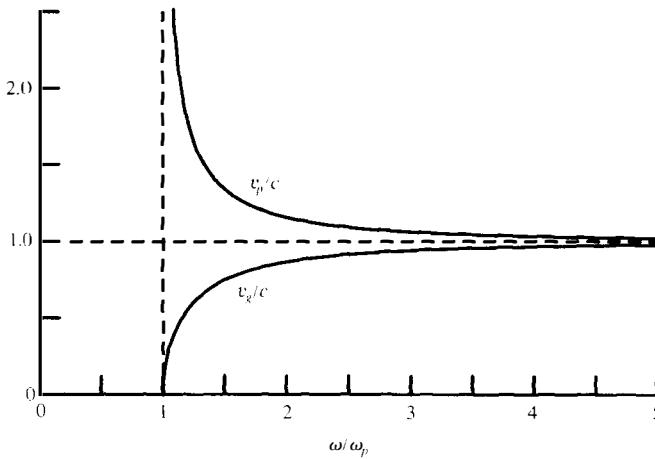


Fig. 29-4. The phase and group velocities v_p and v_g of an electromagnetic wave in a plasma.

We may set the signal velocity equal to the group velocity (App. C). Then

$$v_g = -\frac{1}{dk/d\omega} = \frac{c^2}{v_p}, \quad v_g v_p = c^2. \quad (29-46)$$

Figure 29-4 shows the group velocity as a function of the ratio ω/ω_p .

29.2.5 The Field at $f = f_p$

If the frequency of the applied field is equal to the plasma frequency f_p , then $k = 0$, $\delta \rightarrow \infty$ in Eqs. 29-11 and 29-13 for E and H , and

$$E = E_m \cos \omega t, \quad H = H_m \cos \left(\omega t - \frac{\pi}{4} \right). \quad (29-47)$$

There is no wave, and $\partial \mathbf{D}/\partial t + \sigma \mathbf{E}$ is zero.

29.2.6 The Field at $f < f_p$

At low frequencies, k is imaginary. Setting $k = -jk'$ gives

$$E = E_m \exp(-k'z) \exp j\omega t, \quad H = H_m \exp(-k'z) \exp j\omega t. \quad (29-48)$$

There is no wave, and the field decreases exponentially with z .

The average Poynting vector is clearly zero because there is zero power flow. We can also prove this formally as follows. From Eq. 29-43,

$$\frac{E}{H} = \frac{j\omega\mu_0}{k'} = j \frac{\omega\mu_0\lambda_0}{(\omega_p^2/\omega^2 - 1)^{1/2}} = j \left(\frac{\mu_0}{\epsilon_0} \right)^{1/2} \frac{1}{(\omega_p^2 - \omega^2)^{1/2}}. \quad (29-49)$$

Then

$$\mathcal{S}_{av} = \frac{1}{2} \operatorname{Re} \left(\frac{j\omega\mu_0 H_m^2}{k'} \right) \exp(-2k'z) = 0. \quad (29-50)$$

An electromagnetic wave therefore travels through a low-density plasma without attenuation at $f > f_p$. At lower frequencies there is no propagation.

Example | THE TELECOMMUNICATION BLACKOUT UPON THE REENTRY OF A SPACE SHUTTLE

In the course of one reentry there was complete radio blackout at a certain moment at all frequencies up to 10 gigahertz. Then f_p was 10 gigahertz and the electron density was therefore about 10^{18} electrons per cubic meter.

Example | THE IONOSPHERE

In the upper atmosphere, at altitudes ranging from about 50 to several thousand kilometers, the free electron density is sufficient to interfere with the propagation of radio waves. The main source of the ionization is the ultraviolet radiation of the sun. On the whole, the electron density increases with altitude up to about 300 kilometers, but it shows four ledges where the electron density increases more slowly. These ledges are commonly called *layers*. Beyond about 300 kilometers the electron density decreases slowly. The existence of these layers is ascribable to the fact that both the spectrum of the solar radiation and the chemical composition of the atmosphere change with altitude.

Both the heights and the ionization densities of the layers change with latitude, longitude, hour of the day, season, and sunspot cycle. At the lower altitudes, part of the conductivity comes from the presence of positive ions.

The free electron density varies from about 10^8 to 10^{12} per cubic meter. The degree of ionization increases rapidly with altitude, but it remains low, about 0.1% for the highest layer. For $N = 10^{10}$, the plasma frequency is about 1 megahertz.

Up to about 300 kilometers the electron density and the phase velocity increase with height. Amplitude-modulated radio waves (535 kilohertz to 1.605 megahertz) and the lower-frequency short waves (6 to 26 megahertz) bend back toward the earth in the same way that light waves bend in a mirage. However, frequency-modulated (88 to 108 megahertz) and television (54 to 890 megahertz) waves bend slightly in going through the ionosphere, and they escape.

Our assumption that there are no collisions between the electrons and the gas atoms or molecules is not satisfactory in the lower regions of the ionosphere, where the pressure is highest, at frequencies of the order of 1 megahertz or lower.

The earth's magnetic field renders the plasma doubly refracting, and there are two distinct phase velocities, depending on whether the \mathbf{E} vector of the wave is parallel or perpendicular to the \mathbf{B} of the earth.

29.3 SUMMARY

We define a *good conductor* as a medium for which

$$\mathcal{D} \equiv \frac{\sigma}{\omega\epsilon} \equiv \left| \frac{\partial \mathbf{E}}{\partial t} \right| \geq 50. \quad (29-3)$$

Then

$$k = \beta - j\alpha = \frac{1-j}{\delta}, \quad n = \frac{\lambda_0}{\delta}(1-j), \quad (29-6)$$

$$\delta = \left(\frac{\omega\sigma\mu}{2} \right)^{1/2}, \quad (29-7)$$

$$v_p = \frac{\omega}{\beta} = \left(\frac{2\omega}{\sigma\mu} \right)^{1/2}, \quad v_g = 2v_p. \quad (29-14), (29-15)$$

Most of the field energy is in magnetic form.

The conductivity of a low-density, and hence lossless, *plasma* is

$$\sigma = \frac{Ne^2}{j\omega m} \approx -\frac{4.8 \times 10^{-9} jN}{f} \quad \text{siemens/meter}, \quad (29-28)$$

where N is the number of electrons per cubic meter and m is the mass of the electron. Also,

$$\frac{\partial \mathbf{D}}{\partial t} + \sigma \mathbf{E} = j\omega\epsilon_0 \mathbf{E} \left(1 - \frac{\omega_p^2}{\omega^2} \right), \quad (29-35)$$

where

$$\omega_p = \left(\frac{Ne^2}{\epsilon_0 m} \right)^{1/2} \quad (29-36)$$

is the *plasma angular frequency*.

In a lossless plasma,

$$k^2 = \frac{1}{\lambda_0^2} \left(1 - \frac{\omega_p^2}{\omega^2} \right). \quad (29-42)$$

If $\omega > \omega_p$, k is real, there is no attenuation, and the phase velocity is larger than c . The phase velocity increases with increasing electron density. If $\omega < \omega_p$, there is no wave and the field attenuates exponentially.

PROBLEMS

29-1. (29.1) Good conductors

Show that for a good conductor

$$(a) \delta/\lambda_0 \ll 1, \quad (b) \delta^2/\mathcal{D} = 2\epsilon/(\mu\sigma^2), \quad (c) \delta^2\mathcal{D} = 2\lambda_0^2/(\epsilon_r\mu_r).$$

29-2. (29.1) The damped transmitted wave

Draw curves for E , similar to those of Fig. 29-1, for $\omega t = 0$ to 2π at intervals of $\pi/4$.

29-3. (29.1) Designing bus bars

You are asked to design copper bus bars that can carry 5000 amperes at 60 hertz over a distance of 5 meters. The total length of bus is 10 meters. The power dissipation in the line should not exceed 1 kilowatt. Suggest a plausible cross section.

29-4. (29.1) Heat propagation

It is interesting to draw a parallel between the flow of heat in a thermally conducting medium and the propagation of an electric or magnetic field in an electrically conducting medium.

Let Φ be the heat flux density in watts per square meter and $\Phi = -\lambda \nabla T$, where λ is the thermal conductivity in watts per meter-kelvin and T is the temperature in kelvins. Then, for conservation of energy,

$$\nabla \cdot \Phi = -\rho c \frac{\partial T}{\partial t} + Q,$$

where ρ is the mass density in kilograms/meter³, c is the specific heat in joules/kilogram-kelvin, and Q is the heat produced within the medium in watts/meter³.

For $Q = 0$,

$$\nabla^2 T - \frac{\rho c}{\lambda} \frac{\partial T}{\partial t} = 0.$$

This equation is identical in form to that for an electromagnetic wave in a good conductor, with $\rho c / \lambda$ corresponding to $\mu\sigma$. Its solution for heat flow in one dimension is similar:

Table 29-2

PROPERTY	COPPER	IRON
σ	5.8×10^7	1.0×10^7
c	385	460
ρ	8.9×10^3	7.9×10^3
λ	41.8	6.27
μ_r	1	100

$$T = T_m \exp \left[j \left(\omega t - \frac{z}{\delta_{th}} \right) - \frac{z}{\delta_{th}} \right], \quad \delta_{th} = \left(\frac{2\lambda}{\omega \rho c} \right)^{1/2}.$$

Compare the velocities of propagation of T and of B in copper and in iron. See Table 29-2.

19-5. (29.1) The surface impedance of a conductor

By definition, the surface impedance of a conductor is the ratio E_t/H_t , at the surface, the subscript t indicating a tangential component. It is shown in Prob. 19-4 that H_t is numerically equal to the current per unit width in the conductor.

(a) Show that the surface impedance of a good conductor is

$$(1 + j) \left(\frac{\omega \mu}{2\sigma} \right)^{1/2} \quad \text{or} \quad \frac{1 + j}{\sigma \delta},$$

where σ is the conductivity and δ is the skin depth. The quantity $1/\sigma\delta$ is the *surface resistance*. The surface impedance and the surface resistance are expressed in ohms/square. See Prob. 4-9. For example, the surface resistance of copper at 3 gigahertz is 14.4 miliohms/square.

(b) Show that, if the tangential magnetic field is H_t , then the power dissipated per square meter in the conductor is given by $H_{t,\text{rms}}^2/(\sigma\delta)$. This means that the power dissipated is the same as if the surface current (of density numerically equal to H_t) were distributed uniformly over a thickness δ of the conductor.

19-6. (29.1) Induction heating.

Induction heating consists in exciting eddy currents in a conductor by exposing it to an alternating magnetic field. The method serves for melting, in an *induction furnace*, for heating before a forging operation, or for hardening. An induction furnace comprises a crucible surrounded by a coil. The largest furnaces have capacities of tens of tons and powers up to a few megawatts. Once the load has melted, magnetic forces within the liquid provide stirring.

The coil is usually a single layer of water-cooled copper pipe that surrounds the object to be heated. If the object is ferromagnetic, a small part of the heating comes from hysteresis losses.

Induction heating has the advantage of convenience and of not contaminating the metal with combustion gases. Also, by choosing the

frequency correctly, it is possible to apply a heat treatment down to a known depth. For example, plowshares require a hard, heat-treated skin that resists abrasion and a soft core that resists breakage.

Metals can also be heat treated with a laser beam, but usually to a depth of only a fraction of a millimeter

Here is a simple example of induction heating. A steel rod of circular cross section lies inside a solenoid that applies an axial and tangential magnetic field H_r . We showed in Prob. 19-4 that the net surface current density is equal to H_r . A wave penetrates normally to the conductor, and the power dissipation in the conductor is the same as if the current were distributed uniformly over a thickness equal to the skin depth. See the preceding problem.

The solenoid has 100 turns per meter and carries a current of 600 amperes rms at 100 kilohertz.

Calculate the skin depth and the power P' dissipated in the iron per square meter. Set $\sigma = 10^7$ and $\mu_r = 100$. Neglect end effects, and neglect the fact that the relative permeability decreases to unity when the steel becomes red-hot.

29-7. (29.2) Wave propagation in a plasma

Show that, in a plasma, $vB \ll E$, or $E/B \gg v$, where v is the velocity of an electron.

29-8. (29.2.5) Plasmas compared to metallic conductors

(a) Find the value of \mathcal{D} , as defined in Sec. 29.1, for a low-density plasma.

(b) Show that the value of k^2 that we found for a conductor in Sec. 28.2.3 agrees with that of Sec. 29.2.5.

The values of α and of β that we found in Sec. 28.5.1 are not valid, however, for imaginary values of \mathcal{D} .

29-9. (29.2.4) The phase and group velocities in a plasma

Two uniform plane electromagnetic waves of equal amplitudes propagate in the ionosphere where the free electron density is N per cubic meter. One wave has a circular frequency ω_1 and a corresponding wavelength λ_1 ; the other has a slightly different circular frequency ω_2 and a wavelength λ_2 .

(a) At a given time t there exist values of z for which the two waves are in phase and other values of z for which they are opposite in phase. What is the distance between the maxima?

(b) What is their velocity? This is the *group velocity* v_g .

(c) Show that, in the limit, $v_g = 1/(dk/d\omega)$.

(d) Calculate the phase velocities and the group velocity for $f_1 = 5.3$ megahertz, $f_2 = 5.4$ megahertz, and $N = 5 \times 10^{10}$ electrons/meter³.

(e) Calculate the distance and the number of waves between two minima.

29-10. (29.2.4) The ω - β diagram for a low-density ionized gas

The ω - β diagram is a curve of ω as a function of β . The ratio ω/β is equal to the phase velocity, while the slope $d\omega/d\beta$ is equal to the group velocity.

Draw the ω - β diagram for a low-density ionized gas with $f_p = 3$ megahertz for frequencies ranging from 3 to 30 megahertz.

29-11. (29.2.4) Pulsars and satellites

Pulsars are stars that have suffered gravitational collapse, or *neutron stars*, and that rotate rapidly while emitting a narrow beam of radiation. The pulse lengths, at the earth, are of the order of 1 millisecond, and the periods of the order of 1 second. Neutron stars consist mostly of neutrons with some electrons and some ions. Their masses are of the order of that of the sun, but their radii are only of the order of 10 kilometers.

Within a few months after the discovery of pulsars, distance estimates were obtained in the following manner. It was observed that the arrival time of a pulse depends on the frequency of observation, the arrival time being later at lower frequencies. This delay is attributed to dispersion in the interstellar medium, which is ionized hydrogen with an electron density N of about $10^5/\text{meter}^3$.

(a) Show that, if $\omega^2 \gg \omega_p^2$, a plot of the time delay Δt as a function of $1/(f^2) - 1/[(f + \Delta f)^2]$ is a straight line whose slope is a measure of the distance to the pulsar.

(b) In the case of pulsar CP 0328, arrival times measured at 151, 408, and 610 megahertz gave the following results: between 610 and 408 megahertz, the delay was 0.367 second; between 408 and 151 megahertz, the delay was 4.18 seconds.

Find the distance to CP 0328 in parsecs where 1 parsec is 3.086×10^{16} meters. It is the distance from which the radius of the earth's orbit, 1.495×10^{11} meters, subtends an angle of $1'$.

The fact that such plots give straight lines passing through the origin indicates that the assumption that $\omega^2 \gg \omega_p^2$ is correct. The delay therefore occurs over large distances in interstellar space, and not inside the pulsar itself.

Similar methods are used to reduce satellite ranging errors due to the ionosphere.

29-12. (29.2.4) The energy densities, the Poynting vector, and the group velocity in a plasma

An electromagnetic wave travels in a low-density plasma. The electric field strength is $E_m \cos \omega t$.

(a) Calculate the sum of the electric, magnetic, and kinetic energy densities.

(b) Calculate the Poynting vector.

(c) Calculate the group velocity.

(d) Find the relations between these quantities.

CHAPTER 30

PLANE ELECTROMAGNETIC WAVES III

Reflection and Refraction A: The Basic Laws and Fresnel's Equations

30.1	REFLECTION AND REFRACTION	555
30.2	SNELL'S LAW	557
30.3	FRESNEL'S EQUATIONS	558
30.3.1	<i>E</i> NORMAL TO THE PLANE OF INCIDENCE	559
30.3.2	<i>E</i> PARALLEL TO THE PLANE OF INCIDENCE	561
30.4	REFLECTION AND REFRACTION AT THE INTERFACE BETWEEN TWO NONMAGNETIC NONCONDUCTORS	561
30.5	THE BREWSTER ANGLE	565
	<i>Example:</i> MEASURING THE RELATIVE PERMITTIVITY OF THE MOON'S SURFACE AT RADIO FREQUENCIES	567
30.6	THE COEFFICIENTS OF REFLECTION <i>R</i> AND OF TRANSMISSION <i>T</i>	567
*30.7	REFLECTION BY AN IONIZED GAS	569
30.8	SUMMARY	572
	PROBLEMS	573

Chapters 28 and 29 concerned the propagation of electromagnetic waves in unbounded media. We now investigate the behavior of a wave encountering a discontinuity, as in Fig. 30-1. The media will be the same as those of Chaps. 28 and 29, namely dielectrics, good conductors, and ionized gases. As previously, we define dielectrics as nonconductors, whether magnetic or not.

In this first chapter out of three on reflection and refraction, we make several simplifying assumptions:

- (1) The media extend to infinity on either side of the interface. This avoids multiple reflections.

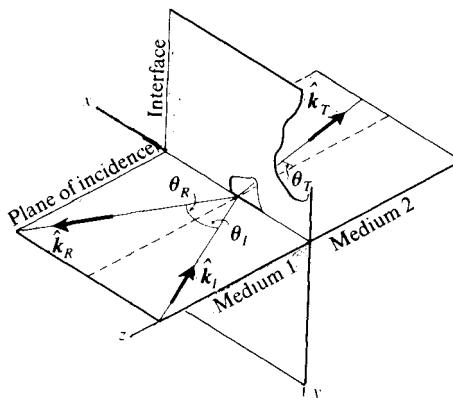


Fig. 30-1. An electromagnetic wave incident on the interface between media 1 and 2 gives both a reflected and a transmitted wave. The vectors \mathbf{k} are normal to their respective wave fronts and point in the direction of propagation. The angles θ_I , θ_R , θ_T are, respectively, the angles of incidence, reflection, and refraction.

- (2) The media are homogeneous, isotropic, linear, stationary (HILS) and lossless.
- (3) The interface is infinitely thin. In other words, the reflection is specular.
- (4) The incident wave is plane and uniform.

An electromagnetic wave incident on an interface usually gives rise to both a reflected and a transmitted wave. This type of phenomenon is common. For example, a sound wave incident upon a wall gives both a reflected wave that comes back into the room and another that proceeds into the wall. A wave propagating along a transmission line is partly reflected and partly transmitted at a discontinuity. Waves on strings show the same type of behavior, as shown in App. C.

30.1 REFLECTION AND REFRACTION

Medium 1 carries the incident and reflected waves. Medium 2 carries the refracted wave. For simplicity, we assume in Secs. 30-1 and 30-2 that the incident wave is linearly polarized. Then, in the incident wave,

$$\mathbf{E}_I = \mathbf{E}_{Im} \exp j(\omega_I t - \mathbf{k}_I \cdot \mathbf{r}), \quad (30-1)$$

where the vector wave number \mathbf{k}_I is real and points in the direction of

propagation of the incident wave. The magnitude of \mathbf{k}_I is $n_1 k_0$, or n_1/λ_0 , n_1 being the index of refraction of medium 1 and λ_0 the radian length of a wave of the same frequency in a vacuum. For convenience, we set the origin of \mathbf{r} in the interface, as in Fig. 30-1, and we take \mathbf{E}_{Im} to be real.

This equation defines a plane wave for all values of t and \mathbf{r} , and thus a wave that extends throughout all time and space. However, it applies only in medium 1.

Since the incident wave is plane, all the incident rays are parallel. By hypothesis, the interface is plane. Now the laws of reflection and of refraction must be the same at all points on the interface. It follows that the reflected rays are parallel to each other. Similarly, the refracted rays are parallel to each other. Further, since a wave front is by definition perpendicular to a ray, we can expect the reflected and transmitted waves to be of the form

$$\mathbf{E}_R = \mathbf{E}_{Rm} \exp j(\omega_R t - \mathbf{k}_R \cdot \mathbf{r}), \quad (30-2)$$

$$\mathbf{E}_T = \mathbf{E}_{Tm} \exp j(\omega_T t - \mathbf{k}_T \cdot \mathbf{r}). \quad (30-3)$$

What do we know about \mathbf{k}_R and \mathbf{k}_T ? From the wave equation 27-72 applied to medium 1, with $\sigma = 0$, $\rho_f = 0$,

$$\nabla^2 \mathbf{E}_R + \epsilon_1 \mu_1 \omega^2 \mathbf{E}_R = \nabla^2 \mathbf{E}_R + k_1^2 \mathbf{E}_R = 0, \quad (30-4)$$

where

$$k_1 = \frac{1}{\lambda_1} = \frac{n_1}{\lambda_0} = n_1 k_0 = \omega (\epsilon_1 \mu_1)^{1/2}. \quad (30-5)$$

A similar string of equations applies to k_2 . Also,

$$k_{Ix}^2 + k_{Iy}^2 + k_{Iz}^2 = k_{Rx}^2 + k_{Ry}^2 + k_{Rz}^2 = k_1^2, \quad k_{Tx}^2 + k_{Ty}^2 + k_{Tz}^2 = k_2^2. \quad (30-6)$$

The wave numbers k_1 and k_2 are real, but \mathbf{k}_R and \mathbf{k}_T are vectors that can be complex.

The tangential component of \mathbf{E} is continuous at the interface. This means that the tangential component of $\mathbf{E}_I + \mathbf{E}_R$ in medium 1, at the interface, is equal to the tangential component of \mathbf{E}_T in medium 2, at the interface. The same applies to \mathbf{H} . These continuity conditions will permit us to find all the unknowns in Eqs. 30-2 and 30-3.

Some relation must exist between \mathbf{E}_I , \mathbf{E}_R , \mathbf{E}_T at the interface for all t and for all points r_{int} on the interface. Such a relation is possible only if the three vectors are identical functions of t and r_{int} . Then

$$\omega_I = \omega_R = \omega_T. \quad (30-7)$$

All three waves are of the same frequency. This is obvious because the waves are all superpositions of the wave emitted by the source and of those waves emitted by the electrons executing forced vibrations in media 1 and 2. Recall from mechanics that forced vibrations are of the same frequency as the applied force.

Also, from the above equations for the \mathbf{E} 's,

$$\mathbf{k}_I \cdot \mathbf{r}_{\text{int}} = \mathbf{k}_R \cdot \mathbf{r}_{\text{int}} = \mathbf{k}_T \cdot \mathbf{r}_{\text{int}}. \quad (30-8)$$

Then the \mathbf{k} 's are oriented in such a way that their components parallel to the interface are equal. In particular, if $k_{Iy} = 0$ as in Fig. 30-1, then

$$k_{Ry} = 0, \quad k_{Ty} = 0, \quad (30-9)$$

and \mathbf{k}_I , \mathbf{k}_R , \mathbf{k}_T are *coplanar*. The plane containing these three vectors is called the *plane of incidence*. The x components of the \mathbf{k} 's are thus all equal:

$$k_{Rx} = k_{Tx} = k_{Ix} = k_1 \sin \theta_I, \quad (30-10)$$

where θ_I is the angle of incidence shown in Fig. 30-1.

It is now easy to find \mathbf{k}_R :

$$k_{Rx}^2 + k_{Rz}^2 = k_{Ix}^2 + k_{Iz}^2 = k_1^2 \quad (30-11)$$

and

$$k_{Rz}^2 = k_{Iz}^2, \quad k_{Rz} = -k_{Iz}. \quad (30-12)$$

We choose the negative sign because the reflected wave travels away from the interface. It follows that, if \mathbf{k}_I is real, as we assumed at the beginning (there is zero attenuation in medium 1), then \mathbf{k}_R is also real, the reflected wave is uniform, and

$$\theta_I = \theta_R. \quad (30-13)$$

The angle of reflection is equal to the angle of incidence.

Therefore the incident, reflected, and transmitted rays are coplanar, and the angle of reflection is equal to the angle of incidence. These are the *laws of reflection*.

30.2 SNELL'S LAW

Now return to Eq. 30-10. It says that

$$k_{Tx} = k_1 \sin \theta_I. \quad (30-14)$$

Then

$$k_{Tz}^2 = k_2^2 - k_{Tx}^2 = k_2^2 - k_1^2 \sin^2 \theta_I = k_0^2(n_2^2 - n_1^2 \sin^2 \theta_I). \quad (30-15)$$

If the term in parentheses is negative, then there is total reflection. We disregard this possibility until Chap. 31. Otherwise, k_{Tz} is real, k_T is real, and the transmitted wave is plane and uniform. If θ_T is the angle of refraction as in the figure,

$$k_{Tz} = -k_2 \cos \theta_T, \quad k_{Tz} = k_2 \sin \theta_T. \quad (30-16)$$

From Eqs. 30-14 and 30-16,

$$k_2 \sin \theta_T = k_1 \sin \theta_I, \quad \text{or} \quad n_2 \sin \theta_T = n_1 \sin \theta_I. \quad (30-17)$$

When an electromagnetic wave crosses an interface, there is conservation of the quantity $n \sin \theta$. This is *Snell's law*.

Therefore, choosing axes as in Fig. 30-1, we find that

$$\mathbf{E}_I = \mathbf{E}_{Im} \exp j[\omega t - k_1(x \sin \theta_I - z \cos \theta_I)]. \quad (30-18)$$

$$\mathbf{E}_R = \mathbf{E}_{Rm} \exp j[\omega t - k_1(x \sin \theta_I + z \cos \theta_I)], \quad (30-19)$$

$$\mathbf{E}_T = \mathbf{E}_{Tm} \exp j[\omega t - k_2(x \sin \theta_T - z \cos \theta_T)]. \quad (30-20)$$

The laws of reflection and Snell's law are general. They apply to any two homogeneous, isotropic, linear, and stationary (HILS) media, whether conducting or not, with either real or complex k 's, provided that one allows complex angles as in the next chapter.

30.3 FRESNEL'S EQUATIONS

We now require relations between \mathbf{E}_{Im} , \mathbf{E}_{Rm} , and \mathbf{E}_{Tm} that will ensure continuity of the tangential components of \mathbf{E} and \mathbf{H} at the interface.[†]

In this chapter \mathbf{k}_I , \mathbf{k}_R , \mathbf{k}_T are all real. Thus all three waves are plane and uniform, and hence transverse (Sec. 28.2.1). The \mathbf{E} vector of the incident wave can point in any direction perpendicular to \mathbf{k}_I .

Then the conditions of continuity at the interface require that

$$E_{Ix} + E_{Rx} = E_{Tx}, \quad E_{Iy} + E_{Ry} = E_{Ty}, \quad (30-21)$$

[†] We could also apply the continuity of the normal components of \mathbf{D} and of \mathbf{B} . But then our results would be applicable to reflection from the surface of a conductor only on the condition of taking into account the surface charge density, which would be another unknown.

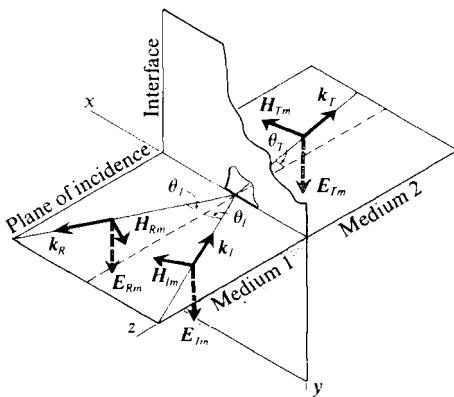


Fig. 30-2. The incident, reflected, and transmitted waves for an incident wave polarized with its \mathbf{E} field *normal* to the plane of incidence. The arrows show the directions in which the vectors are taken to be positive *at the interface*. The vectors $\mathbf{E} \times \mathbf{H}$ point everywhere in the direction of propagation.

$$H_{Ix} + H_{Rx} = H_{Tx}, \quad H_{Iy} + H_{Ry} = H_{Ty}. \quad (30-22)$$

Since the relation

$$\mathbf{H} = \frac{\mathbf{k} \times \mathbf{E}}{\omega\mu} \quad (30-23)$$

of Sec. 28.2.1 applies to all three waves, we first find \mathbf{E}_R and \mathbf{E}_T and then deduce \mathbf{H}_R and \mathbf{H}_T .

It will be convenient to divide the discussion into two parts. We consider successively incident waves polarized with their \mathbf{E} vectors normal and then parallel to the plane of incidence. Any uniform plane incident wave is the sum of two such components.

We now define our sign conventions. See Figs. 30-2 and 30-3. Observe that the two figures agree at normal incidence. We utilize the continuity of E_y and H_x in Fig. 30-2, and the continuity of E_x and H_y in Fig. 30-3. This will yield relations that apply again to any pair of HILS media and to any angle of incidence.

30.3.1 \mathbf{E} Normal to the Plane of Incidence

The \mathbf{E} and \mathbf{H} vectors of the incident wave point as in Fig. 30-2. The media being isotropic, the \mathbf{E} vectors of the other two waves are also normal to the plane of incidence. This is because the electrons in both media oscillate in the direction normal to the plane of incidence and radiate waves polarized with \mathbf{E} normal to the plane of incidence.

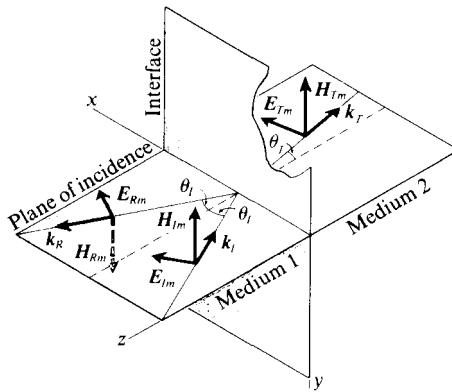


Fig. 30-3. This figure is similar to Fig. 30-2, except that now the \mathbf{E} fields are all parallel to the plane of incidence.

If the \mathbf{E} vectors point in the directions shown, at the interface, then the \mathbf{H} vectors point as shown, to orient the Poynting vectors $\mathbf{E} \times \mathbf{H}$ (Sec. 28.6) in the proper directions.

The continuity of the tangential component of \mathbf{E} at the interface requires that

$$E_{Im} + E_{Rm} = E_{Tm} \quad (30-24)$$

at any given point on the interface. Similarly, for continuity of the tangential component of \mathbf{H} ,

$$H_{Im} \cos \theta_I - H_{Rm} \cos \theta_I = H_{Tm} \cos \theta_T \quad (30-25)$$

or, from Sec. 28.4,

$$\frac{(E_{Im} - E_{Rm}) \cos \theta_I}{Z_1} = \frac{E_{Tm} \cos \theta_T}{Z_2}, \quad (30-26)$$

where Z is the characteristic impedance of a medium

$$Z = \frac{E}{H} = \frac{\omega \mu}{k} = \frac{\omega \mu}{n k_0} = \frac{\omega \mu}{n(\omega/c)} = \frac{c \mu}{n}, \quad (30-27)$$

n being the index of refraction.

Solving,

$$\left(\frac{E_{Rm}}{E_{Im}} \right)_\perp = \frac{Z_2 \cos \theta_I - Z_1 \cos \theta_T}{Z_2 \cos \theta_I + Z_1 \cos \theta_T}, \quad (30-28)$$

$$\left(\frac{E_{Tm}}{E_{Im}} \right)_\perp = \frac{2Z_2 \cos \theta_I}{Z_2 \cos \theta_I + Z_1 \cos \theta_T}. \quad (30-29)$$

The subscript \perp indicates that the E vectors are perpendicular to the plane of incidence.[†] These are two of *Fresnel's equations*.

30.3.2 E Parallel to the Plane of Incidence

The E 's are now all in the plane of incidence, as in Fig. 30-3, and

$$H_{Im} - H_{Rm} = H_{Tm}, \quad (30-30)$$

or

$$\frac{E_{Im} - E_{Rm}}{Z_1} = \frac{E_{Tm}}{Z_2}. \quad (30-31)$$

Also,

$$(E_{Im} + E_{Rm}) \cos \theta_I = E_{Tm} \cos \theta_T. \quad (30-32)$$

Then

$$\left(\frac{E_{Rm}}{E_{Im}} \right)_{\parallel} = \frac{Z_2 \cos \theta_T - Z_1 \cos \theta_I}{Z_2 \cos \theta_T + Z_1 \cos \theta_I}, \quad (30-33)$$

$$\left(\frac{E_{Tm}}{E_{Im}} \right)_{\parallel} = \frac{2Z_2 \cos \theta_I}{Z_2 \cos \theta_T + Z_1 \cos \theta_I}. \quad (30-34)$$

This is the second pair of *Fresnel's equations*.

At normal incidence $\theta_I = \theta_R = \theta_T = 0$, the plane of incidence is undefined, and the two pairs of Fresnel's equations are identical:

$$\frac{E_{Rm}}{E_{Im}} = \frac{Z_2 - Z_1}{Z_2 + Z_1}, \quad (30-35)$$

$$\frac{E_{Tm}}{E_{Im}} = \frac{2Z_2}{Z_2 + Z_1}. \quad (30-36)$$

30.4 REFLECTION AND REFRACTION AT THE INTERFACE BETWEEN TWO NONMAGNETIC NONCONDUCTORS

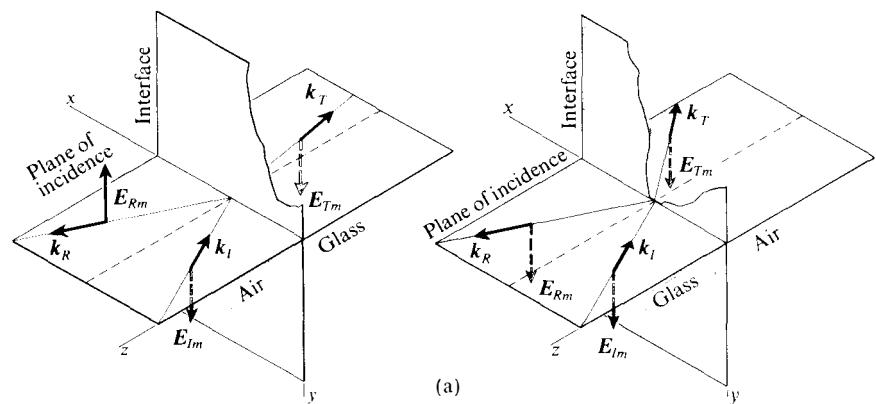
We continue to disregard total reflection. For nonmagnetic nonconductors,

$$Z_1 = \frac{\omega \mu_0}{k_1} = \frac{c \mu_0}{n_1}, \quad Z_2 = \frac{c \mu_0}{n_2}, \quad (30-37)$$

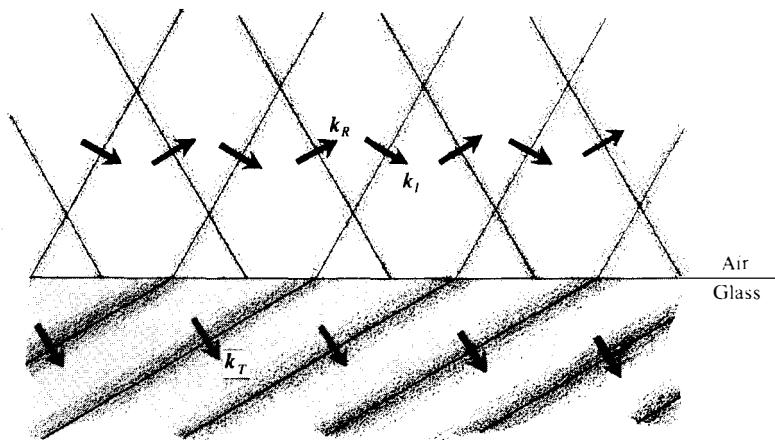
and

$$\left(\frac{E_{Rm}}{E_{Im}} \right)_{\perp} = \frac{(n_1/n_2) \cos \theta_I - \cos \theta_T}{(n_1/n_2) \cos \theta_I + \cos \theta_T}, \quad (30-38)$$

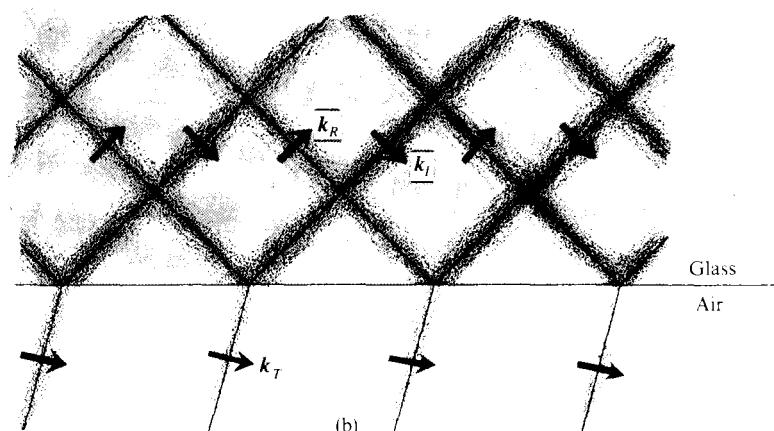
[†] We use the subscripts \perp and \parallel , but some authors use s and p .



(a)



(b)



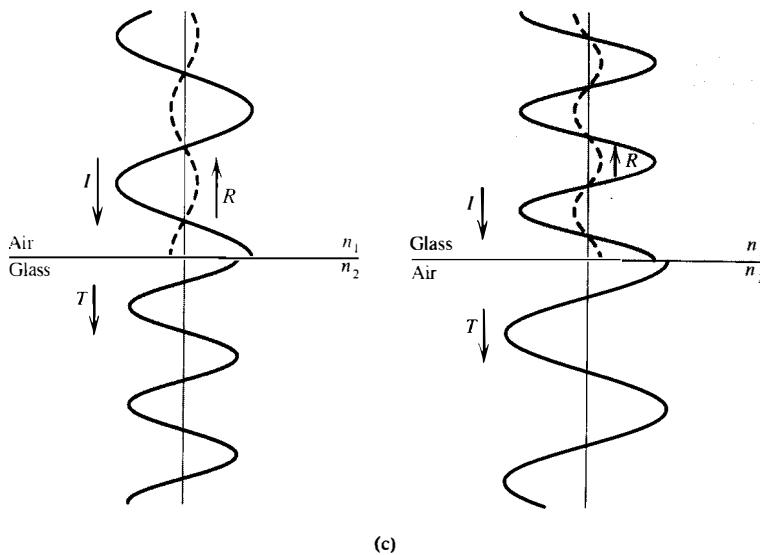


Fig. 30-4. (a) The relative phases, at the interface, of the \mathbf{E} 's in the reflected and transmitted waves for $n_1 < n_2$ and for $n_1 > n_2$, with \mathbf{E}_{lm} normal to the plane of incidence. On the left, the reflected wave is π radians out of phase with respect to the incident wave. The transmitted wave is in phase, in both instances. (b) "Crests" of the field \mathbf{E} at some particular time. The crests are one wavelength apart and travel in the directions of the arrows. Note the phase shift of π upon reflection from a glass surface. Note also the interference pattern that results from the superposition of the incident and reflected waves. (c) The \mathbf{E} vectors at a given instant in the incident, reflected, and transmitted waves at normal incidence on a glass-air interface. On the right, E_{Tm} is larger than E_{lm} . However, conservation of energy still applies.

$$\left(\frac{E_{Tm}}{E_{lm}}\right)_\perp = \frac{2(n_1/n_2) \cos \theta_I}{(n_1/n_2) \cos \theta_I + \cos \theta_T}. \quad (30-39)$$

Observe that the second ratio is always real and positive. This means that, at the interface, the transmitted wave is always in phase with the incident wave.

The first ratio can, however, be either positive or negative, depending on the value of n_1/n_2 . If $n_1/n_2 < 1$, then $\theta_T < \theta_I$ and $\cos \theta_I < \cos \theta_T$ whereas if $n_1/n_2 > 1$, then $\theta_T > \theta_I$ and $\cos \theta_I > \cos \theta_T$. The reflected wave is thus either π radians out of phase with the incident wave at the interface if $n_1 < n_2$, or in phase if $n_1 > n_2$.

Figure 30-4 illustrates the \mathbf{E} vectors for both types of reflection. Figure 30-5 shows the above ratios for $n_1/n_2 = 1/1.5$. This corresponds to a light wave incident in air on a glass whose index of refraction is 1.5.

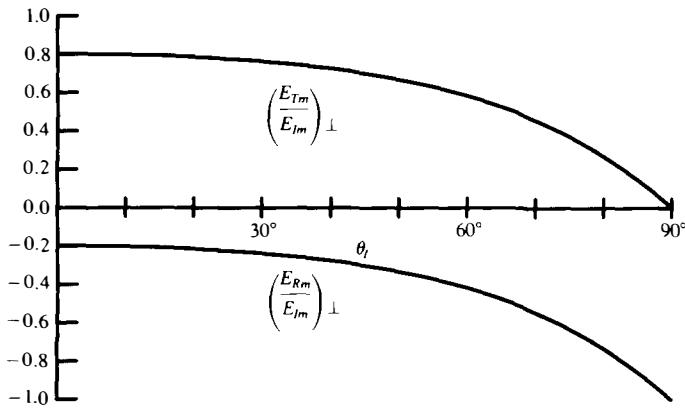


Fig. 30-5. Reflection and refraction when $n_1/n_2 = 1/1.5$, for example, when light falls on a glass of $n = 1.5$. The \mathbf{E} field is *normal* to the plane of incidence.

For an incident wave polarized with its \mathbf{E} vector parallel to the plane of incidence,

$$\left(\frac{E_{Rm}}{E_{Im}}\right)_\parallel = \frac{-\cos \theta_I + (n_1/n_2) \cos \theta_T}{\cos \theta_I + (n_1/n_2) \cos \theta_T}, \quad (30-40)$$

$$\left(\frac{E_{Tm}}{E_{Im}}\right)_\parallel = \frac{2(n_1/n_2) \cos \theta_I}{\cos \theta_I + (n_1/n_2) \cos \theta_T}. \quad (30-41)$$

See Fig. 30-6.

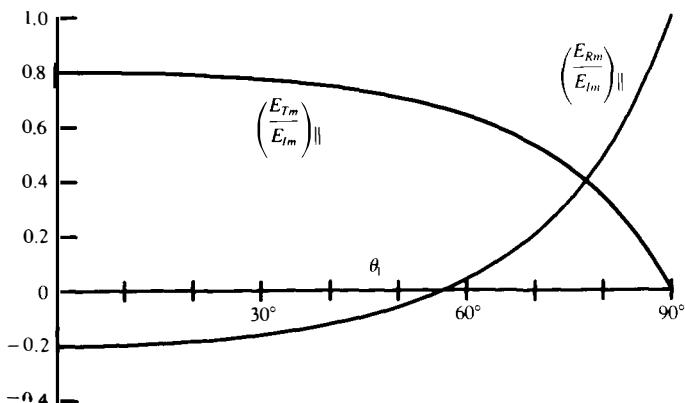


Fig. 30-6. Reflection and refraction when $n_1/n_2 = 1/1.5$, as in Fig. 30-5, but with \mathbf{E} parallel to the plane of incidence.

The second ratio is always positive. Then E_{Tm} and E_{Im} are in phase at the interface, as in Fig. 30-3, which we used to arrive at this result.

However, the ratio for E_{Rm} can be either positive or negative, which indicates that E_{Rm} can point either as in Fig. 30-3 or in the opposite direction. The tangential components of E_{Im} and of E_{Rm} can thus be either in phase or π radians out of phase. They are in phase if

$$\frac{n_1}{n_2} \cos \theta_T > \cos \theta_I, \quad (30-42)$$

or if

$$\sin \theta_T \cos \theta_T - \sin \theta_I \cos \theta_I > 0, \quad (30-43)$$

$$\sin 2\theta_T - \sin 2\theta_I > 0, \quad (30-44)$$

$$\sin(\theta_T - \theta_I) \cos(\theta_T + \theta_I) > 0. \quad (30-45)$$

This inequality requires that either

$$\theta_T > \theta_I \quad \text{and} \quad \theta_T + \theta_I < \frac{\pi}{2} \quad (30-46)$$

or

$$\theta_T < \theta_I \quad \text{and} \quad \theta_T + \theta_I > \frac{\pi}{2}. \quad (30-47)$$

The phase of the reflected wave when \mathbf{E} is parallel to the plane of incidence does not therefore depend only on the ratio n_1/n_2 ; it depends on both θ_I and θ_T . The ratio E_{Rm}/E_{Tm} can be either positive or negative, both for $n_2 > n_1$ and for $n_2 < n_1$. Figure 30-6 shows the above ratios for $n_1/n_2 = 1/1.5$.

30.5 THE BREWSTER ANGLE

We have seen that, when \mathbf{E} is parallel to the plane of incidence, E_{Rm} is either in phase or π radians out of phase with the incident wave, depending on whether $\sin(\theta_T - \theta_I) \cos(\theta_T + \theta_I)$ is greater or less than zero. It follows that there is no reflected wave when this expression is equal to zero, that is, when $\theta_I = \theta_T = 0$ or when $\theta_I + \theta_T = \pi/2$. The first condition is incorrect. It arises from the fact that we have multiplied the inequality

$$\frac{n_1}{n_2} \cos \theta_T - \cos \theta_I > 0 \quad (30-48)$$

by $\sin \theta_I$, which is equal to zero at $\theta_I = 0$.

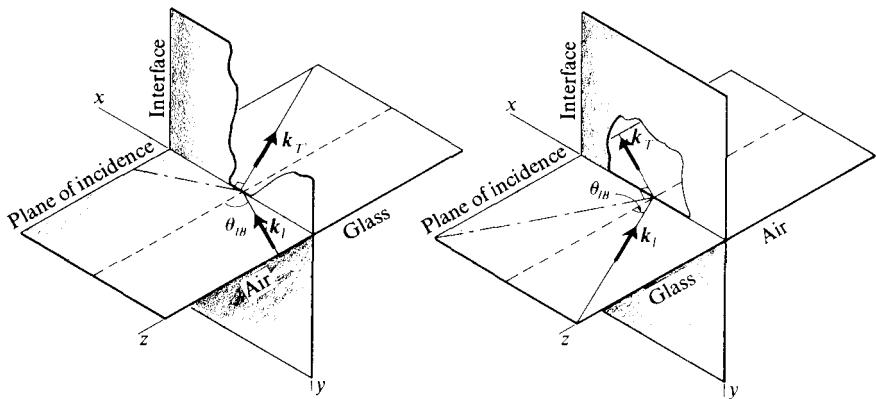


Fig. 30-7. If the angle of incidence is equal to the Brewster angle and if \mathbf{E} lies in the plane of incidence, there is no reflected wave. The position of the missing reflected ray is at 90° to the transmitted ray.

Thus for

$$\theta_I + \theta_T = \frac{\pi}{2} \quad (30-49)$$

there is no reflected wave when the incident wave is polarized with its \mathbf{E} vector parallel to the plane of incidence. This is remarkable because the wave then goes through an interface without reflection. See Fig. 30-7.

This angle of incidence is the *Brewster angle*. It is also called the *polarizing angle*, since an unpolarized wave incident on an interface at this angle is reflected as a polarized wave with its \mathbf{E} vector normal to the plane of incidence.[†]

At the Brewster angle,

$$\frac{n_1}{n_2} = \frac{\sin \theta_T}{\sin \theta_{IB}} = \frac{\sin (\pi/2 - \theta_{IB})}{\sin \theta_{IB}} = \cot \theta_{IB}. \quad (30-50)$$

For light incident in air on a glass whose index of refraction is 1.5, $\theta_{IB} = 56.3^\circ$.

[†]The Brewster angle is often explained incorrectly as follows. For this particular angle of incidence, the missing reflected ray is at 90° to the transmitted ray. It is argued that the electrons excited in medium 2 do not radiate in their direction of oscillation (Sec. 38.2) and hence cannot give rise to a reflected ray in medium 1. This explanation is incorrect, since there is a Brewster angle even when medium 2 is a vacuum. Also, with magnetic media, there can exist a Brewster angle when \mathbf{E}_{lm} is perpendicular to the plane of incidence.

The Brewster angle is commonly used to measure the index of refraction of a substance by reflecting a ray of light from its surface. The measurement can be accurate to five significant figures.

A plane wave incident on a *plate* of glass at the Brewster angle meets the second interface also at its Brewster angle. So there is no reflection, either at the first or at the second interface.

Example | MEASURING THE RELATIVE PERMITTIVITY OF THE MOON'S SURFACE AT RADIO FREQUENCIES

The nature of the moon's surface can be inferred, to some extent, from the value of its relative permittivity $\epsilon_r = n^2$ (Sec. 28.4) at radio frequencies. The Brewster angle can serve to measure this quantity in the following way.

If a radio wave originating from a satellite in lunar orbit illuminates the moon, the reflection observed at the earth is similar to the reflection of sunlight from the surface of a lake: most of the light comes from the regions that happen to be correctly oriented for specular reflection. The surface of the moon thus glistens over an area of the order of 100 kilometers in diameter, the area depending on the height of the satellite above the surface of the moon and on the roughness of the surface.

If the detector on the earth receives both the reflected wave and a direct wave from the satellite, it is possible to discriminate between the two by using the fact that the Doppler effect makes the two radio frequencies slightly different. A plot of the intensity of the reflected wave as a function of the angle of incidence when the E vector lies in the plane of incidence shows zero reflection at the Brewster angle.

In one such measurement, performed at a frequency of 140 megahertz, the Brewster angle was $60 \pm 1^\circ$ in the mare northwest of Hanstein. This gives an ϵ_r of 3.0 ± 0.2 .

It is possible to perform similar measurements at other points on the surface of the moon because of the relative motions of the three bodies involved, namely the satellite, the moon, and the earth.

30.6 THE COEFFICIENTS OF REFLECTION R AND OF TRANSMISSION T

The coefficients of reflection and of transmission concern the flow of energy across the interface. The average energy flux per unit area in a wave is equal to the average value of the Poynting vector, as in Eq. 28-31. We exclude total reflection as well as reflection from conducting media. Setting $\mu_r = 1$, we find that

$$\mathcal{S}_{I,\text{av}} = \frac{1}{2} \left(\frac{\epsilon_1}{\mu_0} \right)^{1/2} E_{Im}^2 \hat{n}_I, \quad (30-51)$$

$$\mathcal{S}_{R,\text{av}} = \frac{1}{2} \left(\frac{\epsilon_1}{\mu_0} \right)^{1/2} E_{Rm}^2 \hat{n}_R, \quad (30-52)$$

$$\mathcal{S}_{T,\text{av}} = \frac{1}{2} \left(\frac{\epsilon_2}{\mu_0} \right)^{1/2} E_{Tm}^2 \hat{n}_T, \quad (30-53)$$

where \hat{n}_I is normal to a wave front of the incident wave:

$$\hat{n}_I = \frac{\mathbf{k}_I}{k_1}, \quad (30-54)$$

and similarly for \hat{n}_R and \hat{n}_T .

The coefficients of reflection R and of transmission T are the ratios of the average energy fluxes per unit time and per unit area at the interface:

$$R = \left| \frac{\mathcal{S}_{R,\text{av}} \cdot \hat{n}}{\mathcal{S}_{I,\text{av}} \cdot \hat{n}} \right| = \frac{E_{Rm}^2}{E_{Im}^2}, \quad (30-55)$$

where \hat{n} is the unit vector normal to the interface;

$$T = \left| \frac{\mathcal{S}_{T,\text{av}} \cdot \hat{n}}{\mathcal{S}_{I,\text{av}} \cdot \hat{n}} \right| = \left(\frac{\epsilon_{r2}}{\epsilon_{r1}} \right)^{1/2} \frac{E_{Tm}^2 \cos \theta_T}{E_{Im}^2 \cos \theta_I} = \frac{n_2 E_{Tm}^2 \cos \theta_T}{n_1 E_{Im}^2 \cos \theta_I}. \quad (30-56)$$

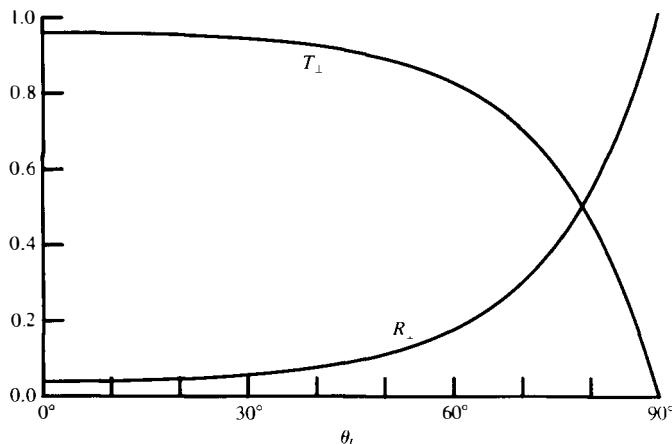


Fig. 30-8. The coefficients of reflection R_{\perp} and of transmission T_{\perp} as functions of the angle of incidence θ_I for $n_1/n_2 = 1/1.5$.

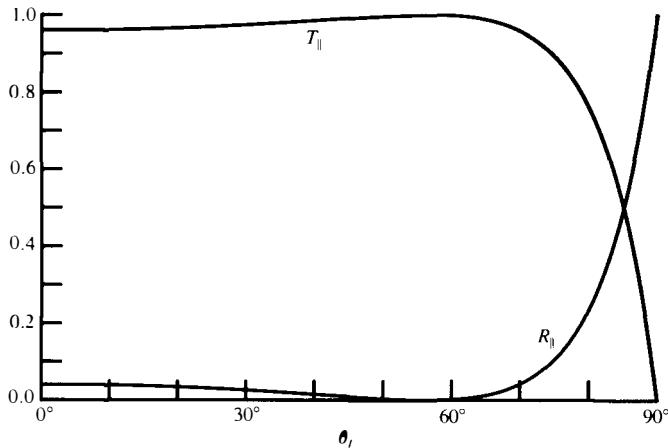


Fig. 30-9. The coefficients of reflection R_{\parallel} and of transmission T_{\parallel} as functions of the angle of incidence θ_I for $n_1/n_2 = 1/1.5$. Note the Brewster angle at 56.3° .

Then, from Fresnel's equations for nonconductors,

$$R_{\perp} = \left[\frac{(n_1/n_2) \cos \theta_I - \cos \theta_T}{(n_1/n_2) \cos \theta_I + \cos \theta_T} \right]^2, \quad (30-57)$$

$$T_{\perp} = \frac{4(n_1/n_2) \cos \theta_I \cos \theta_T}{[(n_1/n_2) \cos \theta_I + \cos \theta_T]^2}, \quad (30-58)$$

$$R_{\parallel} = \left[\frac{-\cos \theta_I + (n_1/n_2) \cos \theta_T}{\cos \theta_I + (n_1/n_2) \cos \theta_T} \right]^2, \quad (30-59)$$

$$T_{\parallel} = \frac{4(n_1/n_2) \cos \theta_I \cos \theta_T}{[\cos \theta_I + (n_1/n_2) \cos \theta_T]^2}. \quad (30-60)$$

In both instances, $R + T = 1$, as expected from the conservation of energy. At the Brewster angle defined above, $R_{\parallel} = 0$ and $T_{\parallel} = 1$, again as expected. See Figs. 30-8 to 30-10.

*30.7 REFLECTION BY AN IONIZED GAS

We saw in Sec. 29.2.4 that, in a low-density ionized gas, the *phase* velocity is *larger* than in free space and the index of refraction n is less than unity:

$$n = \frac{c}{v_p} = \left[1 - \left(\frac{\omega_p}{\omega} \right)^2 \right]^{1/2} = \left(\frac{1 - 80.6N}{f^2} \right)^{1/2}, \quad (30-61)$$

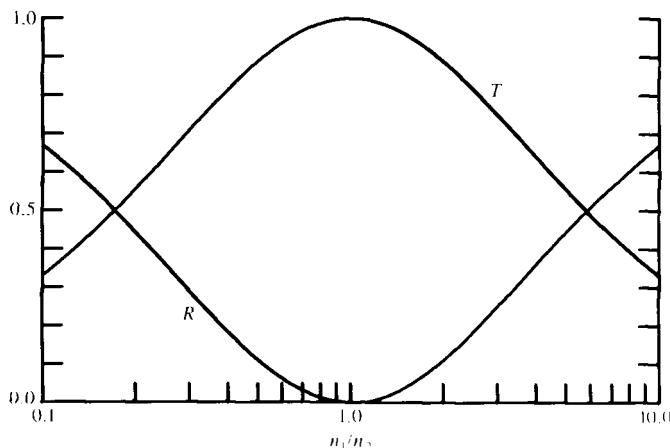


Fig. 30-10. The coefficients of reflection R and of transmission T at normal incidence, as functions of the ratio n_1/n_2 .

where c is the speed of light, v_p is the phase velocity, ω_p is the plasma angular frequency, ω is the angular frequency, N is the number of free electrons per cubic meter, and f is the frequency.

If the ionized gas could have a definite boundary and a uniform free electron density, then reflection and refraction at its surface would be simple: the gas would act as a dielectric with n less than unity. In practice, neither assumption is valid and reflection occurs gradually, as in a mirage.

We could calculate the path of a ray as in Prob. 30-1. However, Snell's law (Sec. 30.2) provides the main features of the reflection.

We select coordinates as in Fig. 30-11 and assume that the index of refraction n varies slowly with z , but not with the other two coordinates. To be more specific, we assume that n varies by a negligible amount over one wavelength. Then a given ray gradually bends down to an angle θ , at a point where the index of refraction is n .

We can calculate θ in the following way. When refraction occurs at the interface between any two media n_1 and n_2 , the quantity $n \sin \theta$ is conserved in going from one side of the interface to the other. This is Snell's law. If n varies gradually with z but not with x and y , then we can imagine the medium to be stratified in thin layers, and the value of $n \sin \theta$ remains the same all along the ray. Thus

$$n \sin \theta = n_1 \sin \theta_i. \quad (30-62)$$

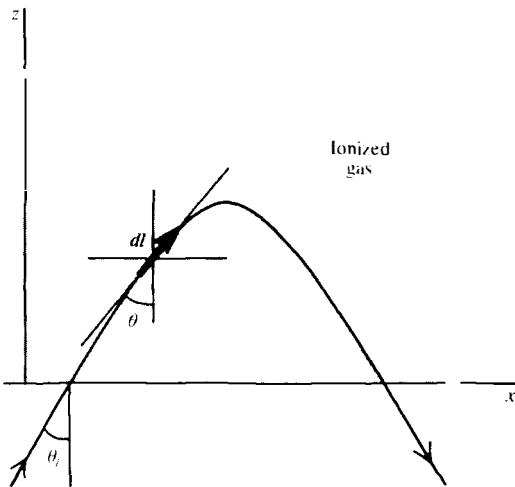


Fig. 30-11. Reflection from an ionized gas in which the electron density increases with increasing z , or upward.

If $n_1 = 1$, then

$$n \sin \theta = \sin \theta_I. \quad (30-63)$$

Differentiating with respect to the distance l measured along a ray, we find that

$$\frac{d\theta}{dl} = -\frac{1}{n} \frac{dn}{dl} \tan \theta. \quad (30-64)$$

If the ray penetrates into an ionized region where the ion density *increases* with z , the index of refraction n *decreases* with l and the derivative dn/dl is negative, so the angle θ *increases* with distance, as in Fig. 30-11. After some distance, if N becomes sufficiently large, θ becomes equal to 90° , the tangent of θ becomes infinite, and dn/dl becomes zero. After this, $\tan \theta$ becomes negative, whereas the derivative dn/dl becomes positive, and θ keeps increasing until the ray escapes back from the ionized region at an angle equal to the angle of incidence θ_I .

- At the top of the trajectory

$$\sin \theta = 1, \quad n_{\theta=90^\circ} = \sin \theta_I. \quad (30-65)$$

This is the index of refraction required for reflection when the angle of incidence is θ_I .

30.8 SUMMARY

At the plane interface between two homogeneous, isotropic, linear, stationary (HILS) and lossless media, (1) the incident, reflected, and transmitted waves all have the same frequency, (2) the three vector wave numbers \mathbf{k}_I , \mathbf{k}_R , \mathbf{k}_T are coplanar, and (3) the angle of reflection is equal to the angle of incidence. These are the *laws of reflection*.

Snell's law states that

$$n_2 \sin \theta_T = n_1 \sin \theta_I, \quad (30-17)$$

where n_1 is the index of refraction of the first medium and n_2 is that of the second.

Fresnel's equations are as follows:

$$\left(\frac{E_{Rm}}{E_{Im}} \right)_\perp = \frac{(n_1/n_2) \cos \theta_I - \cos \theta_T}{(n_1/n_2) \cos \theta_I + \cos \theta_T}, \quad (30-38)$$

$$\left(\frac{E_{Tm}}{E_{Im}} \right)_\perp = \frac{2(n_1/n_2) \cos \theta_I}{(n_1/n_2) \cos \theta_I + \cos \theta_T}, \quad (30-39)$$

$$\left(\frac{E_{Rm}}{E_{Im}} \right)_\parallel = \frac{-\cos \theta_I + (n_1/n_2) \cos \theta_T}{\cos \theta_I + (n_1/n_2) \cos \theta_T}, \quad (30-40)$$

$$\left(\frac{E_{Tm}}{E_{Im}} \right)_\parallel = \frac{2(n_1/n_2) \cos \theta_I}{\cos \theta_I + (n_1/n_2) \cos \theta_T}. \quad (30-41)$$

At the *Brewster angle* of incidence θ_{IB} ,

$$\frac{n_1}{n_2} = \cot \theta_{IB}, \quad (30-50)$$

and there is no reflected wave if E lies in the plane of incidence.

The *coefficients of reflection R* and *of transmission T* are the ratios of the average energy fluxes per unit time and per unit area at the interface:

$$R_\perp = \left[\frac{(n_1/n_2) \cos \theta_I - \cos \theta_T}{(n_1/n_2) \cos \theta_I + \cos \theta_T} \right]^2, \quad (30-57)$$

$$T_\perp = \frac{4(n_1/n_2) \cos \theta_I \cos \theta_T}{[(n_1/n_2) \cos \theta_I + \cos \theta_T]^2}, \quad (30-58)$$

$$R_\parallel = \left[\frac{-\cos \theta_I + (n_1/n_2) \cos \theta_T}{\cos \theta_I + (n_1/n_2) \cos \theta_T} \right]^2, \quad (30-59)$$

$$T_{\parallel} = \frac{4(n_1/n_2) \cos \theta_I \cos \theta_T}{[\cos \theta_I + (n_1/n_2) \cos \theta_T]^2}. \quad (30-60)$$

Reflection from an ionized gas occurs gradually as in a mirage.

PROBLEMS

30-1. (30.2) The ray equation

A wave travels in a stratified medium whose index of refraction is a function only of the coordinate y .

(a) Show that the angle θ between a ray and the y -axis obeys the following law:

$$\frac{d\theta}{dl} = -\frac{1}{n} \frac{dn}{dy} \sin \theta,$$

where the distance l is measured along the ray.

(b) You can now verify the *ray equation*

$$\frac{d}{dl}(n\hat{i}) = \nabla n,$$

where \hat{i} is a unit vector tangent to the ray at a point where the index of refraction is n .

30-2. (30.3) Reflection and refraction at the surface of a dense medium

Write down Fresnel's equations for the case where $\mu_{r1} = 1$, $\mu_{r2} = 1$, $n_2 \gg n_1$. You will find a surprising result: if the E vector of the incident wave is parallel to the plane of incidence, the amplitude of the reflected wave is independent of the angle of incidence!

For what range of θ_I are your formulas valid?

30-3. (30.3) Fresnel's equations expressed in terms of θ_I and θ_T alone

(a) First show that

$$\sin(\theta_I - \theta_T) \cos(\theta_I + \theta_T) = \sin \theta_I \cos \theta_I - \sin \theta_T \cos \theta_T,$$

$$\sin(\theta_I + \theta_T) \cos(\theta_I - \theta_T) = \sin \theta_I \cos \theta_I + \sin \theta_T \cos \theta_T.$$

(b) Show that, for nonmagnetic nonconductors,

$$(i) \left(\frac{E_{Rm}}{E_{Im}} \right)_{\perp} = -\frac{\sin(\theta_I - \theta_T)}{\sin(\theta_I + \theta_T)}, \quad (ii) \left(\frac{E_{Tm}}{E_{Im}} \right)_{\perp} = \frac{2 \cos \theta_I \sin \theta_T}{\sin(\theta_I + \theta_T)},$$

$$(iii) \left(\frac{E_{Rm}}{E_{Im}} \right)_{\parallel} = \frac{\tan(\theta_I - \theta_T)}{\tan(\theta_I + \theta_T)}, \quad (iv) \left(\frac{E_{Tm}}{E_{Im}} \right)_{\parallel} = \frac{2 \cos \theta_I \sin \theta_T}{\sin(\theta_I + \theta_T) \cos(\theta_I - \theta_T)}.$$

30-4. (30.3) Measuring an index of refraction

Set

$$p = \left(\frac{E_{Rm}}{E_{Im}} \right)_{\parallel}, \quad s = \left(\frac{E_{Rm}}{E_{Im}} \right)_{\perp}.$$

Show that, with a laser beam incident at 45° in air on a medium of index of refraction n ,

$$n^2 = \frac{(1-p)(1-s)}{(1+p)(1+s)}.$$

Here, s is negative, from Fig. 30-5. The ratio p is also negative, from Fig. 30-6.

In practice, instruments measure a beam *power*. So p is equal to *minus* the square root of the reflected to incident powers with parallel polarization, and similarly for s .

30-5. (30.5) The Brewster angle

Calculate the Brewster angles for the following cases:

- (a) light incident on a glass whose index of refraction is 1.6,
- (b) light emerging from the same type of glass,
- (c) a radio frequency wave incident on water ($n = 9$ at radio frequencies).

30-6. (30.5) The Brewster angle and the ratio n_2/n_1

- (a) Show that, if $n_2 > n_1$, then $\theta_{IB} > 45^\circ$.
- (b) Show that, if $n_2 < n_1$, then $\theta_{IB} < 45^\circ$.

30-7. (30.5) The Brewster angle

- (a) Show that Brewster's angle is also given by

$$\sin^2 \theta_{IB} = \frac{1}{1 + n_1^2/n_2^2}.$$

It follows that there exists a Brewster angle only if the ratio n_1/n_2 is real.

- (b) Show that $\sin \theta_{IB} = \cos \theta_T$.

30-8. (30.5) Brewster windows for lasers

The mirrors of some gas lasers are outside the glass tube that contains the discharge. Then the tube is sealed at both ends with windows set at the Brewster angle.

Show that there is no reflection from such a window as long as the E vector of the incident wave lies in the plane of incidence.

30-9. (30.5) The value of R_\perp at the faces of a dielectric plate set at the Brewster angle

A beam of light in a medium of index of refraction n_1 falls on a plate of dielectric n_2 at the Brewster angle.

- (a) Show that, at the first interface,

$$R_\perp = \cos^2 2\theta_{IB} = \left(\frac{1 - n_2^2/n_1^2}{1 + n_2^2/n_1^2} \right)^2.$$

- (b) Show that R_\perp has the same numerical value at the second interface.
- (c) Find the value of R_\perp for glass whose n is 1.5, in air.

30-10. (30.5) A “pile of plates” polarizer with pellicles

A *pellicle* is a very thin film of cellulose nitrate that is stretched taut over a flat ring. The cellulose nitrate is transparent and can serve as a

support for various types of coating. The film is so thin that multiple reflections inside it do not give rise to ghost images.

Now it has long been known that a series of parallel glass plates set at Brewster's angle filters out waves polarized with \mathbf{E} normal to the plane of incidence. See the two preceding problems. The same can be done with pellicles in less space and without the inconvenience of ghost images. Also, pellicles are virtually lossless and can thus polarize high-power laser beams.

(a) Find $(R/T)_\perp$ for a pellicle set at Brewster's angle. Take both interfaces into account, but disregard multiple reflections.

(b) Calculate this ratio for a pellicle whose n is 1.5 in air.

(c) Find a general expression for the ratio $(R/T)_\perp$ for N interfaces.

This result is grossly wrong because we have neglected multiple reflections inside the pellicles. In actual fact, the ratio is approximately equal to NR . With 40 interfaces, the above result is too large by 2 orders of magnitude!

30-11. (30.5) The Brewster angle for magnetic media

A wave is incident in air on a nonconducting magnetic medium such as ferrite.

(a) Show that the ratio $(E_{Rm}/E_{Im})_\parallel$ is zero for

$$\tan^2 \theta_I = \frac{\epsilon_r(\epsilon_r - \mu_r)}{\epsilon_r \mu_r - 1}.$$

There is a Brewster angle only if $\epsilon_r > \mu_r$.

(b) Show that $(E_{Rm}/E_{Im})_\perp$ is zero when

$$\tan^2 \theta_I = \frac{\mu_r(\mu_r - \epsilon_r)}{\epsilon_r \mu_r - 1}.$$

Now there is a Brewster angle, but only if $\mu_r > \epsilon_r$.

30-12. (30.6) The condition that makes $R = T$ at normal incidence

Find the ratio n_1/n_2 that makes $R = T = 0.5$ at normal incidence.

30-13. (30.6) E , H , R , and T at normal incidence on a water surface

A 60-watt light bulb is situated in air 1 meter above a water surface.

(a) Calculate the root mean square (rms) values of E and H for the incident, reflected, and refracted rays at the surface of the water directly under the bulb. Assume that all the power is dissipated as electromagnetic radiation. The index of refraction of water is 1.33.

(b) Calculate the coefficients of reflection and transmission.

30-14. (30.6) Antireflection coatings for photographic lenses and solar cells

There are instances where the reflection coefficient of a dielectric must be close to zero. The best known examples are photographic lenses and solar cells.

Clearly, the way to eliminate the reflected wave is by interference. Coating the dielectric with a thin film of another type of dielectric provides two reflected waves that can cancel. The situation is, however,

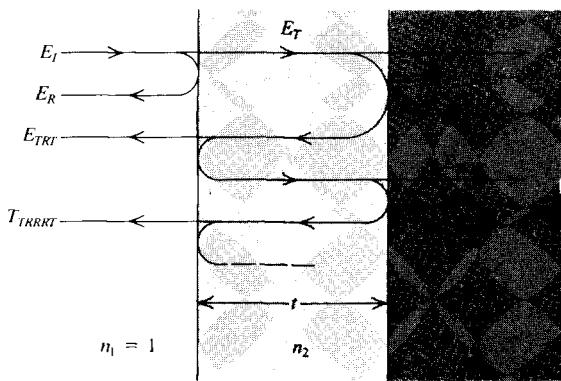


Fig. 30-12.

complicated by the presence of multiple reflections in the film. Also, the degree of cancellation varies with the angle of incidence and with the wave length.

(a) Show that there is no reflected wave at normal incidence in air ($n_1 = 1$) when the dielectric of index of refraction n_3 is coated with a quarter-wavelength film of a dielectric $n_2 = n_3^{1/2}$. Take multiple reflections into account, and use the notation of Fig. 30-12.

(b) Calculate and sum the amplitudes of the first four reflected waves when $n_3 = 4$, to four significant figures.

(c) A silicon solar cell has an index of refraction of 3.9 at 600 nanometers. Calculate the reflection coefficient for normal incidence at that wavelength.

(d) Calculate the thickness and the index of refraction of a coating that would eliminate reflection at normal incidence at that wavelength.

At the interface between air and glass, $R = 0.04$. In complex optical systems with many interfaces, the loss is important. Moreover, stray reflections reduce contrast in the image. Good-quality lenses are coated with magnesium fluoride ($n = 1.38$ at 550 nanometers). This reduces R to 0.015, on average, over the visible spectrum.

- 30-15. (30.6) A simple and accurate method for measuring an index of refraction

Possibly the most practical and most accurate way of measuring an index of refraction is to measure the ratio R_\perp/R_{\parallel} for a beam incident on the material in air at 45° .

(a) Show that, if $\theta_T = 45^\circ$,

$$R_\perp = \frac{1 - \sin 2\theta_T}{1 + \sin 2\theta_T}, \quad R_{\parallel} = \left(\frac{1 - \sin 2\theta_T}{1 + \sin 2\theta_T} \right)^2.$$

It follows that

$$\frac{R_\perp}{R_{\parallel}} = \frac{1 + \sin 2\theta_T}{1 - \sin 2\theta_T}.$$

This ratio is much larger than unity. For example, with $\theta_T = 30^\circ$ ($n = 1.414$), it is equal to 13.93.

(b) Show that

$$n = \frac{\{1 + (1 - f^2)^{1/2}\}^{1/2}}{f},$$

where

$$f = \sin 2\theta_T = \frac{R_\perp/R_\parallel - 1}{R_\perp/R_\parallel + 1}.$$

The signs before the square roots are positive.

30-16. (30.7) Ducting in the ionosphere

Under certain circumstances, the index of refraction of the ionosphere varies with altitude in such a way that a ray that starts out horizontally follows a path at a constant altitude above the earth's surface. The ionosphere then acts as a *duct*, and the phenomenon is called *ducting*. Of course, the required condition applies only over a certain distance. When the ray emerges from this region, it is deflected either upward or downward. Radar signals occasionally travel over large distances in this way.

- (a) How must the index of refraction vary with altitude?
- (b) How must the plasma frequency vary with altitude?

CHAPTER 31

PLANE ELECTROMAGNETIC WAVES IV

Reflection and Refraction B: Nonuniform Plane Waves and Total Reflection[†]

31.1 NONUNIFORM PLANE WAVES	579
31.1.1 COMPLEX ANGLES	580
31.2 TOTAL REFLECTION	581
31.2.1 THE REFLECTED WAVE	583
31.2.2 THE TRANSMITTED WAVE	583
Example: LIGHT EMISSION FROM A CATHODE RAY TUBE	588
Example: THE CRITICAL ANGLE AND THE BREWSTER ANGLE	589
31.3 SUMMARY	590
PROBLEMS	591

In Chap. 30 we established the laws of reflection and of refraction as well as Fresnel's equations. The laws concern the relative orientations of the vector wave numbers \mathbf{k}_I , \mathbf{k}_R , and \mathbf{k}_T , while Fresnel's equations concern the relative amplitudes and phases of the incident, reflected, and transmitted waves. We excluded total reflection as well as reflection by a conducting medium.

This chapter concerns total reflection, and the next one reflection by a conductor. The laws and equations that we found in Chap. 30 also apply to those cases, even though the wave number \mathbf{k}_T of the transmitted wave is then complex.

Total reflection occurs at large angles of incidence, when the incident wave lies in a medium whose index of refraction is larger than that of the second medium. More precisely, total reflection occurs when

[†] All the material in this chapter is essential only if you wish to study optical wave guides in Chaps. 35 and 36.

$$\sin \theta_T = \frac{n_1}{n_2} \sin \theta_I > 1. \quad (31-1)$$

Then the coefficient of reflection is equal to unity.

Total reflection is fascinating because of the unusual features of the *transmitted* wave. Also, optical waveguides (Chaps. 35 and 36) use the fact that total reflection is lossless. The analysis of those waveguides rests on some of the key results found here.

But first we must study briefly nonuniform plane waves and complex angles.

31.1 NONUNIFORM PLANE WAVES

In a *plane* wave the equiphase surfaces are plane. In a *uniform plane wave*, the amplitude is uniform throughout any given plane equiphase surface. The waves that we have referred to until now are of this type.

In a *nonuniform plane wave* the equiphase surfaces are again plane. However, the *amplitude* over a given *equiphase* surface is not uniform. Indeed, propagation occurs in one direction, and the amplitude of the wave decreases exponentially in another direction. In Fig. 31-5, for example, the transmitted wave travels from left to right, but its amplitude decreases exponentially downward.

With nonuniform plane waves we can still write that

$$\mathbf{E} = \mathbf{E}_m \exp j(\omega t - \mathbf{k} \cdot \mathbf{r}), \quad (31-2)$$

$$\mathbf{H} = \mathbf{H}_m \exp j(\omega t - \mathbf{k} \cdot \mathbf{r}), \quad (31-3)$$

where the amplitudes \mathbf{E}_m and \mathbf{H}_m may be complex. However, the wave vector then has the form

$$\mathbf{k} = \boldsymbol{\beta} - j\boldsymbol{\alpha}, \quad (31-4)$$

where the two real vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ point in different directions. Then

$$\mathbf{E} = \mathbf{E}_m \exp(-\boldsymbol{\alpha} \cdot \mathbf{r}) \exp j(\omega t - \boldsymbol{\beta} \cdot \mathbf{r}), \quad (31-5)$$

$$\mathbf{H} = \mathbf{H}_m \exp(-\boldsymbol{\alpha} \cdot \mathbf{r}) \exp j(\omega t - \boldsymbol{\beta} \cdot \mathbf{r}). \quad (31-6)$$

These equations define a wave that propagates in the positive direction of the vector $\boldsymbol{\beta}$ at the phase velocity

$$v = \frac{\omega}{\beta} \quad (31-7)$$

and whose wavelength is

$$\lambda = \frac{2\pi}{\beta}, \quad \text{hence} \quad \beta = \frac{1}{\lambda}. \quad (31-8)$$

The amplitude of the wave decreases exponentially in the positive direction of α , and the *attenuation distance* δ , over which the amplitude decreases by the factor of e , is given by

$$\delta = \frac{1}{\alpha}. \quad (31-9)$$

Equiphasic surfaces are perpendicular to β , and equal-amplitude surfaces are perpendicular to α . In the transmitted wave of Fig. 31-5, β points to the right, and α points downward.

If we substitute the above expression for E in the general wave equation 28-14, we find that

$$k^2 = (\beta - j\alpha)^2 = \beta^2 - \alpha^2 - 2j\alpha \cdot \beta = \omega^2 \epsilon \mu - j\omega \sigma \mu, \quad (31-10)$$

$$\beta^2 - \alpha^2 = \omega^2 \epsilon \mu, \quad 2\alpha \cdot \beta = \omega \sigma \mu, \quad (31-11)$$

where ϵ , μ , σ , and ω are all real and positive. These last two equations are interesting.

(1) The first one shows that $\beta > \alpha$. Then

$$\lambda < \delta. \quad (31-12)$$

(2) According to the second one, $\beta \cdot \alpha \geq 0$. Thus the angle θ between the direction of propagation and the direction of attenuation is at most 90° .

(3) In conductors, $\sigma \neq 0$. Then $\alpha \neq 0$. In good conductors $\beta^2 - \alpha^2 \ll 2\alpha \cdot \beta$.

(4) In nonconductors $\sigma = 0$ and θ is equal to 90° . The attenuation of a nonuniform plane wave in a dielectric can occur only in a direction perpendicular to the propagation.

Nonuniform plane waves are *not* transverse. If E is transverse, then H is not, and inversely. See Prob. 31-1.

31.1.1 Complex Angles

The magnitude of an angle A , expressed in radians, is equal to the pure number a/R , where a is the length of the arc of a circle of radius R whose center is at the apex.

By extension, A can be any complex number. If we apply Eq. 2-7,

$$\exp jA = \cos A + j \sin A, \quad (31-13)$$

then

$$\sin A = \frac{\exp jA - \exp (-jA)}{2j}, \quad \cos A = \frac{\exp jA + \exp (-jA)}{2}. \quad (31-14)$$

Clearly, $\sin A$ can now be any complex number. The same applies to $\cos A$. As with real angles,

$$\sin^2 A + \cos^2 A = 1. \quad (31-15)$$

We have generalized the concept of angle so as to render it more useful. But we now have an abstract quantity that does not lend itself to a simple geometric interpretation.

Remember that $\exp 2\pi j$ is equal to unity. You can easily check that

$$\sin(A + 2\pi) = \sin A, \quad \cos(A + 2\pi) = \cos A. \quad (31-16)$$

We select that value of A whose real part lies between 0 and 2π .

31.2 TOTAL REFLECTION

As we saw in Eq. 31-1, total reflection leads to a value of $\sin \theta_T$ that is real and larger than unity. Then θ_T is complex. Snell's law and Fresnel's equations still apply.

The *critical angle of incidence*, for which $\theta_T = 90^\circ$, is given by

$$\sin \theta_{lc} = \frac{n_2}{n_1}. \quad (31-17)$$

For a glass whose index of refraction n_1 is equal to 1.5, and with $n_2 = 1$, the critical angle is 41.8° .

At angles $\theta_I > \theta_{lc}$, $\sin \theta_T$ is larger than unity and θ_T is complex. Then the wave is *totally reflected* as in Fig. 31-1. This phenomenon is independent of the orientation of the E vector in the incident wave. There nonetheless exists a transmitted wave. Medium 2 acts like a pure inductance fed by a source of alternating voltage: the average power flow is zero, with power flowing alternately one way and then the other.

Total reflection has many uses, mostly based on the fact that the coefficient of reflection (Sec. 30.6) is then equal to unity if the interface is clean. (See Chaps. 35 and 36 on optical wave guides.) One relatively little-known application is *internal-reflection spectroscopy*, in which one

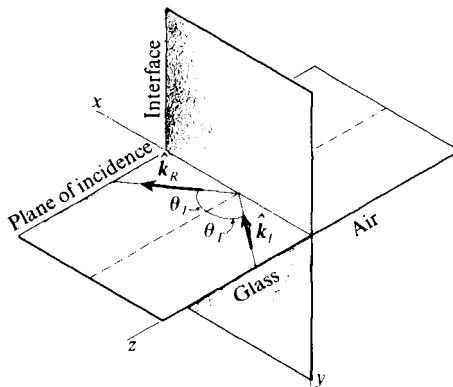


Fig. 31-1. For angles of incidence θ_i larger than the critical angle θ_{ic} , the interface acts as a perfect mirror. This is total reflection.

observes the spectrum of the reflected wave when the second medium is weakly absorbing. The method is convenient for investigating the second medium, for materials that do not lend themselves to conventional transmission or reflection spectroscopy.

Set

$$\theta_T = a + jb. \quad (31-18)$$

Then

$$\begin{aligned} \sin \theta_T &= \frac{\exp j(a+jb) - \exp(-j)(a+jb)}{2j} \\ &= \frac{\exp ja \exp(-b) - \exp(-ja) \exp b}{2j}. \end{aligned} \quad (31-19)$$

Since this quantity is real, a must be equal to $\pi/2$ and

$$\sin \theta_T = \frac{\exp b + \exp(-b)}{2} = \cosh b. \quad (31-20)$$

So

$$\theta_T = \frac{\pi}{2} + jb. \quad (31-21)$$

Then

$$\begin{aligned} \cos \theta_T &= \frac{\exp j(a+jb) + \exp(-j)(a+jb)}{2} \\ &= \frac{j[\exp(-b) - \exp b]}{2} = -j \sinh b. \end{aligned} \quad (31-22)$$

Also,

$$\cos \theta_T = (1 - \sin^2 \theta_T)^{1/2} = -j(\sin^2 \theta_T - 1)^{1/2}. \quad (31-23)$$

We have chosen a negative sign before j to agree with Eq. 31-22.

For example, if $n_1 = 1.5$ and $n_2 = 1.0$, then the critical angle of incidence is 41.8° . If now $\theta_I = 60^\circ$, then $\sin \theta_T = 1.299$, $b = 0.755$, $\theta_T = (\pi/2) + 0.755j$, and $\cos \theta_T = -0.829j$.

The incident, reflected, and transmitted waves are of the same form as in Sec. 30.2:

$$E_I = E_{Im} \exp j[\omega t - k_1(x \sin \theta_I - z \cos \theta_I)], \quad (31-24)$$

$$E_R = E_{Rm} \exp j[\omega t - k_1(x \sin \theta_I + z \cos \theta_I)], \quad (31-25)$$

$$E_T = E_{Tm} \exp j[\omega t - k_2(x \sin \theta_T - z \cos \theta_T)]. \quad (31-26)$$

31.2.1 The Reflected Wave

Applying Fresnel's equations for dielectrics given in Sec. 30.4, we find that

$$\left(\frac{E_{Rm}}{E_{Im}} \right)_\perp = \frac{(n_1/n_2) \cos \theta_I + j(\sin^2 \theta_T - 1)^{1/2}}{(n_1/n_2) \cos \theta_I - j(\sin^2 \theta_T - 1)^{1/2}} = \exp j\Phi_\perp, \quad (31-27)$$

where

$$\Phi_\perp = 2 \arctan \frac{(\sin^2 \theta_T - 1)^{1/2}}{(n_1/n_2) \cos \theta_I}. \quad (31-28)$$

This is the phase of the reflected wave with respect to the incident wave, at any point on the interface. The reflected wave *leads* the incident wave for this polarization. See Fig. 31-2.

Observe that the incident and reflected waves are of the same amplitude: *total reflection is lossless*.

The phase shift is different when E is parallel to the plane of incidence. (See Prob. 31-6.) Total reflection of a wave that is polarized in an arbitrary direction yields a reflected wave that is elliptically polarized.

For the above example, $\Phi_\perp = 95.7^\circ$.

31.2.2 The Transmitted Wave

The vector wave number for the transmitted wave is

$$\mathbf{k}_T = \boldsymbol{\beta}_T - j\alpha_T = k_2(\sin \theta_T \hat{x} - \cos \theta_T \hat{z}) \quad (31-29)$$

$$= k_2(\sin \theta_T \hat{x} + j \sinh b \hat{z}). \quad (31-30)$$

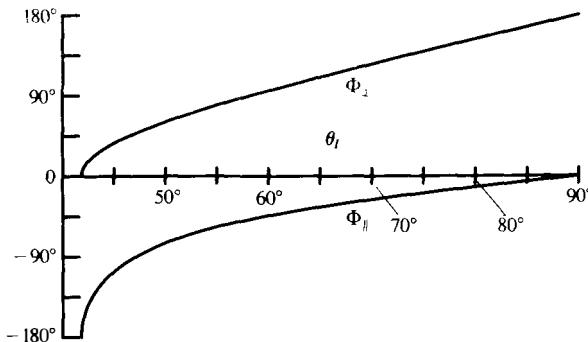


Fig. 31-2. The phases Φ_{\perp} and Φ_{\parallel} of the reflected wave with respect to the incident wave at a point on the interface, for total reflection when the \mathbf{E} of the incident wave is normal and parallel to the plane of incidence. The ratio n_1/n_2 is equal to 1.50.

The first term inside the parenthesis shows that the wave travels in the positive direction of the x -axis. The second term provides attenuation in the negative direction of the z -axis if b is positive. If b were negative, the wave amplitude would grow exponentially with depth inside the second medium, which is absurd.

The attenuation distance in the direction perpendicular to the interface is

$$\delta_z = \frac{1}{k_2 \sinh b} = \frac{\lambda_0}{n_2 \sinh b} = \frac{\lambda_0}{2\pi n_2 \sinh b} = \frac{\lambda_2}{2\pi \sinh b}. \quad (31-31)$$

For the above example, $\delta_z = \lambda_2/5.21$. Then the wave amplitude decreases by a factor of e over a distance of about $\lambda_2/5$. The transmitted wave barely penetrates into the second medium. The amplitude decreases by a factor of 183 over a distance equal to λ_2 ! See Fig. 31-3.

Applying again Fresnel's equations from Sec. 30.4,

$$\left(\frac{E_{Tm}}{E_{Im}} \right)_{\perp} = \frac{2 \cos \theta_I}{\cos \theta_I - j(n_2/n_1)(\sin^2 \theta_T - 1)^{1/2}} \quad (31-32)$$

$$= \frac{2 \cos \theta_I}{[\cos^2 \theta_I + (n_2/n_1)^2(\sin^2 \theta_T - 1)]^{1/2}} \exp j \frac{\Phi_{\perp}}{2} \quad (31-33)$$

$$= \frac{2 \cos \theta_I}{[1 - (n_2/n_1)^2]^{1/2}} \exp j \frac{\Phi_{\perp}}{2}. \quad (31-34)$$

Figure 31-4 shows how this ratio varies in amplitude and in phase with the angle of incidence θ_I . Its magnitude is equal to 1.34 for the above case.

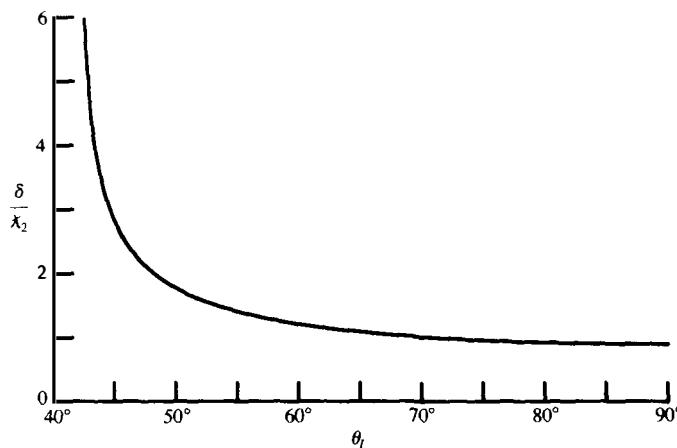


Fig. 31-3. The ratio δ/λ_2 , where δ is the depth of penetration for the transmitted wave and λ_2 is the wavelength in medium 2 divided by 2π , as a function of the angle of incidence θ_i .

Figure 31-5 shows “crests” of \mathbf{E} for the three waves in one specific case.

Let us calculate \mathbf{H}_T . Since the transmitted wave is not uniform, \mathbf{H}_T is not transverse (Prob. 31-1). Then Fig. 30-2 is of no use to calculate \mathbf{H}_T . The vector \mathbf{E} is normal to the plane of incidence. From Secs. 28.2.1 and 28.2.2, and with axes chosen as in Fig. 31-1,

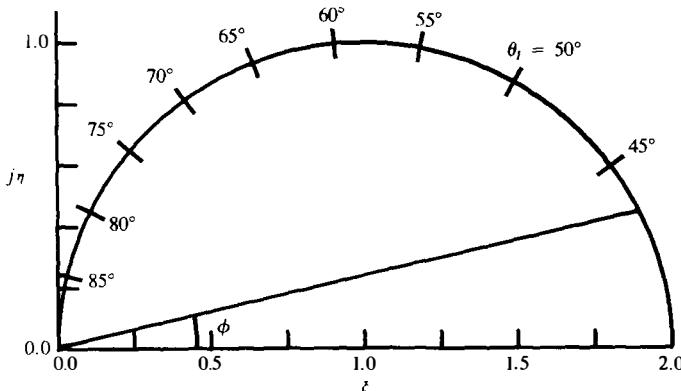


Fig. 31-4. The ratio $(E_{Tm}/E_{Im})_\perp = \xi + j\eta$, plotted in the complex plane for various angles of incidence θ_i larger than the critical angle, and for $n_1/n_2 = 1.50$. The amplitude of the transmitted wave is largest at the critical angle. The transmitted wave leads the incident wave by the angle ϕ .

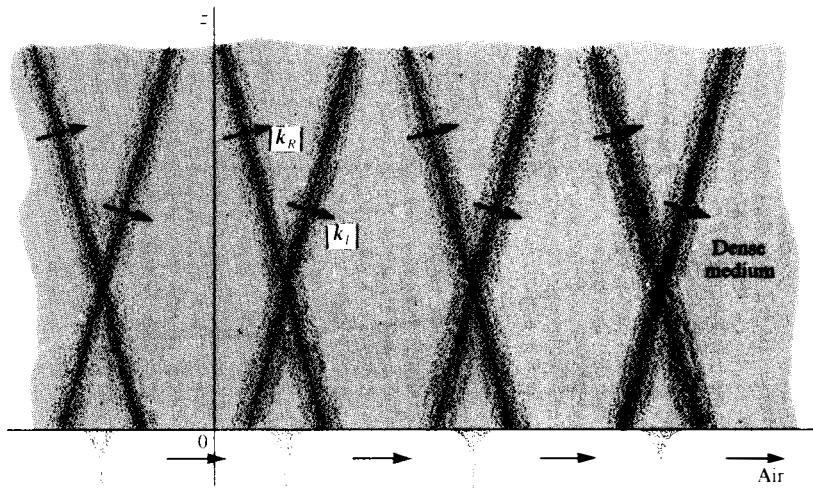


Fig. 31-5. “Crests” of E for the incident, reflected, and transmitted waves for $n_1 = 3.0$, $n_2 = 1.0$, $\theta_I = 75^\circ$.

$$\mathbf{H}_T = \frac{\mathbf{k}_T \times \mathbf{E}_T}{\omega \mu_2} = \frac{\mathbf{k}_T \times \mathbf{E}_T}{Z_2 k_2} = \frac{\mathbf{k}_T \times E_T \hat{\mathbf{y}}}{Z_2 k_2} \quad (31-35)$$

$$= \frac{(\sin \theta_T \hat{\mathbf{x}} - \cos \theta_T \hat{\mathbf{z}}) \times E_T \hat{\mathbf{y}}}{Z_2} = \frac{(\cos \theta_T \hat{\mathbf{x}} + \sin \theta_T \hat{\mathbf{z}}) E_T}{Z_2}. \quad (31-36)$$

Then

$$H_{Tmx} = \frac{E_{Tm}}{Z_2} \cos \theta_T, \quad H_{Tmz} = \frac{E_{Tm}}{Z_2} \sin \theta_T. \quad (31-37)$$

Since

$$H_{Tm} = \frac{E_{Tm}}{Z_2}, \quad (31-38)$$

then

$$H_{Tmx} = H_{Tm} \cos \theta_T, \quad H_{Tmz} = H_{Tm} \sin \theta_T, \quad (31-39)$$

precisely as when θ_T is real and \mathbf{H}_T transverse. Also,

$$H_{Tmx}^2 + H_{Tmz}^2 = H_{Tm}^2 (\cos^2 \theta_T + \sin^2 \theta_T) = H_{Tm}^2. \quad (31-40)$$

If \mathbf{E}_I is parallel to the plane of incidence, the formalism is again the same as when θ_T is real.

Since $\sin \theta_T$ is real and positive while $\cos \theta_T$ is imaginary and negative, the x component of \mathbf{H}_T lags the z component by $\pi/2$. The \mathbf{H}_T vector rotates at the angular velocity ω , as in Fig. 31-6.

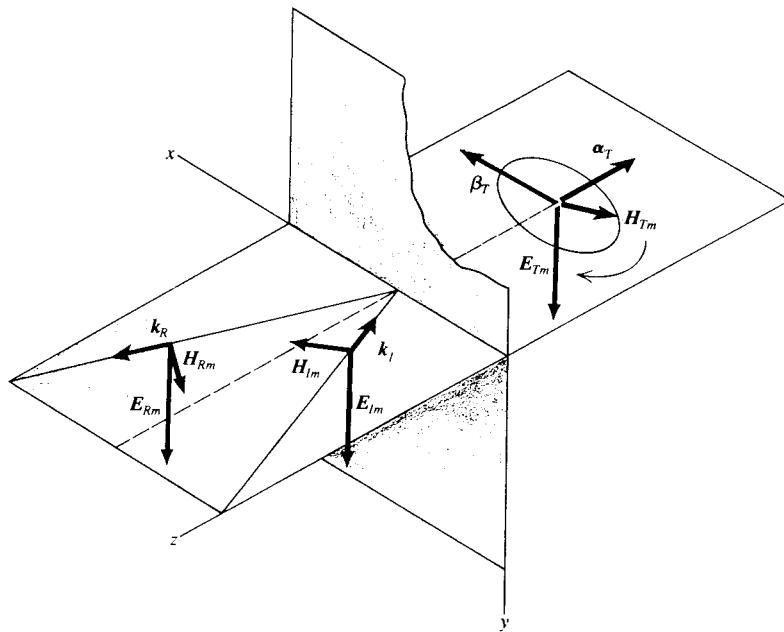


Fig. 31-6. The \mathbf{H} vector of the transmitted wave rotates in the direction shown. The transmitted wave travels to the left, and the rotation is the reverse of that of a wheel rolling on the interface.

There exists a power flow parallel to the interface in the positive direction of the x -axis. The flux is a function of θ_I , as in Fig. 31-7, for \mathbf{E} normal to the plane of incidence: for $n_2 = 1$,

$$\mathcal{G}_{Tav\perp} = \left(\frac{\epsilon_0}{\mu_0} \right)^{1/2} \frac{n_1 \cos \theta_I \sin 2\theta_I}{1 - (1/n_1)^2} E_{Im}^2 \exp \left[\frac{2}{\lambda_0} (n_1^2 \sin^2 \theta_I - 1)^{1/2} z \right] \hat{x}. \quad (31-41)$$

All this applies to an incident wave of infinite extent, but what happens if the incident wave has a finite cross section? Our analysis cannot provide an answer to this question. What happens is this: an incident ray penetrates into medium 2 and returns to medium 1 a bit farther along the x -axis. This is the *Goos-Haenchen shift*.[†]

[†] See Helmut K. V. Lotsch, *Optik*, vol. 32, pp. 116, 189, 299, 553 (1970 and 1971). There exists an analogous phenomenon in acoustics. See H. L. Bertoni and T. Tamir, *Applied Physics*, vol. 2, p. 157 (1973).

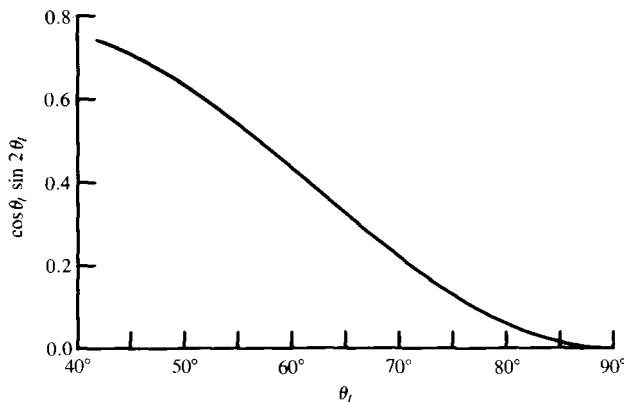


Fig. 31-7. The time-averaged Poynting vector parallel to the interface in medium 2 is proportional to $\cos \theta_i \sin 2\theta_i$, for total reflection. The vector \mathbf{E} is *normal* to the plane of incidence. We have set $n_1 = 1.50$ and $n_2 = 1.00$.

Example

LIGHT EMISSION FROM A CATHODE RAY TUBE

In a cathode ray tube the electron beam generates light in a fluorescent coating deposited on the back of the tube face. A given point in the fluorescent material radiates in all directions, but a thin layer of aluminum, as in Fig. 31-8, doubles the light output. Even then, most of the light stays trapped inside the glass by total reflection and travels back to the gun end of the tube.

What fraction F of the light comes out through the tube face? This is easy to calculate if we assume that inside the cone of angle

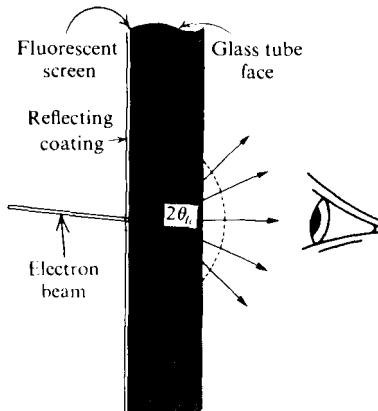


Fig. 31-8. Section through the face of a cathode ray tube.

θ_{lc} all the light crosses the glass-air interface. This is not a bad approximation because the coefficient of transmission is close to unity, except near the critical angle. Then F is the solid angle corresponding to θ_{lc} , divided by 2π (not 4π , because of the mirror). Since the cone defines a solid angle equal to the area of the spherical segment, shown as a dashed line in the figure, divided by R^2 ,

$$F = \frac{1}{2\pi} \int_0^{\theta_{lc}} \frac{2\pi R \sin \theta \, d\theta}{R^2} = 1 - \cos \theta_{lc}. \quad (31-42)$$

For a glass whose index of refraction is 1.5, $\theta_{lc} = 41.8^\circ$ and $F = 0.255$. The fraction F is, in fact, even smaller because of our approximation.

Example

THE CRITICAL ANGLE AND THE BREWSTER ANGLE

The critical angle (Sec. 31.2) is somewhat larger than the Brewster angle (Sec. 30.5). For example, again for light propagating inside a glass with an index of refraction n_1 of 1.5, the wave is totally transmitted into the air at a glass-air interface when the angle of incidence is the Brewster angle, 33.7° , and it is totally reflected back into the glass beyond the critical angle of 41.8° .

Figure 31-9 shows these two angles as functions of the ratio n_1/n_2 . For large values of n_1/n_2 , that is, for light incident in a

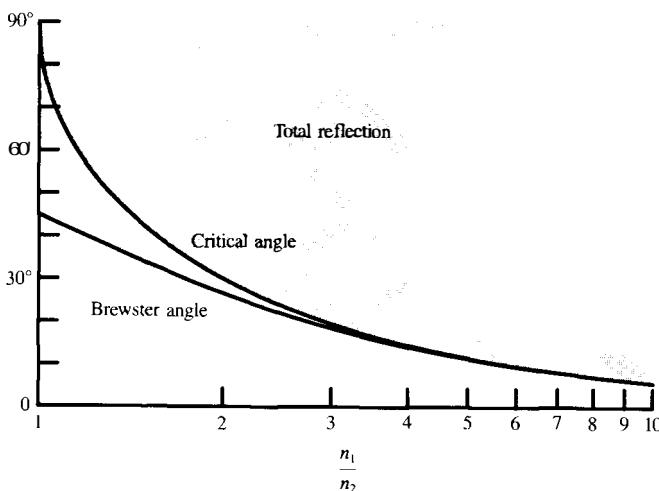


Fig. 31-9. The critical angle and the Brewster angle as functions of the ratio n_1/n_2 . The incident wave is linearly polarized with the E vector parallel to the plane of incidence for the Brewster angle curve.

relatively "dense" medium, θ_{lc} is nearly equal to θ_{IB} . For media with more similar indices of refraction, the Brewster angle approaches 45° whereas the critical angle approaches 90° .

For a wave polarized with its E vector parallel to the plane of incidence, the amplitude of the reflected wave changes rapidly when the angle of incidence lies between the Brewster angle and the critical angle. This peculiar behavior of the reflected and transmitted waves could be useful for measuring small angular displacements.

31.3 SUMMARY

In a *nonuniform plane wave*, propagation occurs in the direction perpendicular to the equiphase planes, and the amplitude decreases exponentially in a different direction. Then the *vector wave number* is of the form

$$\mathbf{k} = \boldsymbol{\beta} - j\boldsymbol{\alpha}, \quad (31-4)$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ point in different directions, with

$$\lambda = \frac{1}{\beta}, \quad \delta = \frac{1}{\alpha}, \quad (31-8), (31-9)$$

δ being the distance over which the amplitude decreases by a factor of e .

Normally, the magnitude A of an angle is a real number. By extension, A can be any complex number. Then

$$\sin A = \frac{\exp jA - \exp (-jA)}{2j}, \quad \cos A = \frac{\exp jA + \exp (-jA)}{2}, \quad (31-14)$$

$$\sin^2 A + \cos^2 A = 1, \quad (31-15)$$

Total reflection occurs at angles of incidence larger than the *critical angle* given by

$$\sin \theta_{lc} = \frac{n_2}{n_1}. \quad (31-17)$$

Then θ_T is complex:

$$\theta_T = \frac{\pi}{2} + jb, \quad (31-21)$$

where b is defined by

$$\sin \theta_T = \cosh b \quad (31-20)$$

Then

$$\cos \theta_T = -j \sinh b = -j(\sin^2 \theta_T - 1)^{1/2}. \quad (31-22), (31-23)$$

Total reflection is lossless, but there is a phase shift

$$\Phi_{\perp} = 2 \arctan \frac{(\sin^2 \theta_T - 1)^{1/2}}{(n_1/n_2) \cos \theta_I}. \quad (31-28)$$

The transmitted wave is nonuniform. It propagates parallel to the interface and its amplitude decreases exponentially perpendicular to the interface with an attenuation distance

$$\delta_z = \frac{\lambda_2}{2\pi \sinh b}. \quad (31-31)$$

Also,

$$\left(\frac{E_{Tm}}{E_{Im}} \right)_{\perp} = \frac{2 \cos \theta_I}{[1 - (n_2/n_1)^2]^{1/2}} \exp j \frac{\Phi_{\perp}}{2} \quad (31-34)$$

PROBLEMS

- 31-1. (31.1) A nonuniform plane electromagnetic wave is not transverse.

We define a transverse wave as one in which \mathbf{E} and \mathbf{H} are both perpendicular to the two vectors α and β .

(a) Write out Maxwell's equations for a plane sinusoidal wave in free space, replacing ∇ by $j\mathbf{k}$ and $\partial/\partial t$ by $j\omega$. Note that $\mathbf{k} \cdot \mathbf{E}$ is always equal to zero. Since the vector \mathbf{k} is complex, it has no specific orientation in space, except that it lies in the plane defined by the vectors α and β .

(b) Suppose that \mathbf{H} is transverse: $\mathbf{H} = H\hat{\mathbf{n}}$, where $\hat{\mathbf{n}} \cdot \alpha = 0$ and $\hat{\mathbf{n}} \cdot \beta = 0$.

Show that \mathbf{E} is then not transverse. Similarly, if \mathbf{E} is transverse, then \mathbf{H} is not transverse.

- 31-2. (31.2) Total reflection as in Fig. 31-5

An electromagnetic wave polarized with its \mathbf{E} vector normal to the plane of incidence is totally reflected as in Fig. 31-5 at the interface between a dielectric whose index of refraction is 3.0 and air. The angle of incidence is 75° .

(a) Calculate δ_z/λ_2 and δ_z/λ_1 .

(b) Calculate the phases of the reflected and transmitted waves with respect to the incident wave at any point on the interface.

(c) Check the continuity of \mathbf{E} across the interface.

- 31-3. (31.2) The Poynting vector for the transmitted wave

Check the value of $\mathcal{S}_{T,\text{av}\perp}$ given in Sec. 31.2.2, for $n_2 = 1$.

- 31-4. (31.2) Total reflection on a plasma when $\omega/\omega_p < 1$

Show that a wave incident on an ionized region is totally reflected if $\omega < \omega_p$, where ω_p is the plasma angular frequency (Sec. 29.2.3).

31-5. (31.2) The phase shifts Φ_{\parallel} and Φ_{\perp} in total reflection

(a) Show that

$$\Phi_{\parallel} = -2 \arctan \frac{\cos \theta_I}{(n_1/n_2)^2 (\sin^2 \theta_I - n_2^2/n_1^2)^{1/2}}.$$

(b) Show that

$$\tan \left(\frac{\Phi_{\perp}}{2} - \frac{\Phi_{\parallel}}{2} \right) = \frac{\sin^2 \theta_I}{\cos \theta_I (\sin^2 \theta_I - n_2^2/n_1^2)^{1/2}}.$$

(c) Plot $\Phi_{\perp} - \Phi_{\parallel}$ as a function of θ_I between 40° and 90° for $n_1 = 1.5$ and $n_2 = 1$.

31-6. (31.2) The transmitted wave when E is parallel to the plane of incidence
Show that

$$\left(\frac{E_{Tm}}{E_{Im}} \right)_{\parallel} = \frac{2}{[(n_2/n_1)^2 + 1]^{1/2}} \exp j \left(\frac{\Phi_{\parallel}}{2} - \frac{\pi}{2} \right).$$

31-7. (31.2) Scintillation particle detector

Figure 31-10 shows one type of *scintillation particle detector*. A scintillator S , usually made out of a single crystal of sodium iodide or of a suitable transparent plastic embedded in a reflector R , emits light when it is traversed by an ionizing particle such as an electron. A photomultiplier PM detects the emitted light.

The scintillator has an index of refraction n_1 and is fixed to the face of the photomultiplier with a cement C of index $n_2 < n_1$. Light is emitted in all directions in the scintillator, but only a fraction F reaches the photomultiplier.

(a) Calculate F as a function of n_1/n_2 , assuming that $T = 1$ for angles of incidence smaller than the critical angle and that the scintillator is surrounded by a nonreflecting substance.

(b) Draw a graph of F for values of n_2/n_1 ranging from 0.1 to 1.0.

31-8. (31.2) Total reflection in light-emitting diodes

In *light-emitting diodes* (LEDs), radiation occurs in a junction plane within a semiconductor whose index of refraction is quite large. For example, with GaAsP, $n = 3.5$. Total reflection at the semiconductor-air interface limits the efficiency of LEDs to a few percent.

(a) Calculate the critical angle.

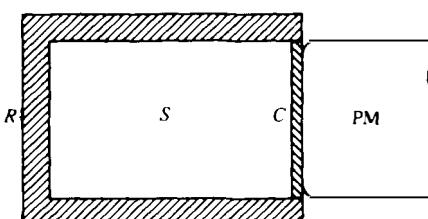


Fig. 31-10.

(b) Plot curves of T_{\parallel} and T_{\perp} for $n_1/n_2 = 3.5$. Note that the transmission coefficients are equal and approximately independent of the angle of incidence when θ_i is small. Identify the Brewster angle.

(c) Assume that the face of the semiconductor is flat and parallel to the junction. The index of refraction is n .

Calculate the fraction F of the light emitted at the source that reaches the surface at an angle smaller than the critical angle. Show that $F \approx 1/(4n^2)$.

(d) Show that $FT \approx 1/[n(n + 1)^2]$.

(e) Calculate F , T , and FT for $n = 3.5$.

(f) Calculate FT for an LED situated at the center of a hemisphere whose index of refraction is the same as that of the semiconductor. This is impractical because shaping the semiconductor is expensive.

(g) LEDs are usually covered with a hemispherical transparent resin whose n is about 1.6.

Calculate the two transmission coefficients and the efficiency. The efficiency is improved, but it is still very low.

31-9. (31.2) The \mathbf{H} vector of the transmitted wave

We found that the x and z components of the \mathbf{H}_T vector of the transmitted wave are in quadrature when there is total reflection with the \mathbf{E} vector normal to the plane of incidence.

Show that the vector rotates in the direction shown in Fig. 31-6.

CHAPTER 32

PLANE ELECTROMAGNETIC WAVES V

*Reflection and Refraction C: Reflection and
Refraction at the Surface of a Good Conductor[†]*

32.1	REFLECTION AND REFRACTION AT THE SURFACE OF A GOOD CONDUCTOR	595
32.1.1	<i>E</i> NORMAL TO THE PLANE OF INCIDENCE	597
32.1.2	<i>E</i> PARALLEL TO THE PLANE OF INCIDENCE	597
	Example: COMMUNICATING WITH SUBMARINES AT SEA	598
	Example: STANDING WAVES AT NORMAL INCIDENCE ON A GOOD CONDUCTOR	598
*32.2	RADIATION PRESSURE ON A NONMAGNETIC GOOD CONDUCTOR	599
*32.2.1	<i>E</i> NORMAL TO THE PLANE OF INCIDENCE	599
*32.2.2	<i>E</i> PARALLEL TO THE PLANE OF INCIDENCE	602
*32.2.3	THE MOMENTUM FLUX DENSITY AND THE MOMENTUM DENSITY IN AN ELECTROMAGNETIC WAVE	602
*32.2.4	THE $\cos^2 \theta$, TERM IN THE EXPRESSION FOR THE RADIATION PRESSURE	603
	Examples	603
*32.3	THE ELECTROMAGNETIC MOMENTUM OF STATIC FIELDS	604
32.4	SUMMARY	604
	PROBLEMS	605

Reflection and refraction at the surface of a good conductor are somewhat similar to total reflection in that the angle of refraction is again complex. However, with good conductors, the imaginary part of θ_T is negligible and $\theta_T \approx 0$. Then the transmitted wave is approximately uniform, but highly attenuated.

Radiation pressure on a conductor results from a Hall effect on the conduction electrons moving in the magnetic field of the wave.

[†]This chapter is a prerequisite only for Chap. 34.

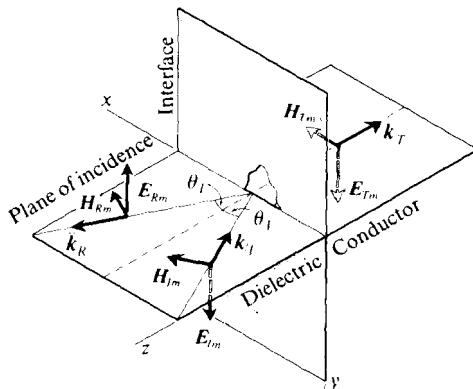


Fig. 32-1. The incident, reflected, and transmitted waves at the interface between a dielectric and a good conductor. The \mathbf{E} vector of the incident wave is *normal* to the plane of incidence.

32.1 REFLECTION AND REFRACTION AT THE SURFACE OF A GOOD CONDUCTOR

As previously, the incident and reflected waves lie in medium 1. Here, medium 1 is a dielectric, while medium 2 is a good conductor, as in Figs. 32-1 and 32-2.

Equations 30-17 to 30-20 apply to any pair of linear and isotropic media. We may therefore write that

$$\mathbf{E}_i = \mathbf{E}_{im} \exp j(\omega t - k_i x \sin \theta_i + k_i z \cos \theta_i), \quad (32-1)$$

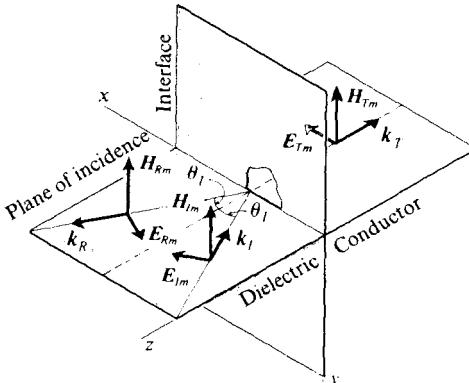


Fig. 32-2. Reflection and refraction at the surface of a good conductor. The \mathbf{E} of the incident wave is *parallel* to the plane of incidence.

$$\mathbf{E}_R = \mathbf{E}_{Rm} \exp j(\omega t - k_1 x \sin \theta_I - k_1 z \cos \theta_I), \quad (32-2)$$

$$\mathbf{E}_T = \mathbf{E}_{Tm} \exp j(\omega t - k_1 x \sin \theta_I + k_2 z \cos \theta_T) \quad (32-3)$$

$$= \mathbf{E}_{Tm} \exp j \left\{ \omega t - k_1 x \sin \theta_I \pm k_2 z \left[1 - \left(\frac{n_1}{n_2} \right)^2 \sin^2 \theta_I \right]^{1/2} \right\}. \quad (32-4)$$

Now, from Sec. 29.1,

$$\frac{n_1}{n_2} = \frac{k_1}{k_2} = \frac{\omega(\epsilon_1 \mu_1)^{1/2} \delta}{1-j} \quad (32-5)$$

$$= \frac{\delta}{(1-j)\lambda_1}. \quad (32-6)$$

We shall assume that

$$\frac{\delta}{2^{1/2}\lambda_1} \ll 1 \quad \text{or that} \quad \left| \frac{n_1}{n_2} \right| \ll 1. \quad (32-7)$$

See Sec. 29.1. Then the expression in brackets in Eq. 32-4 is approximately equal to unity. Then

$$\mathbf{E}_T = \mathbf{E}_{Tm} \exp j \left(\omega t - k_1 x \sin \theta_I \pm \frac{(1-j)z}{\delta} \right) \quad (32-8)$$

$$= \mathbf{E}_{Tm} \exp \left[j(\omega t - k_1 x \sin \theta_I) \pm \frac{(1+j)z}{\delta} \right]. \quad (32-9)$$

We select the plus sign before the z term so that \mathbf{E}_T will tend to zero as z tends to minus infinity. So we need a plus sign in Eq. 32-4, and for reflection from a good conductor,

$$\cos \theta_T = + \left[1 - \left(\frac{n_1}{n_2} \right)^2 \sin^2 \theta_I \right]^{1/2} \approx 1, \quad \theta_T \approx 0. \quad (32-10)$$

Therefore

$$\mathbf{E}_T \approx \mathbf{E}_{Tm} \exp \left[j \left(\omega t - k_1 x \sin \theta_I + \frac{z}{\delta} \right) + \frac{z}{\delta} \right]. \quad (32-11)$$

But

$$\frac{k_1 \sin \theta_I}{1/\delta} \leq k_1 \delta = k_1 \left(\frac{2}{\omega \sigma \mu} \right)^{1/2} = 2^{1/2} \left| \frac{k_1}{k_2} \right| \ll 1. \quad (32-12)$$

Then

$$\mathbf{E}_T \approx \mathbf{E}_{Tm} \exp \left[j \left(\omega t + \frac{z}{\delta} \right) + \frac{z}{\delta} \right], \quad (32-13)$$

and the transmitted wave propagates into the conductor along the normal to the interface, whatever the angle of incidence. Also, the amplitude of the transmitted wave decreases by a factor of e over one skin depth δ .

32.1.1 \mathbf{E} Normal to the Plane of Incidence

Refer to Fig. 32-1. From Fresnel's equations with $|n_1/n_2| \ll 1$,

$$\left(\frac{E_{Rm}}{E_{Im}}\right)_\perp = \frac{(n_1/n_2) \cos \theta_I - \cos \theta_T}{(n_1/n_2) \cos \theta_I + \cos \theta_T} \approx -1 \quad (32-14)$$

for any angle of incidence. Reflection from a superconductor ($n_2 \rightarrow \infty$) is lossless. The negative sign means that the \mathbf{E} vector of the reflected wave is in the direction shown in Fig. 32-1, opposite to that shown in Fig. 30-2, because we based our calculation on the latter figure.

Also, from Sec. 30.3 and from the fact that $\cos \theta_T \approx 1$,

$$\left(\frac{E_{Tm}}{E_{Im}}\right)_\perp = \frac{2(n_1/n_2) \cos \theta_I}{(n_1/n_2) \cos \theta_I + \cos \theta_T} \approx 2\frac{n_1}{n_2} \cos \theta_I \approx 0, \quad (32-15)$$

again for any angle of incidence θ_I .

At the surface of a dielectric such that $n_2 \gg n_1$ one also has that

$$\left(\frac{E_{Rm}}{E_{Im}}\right)_\perp \approx -1, \quad \left(\frac{E_{Tm}}{E_{Im}}\right)_\perp \approx 0. \quad (32-16)$$

32.1.2 \mathbf{E} Parallel to the Plane of Incidence

Refer now to Fig. 32-2. From Fresnel's equations

$$\begin{aligned} \left(\frac{E_{Rm}}{E_{Im}}\right)_\parallel &= \frac{(n_1/n_2) \cos \theta_T - \cos \theta_I}{(n_1/n_2) \cos \theta_T + \cos \theta_I} \\ &\approx \frac{(n_1/n_2) - \cos \theta_I}{(n_1/n_2) + \cos \theta_I} \approx -1. \end{aligned} \quad (32-17)$$

The last approximation is *not* valid at grazing incidence, where θ_I is close to 90° . Also,

$$\left(\frac{E_{Tm}}{E_{Im}}\right)_\parallel = \frac{2(n_1/n_2) \cos \theta_I}{\cos \theta_I + (n_1/n_2) \cos \theta_T} \quad (32-18)$$

$$\approx \frac{2n_1}{n_2} = \frac{2n_1\delta}{1-j} = n_1\delta(1+j). \quad (32-19)$$

The approximation is again invalid at grazing incidence.

Example

COMMUNICATING WITH SUBMARINES AT SEA

For shore-to-ship communication with the submarine antenna submerged, the efficiency is very low, first, because of the large coefficient of reflection at the surface of the sea and, second, because of the high attenuation in seawater. The attenuation in seawater is about 172 decibels/meter at 20 megahertz, 5.5 at 20 kilohertz, and 0.33 at 75 hertz. One solution is to operate at low frequencies (about 75 hertz and 17 to 25 kilohertz) and very high power, with huge transmitting antennas, many kilometers on the side.

Another solution for shore-to-ship communication is to modulate a laser beam emitted by a satellite, seawater being quite transparent to blue-green light. Remember that our discussion on the propagation of electromagnetic waves in conductors disregards atomic and molecular phenomena and is valid only up to roughly 1 gigahertz. Optical frequencies are of the order of 10^{15} hertz.

Ship-to-shore communication at low frequencies is impossible with long radio waves because a submarine can neither supply the required power nor deploy a long enough antenna. Two-way communication takes place at a few megahertz with the submarine antenna projecting above the water.

Example

STANDING WAVES AT NORMAL INCIDENCE ON A GOOD CONDUCTOR

Figure 32-3 shows the incident, reflected, and transmitted waves. Since the direction of propagation of the reflected wave is opposite to that of the incident wave, and since $\mathbf{E} \times \mathbf{H}$ points in the direction of propagation, the \mathbf{H} of the reflected wave is in phase with that of the incident wave at the interface, as in the figure. At the reflecting surface, the electric fields nearly cancel and there is a node of \mathbf{E} ; the magnetic fields add, and there is a loop of \mathbf{H} , as

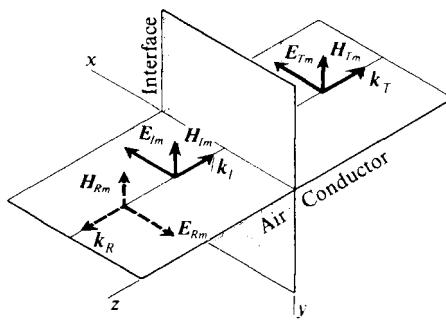


Fig. 32-3. Reflection at normal incidence from the surface of a good conductor: $E_{Rm} \approx -E_{Im}$ and $E_{Tm} \ll E_{Im}$.

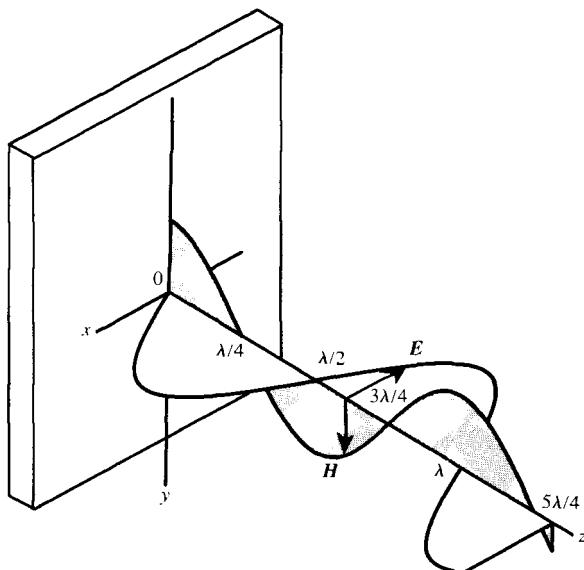


Fig. 32-4. The standing wave pattern for reflection at normal incidence on a good conductor, at a particular time. Nodes of \mathbf{E} and of \mathbf{H} are spaced $\lambda/4$ apart.

in Fig. 32-4. The nodes of \mathbf{E} and \mathbf{H} are thus one-quarter wave-length apart. The energy density is uniform.

A similar situation exists for reflection from any surface. Either the \mathbf{E} or the \mathbf{H} vector must change direction on reflection, in order to change the direction of the Poynting vector $\mathbf{E} \times \mathbf{H}$.

*32.2 RADIATION PRESSURE ON A NONMAGNETIC GOOD CONDUCTOR

*32.2.1 \mathbf{E} Normal to the Plane of Incidence

With \mathbf{E} normal to the plane of incidence, \mathbf{E} is tangent to the interface, there are no surface charges, and there is no surface force.

In the body of the conductor the current density \mathbf{J} is $\sigma \mathbf{E}_T$. Thus \mathbf{J} is parallel to \mathbf{E}_T and perpendicular to \mathbf{H}_T . It turns out, as we shall see, that the $Q\mathbf{v} \times \mu_0 \mathbf{H}$ force pushes the conduction electrons away from the surface. This is just another manifestation of the Hall effect of the first example in Sec. 22.1.1. The resulting electric force per unit area is the *radiation pressure*.

The situation would be different if the conduction electrons were entirely free to move through the metal: the conductivity would then be imaginary, as in the low-pressure ionized gases of Chap. 29, the conduction current would lag the field by $\pi/2$ radians, and there would be zero radiation pressure. In a superconductor σ is real and tends to infinity, δ tends to zero, and the following discussion applies.

We set

$$\left| \frac{n_1}{n_2} \right| \ll 1 \quad (32-20)$$

as in Sec. 32.1 except that, now, $\mu_{r2} = 1$. We consider an element of volume, as in Fig. 32-5, parallel to the interface, of area ab and thickness dz . It carries a current $\sigma E_T b dz$ and is subjected to a magnetic force $\sigma E_T ab dz \mu_0 H_T$ in the negative direction of the z -axis. Then the instantaneous pressure exerted on the element of thickness dz is

$$dp = \sigma E_T \mu_0 H_T dz, \quad (32-21)$$

with the positive directions for E_T and H_T chosen as in Fig. 32-5. A positive result will show that the incident wave pushes on the conductor.

From Eq. 32-13, the phasor for E_T is

$$\mathbf{E}_T = \mathbf{E}_{Tm} \exp \left[j \left(\omega t + \frac{z}{\delta} \right) + \frac{z}{\delta} \right], \quad (32-22)$$

while the corresponding phasor for H_T follows from Sec. 29.1:

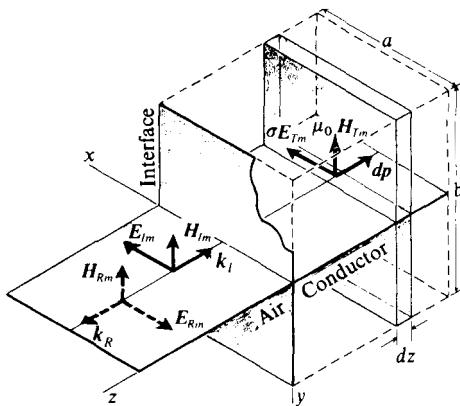


Fig. 32-5. Element of volume of thickness dz inside a conductor.

$$\mathbf{H}_T \approx \left(\frac{\sigma}{\omega \mu_0} \right)^{1/2} \exp \left(-j \frac{\pi}{4} \right) \mathbf{E}_T = \frac{1-j}{\omega \mu_0 \delta} \mathbf{E}_T. \quad (32-23)$$

The value of \mathbf{E}_{Tm} depends on the polarization of the incident wave.

The time-averaged value of the element of pressure on a sheet of thickness dz inside the conductor is thus

$$dp_{av,\perp} = \frac{1}{2} \sigma \mu_0 \operatorname{Re} (E_T H_T^*) dz = \frac{1}{2} \sigma \mu_0 \operatorname{Re} \left(E_T \frac{1+j}{\omega \mu_0 \delta} E_T^* \right) dz \quad (32-24)$$

$$= \frac{\sigma}{2\omega\delta} |E_{Tm}|^2 \exp \frac{2z}{\delta} dz \quad (32-25)$$

and

$$p_\perp = \frac{\sigma}{2\omega\delta} |E_{Tm}|^2 \int_{-\infty}^0 \exp \frac{2z}{\delta} dz = \frac{\sigma}{2\omega\delta} |E_{Tm}|^2 \frac{\delta}{2} = \frac{\sigma}{4\omega} |E_{Tm}|^2. \quad (32-26)$$

However, we need p_\perp in terms of the input power flux (Sec. 28.4)

$$\mathcal{S}_{I,av} = \frac{1}{2} v \epsilon_1 E_{Im}^2 = \frac{1}{2} c \epsilon_0 \epsilon_{r1}^{1/2} E_{Im}^2. \quad (32-27)$$

So we express p_\perp first in terms of E_{Im} and then in terms of $\mathcal{S}_{I,av}$. From Sec. 32.1.1,

$$\left(\frac{E_{Tm}}{E_{Im}} \right)_\perp \approx 2 \frac{n_1}{n_2} \cos \theta_I, \quad (32-28)$$

even at grazing incidence. Then

$$p_{\perp,av} = \frac{\sigma}{4\omega} \left| 2 \frac{n_1}{n_2} \cos \theta_I E_{Im} \right|^2 \quad (32-29)$$

with

$$\frac{n_1}{n_2} = \left(\frac{\omega \epsilon_1}{\sigma} \right)^{1/2}, \quad (32-30)$$

both media 1 and 2 being nonmagnetic. Then

$$\begin{aligned} p_{\perp,av} &= \frac{\sigma}{\omega} \left(\frac{\omega \epsilon_1}{\sigma} \right)^{1/2} \cos^2 \theta_I E_{Im}^2 = 2 \epsilon_1 \cos^2 \theta_I \left(\frac{\mu_1}{\epsilon_1} \right)^{1/2} \mathcal{S}_{I,av} \\ &= 2(\epsilon_1 \mu_1)^{1/2} \cos^2 \theta_I \mathcal{S}_{I,av}. \end{aligned} \quad (32-31)$$

Substituting the index of refraction n_1 of medium 1 for $\epsilon_{r1}^{1/2}$ and setting $\mu_1 = \mu_0$ yields

$$p_{\perp,\text{av}} = \frac{2n_1}{c} \cos^2 \theta_I \mathcal{S}_{I,\text{av}} = 2\mathcal{E}' \cos^2 \theta_I, \quad (32-32)$$

where \mathcal{E}' is the energy density in the wave.

*32.2.2 *E* Parallel to the Plane of Incidence

We do not go through the calculation here. As one might guess, the pressure is the same as above. However, the calculation is somewhat tricky: there are now two forces, a magnetic force (as above) plus an electric force exerted by the electric field in the dielectric on the surface charge.

*32.2.3 The Momentum Flux Density and the Momentum Density in an Electromagnetic Wave

Suppose medium 1 is a vacuum. Then at normal incidence the radiation pressure exerted on medium 2, the conductor, is $2\mathcal{S}_{I,\text{av}}/c$. Since the conducting surface acts as a near-perfect reflector, the pressure corresponds to a change in the momentum of the wave of $2\mathcal{S}_{I,\text{av}}/c$ per unit time and per unit area. Then, in the incident wave,

$$\text{Momentum flux density} = \frac{\mathcal{S}_{I,\text{av}}}{c}. \quad (32-33)$$

Now

$$\text{Momentum flux density} = \text{momentum volume density} \times c. \quad (32-34)$$

It follows that, in a uniform plane electromagnetic wave propagating in a vacuum,

$$\text{Momentum volume density} = \frac{\mathcal{S}_{I,\text{av}}}{c^2} = \frac{E_{\text{rms}} H_{\text{rms}}}{c^2} = \frac{\text{energy density}}{c}. \quad (32-35)$$

These results agree with those of atomic physics, where we associate electromagnetic waves with photons of energy $\hbar\omega$ ($\hbar = 1.05 \times 10^{-34}$ is Planck's constant divided by 2π) and momentum \hbar/λ traveling at a speed c . Thus, for one photon,

$$\frac{\text{Momentum}}{\text{Energy}} = \frac{\hbar/\lambda}{\hbar\omega} = \frac{1}{\lambda\omega} = \frac{1}{c}. \quad (32-36)$$

At the beginning of Sec. 32.2 we restricted ourselves to nonmagnetic media and assumed a magnetic force of the form $Q\mathbf{v} \times \mu_0\mathbf{H}$. What if the conductor is magnetic? The pressure is the same, for otherwise there

would not be conservation of momentum. This is a proof that the magnetic force is independent of μ , for slow electrons.

One interesting application of the momentum flux in an electromagnetic wave is the levitation of transparent particles in a vertical light beam. See Prob. 32-14.[†]

*32.2.4 The $\cos^2 \theta_I$ Term in the Expression for the Radiation Pressure

It is easy to explain the presence of the $\cos^2 \theta_I$ term in the radiation pressure if one imagines photons raining on the surface of the conductor. Let medium 1 be a vacuum. Say their energy density is \mathcal{E}' joules/meter³. Then \mathcal{S} is equal to $\mathcal{E}'c$, their momentum density is \mathcal{E}'/c , and their momentum flux density is \mathcal{E}' . The change in the component of the momentum flux density normal to the interface is $2\mathcal{E}' \cos \theta_I$.

Suppose the incident beam has a cross-sectional area of 1 meter². It illuminates an area of $1/\cos \theta_I$ meters². Then

$$p_{av} = (2\mathcal{E}' \cos \theta_I) \cos \theta_I = 2\mathcal{E}' \cos^2 \theta_I = 2 \frac{\mathcal{S}_{I,av}}{c} \cos^2 \theta_I. \quad (32-37)$$

Examples

Under ordinary circumstances, radiation pressure is weak and difficult to observe. In *sunlight*, at the top of the atmosphere, $\mathcal{S}_{I,av}$ is about 1.4 kilowatts/meter², and the radiation pressure on a metallic reflector is about 10^{-5} pascal, or about 10^{-10} atmosphere.

At the *surface of the sun*, the radiation pressure is larger by the factor

$$\left(\frac{\text{Distance from sun to earth}}{\text{Radius of sun}} \right)^2 = \left(\frac{1.5 \times 10^{11}}{7 \times 10^8} \right)^2 = 4.6 \times 10^4. \quad (32-38)$$

This gives a radiation pressure of only 5×10^{-6} atmosphere. Radiation pressure is unimportant even in the interior of the sun, but it may play an important role in the more luminous stars.

Comet tails point predominantly away from the sun. This phenomenon results partly from radiation pressure and partly from the solar wind (Prob. 28-12).

A *photon-drag detector* consists of a crystal of germanium with electrodes plated on each end. When the beam of a powerful CO₂ laser passes through the crystal along its axis, the conduction electrons drift forward. The voltage difference between the electrodes is a measure of the beam power. See Prob. 32-12.

[†] A. Askin, *Science*, vol. 210, p. 1081 (1980).

*32.3 THE ELECTROMAGNETIC MOMENTUM OF STATIC FIELDS

We have seen above that an electromagnetic wave possesses a momentum density that is proportional to $\mathbf{E} \times \mathbf{H}$, at least in a vacuum. Two questions come to mind. (1) What if the frequency is zero? The vector product makes no reference to frequency. Does this expression apply to static fields? (2) In an electromagnetic wave, there exists a mathematical relation between \mathbf{E} and \mathbf{H} : given \mathbf{E} , \mathbf{H} follows, and inversely. However, in static fields, \mathbf{E} bears no relation to \mathbf{H} . For example, the electric field might be that of some charged body and the magnetic field that of a permanent magnet. Does the above expression apply to *any* pair of unrelated \mathbf{E} and \mathbf{H} fields? Say one brings a permanent magnet near a charged capacitor. Does the field possess momentum? What if the voltage across the capacitor is time-dependent? Is the field momentum also time-dependent?

The answer to these two questions is by now well established: the momentum density of *any* electromagnetic field in a vacuum is $\mathbf{E} \times \mathbf{H}/c^2$. This fact has been demonstrated by thought experiments in which one establishes a field where $\mathbf{E} \times \mathbf{H}$ is not zero, taking into account the magnetic forces and torques, as well as the mechanical forces and torques required to keep the system immobile. To satisfy the law of conservation of angular momentum, the mechanical torque integrated over time is equal to the angular momentum of the field. In this way one always finds that the electromagnetic momentum density is $\mathbf{E} \times \mathbf{H}/c^2$ in a vacuum. See Prob. 32-15.

32.4 SUMMARY

Reflection at the surface of a *good conductor* is slightly lossy. The transmitted wave is weak, highly damped, and travels in a direction nearly perpendicular to the interface.

When electromagnetic radiation illuminates a conductor, the transmitted wave's electric field gives rise to a conduction current that flows in the wave's magnetic field, with the result that the cloud of conduction electrons is pushed back. This is essentially a Hall effect and shows up as radiation pressure:

$$p_{\perp,\text{av}} = p_{\parallel,\text{av}} = \frac{2n_1}{c} \cos^2 \theta_I \mathcal{S}_{I,\text{av}}, \quad (32-32)$$

Alternatively, radiation pressure results from the fact that the normal component of the momentum of the incident wave reverses upon reflection.

In any electromagnetic field, static or not, the momentum density in a vacuum is $\mathbf{E} \times \mathbf{H}/c^2$.

PROBLEMS

32-1. (32.1) Reflection from a good conductor

Draw two figures similar to those of Fig. 30-4, showing \mathbf{E} and \mathbf{H} for an electromagnetic wave incident on a good conductor. You will, of course, have to exaggerate the values of E_m and of λ in the conductor. Be sure to show the phases correctly. Show x -, y -, z -axes on both figures to relate one with the other.

32-2. (32.1) Reflection from a good conductor

Show that, for a good nonmagnetic conductor in air,

$$(a) \left| \frac{E_{Rm}}{E_{Im}} \right|_{\perp} \approx 1 - \frac{\delta}{\lambda_0} \cos \theta_i, \quad (b) \left| \frac{E_{Rm}}{E_{Im}} \right|_{\parallel} \approx 1 - \frac{\delta}{\lambda_0 \cos \theta_i}.$$

This latter relation is not valid at grazing incidence, where $\cos \theta_i$ tends to zero.

A good conductor is a better reflector when \mathbf{E} is normal to the plane of incidence. High-quality metallic reflectors have coefficients of reflection of about 90% near normal incidence in the visible, with unpolarized light.

32-3. (32.1) $|E_{Rm}/E_{Im}|$ as a function of the angle of incidence for reflection on a conductor

For a good conductor, $\sigma/\omega\epsilon \geq 50$. Then

$$\frac{\lambda_0}{\delta} = \frac{c}{\omega} \left(\frac{\omega\sigma\mu}{2} \right)^{1/2} = c \left(\frac{\sigma\mu\epsilon}{2\omega\epsilon} \right)^{1/2} \geq c(25\epsilon\mu)^{1/2} = 5(\epsilon_r\mu_r)^{1/2}.$$

So $\lambda_0/\delta \geq 10$ if $\epsilon_r = 4$ and $\mu_r = 1$.

Plot $|E_{Rm}/E_{Im}|_{\perp}$ and $|E_{Rm}/E_{Im}|_{\parallel}$ as functions of θ_i , for a nonmagnetic good conductor in air and for $\lambda_0/\delta = 10$. You will find that, when \mathbf{E} is in the plane of incidence, there exists a pseudo-Brewster angle for which the amplitude of the reflected wave is minimum.

32-4. (32.1) Liquid-crystal displays (LCDs)

In *liquid-crystal displays* the liquid is sandwiched between a transparent multiple electrode in front and a single black electrode in the back. Upon application of a voltage to a portion of the front window, the rodlike molecules of the *nematic* fluid in that region stand perpendicular to the window, and one can see the black electrode in the back. Elsewhere, the molecules reflect light because their orientations are haphazard.

The transparent multiple electrode is a thin coating either of semiconducting metal oxide, such as tin oxide, or of gold. The surface resistance (Prob. 4-9) is of the order of 10 to 100 ohms per square. At 600 nanometers, and for gold, $\beta\lambda_0 = 1.29$ and $\alpha\lambda_0 = 2.59$. The conductivity of gold in the form of a thin film is 4.26×10^7 siemens/meter.

- (a) Calculate the skin depth δ .
- (b) By what factor does the amplitude decrease in the gold film if its thickness s is 0.05δ ?
- (c) What is the surface resistance?
- (d) Calculate the thickness of the film in wavelengths λ_0 .

A proper calculation of the transmission would take into account multiple reflections. The effect of multiple reflections is, however, much less than in Prob. 30-10 because of the attenuation in the film.

32-5. (32.1) The surface impedance of a conductor

By definition, the *surface impedance* of a conductor is equal to the ratio of the tangential components of \mathbf{E} and \mathbf{H} at the surface, or to E_t/H_t .

- (a) Show that the surface impedance of a good conductor is given by

$$Z_s = \left(\frac{\omega\mu}{2\sigma} \right)^{1/2} (1+j) = \frac{1+j}{\sigma\delta}.$$

The quantity $1/\sigma\delta$ corresponds to the surface resistance of Prob. 4-9. For copper, $1/\sigma\delta$ is equal to 0.261 miliohm per square at 1 megahertz, from Table 29-1.

- (b) Show that the power dissipated per square meter in the conductor is $H_{t,\text{rms}}^2/\sigma\delta$.

Now we saw in Prob. 19-4 that H_t is equal to the current per unit width in the conductor. It follows that the power dissipated in the conductor is the same as if the current were uniformly distributed throughout the thickness δ .

32-6. (32.1) Cutting steel plate with a laser beam

Figure 32-6 shows a laser beam cutting a steel plate,

- (a) Why does the beam cut at a faster rate when the \mathbf{E} vector lies in the plane of the paper than when it is perpendicular?
- (b) Roughly what percentage of the beam power serves to heat the steel in the former case?
- (c) Can you explain why the kerf is narrower and more even when \mathbf{E} is in the plane of the paper?

If the required kerf is not straight, then the laser should rotate to keep the \mathbf{E} vector of the beam parallel to the path. A simpler solution is to use a circularly polarized beam.

32-7. (32.1) The standing wave at normal incidence on a good conductor

Show that the electromagnetic energy density in a plane standing wave at normal incidence on a good conductor is uniform.

32-8. (32.1) Multiple reflections in a dielectric plate backed by a conductor

An electromagnetic wave falls at an angle of θ_l on a slab of dielectric that is backed by a good conductor.

Under what condition is there a single reflected wave?

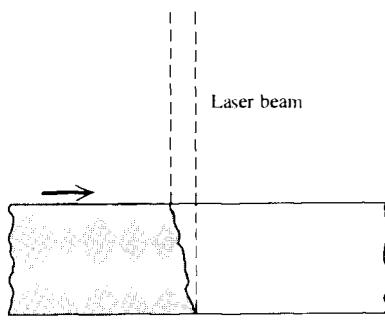


Fig. 32-6.

32-9. (32.2) The radiation force on a sphere

Calculate the radiation force on a reflecting sphere of radius R in terms of the Poynting vector of the incident radiation.

32-10. (32.2) The radiation force on a cylinder

Calculate the radiation force per unit length on a cylinder of radius R whose axis is perpendicular to the Poynting vector of the incoming radiation.

32-11. (32.2) Radiation pressure and comet tails

(a) Compare the gravitational and radiation forces exerted by the sun on a spherical particle of radius a whose density is 5000 kilograms/meter³. The sun radiates 3.8×10^{26} watts. See the Table of Physical Constants at the end of the book. Assume that the particle is black.

(b) Calculate the value of a for which the two forces are equal.

You should find that particles smaller than about 0.1 micrometer in radius are repelled at any distance from the sun. This explains why comet tails that consist of fine particles point away from the sun. Such comets are said to be Type 2. We have disregarded diffraction, which is important when $a \gg \lambda$.

The tails of Type 1 comets are gaseous. They also point away from the sun, but for a different reason. The force then arises from an interaction between this gas and the solar wind (Prob. 28-12), which consists of ionized hydrogen that evaporates from the sun.

32-12. (32.2) Photon-drag radiation monitor

Figure 32-7 shows a schematic diagram of a *photon-drag radiation monitor*. These devices are used to monitor the intensity of powerful laser pulses. The beam enters on the left, through an antireflection coating (Prob. 30-14) and a transparent electrode (Prob. 32-4). The body of the monitor is a single crystal of semiconductor that is quite transparent at the wavelength used. The beam exits on the right where there is, again, an antireflection coating and a transparent electrode, in that order.

Radiation pressure in the semiconductor propels the charge carriers to the right. If the carriers are electrons, the electrodes become charged, as in the figure, and the voltage V is a measure of the beam power.

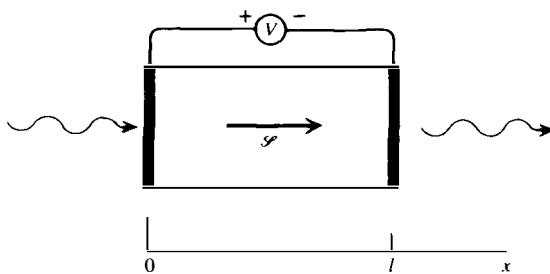


Fig. 32-7.

These monitors have a short response time, of the order of 1 nanosecond. They are made in various sizes, with crystals of the order of 1 centimeter in diameter and a few centimeters long. The crystal absorbs about one-quarter of the pulse energy. The peak power density can be as high as 20 megawatts/centimeter².

Find the ratio V/\mathcal{S}_m . Set $\mathcal{S} = \mathcal{S}_m \exp(-ax)$ inside the crystal.

32-13. (32.2.2) Radiation pressure with \mathbf{E} in the plane of incidence

Show that the radiation pressure on a nonmagnetic good conductor, when \mathbf{E} lies in the plane of incidence, is the same as in Sec. 32.2.1. In this instance there is both a magnetic force within the conductor and an electric force on the surface charges. Use Gauss's law to find σ_r .

32-14. (32.2.3) The levitation of transparent particles in a laser beam

Figure 32-8 shows a simplified diagram of a device for levitating transparent particles in a laser beam. The particles can range from 1 to 100 micrometers in diameter. The light intensity is maximum on the axis of the beam and tapers off on either side. The particle stays on the axis of the beam at a fixed height.

(a) Show qualitatively that, if the particle strays away from the axis, it suffers a restoring force. The axis is therefore a position of equilibrium. You can show this by sketching the paths of two rays that enter the

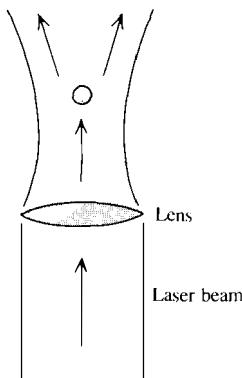


Fig. 32-8.

particle from below, one to the left of the particles' center and one to the right. Refraction deflects the rays and hence changes their momenta.

(b) Show that for a given beam intensity the vertical position of the beam is also stable. A reflecting particle is ejected laterally.

- 32-15. (32.3) The angular momentum of an electrically charged permanent magnet

The field of an electrically charged permanent magnet possesses an angular momentum because \mathbf{E} is radial, while \mathbf{H} points approximately in the $\hat{\theta}$ direction, so that $\mathbf{E} \times \mathbf{H}$ is azimuthal.

We first calculate the value of the momentum from the known values of \mathbf{E} and \mathbf{H} , and we then show that its existence follows from the law of conservation of momentum.

Imagine a conducting sphere of radius R whose magnetization \mathbf{M} is uniform. You may take for granted that outside the sphere the magnetic field is the same as that of a small magnetic dipole of moment $\frac{4}{3}\pi R^3 \mathbf{M}$ situated at the center. The sphere carries a charge Q .

- (a) Find the angular momentum of the field.
- (b) Calculate the value of the angular momentum L for $R = 20$ millimeters and $M = 10^6$ amperes/meter when the sphere is charged to a potential of 1000 volts. Could the sphere be useful as a gyroscope?
- (c) Now let us start with an uncharged sphere and gradually deposit charge on it by means of an axial ion beam. Charge flows in at the north pole and distributes itself uniformly over the surface of the sphere. The magnetic field exerts a torque T_{mag} on the charging current. To prevent the sphere from turning, the support exerts an opposing mechanical torque T_{mech} such that

$$T_{\text{mech}} = -T_{\text{mag}} = \frac{dL}{dt}.$$

Choose polar coordinates with the north pole at $\theta = 0$.

Show that the downward surface current density at θ is

$$\alpha = \frac{1 + \cos \theta}{4\pi R \sin \theta} \frac{dQ}{dt}.$$

- (d) Now show that $T_{\text{mag}} = -dL/dt$, as above.

We have calculated the torque exerted by the magnet on the current. There is no torque exerted by the current on the magnet for the following reason. The equivalent currents on the spherical surface of the magnet are azimuthal, and any force exerted on them has a zero azimuthal component.

CHAPTER 33

GUIDED WAVES I

General Principles. The Coaxial and Microstrip Lines

33.1	GENERAL PROPERTIES OF AN ELECTROMAGNETIC WAVE PROPAGATING IN A STRAIGHT LINE	611
33.1.1	THE TRANSVERSE COMPONENTS ARE FUNCTIONS OF THE LONGITUDINAL COMPONENTS	612
33.1.2	TE AND TM WAVES	614
33.1.3	TEM WAVES	615
33.1.3.1	V IN TEM WAVES	616
33.1.3.2	A IN TEM WAVES	617
33.1.4	BOUNDARY CONDITIONS AT THE SURFACE OF A STRAIGHT METALLIC WAVEGUIDE	618
33.2	THE COAXIAL LINE	619
33.3	THE MICROSTRIP LINE	621
33.4	SUMMARY	622
	PROBLEMS	624

In Chaps. 28 and 29 we studied the propagation of electromagnetic waves in an unbounded region. Then in Chaps. 30 to 32 we investigated the reflection and the refraction of plane waves at the interface between two media.

We now study how electromagnetic waves can be guided in prescribed directions by waveguides, first metallic guides in Chaps. 33 and 34, and then dielectric guides in Chaps. 35 and 36.

In this chapter we first investigate some general properties of waves propagating in a straight line, but without making any assumption as to the way in which the field depends on the transverse coordinates. Then we study the coaxial and microstrip lines.

There exist many types of waveguides. The most common is the coaxial line. It serves to interconnect electronic instruments and, in older systems, is used for long-distance telephony. Hollow metallic guides can operate at high power. They serve mostly for connecting antennas to

transmitters and receivers, in radar sets, for example. Metallic guides are lossy. Dielectric waveguides are nearly lossless and serve to transmit information over long distances.

As we shall see, the coaxial and microstrip lines can transmit waves of any frequency, from zero to about 10^{10} hertz, while hollow metallic guides are narrow-band devices that operate at frequencies of the order of 10^9 hertz and higher.

33.1 GENERAL PROPERTIES OF AN ELECTROMAGNETIC WAVE PROPAGATING IN A STRAIGHT LINE

To simplify, we assume the six following conditions.

- (1) The medium of propagation is homogeneous, isotropic, linear, and stationary (HILS).
- (2) It is nonconducting. This does *not* exclude metallic guides, because the wave propagates *along* a metallic guide.
- (3) The free charge density is zero. This makes $\nabla \cdot \mathbf{E} = 0$.
- (4) Propagation occurs in a straight line, in the positive direction of the z -axis. There is no reflected wave traveling in the $-z$ direction.
- (5) The wave is sinusoidal.
- (6) There is zero attenuation. If the guide is metallic, then its conductivity must be infinite to avoid Joule losses. We shall see in Sec. 34.8 how to calculate attenuation with real conductors.

We may therefore write that

$$\mathbf{E} = \mathbf{E}_m \exp j(\omega t - k_z z) = (E_{mx} \hat{x} + E_{my} \hat{y} + E_{mz} \hat{z}) \exp j(\omega t - k_z z), \quad (33-1)$$

$$\mathbf{H} = \mathbf{H}_m \exp j(\omega t - k_z z) = (H_{mx} \hat{x} + H_{my} \hat{y} + H_{mz} \hat{z}) \exp j(\omega t - k_z z), \quad (33-2)$$

where the coefficients E_{mx} , E_{my} , E_{mz} , H_{mx} , . . . are unspecified functions of x and y . The dependence on z and t appears only in the exponential function. The wave number k_z for the guided wave is real, since there is zero attenuation. It is equal to $2\pi/\lambda_z$, where λ_z is the wavelength of the guided wave.

Let us substitute the above expressions for \mathbf{E} and \mathbf{H} into Maxwell's equations. Since $\nabla \cdot \mathbf{E} = 0$,

$$\frac{\partial E_{mx}}{\partial x} + \frac{\partial E_{my}}{\partial y} - jk_z E_{mz} = 0. \quad (33-3)$$

Similarly, $\nabla \cdot \mathbf{B} = 0$, and

$$\frac{\partial H_{mx}}{\partial x} + \frac{\partial H_{my}}{\partial y} - jk_z H_{mz} = 0. \quad (33-4)$$

From the fact that $\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t$,

$$\frac{\partial E_{mz}}{\partial y} + jk_z E_{my} = -j\omega \mu H_{mx}, \quad (33-5)$$

$$-jk_z E_{mx} - \frac{\partial E_{mz}}{\partial x} = -j\omega \mu H_{my}, \quad (33-6)$$

$$\frac{\partial E_{my}}{\partial x} - \frac{\partial E_{mx}}{\partial y} = -j\omega \mu H_{mz}. \quad (33-7)$$

From $\nabla \times \mathbf{H} = \partial \mathbf{D} / \partial t$,

$$\frac{\partial H_{mz}}{\partial y} + jk_z H_{my} = j\omega \epsilon E_{mx}, \quad (33-8)$$

$$-jk_z H_{mx} - \frac{\partial H_{mz}}{\partial x} = j\omega \epsilon E_{my}, \quad (33-9)$$

$$\frac{\partial H_{my}}{\partial x} - \frac{\partial H_{mx}}{\partial y} = j\omega \epsilon E_{mz}. \quad (33-10)$$

33.1.1 The Transverse Components are Functions of the Longitudinal Components

We can now show that the four transverse components E_{mx} , E_{my} , H_{mx} , H_{my} are functions of the longitudinal components E_{mz} , H_{mz} . From Eqs. 33-6 and 33-8,

$$E_{mx} = \frac{j}{k_z^2 - k^2} \left(k_z \frac{\partial E_{mz}}{\partial x} + \omega \mu \frac{\partial H_{mz}}{\partial y} \right). \quad (33-11)$$

Here

$$k = \omega(\epsilon \mu)^{1/2} = \frac{1}{\lambda} \quad (33-12)$$

is the wave number of a uniform plane wave of wavelength λ traveling in the medium.

We have assumed that $k_z \neq k$ for the moment. Both k and k_z are real and positive. Similarly,

$$E_{my} = \frac{j}{k_z^2 - k^2} \left(k_z \frac{\partial E_{mz}}{\partial y} - \omega \mu \frac{\partial H_{mz}}{\partial x} \right), \quad (33-13)$$

$$H_{mx} = \frac{j}{k_z^2 - k^2} \left(-\omega \epsilon \frac{\partial E_{mz}}{\partial y} + k_z \frac{\partial H_{mz}}{\partial x} \right), \quad (33-14)$$

$$H_{my} = \frac{j}{k_z^2 - k^2} \left(\omega \epsilon \frac{\partial E_{mz}}{\partial x} + k_z \frac{\partial H_{mz}}{\partial y} \right). \quad (33-15)$$

We use the subscript \perp to identify components that are perpendicular to the direction of propagation. Thus

$$\mathbf{E}_{m\perp} = E_{mx} \hat{\mathbf{x}} + E_{my} \hat{\mathbf{y}}, \quad \mathbf{H}_{m\perp} = H_{mx} \hat{\mathbf{x}} + H_{my} \hat{\mathbf{y}}. \quad (33-16)$$

More succinctly,

$$\mathbf{E}_{m\perp} = \frac{j}{k_z^2 - k^2} (k_z \nabla_{\perp} E_{mz} + \omega \mu \nabla \times H_{mz} \hat{\mathbf{z}}), \quad (33-17)$$

$$\mathbf{H}_{m\perp} = \frac{j}{k_z^2 - k^2} (k_z \nabla_{\perp} H_{mz} - \omega \epsilon \nabla \times E_{mz} \hat{\mathbf{z}}). \quad (33-18)$$

So we need to solve the wave equation and apply the boundary conditions only for the two longitudinal components. Once that is done, the other four components will follow immediately.

The longitudinal component of \mathbf{E} satisfies the wave equation 27-70 with $\rho_f = 0$, $\mathbf{J}_f = 0$. So

$$\frac{\partial^2 E_{mz}}{\partial x^2} + \frac{\partial^2 E_{mz}}{\partial y^2} - k_z^2 E_{mz} = -\epsilon \mu \omega^2 E_{mz} = -k^2 E_{mz}, \quad (33-19)$$

or

$$\frac{\partial^2 E_{mz}}{\partial x^2} + \frac{\partial^2 E_{mz}}{\partial y^2} + (k^2 - k_z^2) E_{mz} = 0, \quad (33-20)$$

$$(\nabla_{\perp}^2 + k^2 - k_z^2) E_{mz} = 0. \quad (33-21)$$

Similarly,

$$\frac{\partial^2 H_{mz}}{\partial x^2} + \frac{\partial^2 H_{mz}}{\partial y^2} + (k^2 - k_z^2) H_{mz} = 0, \quad (33-22)$$

$$(\nabla_{\perp}^2 + k^2 - k_z^2) H_{mz} = 0. \quad (33-23)$$

Applying the proper boundary conditions (Sec. 33.1.4) yields the value of k_z .

33.1.2 TE and TM Waves

It is convenient to consider separately three types of wave: (1) transverse electric (TE) waves, in which $E_{mz} = 0$; (2) transverse magnetic (TM) waves, with $H_{mz} = 0$; (3) transverse electric and magnetic (TEM) waves, with $E_{mz} = 0$, $H_{mz} = 0$.

With either TE or TM waves, it follows from Eqs. 33-11 to 33-15 that

$$\frac{E_{mx}}{H_{my}} = -\frac{E_{my}}{H_{mx}}. \quad (33-24)$$

If k_z is real and positive, as we have assumed, these ratios are also real and positive. Then the components

$$E_x = E_{mx} \exp j(\omega t - k_z z) \quad \text{and} \quad H_y = H_{my} \exp j(\omega t - k_z z) \quad (33-25)$$

are in phase, and so are E_y and $-H_x$. This fact, together with Eq. 33-24, implies that

$$\frac{\operatorname{Re} E_x}{\operatorname{Re} H_y} = \frac{\operatorname{Re} E_y}{\operatorname{Re} (-H_x)}, \quad (33-26)$$

and that the real parts of E_\perp and H_\perp are mutually orthogonal in both TE and TM waves.

The ratio $E_{m\perp}/H_{m\perp}$ is the *wave impedance*. This is a real positive quantity if there is no dissipation:

$$Z_{\text{TE}} = \frac{E_{m\perp}}{H_{m\perp}} = \frac{\omega\mu}{k_z} = \left(\frac{\mu}{\epsilon}\right)^{1/2} \frac{\lambda_z}{\lambda_0} \quad (33-27)$$

$$= 3.76731 \times 10^2 \frac{\lambda_z}{\lambda_0} \approx 377 \frac{\lambda_z}{\lambda_0} \quad \text{ohms} \quad (\epsilon_r = 1, \mu_r = 1), \quad (33-28)$$

$$Z_{\text{TM}} = \frac{E_{m\perp}}{H_{m\perp}} = \frac{k_z}{\omega\epsilon} = \left(\frac{\mu}{\epsilon}\right)^{1/2} \frac{\lambda_0}{\lambda_z} \quad (33-29)$$

$$\approx 377 \frac{\lambda_0}{\lambda_z} \quad \text{ohms} \quad (\epsilon_r = 1, \mu_r = 1). \quad (33-30)$$

Here λ and λ_0 are the wavelengths v/f of a plane wave of the same frequency f , and λ_z is the wavelength of the guided wave.

33.1.3 TEM Waves

If $k_z = k$ in Eqs. 33-11 to 33-15, the items in parentheses must be zero. The simplest way of satisfying this condition is to set both E_{mz} and H_{mz} equal to zero. We then have a TEM wave.

With TEM waves the wavelength λ_z of the guided wave is the same as that of a uniform plane wave in the same medium of propagation because k_z is equal to k , so

$$\lambda_z = \lambda. \quad (33-31)$$

If the medium is air, then the phase velocity is c , whatever the geometry of the guide and whatever the frequency. Such a guide is *distortionless* because the various frequency components of a complex waveform all travel at the same velocity.[†]

Setting $E_{mz} = 0$, $H_{mz} = 0$ in Eqs. 33-6 and 33-9 gives

$$E_{mx} = \left(\frac{\mu}{\epsilon}\right)^{1/2} H_{my}, \quad E_{my} = -\left(\frac{\mu}{\epsilon}\right)^{1/2} H_{mx}. \quad (33-32)$$

The wave impedance is now

$$\frac{E_m}{H_m} = \frac{(E_{mx}^2 + E_{my}^2)^{1/2}}{(H_{mx}^2 + H_{my}^2)^{1/2}} = \left(\frac{\mu}{\epsilon}\right)^{1/2} \quad (33-33)$$

$$\approx 377 \text{ ohms} \quad (\epsilon_r = 1, \mu_r = 1). \quad (33-34)$$

The ratio $(\mu/\epsilon)^{1/2}$ is the *characteristic impedance of the medium* (Sec. 28.5.2).

The electric and magnetic energy densities are equal:

$$\frac{\epsilon E^2}{2} = \frac{\mu H^2}{2}. \quad (33-35)$$

Also, the average Poynting vector is

$$\mathcal{S}_{av} = \frac{1}{2} \operatorname{Re}(\mathbf{E} \times \mathbf{H}^*) = \frac{1}{2} \left(\frac{\epsilon}{\mu}\right)^{1/2} E_m^2 \hat{\mathbf{z}} \quad (33-36)$$

$$= \left(\frac{\epsilon}{\mu}\right)^{1/2} E_{rms}^2 \hat{\mathbf{z}} = 2.65441 \times 10^{-3} \left(\frac{\epsilon_r}{\mu_r}\right)^{1/2} E_{rms}^2 \hat{\mathbf{z}} \quad \text{watts/meter}^2 \quad (33-37)$$

$$= v\epsilon E_{rms}^2 \hat{\mathbf{z}} = v\mu H_{rms}^2 \hat{\mathbf{z}}, \quad (33-38)$$

[†]This is only approximately true. The conductivity of metallic waveguides being finite, there is attenuation and dispersion.

where

$$v = \frac{1}{(\epsilon\mu)^{1/2}} = \frac{c}{(\epsilon_r\mu_r)^{1/2}} \quad (33-39)$$

is the *phase velocity*.

The magnitude of the time-averaged Poynting vector is equal to the energy density multiplied by the phase velocity.

33.1.3.1 V in TEM waves

The uniform plane waves that we studied in Chap. 28 are TEM waves. We now study TEM waves that follow a conducting guide.

In a TEM guided wave the electric field in a plane perpendicular to the direction of propagation is derivable from a potential, in the same way as an electrostatic field. We can show this as follows.

Inside a conductor of infinite conductivity, $\mathbf{E} = 0$, for otherwise \mathbf{J} would be infinite. But $\nabla \times \mathbf{E} = -j\omega\mathbf{B}$, and thus \mathbf{B} and \mathbf{H} are also zero inside the conductor. Now let C be an arbitrary closed curve situated in a plane perpendicular to the z -axis. Since \mathbf{H} is transverse in the medium of propagation and zero inside the conductor, there is zero longitudinal \mathbf{H} and the magnetic flux linking C is zero. Thus

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = \exp j(\omega t - k_z z) \oint_C \mathbf{E}_m \cdot d\mathbf{l} = 0, \quad (33-40)$$

$$\oint_C \mathbf{E}_m \cdot d\mathbf{l} = 0, \quad (33-41)$$

and \mathbf{E} is derivable from a potential:

$$\mathbf{E}_m = -\nabla V_m = -\frac{\partial V_m}{\partial x} \hat{x} - \frac{\partial V_m}{\partial y} \hat{y}, \quad (33-42)$$

$$\mathbf{E} = -\nabla V = -\frac{\partial V}{\partial x} \hat{x} - \frac{\partial V}{\partial y} \hat{y}, \quad (33-43)$$

with

$$V = V_m \exp j(\omega t - k_z z). \quad (33-44)$$

If the waveguide is a hollow conducting tube, as in Fig. 33-1, the tangential component of \mathbf{E} at its surface is zero, V_m is a constant all around the tube, and the only possible solution inside is $V_m = \text{constant}$. Now, if V_m is constant throughout the inside of the guide, E_m is zero, $\mathbf{E} = 0$, and since $\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t$, there is no \mathbf{H} wave either. Therefore TEM waves cannot travel inside a hollow conducting tube.

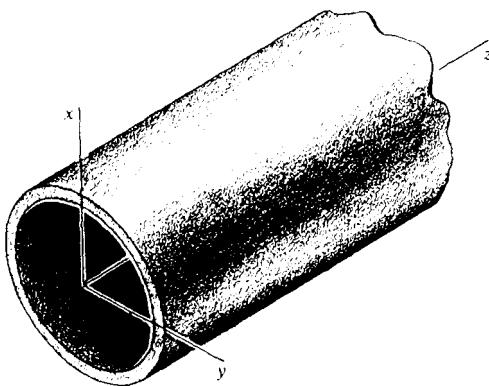


Fig. 33-1. Hollow conducting waveguide.

This is not rigorously true because TEM waves are allowed if the wavelength is much less than the cross-sectional dimensions. For example, light goes through a straight length of metal pipe. We shall see in Sec. 34.4 that the TEM wave is then a limiting case of a TE wave.

In the shielded-pair and parallel-wire lines of Fig. 33-2 and in the coaxial line of Fig. 33-4 the conductors need not all be at the same potential, V_m is a function of x and y as when the field is static, and \mathbf{E} need not be zero, so TEM waves are possible.

33.1.3.2 \mathbf{A} in TEM waves

Applying the general expression for \mathbf{E} that we found in Secs. 17.7 and 23-5 gives

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t} = -\frac{\partial V}{\partial x} \hat{x} - \frac{\partial V}{\partial y} \hat{y} - \left(\frac{\partial V}{\partial z} \hat{z} + \frac{\partial \mathbf{A}}{\partial t} \right). \quad (33-45)$$

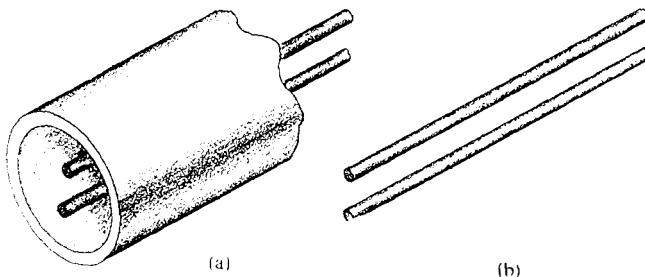


Fig. 33-2. (a) Shielded-pair line. The outer cylinder is grounded, and there is a potential difference between the two wires. (b) Parallel-wire line.

where the expression between parentheses is equal to zero, from Eq. 33-43. Since \mathbf{A} is of the same form as V (Eq. 33-44),

$$-jk_z V \hat{\mathbf{z}} + j\omega \mathbf{A} = 0 \quad \text{or} \quad \mathbf{A} = \frac{k_z}{\omega} V \hat{\mathbf{z}}. \quad (33-46)$$

The vector potential \mathbf{A} is therefore proportional to the scalar potential V in a TEM wave. Also, \mathbf{A} is longitudinal, which means that the currents in a conducting guide for TEM waves are longitudinal.

33.1.4 Boundary Conditions at the Surface of a Straight Metallic Waveguide

By hypothesis, the guiding structure comprises straight conductors of infinite conductivity, parallel to the z -axis. The cross section of the guide is therefore uniform.

(1) With any type of electromagnetic wave, \mathbf{E} vanishes inside a perfect conductor. Then, because of the continuity of the tangential component of \mathbf{E} at an interface (Sec. 10.2.3), that component of \mathbf{E} is zero close to the guide. For guides of finite conductivity, see Sec. 34.8.

(2) Again for any type of wave, $\nabla \times \mathbf{E} = 0$ inside the conductor, $j\omega \mathbf{B} = 0$ and $\mathbf{B} = 0$, $\mathbf{H} = 0$. Then, because of the continuity of the normal component of \mathbf{B} at an interface (Sec. 20.8), the normal component of \mathbf{B} , close to a perfectly conducting guide, is zero.

Once the tangential \mathbf{H} is known, the surface current density α follows from the relation $\mathbf{H} = \alpha \times \hat{\mathbf{n}}$ (Prob. 19-4), where $\hat{\mathbf{n}}$ is the unit normal vector pointing away from the guide. The surface current density can be different from zero with $\mathbf{E} = 0$ and $\sigma \rightarrow \infty$.

In other words, close to the guide, \mathbf{H} is tangent to the surface, orthogonal to α , and equal in magnitude to α . For example, with TM and TEM waves, \mathbf{H} is everywhere transverse and the currents in the guide are longitudinal.

(3) Finally, for TE waves, even with imperfectly conducting guides,

$$\mathbf{E}_{m\perp} = E_{mx} \hat{\mathbf{x}} + E_{my} \hat{\mathbf{y}} = \frac{j\omega\mu}{k_z^2 - k^2} \left(\frac{\partial H_{mz}}{\partial y} \hat{\mathbf{x}} - \frac{\partial H_{mz}}{\partial x} \hat{\mathbf{y}} \right) \quad (33-47)$$

$$= \frac{j\omega\mu}{k_z^2 - k^2} (\nabla H_{mz}) \times \hat{\mathbf{z}}. \quad (33-48)$$

With a perfectly conducting guide, $\mathbf{E}_{m\perp}$ is normal to the surface and ∇H_{mz} is tangent, as in Fig. 33-3. Thus, the rate of change of H_{mz} in the direction normal to the surface, at the surface, is zero.

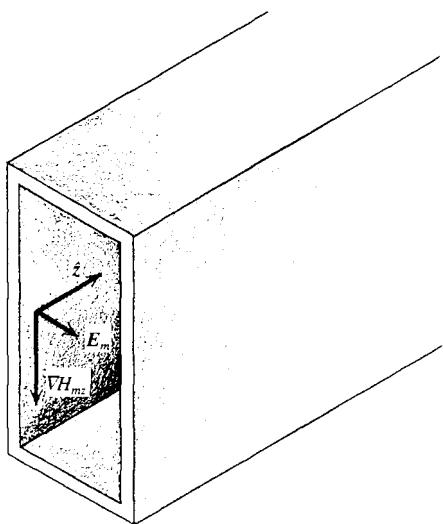


Fig. 33-3. Portion of a rectangular waveguide. With TE waves ∇H_{mz} is tangent to the wall.

33.2 THE COAXIAL LINE

In the coaxial line illustrated in Fig. 33-4, the wave propagates in the annular space between the conductors, and there is zero field outside. The medium of propagation is usually a low-loss dielectric.

This type of guide normally carries TEM waves. Various TE and TM modes are also allowed, but only at wavelengths that are of the same

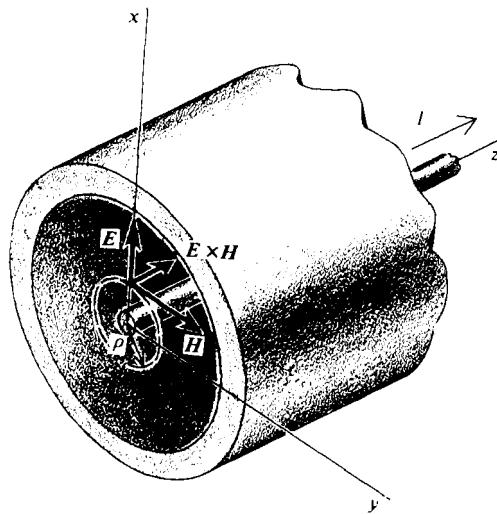


Fig. 33-4. The E , H , and $E \times H$ vectors inside a coaxial line.

order of magnitude as, or smaller than, the diameter of the line ($\lambda \approx 10^{-2}$ meter, $f \approx 10^{10}$ hertz). Then the various modes travel at different velocities, and the line distorts complex waveforms. In practice, one operates at frequencies well below the threshold for the TE and TM modes, and coaxial lines are then nearly distortionless.

Thus

$$E_{mz} = 0, \quad H_{mz} = 0, \quad (33-49)$$

and, from Eq. 33-42,

$$\mathbf{E} = -\left(\frac{\partial V_m}{\partial x}\hat{x} + \frac{\partial V_m}{\partial y}\hat{y}\right) \exp j(\omega t - k_z z). \quad (33-50)$$

For a given t and a given z , \mathbf{E} varies with x and y exactly as when the field is static. It is radial and varies as $1/\rho$ (Prob. 3-8):

$$\mathbf{E} = \frac{C}{\rho} \exp j(\omega t - k_z z) \hat{\rho}, \quad (33-51)$$

where C is a constant and

$$k_z = \frac{1}{\lambda_z} = \omega(\epsilon\mu_0)^{1/2} = \frac{1}{\lambda}, \quad (33-52)$$

as for a uniform plane wave in the medium of propagation. We have set $\mu_r = 1$ since magnetic materials are lossy.

The velocity of propagation is equal to ω/k_z .

The potential of the inner conductor with respect to the outer conductor is

$$\mathcal{V} = \int_{\rho_1}^{\rho_2} E d\rho = C \ln \frac{\rho_2}{\rho_1} \exp j(\omega t - k_z z), \quad (33-53)$$

where ρ_1 and ρ_2 are the inner and outer radii of the annular region between the conductors. This is the *line voltage*.

From the previous section, \mathbf{H} is orthogonal to \mathbf{E} and

$$\mathbf{H} = \left(\frac{\epsilon}{\mu_0}\right)^{1/2} \frac{C}{\rho} \exp j(\omega t - k_z z) \hat{\phi} \quad (\mu_r = 1). \quad (33-54)$$

The vectors \mathbf{E} and \mathbf{H} are as in Fig. 33-4.

The line current flowing along the surface of the inner conductor is related to \mathbf{H} through the circuital law of Sec. 20.6:

$$I = I_m \exp j(\omega t - k_z z) = 2\pi\rho_1 H_{\rho=\rho_1}$$

$$I = 2\pi\rho_1 \left(\frac{\epsilon}{\mu_0} \right)^{1/2} \frac{C}{\rho_1} \exp j(\omega t - k_z z) \quad (33-55)$$

$$\approx \frac{C\epsilon_r^{1/2}}{60} \exp j(\omega t - k_z z) \quad (\mu_r = 1). \quad (33-56)$$

An equal current flows in the opposite direction along the inner surface of the outer conductor. If the medium of propagation is air, then $C = 60I_m$.

The Poynting vector $\mathbf{E} \times \mathbf{H}$ points in the direction of propagation, as in Fig. 33-4, and the average transmitted power is

$$P_T = \int_{\rho_1}^{\rho_2} \mathcal{S}_{av} 2\pi\rho d\rho, \quad (33-57)$$

where

$$\mathcal{S}_{av} = \frac{1}{2} \operatorname{Re} (\mathbf{E} \times \mathbf{H}^*) = \left(\frac{\epsilon}{\mu_0} \right)^{1/2} \frac{C^2}{2\rho^2} \hat{z}. \quad (33-58)$$

Thus

$$P_T \approx \epsilon_r^{1/2} \frac{C^2}{120} \ln \frac{\rho_2}{\rho_1} \quad \text{watts} \quad (\mu_r = 1). \quad (33-59)$$

Also,

$$P_T = \frac{1}{2} \operatorname{Re} (\mathcal{V} I^*) = \frac{1}{2} \mathcal{V}_m I_m = \mathcal{V}_{rms} I_{rms}. \quad (33-60)$$

The *characteristic impedance* of a coaxial line is the ratio \mathcal{V}/I when there is no reflected wave traveling in the $-z$ direction:

$$Z = \frac{\mathcal{V}}{I} \approx \frac{60}{\epsilon_r^{1/2}} \ln \frac{\rho_2}{\rho_1} \quad \text{ohms} \quad (\mu_r = 1). \quad (33-61)$$

33.3 THE MICROSTRIP LINE

Figure 33-5 shows a cross section of a *microstrip line*. It comprises a grounded conducting plane, an insulating sheet, and a conducting strip. The strip is gold or copper, plated onto the insulator, which is either alumina or fused quartz.

Microstrip lines carry TEM waves.[†] They are particularly useful in printed and integrated circuits that operate at frequencies of about one to

[†] This is an approximation; \mathbf{E} and \mathbf{H} also have longitudinal components.

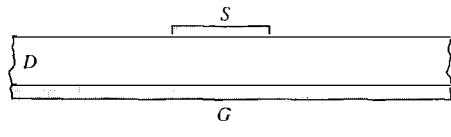


Fig. 33-5. Cross-section of a microstrip line:
G, ground-plane layer;
D, dielectric substrate;
S, conducting strip.

tens of gigahertz. They have the advantage of being much less costly than either the coaxial line or the rectangular waveguide. Their main disadvantage lies in the fact that their field is not strictly limited to the region immediately below the upper electrode. Microstrip lines are therefore lossy and can interact with other elements in a circuit, unless they are either spaced or shielded properly. Losses are high, of the order of 0.5 decibel/millimeter.

33.4 SUMMARY

In any electromagnetic wave propagating in a straight line, the transverse components of \mathbf{E} and \mathbf{H} are simple functions of the longitudinal components. Using the subscript \perp for components that are perpendicular to the direction of propagation,

$$\mathbf{E}_{m\perp} = \frac{j}{k_z^2 - k^2} (k_z \nabla_{\perp} E_{mz} + \omega \mu \nabla \times H_{mz} \hat{\mathbf{z}}), \quad (33-17)$$

$$\mathbf{H}_{m\perp} = \frac{j}{k_z^2 - k^2} (k_z \nabla_{\perp} H_{mz} - \omega \epsilon \nabla \times E_{mz} \hat{\mathbf{z}}). \quad (33-18)$$

The longitudinal components themselves satisfy the wave equations:

$$(\nabla_{\perp}^2 + k^2 - k_z^2) E_{mz} = 0, \quad (33-21)$$

$$(\nabla_{\perp}^2 + k^2 - k_z^2) H_{mz} = 0. \quad (33-23)$$

There exist three types of wave: TE ($E_{mz} = 0$), TM ($H_{mz} = 0$), and TEM ($E_{mz} = 0, H_{mz} = 0$) waves.

In *TE* and *TM* waves, the real parts of \mathbf{E}_{\perp} and of \mathbf{H}_{\perp} are mutually orthogonal.

The *wave impedances* are

$$Z_{TE} = \frac{E_{m\perp}}{H_{m\perp}} = \left(\frac{\mu_0}{\epsilon} \right)^{1/2} \frac{\lambda_z}{\lambda} \approx \frac{377 \lambda_z}{\epsilon_r^{1/2} \lambda}, \quad (\mu_r = 1) \quad (33-27)$$

$$Z_{TM} = \left(\frac{\mu_0}{\epsilon} \right)^{1/2} \frac{\lambda}{\lambda_z} \approx \frac{377 \lambda}{\epsilon_r^{1/2} \lambda_z}. \quad (\mu_r = 1) \quad (33-29)$$

Here λ_z is the wavelength of the guided wave, and λ is the wavelength of a plane wave of the same frequency in the medium of propagation.

For *TEM waves* the wavelength λ_z of the guided wave is the same as that of a uniform plane wave in the same medium of propagation,

$$\lambda_z = \lambda, \quad (33-31)$$

and the wave impedance is

$$\frac{E_m}{H_m} = \left(\frac{\mu_0}{\epsilon} \right)^{1/2} \approx \frac{377}{\epsilon_r^{1/2}} \quad \text{ohms} \quad (\mu_r = 1). \quad (33-33)$$

The electric and magnetic energies are equal, and the magnitude of the Poynting vector is equal to the total energy density multiplied by the phase velocity.

In a TEM wave, \mathbf{E} is derivable from a potential: $\mathbf{E} = -\nabla V$. Also, the currents flowing in a conducting guide are longitudinal.

The *boundary conditions* at the surface of a perfectly conducting guide are as follows: $E_{\text{tangential}} = 0$, $B_{\text{normal}} = 0$, $H_{\text{tangential}} = \alpha \times \hat{n}$, ∇H_{mz} is tangential.

Figure 33-3 shows a portion of a *coaxial line*. In the field,

$$\mathbf{E} = \frac{C}{\rho} \exp j(\omega t - k_z z) \hat{\rho}, \quad (33-51)$$

$$\mathbf{H} = \left(\frac{\epsilon}{\mu_0} \right)^{1/2} \frac{C}{\rho} \exp j(\omega t - k_z z) \hat{\phi}, \quad (33-54)$$

where C is a constant that fixes the amplitude of the field.

The voltage, current, and power flow are

$$\mathcal{V} = C \ln \frac{\rho_2}{\rho_1} \exp j(\omega t - k_z z), \quad (33-53)$$

$$I = 2\pi C \left(\frac{\epsilon}{\mu_0} \right)^{1/2} \exp j(\omega t - k_z z), \quad (33-55)$$

$$P_T = \mathcal{V}_{\text{rms}} I_{\text{rms}}. \quad (33-60)$$

The characteristic impedance of a coaxial line is

$$Z \approx \frac{60}{\epsilon_r^{1/2}} \ln \frac{\rho_2}{\rho_1} \quad \text{ohms}. \quad (33-61)$$

The velocity of propagation of a TEM wave in a coaxial line is the same as that of a uniform plane wave in the medium of propagation.

PROBLEMS

33-1. (33.2) The field inside a coaxial line

- (a) Sketch a rather large cross-sectional view of a coaxial line in a plane containing the axis. Show lines of \mathbf{E} and of \mathbf{H} at a given instant over at least one wavelength. The lines should be most closely spaced where the field is strongest. Indicate the directions of the fields by means of arrow heads. The direction of propagation should point to the right.
- (b) Add arrows at various points to represent Poynting vectors, using longer arrows where the power flow is larger. Assume that the length of the arrow represents the magnitude of the Poynting vector at its midpoint.
- (c) Sketch a cross-sectional view of the coaxial line in a plane perpendicular to the axis, and show lines of \mathbf{E} and of \mathbf{H} at a particular instant. Relate this plane to the figure you drew in (a).
- (d) Add plus and minus signs to both figures to show the surface charges. The spacing between the signs should indicate qualitatively the relative magnitude of the surface charge density.
- (e) Now add arrows of various lengths to your first figure to represent surface current densities.
- (f) How do the current patterns change with time?

33-2. (33.2) The current and the charge density in a coaxial line

Show that, in a coaxial line of infinite conductivity, the current is equal to the linear charge density multiplied by the speed of propagation.

33-3. (33.2) The characteristic impedance of a coaxial line

It is known from transmission-line theory that the characteristic impedance of a line is given by $Z_c = (L'/C')^{1/2}$, where L' and C' are, respectively, the inductance and capacitance per meter. Show that this applies to the coaxial line.

33-4. (33.2) Eliminating reflection at the end of a coaxial line

An air-insulated coaxial line is terminated by a sheet whose surface resistance is 377 ohms per square (Prob. 4-9).

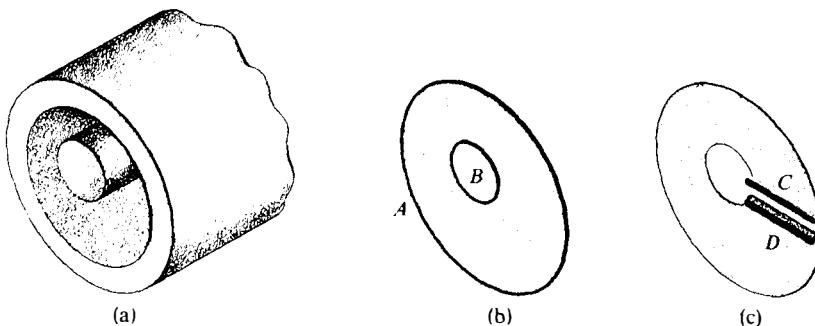
Show that the resistance of the termination is equal to the characteristic impedance of the line. There is then no reflection at the end of the line.

33-5. (33.1.3) Predicting the characteristic impedance of a TEM guide

One important parameter of a TEM waveguide is its characteristic impedance $Z_c = (L'/C')^{1/2}$, where L' and C' are, respectively, the inductance and the capacitance per meter of guide (Prob. 33-3).

In designing such lines it is important to predict the values of L' and of C' . If the geometry is such that these quantities are difficult to calculate, as in Fig. 33-6(a), one can perform the following measurements on a resistance-sheet analog.

(a) One can find the value of C' by cutting out a sheet of resistive material in the shape of the cross section of the dielectric as in Fig. 33-6(b) and by measuring the resistance R_1 between electrodes A and B . See Prob. 9-10.

**Fig. 33-6.**

Show that, if the material has a conductivity σ and a thickness s , and if the permittivity of the dielectric is ϵ , then $R_1 C' = \epsilon / s\sigma$.

(b) One can measure L' as in Fig. 33-6(c) by measuring the resistance R_2 between electrodes C and D .

Show that $R_2 L' = \mu_0 / s\sigma$. Thus $Z_c = (\mu_0 / \epsilon)^{1/2} (R_1 / R_2)^{1/2}$.

33-6. (33.3) The microstrip line

Figure 33-5 shows a cross section of a *microstrip* line.

(a) Sketch lines of E and of H . Use arrows to show the directions of E and H at a given time. Show the direction of propagation.

(b) In practice, the width b of the strip is much larger than its distance h to the grounded plane, and edge effects are small. Show that the instantaneous value $I\mathcal{V}$ of the transmitted power is equal to the Poynting vector integrated over the cross section bh . Assume that there is no reflected wave.

(c) Show that the characteristic impedance \mathcal{V}/I is equal to $(\mu/\epsilon)^{1/2} h/b$.

(d) Show that one arrives at the same result if one defines the characteristic impedance as in Prob. 33-3.

(e) Show that the addition of a second grounded plane placed symmetrically with the first reduces the characteristic impedance by a factor of 2.

33-7. (33.3) The microstrip line.

Calculate the transmitted power when the voltage across the microstrip line of Fig. 33-6 is \mathcal{V} . Disregard edge effects, and assume that there is no reflected wave.

CHAPTER 34

GUIDED WAVES II *The Hollow Rectangular Waveguide*

34.1 THE FIELD COMPONENTS OF A TE WAVE IN A RECTANGULAR METALLIC WAVEGUIDE	628
34.2 THE CUTOFF WAVELENGTH, NONPROPAGATING FIELDS	630
34.3 THE TE ₁ MODE	631
34.4 MULTIPLE REFLECTIONS	633
34.5 THE PHASE, SIGNAL, AND GROUP VELOCITIES	635
34.6 THE TRANSMITTED POWER	636
34.7 THE STORED ENERGY DENSITY	637
34.8 ATTENUATION	637
34.9 SUMMARY	641
PROBLEMS	642

A hollow metallic waveguide is simply a metallic pipe inside which an electromagnetic wave can propagate by reflection on the inner surfaces, in much the same way as a sound wave.

There exist many types of hollow metallic waveguides, but we concentrate on guides of rectangular cross section, as in Fig. 34-1. This is the most common type, and it is also the simplest one. The figure shows

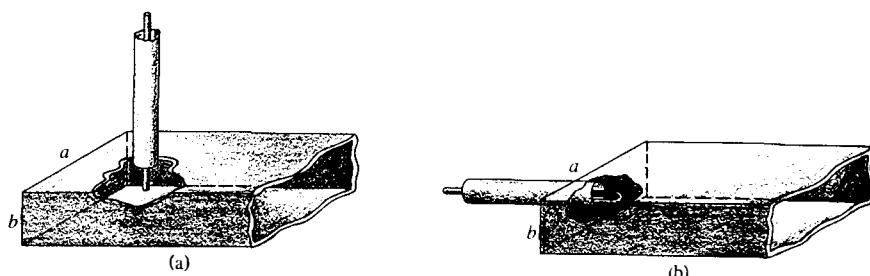


Fig. 34-1. Rectangular metallic waveguides fed by a coaxial line. (a) A quarter-wave antenna injects an electric field at a distance of $\lambda_z/4$ from the closed end of the guide. (b) A small loop antenna injects a magnetic field.

Table 34-1 Characteristics of a few standard rectangular waveguides (TE₁ mode)

INSIDE DIMENSIONS $a \times b$	CUTOFF WAVELENGTH	OPERATING WAVELENGTH RANGE	ATTENUATION	MEAN POWER RATING
millimeters			dB/meter	megawatts
72.1 × 34.0	144	75.9–115	0.0247–0.0362	2.7
47.5 × 22.1	95.1	51.2–75.9	0.0472–0.0682	1.9
34.8 × 15.8	69.7	36.6–51.2	0.0755–0.0942	0.64
22.9 × 10.2	45.7	24.2–36.6	0.147–0.212	0.25
15.8 × 7.9	31.6	16.7–24.2	0.273–0.312	

how such a guide can be connected to a coaxial line. Table 34-1 lists some common standard sizes, but there are many more, down to a width of a few millimeters.

We assume that the medium of propagation is air, as is usually the case. We also assume, for the moment, that the guide is perfectly conducting.

We consider only the Transverse Electric (TE) mode ($E_{mz} = 0$) that results from the multiple reflection of a plane wave on the faces parallel to the yz -plane in Fig. 34-2.

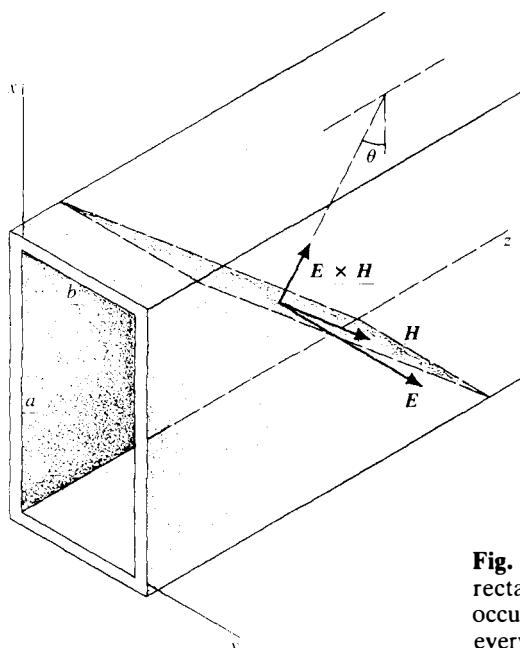


Fig. 34-2. Typical wave front inside a rectangular waveguide. Reflection occurs on the narrow faces, and E is everywhere transverse.

The Transverse Magnetic (TM) waves in hollow rectangular waveguides are not useful because they require larger guides. As we saw in Sec. 33.1.3.1, TEM modes are forbidden.

34.1 THE FIELD COMPONENTS OF A TE WAVE IN A RECTANGULAR METALLIC WAVEGUIDE

We use the coordinate system of Fig. 34-2. The wave propagates in the positive direction of the z -axis by multiple reflection on the upper and lower walls. The figure also shows a wave front of a plane wave incident on the top face at the angle θ .

With *this* mode of propagation,

$$E_{mz} = 0, \quad E_{mx} = 0, \quad H_{my} = 0, \quad \frac{\partial}{\partial y} = 0. \quad (34-1)$$

We require the three other components E_{my} , H_{mx} , H_{mz} .

We proceed as indicated at the end of Sec. 33.1.1. First we solve the wave equation for H_{mz} for the given boundary conditions. This will give us both H_{mz} and k_z . Then the values of E_{my} and H_{mx} will follow, from Eqs. 33-13 and 33-14.

From Eq. 33-22,

$$\frac{\partial^2 H_{mz}}{\partial x^2} = (k_z^2 - k_0^2)H_{mz}, \quad (34-2)$$

where

$$k_0 = \frac{1}{\lambda_0} = \frac{2\pi}{\lambda_0} = \frac{2\pi f}{c} = \frac{\omega}{c} \quad (34-3)$$

is known, for a given frequency. However,

$$k_z = \frac{1}{\lambda_z} \quad (34-4)$$

is unknown, λ_z being the wavelength of the guided wave.

We expect an interference pattern of some sort in the x direction. So H_{mz} is a sinusoidal function of x , and this requires that the expression in parentheses in Eq. 34-2 be negative. So we know that $k_z < k_0$ and hence that $\lambda_z > \lambda_0$, or that the wavelength measured along the guide is longer than the wavelength of a plane wave in air. This makes the phase velocity larger than c , which is correct; the group velocity will turn out to be smaller than c .

Thus

$$H_{mz} = M \cos(k_x x + \alpha), \quad k_x = +(\kappa_0^2 - k_z^2)^{1/2}, \quad (34-5)$$

where M is an arbitrary constant that defines the amplitude of the wave.

We now apply the boundary conditions of Sec. 33.1.4:

$$\frac{\partial H_{mz}}{\partial x} = 0 \quad \text{at } x = 0, a, \quad (34-6)$$

$$\frac{\partial H_{mz}}{\partial y} = 0 \quad \text{at } y = 0, b. \quad (34-7)$$

The second condition is already satisfied because $\partial/\partial y = 0$. From the first condition,

$$k_z \sin \alpha = 0 \quad \text{and} \quad k_x \sin(k_x a + \alpha) = 0. \quad (34-8)$$

Now k_x is a positive number. Therefore

$$\alpha = 0 \quad \text{and} \quad k_x a = n\pi, \quad (34-9)$$

where n is an integer.

Observe that $k_x a$ can take on only discrete or *eigen* values and that $n = 0$ is forbidden:

$$n = 1, 2, 3, \dots \quad (34-10)$$

So

$$H_{mz} = M \cos \frac{n\pi x}{a} \quad (34-11)$$

and, from Eq. 34-5,

$$k_z = \left[k_0^2 - \left(\frac{n\pi}{a} \right)^2 \right]^{1/2} = \frac{\{1 - [n\lambda_0/(2a)]^2\}^{1/2}}{\lambda_0}. \quad (34-12)$$

Then, from Eqs. 33-13 and 33-14, remembering that both E_{mz} and $\partial/\partial y$ are zero,

$$E_{my} = \frac{-j\omega\mu_0}{-k_x^2} \frac{\partial H_{mz}}{\partial x} = \frac{j\omega\mu_0}{k_x^2} (-k_x M \sin k_x x) \quad (34-13)$$

$$= -\frac{j\omega\mu_0}{k_x} M \sin k_x x = -\frac{j\omega\mu_0 a}{n\pi} M \sin \frac{n\pi x}{a}, \quad (34-14)$$

$$H_{mx} = \frac{j k_z a}{n\pi} M \sin \frac{n\pi x}{a}. \quad (34-15)$$

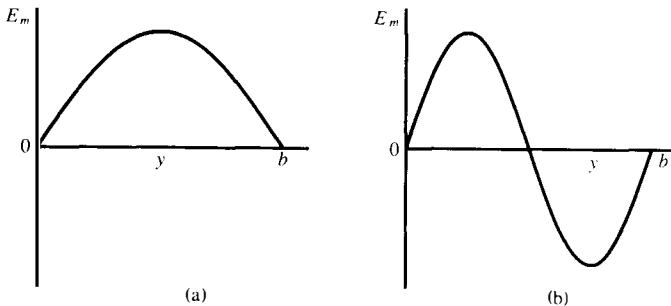


Fig. 34-3. The amplitude $E_m = E_{my}$ of the electric field strength for a TE wave.
(a) TE_1 mode. (b) TE_2 mode.

Figure 34-3(a) shows E_{my} as a function of x for $n = 1$: \mathbf{E} is zero along the walls and maximum in the plane $x = a/2$. With $n = 2$, E_{my} is zero at $x = a/2$.

The various values of n thus correspond to different modes of propagation, denoted as TE_1 , TE_2 , etc. As we shall see below, TE_1 is the only useful mode.

Summarizing,

$$E_{mx} = 0, \quad E_{my} = -\frac{j\omega\mu_0 a}{n\pi} M \sin \frac{n\pi x}{a}, \quad E_{mz} = 0, \quad (34-16)$$

$$H_{mx} = \frac{jk_z a}{n\pi} M \sin \frac{n\pi x}{a} = \frac{k_z}{\omega\mu_0} E_{my}, \quad H_{my} = 0, \quad H_{mz} = M \cos \frac{n\pi x}{a}, \quad (34-17)$$

$$k_z = \frac{\{1 - [n\lambda_0/(2a)]^2\}^{1/2}}{\lambda_0}. \quad (34-18)$$

34.2 THE CUTOFF WAVELENGTH. NONPROPAGATING FIELDS

From Eq. 34-18,

$$k_z = \left[k_0^2 - \left(\frac{n\pi}{a} \right)^2 \right]^{1/2} = \left[\left(\frac{\omega}{c} \right)^2 - \left(\frac{n\pi}{a} \right)^2 \right]^{1/2}. \quad (34-19)$$

For

$$\omega > \frac{n\pi c}{a}, \quad \text{or} \quad \lambda_0 < \frac{2a}{n}, \quad (34-20)$$

k_z is real and a wave can propagate unattenuated down the guide.

The wavelength $2a/n$ is the *cutoff wavelength* for the TE _{n} mode. This corresponds to the condition $\omega = \omega_p$ for propagation in an ionized gas. At that wavelength $k_x = 0$ and $\lambda_z \rightarrow \infty$.

At wavelengths larger than $2a/n$, k_z is imaginary, there is no wave, and the field decreases exponentially with z . There is zero power flow once the field is established. At these longer wavelengths the field amplitude decreases rapidly with z . For example, at twice the cutoff wavelength, where the frequency is too low by a factor of 2,

$$\frac{\omega}{c} = \frac{n\pi}{2a}, \quad k_z = \left[\left(\frac{\omega}{c} \right)^2 - \left(\frac{2\omega}{c} \right)^2 \right]^{1/2} = -j \frac{\omega}{c} 3^{1/2}. \quad (34-21)$$

We choose the negative sign before the square root so that the amplitude will decrease exponentially with z , and then

$$\exp(-jk_z z) = \exp\left(-\frac{2\pi 3^{1/2} z}{\lambda_0}\right) = \exp\left(-10.88 \frac{z}{\lambda_0}\right). \quad (34-22)$$

The amplitude decreases by a factor of 5×10^4 in one free-space wavelength λ_0 !

The waveguide thus acts as a *high-pass* filter, with the lower frequency limit fixed by the width a , and not by b .

The free-space wavelength λ_0 must be shorter than twice the distance between the reflecting walls. For example, if $a = 100$ millimeters, then λ_0 must be less than 200 millimeters and the frequency must be higher than 1.5 gigahertz (1.5×10^9 hertz).

34.3 THE TE₁ MODE

In practice, one selects first the operating frequency, and then a guide whose dimensions are such that it can carry only the $n = 1$ mode. This condition requires that $2a$ be larger than λ_0 , as above. But a must be less than λ_0 to make TE₂, TE₃, . . . forbidden modes. Thus the dimension a must be such that

$$a < \lambda_0 < 2a. \quad (34-23)$$

With single-mode propagation the field configuration is well defined. Rectangular metallic waveguides are narrow band devices: for a given a , λ_0 can vary by at most a factor of 2.

The antennas of Fig. 34-1 launch an assortment of modes, but only the TE₁ survives.

We now write out the field components for the $n = 1$ mode. We simplify the notation by setting

$$E'_{my} = -\frac{j\omega\mu_0 a M}{\pi}. \quad (34-24)$$

Then

$$E_x = 0, \quad E_y = E'_{my} \sin \frac{\pi x}{a} \exp j(\omega t - k_z z), \quad E_z = 0, \quad (34-25)$$

$$H_x = -\frac{E'_{my} k_z}{\omega \mu_0} \sin \frac{\pi x}{a} \exp j(\omega t - k_z z), \quad H_y = 0, \quad (34-26)$$

$$H_z = \frac{\pi E'_{my}}{\omega \mu_0 a} \cos \frac{\pi x}{a} \exp j\left[\omega t - k_z \left(z - \frac{\lambda_z}{4}\right)\right], \quad (34-27)$$

where

$$k_z = \frac{2\pi}{\lambda_z} = \frac{\{1 - [\lambda_0/(2a)]^2\}^{1/2}}{\lambda_0} = \left[1 - \left(\frac{\lambda_0}{2a}\right)^2\right]^{1/2} k_0, \quad (34-28)$$

$$\lambda_z = \frac{\lambda_0}{\{1 - [\lambda_0/(2a)]^2\}^{1/2}}. \quad (34-29)$$

Figure 34-4 shows lines of \mathbf{E} and of \mathbf{H} for this TE_1 mode.

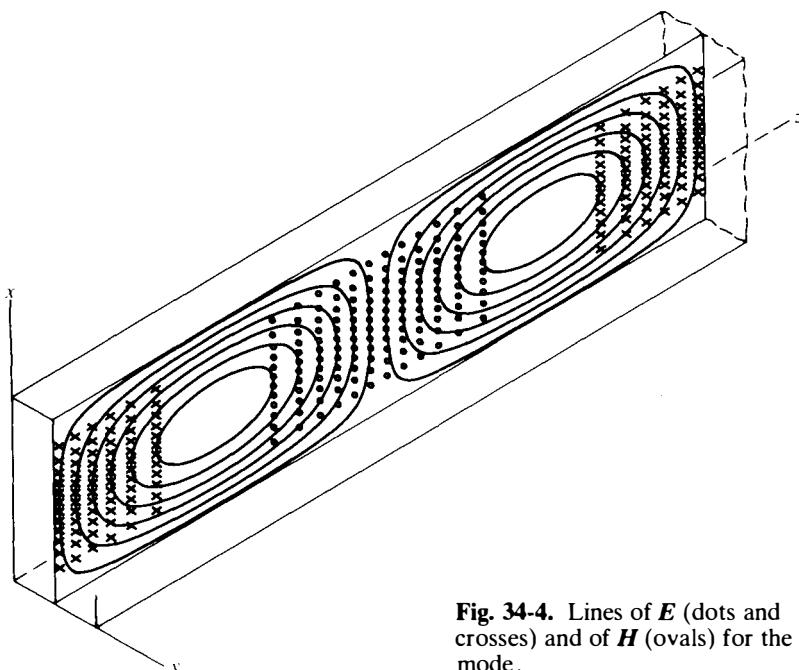


Fig. 34-4. Lines of \mathbf{E} (dots and crosses) and of \mathbf{H} (ovals) for the TE_1 mode.

34.4 MULTIPLE REFLECTIONS

The multiple reflections that take place between the top and bottom walls are interesting. We need only consider the \mathbf{E} field. We substitute k_x for π/a , in agreement with Eq. 34-9.

First recall that

$$\sin k_x x = \frac{\exp jk_x x - \exp (-jk_x x)}{2j}. \quad (34-30)$$

Then we can rewrite the expression for E_y as follows:

$$E_y = \frac{E'_{my}}{2j} [\exp j(\omega t + k_x x - k_z z) - \exp j(\omega t - k_x x - k_z z)]. \quad (34-31)$$

The first term between the brackets represents a wave whose vector wave number is $-k_x \hat{x} + k_z \hat{z}$ (App. C). It is oriented as in Fig. 34-5(a). The second term corresponds to the wave in Fig. 34-5(b). So the sine term in Eq. 34-25 for E_y does come from the interference between an up-going and a down-going wave.

Similarly, the \mathbf{H} vector is the sum of two terms of the form $\mathbf{k} \times \mathbf{E}/(\omega\mu_0)$, corresponding to the up-going and down-going waves.

Mathematically, there is a single up-going wave and a single down-going wave.

The angle of incidence θ is given by

$$\sin \theta = \frac{k_z}{k_0} = \left[1 - \left(\frac{\lambda_0}{2a} \right)^2 \right]^{1/2}, \quad (34-32)$$

$$\cos \theta = \frac{\lambda_0}{2a}. \quad (34-33)$$

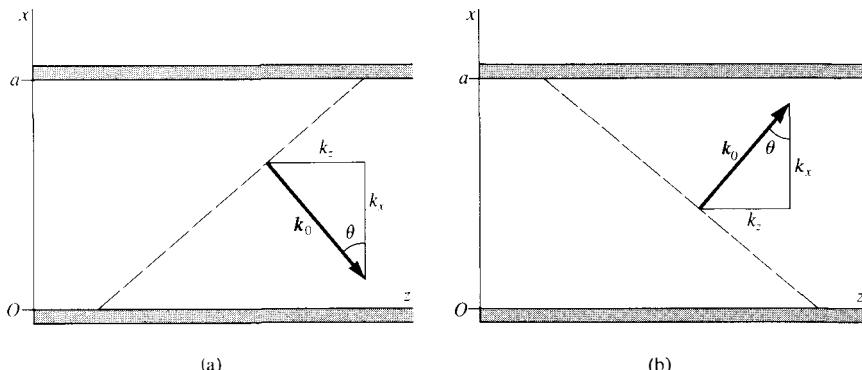


Fig. 34-5. Wave fronts and the \mathbf{k} 's for the (a) down-going and (b) up-going waves. Observe that $k_0^2 = k_x^2 + k_z^2$, as in Eq. 34-5. The angle of incidence is θ .

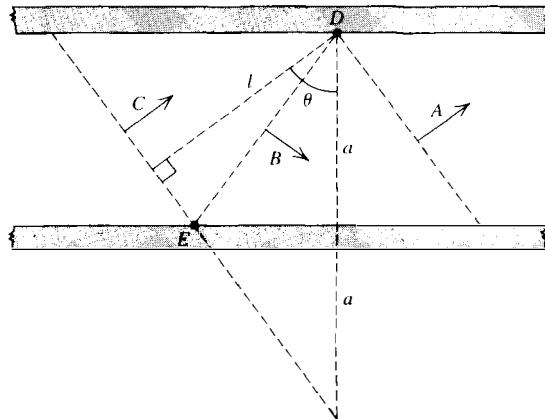


Fig. 34-6. Geometric construction for deducing Eq. 34-33.

For a given guide width a and a given free-space wavelength λ_0 , propagation in the TE₁ mode occurs through reflection only at that particular angle. For $\lambda_0 \ll 2a$, $\theta \approx \pi/2$ and the wave is nearly TEM.

In practice, $a < \lambda_0 < 2a$ as in Eq. 34-23 and θ lies between 0 and 60°. Typically (Table 34-1), θ is about 50°.

Equation 34-33 has a simple explanation. Refer to Fig. 34-6. Lines A and C are wave fronts of the wave that travels upward. At a given time t , set $E_A = \exp j\omega t$ for that wave along A . Then, along line C ,

$$E_C = \exp j\left(\omega t + \frac{2\pi l}{\lambda_0}\right). \quad (34-34)$$

Because of the phase shift upon reflection at D , the E -field of the downward wave front B must lead that of the upward wave front A by π radians and $E_B = \exp j(\omega t + \pi)$.

Similarly, because of the phase shift upon reflection at E ,

$$E_C = \exp j(\omega t + 2\pi). \quad (34-35)$$

Equating the two expressions for E_C and setting

$$l = 2a \cos \theta, \quad (34-36)$$

then

$$2a \frac{\cos \theta}{\lambda_0} = 1, \quad \cos \theta = \frac{\lambda_0}{2a}. \quad (34-37)$$

34.5 THE PHASE, SIGNAL, AND GROUP VELOCITIES

The phase velocity is

$$v_p = f\lambda_z = \omega\lambda_z = \frac{\omega}{k_z} \quad (34-38)$$

$$= \frac{c}{\{1 - [\lambda_0/(2a)]^2\}^{1/2}} = \frac{c}{\sin \theta} > c. \quad (34-39)$$

This is the velocity at which the phase propagates down the guide. It is larger than the speed of light c because the TE₁ wave is the superposition of two plane waves whose \mathbf{k} 's are inclined as in Fig. 34-2. For example, when the wave front AB of Fig. 34-7 moves downward at speed c to $A'B'$, point A moves to the right at a speed $v_p > c$.

At what velocity does a signal progress down the line? From Fig. 34-8 this is $c \sin \theta$. So the *signal velocity* is

$$v_s = c \sin \theta < c, \quad v_p v_s = c^2. \quad (34-40)$$

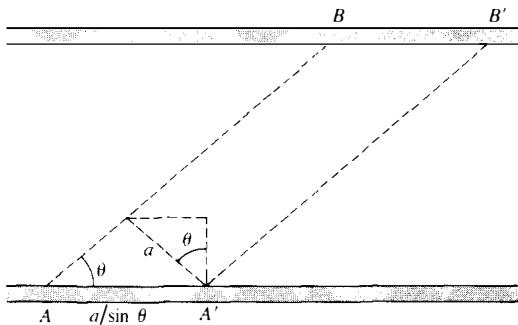


Fig. 34-7. As a wave front AB moves to $A'B'$ through the distance a , the point A moves to A' through a larger distance $a/\sin \theta$, where θ is the angle of incidence.

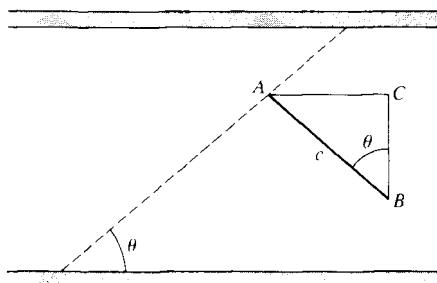


Fig. 34-8. As a wave front travels from A to B at the speed c , the signal progresses along the guide through the shorter distance $AB \sin \theta$.

This is in agreement with the fact that a signal cannot propagate at a speed larger than c (Sec. 13.5.3).

The group velocity v_g is equal to the signal velocity v_s . See Prob. 34-8.

34.6 THE TRANSMITTED POWER

To calculate the average power transmitted through the guide we integrate the time-averaged Poynting vector over the cross section of the guide:

$$P_{T,\text{av}} = \int_0^b \int_0^a \mathcal{S}_{\text{av}} dx dy = \int_0^b \int_0^a \frac{1}{2} \operatorname{Re} |(\mathbf{E} \times \mathbf{H}^*)| dx dy. \quad (34-41)$$

Here

$$\mathcal{S}_{\text{av}} = \frac{1}{2} \operatorname{Re} \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ 0 & E_y & 0 \\ H_x^* & 0 & H_z^* \end{vmatrix} = \frac{1}{2} \operatorname{Re} (E_y H_z^* \hat{\mathbf{x}} - E_y H_x^* \hat{\mathbf{z}}), \quad (34-42)$$

where the components of \mathbf{E} and of \mathbf{H} are as in Eqs. 34-25 to 34-27. After substituting 2π for $k_z \lambda_z$, the expression for H_z becomes

$$H_z = \frac{j\pi E'_{my}}{\omega \mu_0 a} \cos \frac{\pi x}{a} \exp j(\omega t - k_z z). \quad (34-43)$$

Thus the x component of the time-averaged Poynting vector is zero. The net power flows in the direction of the z -axis and

$$\mathcal{S}_{\text{av}} = \frac{E'^2_{my} k_z}{2\omega \mu_0} \sin^2 \frac{\pi x}{a} \hat{\mathbf{z}}. \quad (34-44)$$

We assume that E'_{my} is real.

The average power density \mathcal{S}_{av} is independent of y , as expected, since both \mathbf{E} and \mathbf{H} are independent of y . It is zero at $x = 0$, $x = a$ where \mathbf{E} is zero and maximum at $x = a/2$.

The time-averaged transmitted power is thus

$$P_{T,\text{av}} = \frac{E'^2_{my} k_z}{2\omega \mu_0} \int_0^a \sin^2 \frac{\pi x}{a} b dx \quad (34-45)$$

$$= \frac{E'^2_{my} k_z ab}{4\omega \mu_0} = \frac{E'^2_{my} k_0 ab}{4\omega \mu_0} \left[1 - \left(\frac{\lambda_0}{2a} \right)^2 \right]^{1/2} \quad (34-46)$$

$$P_{T,\text{av}} = E_{my}^{\prime 2} \frac{ab}{4} \left(\frac{\epsilon_0}{\mu_0} \right)^{1/2} \left[1 - \left(\frac{\lambda_0}{2a} \right)^2 \right]^{1/2}. \quad (34-47)$$

34.7 THE STORED ENERGY DENSITY

The instantaneous electric energy density is $\frac{1}{2}\epsilon_0|\text{Re } E|^2$, and its time-averaged value is

$$\mathcal{E}' = \frac{\epsilon_0}{4} E_m^2 = \frac{\epsilon_0}{4} E_{my}^{\prime 2} \sin^2 \frac{\pi x}{a}. \quad (34-48)$$

The time-averaged electric energy stored per unit length of guide is thus

$$\mathcal{E}'' = \frac{\epsilon_0}{4} E_{my}^{\prime 2} \int_0^a \sin^2 \left(\frac{\pi x}{a} \right) b \, dx = \frac{\epsilon_0}{8} ab E_{my}^{\prime 2}. \quad (34-49)$$

Similarly, the instantaneous magnetic energy density is $\frac{1}{2}\mu_0|\text{Re } H|^2$, with a time-averaged value of

$$\mathcal{E}'_{\text{mag}} = \frac{\mu_0}{4} H_m^2 = \frac{\mu_0}{4} (|H_{mx}|^2 + |H_{mz}|^2) \quad (34-50)$$

$$= \frac{\mu_0}{4} \left[\left(\frac{E'_{my} k_z}{\omega \mu_0} \right)^2 \sin^2 \frac{\pi x}{a} + \left(\frac{E'_{my} \pi}{\omega \mu_0 a} \right)^2 \cos^2 \frac{\pi x}{a} \right]. \quad (34-51)$$

The electric energy density is highest near the center, while the magnetic energy density is highest near the sides. Upon integrating over the cross section of the guide as above, one finds that, per unit length of guide,

$$\mathcal{E}''_{\text{mag}} = \mathcal{E}''_{\text{el}} = \frac{\epsilon_0}{8} ab E_{my}^{\prime 2}. \quad (34-52)$$

As one might expect, the transmitted power is equal to the energy per unit length multiplied by the signal velocity:

$$P_{T,\text{av}} = (\mathcal{E}''_{\text{el}} + \mathcal{E}''_{\text{mag}}) v_s. \quad (34-53)$$

34.8 ATTENUATION

We have assumed until now that the waveguide walls are perfectly conducting; it is time to consider real waveguides of finite conductivity.

In the process of guiding electromagnetic waves, conductors dissipate part of the wave energy in the form of Joule losses, because the wave induces electric currents in the guide. A rigorous calculation of the field for a guide of finite conductivity is difficult, but fortunately unnecessary.

The procedure for calculating the Joule losses is the following. As a first approximation, we consider the guide conductivity σ to be infinite. Then the field is the one that we found above. As a second approximation, we make σ large but not infinite. A refracted wave penetrates the conductor, and a tiny part of the electromagnetic energy is lost as heat. As a result, the amplitude of the guided wave decreases with increasing z .

The attenuation is exponential for the following reason. The amplitude of the wave refracted into the guide is proportional to the amplitude of the incident wave, or to the amplitude of the guided wave. Then the power lost in the guide is proportional to the power in the wave, which is the condition required for an exponential attenuation. See also below.

In this second approximation, with σ large but not infinite, \mathbf{E} and \mathbf{H} are the same as previously, except that k_z is now complex:

$$k_z = \beta - j\alpha. \quad (34-54)$$

Let us calculate α . The vector \mathbf{H} is tangential at the surface of the guide. Since the guide material has a finite σ , the skin depth is finite, there are no true surface currents, and the tangential \mathbf{H} just inside the guide is the same as the tangential \mathbf{H} just outside. A wave penetrates into the conductor and, from Sec. 29.1,

$$\frac{E}{H} = \left(\frac{\mu_0 \omega}{\sigma} \right)^{1/2} \exp j \frac{\pi}{4}. \quad (34-55)$$

This small \mathbf{E} exists on both sides of the interface. We assume that the guide is nonmagnetic. This \mathbf{E} is a *perturbation* of the ideal field that we derived for perfect conductors. The method is satisfactory because this small \mathbf{E} hardly disturbs the wave.

We thus have a tangential \mathbf{E} , a tangential \mathbf{H} , and a Poynting vector that is normal to the conducting surface and directed into the metal.

We can thus calculate the average power P'_L that leaks into the conductor per meter of length. We then require the attenuation constant α . This constant must be such that multiplying both the \mathbf{E} and the \mathbf{H} of the guided wave that we found above by $\exp(-\alpha z)$ decreases the average transmitted power $P_{T,av}$ by a factor of

$$\exp(-2\alpha \Delta z) \approx 1 - 2\alpha \Delta z \quad (34-56)$$

in a distance Δz . The approximation is excellent for standard guides. Then

$$P'_L \Delta z = (2\alpha \Delta z) P_T, \quad \alpha = \frac{P'_L}{2P_T}. \quad (34-57)$$

The real part β of k_z is approximately equal to the k_z obtained on the assumption of perfectly conducting walls. You will recall that we calculated $P_{T,\text{av}}$ in Sec. 34.6.

We can calculate P'_L as follows. Along the face $x = 0$, and at $z = 0$, from Eq. 34-27,

$$H = H_z = -\frac{\pi E'_{my}}{\mu_0 \omega a} j \exp j\omega t. \quad (34-58)$$

This is independent of y . Now, from Prob. 29-5, the power dissipated per square meter is the same as if one had a current I of the same magnitude as H , spread uniformly over a cross section $1 \times \delta$, where δ is the skin depth. So the power dissipation per square meter is $I_{\text{rms}}^2 \times 1/(\sigma \times 1 \times \delta)$, or $(H_m^2/2)/(\sigma \delta)$. Then, for the face $x = 0$ and at $z = 0$,

$$\mathcal{P}'_{\text{av}} = \frac{1}{2} \left(\frac{\pi E'_{my}}{\mu_0 \omega a} \right)^2 \frac{1}{\sigma \delta} = \frac{1}{2\sigma} \left(\frac{\pi E'_{my}}{\mu_0 \omega a} \right)^2 \left(\frac{\mu_0 \sigma \omega}{2} \right)^{1/2} = \frac{(\pi E'_{my}/a)^2}{\sigma^{1/2} (2\mu_0 \omega)^{3/2}}. \quad (34-59)$$

The power lost over a length of 1 meter in one face is b times larger, and for the two faces at $x = 0$ and $x = a$, it is

$$P'_{L,\text{av}} = \frac{(\pi E'_{my}/a)^2 2b}{\sigma^{1/2} (2\mu_0 \omega)^{3/2}} \quad (\text{faces } x = 0 \text{ and } x = a). \quad (34-60)$$

For the faces at $y = 0$ and $y = b$, \mathbf{H} has two components. The power lost over 1 meter of guide in these two faces proves to be

$$P'_{L,\text{av}} = \left[\frac{(\pi E'_{my})^2}{\sigma^{1/2} (2\mu_0 \omega)^{3/2} a} \right] \left[\frac{2a}{\lambda_0} \right]^2 \quad (\text{faces } y = 0 \text{ and } b). \quad (34-61)$$

Summing these two powers,

$$P'_{L,\text{av}} = \left[\frac{(\pi E'_{my})^2}{\sigma^{1/2} (2\mu_0 \omega)^{3/2} a} \right] \left[\frac{2b}{a} + \left(\frac{2a}{\lambda_0} \right)^2 \right]. \quad (34-62)$$

Finally, using Eqs. 34-47 and 34.57,

$$\alpha = \frac{1}{b} \left[\frac{\pi (\epsilon_0 / \mu_0)^{1/2}}{\sigma \lambda_0} \right]^{1/2} \frac{1 + (2b/a)[\lambda_0/(2a)]^2}{\{1 - [\lambda_0/(2a)]^2\}^{1/2}} \quad (34-63)$$

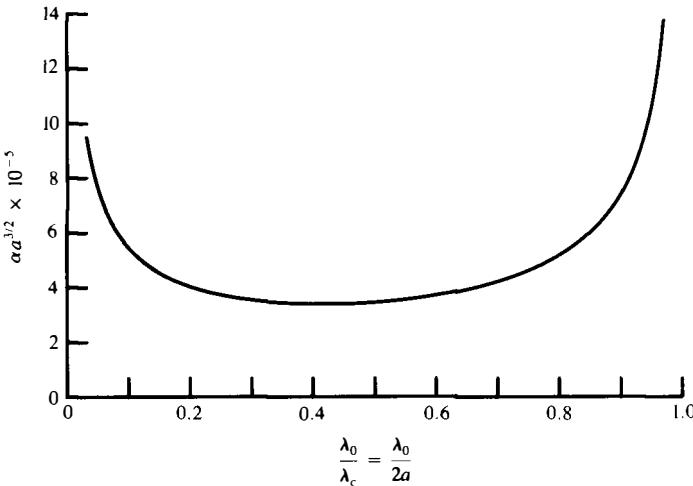


Fig. 34-9. The product $\alpha a^{3/2}$ as a function of the ratio $\lambda_0/(2a)$ for $b = a/2$ and for copper.

$$\alpha \approx \frac{1}{b(120\sigma\lambda_0)^{1/2}} \frac{1 + (2b/a)[\lambda_0/(2a)]^2}{\{1 - [\lambda_0/(2a)]^2\}^{1/2}}. \quad (34-64)$$

In practice, $2b/a \approx 1$. This makes the magnitudes of the power densities about equal on all four sides. This also excludes a TE_1 mode with reflection on the other pair of sides.

Figure 34-9 shows $\alpha a^{3/2}$ as a function of $\lambda_0/(2a)$ for copper and for $b = a/2$. In the range $\lambda_0 = 0.2 \times 2a$ to $0.7 \times 2a$,

$$\alpha \approx \frac{4 \times 10^{-5}}{a^{3/2}}. \quad (34-65)$$

The optimum value of λ_0 is about $0.4 \times 2a$, but actual values are somewhat larger so as to achieve strong attenuation for the $n = 2$ mode.

The attenuation of the $n = 1$ mode in commercial guides is of the order of 0.1 decibel/meter[†] at frequencies of a few gigahertz.

This is a high rate of attenuation. For example, over 100 meters the attenuation is about 10 decibels, which means that the amplitude of the wave decreases by a factor of 3. Such guides transmit high-power signals over distances of the order of tens of meters or less, usually for

[†] $20 \log [(\text{amplitude at the output})/(\text{amplitude at the input})] = -0.1$.

connecting a transmitter or receiver to an antenna, either for radar or for microwave links.

Table 34-1 gives the main characteristics of a few standard types of rectangular waveguide.

34.9 SUMMARY

A *transverse electric* (TE) wave can propagate inside a rectangular waveguide by reflection from two opposite sides, as in Fig. 34-2. Then

$$E_{mx} = 0, \quad E_{my} = -\frac{j\omega\mu_0 a}{n\pi} M \sin \frac{n\pi x}{a}, \quad E_{mz} = 0, \quad (34-16)$$

$$H_{mx} = \frac{jk_z a}{n\pi} M \sin \frac{n\pi x}{a}, \quad H_{my} = 0, \quad H_{mz} = M \cos \frac{n\pi x}{a}, \quad (34-17)$$

$$k_z = \frac{\{1 - [n\lambda_0/(2a)]^2\}^{1/2}}{\lambda_0}. \quad (34-18)$$

Here n is the mode order, and, as usual, the subscript m stands for “maximum value of.”

In practice, one uses the TE₁ mode, and then

$$E_x = 0, \quad E_y = E'_{my} \sin \frac{\pi x}{a} \exp j(\omega t - k_z z), \quad E_z = 0, \quad (34-25)$$

$$H_x = -\frac{E'_{my} k_z}{\omega \mu_0} \sin \frac{\pi x}{a} \exp j(\omega t - k_z z), \quad H_y = 0, \quad (34-26)$$

$$H_z = \frac{\pi E'_{my}}{\omega \mu_0 a} \cos \frac{\pi x}{a} \exp j\left[\omega t - k_z \left(z - \frac{\lambda_z}{4}\right)\right], \quad (34-27)$$

$$k_z = \frac{2\pi}{\lambda_z} = \frac{\{1 - [\lambda_0/(2a)]^2\}^{1/2}}{\lambda_0}, \quad (34-28)$$

where λ_z is the wavelength of the guided wave.

With this mode the *angle of incidence* (Fig. 34-2) is given by

$$\sin \theta = \left[1 - \left(\frac{\lambda_0}{2a}\right)^2\right]^{1/2}. \quad (34-32)$$

The phase and group velocities are

$$v_p = \frac{c}{\sin \theta} > c, \quad v_g = v_s = c \sin \theta < c. \quad (34-39), (34-40)$$

The time-averaged *Poynting vector* is

$$\mathcal{S}_{av} = \frac{E_{my}^{\prime 2} k_z}{2\omega\mu_0} \sin^2 \frac{\pi x}{a} \hat{z} \quad (34-44)$$

and the time-averaged transmitted power is

$$P_{T,av} = E_{my}^{\prime 2} \frac{ab}{4} \left(\frac{\epsilon_0}{\mu_0} \right)^{1/2} \left[1 - \left(\frac{\lambda_0}{2a} \right)^2 \right]^{1/2}, \quad (34-47)$$

where $(\mu_0/\epsilon_0)^{1/2} \approx 377$.

In real waveguides of finite conductivity σ , the field is hardly different, except that

$$k_z = \beta - j\alpha \quad (34-54)$$

with β equal to the above k_z and

$$\alpha \approx \frac{1}{b(120\sigma\lambda_0)^{1/2}} \frac{1 + (2b/a)[\lambda_0/(2a)]^2}{\{1 - [\lambda_0/(2a)]^2\}^{1/2}}. \quad (34-64)$$

PROBLEMS

- 34-1. (34.3) Wavelength and frequency in a rectangular waveguide

An electromagnetic wave propagating in the TE_1 mode in a 34.0×72.1 millimeter rectangular waveguide has a wavelength λ_z of 138 millimeters. Calculate its frequency.

- 34-2. (34.3) The surface currents and the Poynting vector in a rectangular metallic wave guide

Figure 34-10 shows three sides of a rectangular guide that is split open and flattened. Draw a figure like this and show, on face B , lines of \mathbf{H} , electric charges, and vectors $\mathbf{E} \times \mathbf{H}$ at a given instant. Then add lines of current on all three faces.

- 34-3. (34.3) The angle of incidence θ in a rectangular guide

According to Table 34-1, the recommended range of operating frequencies for a rectangular waveguide with a cross section of $34.0 \times$

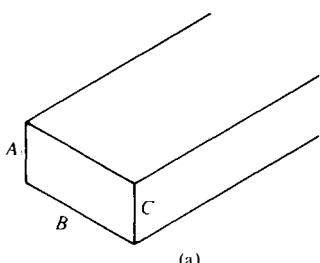


Fig. 34-10.

72.1 millimeters extends from 2.61 to 3.95 gigahertz.
Calculate the values of θ at both ends of this range.

34-4. (34.3) Measuring the standing wave ratio in a rectangular guide

If a load is not properly matched to a waveguide, part of the incident wave turns back and there is a standing wave along the guide. Then only a fraction of the power available at the source reaches the load. It is therefore useful to be able to move a small probe along a longitudinal slot to sample the field inside the guide. The voltage standing wave ratio (VSWR) is the ratio of the maximum to the minimum time-averaged rms voltage measured at the probe. Under ideal conditions there is no reflected wave, and the VSWR is equal to unity.

The probe can be a short length of wire that responds to the electric field or a small loop coupled to the magnetic field. The probe projects into the field by about 1 millimeter.

- (a) With the TE_1 mode, where should the slot be cut to disturb the wave as little as possible?
- (b) If the probe is a loop, how should it be oriented?
- (c) How would you proceed to measure the wavelength of the guided wave?
- (d) If the VSWR is equal to 2, what is the value of the ratio $E_{\text{reflected}}/E_{\text{incident}}$?

34-5. (34.3) An artificial dielectric

Figure 34-11 shows a set of parallel conducting plates uniformly spaced by a distance s . If s/λ has the correct value and if the incident wave is correctly polarized, this medium acts as an *artificial dielectric* whose index of refraction is *less* than unity.

- (a) What must be the orientation of \mathbf{E} ?
- (b) Find the index of refraction as a function of the ratio s/λ_0 .
- (c) Show rays deflected by (i) a prism and (ii) a converging cylindrical lens made in this way.

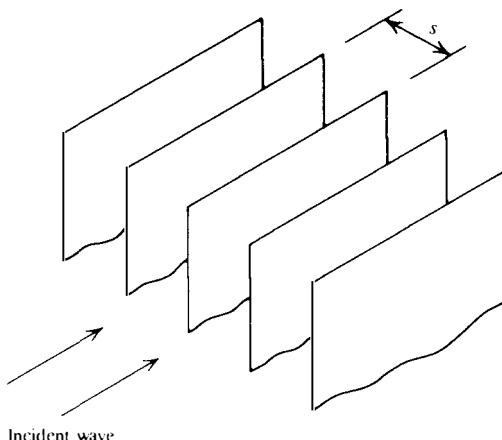


Fig. 34-11.

34-6. (34.3) Dielectric-filled rectangular waveguides

A rectangular metallic waveguide A is air-filled, and its cross section is a, b . Another rectangular metallic waveguide D is filled with a dielectric ϵ_r , and its cross section is $a/\epsilon_r^{1/2}, b/\epsilon_r^{1/2}$.

(a) Show that waveguides A and D have the same cutoff frequency and that, at a given frequency, $\lambda_{zD} = \lambda_{zA}/\epsilon_r^{1/2}$. Thus, for a given operating frequency, a dielectric-filled guide is smaller than an air-filled one. With Teflon ($\epsilon_r = 2.1$), both dimensions a and b are smaller by a factor of 1.45. Also, the phase and signal velocities are smaller by a factor of 1.45.

(b) Compare the power ratings of A and D at a given frequency.

The dielectric strength of a material is the maximum permissible value of E before breakdown. The dielectric strength of a good dielectric such as Teflon is of the order of 10 times that of air. Set this ratio equal to R . The value of R increases as the thickness decreases.

34-7. (34.5) The phase and group velocities in a rectangular waveguide

(a) Find λ_0/λ_c as a function of λ_0/λ_c for a rectangular waveguide, where λ_c is the cutoff wavelength.

(b) Find $\omega\lambda_c/c$ as a function of $k_z\lambda_c$. For a given frequency,

$$\frac{v_p}{c} = \frac{\omega}{k_z} = \frac{\omega\lambda_c}{c} k_z \lambda_c.$$

The group, or signal, velocity is given by

$$\frac{v_g}{c} = \frac{d(\omega\lambda_c/c)}{d(k_z\lambda_c)}.$$

34-8. (34.5) The group or signal velocity in a rectangular waveguide

Show that, in a hollow rectangular waveguide,

$$v_g = \frac{1}{dk_z/d\omega}.$$

34-9. (34.6) Measuring the transmitted power in a hollow rectangular guide

One can measure the power transmitted down a rectangular metallic waveguide by reading the voltage induced in a tiny loop projecting into the guide, as in Fig. 34-12. The loop is situated at $x = a/2, y = 0$, and it lies in the yz -plane.

Show that

$$P_{T,av} = 8.40 \times 10^{-6} \frac{\lambda_0^2 ab}{[1 - \lambda_0^2/(4a^2)]^{1/2}} \frac{\mathcal{V}^2}{\mathcal{A}^2},$$

where \mathcal{V} is the rms voltage induced in the loop and \mathcal{A} is the area of the loop. The effective value of \mathcal{A} is unknown, but the method is satisfactory for measuring relative values of P_T .

34-10. (34.8) Attenuation in rectangular guides

(a) Show that, upon reflection from a good nonmagnetic conductor, in air, the amplitude of an electromagnetic wave decreases by a factor of

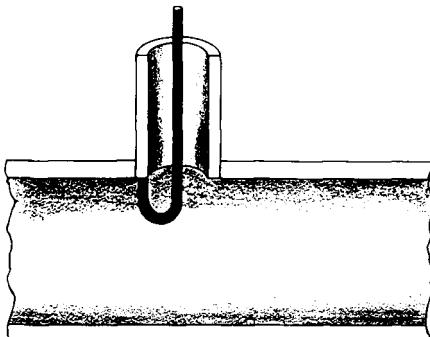


Fig. 34-12.

approximately $1 - (2\epsilon_0\omega/\sigma)^{1/2} \cos \theta$, if \mathbf{E} is normal to the plane of incidence.

(b) Show that this is in agreement with the attenuation calculated in Sec. 34.8.

- 34-11. (34.8) The relative power ratings of a coaxial line and of a rectangular guide

(a) Calculate the maximum power that can be carried by a coaxial line and by a rectangular metallic waveguide at 3.00 gigahertz. The coaxial line has a diameter ρ_2 of 25.0 millimeters, and the guide has an inside cross section of 37.5×75.0 millimeters. The coaxial line satisfies the condition for maximum power transfer, namely, $\rho_2/\rho_1 = 1.65$, and its outside radius is small enough to ensure the attenuation of higher-order modes. Both lines are air-filled, and the current-carrying surfaces are silver-plated. The maximum allowed E is 1.5 megavolts/meter. (Under ideal conditions the breakdown field at 3 gigahertz is about 10^8 volts/meter.) There is no reflected wave.

(b) Calculate the power dissipation per meter in both cases. Clearly, these lines can operate at these power levels only during short pulses.

(c) Calculate the rms voltage and current at the input end of the coaxial line.

- 34-12. (34.8) Decibels and nepers

Attenuation on a transmission line is expressed in decibels per meter. The number of decibels per meter is 20 times the logarithm to the base 10 of the ratio of the E 's (or the H 's, or the voltages, or the currents) at the two ends of a line 1 meter long. The degree of attenuation is also expressed in nepers per meter, and this is simply the value of β .

Show that 1 neper/meter is equivalent to 8.686 decibels/meter.

CHAPTER 35

GUIDED WAVES III

The Planar Optical Waveguide A. The Eigenvalue Equation

35.1	THE PLANAR DIELECTRIC GUIDE	648
35.2	RELATIONS BETWEEN THE FIELD COMPONENTS	650
35.3	THE FIELD COMPONENTS IN THE THREE MEDIA	651
35.4	THE CONDITIONS OF CONTINUITY AT THE INTERFACES	654
35.4.1	CONTINUITY AT THE $x = -a$ INTERFACE	654
35.4.2	CONTINUITY AT THE $x = a$ INTERFACE	655
35.5	THE FIELD COMPONENTS IN THE THREE MEDIA	655
35.6	THE PHASE SHIFTS ON TOTAL REFLECTION	656
35.7	THE EIGENVALUE EQUATION	657
35.8	A THIRD EXPRESSION FOR α	659
35.9	SUMMARY	659
	PROBLEMS	660

Chapters 35 and 36 are the last two chapters on guided waves. They concern a guiding structure that comprises only dielectrics. The wave again follows a zigzag path, as in the hollow conducting waveguides of the last chapter, but guidance results from total reflection.

In optical waveguides total reflection occurs inside a dielectric flanked by another one of a lower index of refraction. Although most of the power flows through the inner medium, some power also flows through the outer one, and the wave is not tightly confined as with a hollow metallic guide.

Dielectric waveguides serve mostly in the form of optical fibers for the transmission of information. However, there are many other forms. *Clad rods*, either straight or curved and up to a few millimeters in diameter, serve to convey light from one point to another. *Light guides* consist of bundles of fibers inside a flexible sheath. They serve the same purpose.

Fiber bundles are similar to light guides except that their ends are coordinated, which makes them suitable for transmitting images. *Fiberscopes* are useful, among other things, for medical examinations. They comprise a light guide, for illumination, and a fiber bundle equipped with lenses at each end. Some oscilloscopes have *fiber-optic*, or *microchannel, face plates*. These consist of a stack of parallel fibers, with both faces polished, and phosphor on the inner side. The fibers convey the light emitted by the phosphor more efficiently than a glass plate. See the first example in Sec. 31.2.2. This permits the observation of very fast transients. One commercial oscilloscope has a writing speed of 2×10^8 meters/second.

Optical fibers are made of either silica or plastic, and their diameters range from that of a human hair to about 0.5 millimeter. The index of refraction decreases with the radius, either gradually or abruptly. The smaller diameter fibers can transmit digital information at rates up to gigabits per second.

Total reflection has the advantage of being lossless. For example, the attenuation of a light wave in an optical fiber can be as low as 0.1 decibel (98.9% transmission) per *kilometer*, while the attenuation in a hollow conducting guide is of the same order over 1 *meter*. The attenuation in optical fibers comes largely from absorption, since no material is perfectly transparent.

Metallic guides serve at frequencies of several gigahertz, usually at high power, while dielectric guides serve at optical frequencies and low power. The two types are therefore complementary.

Rather than go through the fairly abstruse mathematics of fibers, we study the simplest form of optical guide, namely, the *planar optical waveguide*, composed of three layers of dielectric: a *substrate*, a *sheet*, and a *cover*. The indices of refraction of the substrate and of the cover are only slightly lower than that of the sheet. Similar planar guides are used in millimeter-wave integrated circuits. The physics of planar guides is basically the same as that of fibers.

Planar optical waveguides have overall thicknesses of the order of 10 micrometers and widths about 10 times larger. These guides are the basic components of various *integrated optics* devices that are, in a way, the optical analogs of integrated circuits.

Figure 35-1 shows two methods for launching a light wave into a planar dielectric guide.

Our discussion parallels that of the rectangular guides of Chap. 34. We study only the transverse electric (TE) waves in which the electric vector is normal to the direction of propagation.

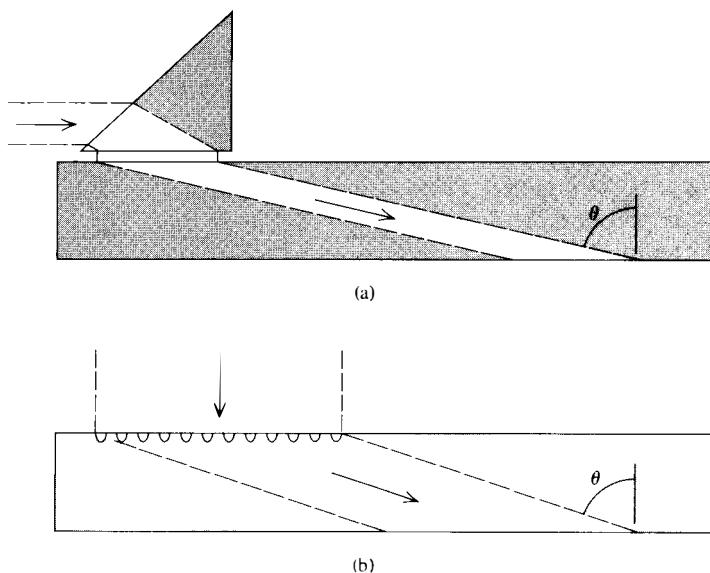


Fig. 35-1. Two methods for launching an electromagnetic wave in a planar optical waveguide. In both cases the coupling efficiency is about 80%. (a) The prism coupler uses frustrated total reflection. The distance between the prism and the guide is a fraction of a wavelength. The transmitted wave couples into the guide. The angle θ is, in fact, close to 90° . (b) Grating coupler. The grooves on the surface of the guide diffract a beam at the appropriate angle θ .

35.1 THE PLANAR DIELECTRIC GUIDE

We select axes as in Fig. 35-2. Medium 1 is the substrate. The sheet, or slab, is medium 2. Medium 3, the cover, can be air. Total reflection occurs at the interfaces 2,1 and 2,3.

We assume that all three media are homogeneous, isotropic, linear, and stationary (HILS) and that they are nonmagnetic ($\mu_r = 1$), lossless ($\sigma = 0$), and nondispersive (the index of refraction n is independent of the frequency).

To simplify the calculations, we also assume that medium 1 extends downward to infinity, while medium 3 extends upward to infinity. In practice, media 1 and 3 need only be several wavelengths thick.

We disregard reflected waves, which means that the guide extends to infinity in the $+z$ direction.

Finally, we assume that all three media extend to infinity in the plus and minus y directions. The reason for this assumption is the following.

Figure 35-3 shows a cross section of a guide whose core occupies only the region between $x = \pm a$ and $y = \pm b$. The cladding has a lower index

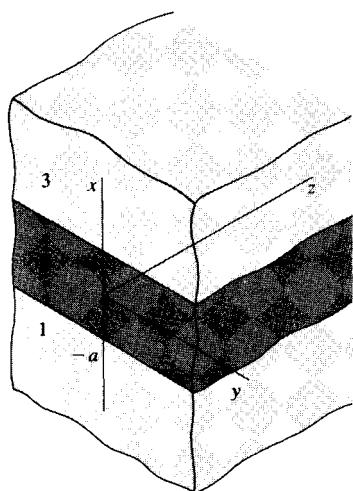


Fig. 35-2. Coordinate system for the planar optical waveguide. The three media are infinite in the y and z directions. Media 1 and 3 extend to infinity downward and upward, respectively. Total reflection at the 2,1 and 2,3 interfaces makes the wave propagate in zig zag fashion along the positive direction of the z -axis.

of refraction, and total reflection occurs at the $x = -a$ and $x = a$ interfaces. Then the wave above $x = a$ and below $x = -a$ spills over into the shaded region. The field in the core will also penetrate the $y = -b$ and $y = b$ interfaces, so as to satisfy the continuity conditions for the normal and tangential components of \mathbf{E} and of \mathbf{H} . So the field of a rectangular dielectric waveguide of finite cross section is quite complex. It is wiser for us to disregard these edge effects and make the guide infinite in the $\pm y$ direction. Our analysis will be satisfactory as long as the breadth of the core is much larger than its height. That is usually the case.

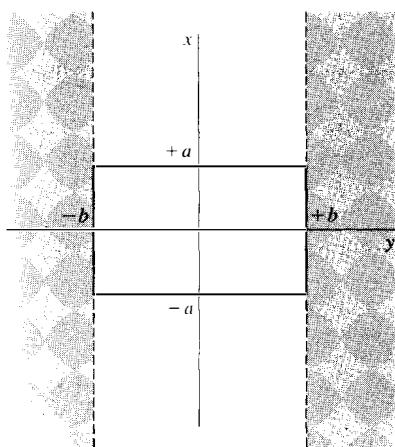


Fig. 35-3. Cross section of an optical waveguide of rectangular cross section.

The electromagnetic field propagates to the right at some phase velocity v_p , and the field components in all three media therefore have the form

$$\mathbf{E} = E_m \exp j(\omega t - k_z z) \hat{\mathbf{y}}, \quad (35-1)$$

where E_m is a function of x only and

$$k_z = \frac{1}{\lambda_z} = \frac{2\pi}{\lambda_z} = \frac{\omega}{v_p}, \quad (35-2)$$

λ_z being the wavelength of the guided wave. The wave number k_z is real because we have assumed zero attenuation.

We set

$$k_0 = \frac{1}{\lambda_0} = \frac{\omega}{c} = \omega(\epsilon_0 \mu_0)^{1/2}, \quad (35-3)$$

$$k_1 = \frac{1}{\lambda_1} = \frac{\omega}{v_{p1}} = \frac{\omega}{c} n_1 = n_1 k_0 = \omega(\epsilon_r \epsilon_0 \mu_0)^{1/2}, \quad (35-4)$$

and similarly for k_2 and k_3 . The wavelengths $\lambda_0, \lambda_1, \lambda_2, \lambda_3$ apply to uniform plane waves in a vacuum and in media 1, 2, and 3, respectively.

35.2 RELATIONS BETWEEN THE FIELD COMPONENTS

As stated previously, we consider only TE waves. You will recall from Sec. 33.1.1 that the transverse components of the field in a guided wave are simple functions of the longitudinal components. Here k_z is equal to neither k_1 nor k_2 nor k_3 .

By hypothesis, $\partial/\partial y = 0$. Also, in TE waves, $E_{mz} = 0$. Then, from Eqs. 33-11 to 33-15, setting $\mu_r = 1$,

$$E_{mx} = 0, \quad (35-5)$$

$$E_{my} = \frac{j\omega\mu_0}{k^2 - k_z^2} \frac{\partial H_{mz}}{\partial x}, \quad (35-6)$$

$$H_{mx} = -\frac{jk_z}{k^2 - k_z^2} \frac{\partial H_{mz}}{\partial x} = -\frac{k_z}{\omega\mu_0} E_{my}, \quad (35-7)$$

$$H_{my} = 0. \quad (35-8)$$

These equations apply to all three media, with k equal to k_1, k_2 , or k_3 .

To find H_z , we must solve its wave equation

$$\frac{\partial^2 H_z}{\partial x^2} + \frac{\partial^2 H_z}{\partial z^2} = \epsilon_r \epsilon_0 \mu_0 \frac{\partial^2 H_z}{\partial t^2} = -\omega^2 \epsilon_r \epsilon_0 \mu_0 H_z = -k^2 H_z, \quad (35-9)$$

where

$$\frac{\partial^2 H_z}{\partial z^2} = -k_z^2 H_z. \quad (35-10)$$

We must therefore solve the differential equation

$$\frac{\partial^2 H_z}{\partial x^2} = (k_z^2 - k^2) H_z \quad (35-11)$$

for each medium in succession with, again, k equal to k_1 , k_2 , or k_3 .

35.3 THE FIELD COMPONENTS IN THE THREE MEDIA

In medium 1 we require an exponential decrease in the $-x$ direction. Therefore, in medium 1, the expression between parentheses in Eq. 35-11 is positive, $k_z > k_1$, and

$$H_{mz} = K \exp(k_z^2 - k_1^2)^{1/2} x = K \exp k_{1x} x, \quad (35-12)$$

where K is a constant and k_{1x} is real and positive:

$$k_{1x} = +(k_z^2 - k_1^2)^{1/2}. \quad (35-13)$$

The value of k_z is as yet unknown.

Then, from Eqs. 35-6 and 35-7, in medium 1,

$$E_{my} = -\frac{j\omega\mu_0}{k_{1x}} H_{mz}, \quad (35-14)$$

$$H_{mx} = -\frac{k_z}{\omega\mu_0} E_{my}. \quad (35-15)$$

In medium 3 we require an exponential decrease in the $+x$ direction. Then the parenthesis in Eq. 35-11 is again positive, $k_z > k_3$, and we set

$$H_{mz} = L \exp[-(k_z^2 - k_3^2)^{1/2} x] = L \exp(-k_{3x} x). \quad (35-16)$$

Here L is another constant, and k_{3x} is real and positive:

$$k_{3x} = +(k_z^2 - k_3^2)^{1/2}. \quad (35-17)$$

Also, in medium 3, from Eqs. 35-6 and 35-7,

$$E_{my} = + \frac{j\omega\mu_0}{k_{3x}} H_{mz}, \quad (35-18)$$

$$H_{mx} = - \frac{k_z}{\omega\mu_0} E_{my}. \quad (35-19)$$

In medium 2,

$$\frac{\partial^2 H_{mz}}{\partial x^2} = (k_z^2 - k_2^2) H_{mz} = -k_{2x}^2 H_{mz}. \quad (35-20)$$

We have put a negative sign before the last term for the following reason. Both k_z and k_2 are real. Then k_{2x}^2 is real. If k_{2x} is real with $k_z^2 < k_2^2$, then H_{mx} is a sinusoidal function of x and there are two uniform plane waves in medium 2, zigzagging up and down the guide as in Sec. 34.4. This is the type of propagation that we are interested in here. So

$$k_{2x}^2 = k_2^2 - k_z^2, \quad k_{2x} = +(k_2^2 - k_z^2)^{1/2}. \quad (35-21)$$

We arbitrarily choose a plus sign before the parenthesis.

Observe that the k 's are related as in Fig. 35-4. Then

$$H_{mz} = M \cos(k_{2x}x + \alpha). \quad (35-22)$$

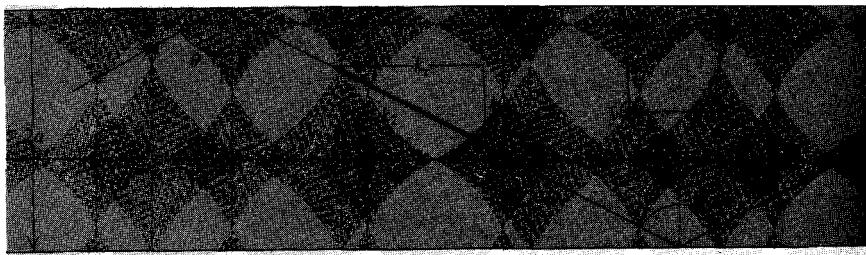


Fig. 35-4. Ray of light propagating down an optical waveguide by total reflection. Actual values of θ are close to 90° .

We have added a phase angle α to satisfy the continuity conditions at the interface. From Sec. 35.2,

$$E_{my} = -\frac{j\omega\mu_0}{k_{2x}} M \sin(k_{2x}x + \alpha), \quad (35-23)$$

$$H_{my} = -\frac{k_z}{\omega\mu_0} E_{my}. \quad (35-24)$$

Grouping Eqs. 35-13, 35-21, and 35-17 gives

$$k_1^2 = k_z^2 - k_{1x}^2, \quad k_2^2 = k_z^2 + k_{2x}^2, \quad k_3^2 = k_z^2 - k_{3x}^2. \quad (35-25)$$

All the k 's are real and positive. Therefore $k_1 < k_2$ and $k_3 < k_2$, in agreement with the requirements $n_1 < n_2$ and $n_3 < n_2$ for total reflection at the interfaces.

The situation is as follows.

$$\left. \begin{array}{l} H_{mz} = L \exp(-k_{3x}x), \\ E_{my} = +\frac{j\omega\mu_0}{k_{3x}} H_{mz}, \\ H_{mx} = -\frac{k_z}{\omega\mu_0} E_{my} = -\frac{jk_z}{k_{3x}} H_{mz}. \end{array} \right\} \quad (35-26)$$

$$\left. \begin{array}{l} H_{mz} = M \cos(k_{2x}x + \alpha), \\ E_{my} = -\frac{j\omega\mu_0}{k_{2x}} M \sin(k_{2x}x + \alpha), \\ H_{mx} = -\frac{k_z}{\omega\mu_0} E_{my}. \end{array} \right\} \quad (35-27)$$

$$\left. \begin{array}{l} H_{mz} = K \exp k_{1x}x, \\ E_{my} = -\frac{j\omega\mu_0}{k_{1x}} H_{mz}, \\ H_{mx} = -\frac{k_z}{\omega\mu_0} E_{my} = \frac{jk_z}{k_{1x}} H_{mz}. \end{array} \right\} \quad (35-28)$$

$$\left. \begin{array}{l} H_{mz} = M \cos(k_{2x}x + \alpha), \\ E_{my} = -\frac{j\omega\mu_0}{k_{2x}} M \sin(k_{2x}x + \alpha), \\ H_{mx} = -\frac{k_z}{\omega\mu_0} E_{my}. \end{array} \right\} \quad (35-29)$$

$$\left. \begin{array}{l} H_{mz} = K \exp k_{1x}x, \\ E_{my} = -\frac{j\omega\mu_0}{k_{1x}} H_{mz}, \\ H_{mx} = -\frac{k_z}{\omega\mu_0} E_{my} = \frac{jk_z}{k_{1x}} H_{mz}. \end{array} \right\} \quad (35-30)$$

$$\left. \begin{array}{l} H_{mz} = M \cos(k_{2x}x + \alpha), \\ E_{my} = -\frac{j\omega\mu_0}{k_{2x}} M \sin(k_{2x}x + \alpha), \\ H_{mx} = -\frac{k_z}{\omega\mu_0} E_{my}. \end{array} \right\} \quad (35-31)$$

$$\left. \begin{array}{l} H_{mz} = K \exp k_{1x}x, \\ E_{my} = -\frac{j\omega\mu_0}{k_{1x}} H_{mz}, \\ H_{mx} = -\frac{k_z}{\omega\mu_0} E_{my} = \frac{jk_z}{k_{1x}} H_{mz}. \end{array} \right\} \quad (35-32)$$

$$\left. \begin{array}{l} H_{mz} = M \cos(k_{2x}x + \alpha), \\ E_{my} = -\frac{j\omega\mu_0}{k_{2x}} M \sin(k_{2x}x + \alpha), \\ H_{mx} = -\frac{k_z}{\omega\mu_0} E_{my}. \end{array} \right\} \quad (35-33)$$

$$\left. \begin{array}{l} H_{mz} = K \exp k_{1x}x, \\ E_{my} = -\frac{j\omega\mu_0}{k_{1x}} H_{mz}, \\ H_{mx} = -\frac{k_z}{\omega\mu_0} E_{my} = \frac{jk_z}{k_{1x}} H_{mz}. \end{array} \right\} \quad (35-34)$$

The wave numbers k_{1x} , k_{2x} , and k_{3x} are all simple functions of k_z , as we have seen.

35.4 THE CONDITIONS OF CONTINUITY AT THE INTERFACES

There are two interfaces, and at each one the tangential components H_{mz} and E_{my} are continuous. This provides us with four independent equations. The normal component H_{mx} is also continuous, but its continuity is already ensured by that of E_{my} , which is a multiple of H_{mx} .

These four conditions of continuity will first give us the ratios K/M and L/M , in terms of the k 's and of α . They will also provide two independent expressions for α . Later these two expressions for α will lead us to the values of the four k 's. We use M as a measure of the amplitude of the wave, which is, of course, arbitrary.

35.4.1 Continuity at the $x = -a$ Interface

To simplify the calculations, set

$$k_{2x}a = b. \quad (35-35)$$

Because of the continuity of H_{mz} at $x = -a$,

$$M \cos(-b + \alpha) = K \exp(-k_{1x}a), \quad K = M \cos(b - \alpha) \exp k_{1x}a. \quad (35-36)$$

Also, from the continuity of E_{my} ,

$$-\frac{j\omega\mu_0}{k_{2x}} M \sin(-b + \alpha) = -\frac{j\omega\mu_0}{k_{1x}} K \exp(-k_{1x}a), \quad (35-37)$$

$$K = -M \frac{k_{1x}}{k_{2x}} \sin(b - \alpha) \exp k_{1x}a. \quad (35-38)$$

Equating the two values of K , we find that

$$\tan(b - \alpha) = -\frac{k_{2x}}{k_{1x}}. \quad (35-39)$$

Since both k 's are positive, the angle $b - \alpha$ is in either the second or the fourth quadrant. Also,

$$\alpha = \arctan \frac{k_{2x}}{k_{1x}} + b + m'\pi, \quad (35-40)$$

where m' is any integer. Setting $m' = 0$ gives

$$\alpha = \arctan \frac{k_{2x}}{k_{1x}} + b. \quad (35-41)$$

35.4.2 Continuity at the $x = a$ Interface

Because of the continuity of H_{mz} ,

$$M \cos(b + \alpha) = L \exp(-k_{3x}a), \quad L = M \cos(b + \alpha) \exp k_{3x}a. \quad (35-42)$$

Proceeding similarly for E_{my} ,

$$-\frac{j\omega\mu_0}{k_{2x}} M \sin(b + \alpha) = \frac{j\omega\mu_0}{k_{3x}} L \exp(-k_{3x}a), \quad (35-43)$$

$$L = -M \frac{k_{3x}}{k_{2x}} \sin(b + \alpha) \exp k_{3x}a. \quad (35-44)$$

Equating now the two expressions for L ,

$$\tan(b + \alpha) = -\frac{k_{2x}}{k_{3x}}. \quad (35-45)$$

Remember that all the k 's are real and positive. Then the angle $b + \alpha$ lies in either the second or the fourth quadrant, like $b - \alpha$.

Also,

$$\alpha = -\arctan \frac{k_{2x}}{k_{3x}} - b + m''\pi, \quad (35-46)$$

m'' being again an integer. We *cannot* set m'' arbitrarily equal to zero, now that we have set $m' = 0$.

35.5 THE FIELD COMPONENTS IN THE THREE MEDIA

Now the field components are as follows.

$$H_{mz} = M \cos(b + \alpha) \exp[k_{3x}(a - x)], \quad (35-47)$$

$$\text{Medium 3: } \left. \begin{aligned} E_{my} &= \frac{j\omega\mu_0}{k_{3x}} H_{mz}, \\ H_{mx} &= -\frac{k_z}{\omega\mu_0} E_{my} = -\frac{jk_z}{k_{3x}} H_{mz}. \end{aligned} \right\} \quad (35-48)$$

$$\left. \begin{aligned} H_{mz} &= M \cos(k_{2x}x + \alpha), \\ E_{my} &= -\frac{j\omega\mu_0}{k_{2x}} M \sin(k_{2x}x + \alpha), \end{aligned} \right\} \quad (35-50)$$

$$\text{Medium 2: } \left. \begin{aligned} H_{mx} &= -\frac{k_z}{\omega\mu_0} E_{my}. \end{aligned} \right\} \quad (35-51)$$

$$\left. \begin{aligned} H_{mz} &= M \cos(b + \alpha) \exp[k_{3x}(a - x)], \\ E_{my} &= \frac{j\omega\mu_0}{k_{3x}} H_{mz}, \\ H_{mx} &= -\frac{k_z}{\omega\mu_0} E_{my}. \end{aligned} \right\} \quad (35-52)$$

$$\text{Medium 1: } \left\{ \begin{array}{l} H_{mz} = M \cos(b - \alpha) \exp[k_{1x}(a + x)], \\ E_{my} = -\frac{j\omega\mu_0}{k_{1x}} H_{mz}, \end{array} \right. \quad (35-53)$$

$$\left\{ \begin{array}{l} H_{mz} = M \cos(b - \alpha) \exp[k_{1x}(a + x)], \\ E_{my} = -\frac{j\omega\mu_0}{k_{1x}} H_{mz}, \end{array} \right. \quad (35-54)$$

$$\left\{ \begin{array}{l} H_{mz} = M \cos(b - \alpha) \exp[k_{1x}(a + x)], \\ E_{my} = -\frac{k_z}{\omega\mu_0} E_{mx} = \frac{jk_z}{k_{1x}} H_{mz}. \end{array} \right. \quad (35-55)$$

We set M equal to a real number. The wave number k and m'' are now the only unknowns.

35.6 THE PHASE SHIFTS ON TOTAL REFLECTION

We shall need the phase shifts at the two interfaces in a moment.

Figure 35-4 shows a ray that zigzags down the guide. The angle of incidence is θ (the θ_I of Chap. 30). Also,

$$\frac{k_z}{k_2} = \sin \theta, \quad \frac{k_{2x}}{k_2} = \cos \theta. \quad (35-56)$$

Observe that $k_z < k_2$. Since $k = 1/\lambda$, $\lambda_z > \lambda_2$, where λ_z is the wavelength of the guided wave and λ_2 is the wavelength of a uniform plane wave in medium 2. Then the phase velocity of the guided wave is larger than c/n_2 .

We found the phase shift upon reflection in Sec. 31.2.1. Remember that the first medium is now medium 2, while the second is medium 3. Substituting $(n_2/n_3) \sin \theta$ for $\sin \theta_T$, the phase shift at the 2,3 interface is

$$\Phi_{2,3} = 2 \arctan \frac{[\sin^2 \theta - (n_3/n_2)^2]^{1/2}}{\cos \theta} = 2 \arctan \frac{[(k_z/k_2)^2 - (k_3/k_2)^2]^{1/2}}{k_{2x}/k_2} \quad (35-57)$$

$$= 2 \arctan \frac{k_{3x}}{k_{2x}}, \quad (35-58)$$

$$\tan \frac{\Phi_{2,3}}{2} = \frac{k_{3x}}{k_{2x}}. \quad (35-59)$$

Remember that the incident ray lies in medium 2, here. Similarly,

$$\Phi_{2,t} = 2 \arctan \frac{k_{1x}}{k_{2x}}, \quad \tan \frac{\Phi_{2,t}}{2} = \frac{k_{1x}}{k_{2x}}. \quad (35-60)$$

35.7 THE EIGENVALUE EQUATION

Equating the two expressions for α from Eqs. 35-41 and 35-46 gives

$$\arctan \frac{k_{2x}}{k_{1x}} + b = -\arctan \frac{k_{2x}}{k_{3x}} - b + m''\pi, \quad (35-61)$$

$$2b = -\arctan \frac{k_{2x}}{k_{1x}} - \arctan \frac{k_{2x}}{k_{3x}} + m''\pi. \quad (35-62)$$

But, from the definition of the tangent function,

$$\arctan \frac{k_{2x}}{k_{1x}} = \frac{\pi}{2} - \arctan \frac{k_{1x}}{k_{2x}}, \quad (35-63)$$

$$\arctan \frac{k_{2x}}{k_{3x}} = \frac{\pi}{2} - \arctan \frac{k_{3x}}{k_{2x}}. \quad (35-64)$$

Setting

$$m = m'' - 1, \quad (35-65)$$

where m is again an integer,

$$2b = \arctan \frac{k_{1x}}{k_{2x}} + \arctan \frac{k_{3x}}{k_{2x}} + m\pi. \quad (35-66)$$

Now

$$b = k_{2x}a = k_2 a \cos \theta = n_2 k_0 a \cos \theta = 2\pi \frac{a}{\lambda_2} \cos \theta, \quad (35-67)$$

where $\lambda_2 = \lambda_0/n_2$ is the wavelength of a uniform plane wave in medium 2. Substituting the Φ 's from Eqs. 35-58 and 35-60 into Eq. 35-66, and multiplying both sides by 2, we end up with the relation

$$8\pi \frac{a}{\lambda_2} \cos \theta = \Phi_{2,1} + \Phi_{2,3} + 2m\pi. \quad (35-68)$$

This is the *eigenvalue equation* for θ , the Φ 's being functions of θ . The angle of incidence θ of Fig. 35-4 can take only those values that satisfy the above equation, each integral value of m corresponding to an *eigenvalue* of θ and to a particular propagation mode. Thus m is the *mode order*. As we shall see in Sec. 36.1, $m \geq 0$. The various modes are denoted as TE_0 , TE_1 , TE_2 , etc.

We can also deduce the eigenvalue equation directly, as in Sec. 34.4. Figure 35-5 shows two wave fronts, A and C , of the up-going wave. At a

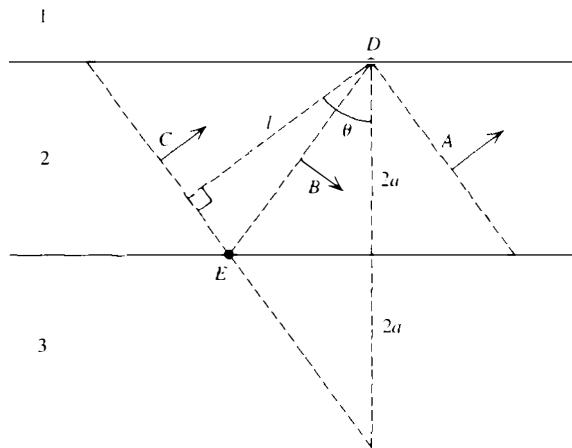


Fig. 35-5. Geometric construction for deducing the eigenvalue equation.

given time t , set $E_A = \exp j\omega t$ for that wave along A . At the same time t ,

$$E_C = \exp j\left(\omega t + \frac{2\pi l}{\lambda_2}\right). \quad (35-69)$$

At point D , reflection on the upper interface gives a phase shift of $\Phi_{2,3}$. Thus, along B , $E_B = \exp j(\omega t + \Phi_{2,3})$.

Similarly, as a result of the reflection at E ,

$$E_C = \exp j(\omega t + \Phi_{2,3} + \Phi_{2,1}) \quad (35-70)$$

and

$$\frac{2\pi l}{\lambda_2} = \Phi_{2,3} + \Phi_{2,1} + 2m\pi. \quad (35-71)$$

If $m = 0$, then l is minimum and the angle of incidence θ is maximum. As m increases to 1, 2, 3, ..., l becomes progressively larger and the angle of incidence decreases.

Now, from Fig. 35-5,

$$l = 4a \cos \theta. \quad (35-72)$$

Substituting into Eq. 35-71 gives the eigenvalue equation.

We shall solve the eigenvalue equation graphically in the next chapter. This will yield the values of θ for the various possible modes. Then the values of α and of the various k 's will follow.

Before going on to that, let us find a third expression for the phase angle α .

35.8 A THIRD EXPRESSION FOR α

Combining Eqs. 35-41, 35-63, 35-66, 35-59, and 35-60, we find that

$$\alpha = \arctan \frac{k_{2x}}{k_{1x}} + b \quad (35-73)$$

$$= \frac{\pi}{2} - \frac{\Phi_{2,1}}{2} + \frac{\Phi_{2,1}}{4} + \frac{\Phi_{2,3}}{4} + \frac{m\pi}{2} \quad (35-74)$$

$$= \frac{\Phi_{2,3}}{4} - \frac{\Phi_{2,1}}{4} + (m+1)\frac{\pi}{2}. \quad (35-75)$$

In a *symmetric guide*, medium 1 and medium 3 are identical, the Φ 's are equal, and

$$\alpha = (m+1)\frac{\pi}{2}. \quad (35-76)$$

35.9 SUMMARY

A *planar optical waveguide* comprises a *substrate*, a *sheet*, and a *cover*, with $n_1 < n_2$, $n_3 < n_2$, n being an index of refraction. Total reflection occurs at the interfaces 2,1 and 2,3.

We proceed as follows to find the field equations for a traveling wave. (1) We apply the theory of Sec. 33.1. (2) The field in medium 1 must decrease exponentially with $-x$, and the field in medium 3 exponentially with x . (3) In medium 2 we choose uniform plane waves that zigzag down the guide. (4) We apply the conditions of continuity for the tangential components of \mathbf{E} and \mathbf{H} at the interfaces.

This leads to Eqs. 35-47 to 35-55 for the field.

The factor M is a measure of the amplitude of the wave,

$$b = k_{2x}a, \quad (35-67)$$

$$\alpha = \arctan \frac{k_{2x}}{k_{1x}} + b = -\arctan \frac{k_{2x}}{k_{3x}} - b + m''\pi, \quad (35-41), (35-46)$$

where m'' is an integer,

$$k_{1x} = +(\mathbf{k}_z^2 - \mathbf{k}_1^2)^{1/2}, \quad k_{2x} = +(\mathbf{k}_2^2 - \mathbf{k}_z^2)^{1/2}, \quad k_{3x} = +(\mathbf{k}_z^2 - \mathbf{k}_3^2)^{1/2}, \\ (35-13), (35-21), (35-17)$$

$$k_1 = \frac{1}{\lambda_1} = \omega(\epsilon_r \epsilon_0 \mu_0)^{1/2}, \quad (35-4)$$

and similarly for k_2 and k_3 .

The constants k_{1x} , k_{2x} , k_{3x} , and α are all functions of k_z , which is itself related to the angle of incidence θ :

$$k_z = k_x \sin \theta. \quad (35-56)$$

Equating the two values of α leads to the *eigenvalue equation*

$$8\pi \frac{a}{\lambda_2} \cos \theta = \Phi_{2,1} + \Phi_{2,3} + 2m\pi \quad (35-68)$$

that fixes the values of θ for the various modes m . Here

$$\Phi_{2,1} = 2 \arctan \frac{k_{1x}}{k_{2x}} \quad \text{and} \quad \Phi_{2,3} = 2 \arctan \frac{k_{3x}}{k_{2x}} \quad (35-60), (35-58)$$

are the phase shifts upon total reflection.

The phase angle is

$$\alpha = \frac{\Phi_{2,3}}{4} - \frac{\Phi_{2,1}}{4} + (m+1) \frac{\pi}{2}. \quad (35-75)$$

PROBLEMS

35-1. (35.1) The numerical aperture of an optical fiber

Figure 35-6 shows a longitudinal section through an optical fiber. A ray emanating from a source S enters the fiber at an angle such that θ is the critical angle.

(a) Show that $\sin \phi = (n_2^2 - n_1^2)^{1/2}$. The quantity $\sin \phi$ is the *numerical aperture* (NA) of the fiber. This expression is valid only if $n_2^2 - n_1^2 \leq 1$. The maximum possible value of ϕ is 90° . If the angle ϕ increases beyond the value defined by the above equation, then θ becomes smaller than the critical angle and the ray does not propagate down the guide. This equation therefore defines an upper limit for ϕ .

(b) Show that, if the source radiates isotropically, then the fraction of the total available light that is collected by the fiber is about $(NA)^2/4$, or about $n_2 \Delta n/2$, where $\Delta n = n_2 - n_1$. If $n_2 = 2$ and $n_1 = 1.98$, then $\phi = 11.5^\circ$ and $F = 0.01$. The light collection efficiency is thus very low.

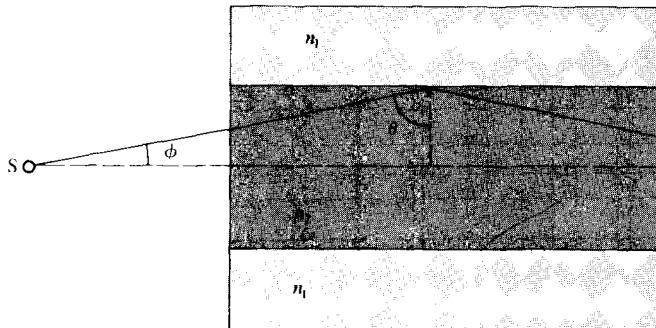


Fig. 35-6.

35-2. (35.3) The k 's

(a) Show that $k_{2x}^2 + k_{1x}^2 = (n_2^2 - n_1^2)k_0^2$.

(b) Show also that if, in a symmetric guide, $n_1 = n_3 = n$, $n_2 = n + \Delta n$, $\Delta n \ll n$, then $k_{2x}^2 + k_{1x}^2 \approx (80n\Delta n)/\lambda_0^2$.

35-3. (35.5) Multiple reflections in the sheet

Show that the E field in medium 2 is the superposition of an up-going and a down-going wave, at the correct angle of incidence.35-4. (35.7) $\tan 2k_{2x}a$ in terms of the k 's

Show that

$$\tan 2k_{2x}a = (-1)^m \frac{k_{2x}(k_{1x} + k_{3x})}{k_{2x}^2 - k_{1x}k_{3x}}.$$

Recall that $k_{2x} = k_2 \cos \theta$, where $\cos \theta \ll 1$ and $k_2 = n_2 k_0 = n_2/\lambda_0$.

35-5. (35.8) The field components in a symmetric optical guide

(a) Show that, for even modes in a symmetric guide,

$$\begin{aligned} E_{my} &= M' \cos b \exp [k_{3x}(a - x)] && \text{in medium 3,} \\ &= M' \cos k_{2x}x && \text{in medium 2,} \\ &= M' \cos b \exp [k_{1x}(a + x)] && \text{in medium 1,} \end{aligned}$$

where $M' = (-1)^{m/2+1} j \omega \mu_0 M / (k_{2x})$.

(b) Show that, for odd modes in a symmetric guide,

$$\begin{aligned} E_{mv} &= -M'' \sin b \exp [k_{3x}(a - x)] && \text{in medium 3,} \\ &= -M'' \sin k_{2x}x && \text{in medium 2,} \\ &= +M'' \sin b \exp [k_{1x}(a + x)] && \text{in medium 1,} \end{aligned}$$

where $M'' = (-1)^{(m+1)/2} j \omega \mu_0 M / (k_{2x})$.

CHAPTER 36

GUIDED WAVES IV

The Planar Optical Waveguide B. The Guided Wave

36.1 SOLVING THE EIGENVALUE EQUATION FOR θ	662
<i>Example: A SYMMETRIC GUIDE</i>	663
<i>Example: AN ASYMMETRIC GUIDE</i>	666
36.2 THE ELECTRIC FIELD STRENGTH E AS A FUNCTION OF x	666
36.3 LINES OF E AND OF H FOR THE $m = 0$ MODE IN THE SYMMETRIC GUIDE	666
36.4 THE PHASE VELOCITY v_p	668
36.5 THE GROUP VELOCITY v_g	670
36.6 THE FIELD ENERGY	670
36.7 THE TRANSMITTED POWER	671
36.8 SUMMARY	671
PROBLEMS	672

We terminate here our study of the planar optical waveguide and of guided waves. The next, and last, subject will be the radiation of electromagnetic waves.

You have probably noticed that the field of a planar optical waveguide is more complex than that of a hollow rectangular metallic guide. There are two reasons. First, the field extends over three different regions, instead of one. Second, the phase shifts at the interfaces are complex functions of the indices of refraction and of the angle of incidence.

We do not show the calculations in detail in this chapter, because they are rather involved.

36.1 SOLVING THE EIGENVALUE EQUATION FOR θ

In Sec. 35.7 we found that the angle of incidence θ of Fig. 35.4 satisfies the equation

$$8\pi \frac{a}{\lambda_2} \cos \theta = \Phi_{2,1} + \Phi_{2,3} + 2m\pi. \quad (36-1)$$

Here a is half the thickness of medium 2 (Fig. 35-4), λ_2 is the wavelength of a uniform plane wave in medium 2, the Φ 's are the phase shifts upon total reflection (Sec. 35.6),

$$\Phi_{2,1} = 2 \arctan \frac{(\sin^2 \theta - n_1^2/n_2^2)^{1/2}}{\cos \theta}, \quad (36-2)$$

$$\Phi_{2,3} = 2 \arctan \frac{(\sin^2 \theta - n_3^2/n_2^2)^{1/2}}{\cos \theta}, \quad (36-3)$$

and the *mode order* m is an integer. For each allowed value of m there corresponds one mode of propagation.

Equation 36-1 does not seem to possess an analytical solution. We shall solve it graphically, but let us first examine it carefully.

(1) Total reflection occurs if both

$$\theta > \arcsin \frac{n_1}{n_2} \quad \text{and} \quad \theta > \arcsin \frac{n_3}{n_2}. \quad (36-4)$$

(2) Since $\theta \leq \pi/2$, the left-hand side of Eq. 36-1 is positive.

(3) From Sec. 31.2.1, the phase shifts Φ , with E normal to the plane of incidence, lie between 0 and π . Thus

$$0 < \Phi_{2,1} + \Phi_{2,3} < 2\pi. \quad (36-5)$$

Then the integer m cannot be negative.

(4) The Φ 's increase with θ , while $\cos \theta$ decreases. Then for a given m there is only one possible θ .

(5) As m increases from 1 to 2 to 3, etc., the corresponding value of θ decreases.

We solve Eq. 36-1 below by plotting both sides as functions of θ on the same graph, and noting where the curves cross. That is less accurate, but more instructive, than a numerical solution. The values shown in Tables 36-1 and 36-2 were calculated numerically.

Example A SYMMETRIC GUIDE

Figure 36-1 shows a set of such curves for the symmetric ($n_1 = n_3$) guide of Table 36-1. The curves marked $m = 0, 1, 2, 3$ are curves of the right-hand side of Eq. 36-1. They start at

$$\theta \arcsin \frac{n_1}{n_2} = \arcsin \frac{n_3}{n_2}, \quad (36-6)$$

n_1 being equal to n_3 in this case.

Table 36-1 Parameters for a symmetric guide

CHARACTERISTICS	QUANTITY	MODE ORDER		
		0	1	2
$a = 1 \mu\text{m}$	$\theta^{\dagger \ddagger}$	1.5223	1.4752	1.4345
$n_1 = 1.9800$	$\sin \theta$	10.99882	0.99543	0.99073
$n_2 = 2.0000$	k_z	2.5103×10^7	2.5018×10^7	2.4900×10^7
$n_3 = 1.9800$	λ	2.5030×10^{-7}	2.5118×10^{-7}	2.5234×10^{-7}
$\lambda_0 = 500 \text{ nm}$	k_{1x}	3.3289×10^6	2.6106×10^6	9.5415×10^5
$k_0 = 1.2566 \times 10^7$	k_{2x}	1.2196×10^6	2.3989×10^6	3.4146×10^6
$k_1 = 2.4881 \times 10^7$	k_{3x}	3.3289×10^6	2.6106×10^6	9.5415×10^5
$k_2 = 2.5133 \times 10^7$	b^{\dagger}	1.2196	2.3989	3.4146
$k_3 = 2.4881 \times 10^7$	$\Phi_{2,1}^{\dagger \ddagger}$	2.4392	1.6553	0.54496
$\sin \theta_c = 0.99000$	$\Phi_{2,3}^{\dagger \ddagger}$	2.4392	1.6553	0.54496
$\theta_c = 81.890^\circ$	α^{\dagger}	$\pi/2 = 1.5708$	$\pi = 3.1416$	$3\pi/2 = 4.7124$
	$b - \alpha^{\dagger}$	-0.35120	-0.74269	-1.2978
	$b + \alpha^{\dagger}$	2.7904	5.5405	8.1270
	v_p	1.5007×10^8	1.5058×10^8	1.5130×10^8
	v_g	1.4980×10^8	1.4959×10^8	1.4992×10^8

[†] We found these values by solving the eigenvalue equation numerically.

[‡] radians

Table 36-2 Parameters for an asymmetric guide

CHARACTERISTICS	QUANTITY	MODE ORDER	
		0	1
$a = 1 \mu\text{m}$	θ^+	1.5172	1.4649
$n_1 = 1.9800$	$\sin \theta$	0.99856	0.99440
$n_2 = 2.0000$	k_z	2.5097×10^7	2.4992×10^7
$n_3 = 1.0000$	λ	2.5036×10^{-7}	2.5141×10^{-7}
$\lambda_0 = 500 \text{ nm}$	k_{1x}	3.2801×10^6	2.3471×10^6
$k_0 = 1.2566 \times 10^7$	k_{2x}	1.3458×10^6	2.6573×10^6
$k_1 = 2.4881 \times 10^7$	k_{3x}	2.1724×10^7	2.1603×10^7
$k_2 = 2.5133 \times 10^7$	b^+	1.3458	2.6573
$k_3 = 1.2566 \times 10^7$	$\Phi_{2,1}^+$	2.3629	1.4470
$\sin \theta_{c1} = 0.9900$	$\Phi_{2,3}^+$	3.0179	2.8968
$\theta_{c1} = 81.890^\circ$	α^+	1.7351	3.5046
$\theta_{c2} = 30.000^\circ$	$b - \alpha^+$	-0.3893	-0.8473
	$b + \alpha^+$	3.0809	6.1619

† radians

Curve A is a cosine function of amplitude $8\pi(a/\lambda_2)$, near $\pi/2$. Decreasing the ratio a/λ_2 sweeps A down in the direction of the arrow, its right-hand extremity remaining fixed [$\cos(\pi/2) = 0$]. This eliminates the higher modes ($m = 3, 2, 1$) one by one. However, mode $m = 0$ remains, even at long wavelengths.

Thus a very thin guide is monomode: it supports only the $m = 0$ mode. Then the angle of incidence is only slightly larger than the critical angle.

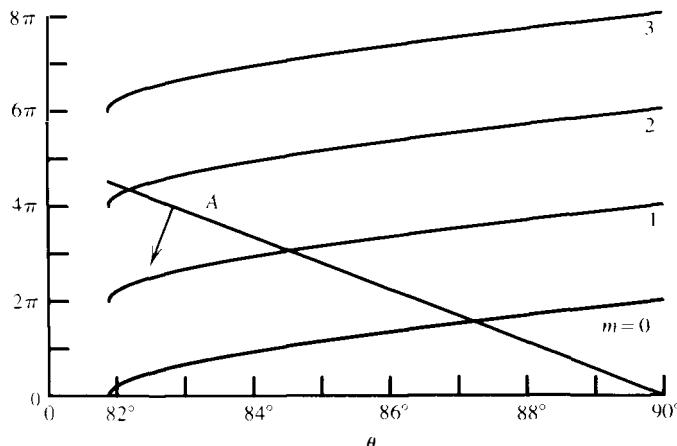


Fig. 36-1. Graphical solution of the eigenvalue equation for the symmetric guide of Table 36-1. Curve A is a plot of the left-hand side of the equation. The other curves are plots of the right-hand side for various values of the mode order m . The curves intersect at the values of θ shown in Table 36-1.

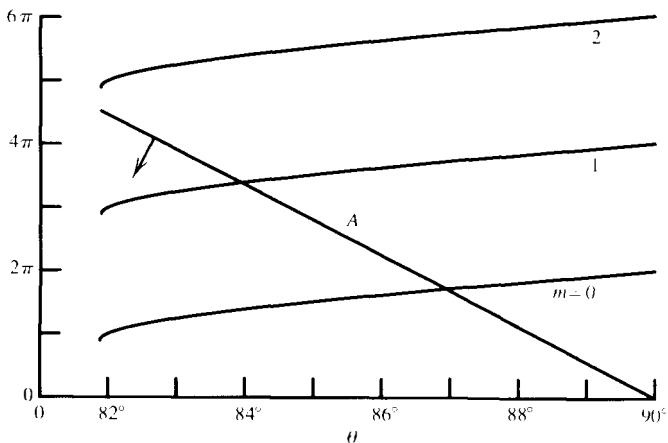


Fig. 36-2. Graphical solution of the eigenvalue equation for the asymmetric guide of Table 36-2.

Table 36-1 shows the guide parameters for the allowed modes $m = 0, 1, 2$.

Example

AN ASYMMETRIC GUIDE

Now let medium 3 be air. Then $n_3 = 1$ as in Table 36-2, with $n_2 = 2$ and $n_1 = 1.98$ as above. The critical angle at the 2,3 interface is $\arcsin 0.5$, or 0.52 radian, but total reflection at the 2,1 interface can only occur for $\theta \geq \arcsin 0.99$, or 1.43 radians. The allowed range for θ is therefore the same as for the symmetric guide.

Figure 36-2 shows the graphical solution for this guide. Only the $m = 0$ and $m = 1$ modes are allowed. Notice that, as the ratio a/λ_2 decreases, curve A swings along the arrow, as previously, and mode 0 eventually disappears. So an asymmetric guide filters out long wavelengths.

36.2 THE ELECTRIC FIELD STRENGTH E AS A FUNCTION OF x

Figure 36-3(a) to (e) shows E_{my} as a function of the vertical coordinate x for the five modes of Tables 36-1 and 36-2 and Figs. 36-1 and 36-2.

Recall that, for the TE modes that we are considering, \mathbf{E} is everywhere parallel to the y -axis of Fig. 35-2.

36.3 LINES OF E AND OF H FOR THE $m = 0$ MODE IN THE SYMMETRIC GUIDE

Let us see what the lines of \mathbf{E} and \mathbf{H} look like with the $m = 0$ mode in the symmetric guide. Starting with Eqs. 35-47 to 35-55, we set $\alpha = \pi/2$, from

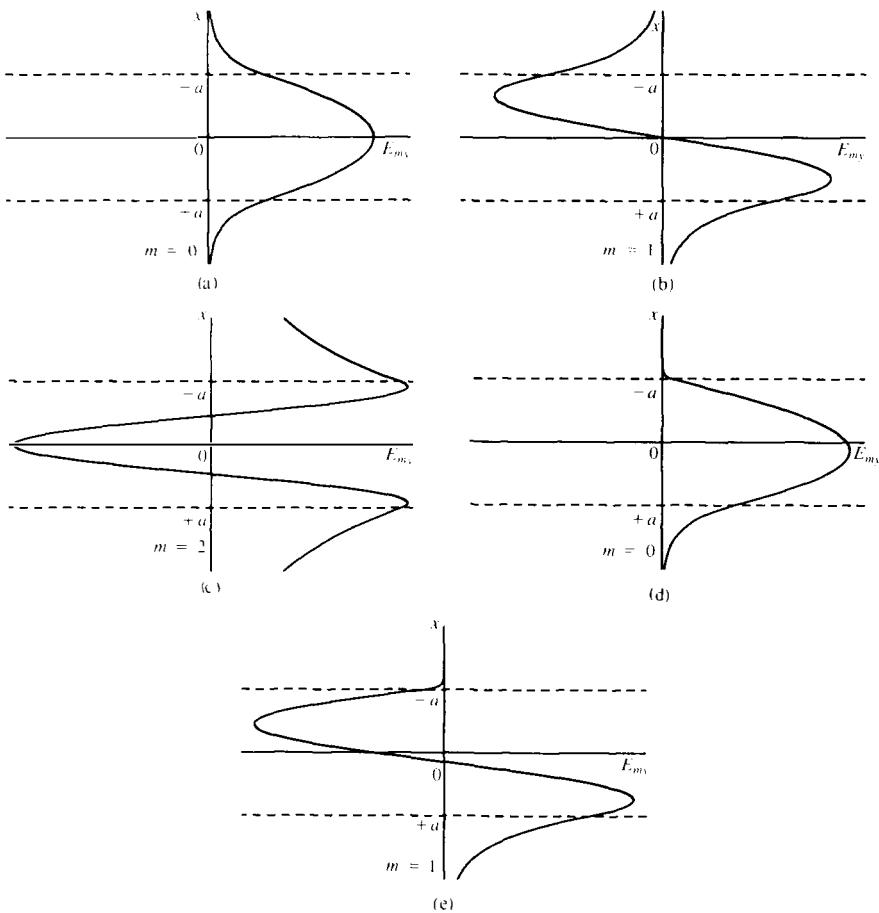


Fig. 36-3. (a), (b), (c) Curves of E_{my} as a function of x for the symmetric guide of Table 36-1. (d), (e) Curves of E_{my} for the asymmetric guide of Table 36-2.

Table 36-1. We set $M = j$. Then we multiply each equation by $\exp j(\omega t - k_z z)$ to obtain the phasors, take the real parts, and disregard constant factors, while preserving the signs. The field components are now as follows.

$$\text{Medium 3: } \begin{cases} H_z \propto \exp [k_{3x}(a-x)] \sin (\omega t - k_z z), \\ E_y \propto \exp [k_{3x}(a-x)] \cos (\omega t - k_z z), \\ H_x \propto -\exp [k_{3x}(a-x)] \cos (\omega t - k_z z). \end{cases} \quad (36-7)$$

$$\text{Medium 2: } \begin{cases} H_z \propto \sin k_{2x}x \sin (\omega t - k_z z), \\ E_y \propto \cos k_{2x}x \cos (\omega t - k_z z), \\ H_x \propto -\cos k_{2x}x \cos (\omega t - k_z z). \end{cases} \quad (36-8)$$

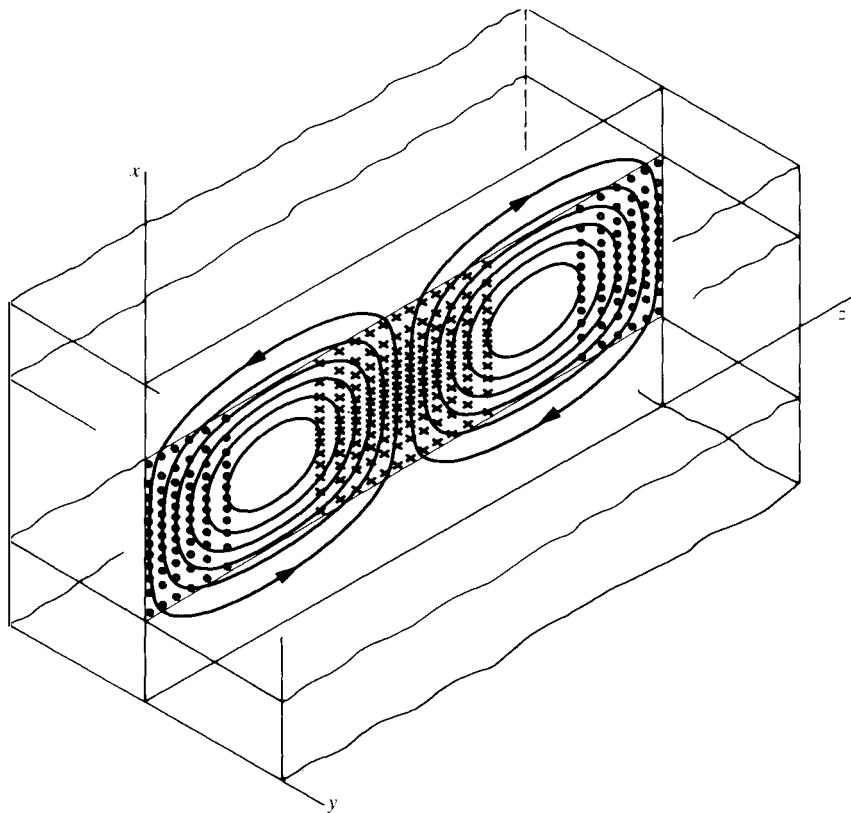


Fig. 36-4. Lines of \mathbf{E} (dots and crosses) and of \mathbf{H} for the symmetrical optical waveguide of Table 36-1. The wave travels from left to right. The figure does not show lines of \mathbf{E} in media 1 and 3 because the electric field there is much weaker than in medium 2. Compare with Fig. 34-4.

$$\text{Medium 1: } \begin{cases} H_z \propto -\exp [k_{1x}(a+x)] \sin (\omega t - k_z z), \\ E_y \propto \exp [k_{1x}(a+x)] \cos (\omega t - k_z z), \\ H_x \propto -\exp [k_{1x}(a+x)] \cos (\omega t - k_z z). \end{cases} \quad (36-9)$$

Figure 36-4 shows lines of \mathbf{E} and lines of \mathbf{H} . Compare with Fig. 34-4.

36.4 THE PHASE VELOCITY v_p

The wavelength of the guided wave is given by

$$\frac{1}{\lambda_z} = k_z = k_2 \sin \theta = n_2 k_0 \sin \theta = n_2 \frac{\omega}{c} \sin \theta, \quad (36-10)$$

so $\lambda_z = \lambda_0 / (n_2 \sin \theta) > \lambda_0$. The phase velocity is given by

$$v_p = f\lambda_z = \frac{c}{n_2 \sin \theta} > \frac{c}{n_2}. \quad (36-11)$$

For the $m = 0$ mode of the symmetric guide of Table 36-1, $\sin \theta = 0.9988$.

The *dispersion relation* for a wave is ω expressed as a function of k_z . The dispersion relation for a planar optical waveguide is hidden in the eigenvalue equation, 36-1. Its left-hand side is $4an_2(\omega/c) \cos \theta$, while its right-hand side is a function of θ , n_1 , n_2 , n_3 , and m . Thus, for a given mode in a given guide, Eq. 36-1 expresses ω as a function of θ in implicit form. But θ is itself a function of k_z and ω : from Eq. 35-56,

$$\sin \theta = \frac{k_z}{k_2} = \frac{k_z}{n_2(\omega/c)}. \quad (36-12)$$

So the eigenvalue equation gives ω as a function of k_z , and inversely.

To plot ω as a function of k_z we proceed as follows. We first select a value for m . Then, for each value of ω , we solve Eq. 36-1 for θ , and then we deduce k_z from the above equation.[†]

For the symmetric guide of Table 36-1,

$$\omega = 2\pi f = \frac{2\pi c}{\lambda_0}, \quad (36-13)$$

$$\beta = k_z = n_2 \frac{\omega}{c} \sin \theta = n_2 \frac{2\pi}{\lambda_0} \sin \theta. \quad (36-14)$$

Now θ can vary only between θ_c and 90° , with $\sin \theta_c = 1.98/2 = 0.99$. So, throughout the permissible range of θ , $\sin \theta$ varies at most between 0.99 and 1.0, and

$$\frac{\omega}{\beta} \approx \frac{2\pi c}{\lambda_0} \frac{\lambda_0}{4\pi} = \frac{c}{2}. \quad (36-15)$$

The phase velocity of the guided wave is approximately the same as that of a uniform plane wave in medium 2.

For this symmetric guide, the curve of ω as a function of k_z is very nearly a straight line through the origin with a slope of $c/n_2 = c/2$.

[†]The wave number k_z is the β of Chap. 28 and the curve of ω as a function of β is then referred to as an ω - β diagram.

36.5 THE GROUP VELOCITY v_g

The group velocity (Sec. 29.2.4 and App. C) is defined as

$$v_g = \frac{d\omega}{dk_z} = \frac{1}{dk_z/d\omega}. \quad (36-16)$$

This is the slope of the curve of ω as a function of k_z . It is shown in Prob. 36.5 that,[†] for a symmetric guide with $n_3 = n_1$,

$$v_g = \frac{c}{n_2} \sin \theta \frac{(\sin^2 \theta - n_1^2/n_2^2)^{1/2} + (\lambda_2/a)}{(\sin^2 \theta - n_1^2/n_2^2)^{1/2} + (\lambda_2/a) \sin^2 \theta}. \quad (36-17)$$

For the symmetric guide of Table 36-1,

$$v_g \approx v_p \approx \frac{c}{2}. \quad (36-18)$$

36.6 THE FIELD ENERGY

We now calculate the time-averaged electromagnetic field energy in the guide per unit length and per unit width:

$$\mathcal{E}' = \frac{1}{2} \int_{-\infty}^{+\infty} (\epsilon_r \epsilon_0 E_{\text{rms}}^2 + \mu_0 H_{\text{rms}}^2) 1 \times 1 \times dx \quad (36-19)$$

with

$$E_{\text{rms}}^2 = \frac{1}{2} |E_{my}|^2, \quad H_{\text{rms}}^2 = \frac{1}{2} (|H_{mx}|^2 + |H_{mz}|^2). \quad (36-20)$$

The average energies per unit length and per unit width in the three media are, respectively,

$$\mathcal{E}'_1 = \frac{\mu_0 M^2 k_z^2}{4k_{1z}(2, 1)}, \quad (36-21)$$

$$\mathcal{E}'_2 = \frac{\mu_0 M^2}{4k_{2x}^2} \left[2ak_z^2 + \frac{k_{1x} k_x^2}{(2, 1)} + \frac{k_{3x} k_z^2}{(2, 3)} \right], \quad (36-22)$$

$$\mathcal{E}'_3 = \frac{\mu_0 M^2 k_z^2}{4k_{3x}(2, 3)}. \quad (36-23)$$

[†]When we discussed the group, or signal, velocity in a rectangular metallic waveguide, we saw that v_g was simply the axial component of an individual plane wave that zigzags down the guide. In that case, $v_g = (c/n_2) \sin \theta$. Here the group velocity is slightly larger than this axial component, because of the Goos-Haenchen shift (Sec. 31.2.2). A given ray travels part of the time in media 1 and 3. This effect *increases* when θ *decreases* or when the mode order *increases*.

We have set

$$(1, 2) = k_{1x}^2 + k_{2x}^2, \quad (2, 3) = k_{2x}^2 + k_{3x}^2 \quad (36-24)$$

and we have assumed that M is a real number.

The total time-averaged energy per unit length and per unit width is thus

$$\mathcal{E}' = \mathcal{E}'_1 + \mathcal{E}'_2 + \mathcal{E}'_3 \quad (36-25)$$

$$= \frac{\mu_0 M^2}{4} \left[\frac{k_z^2}{k_{2x}^2} \left(\frac{1}{k_{1x}} + \frac{1}{k_{3x}} + 2a \right) + 2a \right]. \quad (36-26)$$

This energy is half electric and half magnetic, with a slight excess of magnetic energy in media 1 and 3 and a slight deficiency in medium 2. For the symmetric guide of Table 36-1 and for the mode $m=0$, $\mathcal{E}'_1 = \mathcal{E}'_3 = 0.0136\mathcal{E}'_2$.

36.7 THE TRANSMITTED POWER

To obtain the power transmitted per unit width, we integrate the time-averaged Poynting vector over a width of 1 meter:

$$P' = \frac{1}{2} \operatorname{Re} \int_{-\infty}^{+\infty} -E_y H_x^* dx. \quad (36-27)$$

Integrating and simplifying as in Prob. 36-8, the values of P' for the three media are, respectively,

$$P'_1 = \frac{\omega \mu_0 k_z M^2}{4k_{1x}(2, 1)}, \quad (36-28)$$

$$P'_2 = \frac{\omega \mu_0 k_z M^2}{4k_{2x}^2} \left(2a + \frac{k_{1x}}{(2, 1)} + \frac{k_{3x}}{(2, 3)} \right), \quad (36-29)$$

$$P'_3 = \frac{\omega \mu_0 k_z M^2}{4k_{3x}(2, 3)}. \quad (36-30)$$

Then

$$P' = P'_1 + P'_2 + P'_3 = \frac{\omega \mu_0 k_z M^2}{4k_{2x}^2} \left(\frac{1}{k_{1x}} + \frac{1}{k_{3x}} + 2a \right). \quad (36-31)$$

36.8 SUMMARY

We can solve the eigenvalue equation for θ , Eq. 36-1, by plotting the two sides separately as functions of θ and noting where the curves cross.

With a symmetric guide whose parameters are those of Table 36-1, we have the curves of Figs. 36-1 and 36-3(a), (b), (c).

With the asymmetric guide of Table 36-2 we have Figs. 36-2 and 36-3(d) and (e).

The *lines of E and of H* (Fig. 36-4) are reminiscent of those for a hollow rectangular metallic guide (Fig. 34-4).

The *phase velocity*

$$v_p = \frac{c}{n_2 \sin \theta} \quad (36-11)$$

is only slightly larger than that of a uniform plane wave in medium 2.

The *dispersion relation* is the circular frequency ω expressed as a function of k_z .

The *group velocity* is

$$v_g = \frac{d\omega}{dk_z} = \frac{c}{n_2} \sin \theta \frac{(\sin^2 \theta - n_1^2/n_2^2)^{1/2} + (\lambda_2/a)}{(\sin^2 \theta - n_1^2/n_2^2)^{1/2} + (\lambda_2/a) \sin^2 \theta} \quad (36-17)$$

for a symmetric guide.

The *field energy* per unit length and per unit width is

$$\mathcal{E}' = \frac{\mu_0 M^2}{4} \left[\frac{k_z^2}{k_{2x}^2} \left(\frac{1}{k_{1x}} + \frac{1}{k_{3x}} + 2a \right) + 2a \right], \quad (36-26)$$

and the *transmitted power* per unit width is

$$P' = \frac{\omega \mu_0 k_z M^2}{4k_{2x}^2} \left(\frac{1}{k_{1x}} + \frac{1}{k_{3x}} + 2a \right). \quad (36-31)$$

PROBLEMS

36-1. (36.1) Negative mode orders are forbidden

In Fig. 36-1, curves for $m = -2, -3, -4, \dots$ would intersect curve A at angles of incidence larger than 90° , which is absurd. So those modes are forbidden. But for mode $m = -1$ the curves would intersect at $\theta = \pi/2$, which is sensible.

We have shown that all negative values of m are forbidden. Show in a different way that the mode $m = -1$ is forbidden.

36-2. (36.1) The maximum value of the free-space wavelength λ_0 as a function of the mode order m

Find the maximum permissible value of λ_0 as a function of m for a symmetric optical waveguide.

36-3. (36.1) Modal dispersion and the numerical aperture in a symmetric guide

In multimode propagation, each mode has its own group velocity. Then a

narrow light pulse broadens as it travels down the guide. This is *modal dispersion*. We saw in Prob. 35-1 that the numerical aperture is only of the order of 1% for a symmetric optical guide with n_1 and n_3 slightly smaller than n_2 . How would the numerical aperture and the modal dispersion for the guide of Table 36-1 be affected if media 1 and 3 were both air?

- 36-4. (36.4) The phase velocity v_p .

Let medium 1 be denser than medium 3.

Show that, if θ is only slightly larger than the critical angle at the interface 2,1, then the phase velocity of the guided wave is approximately equal to that of a uniform plane wave traveling in medium 1.

- 36-5. (36.5) The group velocity v_g .

Show that the group velocity in a symmetric planar optical waveguide is

$$v_g = \frac{c}{n_2} \sin \theta \frac{A^{1/2} + (\lambda_z/a)}{A^{1/2} + (\lambda_z/a) \sin^2 \theta} \approx \frac{c}{n_2 \sin \theta} = v_p,$$

where $A = \sin^2 \theta - n_1^2/n_2^2$.

Show that the approximate value applies to the mode $m = 1$ of Table 36-1. You can find the value of $d\theta/d\omega$ by differentiating the eigenvalue equation with respect to ω . Thus

$$v_p v_g = \frac{c^2}{n_2^2 A^{1/2}} \frac{A^{1/2} + (\lambda_z/a)}{A^{1/2} + (\lambda_z/a) \sin^2 \theta}.$$

- 36-6. (36.6) The field energy

(a) Show that, in a planar optical waveguide, the electromagnetic energy per unit width and per unit length, in medium 1, comprises three terms:

$$\mathcal{E}'_1 = \frac{\mu_0 M^2 \cos^2(b - \alpha)}{8k_{1x}^3} (k_1^2 + k_{1x}^2 + k_x^2).$$

The first term is the electric energy; the second is the longitudinal magnetic energy, or the magnetic energy associated with the longitudinal component of \mathbf{H} ; and the third term is the transverse magnetic energy. From Sec. 35.3, $k_1^2 = k_z^2 - k_{1x}^2$. Thus there is more magnetic energy than electric energy in medium 1.

(b) Show that $\mathcal{E}'_1 = \mu_0 M^2 k_z^2 / [4k_{1x}(2, 1)]$, where $(2, 1)$ is defined in Sec. 36.6. In medium 3, by symmetry, $\mathcal{E}'_3 = \mu_0 M^2 k_z^2 / [4k_{3x}(2, 3)]$. Here also there is more magnetic than electric energy.

(c) Show that

$$\mathcal{E}'_2 = \frac{\mu_0 M^2}{4k_{2x}^2} \left[2ak_z^2 + \frac{k_{1x}k_z^2}{(2, 1)} + \frac{k_{3x}k_z^2}{(2, 3)} \right].$$

(d) Show that, for the three media together and for a symmetric guide,

$$\mathcal{E}' = \frac{\mu_0 M^2}{2} \left[\left(\frac{k_z}{k_{2x}} \right)^2 \left(\frac{1}{k_{1x}} + a \right) + a \right].$$

As one would expect, $\mathcal{E}'_r = \mathcal{E}'_m$.

36-7. (36.7) The power density and the electric field strength in an optical guide

One author states that he has transmitted a 150-milliwatt signal through an optical guide with a cross section of 3 by 5 micrometers.

(a) Calculate the space- and time-averaged value of the Poynting vector.

(b) Calculate the peak electric field strength. This is the breakdown field.

The index of refraction is 1.5.

36-8. (36.7) Power transmission in the three media

(a) Show that the powers transmitted per meter of width in media 1 and 2 are, respectively,

$$P'_1 = \frac{\omega\mu_0 k_z}{4k_{1x}(2, 1)} M^2, \quad P'_2 = \frac{\omega\mu_0 k_z}{4k_{2x}^2} \left[2a + \frac{k_{1x}}{(2, 1)} + \frac{k_{3x}}{(2, 3)} \right] M^2,$$

where $(2, 1)$ and $(2, 3)$ are defined in Sec. 36.6. By symmetry, the power transmitted in 3 is

$$P'_3 = \frac{\omega\mu_0 k_z}{4k_{3x}(2, 3)} M^2.$$

(c) Calculate the ratio P'_1/P'_2 for the symmetric guide of Table 36-1 for modes 0, 1, and 2.

36-9. (36.7) The energy transport velocity

By definition, the *energy transport velocity* is equal to the ratio P'/\mathcal{E}' , where P' is the transmitted power per meter of width, and \mathcal{E}' is the electromagnetic energy per meter of width and per meter of length.

Show that the energy transport velocity for a symmetrical guide is equal to the group velocity given in Prob. 36-5.

CHAPTER 37

RADIATION I

The Potentials V and \mathbf{A} , and the Fields \mathbf{E} and \mathbf{B}

37.1 THE LORENTZ CONDITION	676
37.2 THE NONHOMOGENEOUS WAVE EQUATION FOR V	678
37.3 THE NONHOMOGENEOUS WAVE EQUATION FOR \mathbf{A}	678
37.4 THE RETARDED POTENTIALS	680
<i>Example: THE RETARDED POTENTIALS FOR THE OSCILLATING ELECTRIC DIPOLE</i>	681
<i>Example: THE RETARDED POTENTIALS FOR THE OSCILLATING MAGNETIC DIPOLE</i>	684
37.5 A SECOND PAIR OF INTEGRALS FOR \mathbf{E} AND \mathbf{B}	686
37.6 A THIRD PAIR OF INTEGRALS FOR \mathbf{E} AND \mathbf{B}	687
37.7 SUMMARY	688
PROBLEMS	689

We have studied the propagation of electromagnetic waves in considerable detail. In Chaps. 28 through 36 we studied successively their propagation in free space and in various media, across an interface, and then along various guiding structures.

Our final topic, the radiation of electromagnetic fields, will occupy us during the next three chapters.

Here we turn to general considerations on the fields of time-dependent sources. Until now we have disregarded the time taken by the field to propagate from the source of radiation to the point of observation. This is legitimate only if the time delay is a small fraction of one period.

We shall obtain the correct integrals for the retarded field from the nonhomogeneous wave equations for \mathbf{E} and \mathbf{B} . These equations are similar to the usual wave equations, except that they include a source term.

37.1 THE LORENTZ CONDITION

In Secs. 9.4 and 20.3 we found that

$$V = \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\rho}{r} dv', \quad \mathbf{A} = \frac{\mu_0}{4\pi} \int_{v'} \frac{\mathbf{J}}{r} dv', \quad (37-1)$$

where V and \mathbf{A} are, respectively, the scalar and vector potentials at the field point $P(x, y, z)$, $\rho = \rho_f + \rho_b$ is the total charge density at the source point $P'(x', y', z')$, dv' is the element of volume $dx' dy' dz'$ at P' , r is the distance between P and P' , and $\mathbf{J} = \mathbf{J}_f + \partial\mathbf{P}/\partial t + \nabla \times \mathbf{M}$ is the total current density at P' . The volume v' encloses all the charges and all the currents.

Now ρ and \mathbf{J} are not independent quantities because they satisfy the equation for the conservation of charge (Secs. 4.2, 16.4, and 27.5):

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t}. \quad (37-2)$$

Let us therefore seek an equation linking V and \mathbf{A} . We shall need this equation in the next sections. We have already found this equation in Sec. 17.9 for the field of a point charge moving at a constant velocity.

We start with the divergence of \mathbf{A} :

$$\nabla \cdot \mathbf{A} = \frac{\mu_0}{4\pi} \nabla \cdot \int_{v'} \frac{\mathbf{J}}{r} dv' = \frac{\mu_0}{4\pi} \int_{v'} \nabla \cdot \frac{\mathbf{J}}{r} dv'. \quad (37-3)$$

We can invert the order of the ∇ and of the integral sign because the former operates on x, y, z and the latter on x', y', z' . Now

$$\nabla \cdot \frac{\mathbf{J}}{r} = \frac{1}{r} \nabla \cdot \mathbf{J} + \mathbf{J} \cdot \nabla \frac{1}{r}. \quad (37-4)$$

The first term on the right is zero because \mathbf{J} is a function of x', y', z' and not of the coordinates x, y, z of the field point. Also, from identities 15 and 16 from the inside of the front cover, $\nabla(1/r) = -\nabla'(1/r)$. Thus

$$\nabla \cdot \frac{\mathbf{J}}{r} = -\mathbf{J} \cdot \nabla' \frac{1}{r}. \quad (37-5)$$

Also,

$$\nabla' \cdot \frac{\mathbf{J}}{r} = \frac{1}{r} \nabla' \cdot \mathbf{J} + \mathbf{J} \cdot \nabla' \frac{1}{r}. \quad (37-6)$$

Adding these last two equations and rearranging, we find that

$$\nabla \cdot \frac{\mathbf{J}}{r} = \frac{1}{r} \nabla' \cdot \mathbf{J} - \nabla' \cdot \frac{\mathbf{J}}{r} \quad (37-7)$$

and

$$\nabla \cdot \mathbf{A} = \frac{\mu_0}{4\pi} \int_{v'} \frac{1}{r} \nabla' \cdot \mathbf{J} dv' - \frac{\mu_0}{4\pi} \int_{v'} \nabla' \cdot \frac{\mathbf{J}}{r} dv'. \quad (37-8)$$

The second integral is identical to the one that we started from, except that now we have the divergence of \mathbf{J}/r at the source point P' instead of at P . This second integral is zero for the following reason. According to the divergence theorem, it is equal to the integral of $(\mathbf{J}/r) \cdot d\mathcal{A}$ over the surface bounding v' . Now, by definition, v' encloses all the currents. So, on the surface of v' , \mathbf{J} is either zero or tangential, and $(\mathbf{J}/r) \cdot d\mathcal{A} = 0$. Then

$$\nabla \cdot \mathbf{A} = \frac{\mu_0}{4\pi} \int_{v'} \frac{1}{r} \nabla' \cdot \mathbf{J} dv'. \quad (37-9)$$

Here we used a prime on the ∇ that appears under the integral sign because we needed to distinguish between ∇ and ∇' . But $\nabla' \cdot \mathbf{J}$ is really the same as the $\nabla \cdot \mathbf{J}$ that appears in Eq. 37-2. In this latter equation we were concerned with only the source point, and a prime on the ∇ would have been superfluous. Then, applying the law of conservation of charge,

$$\nabla \cdot \mathbf{A} = -\frac{\mu_0}{4\pi} \int_{v'} \frac{\partial \rho / \partial t}{r} dv'. \quad (37-10)$$

Since the distance r between P and P' is not a function of the time, we can remove the time derivative from under the integral sign, and

$$\nabla \cdot \mathbf{A} = -\frac{\mu_0}{4\pi} \frac{\partial}{\partial t} \int_{v'} \frac{\rho}{r} dv' = -\epsilon_0 \mu_0 \frac{\partial V}{\partial t}, \quad (37-11)$$

or

$$\nabla \cdot \mathbf{A} + \epsilon_0 \mu_0 \frac{\partial V}{\partial t} = 0. \quad (37-12)$$

This is the *Lorentz condition*. It is an *identity* with V and \mathbf{A} defined as in Eq. 37-1. The Lorentz condition is a consequence of the conservation of charge.

This result is general, but our proof is not because, as we shall see, the above integrals for V and \mathbf{A} are valid only for slowly varying fields. The Lorentz condition has practical importance. Of course, the six components of \mathbf{E} and \mathbf{B} derive from the four components of V and \mathbf{A} . However, with the Lorentz condition, and if V is a function of t , one

need only know the *three* components of \mathbf{A} to find V and the six components of \mathbf{E} and \mathbf{B} .

For example, one can calculate \mathbf{E} and \mathbf{H} in the field of an antenna solely from \mathbf{A} , hence solely from the currents in the antenna.

37.2 THE NONHOMOGENEOUS WAVE EQUATION FOR V

We can now find the wave equation for V in just a few lines. If ρ is the total charge density $\rho_f + \rho_b$, then

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad (37-13)$$

$$\nabla \cdot \left(-\nabla V - \frac{\partial \mathbf{A}}{\partial t} \right) = \frac{\rho}{\epsilon_0}, \quad (37-14)$$

$$\nabla^2 V + \frac{\partial}{\partial t} \nabla \cdot \mathbf{A} = -\frac{\rho}{\epsilon_0}, \quad (37-15)$$

$$\nabla^2 V - \epsilon_0 \mu_0 \frac{\partial^2 V}{\partial t^2} = -\frac{\rho}{\epsilon_0}. \quad (37-16)$$

This is the *nonhomogeneous wave equation* for V .

At points where $\rho = 0$,

$$\nabla^2 V - \epsilon_0 \mu_0 \frac{\partial^2 V}{\partial t^2} = 0, \quad (37-17)$$

which is the usual *homogeneous wave equation* (App. C).

The phase velocity is

$$c = \frac{1}{(\epsilon_0 \mu_0)^{1/2}}, \quad (37-18)$$

as for \mathbf{E} and \mathbf{B} (Sec. 27.8).

If V is constant, then

$$\nabla^2 V = -\frac{\rho}{\epsilon_0}, \quad (37-19)$$

as in Sec. 4.1.

37.3 THE NONHOMOGENEOUS WAVE EQUATION FOR \mathbf{A}

We can find the corresponding equation for \mathbf{A} just as easily, starting this time from the Maxwell equation for the curl of \mathbf{B} :

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t}. \quad (37-20)$$

Here

$$\mathbf{J} = \mathbf{J}_f + \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M}, \quad (37-21)$$

as in Sec. 27.1. Thus

$$\nabla \times (\nabla \times \mathbf{A}) = \mu_0 \mathbf{J} + \epsilon_0 \mu_0 \frac{\partial}{\partial t} \left(-\nabla V - \frac{\partial \mathbf{A}}{\partial t} \right), \quad (37-22)$$

$$\nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} = \mu_0 \mathbf{J} - \epsilon_0 \mu_0 \left(\nabla \frac{\partial V}{\partial t} + \frac{\partial^2 \mathbf{A}}{\partial t^2} \right), \quad (37-23)$$

$$\nabla \left(-\epsilon_0 \mu_0 \frac{\partial V}{\partial t} \right) - \nabla^2 \mathbf{A} = \mu_0 \mathbf{J} - \epsilon_0 \mu_0 \nabla \frac{\partial V}{\partial t} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{A}}{\partial t^2}, \quad (37-24)$$

$$\nabla^2 \mathbf{A} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{J}. \quad (37-25)$$

We used the vector identity for the Laplacian of a vector from the inside of the front cover and the Lorentz condition of Sec. 37.1. This is the *nonhomogeneous wave equation* for \mathbf{A} .

Outside current distributions, we have the *homogeneous wave equation*

$$\nabla^2 \mathbf{A} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} = 0. \quad (37-26)$$

For static fields, the homogeneous wave equation reduces to

$$\nabla^2 \mathbf{A} = 0, \quad (37-27)$$

as in Sec. 19.2.

Observe that we deduced the existence of the potentials V and \mathbf{A} , defined by

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t} \quad \text{and} \quad \mathbf{B} = \nabla \times \mathbf{A}, \quad (37-28)$$

from Maxwell's equations

$$\nabla \cdot \mathbf{B} = 0 \quad \text{and} \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \quad (37-29)$$

Then we deduced the nonhomogeneous wave equations for V and \mathbf{A} from the other two equations of Maxwell,

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad \text{and} \quad \nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t}. \quad (37-30)$$

37.4 THE RETARDED POTENTIALS

The integrals for V and \mathbf{A} of Eq. 37-1 do not take into account the finite velocity of propagation of electric and magnetic fields. For example, if the charge distribution shifts in one region, then the integrals imply that V and \mathbf{A} change simultaneously throughout all space.

In fact, the potentials at a given point and at a given instant do not correspond to the charge and current distributions at that instant, unless the charges are all fixed in space. The analogy with astronomy is obvious: we cannot see a star as it is now, but only as it was millions or billions of years ago. The star is not even at the place where we see it!

The correct integral for V is the solution of the nonhomogeneous wave equation 37-16 for V . This is

$$V(x, y, z, t) = \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\rho(x', y', z', t - r/c)}{r} dv'. \quad (37-31)$$

This integral is identical to that of Eq. 37-1, except that ρ is the electric charge density at the previous time $t - r/c$, the interval r/c being the time taken by a wave of speed c to travel the distance r . This is the *retarded scalar potential*.

Notice that the speed that we have used is c , the speed of light in a vacuum, and not the speed in the medium of propagation. This comes about as follows. As we shall see in Chap. 38, electromagnetic radiation occurs whenever electric charges accelerate. An electromagnetic wave travels through space at the speed c everywhere. Upon passing through, say, a piece of glass, the fluctuating electric field of the wave polarizes the molecules, which radiate in turn in all directions. The sum of all these waves is a single wave that travels at a speed different from c . If we express V and \mathbf{A} in Eq. 37-1, with

$$\rho = \rho_f + \rho_b \quad \text{and} \quad \mathbf{J} = \mathbf{J}_f + \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M} \quad (37-32)$$

under the integral sign, we consider, in effect, that the medium of propagation itself acts as a source in a vacuum. Then the retardation involves the speed c .

Similarly,

$$\mathbf{A}(x, y, z, t) = \frac{\mu_0}{4\pi} \int_{v'} \frac{\mathbf{J}(x', y', z', t - r/c)}{r} dv'. \quad (37-33)$$

The retarded potentials satisfy the Lorentz condition.

Fortunately, retardation effects are often negligible. Retardation is important whenever the delay is an appreciable fraction of a period $T = 1/f$.

For an infinite, homogeneous, isotropic, linear, and stationary (HILS) medium ϵ, μ ,

$$V(x, y, z, t) = \frac{1}{4\pi\epsilon} \int_{v'} \frac{\rho_f(x', y', z', t - r/v)}{r} dv, \quad (37-34)$$

$$\mathbf{A}(x, y, z, t) = \frac{\mu}{4\pi} \int_{v'} \frac{\mathbf{J}_f(x', y', z', t - r/v)}{r} dv', \quad (37-35)$$

where v is the speed of propagation $1/(\epsilon\mu)^{1/2}$ (Sec. 28.2) in the medium. This is another application of the rule of Sec. 27.1.

Example

THE RETARDED POTENTIALS FOR THE OSCILLATING ELECTRIC DIPOLE

Figure 37-1 shows an electric dipole similar to that of Fig. 5-1, except that now the electric dipole moment \mathbf{p} is a sinusoidal function of t :

$$Q = Q_m \exp j\omega t, \quad (37-36)$$

$$\mathbf{p} = Q_m \mathbf{s} \exp j\omega t = \mathbf{p}_m \exp j\omega t, \quad \mathbf{p}_m = Q_m \mathbf{s}. \quad (37-37)$$

This simple-minded model serves to calculate the radiation field of any charge distribution whose dipole moment is a sinusoidal function of the time.

The upward current through the wire is

$$I = \frac{dQ}{dt} = j\omega Q_m \exp j\omega t = I_m \exp j\omega t. \quad (37-38)$$

Thus

$$I_m = j\omega Q_m, \quad I_m \mathbf{s} = j\omega \mathbf{p}_m, \quad I \mathbf{s} = j\omega \mathbf{p}. \quad (37-39)$$

We could first calculate \mathbf{A} and then deduce V from the Lorentz condition (Sec. 37.1). However, it will be more instructive to calculate V separately. Refer to Fig. 5-1.

First

$$V = \frac{Q_m \exp j\omega(t - r_b/c)}{4\pi\epsilon_0 r_b} - \frac{Q_m \exp j\omega(t - r_a/c)}{4\pi\epsilon_0 r_a}. \quad (37-40)$$

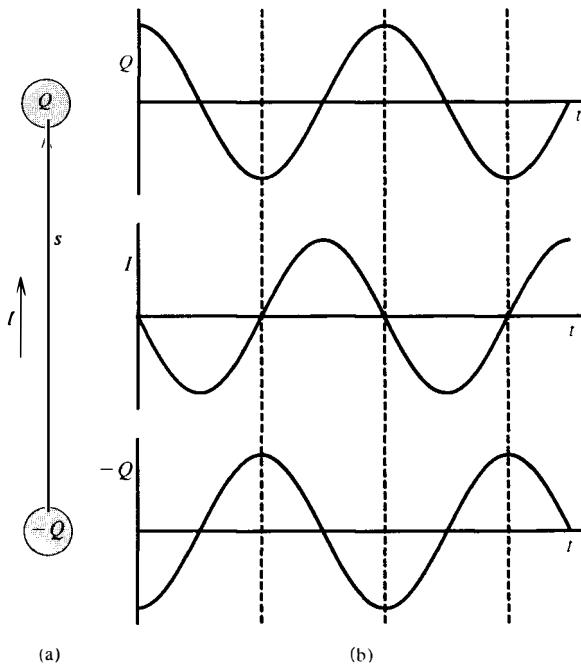


Fig. 37-1. An oscillating electric dipole. (a) The vector s points in the direction shown. (b) The charges and currents as functions of the time.

where the numerators are the charges, as they appear at P of Fig. 5-1 at the time t . Notice that the two components of V differ both in amplitude and in phase.

One can usually set $s^3 \ll r^3$ and $s^3 \ll \lambda^3$. This legitimates a simpler expression for V . Setting

$$r_a \approx r + \frac{s}{2} \cos \theta, \quad r_b \approx r - \frac{s}{2} \cos \theta, \quad (37-41)$$

then

$$\omega\left(t - \frac{r_b}{c}\right) \approx \omega\left(t - \frac{r}{c} + \frac{s \cos \theta}{2c}\right) \approx \omega[t] + \frac{s}{\lambda} \cos \theta, \quad (37-42)$$

where

$$[t] = t - \frac{r}{c} \quad \text{and} \quad \lambda = \frac{\lambda}{2\pi} = \frac{c}{\omega}. \quad (37-43)$$

A similar expression applies to $\omega(t - r_a/c)$, with a negative sign before the cosine term.

From now on we reserve brackets for quantities taken at the time $t - r/c$.

Then

$$V \approx \frac{Q_m \exp j\omega[t]}{4\pi\epsilon_0 r} \left\{ \frac{\exp \{js \cos \theta / (2\lambda)\}}{1 - (s \cos \theta) / (2r)} - \frac{\exp \{-js \cos \theta / (2\lambda)\}}{1 + (s \cos \theta) / (2r)} \right\}. \quad (37-44)$$

Expanding the exponentials and the denominators between the braces as power series, neglecting terms of the third order and higher in s/r and in s/λ , and then substituting p_m for $Q_m s$, we find that

$$V = \frac{p_m}{4\pi\epsilon_0\lambda r} \left(\frac{\lambda}{r} + j \right) \cos \theta \exp j\omega[t] \quad (s^2 \ll r^2, s^3 \ll \lambda^3). \quad (37-45)$$

We have made no assumption as to the relative magnitudes of r and λ . Of course,

$$\frac{\lambda}{r} + j = \left(\frac{\lambda^2}{r^2} + 1 \right)^{1/2} \exp j \left(\arctan \frac{r}{\lambda} \right). \quad (37-46)$$

Thus

$$V = \frac{p_m}{4\pi\epsilon_0\lambda r} \left(\frac{\lambda^2}{r^2} + 1 \right)^{1/2} \cos \theta \exp j \left\{ \omega \left(t - \frac{r}{c} \right) + \arctan \frac{r}{\lambda} \right\}. \quad (37-47)$$

Naturally, $V = 0$ in the equatorial plane at $\theta = 90^\circ$, where the potentials of the charges $+Q$ and $-Q$ cancel exactly. Figure 37-2 shows a radial plot of V as a function of θ and ϕ .

Observe the peculiar way in which the amplitude of V decreases with the distance r . Close to the dipole where $r^2/\lambda^2 \ll 1$, V falls off as $1/r^2$. But farther out, where $r^2/\lambda^2 \gg 1$, V falls off as $1/r$.

At zero frequency, $\lambda \rightarrow \infty$ and $[t] = t$. This returns to the V of a static field (Sec. 5.1).

As to \mathbf{A} , it is simply the vector potential of the element of current Is :

$$\mathbf{A} = \frac{\mu_0[I]s}{4\pi r} = \frac{\mu_0 j\omega[p]}{4\pi r}. \quad (37-48)$$

Since

$$\mu_0 = \frac{1}{\epsilon_0 c^2}, \quad \frac{\omega}{c} = \frac{1}{\lambda}, \quad (37-49)$$

$$\mathbf{A} = \frac{j[p]}{4\pi\epsilon_0 c \lambda r} = \frac{j p_m}{4\pi\epsilon_0 c \lambda r} \exp j\omega \left(t - \frac{r}{c} \right). \quad (37-50)$$

The vector potential propagates at the speed c everywhere. It is parallel to the polar axis, and its magnitude at P depends solely on the distance r to the dipole. See Fig. 37-2.

In polar coordinates,

$$\mathbf{A} = \frac{j[p]}{4\pi\epsilon_0 c \lambda r} (\cos \theta \hat{r} - \sin \theta \hat{\theta}). \quad (37-51)$$

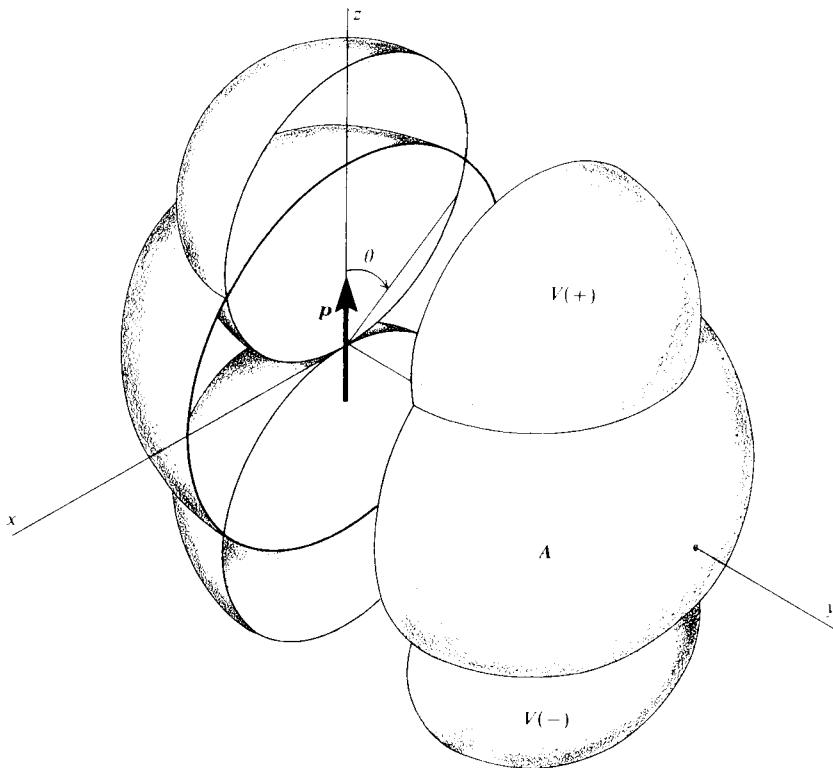


Fig. 37-2. The scalar potential V and the magnitude of the vector potential \mathbf{A} as functions of θ and ϕ , about an oscillating electric dipole oriented as shown. The radial distance from the center of the dipole to the spheres marked V is proportional to the value of V at a fixed distance in that particular direction. The scalar potential is maximum at the poles and vanishes at the equator, where the individual potentials of the charges $-Q$ and $+Q$ of the dipole cancel. It is positive in the northern hemisphere, where the field of $+Q$ is predominant, and negative in the southern hemisphere. The vector potential \mathbf{A} is independent, both in magnitude and in direction, of the coordinates θ and ϕ .

Example

THE RETARDED POTENTIALS FOR THE OSCILLATING MAGNETIC DIPOLE

The magnetic dipole of Fig. 37-3 is identical to that of Fig. 18-9, except that the source now supplies an alternating current. By hypothesis, there is zero net charge and $V = 0$.

We now have that

$$\mathbf{A} = \frac{\mu_0}{4\pi} \int_0^{2\pi} \frac{I_m \exp j\omega(t - r'/c)}{r'} a \cos \phi d\phi \hat{\phi} \quad (37-52)$$

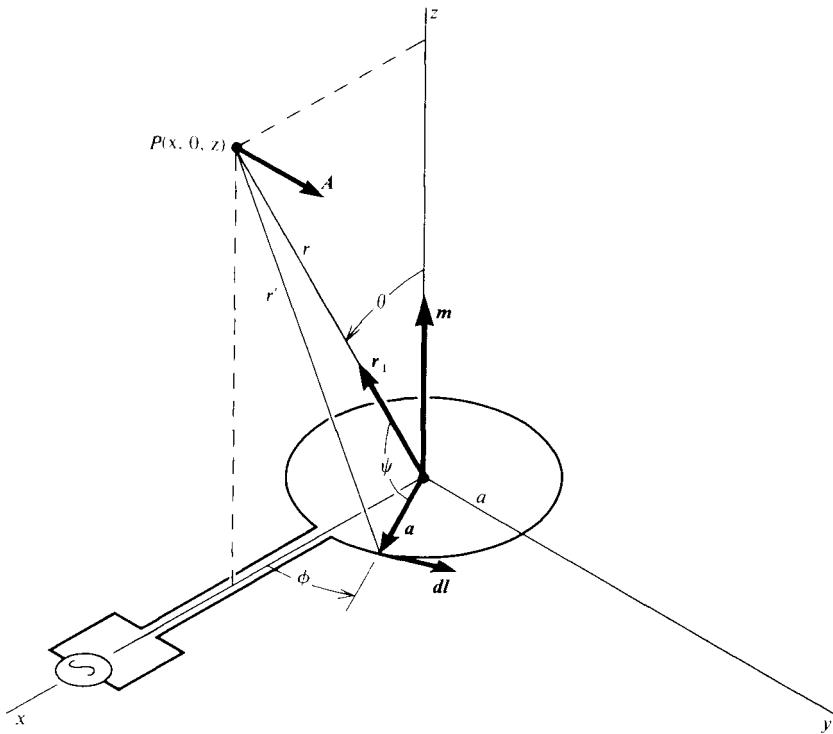


Fig. 37-3. An oscillating magnetic dipole.

$$\mathbf{A} = \frac{\mu_0}{4\pi} a I_m \exp j\omega[t] \int_0^{2\pi} \frac{\exp \{j(r - r')/\lambda\}}{r'} \cos \phi d\phi \hat{\phi}. \quad (37-53)$$

As in the third example in Sec. 18.4, we set

$$\frac{1}{r'} = \frac{1}{r} \left\{ 1 + \frac{a^2}{2r^2} + \frac{ax}{r^2} \cos \phi \right\}. \quad (37-54)$$

Then, within the same approximation $r^3 \gg a^3$,

$$r' \approx r \left(1 + \frac{a^2}{2r^2} - \frac{ax \cos \phi}{r^2} \right), \quad r - r' \approx \frac{ax \cos \phi}{r} - \frac{a^2}{2r}. \quad (37-55)$$

Also,

$$\exp \frac{r - r'}{\lambda} \approx 1 + j \frac{r - r'}{\lambda} \approx 1 + j \frac{ax \cos \phi}{\lambda r} - j \frac{a^2}{2\lambda r}, \quad (37-56)$$

if $a^2 \ll 2\lambda^2$. Then the above integral becomes equal to

$$\frac{1}{r} \int_0^{2\pi} \left(1 - \frac{a^2}{2r^2} + \frac{ax \cos \phi}{r^2} \right) \left(1 + j \frac{ax \cos \phi}{\lambda r} - \frac{a^2}{2\lambda r} \right) \cos \phi d\phi.$$

This integral is easy to evaluate. First we disregard the cosine and the cosine cubed terms because their integrals are zero. Also, the integral of $\cos^2 \phi$ is π . Then the integral is equal to

$$\frac{\pi ax}{r^2} \left(\frac{1}{r} + j \frac{1}{\lambda} \right).$$

Substituting now $r \sin \theta$ for x , and m_m for $\pi a^2 l_m$, m_m being the maximum value of the dipole moment, $m = m_m \exp j\omega t$,

$$\mathbf{A} = \frac{\mu_0 m_m}{4\pi \lambda r} \left(\frac{\lambda}{r} + j \right) \sin \theta \exp j\omega t \hat{\phi} \quad (37-57)$$

$$= \frac{j\mu_0 [m] \times \hat{r}}{4\pi \lambda r} \left(1 - j \frac{\lambda}{r} \right) \quad (a^3 \ll r^3, a^2 \ll 2\lambda^2). \quad (37-58)$$

Compare Eq. 37-57 with Eq. 37-45 for the V of an *electric* dipole.

Since $V = 0$, then $\nabla \cdot \mathbf{A} = 0$, from the Lorentz condition (Sec. 37.1).

For $r \gg \lambda$ the vector potential propagates as a spherical wave of wavelength c/f whose amplitude is inversely proportional to r . However, \mathbf{A} is zero on the axis of symmetry and maximum in the equatorial plane, again as we might expect.

37.5 A SECOND PAIR OF INTEGRALS FOR \mathbf{E} AND \mathbf{B}

We found the following integral for \mathbf{E} in Sec. 9.4:

$$\mathbf{E} = \frac{1}{4\pi \epsilon_0} \int_{v'} \frac{\rho \hat{r}}{r^2} dv', \quad (37-59)$$

where \mathbf{E} is the electric field strength at $P(x, y, z)$, $\rho = \rho_f + \rho_b$ is the total charge density at $P'(x', y', z')$, $dv' = dx' dy' dz'$ is the element of volume at P' , and r is the distance between P and P' . We disregard surface charges here.

In Sec. 20.3 we found that

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int_{v'} \frac{\mathbf{J} \times \hat{r}}{r^2} dv', \quad (37-60)$$

where \mathbf{B} is the magnetic flux density at the point $P(x, y, z)$ and

$$\mathbf{J} = \mathbf{J}_f + \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M}, \quad (37-61)$$

\mathbf{J}_f being the free current density, $\partial\mathbf{P}/\partial t$ the polarization density, and $\nabla \times \mathbf{M}$ the volume density of equivalent currents in magnetized matter at $P'(x', y', z')$.

The above integrals apply only if retardation effects are negligible. To find more general expressions, we start with the nonhomogeneous wave equations for \mathbf{E} and for \mathbf{B} . We found those equations in Sec. 27.8. By analogy with the solution of the corresponding equation for \mathbf{A} , the retarded fields are

$$\mathbf{E} = -\frac{1}{4\pi\epsilon_0} \int_{v'} \frac{[\nabla' \rho + \epsilon_0\mu_0(\partial\mathbf{J}/\partial t)]}{r} dv', \quad (37-62)$$

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int_{v'} \frac{[\nabla' \times \mathbf{J}]}{r} dv'. \quad (37-63)$$

Here \mathbf{E} and \mathbf{B} are the fields at (x, y, z, t) , while the bracketed terms are taken at $(x', y', z', t - r/c)$, r being again the distance between P and P' .

You will remember that the integrals for the retarded potentials involved $[\rho]$ and $[\mathbf{J}]$; the integrals for the retarded fields involve more complex quantities.

We found the above integral for \mathbf{E} , but for static fields, in Sec. 12.2.

The integral for \mathbf{B} is intriguing. For a static field we now have that

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int_{v'} \frac{\nabla' \times \mathbf{J}}{r} dv', \quad (37-64)$$

instead of Eq. 37-60. The source term is now $\nabla' \times \mathbf{J}$, instead of $\mathbf{J} \times \hat{r}$, and the term in the denominator is r instead of r^2 .

37.6 A THIRD PAIR OF INTEGRALS FOR \mathbf{E} AND \mathbf{B}

We can deduce another pair of integrals for \mathbf{E} and \mathbf{B} from the retarded potentials of Eqs. 37-31 and 37-33. Let us start with \mathbf{E} :

$$\mathbf{E} = -\nabla V - \frac{\partial \mathbf{A}}{\partial t} = -\nabla \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{[\rho]}{r} dv' - \frac{\partial}{\partial t} \frac{\mu_0}{4\pi} \int_{v'} \frac{[\mathbf{J}]}{r} dv' \quad (37-65)$$

$$= -\frac{1}{4\pi\epsilon_0} \int_{v'} \nabla \frac{[\rho]}{r} dv' - \frac{\mu_0}{4\pi} \int_{v'} \frac{\partial [\mathbf{J}]/\partial t}{r} dv'. \quad (37-66)$$

We can insert the del operator under the first integral sign because it operates on x, y, z while the integral operates on x', y', z' . Recall that the quantities between brackets are measured at $(x', y', z', t - r/c)$.

Now

$$\nabla \frac{[\rho]}{r} = \frac{1}{r} \nabla [\rho] + [\rho] \nabla \frac{1}{r} = \frac{1}{r} \nabla [\rho] - [\rho] \frac{\hat{r}}{r^2}. \quad (37-67)$$

Here $[\rho]$ is a function of r and hence of x, y, z . Therefore

$$\nabla [\rho] = \frac{\partial [\rho]}{\partial (t - r/c)} \nabla \left(t - \frac{r}{c} \right) = - \left[\frac{\partial \rho}{\partial t} \right] \frac{1}{c} \nabla r = - \left[\frac{\partial \rho}{\partial t} \right] \frac{\hat{r}}{c}. \quad (37-68)$$

Also,

$$\frac{\partial}{\partial t} [\mathbf{J}] = \frac{\partial [\mathbf{J}]}{\partial (t - r/c)} \frac{\partial (t - r/c)}{\partial t} = \frac{\partial [\mathbf{J}]}{\partial [t]} = \left[\frac{\partial \mathbf{J}}{\partial t} \right]. \quad (37-69)$$

Thus

$$\mathbf{E} = - \frac{1}{4\pi\epsilon_0} \int_{v'} \left(- \left[\frac{\partial \rho}{\partial t} \right] \frac{\hat{r}}{rc} - [\rho] \frac{\hat{r}}{r^2} \right) dv' - \frac{\mu_0}{4\pi} \int_{v'} \frac{[\partial \mathbf{J} / \partial t]}{r} dv' \quad (37-70)$$

$$= \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{[\rho] + [\partial \rho / \partial t](r/c)}{r^2} \hat{r} dv' - \frac{\mu_0}{4\pi} \int_{v'} \frac{[\partial \mathbf{J} / \partial t]}{r} dv'. \quad (37-71)$$

The first integral is curious. The integrand is $\rho dv'/r^2$, except that the charge density is ρ at $t - r/c$, extrapolated to the time t , to the *first* order! The term under the second integral is not extrapolated.

Similarly,

$$\mathbf{B} = \nabla \times \mathbf{A} = \nabla \times \frac{\mu_0}{4\pi} \int_{v'} \frac{[\mathbf{J}]}{r} dv' = \frac{\mu_0}{4\pi} \int_{v'} \nabla \times \frac{[\mathbf{J}]}{r} dv' \quad (37-72)$$

$$= \frac{\mu_0}{4\pi} \int_{v'} \frac{[\mathbf{J}] + [\partial \mathbf{J} / \partial t](r/c)}{r^2} \times \hat{r} dv'. \quad (37-73)$$

The numerator under the integral sign is again a first-order extrapolation of \mathbf{J} .

37.7 SUMMARY

The *Lorentz condition*

$$\nabla \cdot \mathbf{A} + \epsilon_0 \mu_0 \frac{\partial V}{\partial t} = 0 \quad (37-12)$$

is an identity if V and \mathbf{A} are defined by the usual integrals, retarded or not (Eq. 37-1 or Eqs. 37-31 and 37-33). If V is time-dependent, the three components of \mathbf{A} yield V , and then the six components of \mathbf{E} and \mathbf{B}

follow. The Lorentz condition is a consequence of the conservation of charge.

The *nonhomogeneous wave equation* for V is

$$\nabla^2 V - \epsilon_0 \mu_0 \frac{\partial^2 V}{\partial t^2} = -\frac{\rho}{\epsilon_0}, \quad (37-16)$$

where ρ is the total charge density $\rho_f + \rho_b$.

The *nonhomogeneous wave equation* for \mathbf{A} is

$$\nabla^2 \mathbf{A} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{J}, \quad (37-25)$$

where \mathbf{J} is $\mathbf{J}_f + \partial \mathbf{P}/\partial t + \nabla \times \mathbf{M}$.

The solutions of these two equations are the *retarded potentials*:

$$V(x, y, z, t) = \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\rho(x', y', z', t - r/c)}{r} dv', \quad (37-31)$$

$$\mathbf{A}(x, y, z, t) = \frac{\mu_0}{4\pi} \int_{v'} \frac{\mathbf{J}(x', y', z', t - r/c)}{r} dv'. \quad (37-33)$$

The solutions of the nonhomogeneous wave equations for \mathbf{E} and \mathbf{B} that we found in Sec. 27.8 are the *retarded fields*

$$\mathbf{E} = -\frac{1}{4\pi\epsilon_0} \int_{v'} \frac{[\nabla' \rho + \epsilon_0 \mu_0 \partial \mathbf{J} / \partial t]}{r} dv', \quad (37-62)$$

$$\mathbf{B} = +\frac{\mu_0}{4\pi} \int_{v'} \frac{[\nabla' \times \mathbf{J}]}{r} dv'. \quad (37-63)$$

The retarded potentials yield still another pair of integrals for \mathbf{E} and \mathbf{B} :

$$\mathbf{E} = \frac{1}{4\pi\epsilon_0} \int_{v'} \frac{[\rho] + [\partial \rho / \partial t](r/c)}{r^2} \hat{\mathbf{r}} dv' - \frac{\mu_0}{4\pi} \int_{v'} \frac{[\partial \mathbf{J} / \partial t]}{r} dv', \quad (37-71)$$

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int_{v'} \frac{([\mathbf{J}] + [\partial \mathbf{J} / \partial t](r/c))}{r^2} \times \hat{\mathbf{r}} dv'. \quad (37-73)$$

PROBLEMS

- 37-1. (37.4) The retarded vector potential near a long, straight wire carrying a time-dependent current

A long, straight wire of length C carries a current that increases linearly with time: $I = Kt$.

Show that, at a distance ρ from the wire such that $4\rho^2 \ll C^2$, and away from the ends,

$$A = \frac{\mu_0 I}{2\pi} \ln \frac{C}{\rho}.$$

Refer to the first example in Sec. 18.4. We have disregarded a term that is independent of both the time and the coordinates and that does not, therefore, affect either \mathbf{E} or \mathbf{B} . This particular A therefore has the same form as if the current were constant.

37-2. (37.4) The propagation speed of V and A in the field of an electric dipole

Find the propagation speed of the scalar and vector potentials V and A in the field of an electric dipole, on the assumption that $s^3 \ll r^3$ and that $s^3 \ll \lambda^3$.

Observe that the addition of two waves of V traveling at the speed c gives a wave whose phase speed is larger than c . This is an interference effect. Observe also that the speed of propagation V , close to the dipole, is a function of the wavelength and hence of the frequency. The dispersion originates, not in the properties of the medium, but rather in the geometry, as in rectangular metallic waveguides and in dielectric waveguides.

37-3. (37.4) Can A be zero in a radiation field?

In the field of a magnetic dipole, $V = 0$, $A \neq 0$. Do there exist radiation fields where the inverse is true?

37-4. (37.5) The alternate integral for \mathbf{B}

We have shown that

$$\mathbf{B} = \frac{\mu_0}{4\pi} \int_{v'} \frac{\nabla' \times \mathbf{J}}{r} dv',$$

if retardation is negligible. The volume v' encloses all the currents.

Show that the term on the right is equal to $\nabla \times \mathbf{A}$ for a finite current distribution. Refer to the identities on the inside of the front cover.

37-5. (37.5) Identity of the first and second integrals for \mathbf{E}

Show that, for any finite charge distribution, the integrals for \mathbf{E} given in Sec. 37.5 are equal:

$$\frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\rho \hat{\mathbf{r}}}{r^2} dv' = -\frac{1}{4\pi\epsilon_0} \int_{v'} \frac{\nabla' \rho}{r} dv'.$$

37-6. (37.5) Identity of the first and second integrals for \mathbf{B}

Show that, for any finite current distribution, the integrals for \mathbf{B} given in Sec. 37.5 are equal:

$$\frac{\mu_0}{4\pi} \int_{v'} \frac{\mathbf{J} \times \hat{\mathbf{r}}}{r^2} dv' = \frac{\mu_0}{4\pi} \int_{v'} \frac{\nabla' \times \mathbf{J}}{r} dv'.$$

37-7. (37.5)

Show that

$$\int_{v'} \nabla' \times \frac{\mathbf{J}}{r} dv' = 0.$$

37-8. (37.6) The electric field of the electric dipole, calculated from the third integral for \mathbf{E}

Calculate \mathbf{E} in the field of the electric dipole, starting from the integral of Sec. 37.6. To simplify the calculation, disregard terms in s^2/r^2 and s^3/r^3 as well as the higher-order terms in s/λ . This will leave you only the radiation term.

37-9. (37.5) The field at the center of a rotating disk of charge.

Suppose you have a disk of radius R and thickness $2s \ll R$. It carries a charge density $Q' = K(R - \rho)(s - z)^2$. We use Q' for the charge density in order to use ρ for the radial coordinate.

You are required to find the value of \mathbf{B} at the center when the disk rotates as a solid at the angular velocity ω . Of course, \mathbf{B} is normal to the plane of the disk. See Prob. 18-4.

Calculate \mathbf{B} from Eq. 37-60 and then from Eq. 37-63.

CHAPTER 38

RADIATION II *The Electric Dipole Transmitting Antenna*

38.1 RADIATION BY AN ACCELERATED CHARGE	692
38.2 ELECTRIC DIPOLE RADIATION	697
38.2.1 THE ELECTRIC FIELD STRENGTH E	698
38.2.2 THE MAGNETIC FIELD STRENGTH H	699
38.2.3 THE POYNTING VECTOR $E \times H$	701
38.2.4 THE RADIATED POWER P	702
Examples: THE COLOR OF THE SKY, OF THE SETTING SUN, AND OF TOBACCO SMOKE	
38.2.5 RADIATION RESISTANCE	703
*38.2.6 THE LINES OF E	704
*38.2.7 THE $K\ell$ SURFACE	705
*38.2.8 THE LINES OF H	708
38.3 SUMMARY	708
PROBLEMS	709

The final two chapters address the radiation of electromagnetic fields by antennas. There exist a seemingly endless variety of antennas, but we study only three basic types, namely, the electric dipole, the half-wave antenna, and the magnetic dipole. We also study briefly some simple antenna arrays that serve to orient the radiation in a given direction.

In this chapter we return to the electric dipole. We found its potentials in the first example in Sec. 37.4, and we now study its field in some detail.

38.1 RADIATION BY AN ACCELERATED CHARGE

We saw in Chap. 17 that, if a charge moves at a constant speed, its field follows as a rigid body. The lines of E are radial. As the speed approaches c , then, for a fixed observer, the field contracts in the longitudinal direction and eventually shrinks to the plane normal to the direction of motion.

If the charge accelerates, then the lines of E become twisted into odd shapes, as in Fig. 38-1. The kink moves out at the speed c .

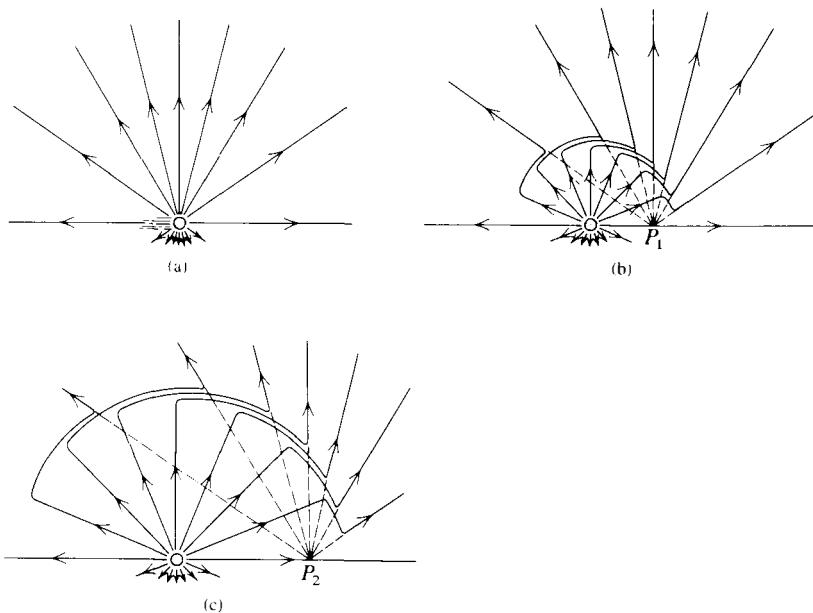


Fig. 38-1. (a) The field of a point charge that travels at a velocity of $0.8c$. (b) The field a short time after the charge has stopped. The kink travels outward at the speed of light. For a distant observer the field is that of a charge at P_1 . (c) The field a short time later.

Let us do a rough calculation of \mathbf{E} and \mathbf{B} in the kink. This will yield all the basic characteristics of radiation fields.

It appears that J. J. Thomson was the first to calculate this \mathbf{E} . He did so in 1903,[†] 2 years before the publication of Einstein's historic paper on relativity. Several authors have since extended the calculation to relativistic velocities and to \mathbf{B} .

Imagine a charge that starts from rest, accelerates at the rate a for a short interval of time τ , and then continues in a straight line at the constant velocity $a\tau$. Figure 38-2(a) and (b) shows the velocity v and the distance traveled z as functions of the time. The interval τ is short and $z \approx (a\tau)t$.

Figure 38-3 shows how a line of \mathbf{E} changes with time. Beyond the circle of radius ct centered on the origin, the field is that of a stationary charge at 0. Inside a circle of radius $c(t - \tau)$ centered on the charge, the field is

[†] J. J. Thomson, *Electricity and Matter*, Yale University Press, New Haven, Conn., 1904, p. 55.

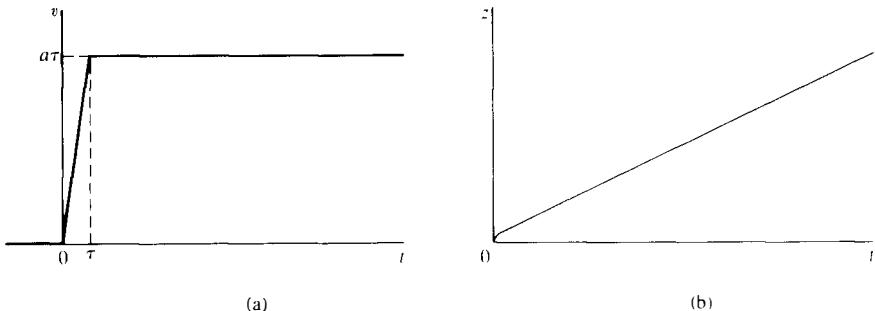


Fig. 38-2. (a) Velocity v of a point charge as a function of the time t . Between $t = 0$ and $t = \tau$ the acceleration is a . (b) Distance z traveled as a function of t . For $t \gg \tau$, $z \approx (a\tau)t$.

that of a point charge moving at the constant velocity $a\tau$. Set again $v^2 \ll c^2$. Then, in the figure, bd is parallel to Oe because the field moves as a solid.

Assume that, in the kink, the line of \mathbf{E} is straight. This is a coarse approximation! Then, in that region, from Fig. 38-4,

$$\frac{E_r}{E_\theta} = \frac{ct}{vt \sin \theta} = \frac{c\tau}{a\tau t \sin \theta} = \frac{c}{at \sin \theta} = \frac{c^2}{ar \sin \theta}, \quad (38-1)$$

where $r = ct = Oe$.

We now apply Gauss's law and set

$$E_r = \frac{Q}{4\pi\epsilon_0 r^2}. \quad (38-2)$$

This is also a coarse approximation because the radial component of \mathbf{E} is not uniform in the region between the spheres centered on O and on b . Then

$$E_\theta = \frac{ar \sin \theta}{c^2} \frac{Q}{4\pi\epsilon_0 r^2} = \frac{Qa \sin \theta}{4\pi\epsilon_0 c^2 r} = \frac{\mu_0 Q a \sin \theta}{4\pi r}. \quad (38-3)$$

We have used the relation $\epsilon_0 \mu_0 = 1/c^2$, from Sec. 27.8.

Observe that this E_θ is proportional to the acceleration, that it varies inversely as the *first* power of r , and that it is maximum in the direction *perpendicular* to the acceleration.

We now calculate the magnetic field in the region of the kink. We expect a \mathbf{B} that is azimuthal around the z -axis, like the magnetic field

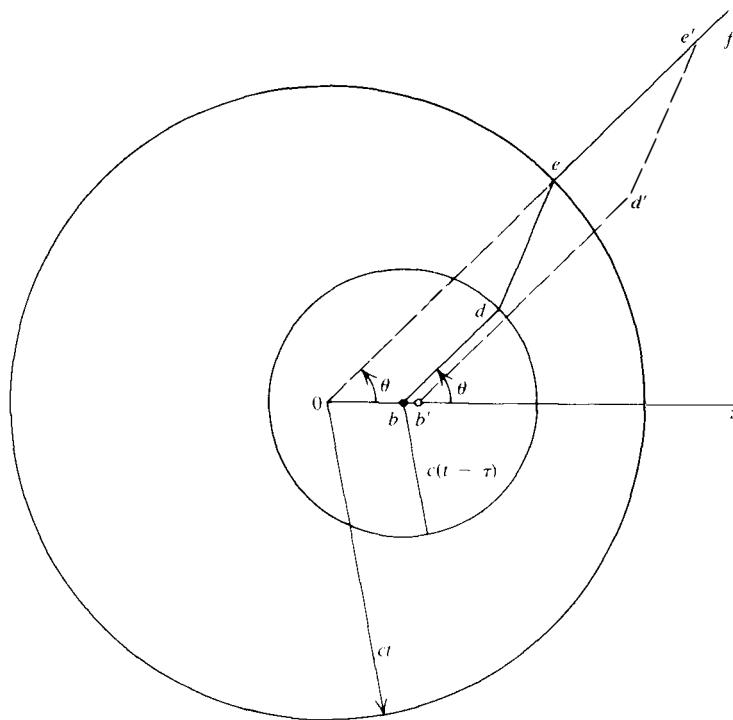


Fig. 38-3. The charge is originally at rest at 0. After a time τ it moves to the right at the constant velocity $a\tau$. When the charge reaches point b , one of its lines of \mathbf{E} has the shape $bdef$. A short time later the charge is at b' , and the line of \mathbf{E} is now $b'd'e'f$.

near a current-carrying wire. If C is a circle whose plane is perpendicular to the paper in Fig. 38-4 and that goes through points e and g ,

$$\mathbf{B} \cdot d\mathbf{l} = 2\pi r \sin \theta B. \quad (38-4)$$

Now apply Eq. 27-31 to the circle C and to the area of the spherical surface bounded on the left by C :

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \epsilon_0 \mu_0 \int_{\mathcal{A}} \frac{\partial \mathbf{E}}{\partial t} \cdot d\mathbf{A} = \epsilon_0 \mu_0 \frac{d}{dt} \int_{\mathcal{A}} \mathbf{E} \cdot d\mathbf{A} = \epsilon_0 \mu_0 \frac{d\Phi_E}{dt}, \quad (38-5)$$

where Φ_E is the flux of \mathbf{E} through C . Thus $d\Phi_E$ is the sudden change in the electric flux that occurs when the region of the kink goes through the point e , and $dt = \tau$.

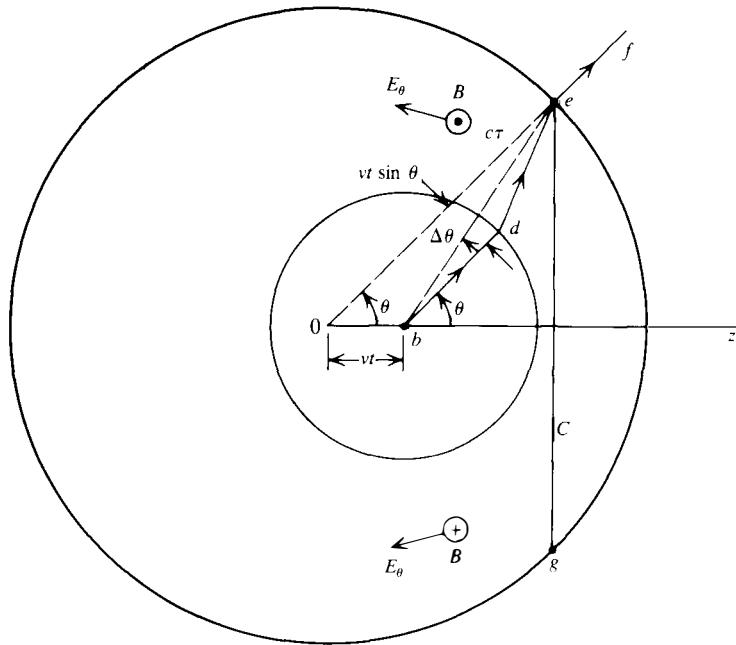


Fig. 38-4. The charge is at *b*, as in Fig. 38-3. *C* is a circle whose plane is perpendicular to the paper and that is centered on the *z*-axis.

In Fig. 38-4 the spherical surface bounded on the left by *C* lies, just before the arrival of the kink, in the field of a stationary charge situated at the origin and

$$\Phi_E = \frac{Q}{\epsilon_0} \frac{\Omega}{4\pi} = \frac{Q}{\epsilon_0} \frac{1 - \cos \theta}{2}, \quad (38-6)$$

where Ω is the solid angle subtended by *C* at the origin.

After τ seconds the kink has swept through *e*. The point *e* then lies in the field of a charge situated slightly to the right of *b*, and Φ_E is larger:

$$\Phi_E + \Delta\Phi_E = \frac{Q}{\epsilon_0} \frac{1 - \cos(\theta + \Delta\theta)}{2}. \quad (38-7)$$

So

$$\frac{\Delta\Phi_E}{\tau} = \frac{Q}{2\epsilon_0\tau} \Delta(-\cos \theta) = \frac{Q}{2\epsilon_0\tau} \sin \theta \Delta\theta. \quad (38-8)$$

Now the two circles in Figs. 38-3 and 38-4 are very close together. We can therefore set $r \Delta\theta = vt \sin \theta$. Then

$$\frac{\Delta\Phi_E}{\tau} = \frac{Q \sin \theta}{2\epsilon\tau} \frac{vt \sin \theta}{r} = \frac{Q \sin^2 \theta a \tau t}{2\epsilon_0 \tau r} = \frac{Qa \sin^2 \theta t}{2\epsilon_0 r}. \quad (38-9)$$

Combining, finally, Eqs. 38-4, 38-5, and 38-9, and setting $r = ct$, gives

$$B = \frac{1}{2\pi r \sin \theta} \epsilon_0 \mu_0 \frac{\Delta\Phi_E}{\tau} = \frac{\mu_0}{4\pi} \frac{Qa \sin \theta}{cr} = \frac{E_\theta}{c}. \quad (38-10)$$

Therefore, both E_θ and B in the kink result from the acceleration, they are proportional to the acceleration a , they vary as $1/r$, and they are maximum in the direction perpendicular to the acceleration. Moreover, E_θ and B are mutually orthogonal, and orthogonal to the direction of propagation of the kink. Finally, $E_\theta = Bc$. All these characteristics apply to the radiation fields that we shall study in this chapter and the next.[†]

38.2 ELECTRIC DIPOLE RADIATION

Figure 38-5 shows an electric dipole. As we saw in the example in Sec. 37.4, this simple model serves to calculate the radiation field of any charge distribution whose dipole moment is a sinusoidal function of the

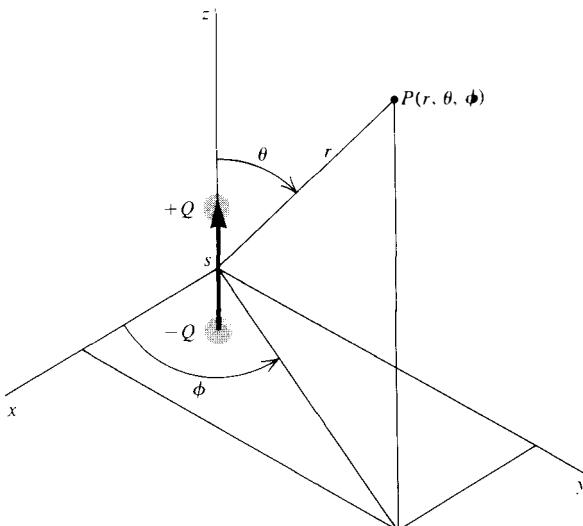


Fig. 38-5. Electric dipole of moment $\mathbf{p} = Q\mathbf{s}$ and a point P in its field.

[†]They do not apply to the field close to a source of radiation.

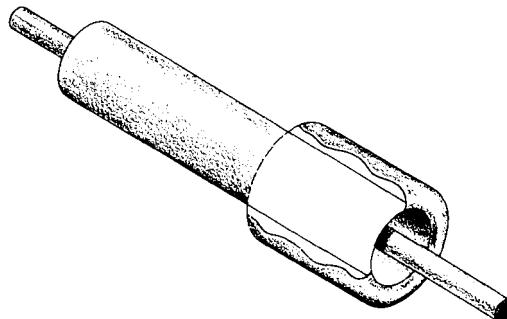


Fig. 38-6. Electric dipole antenna at the end of a coaxial line. The outer conductor is folded back.

time, disregarding quadrupole, octupole, and higher-order fields. In the oscillating dipole, charges oscillate along s , and their acceleration is thus lengthwise.

We calculated V and \mathbf{A} in the example, setting $s^3 \ll r^3$ and $s^3 \ll \lambda^3$, but without limiting the relative magnitudes or r and λ . We now deduce \mathbf{E} , \mathbf{H} , the Poynting vector \mathcal{S} , and the radiated power P .

The dipole is in free space.

Figure 38-6 shows an electric dipole antenna at the end of a coaxial line.

38.2.1 The Electric Field Strength \mathbf{E}

First, in polar coordinates,

$$\nabla V = \frac{\partial V}{\partial r} \hat{\mathbf{r}} + \frac{1}{r} \frac{\partial V}{\partial \theta} \hat{\mathbf{\theta}}, \quad (38-11)$$

V being independent of ϕ . Then, from Eq. 37-45[†]

$$\nabla V = \frac{[p]}{4\pi\epsilon_0\lambda^2 r} \left\{ \left(1 - 2 \frac{\lambda^2}{r^2} - 2j \frac{\lambda}{r} \right) \cos \theta \hat{\mathbf{r}} - \left(\frac{\lambda^2}{r^2} + j \frac{\lambda}{r} \right) \sin \theta \hat{\mathbf{\theta}} \right\}. \quad (38-12)$$

Also, from Eq. 37-51,

$$\frac{\partial \mathbf{A}}{\partial t} = j\omega \mathbf{A} = -\frac{[p]}{4\pi\epsilon_0\lambda^2 r} = -\frac{[p]}{4\pi\epsilon_0\lambda^2 r} (\cos \theta \hat{\mathbf{r}} - \sin \theta \hat{\mathbf{\theta}}) \quad (38-13)$$

and

[†] We again reserve brackets for quantities measured at the previous time $t - r/c$.

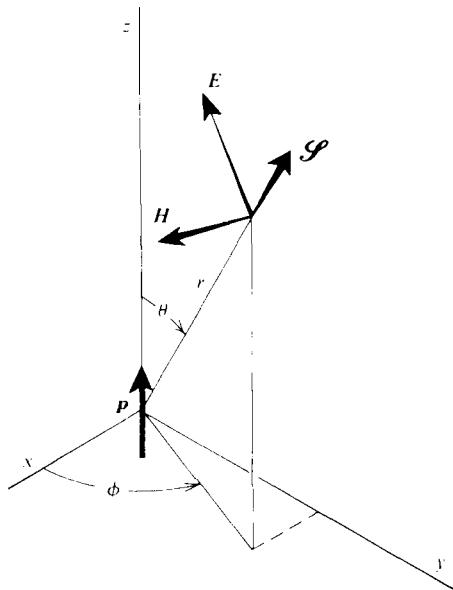


Fig. 38-7. The \mathbf{E} , \mathbf{H} , and $\mathcal{S} = \mathbf{E} \times \mathbf{H}$ vectors for an oscillating electric dipole for $r \gg \lambda$ and $\omega[t] = 0$.

$$\mathbf{E} = \frac{[P]}{4\pi\epsilon_0\lambda^2 r} \left\{ 2\left(\frac{\lambda^2}{r^2} + j\frac{\lambda}{r}\right) \cos \theta \hat{\mathbf{r}} + \left(\frac{\lambda^2}{r^2} - 1 + j\frac{\lambda}{r}\right) \sin \theta \hat{\boldsymbol{\theta}} \right\}. \quad (38-14)$$

Close to the dipole, where $r \ll \lambda$, or at zero frequency ($\lambda \rightarrow \infty$), only two terms survive, and

$$\mathbf{E} = \frac{P}{4\pi\epsilon_0 r^3} (2 \cos \theta \hat{\mathbf{r}} + \sin \theta \hat{\boldsymbol{\theta}}) \quad (r \ll \lambda, \text{ or } f = 0). \quad (38-15)$$

These are the *static terms* of Sec. 5.1. They fall off as $1/r^3$.

Far away from the dipole, for $r \gg \lambda$, only the *radiation term* remains, and it falls off as $1/r$:

$$\mathbf{E} = -\frac{[P]}{4\pi\epsilon_0\lambda^2 r} \sin \theta \hat{\boldsymbol{\theta}} \quad (r \gg \lambda). \quad (38-16)$$

See Figs. 38-7 and 38-8. Remember that we have set $s^3 \ll r^3$ and $s^3 \ll \lambda^3$. Observe that \mathbf{E} is proportional to $1/\lambda^2$, or to ω^2 , or to the second time derivative of p , or to the *acceleration* of the charges.

38.2.2 The Magnetic Field Strength \mathbf{H}

In the example in Sec. 37.4 we found that in the field of an electric dipole

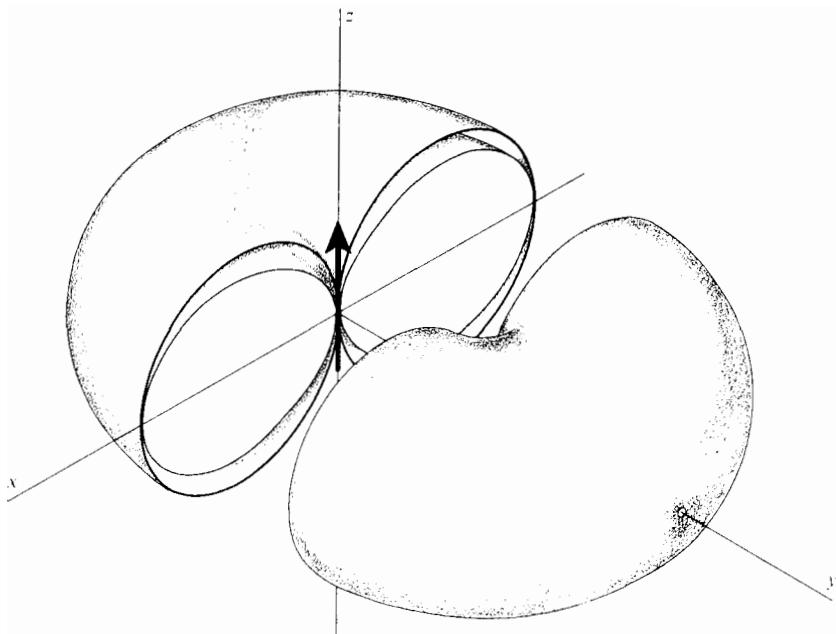


Fig. 38-8. Polar diagrams of $\sin \theta$ (outer surface) and of $\sin^2 \theta$ (inner surface) showing, respectively, the angular distributions of E , or of H , and of S_{av} at a distance $r \gg \lambda$ from an oscillating electric dipole situated at the origin. The radial distance from the dipole to one of the surfaces is proportional to the magnitude of the quantity in the corresponding direction. There is zero field and zero power flow along the axis.

$$\mathbf{A} = \frac{j[p]}{4\pi\epsilon_0 c \lambda r} (\cos \theta \hat{\mathbf{r}} - \sin \theta \hat{\mathbf{\theta}}). \quad (38-17)$$

Then \mathbf{H} has only a ϕ component:

$$\mathbf{H} = \frac{1}{\mu_0} \nabla \times \mathbf{A} = \frac{1}{\mu_0 r} \left\{ \frac{\partial}{\partial r} (r A_\theta) - \frac{\partial A_r}{\partial \theta} \right\} \hat{\phi}. \quad (38-18)$$

To work this out, we require the following two derivatives:

$$\frac{\partial}{\partial r} [p] = \frac{\partial}{\partial r} \left\{ p_m \exp j\omega \left(t - \frac{r}{c} \right) \right\} = -\frac{j\omega [p]}{c} = -\frac{j[p]}{\lambda}, \quad (38-19)$$

$$\frac{\partial}{\partial \theta} [p] = 0. \quad (38-20)$$

Then

$$\mathbf{H} = \frac{c[p]}{4\pi\lambda^2 r} \left(-1 + j \frac{\lambda}{r} \right) \sin \theta \hat{\phi}, \quad (38-21)$$

again with $s^3 \ll r^3$ and $s^3 \ll \lambda^3$.

For $r \ll \lambda$, and using Eqs. 37-39,

$$\mathbf{H} = \frac{j\omega[p]}{4\pi r^2} \sin \theta \hat{\phi} = \frac{[I]s \times \hat{r}}{4\pi r^2} \quad (r \ll \lambda), \quad (38-22)$$

which agrees with the Biot-Savart law of Sec. 18.2.

For $r \gg \lambda$,

$$\mathbf{H} = -\frac{\omega[p]}{4\pi\lambda r} \sin \theta \hat{\phi} = -\frac{\omega[\mathbf{p}] \times \hat{r}}{4\pi\lambda r} = -\frac{c[\mathbf{p}] \times \hat{r}}{4\pi\lambda^2 r} \quad (r \gg \lambda), \quad (38-23)$$

as in Figs. 38-7 and 38-8.

For $r \gg \lambda$, the characteristic impedance is the same as that for a uniform plane wave in free space (Sec. 28.3):

$$Z = \frac{E}{H} = \mu_0 c \approx 377 \text{ ohms} \quad (r \gg \lambda). \quad (38-24)$$

The electric and magnetic energies are equal.

Curiously enough, \mathbf{A} propagates everywhere at the speed c , but \mathbf{H} propagates at that speed only for $r \gg \lambda$. Closer in, its phase speed is larger than c .

At zero frequency $\omega = 0$ and $\mathbf{H} = 0$, as expected.

38.2.3 The Poynting Vector $\mathbf{E} \times \mathbf{H}$

We now calculate the time-averaged Poynting vector and the radiated power.

For any r ,

$$\mathcal{S}_{av} = \frac{1}{2} \operatorname{Re} (\mathbf{E} \times \mathbf{H}^*) = \frac{1}{2} \operatorname{Re} \{ (E_r \hat{r} + E_\theta \hat{\theta}) \times H_\phi^* \hat{\phi} \}. \quad (38-25)$$

Recalling that

$$\hat{r} \times \hat{\phi} = -\hat{\theta}, \quad \hat{\theta} \times \hat{\phi} = \hat{r}, \quad (38-26)$$

then, from Eqs. 38-14 and 38-21,

$$\mathcal{S}_{av} = \frac{1}{2} \operatorname{Re} (-E_r H_\phi^* \hat{\theta} + E_\theta H_\phi^* \hat{r}) \quad (38-27)$$

$$= \frac{\mu_0 \omega^4 p_m^2}{32\pi^2 c r^2} \sin^2 \theta \hat{r} = \frac{\mu_0 \pi^2}{2c} \frac{f^4 p_m^2}{r^2} \sin^2 \theta \hat{r} \quad (38-28)$$

$$= 4.137 \times 10^{-14} \frac{f^4 p_{\text{rms}}^2}{r^2} \sin^2 \theta \hat{r} \quad \text{watts/meter}^2. \quad (38-29)$$

Or, since $I_m s = \omega p_m = 2\pi f p_m$,

$$\begin{aligned} \mathcal{S}_{\text{av}} &= \frac{\mu_0}{8c} \frac{f^2 I_m^2 s^2}{r^2} \sin^2 \theta \hat{r} \\ &= 1.0496 \times 10^{-15} \frac{f^2 I_{\text{rms}}^2 s^2}{r^2} \sin^2 \theta \hat{r} \quad \text{watts/meter}^2 \end{aligned} \quad (38-30)$$

$$\begin{aligned} &= \frac{10^{-7} c}{4\pi} \left(\frac{s}{\lambda}\right)^2 \frac{I_{\text{rms}}^2}{r^2} \sin^2 \theta \hat{r} \\ &= 2.386 \left(\frac{s}{\lambda}\right)^2 \frac{I_{\text{rms}}^2}{r^2} \sin^2 \theta \hat{r} \quad \text{watts/meter}^2. \end{aligned} \quad (38-31)$$

Note the following points about this time-averaged Poynting vector.

- (1) It involves only the radiation terms, despite the fact that our calculation is valid even if r is not much larger than λ . The time-averaged power flux is everywhere radial, at least for $r^3 \gg s^3$, $\lambda^3 \gg s^3$. See Figs. 38-7 and 38-8
- (2) It varies as $1/r^2$ because, under steady conditions, the power flow through any given solid angle must be independent of r to satisfy the conservation of energy. This $1/r^2$ dependence results from the fact that the radiation terms for \mathbf{E} and for \mathbf{H} both vary as $1/r$.
- (3) Since it varies as $\sin^2 \theta$, it is zero along the axis of the dipole and maximum in the equatorial plane, as in Fig. 38-8. *An electric dipole does not radiate along its axis.*

38.2.4 The Radiated Power P

Integrating the time-averaged Poynting vector over a sphere of radius r yields the radiated power:

$$P = \frac{\mu_0 \pi^2 f^4 p_{\text{rms}}^2}{cr^2} \int_0^{2\pi} \int_0^\pi \sin^2 \theta r^2 \sin \theta d\theta d\phi \quad (38-33)$$

$$= \frac{8\mu_0 \pi^3}{3c} f^4 p_{\text{rms}}^2 = 3.466 \times 10^{-13} f^4 p_{\text{rms}}^2 \quad \text{watts.} \quad (38-34)$$

Substituting again $I_{\text{rms}}s/\omega$ for p_{rms} ,

$$P = \frac{2\mu_0\pi}{3c} f^2 I_{\text{rms}}^2 s^2 = 8.779 \times 10^{-15} f^2 I_{\text{rms}}^2 s^2 \quad (38-35)$$

$$= \frac{2 \times 10^{-7} c}{3} \left(\frac{s}{\lambda}\right)^2 I_{\text{rms}}^2 = 19.99 \left(\frac{s}{\lambda}\right)^2 I_{\text{rms}}^2 \quad \text{watts.} \quad (38-36)$$

Thus the radiated power is proportional to: (1) $(s/\lambda)^2$, which is a small quantity as we assumed in the first example in Sec. 37.4; (2) s^2 , where s is the length of the dipole; (3) f^2 for a given I_{rms} ; and (4) f^4 for a given p_{rms} .

Example

THE COLORS OF THE SKY, OF THE SETTING SUN, AND OF TOBACCO SMOKE[†]

Dust particles suspended in the atmosphere scatter the light coming from the sun. This scattering occurs because the electric field of the incident light wave excites electrons present in the particles. These electrons act as small electric dipoles and reradiate. If we disregard resonances, the p_m of an oscillating electron is proportional to the amplitude of the incident wave. Then the reradiated power is proportional to f^4 , and the light scattered by the sky is bluer than sunlight.

If the air were completely dust-free, the sky would still be blue, but darker: atoms and molecules of the air also absorb and reradiate energy, but mostly in the ultraviolet. The light from the sun that reaches the earth, particularly at sunset, is reddish because part of the blue has been diffused out.

It is for the same reason that tobacco smoke is either bluish or reddish, according to the way you look at it, with respect to a source of light. See Fig. 38-9.

38.2.5 Radiation Resistance

We saw above that the radiated power is proportional to the square of the rms current. The factor of proportionality is the *radiation resistance*:

$$R_{\text{rad}} \approx 19.99 \left(\frac{s}{\lambda}\right)^2 \quad \text{ohms} \quad (38-37)$$

if $(s/\lambda)^3 \ll 1$.

Let $(s/\lambda)^3 = 0.01$. Then $(s/\lambda)^2 = 0.46$, and the radiation resistance is about 0.9 ohm.

[†] See M. Minnaert, *The Nature of Light and Color in the Open Air*. Dover, New York, 1954.

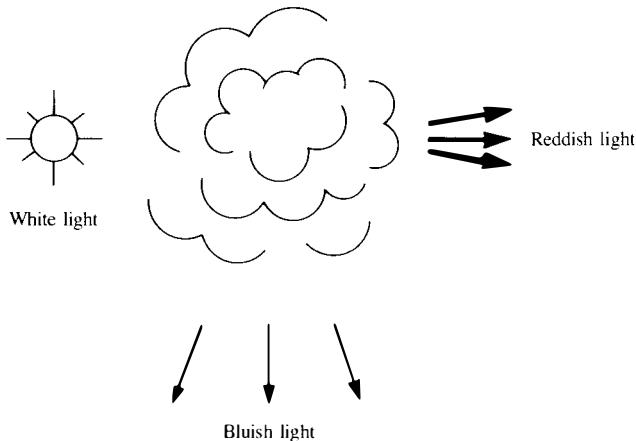


Fig. 38-9. Fine particles, as in tobacco smoke, scatter blue light preferentially. The transmitted light is thus reddish.

*38.2.6 The Lines of \mathbf{E}

Of course, there is no such thing as a distinct line of \mathbf{E} or of \mathbf{H} . All that we know is that the fields \mathbf{E} and \mathbf{H} possess both a magnitude and a direction that vary from one point to another in space, and from one instant to another in time, according to certain laws. Lines of \mathbf{E} and of \mathbf{H} , however, provide the best way to picture a field.

We can find an equation for the lines of \mathbf{E} by setting

$$\frac{E_r}{dr} = \frac{E_\theta}{r d\theta}, \quad (38-38)$$

where dr and $r d\theta$ are the components of an element of a line of \mathbf{E} . Here E_r and E_θ are the real parts of the corresponding phasors. We continue our calculation with phasors.

First we note that, from Maxwell's equation for the curl of \mathbf{H} (Eq. 27-27),

$$\mathbf{E} = \frac{1}{j\omega\epsilon_0} \nabla \times \mathbf{H}. \quad (38-39)$$

Since \mathbf{H} has only a ϕ component, from Sec. 38.2.2,

$$E_r = \frac{1}{j\omega\epsilon_0} \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (H \sin \theta), \quad E_\theta = -\frac{1}{j\omega\epsilon_0} \frac{1}{r} \frac{\partial}{\partial r} (rH). \quad (38-40)$$

Thus the differential equation for the lines of \mathbf{E} becomes

$$\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} (H \sin \theta) d\theta = -\frac{1}{r} \frac{\partial}{\partial r} (rH) dr. \quad (38-41)$$

Multiplying both sides by $r \sin \theta$ gives

$$\frac{\partial}{\partial \theta} (Hr \sin \theta) d\theta + \frac{\partial}{\partial r} (Hr \sin \theta) dr = 0. \quad (38-42)$$

The total differential of $Hr \sin \theta$ is therefore zero, and on a given line of \mathbf{E} , $Hr \sin \theta$ is a constant. Substituting the value of \mathbf{H} from Eq. 38-21 after replacing $j\omega p$ by Is , expressing the complex factor in polar form, and replacing the exponential function by a cosine, we have the equation for a line of \mathbf{E} :

$$\sin^2 \theta \left(\frac{\lambda^2}{r^2} + 1 \right)^{1/2} \cos \left(\omega t - \frac{r}{\lambda} + \arctan \frac{r}{\lambda} \right) = K\lambda. \quad (38-43)$$

The parameter K varies from one line to the next. Remember that we have assumed the conditions $(s/\lambda)^3 \ll 1$ and $(s/r)^3 \ll 1$.

Figure 38-10 shows eight families of lines of \mathbf{E} .[†]

At $r \gg \lambda$,

$$\sin^2 \theta \cos \left(\omega t - \frac{r}{\lambda} + \frac{\pi}{2} \right) = K\lambda, \quad (38-44)$$

and the lines of \mathbf{E} travel outward at speed $\omega\lambda = c$. Closer in, the arctan term is a function of r and the speed of the lines is *larger* than c .

*38.2.7 The $K\lambda$ Surface

It is instructive to plot Eq. 38-43 as a three-dimensional surface as in Fig. 38-11. The loops are both contour lines and lines of \mathbf{E} at $t = 0$. In fact, they are the same lines of \mathbf{E} as those of Fig. 38-10(a). As time goes on, the argument of the cosine function increases, and the ripples move out as a damped wave, carrying the lines of \mathbf{E} with them.

Let us see how the lines behave. Figure 38-12 shows the intersection of the $K\lambda$ surface, again at $t = 0$, with the plane $\theta = \pi/2$. These curves lie inside the envelope

$$K\lambda = \pm \left(\frac{\lambda^2}{r^2} + 1 \right)^{1/2} \quad (\sin \theta = 1). \quad (38-45)$$

[†] See the animated film loop on *Dipole Radiation* by R. H. Good, California State University, Hayward, California.

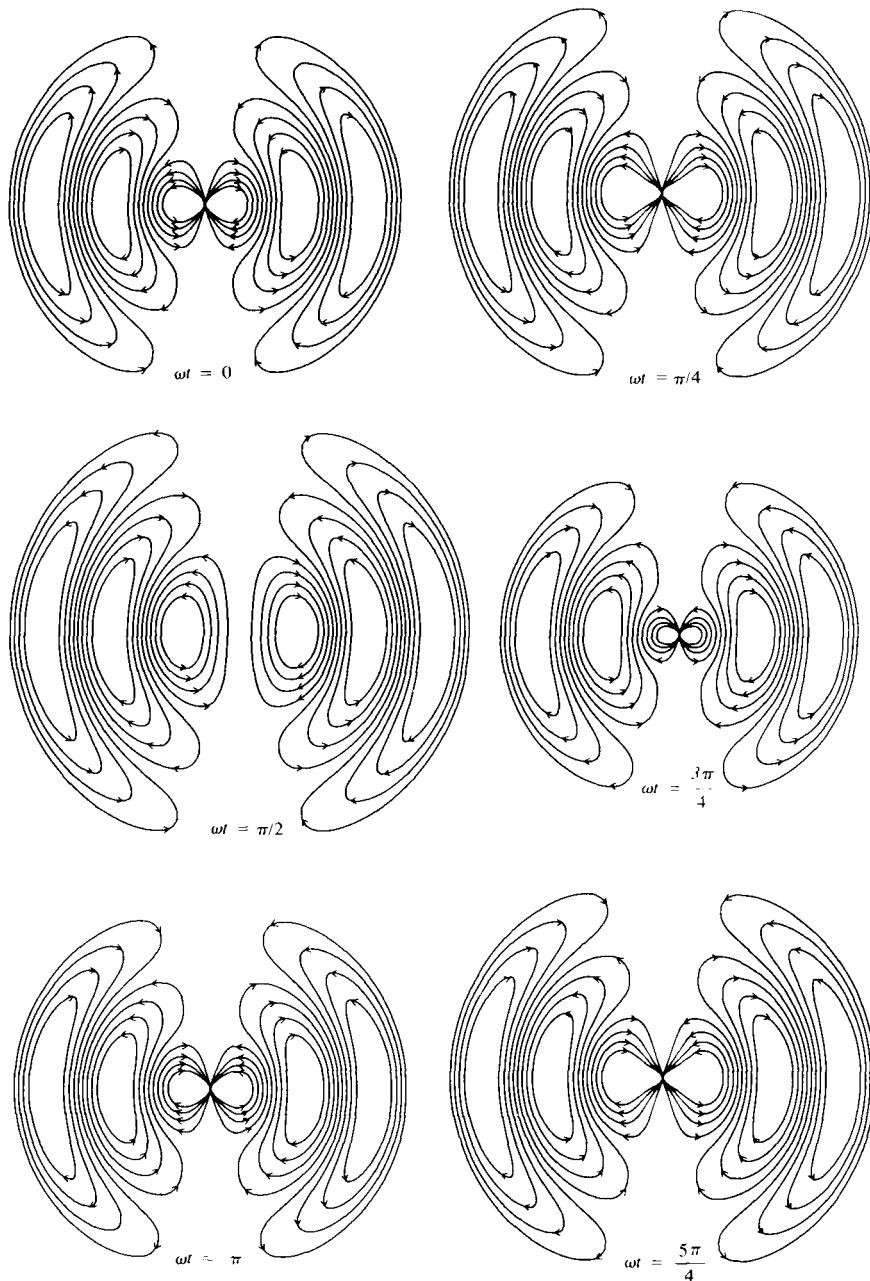
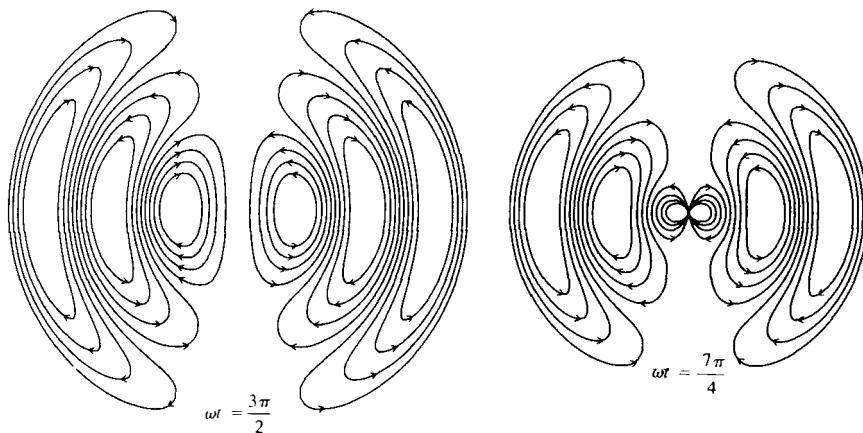


Fig. 38-10. Lines of \mathbf{E} for an oscillating electric dipole for $\omega t = 0, \pi/4, \pi/2, 3\pi/4, \pi, 5\pi/4, 3\pi/2, 7\pi/4$. The dipole is vertical at the center. Note how the wavelength decreases with distance. The lines of \mathbf{H} are circles perpendicular to the paper and centered on the axis of the dipole.



As the ripples move out, their height decreases rapidly and soon approaches unity. Clearly, lines of \mathbf{E} with $K\lambda < 1$ can travel out to infinity. They provide the *radiation field*. It is also clear that, if $K\lambda > 1$, they cannot go far. If $K\lambda$ is only slightly larger than unity, a loop shrinks until it reaches the top of a ripple and then disappears.

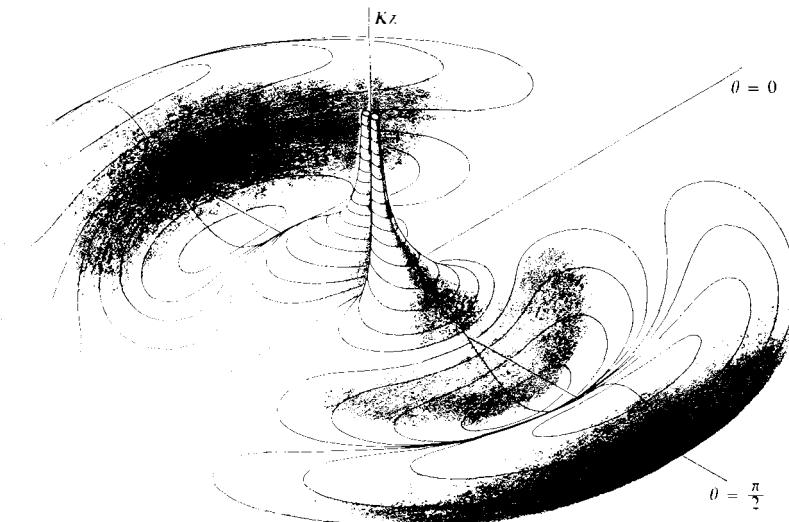


Fig. 38-11. The parameter $K\lambda$ plotted as a function of the coordinates r and θ at $t = 0$. The dipole is at the center, on the $\theta = 0$ axis. As t increases, the central peaks oscillate in unison from $-\infty$ to $+\infty$, and the ripples move out radially. The loops are both level lines and lines of \mathbf{E} .

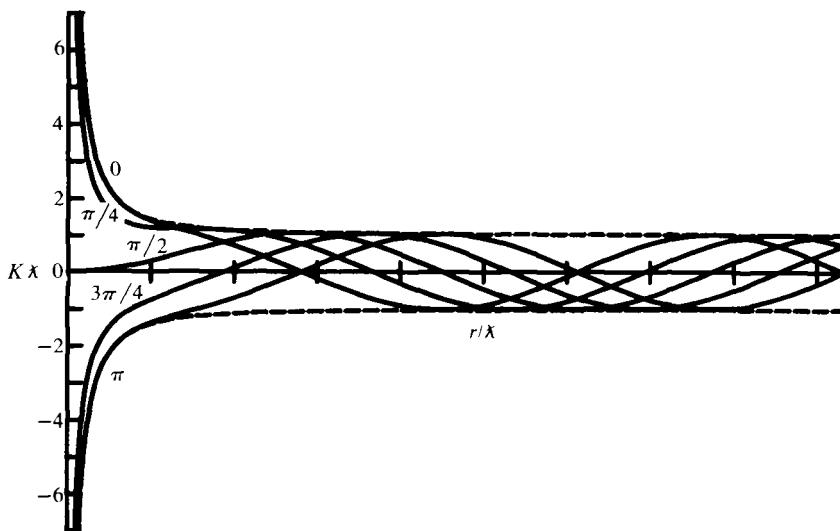


Fig. 38-12. This figure illustrates how the intersection of the $K\lambda$ surface with the plane $\theta = \pi/2$ changes with time. The curves all lie within the envelope shown as a dashed line.

*38.2.8 The Lines of \mathbf{H}

We have seen in Sec. 38.2.2 that \mathbf{H} has only a ϕ component. Thus the lines of \mathbf{H} are circles perpendicular to, and centered on, the axis of the electric dipole.

38.3 SUMMARY

In the field of an oscillating electric dipole,

$$\mathbf{E} = \frac{[p]}{4\pi\epsilon_0\lambda^2 r} \left\{ 2\left(\frac{\lambda^2}{r^2} + j\frac{\lambda}{r}\right) \cos \theta \hat{\mathbf{r}} + \left(\frac{\lambda^2}{r^2} - 1 + j\frac{\lambda}{r}\right) \sin \theta \hat{\mathbf{\theta}} \right\}, \quad (38-14)$$

$$\mathbf{H} = \frac{c[p]}{4\pi\lambda^2 r} \left(-1 + j\frac{\lambda}{r} \right) \sin \theta \hat{\mathbf{\phi}}, \quad (38-21)$$

$$\mathcal{S}_{av} = \frac{1}{2} \operatorname{Re} (\mathbf{E} \times \mathbf{H}^*) = 4.137 \times 10^{-14} \frac{f^4 p_{rms}^2}{r^2} \sin^2 \theta \hat{\mathbf{r}} \quad (38-29)$$

$$= 2.386 \left(\frac{s}{\lambda}\right)^2 \frac{I_{rms}^2}{r^2} \sin^2 \theta \hat{\mathbf{r}} \quad \text{watts/meter}^2. \quad (38-32)$$

Far away from the dipole ($r \gg \lambda$), \mathbf{E} has only a θ component and \mathbf{H} is azimuthal. The time-averaged Poynting vector is everywhere radial.

The *radiation power* is

$$P = 3.466 \times 10^{-13} f^4 p_{\text{rms}}^2 \quad (38-34)$$

$$= 8.799 \times 10^{-15} f^2 I_{\text{rms}}^2 s^2 \quad (38-35)$$

$$= 19.99 \left(\frac{s}{\lambda} \right)^2 I_{\text{rms}}^2 \quad \text{watts.} \quad (38-36)$$

The *radiation resistance* of an electric dipole is about $20(s/\lambda)^2$ ohms.

On a *line of E* ,

$$\sin^2 \theta \left(\frac{\lambda^2}{r^2} + 1 \right)^{1/2} \cos \left(\omega t - \frac{r}{\lambda} + \arctan \frac{r}{\lambda} \right) = K\lambda. \quad (38-43)$$

Lines of H are circles perpendicular to and centered on the axis of the dipole.

PROBLEMS

38-1. (38.1) The radiation field of a long wire carrying a step current

Figure 38-13(a) shows a long wire that carries a current that varies as in Fig. 38-13(b). Beyond $\rho_2 = ct$ there is no field. Inside $\rho_1 = c(t - \tau)$ the field is that of a steady current I_0 . In the shaded region,

$$\mathbf{A} = \frac{\mu_0 I}{2\pi} \ln \frac{C}{\rho} \hat{z},$$

where C is the length of the wire, as in Prob. 37-1.

- (a) Calculate \mathbf{E} and \mathbf{B} in the shaded region.
- (b) Show that Maxwell's equations apply.
- (c) Sketch curves of E and B as functions of ρ at a given instant, between $\rho = 0$ and $\rho = \rho_2$. There are discontinuities at $\rho = \rho_1$ and at $\rho = \rho_2$ because we have assumed that d^2I/dt^2 is infinite at $t = 0$ and at $t = \tau$.

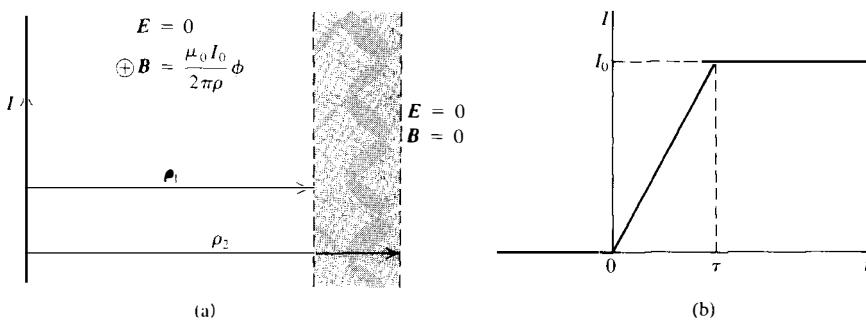


Fig. 38-13.

38-2. (38.2) The dipole moment of an oscillating charge

A charge Q oscillates along the z -axis, and $z = z_m \exp j\omega t$.

What is the dipole moment of an equivalent oscillating dipole? You can find the equivalence in the following way. If the currents are the same, then the \mathbf{A} 's are the same. Then the \mathbf{V} 's are the same, from the Lorentz condition (Sec. 37.1). Then the \mathbf{E} 's and \mathbf{H} 's are the same.

38-3. (38.2.1) The three components of the \mathbf{E} field of an electric dipole

Show that the electric field of an electric dipole has three components: one that depends on the positions of the charges, one that depends on their velocities, and one that depends on their accelerations.

38-4. (38.2.3) The radiation pattern of the electric dipole

What fraction of the total power in the field of an electric dipole is radiated within 45° of the equatorial plane?

38-5. (38.2.3) The electric and magnetic energy densities in the field of an electric dipole

Calculate the ratio of the time-averaged electric energy density to the time-averaged magnetic energy density in the field of an electric dipole (a) for $r \ll \lambda$, (b) for $r = \lambda$, and (c) for $r \gg \lambda$.

38-6. (38.2.3) The Poynting vector and the energy density in the field of an electric dipole

Show that, for $r \gg \lambda$, the magnitude of the Poynting vector in the field of an electric dipole is equal to the energy density multiplied by c .

38-7. (38.2.4) The light source paradox

All light sources should be black, for the following reason. Take the sun, for example. A cone in the retina of the eye collects radiation emanating from a very large number of atoms. These sources are incoherent.

At any given instant there is a near-infinite number of phasors in the complex plane, all of different magnitudes and different phases, rotating at different velocities. Their vector sum is clearly zero.

The same reasoning applies to any object, say a white wall, illuminated with incoherent light. The radiation that reaches a given cone comes from an area that is a large number of wavelengths in diameter. There again, the net field at the cone should be zero, and the wall should appear black.

To explain this paradox, consider N waves of a single frequency and of a given linear polarization but of random amplitudes and phases. The number N is very large. For the i -th wave, $E_i = E_{m_i} \exp j(\omega t - \alpha_i)$ at the cone, and the net E is the sum of the E_i 's.

Now the eye is sensitive, not to E but to \mathcal{S} , and thus to EE^* . Show that $\mathcal{S}_{av} = \sum \mathcal{S}_i$. This means that the net energy flux is equal to the sum of the energy fluxes of the individual waves.[†]

[†] You have probably noticed that laser light diffused by a wall or a sheet of paper has a granular structure. The pattern moves if one moves one's head from side to side. This phenomenon is called *speckle*. It is used for studying surfaces.

Speckle arises in the following way. Each point on the object, say the sheet of paper, produces on the retina a diffraction pattern whose shape and size depend on the optical characteristics of the eye. Since the radiation is coherent, there is interference between these patterns and the field varies from point to point on the retina.

38-8. (38.2.4) Cosmological evolution

Consider a particular class of astronomical objects, say quasars. Assume that they are all identical and distributed uniformly in a Euclidean universe.

Show that, if N is the number of objects whose radio-frequency flux is greater than \mathcal{S} at the earth, then a plot of $\log N$ against $\log \mathcal{S}$ should be a straight line whose slope is -1.5 . The slope for quasars is, in fact, larger. This is possibly a measure of cosmological evolution.

38-9. (38.2.4) The polarization of skylight

The atmosphere scatters sunlight. Draw a sketch showing the sun, the earth, and a vector E on a ray of scattered light. Explain why skylight is polarized. The light is only partially polarized because it is scattered many times.

CHAPTER 39

RADIATION III

The Half-Wave Antenna, Antenna Arrays, and the Magnetic Dipole Antenna

39.1 RADIATION FROM A HALF-WAVE ANTENNA	713
39.1.1 THE ELECTRIC FIELD STRENGTH E	714
39.1.2 THE MAGNETIC FIELD STRENGTH H	715
39.1.3 THE POYNTING VECTOR $E \times H$	716
39.1.4 THE RADIATED POWER P AND THE RADIATION RESISTANCE	716
39.2 ANTENNA ARRAYS	717
<i>Example: PAIR OF PARALLEL ANTENNAS SEPARATED BY ONE-HALF WAVELENGTH</i>	717
39.3 MAGNETIC DIPOLE RADIATION	720
39.3.1 THE ELECTRIC FIELD STRENGTH E	721
39.3.2 THE MAGNETIC FIELD STRENGTH H	722
39.3.3 THE POYNTING VECTOR, THE RADIATED POWER, AND THE RADIATION RESISTANCE	723
39.3.4 ELECTRIC AND MAGNETIC DIPOLE RADIATION COMPARED	724
39.4 THE ELECTRIC DIPOLE AS A RECEIVING ANTENNA	724
39.5 THE MAGNETIC DIPOLE AS A RECEIVING ANTENNA	724
39.6 SUMMARY	725
PROBLEMS	726

The half-wave antenna is a long, straight conductor, one-half wavelength long, that carries a standing wave of current. Its radiation pattern is similar to that of an electric dipole. However, for a given current, it radiates much more energy. This is the building block for assembling arrays of antennas. We deduce its field from that of an electric dipole.

The directivity of a half-wave antenna is hardly better than that of an electric dipole. However, arrays of such antennas, with the proper spacings and the proper phases, can be highly directive. Some arrays comprise a few antennas, but others comprise thousands.

We also calculate E and B in the field of a magnetic dipole, and we discuss briefly electric and magnetic dipoles as receiving antennas.

This chapter ends our study of electromagnetic fields and waves. Obviously we have not exhausted the subject! Indeed, we have done no more than establish a base from which you can explore on your own.

39.1 RADIATION FROM A HALF-WAVE ANTENNA

Figure 39-1 shows a half-wave antenna connected to a transmitter through a parallel-wire line. The half-wave antenna is essentially a pair of wires, each $\lambda/4$ long, fed with a current $I_m \cos \omega t$ at the junction. Here λ is the wavelength of a uniform plane wave in the medium of propagation.

At short wavelengths one can fold back a length $\lambda/4$ of the outer conductor of a coaxial line, as in Fig. 38-6 to obtain a half-wave antenna.

Roof antennas for automobiles are only one-quarter wavelength long; the other half is a reflection in the sheet metal of the roof. Transmitting antennas for AM waves are similarly $\lambda_0/4$ towers standing on conducting ground.

The antenna carries a standing wave of current, with a maximum at the center and nodes at the end. The current at l is thus

$$I = I_m \cos \frac{l}{\lambda} \exp j\omega t. \quad (39-1)$$

Each element of length dl radiates as an electric dipole.

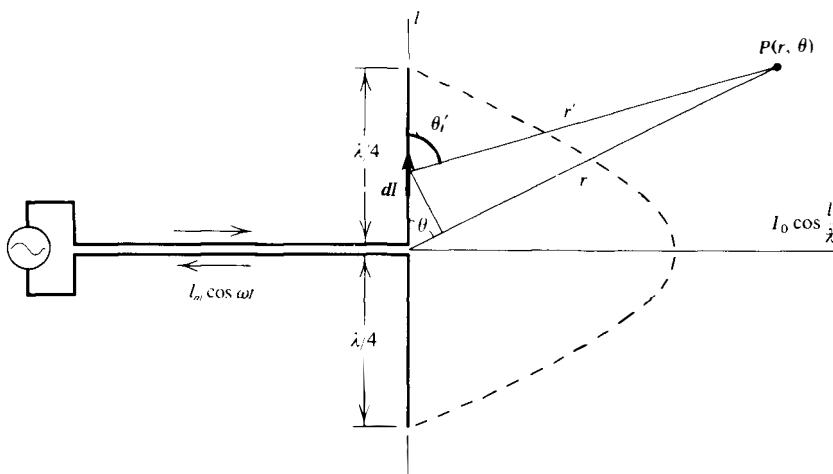


Fig. 39-1. Half-wave antenna. The broken line shows the standing wave of current at $\cos \omega t = 1$.

This description of the half-wave antenna is contradictory because the standing wave along the conductor can be truly sinusoidal only if there is zero energy loss, hence no radiation. In a real antenna the current distribution is not quite sinusoidal, but the distortion hardly affects the field.

The standing wave of current is the sum of two waves, one in the positive direction of l and the other in the negative direction, each of amplitude $I_m/2$:[†]

$$I = \frac{I_m}{2} \left\{ \exp j\left(\omega t - \frac{l}{\lambda}\right) + \exp j\left(\omega t + \frac{l}{\lambda}\right) \right\}. \quad (39-2)$$

39.1.1 The Electric Field Strength \mathbf{E}

We set $r \gg \lambda$. Then $\theta' \approx \theta$, Eq. 38-16 applies, and

$$\mathbf{dE} = -\frac{[dp]}{4\pi\epsilon_0\lambda^2 r'} \sin \theta \hat{\theta} = -\frac{\omega^2 [dp]}{4\pi\epsilon_0 c^2 r'} \sin \theta \hat{\theta} \quad (39-3)$$

$$= \frac{\mu_0 j \omega [I] dl}{4\pi r'} \sin \theta \hat{\theta}, \quad (39-4)$$

where

$$r' \approx r - l \cos \theta, \quad (39-5)$$

as in Fig. 39-1, or

$$\mathbf{dE} = \frac{\mu_0 j \omega I_m}{8\pi r'} \left\{ \exp j\left(\omega[t'] - \frac{l}{\lambda}\right) + \exp j\left(\omega[t'] + \frac{l}{\lambda}\right) \right\} \sin \theta dl \hat{\theta}. \quad (39-6)$$

We now integrate over the length of the antenna to find \mathbf{E} at r, θ . We can replace the r' in the denominator by r since $r \gg \lambda$, hence $r \gg l$. However, we must *not* replace the r' by r in

$$[t'] = t - \frac{r'}{c} \quad (39-7)$$

because the phases of the exponential terms vary rapidly with r' . So we set

$$[t'] \approx t - \frac{r - l \cos \theta}{c} = [t] + \frac{l \cos \theta}{c}. \quad (39-8)$$

[†] As in the previous two chapters, we reserve brackets for quantities evaluated at $t = r/c$.

At a given point in space, the $d\mathbf{E}$'s thus all have about the same amplitude and direction, but their phases differ. All these $d\mathbf{E}$'s point in the direction of the local unit vector $\hat{\theta}$. Then

$$\mathbf{E} = \frac{\mu_0 j \omega I_m}{8\pi r} \sin \theta \exp j\omega[t] \int_{-\lambda/4}^{+\lambda/4} \left\{ \exp j \frac{l(\cos \theta - 1)}{\lambda} + \exp j \frac{l(\cos \theta + 1)}{\lambda} \right\} dl \hat{\theta}. \quad (39-9)$$

Integrating yields

$$\mathbf{E} = \frac{jI_m}{4\pi c \epsilon_0 r} \sin \theta \exp j\omega[t] \left(\frac{\sin \{\pi(\cos \theta - 1)/2\}}{\cos \theta - 1} + \frac{\sin \{\pi(\cos \theta + 1)/2\}}{\cos \theta + 1} \right) \hat{\theta}, \quad (39-10)$$

where

$$\sin \frac{\pi(\cos \theta - 1)}{2} = -\cos \left(\frac{\pi}{2} \cos \theta \right), \quad \sin \frac{\pi(\cos \theta + 1)}{2} = +\cos \left(\frac{\pi}{2} \cos \theta \right). \quad (39-11)$$

Thus

$$\mathbf{E} = \frac{j}{2\pi c \epsilon_0 r} \frac{\cos \{(\pi/2) \cos \theta\}}{\sin \theta} [I] \hat{\theta} \approx 60.0j \frac{\cos \{(\pi/2) \cos \theta\}}{r \sin \theta} [I] \hat{\theta}. \quad (39-12)$$

This expression is indeterminate at $\theta = 0$ and at $\theta = \pi$. But, according to L'Hospital's rule, the limiting value of such a ratio is equal to the limiting value of the ratio of the derivatives. So \mathbf{E} is zero on the axis of a half-wave antenna, in agreement with the fact that the elementary dipoles do not radiate along the axis.

Why should the magnitude of \mathbf{E} be independent of the frequency? The explanation is that the \mathbf{E} of an elementary dipole, for a given current, is proportional to $1/\lambda$, but the antenna is $\lambda/2$ long.

Figure 39-2 shows that the radiation pattern for a half-wave antenna is similar to that of a dipole. This is because the phase differences between the $d\mathbf{E}$'s from the elements of current along the antenna are small near $\theta = \pi/2$, where the $d\mathbf{E}$'s are large, and are large only near the polar axis where the $d\mathbf{E}$'s tend to zero.

39.1.2 The Magnetic Field Strength \mathbf{H}

The value of \mathbf{H} follows immediately. We found in Sec. 38.1.2 that, for the electric dipole, \mathbf{H} is azimuthal and

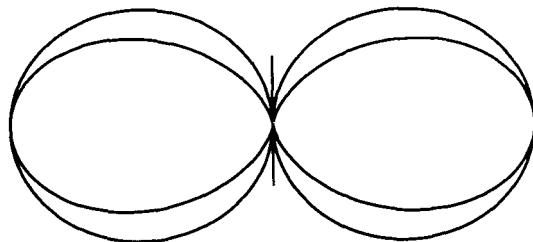


Fig. 39-2. Polar diagrams of the functions $\cos \{(\pi/2) \cos \theta\}$ (outer pair of curves) and of its square (inner pair). On the outer pair, the distance from the origin to a point on a curve is proportional to the magnitude of E , or of H at a fixed distance from a half-wave antenna, in that direction. On the inner pair, this distance is proportional to S_{av} . Compared to the electric dipole, the half-wave antenna radiates a somewhat larger fraction of its power in the region of the equatorial plane.

$$\frac{E}{H} = \left(\frac{\mu_0}{\epsilon_0} \right)^{1/2} \approx 377 \text{ ohms} \quad (r \gg \lambda). \quad (39-13)$$

Therefore, in the field of a half-wave antenna,

$$\mathbf{H} = \frac{j}{2\pi r} \frac{\cos \{(\pi/2) \cos \theta\}}{\sin \theta} [I] \hat{\phi}. \quad (39-14)$$

39.1.3 The Poynting Vector $\mathbf{E} \times \mathbf{H}$

The time-averaged Poynting vector is

$$\mathcal{S}_{av} = \frac{1}{2} \operatorname{Re} (\mathbf{E} \times \mathbf{H}^*) \quad (39-15)$$

$$= \frac{1}{\pi c \epsilon_0} \frac{\cos^2 \{(\pi/2) \cos \theta\}}{\sin^2 \theta} \frac{I_{rms}^2}{4\pi r^2} \hat{r} \quad (39-16)$$

$$= 9.543 \frac{\cos^2 \{(\pi/2) \cos \theta\}}{\sin^2 \theta} \frac{I_{rms}^2}{r^2} \hat{r} \quad \text{watts/meter}^2. \quad (39-17)$$

See Fig. 39-2.

39.1.4 The Radiated Power P and the Radiation Resistance

To obtain the radiated power, we integrate over a sphere of radius r :

$$P = \frac{I_{rms}^2}{4\pi^2 c \epsilon_0 r^2} 2\pi \int_0^\pi \frac{\cos^2 \{(\pi/2) \cos \theta\}}{\sin^2 \theta} r^2 \sin \theta d\theta. \quad (39-18)$$

The integral is equal to 1.2188267, and

$$P = 73.083 I_{\text{rms}}^2 \approx 73 I_{\text{rms}}^2 \quad \text{watts.} \quad (39-19)$$

The radiation resistance of a half-wave antenna is about 73 ohms.

39.2 ANTENNA ARRAYS

The electric dipole and the half-wave antenna are omnidirectional in the equatorial plane: at a given distance, the amplitude of the field is the same in all directions in that plane. Omnidirectional antennas have their uses, but for most applications the radiation field of an antenna should be maximum in a given direction. This is achieved with arrays of half-wave antennas that are properly spaced and properly phased.

Linear arrays comprise several parallel half-wave antennas disposed along a straight line. *Planar arrays* operating at wavelengths of the order of 1 centimeter comprise many more, often thousands, disposed over a rectangular or circular plane surface. Usually the individual antennas are identical, equally spaced, and oriented similarly. Beam steering and pattern control are achieved nearly instantaneously by means of phase shifters next to each element.

The radiation patterns of arrays are typically like the one shown in Fig. 39-3, with one main lobe and several smaller side lobes.

An *adaptive receiving array* adjusts its pattern automatically to optimize the signal-to-noise ratio in the presence of identifiable noise sources.

We illustrate the principle involved in antenna arrays by calculating the field of two half-wave antennas spaced by $\lambda/2$, first when they are in phase and then when they are in opposite phases.

Example PAIR OF PARALLEL ANTENNAS SEPARATED BY ONE-HALF WAVELENGTH

Figure 39-4 shows a pair of parallel half-wave antennas separated by a distance $\lambda/2$. We assume that $r \gg \lambda$.

The antennas are in phase

If the antennas are in phase, then \mathbf{E} at point P is the sum of two terms like that of Eq. 39-12, except that one wave travels a distance $r + (\lambda/4) \cos \psi$ and the other a distance $r - (\lambda/4) \cos \psi$. Therefore one wave leads, relative to an imaginary antenna at the center, by the phase angle

$$\frac{2\pi \lambda \cos \psi}{\lambda} = \frac{\pi}{2} \cos \psi, \quad (39-20)$$

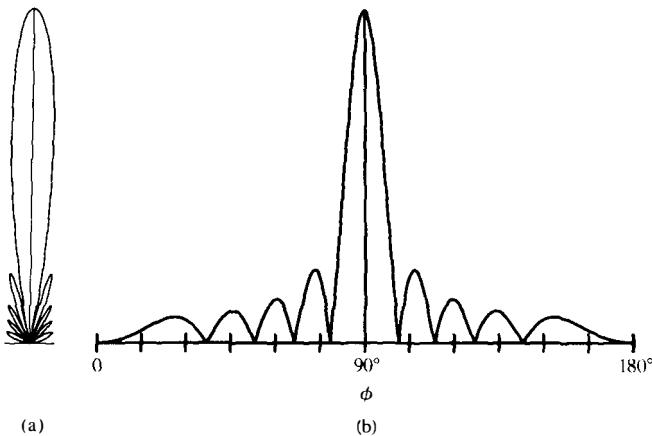


Fig. 39-3. The radiation pattern of an antenna is a plot of E as a function of θ . This is the radiation pattern of a 10-element linear array of in-phase half-wave antennas. (a) Polar diagram. (b) Cartesian diagram.

the other lags by the same amount, and

$$\mathbf{E} = 60j \frac{\cos\{(\pi/2)\cos\theta\}}{r \sin\theta} \left\{ \exp\left(j\frac{\pi}{2}\cos\psi\right) + \exp\left(-j\frac{\pi}{2}\cos\psi\right) \right\} [I]\hat{\theta} \quad (39-21)$$

$$= 120j \frac{\cos\{(\pi/2)\cos\theta\}}{r \sin\theta} \cos\left(\frac{\pi}{2}\cos\psi\right) [I]\hat{\theta}. \quad (39-22)$$

The angle ψ is awkward to use, but we can express it in terms of θ and ϕ , since

$$r \cos\psi = r \sin\theta \cos\phi. \quad (39-23)$$

Then

$$\mathbf{E} \approx 120j \frac{\cos\{(\pi/2)\cos\theta\}}{r \sin\theta} \cos\left(\frac{\pi}{2}\sin\theta \cos\phi\right) [I]\hat{\theta}. \quad (39-24)$$

In the xy -plane, $\theta = \pi/2$ and

$$E \propto \cos\left(\frac{\pi}{2}\cos\phi\right). \quad (39-25)$$

This function is zero at ϕ equal to 0 or π , and maximum at

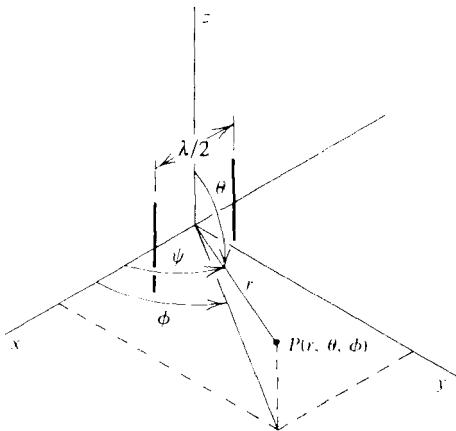


Fig. 39-4. Pair of parallel half-wave antennas separated by a distance of $\lambda/2$. The distances from the centers of the antennas to P are approximately $r - (\lambda/4) \cos \psi$ and $r + (\lambda/4) \cos \psi$.

$\phi = \pi/2$: there is destructive interference along the x -axis and constructive interference along the y -axis.

In the xz -plane, $\phi = 0$ and

$$E \propto \frac{\cos \{(\pi/2) \cos \theta\}}{\sin \theta} \cos \left(\frac{\pi}{2} \sin \theta \right). \quad (39-26)$$

The first term on the right is the angular distribution for a single half-wave antenna; it is zero at $\theta = 0$ and maximum at $\theta = \pi/2$. The second term comes from the interference between the two antennas; it is maximum at $\theta = 0$ and zero at $\theta = \pi/2$. The product of the two is zero both at $\theta = 0$ and at $\theta = \pi/2$.

Finally, in the yz -plane, $\phi = \pi/2$ and

$$E \propto \frac{\cos \{(\pi/2) \cos \theta\}}{\sin \theta}, \quad (39-27)$$

as for a single half-wave antenna. The two waves are in phase, and the total field is twice that of a single antenna when $r \gg \lambda$.

Figure 39-5 shows the radiation pattern.

The antennas are in opposite phases

The antenna at $x = \lambda/4$ now leads by π . Equation 39-21 applies, except that the first term between the pair of braces on the right is negative and

$$E = 120 \frac{\cos \{(\pi/2) \cos \theta\}}{r \sin \theta} \sin \left(\frac{\pi}{2} \sin \theta \cos \phi \right) [I] \hat{\theta}. \quad (39-28)$$

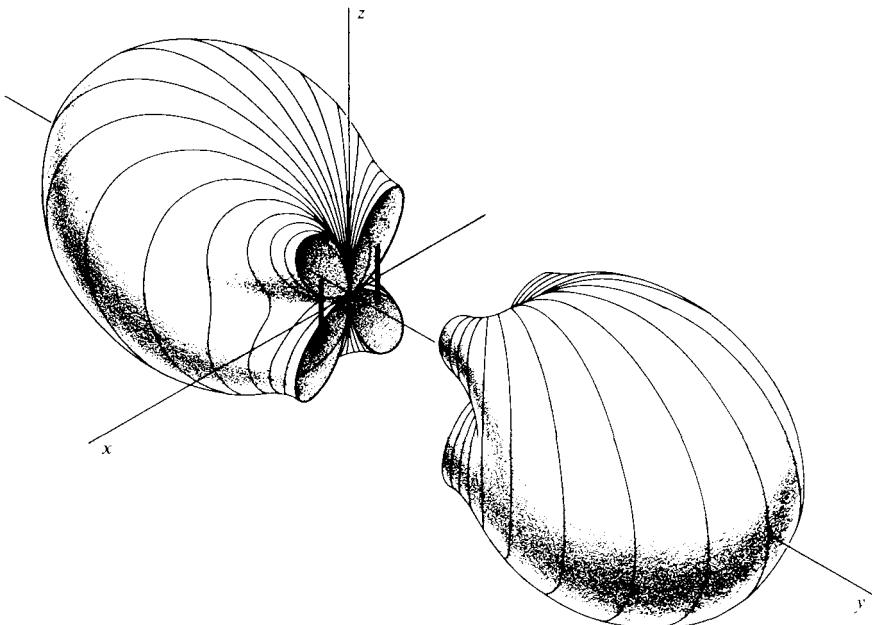


Fig. 39-5. The radiation pattern for the simple antenna array of Fig. 39-4 when the two antennas are excited in phase and for $r \gg \lambda$. Here we have plotted the magnitude of E , or of H , radially as a function of θ and of ϕ . We have split the surface into two parts for clarity. In the yz -plane, the field is twice that of a single antenna. Along the x -axis the waves arrive in opposite phases, for $r \gg \lambda$, and cancel. There is zero field on the z -axis, again for $r \gg \lambda$.

The radiation pattern is now that of Fig. 39-6.

These simple arrays are only slightly more directional than a single half-wave antenna.

Clearly, one can obtain a wide range of radiation patterns by varying either the geometry of an antenna array or the phases of the individual antennas, or both. The main beam sharpens as the size of the array increases.

39.3 MAGNETIC DIPOLE RADIATION

Figure 39-7 shows a magnetic dipole that is similar to that of Fig. 37-4. As in that section, we set

$$a^3 \ll r^3 \quad \text{and} \quad a^2 \ll 2\lambda^2. \quad (39-29)$$

We already know that

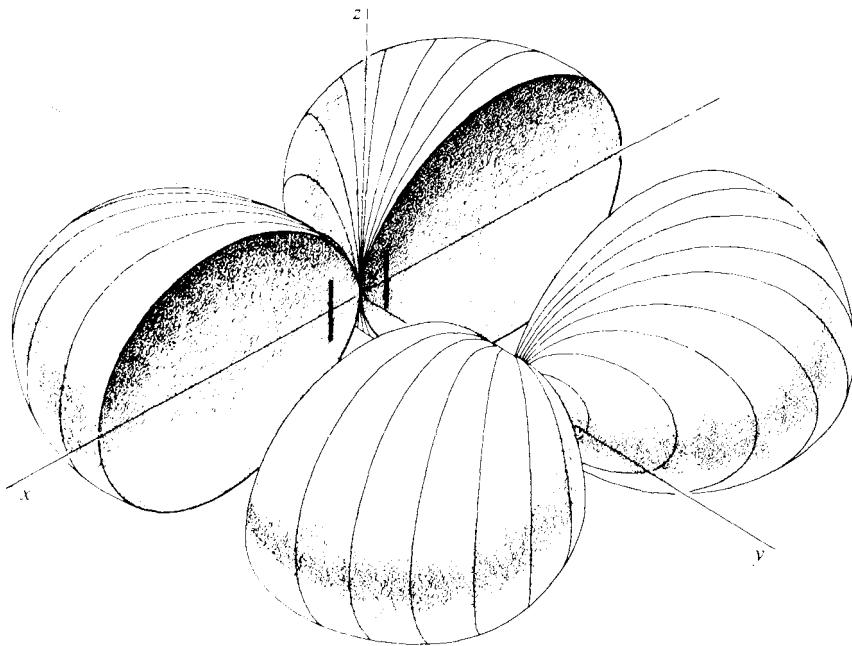


Fig. 39-6. The radiation pattern for the array of Fig. 39-4 with the antennas excited in opposite phases.

$$V = 0, \quad \mathbf{A} = j \frac{\mu_0[m]}{4\pi\lambda r} \left(1 - j \frac{\lambda}{r}\right) \sin \theta \hat{\phi}, \quad (39-30)$$

from the second example in Sec. 37.4.

39.3.1 The Electric Field Strength \mathbf{E}

Since $V = 0$, and since $\omega = c/\lambda$,

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} = -j\omega \mathbf{A} = \frac{\mu_0 c[m]}{4\pi\lambda^2 r} \left(1 - j \frac{\lambda}{r}\right) \sin \theta \hat{\phi}, \quad (39-31)$$

where $\mu_0 c \approx 377$ ohms. Thus \mathbf{E} is azimuthal.

At zero frequency, λ is infinite and \mathbf{E} is zero, as expected.

For $r \gg \lambda$,

$$\mathbf{E} = \frac{\mu_0 c[m]}{4\pi\lambda^2 r} \sin \theta \hat{\phi} \quad (r \gg \lambda). \quad (39-32)$$

Observe that \mathbf{E} is proportional to the time derivative of \mathbf{A} , hence to the time derivative of the current, and thus to the azimuthal acceleration

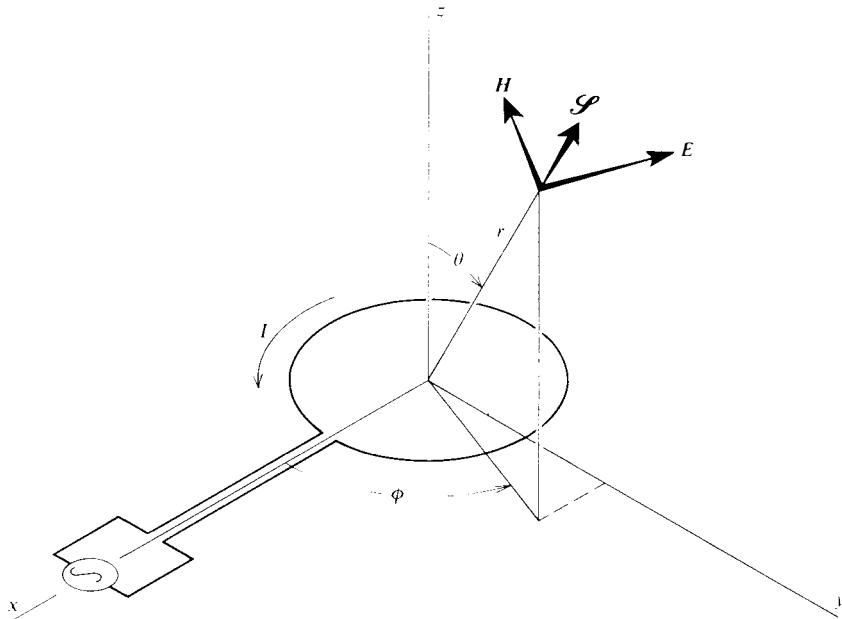


Fig. 39-7. Magnetic dipole antenna fed by an oscillator. The \mathbf{E} and \mathbf{H} vectors point in the directions shown when $\omega[t] = 0$. The Poynting vector $\mathbf{E} \times \mathbf{H}$ always points outward. We assume that $r \gg \lambda$.

of the electrons. The centripetal acceleration is negligible (Prob. 39-8).

The situation is different with high-energy electrons on a circular trajectory, as is a synchrotron accelerator. Then the electron speed is close to c and nearly constant, the acceleration is radial and the radiation is tangential. Because of the headlight effect (Prob. 14-3), the electrons radiate in the forward direction over a small solid angle. This is *synchrotron radiation*.

39.3.2 The Magnetic Field Strength \mathbf{H}

We have that

$$\mathbf{H} = \frac{\nabla \times \mathbf{A}}{\mu_0} = \frac{jm_m}{4\pi\lambda} \nabla \times \left\{ \frac{1}{r} \left(1 - j \frac{\lambda}{r} \right) \sin \theta \exp j\omega \left(t - \frac{r}{c} \right) \hat{\phi} \right\} \quad (39-33)$$

$$= \frac{[m]}{4\pi\lambda^2 r} \left\{ 2 \left(\frac{\lambda^2}{r^2} + j \frac{\lambda}{r} \right) \cos \theta \hat{r} + \left(-1 + \frac{\lambda^2}{r^2} + j \frac{\lambda}{r} \right) \sin \theta \hat{\theta} \right\}. \quad (39-34)$$

At zero frequency, $\lambda \rightarrow \infty$ and

$$\mathbf{H} = \frac{m_m}{4\pi r^3} (2 \cos \theta \hat{\mathbf{r}} + \sin \theta \hat{\theta}) \quad (f = 0), \quad (39-35)$$

as in the fourth example in Sec. 18.4. For $r \gg \lambda$,

$$\mathbf{H} = -\frac{[m]}{4\pi\lambda^2 r} \sin \theta \hat{\theta} \quad (r \gg \lambda). \quad (39-36)$$

39.3.3 The Poynting Vector, the Radiated Power, and the Radiation Resistance

The time-averaged Poynting vector for the magnetic dipole is

$$\mathcal{S}_{av} = \frac{1}{2} \operatorname{Re} (\mathbf{E} \times \mathbf{H}^*) = \frac{\mu_0 \pi^2}{c^3} \frac{m_{rms}^2 f^4}{r^2} \sin^2 \theta \hat{\mathbf{r}} \quad (39-37)$$

$$= 4.603 \times 10^{-31} \frac{m_{rms}^2 f^4}{r^2} \sin^2 \theta \hat{\mathbf{r}} \quad \text{watts/meter}^2. \quad (39-38)$$

The vector \mathcal{S}_{av} is radial, even close to the magnetic dipole. See Sec. 38.2.3.

Integrating over θ to obtain the radiated power, we find that

$$P = \frac{8\mu_0 \pi^3 m_{rms}^2 f^4}{3c^3} \quad (39-39)$$

$$\approx 3.8564 \times 10^{-30} m_{rms}^2 f^4 \quad \text{watts}. \quad (39-40)$$

For a circular loop of radius a and N turns,

$$m_{rms} = N\pi a^2 I_{rms}, \quad (39-41)$$

and

$$P = 3.806 \times 10^{-29} (N a^2 I_{rms})^2 f^4 \quad (39-42)$$

$$= 197.3 \left(\frac{a}{\lambda} \right)^4 (N I_{rms})^2 \quad (39-43)$$

$$\approx 200 \left(\frac{a}{\lambda} \right)^4 (N I_{rms})^2 \quad \text{watts}. \quad (39-44)$$

The radiation resistance is about $200(a/\lambda)^4 N^2$ ohms and is proportional to the *fourth* power of the frequency. We have assumed that $a^2 \ll 2\lambda^2$, and that $a^3 \ll r^3$.

39.3.4 Electric and Magnetic Dipole Radiation Compared

Compare Eqs. 38-16 and 38-23, for electric dipole radiation, with Eqs. 39-32 and 39-36, for magnetic dipole radiation. The \mathbf{E} and \mathbf{H} fields of the electric dipole are similar to the \mathbf{H} and \mathbf{E} fields of the magnetic dipole, respectively. However, the sign of \mathbf{E} for the magnetic dipole is opposite to that of \mathbf{H} for the electric dipole. This keeps the Poynting vector directed outward.

The lines of \mathbf{E} and of \mathbf{H} for the magnetic dipole are identical, respectively, to the lines of \mathbf{H} and of \mathbf{E} for an electric dipole at points remote from the dipoles.

The radiation resistance of an electric dipole is about $20(s/\lambda)^2$, while that of a magnetic dipole is about $200(a/\lambda)^4$.

39.4 THE ELECTRIC DIPOLE AS A RECEIVING ANTENNA

Figure 39-8(a) shows an electric dipole set up as a receiving antenna. The current induced in the dipole is detected at the receiver. If the input impedance of the receiver is infinite, the antenna generates a voltage that is equal to the tangential component of \mathbf{E} , multiplied by the length of the dipole.

39.5 THE MAGNETIC DIPOLE AS A RECEIVING ANTENNA

The electromotance induced in a loop antenna, under proper conditions, is simply equal to $d\Lambda/dt$, where Λ is the flux linkage in the loop.

The output voltage of the magnetic dipole is not necessarily equal to the induced electromotance, because the loop can also operate in the electric dipole mode. For example, with the loop of Fig. 39-8(b), if \mathbf{E} is horizontal and in the plane of the paper, charge oscillates from one end of the circuit to the other and the above relation is correct. However, if \mathbf{E} is vertical, an extra voltage comes from the electric dipole excitation and adds to the induced electromotance.

The induced electromotance is maximum when the axis of the dipole is parallel to the ambient magnetic field. Since the \mathbf{H} vector of a distant transmitter tends to be horizontal near the earth's surface and transverse to the direction of propagation, magnetic dipoles are useful for direction finding. The loop is set up with its axis horizontal and rotated about the

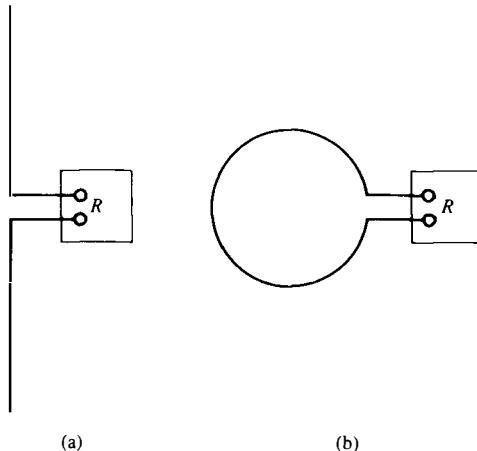


Fig. 39-8. An electric and a magnetic dipole, connected as receiving antennas. The radio receivers R measure the induced voltages.

vertical diameter. The signal is zero when the transmitter lies in the vertical plane containing the axis of the coil.

The antennas of small portable radios are usually coils wound on a rod of ferrite. The ferrite core amplifies the magnetic flux by a factor of about 100. The directivity of these antennas is well known. However, indoors, the antenna is close to many metallic objects that reradiate radio waves, and its directivity is meaningless.

39.6 SUMMARY

In the field of a *half-wave antenna*,

$$E = \frac{j}{2\pi c \epsilon_0 r} \frac{\cos \{(\pi/2) \cos \theta\}}{\sin \theta} [I] \hat{\theta}, \quad (39-12)$$

$$\mathbf{H} = \frac{j}{2\pi r} \frac{\cos \{(\pi/2) \cos \theta\}}{\sin \theta} [I] \hat{\phi}, \quad (39-14)$$

$$\mathcal{G}_{av} = \frac{1}{\pi c \epsilon_0} \frac{\cos^2 \{(\pi/2) \cos \theta\}}{\sin^2 \theta} \frac{I_{rms}^2}{4\pi r^2} \hat{r}, \quad (39-16)$$

$$= 9.543 \frac{\cos^2 \{(\pi/2) \cos \theta\}}{\sin^2 \theta} \frac{I_{rms}^2}{r^2} \hat{r} \quad \text{watts/meter}^2, \quad (39-17)$$

$$P = 73.083 I_{\text{rms}}^2 \quad \text{watts,} \quad (39-19)$$

where $[I]$ is the antenna current at the time $t - r/c$ and P is the *radiated power*. See Fig. 39-1 for the definitions of r and θ .

The *radiation resistance* is 73.08 ohms.

Arrays of antennas can be more directive than single antennas.

In the field on an oscillating *magnetic dipole*,

$$\mathbf{E} = \frac{\mu_0 c [m]}{4\pi\lambda^2 r} \left(1 - j \frac{\lambda}{r} \right) \sin \theta \hat{\phi}, \quad (39-31)$$

$$\mathbf{H} = \frac{[m]}{4\pi\lambda^2 r} \left\{ 2 \left(\frac{\lambda^2}{r^2} + j \frac{\lambda}{r} \right) \cos \theta \hat{r} + \left(-1 + \frac{\lambda^2}{r^2} + j \frac{\lambda}{r} \right) \sin \theta \hat{\theta} \right\}, \quad (39-34)$$

$$\mathcal{S}_{av} = \frac{\mu_0 \pi^2}{c^3} \frac{m_{rms}^2 f^4}{r^2} \sin^2 \theta \hat{r} \quad (39-37)$$

$$= 4.603 \times 10^{-31} \frac{m_{rms}^2 f^4}{r^2} \sin^2 \theta \hat{r} \quad \text{watts/meter}^2, \quad (39-38)$$

$$P = \frac{8\mu_0 \pi^3 m_{rms}^2 f^4}{3c^3} \quad (39-39)$$

$$= 3.856 \times 10^{-30} m_{rms}^2 f^4 \quad \text{watts} \quad (39-40)$$

$$= 197.3 \left(\frac{a}{\lambda} \right)^4 (NI_{rms})^2 \quad \text{watts}, \quad (39-43)$$

if $m = NI\pi a^2$, N being the number of turns and a the radius of the loop.

The *radiation resistance* is $197.3(a/\lambda)^4 N^2$ ohms.

An *electric dipole receiving antenna* feeding a high-resistance receiver generates a voltage equal to the tangential \mathbf{E} multiplied by the length of the dipole. A *magnetic dipole receiving antenna* under similar conditions generates a voltage $d\Lambda/dt$, where Λ is the linking flux. However, see Sec. 39.5.

PROBLEMS

- 39-1. (39.1) The radiation patterns of the electric dipole and of the half-wave antenna

- (a) Show that in the far field of an electric dipole $E_{rms} = 6.71 P^{1/2} \sin \theta / r$.
 (b) Show that for the half-wave antenna

$$E_{rms} = 7.02 \frac{P^{1/2} \cos \{(\pi/2) \cos \theta\}}{r \sin \theta}.$$

39-2. (39.1.4) The electric field of a radio antenna

Calculate E at a distance of 1 kilometer in the equatorial plane of a half-wave radio antenna radiating 1 kilowatt of power. Set $\lambda \ll 1$ kilometer.

39-3. (39.2) The image of a half-wave antenna

An antenna is normally situated near a conductor (the earth, an airborne vehicle, a satellite, etc.). Energy radiated toward the conductor is reflected, and the total field is thus the vector sum of the direct wave plus the reflected wave. It is convenient to consider that the latter is generated, not by reflection, but by an image of the antenna located behind the surface of the conductor.

(a) Show that the current in the image of a horizontal half-wave antenna and the current in the antenna flow in opposite directions.

(b) Show that the current in the image of a vertical half-wave antenna and that in the antenna flow in the same direction.

Both rules apply to oblique half-wave antennas.

(c) We have shown that the radiation resistance of a half-wave antenna is 73.1 ohms. Find the radiation resistance of a quarter-wave antenna perpendicular to a conducting plane.

39-4. (39.2) The radiation pattern of a linear array of half-wave antennas

A *linear array* consists of parallel half-wave antennas lying in a plane. Say there are N antennas, uniformly separated by a distance D and excited in phase.

(a) Show that, in the plane perpendicular to the antennas,

$$E \propto \frac{\sin \{(ND/2\lambda) \cos \phi\}}{\sin \{(D/2\lambda) \cos \phi\}},$$

where ϕ is the angle between the direction of observation and the plane of the array. The best approach is to sum the individual E phasors graphically in the complex plane.

(b) Find the angular positions of the minima and maxima of E . Differentiation yields only the maxima.

(c) Show that, for a given spacing D , the main lobe at $\phi = \pi/2$ becomes narrower as N increases.

(d) Draw a polar diagram of E as a function of θ between 0 and 360° for an array of 30 parallel half-wave antennas that are in phase and spaced by $\lambda/4$.

(e) Now plot the same function, using Cartesian coordinates, between 0 and 180° with a log scale for the E -axis.

(f) Explain why the main lobe is twice as wide as the two neighboring lobes.

(g) Show that its half-width (the angle between the maximum and the first minimum on one side or the other) is approximately equal to λ/l , where l is the length of the array.

39-5. (39.3) The directivity of an antenna

By definition, the *directivity* of an antenna is equal to the ratio of the Poynting vector at the maximum of the radiation pattern to the Poynting vector averaged over a spherical surface surrounding the antenna:

$$D = \frac{\mathcal{S}_{\max}}{P/(4\pi r^2)} = \frac{\mathcal{S}_{\max}}{\frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi \mathcal{S} \sin \theta d\theta d\phi}.$$

- (a) Show that, the directivity of an electric or magnetic dipole is 1.5.
 (b) Show that the directivity of a half-wave antenna is 1.64.

39-6. (39.3) The magnetic dipole antenna

Show that, for a given diameter and for a given mass of copper, the ratio $\omega L/R$ for a magnetic dipole antenna is independent of the number of turns, L and R being, respectively, the inductance and the resistance of the dipole.

Increasing the number of turns increases the resistance, the inductance, and the stray capacitance of the coil. The impedance is thus, in fact, a complicated function of the number of turns.

39-7. (39.3) The fields of electric and magnetic dipoles

You have two receiving antennas. One is an electric dipole of length l , and the other is a single-turn magnetic dipole of diameter l .

(a) Calculate the ratio of the induced voltages far away from an electric dipole and close by, in the equatorial plane.

(b) Repeat the calculation for the field of a magnetic dipole.

39-8. (39.3) The azimuthal and centripetal accelerations in the magnetic dipole; synchrotron radiation

We found that the radiation fields of electric and magnetic dipoles result from the accelerations of the electric charges.

Show that the centripetal acceleration is negligible in the oscillating magnetic dipole. Assume a copper torus of major radius R_1 and minor radius R_2 and a current $I_m \cos \omega t$. Assume also that the current is uniformly distributed over a thickness equal to the skin depth.

Set $R_1 = 200$ millimeters, $R_2 = 10$ millimeters, $I_m = 1$ ampere, $\rho = 1.3 \times 10^{10}$ coulombs/meter³ of conduction electrons, and $f = 1$ megahertz.

39-9. (39.4) The coupling between parallel electric dipoles

Figure 39-9 shows two parallel electric dipoles, one of which acts as a transmitting antenna and the other as a receiving antenna. The distance r is much larger than the lengths of the dipoles. How does the signal at the receiving antenna vary with the distance $r \gg \lambda$ and with the angle θ ?

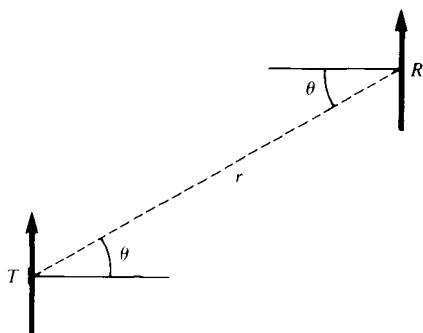
39-10. (39.5) The optimum load for a loop antenna

A loop receiving antenna of inductance L feeds a load resistance R . The resistance of the loop is negligible compared to R .

Show that there is maximum power transfer to the load when $R = \omega L$.

39-11. (39.5) The electric and magnetic dipole antennas

Compare the responses of electric and of magnetic dipole antennas used as receivers in seawater. Assume a frequency of 20 kilohertz, a typical frequency for communicating with submarines. Assume that the loop antenna has a single turn, that its diameter is equal to the length l of the electric dipole, and that $l < \lambda$ in seawater.

**Fig. 39-9.**

39-12. (39.5) Omnidirectional magnetic dipole antenna

Two identical magnetic dipoles are perpendicular and have a common diameter.

(a) Show that the radiation pattern is a circle in the plane perpendicular to the common diameter if one dipole leads the other by $\pi/2$ radians.

(b) Explain the nature of the resulting field.

(c) How would you connect these antennas to a common source?

Such a pair of crossed coils forms an *omnidirectional* transmitting or receiving antenna.

APPENDIX A

SI PREFIXES AND THEIR SYMBOLS

MULTIPLE	PREFIX	SYMBOL	MULTIPLE	PREFIX	SYMBOL
10^{-18}	atto	a	10	deka [†]	da
10^{-15}	femto	f	10^2	hecto	h
10^{-12}	pico	p	10^3	kilo	k
10^{-9}	nano	n	10^6	mega	M
10^{-6}	micro	μ	10^9	giga	G
10^{-3}	milli	m	10^{12}	tera	T
10^{-2}	centi	c	10^{15}	petra	P
10^{-1}	deci	d	10^{18}	exa	E

[†] This prefix is written *déca* in French.

Caution: the symbol for the prefix is written next to that for the unit *without* a dot. For example, mN stands for millinewton, while m · N is a meter-newton, or a joule.

APPENDIX **B**

CONVERSION TABLE

Examples: One meter equals 100 centimeters. One volt = 10^8 electromagnetic units of potential.

QUANTITY	SI	CGS SYSTEMS	
		esu	emu
Length	meter	10^2 centimeters	10^2 centimeters
Mass	kilogram	10^3 grams	10^3 grams
Time	second	1 second	1 second
Force	newton	10^5 dynes	10^5 dynes
Pressure	pascal	10 dynes/centimeter ²	10 dynes/centimeter ²
Energy	joule	10^7 ergs	10^7 ergs
Power	watt	10^7 ergs/second	10^7 ergs/second
Charge	coulomb	3×10^9	10^{-1}
Electric potential	volt	$\frac{1}{300}$	10^8
Electric field strength	volt/meter	$1/(3 \times 10^4)$	10^6
Electric flux	coulomb	$12\pi \times 10^9$	$4\pi \times 10^{-1}$
Electric flux density	coulomb/meter ²	$12\pi \times 10^5$	$4\pi \times 10^{-5}$
Polarization	coulomb/meter ²	3×10^5	10^{-5}
Electric current	ampere	3×10^9	10^{-1}
Conductivity	siemens/meter	9×10^9	10^{-11}
Resistance	ohm	$1/(9 \times 10^{11})$	10^9
Conductance	siemens	9×10^{11}	10^{-9}
Capacitance	farad	9×10^{11}	10^{-9}
Magnetic flux	weber	$\frac{1}{300}$	10^8 maxwells
Magnetic flux density	tesla	$1/(3 \times 10^6)$	10^4 gausses
Magnetic field strength	ampere/meter	$12\pi \times 10^7$	$4\pi \times 10^{-3}$ oersted
Magnetomotance	ampere	$12\pi \times 10^9$	$4\pi/10$ gilberts
Magnetization	ampere/meter	$1/(3 \times 10^{13})$	10^{-3}
Inductance	henry	$1/(9 \times 10^{11})$	10^{-9}
Reluctance	ampere/weber	$36\pi \times 10^{11}$	$4\pi \times 10^{-9}$

Note: We have set $c = 3 \times 10^8$ meters/second.

APPENDIX C

WAVES

If one disturbs a medium in some way, the disturbance travels outward as a *wave*. For example, a vibrating object generates acoustic waves in air.

C.1 PLANE SINUSOIDAL WAVES

Suppose the quantity a propagates at the velocity v_p in the positive direction of the z -axis. At $z = 0$ and for all x and y , it is given by

$$a = a_m \exp j\omega t. \quad (\text{C-1})$$

Then, at any position z ,

$$a = a_m \cos \left[\omega \left(t - \frac{z}{v_p} \right) \right]. \quad (\text{C-2})$$

This equation defines an *unattenuated plane sinusoidal wave*. The *wave fronts* are surfaces of uniform phase, at a given time. Here, the wave fronts are normal to the z -axis. This wave is also *uniform* because its *amplitude* a_m is uniform over a wave front.

At a given z , a is a sinusoidal function of t . At a given t , a is a sinusoidal function of z , as in Fig. C-1.

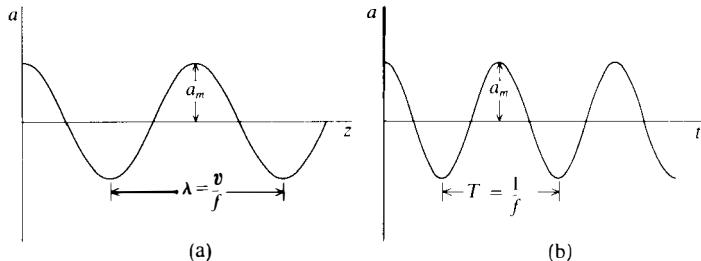


Fig. C-1. The quantity $a = a_m \cos(\omega t - kz)$ as a function of z and as a function of t .

The *phase* of the wave is the quantity between brackets. It is constant for $z = v_p t$. Hence v_p is the *phase velocity*.

It is usually more convenient to write

$$a = a_m \cos(\omega t - kz), \quad (\text{C-3})$$

where

$$k = \frac{\omega}{v_p} = \frac{2\pi f}{v_p} = \frac{2\pi}{\lambda} = \frac{1}{\bar{\lambda}} \quad (\text{C-4})$$

is the *wave number*. Note that this quantity is 2π times $1/\lambda$, or 2π times the wave number used in optics. For this reason, k is called the *circular wave number*.

The *wavelength* λ is the distance over which kz changes by 2π . The quantity $\bar{\lambda} = \lambda/(2\pi)$, read “lambda bar,” is often more convenient to use than λ . This is the *radian length*.

In phasor notation (Chap. 2),

$$a = a_m \exp j(\omega t - kz). \quad (\text{C-5})$$

If the wave travels in the negative direction of the z -axis,

$$a = a_m \exp j(\omega t + kz). \quad (\text{C-6})$$

If there is attenuation, then the wave amplitude decreases exponentially with z and

$$a = a_m \exp(-\alpha z) \exp j(\omega t - \beta z) \quad (\text{C-7})$$

$$= a_m \exp j[\omega t - (\beta - j\alpha)z]. \quad (\text{C-8})$$

The wave number is then complex:

$$k = \beta - j\alpha. \quad (\text{C-9})$$

Note the negative sign. Both α and β are positive:

$$\beta = \frac{1}{\delta}, \quad \alpha = \frac{1}{\delta}, \quad (\text{C-10})$$

where the *attenuation distance* δ is the distance over which the amplitude

decreases by a factor of $e = 2.71828$. The phase velocity is then

$$v_p = \frac{\omega}{k_r}. \quad (\text{C-11})$$

We often refer to waves traveling in a specified direction. Then the practice is to use a vector wave number

$$\mathbf{k} = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}} + k_z \hat{\mathbf{z}}, \quad (\text{C-12})$$

as in Fig. C-2, and

$$a = a_m \exp j(\omega t - \mathbf{k} \cdot \mathbf{r}) \quad (\text{C-13})$$

$$= a_m \exp j(\omega t - k_x x - k_y y - k_z z). \quad (\text{C-14})$$

The vector wave number can be complex.

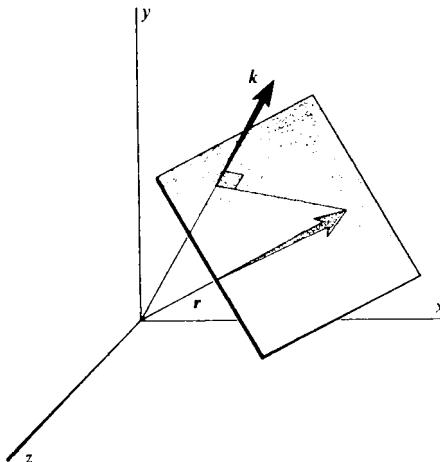


Fig. C-2. The vector wave number \mathbf{k} is normal to a wave front. The scalar product $\mathbf{k} \cdot \mathbf{r}$ has the same value anywhere on a given wave front.

C.2 THE PHASE AND GROUP VELOCITIES

Let us superpose two plane waves of angular frequencies ω and $\omega + \Delta\omega$ and of wave numbers k and $k + \Delta k$. The amplitudes are equal, and the phases are zero at $z = 0$, $t = 0$.

At z, t the phases will be, respectively, $\omega t - kz$ and $(\omega + \Delta\omega)t - (k + \Delta k)z$. The two waves are in phase when

$$(\Delta\omega)t - (\Delta k)z = 0 \quad (C-15)$$

or at points such that

$$\frac{z}{t} = \frac{\Delta\omega}{\Delta k}. \quad (C-16)$$

In other words, the points where the two waves are in phase travel at the *group velocity*

$$v_g = \frac{\Delta\omega}{\Delta k}. \quad (C-17)$$

The superposition of two waves of slightly different frequencies gives an amplitude-modulated wave. The group velocity of the envelope.

Remember that

$$k = \frac{1}{\lambda} = \frac{2\pi}{\lambda} = \frac{\omega}{v_p}. \quad (C-18)$$

If v_p is independent of ω , then k is proportional to ω and the phase and group velocities are equal:

$$v_g = \frac{\Delta\omega}{\Delta k} = \frac{\omega}{k} = v_p. \quad (C-19)$$

If the phase velocity is frequency-dependent, then the phase and group velocities are different. In other words, the wavelets in Fig. C-3 travel either faster or slower than the envelope, depending on the nature of the medium. (For waves at the surface of water, $v_p = 2v_g$.) In the limit,

$$v_g = \frac{d\omega}{dk} = \frac{1}{dk/d\omega}. \quad (C-20)$$

If one draws a curve of ω as a function of k , its *slope* is equal to v_g , while ω/k is equal to v_p .

The above equation is exact either if one has only two waves of circular

frequencies ω and $\omega + \Delta\omega$ or if k is a linear function of ω . In practice, neither condition is strictly true in matter, and the above definition of the group velocity is an approximation. After a while the group spreads out and disappears.

C.3 THE DIFFERENTIAL EQUATION FOR A PLANE SINUSOIDAL WAVE

You can easily check that, for the a of Eq. C-2,

$$\frac{d^2a}{dz^2} = \frac{1}{v_p^2} \frac{d^2a}{dt^2}. \quad (\text{C-21})$$

This is the differential equation for an *unattenuated* plane wave traveling along the z -axis.

If there is attenuation, then the differential equation is

$$\frac{d^2a}{dz^2} = g \frac{d^2a}{dt^2} + h \frac{da}{dt}, \quad (\text{C-22})$$

where

$$g = \frac{\alpha - \beta}{\omega^2}, \quad h = \frac{2\alpha\beta}{\omega}. \quad (\text{C-23})$$

Inversely

$$\beta = \omega \left(\frac{g}{2} \right)^{1/2} \left[\left(1 + \frac{h^2}{\omega^2 g^2} \right)^{1/2} + 1 \right]^{1/2}, \quad (\text{C-24})$$

$$\alpha = \omega \left(\frac{g}{2} \right)^{1/2} \left[\left(1 + \frac{h^2}{\omega^2 g^2} \right)^{1/2} - 1 \right]^{1/2}. \quad (\text{C-25})$$

C.4 WAVE PROPAGATION IN THREE DIMENSIONS

More generally, the differential equation for any *unattenuated* wave propagating in space at the phase velocity v_p is

$$\nabla^2 a = \frac{1}{v_p^2} \frac{d^2a}{dt^2}. \quad (\text{C-26})$$

For an attenuated wave,

$$\nabla^2 a = g \frac{\partial^2 a}{\partial t^2} + h \frac{\partial a}{\partial t}, \quad (\text{C-27})$$

and Eqs. C-23 to C-25 apply.

C.5 THE WAVE PROPAGATION OF A VECTOR QUANTITY

A vector quantity, such as an electric field strength \mathbf{E} , can also propagate as a wave. If the wave is uniform and plane and if it propagates in the positive direction of the z -axis, then

$$\mathbf{E} = \mathbf{E}_m \exp j(\omega t - kz) \quad (\text{C-28})$$

$$= (E_{mx} \hat{x} + E_{my} \hat{y} + E_{mz} \hat{z}) \exp j(\omega t - kz). \quad (\text{C-29})$$

The vector \mathbf{E}_m may depend on x and on y , but it does *not* depend on z : the only dependence on z and t appears in the exponential function. The wave number k can be complex.

C.6 THE NONHOMOGENEOUS WAVE EQUATION

In the absence of attenuation, the *nonhomogeneous wave equation* is of the form

$$\nabla^2 a - \frac{1}{v_p^2} \frac{d^2 a}{dt^2} = f(x', y', z', t), \quad (\text{C-30})$$

where f is the disturbance at the source. Thus $f = 0$ outside the source.

If f is not a function of t , then there is no wave and

$$\nabla^2 a = f, \quad (\text{C-31})$$

which is *Poisson's equation*, and

$$a(x, y, z) = -\frac{1}{4\pi} \int_{v'} \frac{f(x', y', z')}{r} dv', \quad (\text{C-32})$$

where r is the distance between the field point (x, y, z) and the source point (x', y', z') .

If f is a function of t , then one must take into account the time taken by the disturbance to travel from (x', y', z') to (x, y, z) at the velocity v_p and

$$a(x, y, z, t) = -\frac{1}{4\pi} \int_{v'} \frac{f(x', y', z', t - r/v_p)}{r} dv'. \quad (\text{C-33})$$

ANSWERS

Problems often require the demonstration of a given result; the list below provides about half of the remaining answers, to two significant figures.

1-5. $4\pi a^3$.

2-7. 85 watts.

3-2. 440 millimeters.

3-4. (b) 3.3×10^{-3} newton, (d) 6 microseconds.

3-6. (a) (i) 10^{19} atoms/meter², (ii) 1.6 coulombs/meter²,
(iii) 1.8×10^{11} volts/meter. (b) 240 atomic diameters, or 0.07 micrometer.

3-9. (a) $\frac{Q\lambda}{2\pi\epsilon_0 r}$.

3-13. (a) $\frac{GMm}{R} - \frac{Qq}{4\pi\epsilon_0 R}$.

4-1. 1.5×10^{20} newtons. (b) 3.6×10^{22} newtons, or 200 times more.

4-10. $E_p - I \frac{s}{\mathcal{A}\sigma}$.

4-13. 1.9 newtons.

5-3. $\frac{4}{3}\pi R^3 \sigma_0 \hat{z}$.

5-7. $\frac{Q}{4\pi\epsilon_0 r}$, $\frac{Qs \cos \theta}{4\pi\epsilon_0 r^2}$, $Qs^2 \left(\frac{3 \cos^2 \theta}{2} - \frac{1}{2} \right) / 4\pi\epsilon_0 r^3$.

5-9. $\frac{\rho a^3}{4\pi\epsilon_0 r}$, 0, 0.

5-11. 0, $\frac{a^4 Q}{16\pi\epsilon_0 r^5} [35(l^4 + m^4 + n^4) - 21]$.

6-2. (a) $\frac{0.15Q^2}{\pi\epsilon_0 R}$. (b) $\frac{3GM^2}{5R}$. (c) 1.2×10^{29} joules. (d) 0.17 meter.
(e) 1.0×10^{20} volts.

6-5. (b) 56 picofarads/meter.

6-10. (a) 150 kilovolts. (b) 4×10^{-4} atmosphere. (c) 4 kilograms/meter².

7-2. (b) 3000.

7-12. $R_3\omega C_3 = 1$.

8-1. 0.07 ohm.

$$8-5. \frac{V}{4R + 3R_0} R_0.$$

8-10. 0.44.

8-12. 2.9 kilohms.

9-1. (a) 5.7×10^{-37} coulomb-meter. (b) 5.9×10^{-19} meter.

$$9-8. (a) \frac{2\pi\epsilon_r\epsilon_0}{\ln(R_2/R_1)} V.$$

$$10-2. (a) \frac{R_1/(j\omega C_1)}{R_1 + 1/(j\omega C_1)} + \frac{R_2/(j\omega C_2)}{R_2 + 1/(j\omega C_2)}.$$

$$(b) \frac{(\epsilon_{r2}\sigma_{co1} - \epsilon_{r1}\sigma_{co2})\epsilon_0}{s_1\sigma_{co2} + s_2\sigma_{co1}} V.$$

10-4. 4000.

10-9. (b) 10 meters³, 10 tons.

10-13. 25 micrometers.

$$11-2. \frac{-2QD}{4\pi(D^2 + r^2)^{3/2}}.$$

$$12-6. \frac{\rho_0}{12\epsilon_0 r^2} (\alpha^2 + \beta^2)(\alpha + \beta).$$

13-3. According to O' the signals were emitted simultaneously. According to O , B emitted her signal first.

13-4. (a) $\tan \alpha = \gamma \tan \alpha'$. (b) $\pi/2$.

14-2. -1.2×10^9 meters, 0.8c.

14-5. $\$2.16 \times 10^8$.

14-8. (a) 4.7×10^4 electronvolts. (b) 9.33×10^8 electronvolts.

15-2. (a) 7.1×10^{-26} kilogram. (b) 38 millimeters. (c) 10^{-5} second,
 1.28×10^{-10} second.

15-6. 2.6×10^{-27} kilogram, 2.27×10^8 meters/second.

16-1. -4×10^{-6} .

16-5. (a) 2.3×10^{-22} newton. (b) 7.7×10^{-23} newton.

17-3. $-\gamma^2 \frac{\sigma'}{\epsilon_0} (x - \gamma t)$, $-\gamma^2 \frac{\gamma \sigma'}{\epsilon_0 c^2} (x - \gamma t)$, $\frac{\sigma'}{\epsilon_0} \hat{x}$, 0.

17-5. Inside, 0, $\mu_0 N' I \hat{x}$; outside, 0. 0.

18-2. $\frac{\mu_0 I}{2\pi} \left(\frac{1}{D-x} + \frac{1}{D+x} \right)$.

18-5. (a) $\frac{eru}{2}$. (b) 9.3×10^{-24} ampere-meter².

18-9. $\mu_0 N I a^2 \left\{ \frac{1}{2[(z+a/2)^2 + a^2]^{3/2}} + \frac{1}{2[(z-a/2)^2 + a^2]^{3/2}} \right\}$.

19-2. (a) 330 microamperes, (b) 4.2×10^{-10} tesla.

19-4. $\frac{\mu_0 \alpha}{2}$.

19-7. $\frac{L}{R_2 - R_1} \ln \dots$

20-1. 2.

20-5. $\frac{B}{\mu_0}$, B .

21-1. 0.5 tesla.

21-3. 0.14 tesla, 0.085 tesla.

22-1. (a) 3×10^{13} meters. (b) 7.3 days.

22-3. (b) $2.2 \times 10^{18} \hat{y}$ meters/second².

22-6. 1.2, 2.7.

22-8. (c) 1.4×10^{-4} volt.

22-11. $\frac{\mu_0 I^2}{2\pi D} [(4L^2 + D^2)^{1/2} - D]$.

22-14. (a) $\frac{AV}{R} \omega - \frac{A^2}{R} \omega^2$.

23-3. (a) Charge flows to cancel the $v \times B$ field. (b) 31 amperes. (c) 1.5×10^{-2} newton. (d) For a diameter of 10 meters, $I = 30$ microamperes, $P \approx 1$ microwatt, $F \approx 2 \times 10^{-10}$ newton.

23-8. $P_{\max} = \frac{\pi \sigma \omega^2 B_{\text{rms}}^2 \delta a^4}{8}$.

23-10. (a) $B_0 \omega^2 \cos(2\omega t - \phi)$. (b) 2 kilohertz.

24-1. (a) $N\pi R^2 \mu_0 N'$. (b) No if its diameter is much smaller than the length of the solenoid.

24-3. $3.4 \times 10^{-7} R$ henry.

24-6. (a) $R' = \frac{2\pi a}{\sigma b}$, $L' = \mu_0 \pi a^2$. (b) $\left(\frac{4+x^2}{4+4x^2}\right)^{1/2}$, $x = \omega \mu_0 \sigma a b$.

24-8. $L = R_b R_d C$, $R = R_b R_d / R_c$.

24-10. (a) 6 ohms, 0.86 radian, 3.9 ohms, 4.6 ohms. (b) 65 amperes, 76 amperes. (c) 336 microfarads. (d) 65 amperes.

25-1. (a) $R\left(1 + \frac{1}{R^2 \omega^2 C + 1}\right) + j\left(\omega L - \frac{R^2 \omega C}{R^2 \omega^2 C^2 + 1}\right)$. (b) 37 ohms, 57.5° .

(c) 2.7×10^{-2} siemens, -58° . (d) 0.2 watt. (e) No. (f) (i) At all frequencies. (ii) Never. (g) 0.

25-7. $\frac{Z_1 Z_2 + \omega^2 M^2}{Z_1 + Z_2 + 2j\omega M}$.

25-12. (a) $\frac{8N^3 a}{\sigma_c l^2}$. (b) $\frac{8N^3 a}{\sigma_c l^2} + j\omega \frac{\mu_0 N^2 \pi a^2}{l}$. (c) $\frac{2\pi a}{\sigma_b b l}$. (d) $L_2 = \frac{\mu_0 \pi a^2}{l}$,

$Z_2 = \frac{2\pi a}{\sigma_b b l} + j\omega \frac{\mu_0 \pi a^2}{l}$. (e) $\frac{8N^3 a}{\sigma_c l^2} + j\omega \mu_0 N^2 \pi a^2 \frac{1}{l} + \frac{\mu_0^2 \omega^2 N^2 \pi a^3 \sigma_b b}{2l + j\omega \mu_0 \sigma_b b a l}$.

(f) $69 + j50$ ohms, $91 + j36$ ohms.

25-14. (a) $\frac{L}{C} = \frac{R^2}{1 + R^2 \omega^2 C^2}$. (b) $1 + R^2 \omega^2 C^2$. (c) 770 picofarads, 150 millihenrys.

26-6. 2.9×10^4 ampere-turns.

26-9. (a) 22 kilowatt-hours/meter³. (b) 780 atmospheres. (c) 0.083 kilowatt-hour/meter³.

26-12. (a) 4 atmospheres. (b) The same.

26-16. (c) $-\frac{R}{2} \frac{\partial B_z}{\partial z}$, $m \frac{\partial B_z}{\partial z}$. (d) $\frac{I\pi R^2 dB_z}{dz}$.

27-1. (e) 1.2×10^{-8} meter.

27-3. (a) $\frac{\epsilon_0 k b}{s} (a - x)$. (b) $\mathbf{B} = \frac{\epsilon_0 \mu_0 k}{s} (a - x) \hat{\mathbf{y}}$, $\mathbf{A} = \frac{\epsilon_0 \mu_0 k}{s} (a - x) z \hat{\mathbf{x}}$.

(c) 0. Above the upper plate, $\frac{(\epsilon_0 \mu_0 k/s)(a - x)s}{2} \hat{\mathbf{x}}$. Below the upper plate,

\mathbf{A} has the same magnitude and the opposite sign.

27-6. (a) $\frac{\sigma_{co}RC}{\epsilon}$.

27-10. (a) $\frac{\mu_0 Q \mathcal{V}}{4\pi r^2} \sin \theta$. (b) Set $\gamma = 1$, $\mathcal{V}/c = 0$. The two values agree. (d) Q .

28-4. (a) 5.0×10^{-7} meter, 1.5×10^{-8} meter. (b) 2.5×10^{-7} meter.

28-7. $250 + 6.3 \times 10^{-3}j$ ohms.

28-11. (a) (i) $\frac{I^2}{16\pi\epsilon_0 v^2} \left(1 + 4 \ln \frac{R_2}{R_1}\right)$. (ii) $\frac{v^2}{c^2} \mathcal{E}'_E$. (iii) $2\mathcal{E}'_E v$. (iv) $\frac{Imv^2}{2e}$, or IV.

(b) (i) 2.0×10^{-10} joule/meter. (ii) 4.2×10^{-13} joule/meter.
(iii) 5.5×10^3 watt. (iv) 10^3 watt.

29-3. Use 6 vertical plates, about 6 millimeters thick and 6 millimeters apart, with a total width of 1 meter.

29-6. 5×10^{-5} meter; 7.2×10^6 watts/meter².

29-9. (a) $\frac{c}{(f_z^2 - f_p^2)^{1/2} - (f_1^2 - f_p^2)^{1/2}}$. (b) $\frac{\omega_1 - \omega_2}{(\omega_1/v_{p1}) - (\omega_2/v_{p2})}$. (d) $v_{p1} = 3.2 \times$

10^8 meters/second, $v_{p2} = 3.2 \times 10^8$ meters/second, $v_g = 2.8 \times 10^8$ meters/second. (e) 2.8×10^3 meter, 45°.

29-11. (b) 270 parsecs.

30-2. $\frac{E_{Rm}}{E_{Im}} \approx -1$ and $\left(\frac{E_{Tm}}{E_{Im}}\right)_\parallel \approx \frac{2n_1}{n_2}$, except near $\theta_I = 90^\circ$;

$\left(\frac{E_{Tm}}{E_{Im}}\right)_\perp \approx \frac{2n_1}{n_2} \cos \theta_I$ for any θ_I .

30-9. (c) 0.15.

30-13. (a) $E_{I rms} = 42$ volts/meter, $H_{I rms} = 0.11$ ampere/meter, $E_{R rms} = 6.0$ volts/meter, $H_{R rms} = 1.6 \times 10^{-2}$ ampere/meter, $E_{T rms} = 36$ volts/meter, $H_{T rms} = 0.17$ ampere/meter.

30-16. (a) $\frac{c}{K(A+R)}$, where K is a constant, A is the altitude, and R is the radius of the earth. (b) $\omega \left[1 - \frac{c^2}{K^2(A+R)^2}\right]^{1/2}$.

31-2. (a) $\frac{\delta_z}{\lambda_2} = 0.37$, $\frac{\delta_x}{\lambda_1} = 1.1$. (b) 150° , 74° .

31-8. (a) 17° . (e) $F \approx 0.020$, $T \approx 0.69$, $FT \approx 1.4\%$.

32-4. (a) 3.7×10^{-8} meter. (b) 0.95. (c) 13 ohms/square. (d) $3.1 \times 10^{-3} \lambda_0$.

32-8. E parallel to the plane of incidence and θ_I equal to the Brewster angle.

32-10. $\frac{8}{3} \frac{R\mathcal{S}}{c}$.

32-12. $\frac{[1 - \exp(-al)]}{Nec}$.

33-7. $\left(\frac{\epsilon}{\mu}\right)^{1/2} \frac{\mathcal{V}_{\text{rms}}^2 b}{h}$.

34-1. 3.0 gigahertz.

34-3. $58^\circ, 37^\circ$.

34-5. (a) E must be parallel to the plates. (b) $\left(1 - \frac{\lambda_\theta^2}{4s^2}\right)^{1/2}$.

34-7. (a) $(\tilde{\chi}_0/\tilde{\chi}_z)^2 + (\tilde{\chi}_0/\tilde{\chi}_c)^2 = 1$. (b) $\frac{\omega \tilde{\chi}_c}{c} = \{(k_z \tilde{\chi}_z)^2 + 1\}^{1/2}$.

34-11. (a) 3.1 megawatts. (b) $P'_{L_{\text{coax}}} = 8.2$ kilowatts/meter, $P'_{L_{\text{guide}}} = 12$ kilowatts/meter. (c) 130 amperes, 4.0 kilovolts.

36-2. $\frac{4a}{(n_2^2 - n_3^2)^{1/2} m}$.

36-7. (a) 10 gigawatts/meter². (b) 1.6×10^6 volts/meter.

37-2. $c \left(1 + \frac{\tilde{\chi}^2}{r^2}\right)$, c.

37-9. $\frac{2}{3}(\mu_0 K \omega s^3 R^2)$.

38-1. (a) $-\frac{K\mu_0}{2\pi} \ln\left(\frac{C}{\rho}\right) \hat{z}, \frac{K\mu_0 t}{2\pi\rho} \hat{\phi}$.

38-4. 88%.

39-2. 0.22 volt/meter.

39-7. (a) Far away, $6.1 \times 10^7/lf$; close by, $4.3 \times 10^7/lf$. (b) Far away, $6.1 \times 10^7/lf$; close by, $8.6 \times 10^7/lf$.

39-11. $\frac{\mathcal{V}_{\text{el}}}{\mathcal{V}_{\text{mag}}} = \frac{1.43}{l}$.

INDEX

- Accelerator, 208, 280, 435
Admittance, 134, 152
Alternating currents, 32, 127–139, 147
Ampere, 44, 398
Ampère's circuital law, 352, 367, 378, 496
Ampere-turn, 353
Amperian formulation, 494
Amplifier, operational, 141, 144, 145
Angle
 complex, 580
 of incidence, 555, 581, 589
Antenna
 arrays, 717, 727
 directivity, 727
 electric dipole. *See* Electric dipole,
 oscillating
 half-wave, 712–717, 726
 image, 727
 magnetic dipole. *See* Magnetic dipole,
 oscillating
 receiving, 724
Antireflection coatings, 575
Askin, A., 603
Attenuation distance, 524, 580, 733
- Bartlett, D. F., 499
Battery, 82, 151, 167
Bertoni, H. L., 587
Biot–Savart law, 327, 334, 364
Bohm, David, 270
Bohr magneton, 344, 373
Bondi, H., 63, 511
Boundary conditions
 for a metallic waveguide, 618
 for \mathbf{B} and \mathbf{H} , 370
 for \mathbf{E} and \mathbf{J} , 207
 for V , \mathbf{D} , and \mathbf{E} , 197
Brackets, use of, 682, 698, 714
Branch, 120
Brewster angle, 565–567, 574–575, 589
Bridge circuit, 146, 167, 450
- Capacitance, 108–110, 117, 403
Capacitor, parallel-plate. *See* Parallel-plate capacitor
Cathode-ray tube, 588
Causality, 247
Chambers, L. G., 513
Charge. *See also* Electric field
 center of, 87, 173
 conservation, 69, 290, 314
 free, embedded in a dielectric, 186
 invariance of, 284, 290, 306
 magnetic. *See* Monopole, magnetic
Charge density
 at an interface, 80
 free and bound, 174, 182, 183
 in a conductor, 75, 402, 409
 total, 178, 233
Charged particle in \mathbf{E} , \mathbf{B} , 406
Child–Langmuir law, 68
Circuit, electric
 active, 121
 bridge, 146, 167, 450
 delta-star transformation, 163
 differentiating, 143–144
 integrating, 144–145, 422
 linear, 121, 149
 mesh method, 122
 node method, 126
 passive, 121
 Q , 454
 RC , 124, 135, 138, 119–171
 R_L , 467
 R_L , RLC , 452–470
 symbols and definitions, 120
 theorems, 149–167
Circuit, magnetic. *See* Magnetic circuit
Circuital law. *See* Ampère's circuital law
Circular frequency, 32, 130
Clock, time read on a rapidly moving, 255
Coaxial line, 60, 190, 477, 488, 532,
 619–621, 624, 645
Coefficient of coupling, 446, 467–468

- Coercive force, 376
 Collimator paradox, 264
 Comet tails, 603, 607
 Complementary solution, 125
 Complex numbers, 30–38
 Conductance, 135
 Conduction, 69, 72–77, 80, 82
 current, 287, 306
 Conduction electrons, 60, 72–82, 310
 drift velocity, 72, 80, 300
 Conductivity, 70
 ground, anomalies in, 160
 of a plasma, 542
 Conductor, 64, 77. *See also* Electric field;
 Magnetic field
 electric field at surface of, 110
 electric force on, 110
 good, 500
 hollow, 77
 Conjugate, complex, 31
 Conservation laws for colliding particles, 276, 281
 Conservation of charge, 64, 69, 290, 314, 500, 676
 Constant of integration, 125
 Constants, physical. *See inside the back cover*
 Contact potential, 77
 Continuity, conditions. *See Boundary conditions*
 Continuous creation theory, 511, 513
 Coordinates
 Cartesian, 2, 19
 cylindrical, 17–24
 orthogonal curvilinear, 15–17
 spherical, 18–24, 225
 Corle, P. R., 499
 Coulomb, 43, 398
 Coulomb's law, 42, 51, 520
 Coupling, coefficient of, 446, 467–468
 Crab nebula, 405
 Crack detector, 469
 Curl, 11, 22, 290
 of \mathbf{A} , 333
 of \mathbf{B} , 312–314, 352
 of \mathbf{E} , 48, 311, 314, 420, 499
 of \mathbf{H} , 365, 499
 Current
 conduction, 69, 306
 convection, 306
 displacement, 186, 499
 eddy, 425, 432–433, 436
 equivalent, 313, 363, 373
 mesh, 122
 polarization, 176
 Current density, 67–72, 329, 676
 displacement, 186, 499
 equivalent, 313, 363, 373
 four-current density, 285
 polarization, 176, 313
 total, 316
 Current source, 120, 152
 Curve plotters, 141
 Cyclotron frequency, 405–406

 \mathcal{D} , 525, 537
 D'Alembertian, 290
 Decibel, 640, 645
 Del operator, 4, 6, 19
 Delta-star transformation, 163–165, 467
 Depth of penetration, 507, 538. *See also*
 Skin depth
 Diamagnetism, 361
 Dielectric constant. *See* Permittivity,
 relative
 Dielectric strength, 207, 644
 Dielectrics, 172–210
 and magnetic materials compared, 366
 anisotropic, 195
 artificial, 643
 lossy, 193
 nonhomogeneous, 190
 nonlinear, 197, 207
 polarization of, 173
 Digital-to-analog (D/A) conversion, 142
 Diode, vacuum, 66–69
 Dipole. *See* Electric dipole, Magnetic dipole
 Dirac, P. A. M., 236, 327
 Direction cosines, 94
 Displacement current density, 186, 499
 Divergence, 8, 20
 four-dimensional, 289
 of \mathbf{A} , 351, 676
 of \mathbf{B} , 310, 314, 333, 364, 498
 of \mathbf{D} , 179
 of \mathbf{E} , 151, 309, 314, 498
 of \mathbf{H} , 373
 of \mathbf{J} , 69, 500, 676
 Divergence theorem, 9
 Domain, electric, 196
 Domain, magnetic, 374
 Doppler effect, 257, 262, 281, 283
 Duality, 501
 Ducting, 577
 Duinker, Simon, 154

 Earnshaw's theorem, 54
 Earth's electric charge, 60, 212

- Eddy currents, 425, 432–433, 436, 461, 468
 Edwards, T. C., 621
 Eigenvalues, 629
 Einstein, Albert, 254
 Electret, 177, 180, 208, 379
 Electric circuit. *See* Circuit, electric
 Electric dipole, 84–87, 93, 98, 173, 501
 oscillating, 681–684, 690, 697–708,
 710
 Electric displacement, 186. *See also* Electric
 flux density **D**
 Electric energy, 101–107, 118, 488–489
 associated with polarization, 203
 in a wave, 521, 523, 527, 539
 in terms of **E**, 105, 107
 in terms of **E** and **D**, 202
 in terms of ρ and V , 101, 200
 of a charged conducting sphere, 105
 of a continuous charge distribution, 104,
 116
 Electric field
 at the surface of a conductor, 110
 average **E** over a spherical volume, 56
 average V over a spherical surface, 54
 of a charge embedded in dielectric, 186
 of a conducting cylinder in a uniform **E**,
 235
 of a conducting sphere in a uniform **E**,
 228
 of a dielectric sphere in a uniform **E**, 231
 of a polarized dielectric, 176, 189
 of a spherical charge, 52, 65, 234
 of an atomic nucleus, 236
 of an electric dipole. *See* Electric dipole
 of macroscopic bodies, 49
 Electric field strength **E**, 44, 49, 56, 63, 89,
 293, 294, 430, 686–687, 690–691
 and the $v \times \mathbf{B}$ field, 427
 curl of, 46, 311, 314, 420, 430
 divergence of, 51, 309
 flux of, 50
 in terms of V and **A**, 316, 424
 induced, 413, 420
 inside a dielectric, 176
 integrals for, 234, 236
 line integral of, 46, 198, 430, 495
 maximum. in air, 118
 Poisson's equation for, 233
 transformation of, 301
 Electric flux density **D**, 179
 divergence of, 179
 lines of, 232
 Electric force, 61, 293
 and lines of **E**, 112, 117
 in presence of dielectrics, 204, 209–210
 on conductors, 110, 114, 209–210
 Electric polarization, 173
 Electric potential V , 46–49, 54, 89
 at the surface of a star, 62
 average, over a spherical surface, 54
 gradient, 47, 316, 424
 Poisson's equation for, 65, 189, 198, 233
 Electric susceptibility, 181
 Electrolytic tank for plotting magnetic
 fields, 386
 Electromagnet, 383, 450
 Electromagnetic momentum, 602, 604, 609
 Electromagnetic potentials, 314–322, 324
 retarded, 680–686, 689
 Electromagnetic waves, 514–729
 guided. *See* Guided waves
 in a general medium, 515–520
 in conductors, 524, 526, 531, 537–542
 in free space, 520–522
 in plasmas, 542–549
 in nonconductors or poor conductors,
 522, 526, 531
 nonuniform, 579–581, 591
 polarization of, 517, 527
 propagating in a straight line, 611–618
 spectrum, 515
 standing, 598
 TE, 614
 TEM, 615–618, 624
 uniform, 515, 520, 522
 Electromotance, 413, 426
 Electron
 drift velocity inside a conductor, 72, 80
 emission, 66–69
 mass, effective, 74, 408
 pair formation, 275
 Energy. *See also* Electric energy; Magnetic
 energy
 kinetic, 276
 relativistic, 273–275, 294
 rest, 274
 Energy storage
 electric, 207, 489
 magnetic, 472, 489
 Energy theorem, 487
 Equations, differential, 33, 35, 124
 Equations, six key, 430
 Equipotential surface, 50, 86, 230, 232, 318,
 323
 Equivalent currents, 363, 373
 Fabry equation, 345
 Farad, 108

- Faraday induction law, 414, 420, 428, 495, 499
 Faraday, M., 409
 Fermi acceleration, 407
 Ferrites, 425, 725
 Ferroelectricity, 196
 hysteresis, 197
 Ferromagnetism, 361, 368, 374–386
 hysteresis, 375–377, 423
 Feynman, Richard P., 334
 Field, 2. *See also* Electric field; Magnetic field
 conservative, 10, 15, 46
 of a point charge moving at a constant velocity, 291–301, 306, 314–319, 512
 Flaws in metal tubing, 433
 Flowmeter, 409
 Fluorescent lamp, 451
 Flux, 7
 of \mathbf{B} , 312, 329
 of \mathbf{E} , 50
 Flux compression, 490
 Flux, leakage, 381–382
 Flux linkage, 349, 421
 Force. *See also* Electric force; Magnetic force
 Coulomb, 43, 292
 gravitational, 44
 inertial, 239
 Lorentz, 293, 326
 relativistic, 270, 281
 Four-current-density, 285, 320
 Four-divergence, 289
 Four-gradient, 289
 Four-momentum, 271–273, 276, 287, 320
 Four-potential, 319–323
 Four-vector, 265–270
 Fourier series, 219, 228
 Frequency, 32
 circular, 32, 130
 cyclotron, 405, 406
 radio and television, 548
 Fresnel's equations, 558–561, 573, 584
 Friedman, F., 255
 Frish, D., 255
 Functions
 complete set of, 220
 continuously differentiable, 10
 harmonic, 215
 reasonably well-behaved, 220
 Galilean transformation, 224
 Galilei, Galileo, 241
 Gallium arsenide, 74
 Garden-hose effect, 533
 Gas focusing, 405
 Gauss's law
 for \mathbf{B} , 311, 495, 498
 for \mathbf{E} , 50–57, 178, 309, 495, 498
 Gauss's theorem, 10
 Generator, 117, 414–419
 homopolar, 403
 magnetohydrodynamic (MHD), 389–392
 Geophysical prospection, 160
 Gibbs phenomenon, 221
 Good, R. H., 705
 Goos-Haenchen shift, 587, 670
 Gradient, 3, 20
 four-dimensional, 289
 of V , 47, 316, 424
 Gravitational force, 44
 Green, Estill I., 454
 Green's theorem, 10
 Guided waves, 610–674. *See also* Coaxial line; Hollow rectangular waveguide; Microstrip line; Planar optical waveguide; Transmission lines
 Hall effect, 388, 393, 408, 416, 599
 Harmonic functions, 215
 spherical, 225–233
 Haus, Hermann A., 307, 493
 Headlight effect, 263
 Heat conduction, 537, 550
 Heaviside, Oliver, 492
 Henry, 438
 HILS, homogeneous, isotropic, linear, and stationary medium, 181
 Hodoscope, floating wire, 394
 Hole, 70, 82
 Hollow rectangular waveguide, 627–645
 attenuation, 637, 640, 644
 cut-off wavelength, 630
 energy density, 637
 field components, 628
 multiple reflections, 633
 phase, signal, and group velocities, 635
 transmitted power, 636, 644
 Homopolar generator, 403
 Homopolar motor, 399–403, 411
 Hubble constant, 512
 Hurd, C. M., 393
 Hysteresis
 ferroelectric, 197, 207
 ferromagnetic, 375–377, 423, 468

- Images, 212–215, 727
 Impedance, 134, 458
 bridges, 146, 155
 characteristic, 518, 520, 523, 526, 531–
 532, 538, 540
 matching, 465, 470
 of an inductor, 443, 453
 output, 150, 161
 surface, 551, 606. *See also* Resistance,
 surface
 wave, 614, 615
 Incidence, plane of, 557
 Index of refraction, 523, 531, 569, 573, 576
 Inductance
 mutual, 437–441, 448–449, 457, 488
 self-inductance, 441–445, 450, 453, 461,
 477
 Induction heating, 551
 Interface. *See* Boundary conditions
 Internal reflection spectroscopy, 581
 Invariance, 3, 6, 243
 of a physical law, 241
 of a quantity, 243
 of $B^2 - E^2/c^2$ and $\mathbf{E} \cdot \mathbf{B}$, 303, 304, 307
 of electric charge, 284, 290, 306
 of the speed of light in a vacuum, 244,
 260
 Ion beam, 61
 Ion thruster, 59
 Ionized gases. *See* Plasma
 Ionosphere, 548, 577

 Joule effect, 76, 540
 $j\omega$, 33

 Kirchhoff's laws, 121, 138, 453, 457
 Kittel, Charles, 285
 $K\chi$ surface, 705

 Ladder network, 142
 Lambda bar χ , 524, 733
 Laplace's equation, 215–221, 224, 225–235
 Laplacian, 15, 23, 25
 of \mathbf{A} , 351, 678
 of \mathbf{B} , 357, 504, 518
 of \mathbf{E} , 233, 504, 518
 of V , 65, 170, 678
 Laser beam, 522, 524, 606, 710
 Laws, local and nonlocal, 52
 Leakage flux, 381–382
 Legendre polynomials, 90–93, 227–228
 Legendre's equation, 227
 Leighton, Robert B., 334

 Length, proper, 251
 Lenz's law, 419–426, 432
 Levitation, 488, 608
 Light cone, 244
 Light-emitting diodes, 592
 Light source paradox, 710
 Light, speed of, 242, 244, 247, 260, 263,
 504, 520
 Line integral, 10
 of \mathbf{A} , 348
 of \mathbf{B} , 352, 367, 378, 499
 of \mathbf{E} , 46, 198, 413, 426, 430, 495
 of \mathbf{H} , 367, 381
 Line losses, 445
 Lines of
 \mathbf{B} , 295, 329, 356, 371, 378
 \mathbf{D} , 232
 \mathbf{E} , 50, 80, 87, 199, 230, 295, 296
 \mathbf{H} , 378
 Liquid crystal displays (LCDs), 605
 London equations, 507
 Lorentz condition, 321, 325, 511, 676
 Lorentz contraction, 251, 261
 Lorentz force, 293, 300, 387, 391, 542
 Lorentz reciprocity theorem, 502
 Lorentz transformation, 241–244. *See also*
 Magnetic force
 Loss angle, 194, 525
 Loss tangent, 194
 Lotsch, H. K. V., 587
 Loudspeaker, 384
 Lyttleton, R. A., 63, 511

 Magnet
 bar, 364, 377–380
 electromagnet, 383, 450
 permanent, 379
 Magnetic braking force on a satellite, 432
 Magnetic circuits, 381–385
 Magnetic dipole, 337–340, 501
 oscillating, 684–686, 720–724, 728–729
 Magnetic energy density
 in a long solenoid, 476
 in a wave, 521, 523, 527, 539
 in terms of \mathbf{B} , 476
 in terms of \mathbf{B} and \mathbf{H} , 475
 in terms of \mathbf{J} and \mathbf{A} , 474, 488
 in terms of L and I , 467, 472, 488, 489
 in terms of Φ and \mathcal{R} , 487
 Magnetic field
 calculations of, 377
 near a current sheet, 358, 483
 of a circular loop, 331, 359

- Magnetic field (*cont.*)
 of a long cylindrical conductor, 298, 330, 335, 353, 358
 of a magnetic dipole. *See* Magnetic dipole
 of a magnetized rod, 364, 377–380
 of a Maxwell pair, 346
 of a moving charge, 291–301, 512
 of a rotating disk of charge, 691
 of a solenoid, 331, 346, 349, 355, 359, 367
 of a spinning charged sphere, 347
 of a toroidal coil, 358
 of Helmholtz coils, 346
 of magnetized material, 362, 373
 of parallel currents, 336, 343
 of saddle coils, 343
 of the earth, 406, 432
 rotating, 345
 transformation of, 301
 uniform, 417
- Magnetic field strength \mathbf{H} , 365–371
 curl of, 365
 divergence of, 373
 line integral of, 367
- Magnetic flux, 312, 329
 compression, 490
 leakage, 381–382
 linkage, 349, 421
- Magnetic flux density \mathbf{B} , 293, 294, 326, 686–687, 690–691
 average over a sphere, 359
 curl of, 312, 430
 divergence of, 310, 314, 333, 364
 line integral of, 352, 367, 378, 430
 lines of, 295, 329, 356, 371, 378
 saturation, 369
 surface integral of, 333
- Magnetic force, 293–294, 300, 326, 388. *See also* Lorentz force
 and lines of \mathbf{B} , 482
 between coaxial solenoids, 481
 between long parallel currents, 397, 410
 between particles, 410
 between two electric currents, 396, 478, 481, 489
 inside ferromagnetic materials, 392
 on a magnetic dipole, 491
 on a volume distribution of current, 398
 on a wire, 393
 within an isolated circuit, 485, 489
- Magnetic induction. *See* Magnetic flux density \mathbf{B}
- Magnetic materials, 361–384
- Magnetic monopole, 310, 327, 343, 498, 510
- Magnetic pressure, 483–484, 490
- Magnetic separation, 490
- Magnetic shutter, 490
- Magnetic susceptibility, 368
- Magnetic torque, 485, 490
- Magnetization curve, 369. *See also* Hysteresis, ferromagnetic
- Magnetization \mathbf{M} , 361, 373
- Magnetohydrodynamic (MHD) generator, 389–392
- Magnetometer, 434, 436
- Magnetomotance, 367, 381
- Magnetoresistor, 409
- Marsden, J. E., 14
- Mass
 effective, 74
 inertial, 281
 relativistic, 270
 rest, 270
- Maxwell bridge, 450
- Maxwell, James Clerk, 492, 499
- Maxwell's equations, 178, 308–314, 421, 492–500, 511. *See also* inside the back cover.
 invariance, 509, 512
 linearity, 498
 redundancy, 311–313, 500
- McAlister, S. P., 393
- McCuaig, Malcolm, 383
- McKinnon, W. R., 393
- Mesh method, 120, 122
- Metal detector, 469
- Metallic glass, 377
- Mho, 70
- Microstrip line, 621, 625
- Millikan, R. A., 63
- Millman's theorem, 153, 168
- Minkowski diagram, 245, 250
- Minkowski formulation, 495
- Minnaert, M., 703
- Mobility, 73, 82, 408
- Models, 507
- Modes of propagation, 630, 663
- Momentum, 270, 282
 electromagnetic, 602, 604, 609
 flux of, 602
 four-momentum, 271–273, 276, 287, 320
 relativistic, 270
- Monopole, electric, 93
- Monopole, magnetic, 310, 327, 343, 498, 510
- Moon, 116, 567

- Mössbauer effect, 283
 Motion transducer, 145
 Motor, 410, 534
 homopolar, 399–403, 411
 Multipoles, electric, 88–97
- Neper, 645
 Neumann equation, 437
 Nitrobenzene, 195
 Node method, 120, 126
 Norton's theorem, 152, 167, 168
 Nucleus, atomic, 61
 Number, binary, 142
- Observer, 239
 Octupole, electric, 89
 Ohm's law, 70, 134, 453
 Omega-beta diagram, 552, 669
- Panofsky, Wolfgang, K. H., 241
 Paradox of the perpendicular capacitors, 324
 Parallel-plate capacitor, 109, 114, 118, 509, 532, 545
 dielectric-insulated, 184, 187, 190, 193, 210, 203, 205
 magnetic field in circular, 510
 moving, 303, 321, 323
 Paramagnetism, 361
 Particular solution, 125
 Peaking strip, 434
 Penfield, Paul, 154, 307, 493
 Permalloy, 434
 Permeability, 368–370
 of free space, 294, 328, 398, 520
 Permeance, 382
 Permittivity, 181
 complex relative, 194, 525
 frequency and temperature dependence, 195
 of free space, 43, 398, 520
 relative, 75, 181, 194, 544
 Table, 182
 Perturbation method, 638
 Phase angle, 32
 Phase shifter, 146
 Phasors, 32–40
 Philips, Melba, 241
 Photomultiplier, 406
 Photon, 244, 263, 277
 Photon-drag detector, 603, 607
 Pinch effect, 405, 411
- Planar optical waveguide, 646–674
 dispersion relation, 669
 eigenvalue equation, 657, 662–666
 field components, 650–656, 661, 666–668
 field energy, 670, 673
 modal dispersion, 672
 mode order, 657, 672
 numerical aperture, 660, 672
 phase and group velocities, 668–670, 673
 phase shifts on total reflection, 656
 transmitted power, 671, 674
- Planck's constant, 669
 Plane of incidence, 557
 Plasma, 542–553, 569
 frequency, 544, 547
 Point charges, 104
 Point function, 3, 5
 Poisson's equation, 737
 for \mathbf{B} , 357
 for \mathbf{E} , 233
 for V , 65, 178, 189, 233
 Polar dielectrics, 173
 Polarization, electric, 173
 current density, 176, 313
 Polarization, wave. *See* Electromagnetic waves, polarization of
 Polarizer, 574
 Polarizing angle, 566
 Poles, magnetic, 379
 Polyvinylidene fluoride, 177
 Port, 155
 Positive logic, 142
 Positron, 236, 275
 Potential
 electric. *See* Electric potential V
 four-potential, 319, 320, 323
 line integral of vector, 348
 scalar, 316, 333
 scalar magnetic, 386
 vector. *See* Vector potential
 Potential divider, 141, 145
 Potential energy. *See* Electric energy
 Potentials, electromagnetic, 314–322, 324
 retarded, 680–686, 689
 Potentiometer, 141
 Powdered iron, 425
 Power
 factor, 138, 195, 445, 450, 451
 in alternating-current circuits, 138
 Power-transfer theorem, 161
 Poynting theorem, 528
 Poynting vector, 519, 521–523, 532–535,
 539

- Precipitation, electrostatic, 62
 Pressure
 magnetic, 483–484, 490
 radiation, 599–603, 607
 Propagation constant, 517
 Proton beam, 61, 533
 Pulsar, 553
 Pump, electromagnetic, 407
 Purcell, E., 174
- Q* of a circuit, 454
 Quad, 289
 Quadrupole, electric, 87, 89, 94–95, 99
- Radar, 577
 Radian length, 517, 733
 Radiation, 675–729
 by an accelerated charge, 692–697
 by an oscillating charge, 710
 electric dipole, 697–708
 magnetic dipole, 684–686, 720–724
 pattern, 720–721
 resistance, 703
 Radiation pressure, 599–603, 607
 Rasetti, F., 392
 Ray equation, 573
 Re operator, 33
 Reactance, 134
 Reciprocity theorems, 117, 157–161, 168–169, 439
 Red shift, 261, 281
 Reference frame, 239, 258
 inertial, 239, 272
 Reflection and refraction, 554–609
 Brewster's angle, 565–567, 574–575, 589
 coefficients of, 567–569
 conductors, 594–608
 Fresnel's equations, 558–565, 573
 index of refraction, 523, 531, 569, 573, 576
 laws of reflection, 557
 nonconductors, 561
 plane of incidence, 557
 plasmas, 569–571
 radiation pressure, 599–603, 607
 ray equation, 573
 Snell's law, 557, 570
 total reflection, 581–593, 656
 Relativity, 238–325
 general and special, 239
 principle of, 240
- Relaxation time, 76, 525
 Reluctance, 381, 450
 Remanence, 376
 Resistance, 70
 output, 150
 radiation, 703, 716, 723
 surface, 81, 551
 Resistivity, 80, 435. *See also* Conductivity
 Resistojet, 83
 Resonace, 453–456
 Retentivity, 376
 Right-hand screw rule, 2
 RMS, or root-mean square, value, 36, 131
 Rowland ring, 375, 422
 Rutherford experiment, 59
- Sands, Matthew, 334
 Satellites, 432, 553, 567
 Scalar potential. *See* Electric potential
 Scattering, resonant, 283
 Scintillation particle detector, 592
 Seed-sorting device, 58
 Semiconductors, 70, 388, 389
 Separation of variables, 216
 Set, complete, 220
 Shape, apparent, of rapidly moving object, 254
 Shield, 77, 450
 S.I. (*Système International*) units, 730, 731
 Siemens, 70
 Siemens, E. W. von, 70
 Signal velocity, 247
 Simultaneity, 247
 Skin depth, 524, 531, 538, 540
 Skin effect, 359, 443, 488, 510, 537
 Smith, J., 255
 Snell's law, 557, 570
 Solar wind, 533
 Soldering gun, 469
 Solenoid, 324, 331, 346, 349, 355, 359, 367
 Source, current density inside a, 494
 Space shuttle, 548
 Space-time diagram, 244–247, 249, 266
 Spaceship, 282
 Speckle, 710
 Spence, Robert, 154
 Spherical harmonic functions, 215
 Starred sections, 88
 Stokes's theorem, 13
 Submarines, 598
 Substitution theorem, 150

- Sun, 344, 603
 Superconductivity, 507
 Superposition, principle of, 45, 48, 149, 329
 Surface
 integral of \mathbf{B} . *See* Gauss's law for \mathbf{B}
 integral of \mathbf{E} . *See* Gauss's law for \mathbf{E}
 Surface, orientable, 14
 Susceptance, 135
 Susceptibility
 electric, 181
 magnetic, 368
 Synchrotron radiation, 722, 728
- Tamir, T. 587
 Taylor, Edwin F., 239, 242
 Tellegen's theorem, 154–157, 168
 Terrell, James, 254
 Tesla, 328
 Thermal agitation, 72
 Thévenin's theorem, 150, 167, 391
 Tides, 432
 Time
 constant, 125
 dilation, 254, 257
 interval, proper, 258
 proper, 254, 258
 read on a moving clock, 255
 Transformation, 241
 Galilean, 240, 243
 Lorentz, 241–244, 248–250
 of \mathbf{A} , 319
 of a current density, 287
 of a force, 276, 282
 of a four-current density, 287
 of a four-momentum, 273
 of a four-potential, 321, 324
 of a four-vector, 266
 of a length, 251
 of a magnetic flux, 429
 of a magnetic force, 324
 of a mass, 273
 of a mass density, 280
 of a mutual inductance, 458
 of a relative permittivity, 307
 of a time interval, 254, 258
 of a velocity, 259
 of an angle, 249
 of an electric charge density, 287
 of an electromotance, 429
 of an element of area, 253, 261
 of an element of volume, 285
 of an equation, 495
 of \mathbf{E} and \mathbf{B} , 301
 of \mathbf{M} , 307
 of Maxwell's equations, 313
 of \mathbf{P} , 307
 of quad, 289
 of the field of a long solenoid, 324
 of the field of a parallel-plate capacitor, 303, 321, 323
 of the potentials, 319, 324
 of the space and time partial derivatives, 289
 of the speed of light, 244, 260
 of V , 319
 of visible light into high-energy radiation, 263
 Transformer, 457, 459, 469, 535
 auto, 463
 current, 449
 ideal, 461–465
 magnetic-core, 377, 425, 460, 468
 power transfer through a, 464
 Transient, 126
 Transmission lines, 147, 489, 615, 617. *See also* Coaxial line; Hollow rectangular waveguide; Microstrip line; Planar optical waveguide
 Travel, interstellar, 282
 Tromba, A. J., 14
 Tsukada, Masanobu, 81
 Twin paradox, 262
- Uniqueness theorem, 211, 222, 231
 Units, 730, 731
 Universe, expansion of the, 63
- Van Bladel, Jean, 26
 Van de Graff accelerator, 61, 358
 Van der Pauw, 169
 Variables, separation of, 216, 226
 $\mathbf{v} \times \mathbf{B}$ field, 324, 413–419
 Vector, 2–24, 267. *See also* Four-vector definitions, identities, and theorems. *See inside the front cover*
 operators, 3–24
 rotating, 37
 unit, 2, 16, 17, 18, 27
 Vector potential, 314, 323, 333, 347–349, 417
 divergence, 351, 676
 Laplacian, 351, 678
 line integral of, 348

- Velocities, addition of, 259
Velocity
 energy transport, 674
 group, 539, 547, 552, 734
 maximum, 242, 249, 270
 maximum signal, 247, 249
 phase, 517, 523, 525, 538, 536, 552, 569,
 733–734
Virtual work, method of, 114, 479
Volt, 45, 48
Voltage source, 120
VSWR, voltage standing-wave ratio, 643
- Wannier, G. H., 392
Watson, W. H., 513
Wave equations, 290, 737
 for A , 678
 for B and E , 504, 518
 for V , 678
- Wave impedance, 614–615
Wave number, 34, 517, 518, 524–526, 545,
 733–734
Wavefront, 515
Waveguides, 610–674. *See also* Coaxial line;
 Hollow rectangular waveguide;
 Microstrip line; Planar optical wave-
 guide; Transmission lines
Wavelength, 517, 733
 cut-off, 630
Waves, 34, 732–738. *See also*
 Electromagnetic waves
 guided. *See* Waveguides
 uniform and nonuniform, 537, 579, 732
Weber, 328, 349
Weisskopf, V. F., 254
Wheatstone bridge, 155
Wheeler, John A., 239, 242
Wien bridge, 146
World line, 244, 271

Vector definitions, identities, and theorems

Definitions

Rectangular coordinates

$$1. \quad \nabla f = \frac{\partial f}{\partial x} \hat{x} + \frac{\partial f}{\partial y} \hat{y} + \frac{\partial f}{\partial z} \hat{z}$$

$$2. \quad \nabla \cdot A = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

$$3. \quad \nabla \times A = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \hat{x} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \hat{y} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \hat{z}$$

$$4. \quad \nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

$$5. \quad \nabla^2 A = \nabla^2 A_x \hat{x} + \nabla^2 A_y \hat{y} + \nabla^2 A_z \hat{z} = \nabla(\nabla \cdot A) - \nabla \times (\nabla \times A)$$

Cylindrical coordinates

$$6. \quad \nabla f = \frac{\partial f}{\partial \rho} \hat{\rho} + \frac{1}{\rho} \frac{\partial f}{\partial \phi} \hat{\phi} + \frac{\partial f}{\partial z} \hat{z}$$

$$7. \quad \nabla \cdot A = \frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho A_\rho) + \frac{1}{\rho} \frac{\partial A_\phi}{\partial \phi} + \frac{\partial A_z}{\partial z}$$

$$8. \quad \nabla \times A = \left(\frac{1}{\rho} \frac{\partial A_z}{\partial \phi} - \frac{\partial A_\phi}{\partial z} \right) \hat{\rho} + \left(\frac{\partial A_\rho}{\partial z} - \frac{\partial A_z}{\partial \rho} \right) \hat{\phi} + \frac{1}{\rho} \left[\frac{\partial}{\partial \rho} (\rho A_\phi) - \frac{\partial A_\rho}{\partial \phi} \right] \hat{z}$$

$$9. \quad \nabla^2 f = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial f}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2}$$

$$10. \quad \nabla^2 A = \nabla(\nabla \cdot A) - \nabla \times (\nabla \times A) \text{ (Sec. 1.11.6).}$$

Spherical coordinates

$$11. \quad \nabla f = \frac{\partial f}{\partial r} \hat{r} + \frac{\partial f}{r \partial \theta} \hat{\theta} + \frac{1}{r \sin \theta} \frac{\partial f}{\partial \phi} \hat{\phi}$$

$$12. \quad \nabla \cdot A = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 A_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (A_\theta \sin \theta) + \frac{1}{r \sin \theta} \frac{\partial A_\phi}{\partial \phi}$$

$$13. \quad \nabla \times A = \frac{1}{r \sin \theta} \left[\frac{\partial}{\partial \theta} (A_\phi \sin \theta) - \frac{\partial A_\theta}{\partial \phi} \right] \hat{r} + \frac{1}{r} \left[\frac{1}{\sin \theta} \frac{\partial A_r}{\partial \phi} - \frac{\partial (r A_\phi)}{\partial r} \right] \hat{\theta} + \frac{1}{r} \left[\frac{\partial (r A_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta} \right] \hat{\phi}$$

$$14. \quad \nabla^2 f = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2}$$

$$15. \quad \nabla^2 A = \nabla(\nabla \cdot A) - \nabla \times (\nabla \times A) \text{ (Sec. 1.11.6).}$$

Identities

1. $(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$
2. $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$
3. $\nabla(fg) = f\nabla g + g\nabla f$
4. $\nabla(a/b) = (1/b)\nabla a - (a/b^2)\nabla b$
5. $\nabla(\mathbf{A} \cdot \mathbf{B}) = (\mathbf{B} \cdot \nabla)\mathbf{A} + (\mathbf{A} \cdot \nabla)\mathbf{B} + \mathbf{B} \times (\nabla \times \mathbf{A}) + \mathbf{A} \times (\nabla \times \mathbf{B})$
6. $\nabla \cdot (f\mathbf{A}) = (\nabla f) \cdot \mathbf{A} + f(\nabla \cdot \mathbf{A})$
7. $\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B})$
8. $(\nabla \cdot \nabla)f = \nabla^2 f$
9. $\nabla \times (\nabla f) = 0$
10. $\nabla \cdot (\nabla \times \mathbf{A}) = 0$
11. $\nabla \times (f\mathbf{A}) = (\nabla f) \times \mathbf{A} + f(\nabla \times \mathbf{A})$
12. $\nabla \times (\mathbf{A} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla)\mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{B} + (\nabla \cdot \mathbf{B})\mathbf{A} - (\nabla \cdot \mathbf{A})\mathbf{B}$
13. $\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$ (Sec. 1.11.6)
14.
$$(\mathbf{A} \cdot \nabla)\mathbf{B} = \left[A_x \frac{\partial B_x}{\partial x} + A_y \frac{\partial B_x}{\partial y} + A_z \frac{\partial B_x}{\partial z} \right] \hat{x} \\ + \left[A_x \frac{\partial B_y}{\partial x} + A_y \frac{\partial B_y}{\partial y} + A_z \frac{\partial B_y}{\partial z} \right] \hat{y} \\ + \left[A_x \frac{\partial B_z}{\partial x} + A_y \frac{\partial B_z}{\partial y} + A_z \frac{\partial B_z}{\partial z} \right] \hat{z}$$
15. $\nabla'(1/r) = \hat{r}/r^2$. This is the gradient calculated at (x', y', z') , and \mathbf{r} is the vector \mathbf{r} pointing from (x', y', z') to (x, y, z) .
16. $\nabla(1/r) = -\hat{r}/r^2$. This is the gradient calculated at (x, y, z) with the same vector \mathbf{r} .
17. $\mathcal{A} = \frac{1}{2} \oint_C \mathbf{r} \times d\mathbf{l}$, where the surface of area \mathcal{A} is plane. The vector \mathbf{r} extends from an arbitrary origin to a point on the curve C that bounds \mathcal{A} .
18. $\int_v \nabla f dv = \int_{\mathcal{A}} f d\mathcal{A}$
19. $\int_v (\nabla \times \mathbf{A}) dv = -\int_{\mathcal{A}} \mathbf{A} \times d\mathcal{A}$, where \mathcal{A} is the area of the closed surface that bounds the volume v .
20. $\oint_C f d\mathbf{l} = -\int_{\mathcal{A}} \nabla f \times d\mathcal{A}$ where C is the closed curve that bounds the open surface of area \mathcal{A} .

Theorems

1. The divergence theorem. $\int_{\mathcal{A}} \mathbf{A} \cdot d\mathcal{A} = \int_v \nabla \cdot \mathbf{A} dv$ where \mathcal{A} is the area of the closed surface that bounds the volume v .
2. Stokes's theorem: $\oint_C \mathbf{A} \cdot d\mathbf{l} = \int_{\mathcal{A}} (\nabla \times \mathbf{A}) \cdot d\mathcal{A}$.

Physical constants[†]

Elementary charge	$e = 1.602177 \times 10^{-19} \text{ C}$
Electron rest mass	$m = 9.10938 \times 10^{-31} \text{ kg}$ $= 5.10999 \times 10^5 \text{ eV}$
Proton rest mass	$m_p = 1.67262 \times 10^{-27} \text{ kg}$ $= 9.38272 \times 10^8 \text{ eV}$
Speed of light in vacuum	$c = 2.99792458 \times 10^8 \text{ m/s}$
Permittivity of vacuum	$\epsilon_0 = 8.85418782 \times 10^{-12} \text{ F/m}$ $1/4\pi\epsilon_0 = 8.9875518 \times 10^9 \text{ N/C}$
Permeability of vacuum	$\mu_0 = 4\pi \times 10^{-7} \text{ H/m}$
Avogadro constant	$N_A = 6.02213 \times 10^{23} \text{ mol}^{-1}$
Boltzmann constant	$k = 1.3806 \times 10^{-34} \text{ J/K}$
Planck constant	$h = 6.62607 \times 10^{-34} \text{ J} \cdot \text{s}$ $\hbar = h/(2\pi) = 1.054572 \times 10^{-34} \text{ J} \cdot \text{s}$
Gravitational constant	$G = 6.672 \times 10^{-11} \text{ N} \cdot \text{m}^2/\text{kg}^2$
Mass of the sun	$1.98596 \times 10^{30} \text{ kg}$
Radius of the sun	$6.965 \times 10^8 \text{ m}$
Mean sun-earth distance	$1.495 \times 10^{11} \text{ m}$
Earth's mean orbital speed	$2.98 \times 10^4 \text{ m/s}$
Mass of the earth	$5.974 \times 10^{24} \text{ kg}$
Radius of the earth	$6.378 \times 10^6 \text{ m}$
Mass of the moon	$7.3305 \times 10^{22} \text{ kg}$
Radius of the moon	$1.74 \times 10^6 \text{ m}$
Moon-earth distance	$3.84393 \times 10^8 \text{ km}$

[†] Codata Bulletin, November 1986.

Maxwell's equations for stationary media

A. Differential form with \mathbf{E} , \mathbf{B} , \mathbf{P} , \mathbf{M}

$$\begin{aligned}\nabla \cdot \mathbf{E} &= \frac{\rho_f - \nabla \cdot \mathbf{P}}{\epsilon_0}, & \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} &= 0, \\ \nabla \cdot \mathbf{B} &= 0, & \nabla \times \mathbf{B} - \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} &= \mu_0 \left(\mathbf{J}_f + \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M} \right).\end{aligned}$$

B. Integral form

$$\begin{aligned}\int_{\mathcal{A}} \mathbf{E} \cdot d\mathcal{A} &= \frac{1}{\epsilon_0} \int_v (\rho_f - \nabla \cdot \mathbf{P}) dv \\ \int_{\mathcal{A}} \mathbf{B} \cdot d\mathcal{A} &= 0, \\ \int_{\mathcal{A}} (\nabla \times \mathbf{E}) \cdot d\mathcal{A} &= \oint_C \mathbf{E} \cdot d\mathbf{l} = - \frac{d\Lambda}{dt}, \\ \int_{\mathcal{A}} (\nabla \times \mathbf{B}) \cdot d\mathcal{A} &= \oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_0 \int_{\mathcal{A}} \left(\mathbf{J}_f + \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \cdot d\mathcal{A}.\end{aligned}$$

C. Differential form with \mathbf{E} , \mathbf{D} , \mathbf{H} , \mathbf{B}

$$\begin{aligned}\nabla \cdot \mathbf{D} &= \rho_f, & \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} &= 0, \\ \nabla \cdot \mathbf{B} &= 0, & \nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} &= \mathbf{J}_f.\end{aligned}$$

D. Sinusoidal fields with \mathbf{E} , \mathbf{D} , \mathbf{H} , \mathbf{B} and for linear media

$$\begin{aligned}\nabla \cdot \epsilon \mathbf{E} &= \rho_f, & \nabla \times \mathbf{E} + j\omega \mu \mathbf{H} &= 0 \\ \nabla \cdot \mu \mathbf{H} &= 0, & \nabla \times \mathbf{H} - j\omega \epsilon_r \epsilon_0 \mathbf{E} &= \mathbf{J}.\end{aligned}$$